CHARACTERIZATION OF PLANT UDP-GLUCURONIC ACID 4-EPIMERASES

by

XIAOGANG GU

(Under the Direction of Maor Bar-Peled)

ABSTRACT

Galacturonic acid (GalA) is an important sugar residue found in pectic polysaccharides, xylan sequence 1 structure and the arabinogalactan-proteins (AGPs). These GalA containing glycans are essential components of primary and secondary wall, and are critical for proper plant structure, growth and development.

UDP-GalA is a key precursor for the syntheses of GalA containing glycans. In plant, UDP-GalA is synthesized from UDP-glucuronic acid (UDP-GlcA) by UDP-GlcA 4-epimerases (UGlcAE). My research focused on the identification of UGlcAE genes in plants and the characterization of these proteins at the biochemical, cellular and genetic levels. In plants, multiple UGlcAE isoforms exist and each isoform contains two domains: a sub-cellular localization domain at the N-terminus (~120 amino acids long), followed by a highly conserved catalytic domain (300 amino acids long). Based on phylogeny analyses, these isoforms can be further divided into three evolutionary clades: A, B and C. We determined that UGlcAEs from each clade are enzymatically active as dimers, and all have similar enzymatic kinetic properties. Each of the isoforms tested is inhibited by UDP-xylose and UDP but not by UMP, UDP-glucose or UDP-galactose. Interestingly however, the isoforms from different plant species have different

properties. For example, the maize UGlcAE is inhibited to a greater extent by UDP-arabinose when compared with rice and *Arabidopsis* homologs.

We found that all identified isoforms are integral membrane proteins and all the isoforms studied (as UGlcAEs- EYFP fusion) are located in the plant Golgi with their catalytic domains in the Golgi lumen. Since each plant cell contains many Golgi apparatus, we could not distinguish the localization of a specific isoform to a specific set of Golgi within one plant cell.

The control of UDP-GalA flux for the synthesis of various GalA containing glycans is now more complicated with the recent discovery of another UDP-GalA synthesis pathway, the salvage pathway. While the GalA salvage pathway forms UDP-GalA in the cytoplasm, the inter-conversion pathway we studied here forms UDP-GalA inside the Golgi lumen.

INDEX WORDS: UDP-glucuronic acid, UDP-galacturonic acid, UDP-GlcA-4-epimerase isoforms (UGlcAE), Pectin polysaccharide, xylan, GalA glycans, membrane topology, subcellular localization, Golgi

CHARACTERIZATION OF PLANT UDP-GLUCURONIC ACID 4-EPIMERASES

by

XIAOGANG GU

B.S. Nanjing University, China, 2001

M.S. Nanjing University, China, 2003

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

DOCTOR OF PHILIOSOPHY

ATHENS, GEORGIA

2009

CHARACTERIZATION OF PLANT UDP-GLUCURONIC ACID 4-EPIMERASES

by

XIAOGANG GU

Major Professor:	Maor Bar-Peled
Committee:	Alan Darvill Michael Hahn Kelley Moremen Sue Wessler

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2009

DEDICATION

I dedicate this work to my family, for their unconditional support and love, especially to my grandma Fengying Wu.

ACKNOWLEDGEMENTS

I would first like to thank my major professor, Dr. Maor Bar-Peled, for his direction and support for this project. I also appreciate my other committee members: Dr. Alan Darvill, Dr. Michael Hahn, Dr. Kelly Moremen and Dr. Sue Wessler, for their valuable comments in completion of this research.

I would also like to thank other individuals who thought me specifically, Dr. Michael Hahn for helping me in learning the secret of phylogenetic analysis, Dr. Federica Brandizzi at Michigan State University for her hospitality and teaching and guiding me with confocal microscope. I would also like to thank Dr. Andreas Nebenführ at the University of Tennessee for providing me with the ER and Golgi markers. I would like to thank Stephen Eberhardt for his endless efforts in maintaining the growth chambers at the CCRC. In addition, I would like to thank the great people in my lab for their great help.

Last but not least, I would like to thank the many undergraduate students that worked or are working with me. They are Paul, Liron, Kieu, Uzo, Jamie and Chris.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

PART I: GALACTURONIC ACID (GALA) CONTAINING GLYCANS: STRUCTURE, FUNCTION AND BIOSYNTHESIS.

Introduction

There are two different types of walls in plant: the polysaccharide-rich primary wall that surrounds growing plant cells, and the much thicker and stronger secondary wall that is deposited once the cell growth is ceased. The primary walls isolated from higher plant tissues are composed predominantly of various polysaccharides (up to 90% of dry weight), structural glycoproteins (e.g. HGRPs and GRPs), ionically and covalently bound minerals (1-5%), and phenolic esters (<2%). Walls also contain numerous enzymes, including those involved in wall metabolism (e.g. endo and exoglycanases, methyl and acetyl esterases, and trans-glycosylases) and enzymes that may generate cross-linking between wall components (e.g. peroxidases).

One of the major primary wall polysaccharides in flowering plants is pectin, which is rich in galacturonic acid (GalA) residues. Pectin is believed to play critical roles in plant development and growth, and this thesis focuses on enzymes that are involved in producing the precursors (UDP-GalA) for the synthesis of those pectic polysaccharides. The first part of the literature review will summarize the available knowledge of the structure, function and biosynthesis of the pectic polysaccharides.

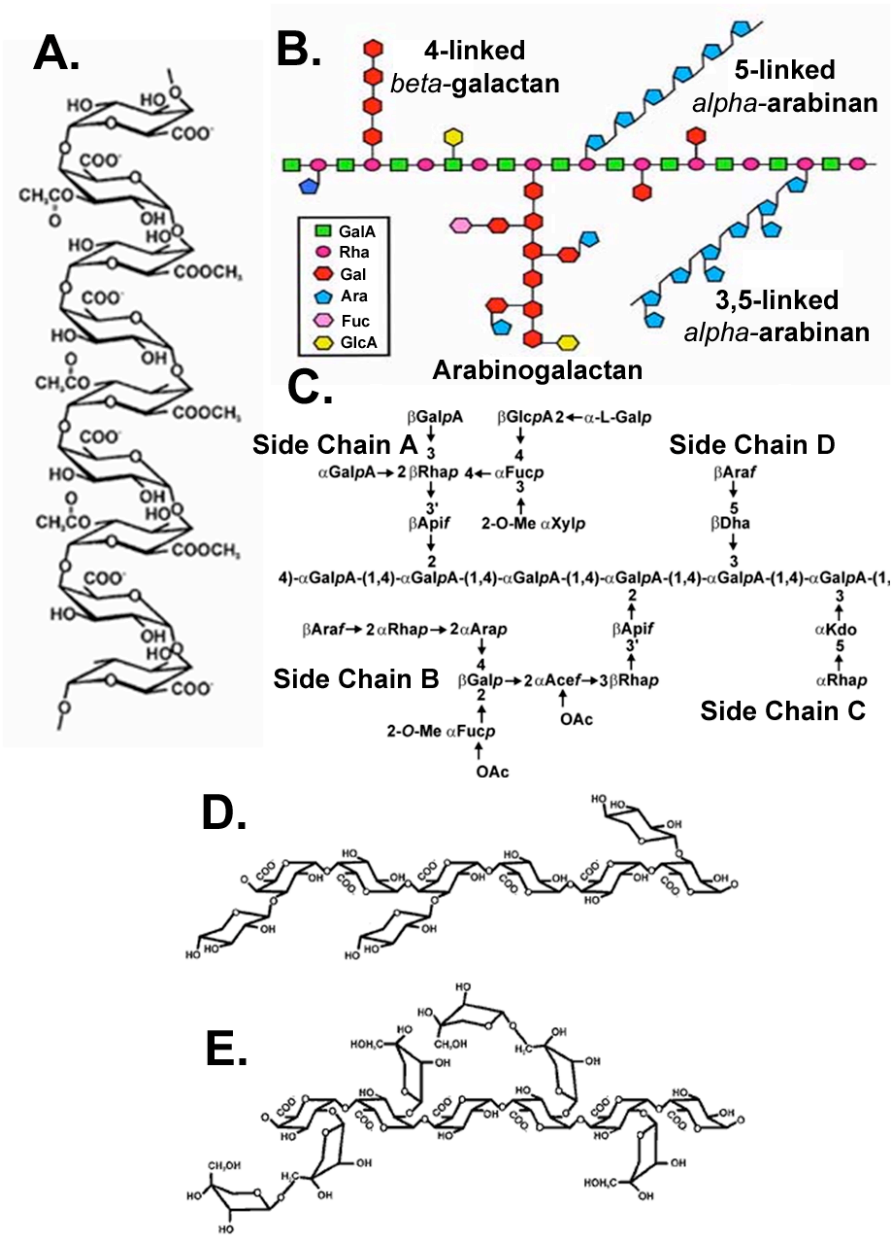


Figure 1.1. The structure of major pectic polysaccharides

A. Homogalacturonan (HGA) is a linear chain of 1, 4-linked α-D-GalA residues, some of which are shown to be methyl esterified at C-6, or *O*-acetylated at C-3 or C-2 (http://www.ccrc.uga.edu/~mao/galact/gala.htm),

B. Rhamnogalacturonan I (RG-I), a backbone of rhamnose and GalA residues with various side chain modifications (http://www.ccrc.uga.edu/~mao/rg1/rg1.htm)

C. Rhamnogalacturonan II (RG-II), a backbone of GalA residues decorated with four conserved side chains. (http://www.ccrc.uga.edu/~mao/rg2/intro.htm)

D. Xylogalacturonan (XGA), a backbone of GalA with xylose side chains (http://www.ccrc.uga.edu/~mao/galact/gala.htm)

E. Apigalacturonan (AGA) a backbone of GalA with apiose side chains (http://www.ccrc.uga.edu/~mao/galact/gala.htm)

The structures of plant GalA containing glycans

GalA is a major sugar residue in plant pectic polysaccharides. Five different types of pectic polysaccharides (homogalacturonan, rhamnogalacturonan I, rhamnogalacturonan II, xylogalacturonan and apigalacturonan) have been isolated from plant primary cell walls and structurally characterized (Ridley et al., 2001, Mohnen 2008).

Homogalacturonan (HGA) is a linear chain of 1, 4-linked α-D-GalA residues (see Fig. 1.1A), which accounts for up to 60% of the pectin in the primary walls of dicotyledons and non-graminaceous monocotyledons and is the predominant anionic polymer (O'Neill and York, 2003). Some of the carboxyl groups of HGA may be methyl esterified at the C-6 or *O*-acetylated at C-3 or C-2 (Fig. 1.1A). In addition, Ishii (1997) reported the presence of *O*-acetylation at C-4 of the non-reducing GalA residue in galacturonide-oligosaccaride (DP 2, 3) obtained from potato tuber cell walls. The pattern and degree of the methyl-esterification and acetyl-esterification may vary depending on the plant source, the specific plant tissue and developmental stage (Willats et al., 2001). The non-esterified C-6 carboxyl group of GalA residues in HGA can bind Ca^{2+} and form Ca^{2+} bridges between GalA-chains (O'Neill and York, 2003). The degree of methyl-esterification, the amount of Ca^{2+} at the wall, and the nature of the HGA-calcium cross-linking thus provide this pectic polymer with different physical properties.

Rhamnogalacturonan I (RG-I) is a family of heterosaccharide polymers with a repeating disaccharide backbone of (1,2)-α-L-Rha-(1,4)-α-D-GalA (Fig 1.1B). Depending on the plant source, the rhamnose (Rha) residues of RG-I backbone are substituted to a greater or lesser degree at C-4 with various side chains, including linear and branched galactans, arabinans and arabinogalactans (Fig 1.1B) (Willats et al., 2001; Mohnen, 2008). Although one report

demonstrated that a single GlcA residue was attached to the backbone GalA residues in sugar beet RG-I (Renard et al., 1999), the GalA residues are usually not substituted with other glycosyl residues. However, they might be *O*-acetylated at C-3 as reported by Ishii (1997) and it is possible that the degree of acetylated pectin is under-estimated since typical extraction at high pH removes the base-labile *O*-acetyl group. In addition, sugar residues like α-L-Fucose, β-D-GlcA and 4-*O*-Me β-D-GlcA may also be present in RG-I (O'Neill and York, 2003). Recently, a novel RG-I structure, which has numerous, single, non-reducing, terminal residues of the rare sugar L-galactose attached at the O-3 position of the rhamnosyl residues instead of the typical O-4 position, was identified in the viscous seed mucilage of flax (*Linum usitatissimum*), suggesting the complexity of RG-I structure in different plant species (Naran et al., 2008).

Rhamnogalacturonan II (RG-II) contains a backbone of at least eight 1,4-linked α-D-GalA residues (Fig 1.1C) and is suggested to exist in the walls of all higher plants (O'Neill and York, 2003). The GalA residues of RG-II backbone are decorated with four structurally distinct side chains (A, B, C and D): two disaccharides (C and D) attached to C-3 of the backbone and two oligosaccharides (A and B) attached to C-2 of the backbone (Fig 1.1C). However, the exact location of the side chains on the backbone with respect to each other has not been unambiguously established. RG-II is a structurally complex pectic polysaccharide in plants. It is composed of at least 12 different glycosyl residues including some rare sugars like Dha and Kdo, and contains more than 20 different glycosidic linkages (O'Neill et al., 2004). Despite this complexity, however, the glycosyl residue composition of RG-II (with a few exceptions) is remarkably conserved among plant species (O'Neill et al., 2004).

Xylogalacturonans (XGA) and **apigalacturonans** (AGA) both contain a backbone of linear 1,4-linked α-D-GalA residues. XGA contains β-D-xylose (Xyl) residues attached to the C-

3 of the backbone (Fig 1.1D) and it has been detected in the walls of specific plant tissues including soybean and pea seeds (O'Neill and York, 2003), mellon fruit (Mort et al., 2008), Arabidopsis leaf (Zandleven et al., 2007) and possibly in apple fruit (Schols et al., 1995). AGA contains β-D-apiose (Api) residues attached to C-2 of the backbone (Fig 1.1E) and is present in the wall of some aquatic monocotyledonous plants, including *Lemma* and *Zostera* (O'Neill and York, 2003).

While GalA is a major sugar residue in pectic polysaccharides (Fig 1.1), it is also found in other plant glycans. For example, the reducing end-group of spruce and Arabidopsis xylan contains GalA within the "xylan sequence 1 structure": β-D-Xyl-(1,3)-α-L-Rha-(1,2)-α-D-GalA-(1,4)-D-Xyl (Andersson et al., 1983; Pena et al., 2007). In addition, GalA residues were also found in the arabinogalactan-proteins (AGPs) from red wine (Pellerin et al., 1995) and in the Arabidopsis AGPs as well (Kieliszewski, personal communication). Kieliszewski also reported that AGP isolated from the media of BY2 tobacco cells has a GalA residue linked β1-6 to Gal.

Recently, a novel type of polysaccharide was isolated from the bark of *Cola cordifolia* (Sterculiaceae) and it is comprised of 20% 2,3- and 2,4-linked rhamnose, 24% 4-linked galacturonic acid, 15% terminal, 3- and 4-linked galactose, and various methylated monosaccharides including 20% terminal and 3-linked 2-*O*-methyl galactose, 18% terminal 4-*O*-methyl glucuronic acid, and 2% terminal 2-*O*-methyl fucose (Togola et al, 2008). In addition, a β-D-GalA residue was recently reported to be present in the sidechains of XXGGG-type xyloglucans extracted from the moss *Physcomitrella patens* (Pena et al., 2008). Therefore, as more polysaccharides are analyzed, more types of GalA glycans may be discovered.

The 4-epimer of GalA, glucuronic acid (GlcA), was also found in various plant glycans. In glucurono(arabino)xylans, some of the backbone xylose residues were 1,2- linked to α-D-GlcA

or 4-*O*-Me α-D-GlcA, the methyl-ester form of GlcA (York and O'Neill, 2008). β-D-GlcA residues and 4-*O*-Me β-D-GlcA were found in arabinogalactan side chains of RG-I (An et al., 1994). In addition, Renard et al. (1999) reported that in sugar beet RG-I, GlcA residues were linked to GalA of the RG-I backbone. In RG-II, β-D-GlcA is α-1,4 linked to the L-fucose in side chain A (O'Neill and York, 2003), and GlcA residues were also found decorating the arabinogalactan chain of AGPs (Johnson et al., 2003). Furthermore, glycoglucuronomannan, an acidic polysaccharide isolated from gum exudates of quaruba (*Vochysia lehmannii*) consists of a β-GlcA-(1-2)-α-Man-(1-2)-β-GlcA-(1-2)-α-Man structure (Wagner et al., 2004).

Function of Plant GalA containing glycans

Numerous reports indicate the roles of GalA containing glycans in determining the rigidity of plant cell walls (O'Neill et al., 2001; Iwai et al., 2002; Jones et al., 2003; Parre and Geitmann, 2005; Jiang et al., 2006; Iwai et al., 2006), in cell-cell adhesion (Iwai et al., 2002; Santiago-Doménech et al., 2008) and in plant signaling (Pilling and Hofte, 2003), demonstrating their critical roles for plant growth and development.

The role of pectin in determining the plant cell wall rigidity

HGA. The degree of methyl esterification of HGA affects the affinity of HGA to calcium cations and is proposed to contribute distinct mechanical and porosity properties to the plant cell wall matrix (Willats et al., 2001). Modifying the degree of HGA-methyl esterification degree can influence the cell wall elasticity and thus affect the growth and development of certain plant tissues. For example, treating the *Solanum chacoense* pollen tubes with pectinase and pectin methyl esterase (PME) alters the physical properties of its pollen tube cell wall by reducing the cellular stiffness. This was proposed to stimulate pollen germination and tube growth (Parre and

Geitmann, 2005). On the other hand, a gene knockout of a putative *Arabidopsis thaliana* PME, *VANGUARD1* (*VGD1*), was reported to greatly retard the growth of the pollen tube in the style and transmitting tract, and led to reduction of male fertility (Jiang et al., 2006).

RG-I. The arabinan side-chain of RG-I and the proposed association between RG-I and HGA are suggested to play roles in regulating the cell wall rigidity and were shown to be essential for guard cell function (Jones et al., 2003; Jones et al., 2005). Jones et al (2003) reported that degradation of RG-I arabinan side chains within the guard cell wall of *Commelina communis* after arabinanase treatment can prevent stomatal opening or closing. Interestingly, this "locking" of guard cell wall movements can be reversed if homogalacturonan is subsequently removed from the wall by a combination of PME and endopolygalacturonase (EPG) treatment. Jones et al. suggested that the arabinan side-chain of RG-I would maintain flexibility in the cell wall by preventing homogalacturonan polymers from forming tight Ca^{2+} cross-link associations (see Fig 1.2). Similar biological roles of RG-I arabinan side-chains and HGA in guard cell function were identified in guard cells of *Vicia faba* and *Zea mays* (Jones et al., 2005).

RG-II. In the primary wall, RG-II exists as a dimer. The dimer is formed via borate crosslinking of two apiose residues (Fig 1.3) and is believed to have an important role in regulating the chemical and physical properties of cell wall pectin (O'Neill et al., 2004). In particular, reducing the amount of apiose or boron available for RG-II synthesis prevents normal plant growth and development (O'Neill et al., 2001). This RG-II dimerization is conserved in the primary walls of numerous plants (O'Neill and York, 2003). Mutant plants with fewer or complete disruption of RG-II-borate complexes always demonstrate boron deficient phenotypes, including dwarfed stature, slowed root growth, increased tissue brittleness and aberrant development of reproductive organs (O'Neill et al., 2001, Iwai et al., 2002, Iwai et al., 2006).

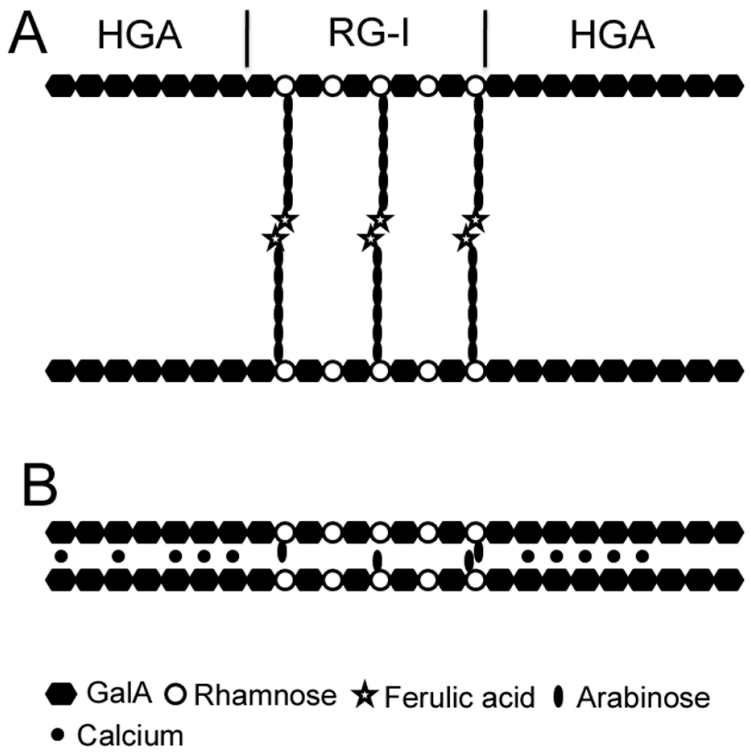


Figure 1.2. The proposed model of HGA/RG-I complex in regulating the stomata cell wall rigidity

Arabinan side chains of RG-1 may be esterified by ferulic acid, which provides hindrance to the association of neighboring domains of HGA (A). Digestion with arabinanase removes the side chains, allowing HGA to form tight Ca^{2+} cross-link associations (B), which leads to "locking" of the wall. (Jones et al., 2003)

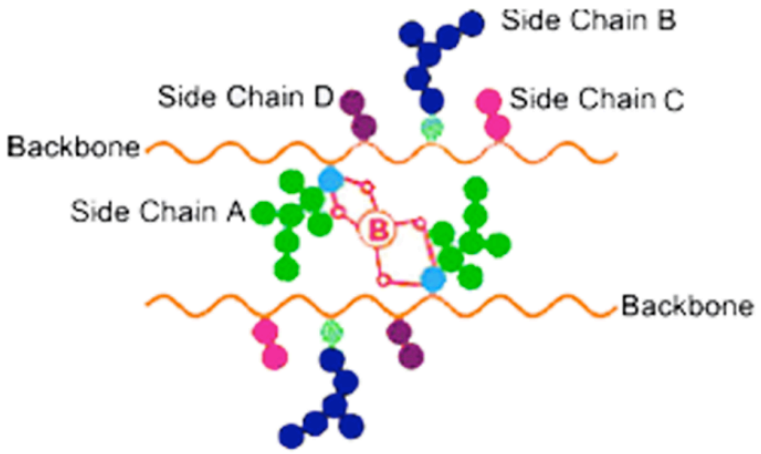


Figure 1.3. The structure of the 1:2 borate-diol ester that cross-links two RG-II molecules.

Apiose residue on each side-chain A of two adjacent RG-II molecules can be cross-linked by one borate ion to form a RG-II dimer. The figure was taken with permission from O'Neill et al., (2004)

In addition, Ahn et al., (2006) reported the reduction of RG-II in *Nicotiana benthamiana* by depletion of its UDP-D-apiose/UDP-D-xylose Synthases 1(NbAXS1), an enzyme producing UDP-apiose for the synthesis of RG-II side chain A (Fig 1.1C), results in abnormal wall structure including excessive cell wall thickening and wall gaps. This observation suggests that the normal dimer formation of RG-II-borate complex is required for the formation of a three-dimensional pectic network, which contributes to the mechanical properties of the primary wall and is required for normal plant growth and development (O'Neill et al., 2004).

The role of pectin in cell-cell adhesion

It was reported that inhibiting the expression of a strawberry pectate lyase gene by antisense transformation resulted in firmer fruits with an extended post-harvest life and the wall extracts from the transformed lines showed a reduction in pectin solubility and decreased depolymerization of HGA (Santiago-Doménech et al., 2008). Microscopic studies further revealed that in the pectate lyase knock-out strawberry lines, the typical ripening-associated loss of cell–cell adhesion was substantially reduced, indicating the role for pectin depolymerization in strawberry fruit softening (Santiago-Doménech et al., 2008). In addition, early studies of reducing the PME activity using antisense technology led scientists to produce a commercially available, edible, transgenic tomato with a longer shelf-life (Tieman et al., 1992). Analysis of the nolac-H18 (nonorganogenic callus with loosely attached cells) mutant further supported the role of pectin in cell-cell adhesion. In this T-DNA insertion mutant, a putative glucuronosyltransferase 1 (NpGUT1), involved in the biosynthesis of RG-II, was not expressed. The authors proposed that nolac-H18 mutation caused defects in linking GlcA to RG-II, thus reducing the formation of RG-II-borate complexes (Iwai et al., 2002). As a consequence, the mutant lost the ability of forming tight intercellular attachments and adventitious shoots.

The role of pectin in signaling

Pectin is considered to be a source of intercellular signaling molecules that function in plant morphogenesis and defense (Ridley et al., 2001). Oligogalacturonan (OGA) fragments, which are released by the activity of polygalacturonases and pectate lyases secreted by either the plant or plant pathogens, induce activation of plant defense responses (Pilling and Hofte, 2003). For example, feeding cell suspension cultures of parsley (*Petroselinum crispum*) with oligogalacturonides solubilized from parsley cell walls was shown to increase the pathogenic related protein, β-1,3-glucanase activity (Davis and Hahlbrock, 1987). The expression of allene oxide synthase, an enzyme involved in jasmonic acid synthesis during the Arabidopsis defense response (Norman et al., 1999) and the expression of aminocyclopropane 1-carboxylic acid oxidase, a key enzyme involved in the ethylene biosynthesis pathway in tomato plants (Simpson et al., 1998), were increased by treatment with exogenous OGA, indicating a role of OGAs in inducing both the JA and ethylene defense responses in plants.

The role of xylan sequence 1 structure in the biosynthesis of plant secondary cell wall

The reducing end of xylan extracted from woody and herbaceous plants (Andersson et al., 1983, Pena et al., 2007) contains an oligosaccharide domain abbreviated as "sequence 1 structure". So far, such a structure was not detected in grasses (Pena et al., personal communication). Recent studies have shown that this structure is critical for the biosynthesis of glucuronoxylan (GX) during secondary wall formation. Knocking-out the *Arabidopsis thaliana* IRREGULAR XYLEM8 (IRX8), a putative GalA transferase that is proposed to be involved in the biosynthesis of Arabidopsis xylan sequence 1 structure, was devoid of the xylan sequence 1 structure and deficient in GX (Pena et al., 2007; Persson et al., 2007). The *irx8* mutant also shows a strong dwarf phenotype and abnormal vessel morphology (Pena et al., 2007; Persson et

al., 2007), indicating that addition of GalA residue early in xylan synthesis is essential for xylan formation.

The role of pectin in human health

Various reports suggested that pectin also exerts a favorable effect on a wide spectrum of human pathological conditions. For example, pectin was reported to induce apoptosis in different types of human prostate cancer cells (Jackson et al., 2007). High-pectin diet can partially decrease gastroesophageal reflux and might improve vomiting and respiratory symptoms in children with cerebral palsy (Miyazawa et al., 2008). More recently, Zhao et al. (2008) reported the ability of modified citrus pectin to chelate harmful metals "consumed" by human. The authors claimed that treatment of Chinese children, who were hospitalized for toxic levels of lead, with modified pectin increased the urinary excretion of lead. Similarly, Khotimchenko and Kolenchenko (2006) reported that low methyl-esterified pectin can reduce the lead level in liver. The exact pectic structure used in these studies is not known, and neither is the mechanism by which pectin exerts these health benefits.

The biosynthesis of GalA-containing glycans

Using antibodies against cell wall carbohydrate epitopes, the synthesis of HGA and RG-I epitopes was proposed to initiate in the *cis*-Golgi and continues into the *medial* Golgi (reviewed by Driouich et al., 1993). Subsequent esterification of HGA was proposed to occur in the *medial* and *trans*-Golgi (Mohnen, 2008). However, since only a few enzymes involved in pectin biosynthesis have been identified or characterized so far, the exact sub-cellular localization where each step of pectin biosynthesis takes place is still unknown.

The synthesis of the pectin is predicted to require the action of at least 53 different enzymatic activities including glycosyltransferases, methyltransferases and acetyltransferases

(Ridley et al, 2001). To date, only three pectin biosynthetic enzymes have been identified and characterized, which are GAUT1 (a galacturonosyltransferase involved in the elongation of HGA) (Sterling et al., 2006), RGXT1 and RGXT2 (putative xylosyltransferases for RG-II) (Egelund et al., 2006), and XGD1 (xylosyltransferase for the synthesis of xylogalacturonan) (Jensen et al., 2008). Using reverse genetics approaches, several other putative glycosyltransferases or methyltransferases, like QUA1 (putative galacturonosyltransferase for HGA, or putative xylosyltransferase for xylan) (Bouton et al., 2002; Orfila et al., 2005), ARAD1 (putative arabinosyltransferase for RG-I) (Harholt et al., 2006), NpGUT1 (putative glucuronosyltransferases for RG-II) (Iwai et al., 2002) and QUA2 (putative HGA methyltransferase) (Mouille et al., 2007) have been proposed. Clearly, identification of all genes involved in pectin synthesis is a major undertaking.

Several groups suggested that the biosynthesis of primary cell wall polysaccharides is controlled by both the glycosyltransferases responsible for assembling the glycans and the level of activated nucleotide-sugar precursors that supply those synthases (Freshour et al., 2003; Seifert, 2004; Lerouxel et al., 2006). Thus, enzymes that synthesize the activated nucleotide-sugars are likely to play important roles in the polysaccharide biosynthesis. Such a scenario is supported by the studies of the *reb1-1* mutant that knocks out the gene encoding an isoform of UDP-glucose 4-epimerase (UGE) involved in the synthesis of UDP-galactose (UDP-Gal) (Nguema-Ona et al., 2006). In the *reb1-1* mutant, only the Gal-containing xyloglucan (XyG) structures, but not Gal-containing pectic polysaccharides, were altered compared to the wild-type plants (Nguema-Ona et al., 2006), indicating that a specific nucleotide-sugar synthase can regulate the synthesis of specific polymers.

In addition, UDP-galacturonic acid (UDP-GalA) has been reported as the activated nucleotide-sugar form of GalA that is required for the syntheses of pectin and the xylan domain sequence 1 structure. It has also been predicted that at least four galacturonosyltransferases (GAUTs) are required to transfer the GalA from UDP-GalA during the synthesis of the plant pectic polymers (Mohnen, 2002; Sterling et al., 2006). Therefore, knowledge of where UDP-GalA is synthesized in the cell, what regulates the amount of UDP-GalA produced and what determines the flux of NDP-sugars to different wall glycans is essential for better understanding the distribution of the GalA containing glycans in different types of plant cells.

PART II: BIOSYNTHESIS OF UDP-SUGARS

Introduction

UDP-sugars have been widely accepted as precursors for the syntheses of various cell wall polysaccharides, glycoproteins, and glycolipids. Three pathways are known to supply UDP-sugars in plants: the sucrose pathway, the salvage pathway, and the interconversion pathway. In the interconversion pathway that is described below, the pre-existing UDP-sugar is converted to a different UDP-sugar by specific enzymes. And the types of enzymatic modifications of UDP-sugars include 4-epimerization (i.e. interconversion of UDP-Glc and UDP-Gal), C-6 oxidation (i.e. conversion of UDP-Glc to UDP-GlcA), C-6 decarboxylation (i.e. conversion of UDP-GlcA to UDP-Xyl), C-4 keto-reductase / C-3,5 epimerization (i.e. conversion of UDP-Glc to UDP-Rha), and the interconversion of UDP-apiose from UDP-GlcA. An overview of this pathway is shown in figure 1.4 and several interconversions are described in the following sections.

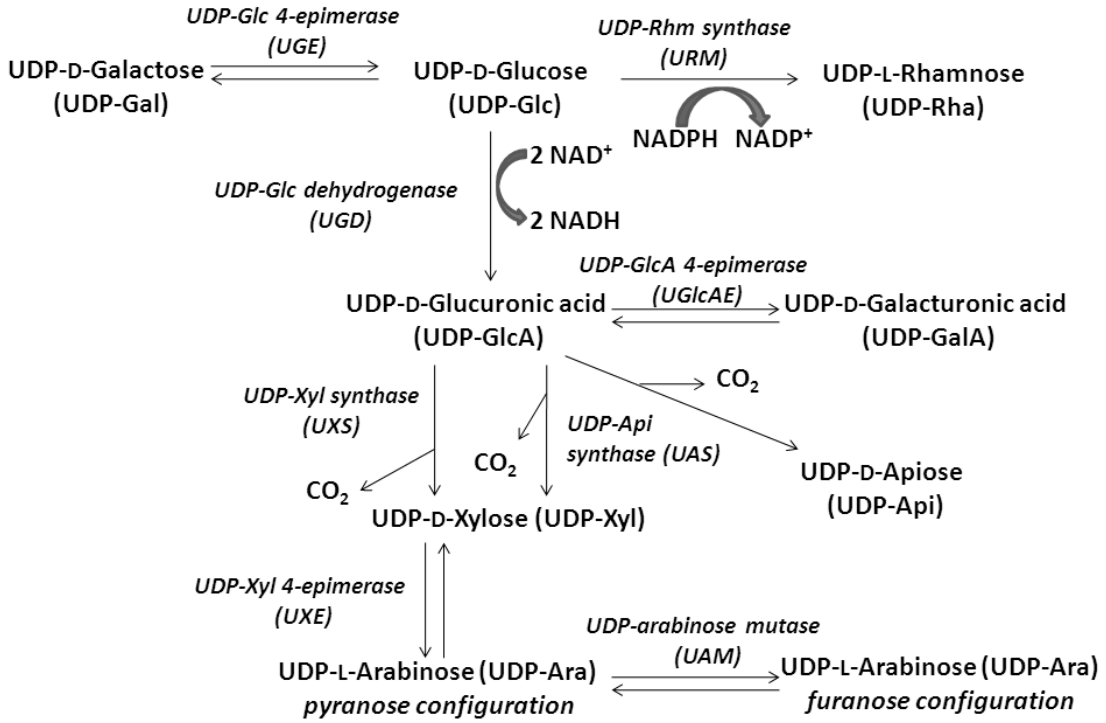


Figure 1.4. The plant UDP-sugar interconversion pathway

UDP-Glc serves as the precursor for the synthesis of UDP-Gal, UDP-GlcA and UDP-Rha. And UDP-GlcA is the precursor for UDP-GalA, UDP-Xyl and UDP-Api biosyntheses.

Interconversion of UDP-Glucose (UDP-Glc) and UDP-Galactose (UDP-Gal) (UDP-Glc↔UDP-Gal)

UDP-Glc is the most abundant UDP-sugar in plants (Feingold and Avigad, 1980) and serves not only as a precursor to UDP-Gal and UDP-GlcA, but also for the syntheses of sucrose and numerous glycosylated secondary metabolites. The interconversion of UDP-Glc and UDP-Gal is catalyzed by the reversible UDP-Glc/UDP-Gal 4-epimerase (UGE), a well-studied enzyme in all organisms. In the presence of the cofactor NAD^+, UGE oxidizes C-4 of the sugar by abstracting its hydride and then reduces the resulting UDP-4-keto-hexose from the opposite face of the sugar, conducting reversible conversion between a *gluco*- and *galacto*-configured pyranose ring (reviewed by Field and Naismith, 2003).

Multiple UGE isoforms have been identified in Arabidopsis (Seifert, 2004; Barber et al., 2006) and barley (Zhang et al., 2006). It has been reported (Seifert, 2004; Barber et al., 2006) that the five Arabidopsis UGE isoforms have distinct enzymatic efficiencies and show different requirements for exogenous NAD^+. In addition, although all UGE isoforms are expressed in every plant organ, subtle variations exist between the expression patterns of different isoforms. For example, while UGE1 and UGE3 expression patterns globally resemble enzymes involved in carbohydrate catabolism, the expression of UGE2, 4, and 5 is more related to carbohydrate biosynthesis (Barber et al., 2006). Interestingly, while UGE1, 2, and 4 are reported to be present in the cytoplasm, UGE4 is observed to be additionally enriched to Golgi stacks (Barber et al., 2006). The different enzymatic properties, gene expression patterns, and subcellular localizations observed for UGE isoforms were suggested to contribute to the differentiation of isoform functions (Seifert, 2004; Barber et al., 2006).

Conversion of UDP-Glc to UDP-GlcA (UDP-Glc→UDP-GlcA)

UDP-GlcA has been demonstrated to be the sugar donor for the syntheses of various GlcA-containing glycans. It is also the precursor for UDP-GalA, UDP-Xyl and UDP-Api biosyntheses, all of which are key donors for the syntheses of various plant cell wall polysaccharides including pectin and xylan.

UDP-GlcA is formed from UDP-Glc by UDP-Glc dehydrogenase (UGD). In the presence of two molecules of NAD^+, UGD oxidizes the C-6 of UDP-Glc, resulting in UDP-GlcA and two molecules of NADH. The gene encoding UGD was first described in soybean (Tenhaken and Thulke, 1996) and subsequently, multiple UGD isoforms were identified in Arabidopsis (Klinghammer et al., 2007) and maize (Karkonen and Fry, 2006). Four UGD isoforms are present in the Arabidopsis genome; each shows a distinct tissue-specific expression pattern during plant development (Klinghammer et al., 2007). The four Arabidopsis UGD isoforms show similar affinities for NAD^+ but different catalytic constants. Klinghammer et al. (2007) therefore suggested that each UGD isoform has a specific biological role (Klinghammer et al., 2007).

Conversion of UDP-GlcA to UDP-Xylose (UDP-Xyl) (UDP-GlcA→UDP-Xyl)

UDP-Xyl is a nucleotide sugar required for the synthesis of diverse xylose-containing plant cell wall polysaccharides including xyloglucan and RG-II (O'Neill and York, 2003; Mohnen et al, 2008). UDP-Xyl is synthesized from UDP-GlcA by UDP-GlcA decarboxylase (also named as UDP-Xylose Synthase, UXS) (Harper and Bar-Peled, 2002). Feingold and Avigad (1980) proposed that in the presence of NAD^+ bound enzyme, the C-4 proton of UDP-GlcA is first abstracted to yield a UDP-4-keto-hexuronate intermediate, which is further decarboxylated and forms a UDP-4-keto-pentose, and releases COOH residue as CO_2.

Subsequently, the NADH bound enzyme from the previous oxidation reaction reduces the UDP-4-keto-pentose to UDP-Xyl.

Several distinct UXS isoforms have been identified in Arabidopsis (Harper and Bar-Peled, 2002), rice (Suzuki et al., 2003) and barley (Zhang et al., 2005). All the characterized plant UXS isoforms are active as dimers and their activities are inhibited by UDP-Xyl. Phylogenetic analysis indicates that plant UXS isoforms can be divided into three distinct evolutionary clades: type A, B or C (reviewed by Mohnen et al, 2008) and they are predicted to be localized either in the cytoplasm or in the Golgi (Harper and Bar-Peled, 2002), which is consistent with the notion that UDP-Xyl is formed in the cytosol and in the endomembrane system. The different subcellular localizations of UXS isoforms were confirmed by expressing the GFP-tagged UXS2 or UXS3 in tobacco leaves, in which UXS2-GFP was targeted to the Golgi while UXS3-GFP was in the cytosol (Pattathil et al., 2005). All the characterized plant UXS isoforms do not require exogenous NAD^+ (Harper and Bar-Peled, 2002) and Zhang et al., (2005) have shown that NAD^+ is tightly bound to the enzyme. Such tight binding, however, may not be the case for UXS isolated from *Cryptococcus laurentii* because this fungal enzyme required exogenous NAD^+ for decarboxylase activity (Ankel and Feingold, 1966).

Conversion of UDP-GlcA to UDP-Apiose (UDP-APi) (UDP-GlcA→UDP-Api)

UDP-Apiose synthase (UAS) catalyzes the synthesis of UDP-Api from UDP-GlcA. Apiose residues serve as the binding site for the borate cross-linking of RG-II in the plant cell wall (O'Neill et al., 2004). Growth defects and cell death were observed in plant mutants lacking apiose biosynthetic genes (Ahn et al., 2006). UDP-Api is formed from UDP-GlcA by a single enzymatic reaction that leads to the decarboxylation of the substrate followed by re-arrangement of the carbon skeleton and ring contraction (Feingold and Avigad, 1980; Mølhøj et al., 2003;

Guyett et al., unpublised). The initial steps of this reaction are believed to be identical to the synthesis of UDP-Xyl, in which, with NAD^+, the C-4 proton of UDP-GlcA is abstracted to yield a UDP-4-keto-hexuronate intermediate and subsequently decarboxylated to form a UDP-4-keto-pentose intermediate and release CO_2. The exact re-arrangement mechanism of the UDP-4-keto-pentose to UDP-Api still remains unclear, and current effort to crystalize the enzyme with substrates is underway to address this issue.

UAS homologs have been identified in different plant species including tobacco (Ahn et al., 2006), Arabidopsis (Molhoj et al., 2003) and potato (Guyett et al., unpublised). *In vitro*, UAS is also reported to convert a portion of the UDP-GlcA to UDP-Xyl. Recent *in vitro* NMR time course assays of the conversion of UDP-GlcA to UDP-pentose by recombinant potato enzyme indicate that UDP-Api synthesis occurs prior to UDP-Xyl synthesis (Guyett et al., unpublised). Therefore, this "dual function" enzyme is also named as UDP-Apiose/UDP-Xylose synthase (AXS). Furthermore, UDP-Api is not a stable molecule and is degraded to cyclic apiose 1,2-phosphate and UMP as shown by NMR studies described above (Guyett et al., unpublised).

Interconversion of UDP-Xyl to UDP-Arabinose (UDP-Ara) (UDP-Xyl↔UDP-Ara)

In plants, L-Ara residue is found in two distinct sugar configurations: the pyranose and furanose forms. Arabinose is an important constituent of cell wall polysaccharides including RG-I, RG-II and the hemicellulose glucuronoarabinoxylans (O'Neill and York, 2003). L-Ara is also a major component of hydroxyproline-rich glycoproteins (HRGPs) (Stafstrom and Staehelin, 1986) and arabinogalactan proteins (AGPs) (Gaspar et al., 2001).

UDP-L-Ara (pyranose configuration) is synthesized de novo from UDP-Xyl via 4-epimerization (Feingold and Avigad, 1980; Seifert, 2004). An Arabiodpsis mutant (*mur4*) showing a 50% reduction of arabinose in the wall was identified and shown to be partially

defective in the last step of L-Ara synthesis (Burget and Reiter, 1999). The corresponding MUR4 gene was cloned and subsequently demonstrated to encode an active UDP-Xyl 4-epmierase (UXE1), which reversibly converts UDP-Xyl to UDP-Ara (Burget et al., 2003). UXE1 was predicted to be a Type II membrane protein with its catalytic domain facing the lumen and reported to be localized in the Golgi apparatus (Burget et al., 2003). So far, two isoforms (UXE1, At1g30620 and UXE2, At2g34850) exist in the Arabidopsis genome and multiple isoforms have been identified in rice and barley (Reviewed by Mohnen et al, 2008). Recently, Konishi et al., (2007) have cloned a rice UDP-arabinose mutase (UAM) gene and demonstrate that the recombinant enzyme can reversibly interconvert the UDP-Ara pyranose form to its furanose form.

Interconversion of UDP-GlcA and UDP-GalA (UDP-GlcA↔UDP-GalA)

In plants, the formation of UDP-GalA by a reversible 4-epimerase (abbreviated herein UGlcAE) from UDP-GlcA was reported in various plants (Feingold et al., 1960; Dalessandro and Northcote, 1977a, 1977b; Liljebjelke et al., 1995; Orellana and Mohnen, 1999) and the activity of plant UGlcAE was reported to reside in the Golgi (Sterling et al., 2001). Similar to the 4-epimerization of UDP-Glc and UDP-Xyl, the conversion of UDP-GlcA to UDP-GalA is believed to proceed via a transient 4-keto intermediate, which is generated by an enzyme-bound pyridine nucleotide cofactor (i.e. NAD^+) (Feingold and Avigad, 1980).

At the time when I started my PhD, none plant UGlcAE proteins were purified and the corresponding genes were unknown. As a consequence, several questions also remain unclear:

1) How specific are these 4-epimerase enzymes?

2) Are activities of these enzymes regulated by other nucleotides or nucleotide-sugars?

3) What determines the amount of UDP-GalA produced in a plant cell?

4) Where is UDP-GalA made in the cell?

5) How does the supply of UDP-GalA regulate the formation of GalA containing glycans?

The first breakthrough leading to the isolation of plant UGlcAEs (that will be described in my thesis) came from the work on a bacterial polysaccharide mutant. A bacterial cap1J was isolated from a capsular polysaccharide deficient strain type I of *Streptococcus pneumoniae* (Munoz et al., 1997). Later work, including cloning, expression and characterization of the cap1J led to the identification of a bacterium UDP-GlcA 4-epimerase that converted UDP-GlcA to UDP-GalA (Munoz et al., 1999). Munoz et al. (1999) also confirmed that the enzyme tightly bound NAD^+.

To gain a better understanding of the inter-conversion between UDP-GalA and UDP-GlcA in plants, my Ph.D. research focused on the identification of plant UGlcAE genes and the characterization of these proteins at the biochemical, cellular and genetic levels, which will be introduced in details in the following chapters.

PART III: THESIS OVERVIEW

Chapter 2 describes the cloning, purification and characterization of an Arabidopsis UGlcAE isoform, AtUGlcAE1, and the confirmation of its epimerase activity by HPLC and NMR analyses. Chapter 3 describes the functional identification of UGlcAE homologs in maize and rice. We conclude in this chapter that there are three different types of UGlcAEs among monocot and Arabidopsis UGlcAE homologs (type A, B, and C). Chapter 3 further compares the enzymatic properties of the monocot and Arabidopsis UGlcAE homologs. Chapter 4 addresses where UDP-GalA is produced and how UDP-GalA is converted from UDP-GlcA. The presented data in Chapter 4 suggest that the catalytic domains of UGlcAE type A, B, and C are facing inside the Golgi lumen. The detailed phylogenetic analyses of plant UGlcAEs are described in

Appendix A and the transcripts of Arabidopsis UGlcAE isoforms are described in Appendix B. Appendix C summarizes the initial experiments performed to verify that the NAD^+ was tightly bound to the UGlcAE.

CHAPTER 2

THE BIOSYNTHESIS OF UDP-GALACTURONIC ACID IN PLANTS. FUNCTIONAL CLONING AND CHARACTERIZATION OF ARABIDOPSIS UDP-D-GLUCURONIC ACID 4-EPIMERASE[1]

[1]Gu X and Bar-Peled M. 2004, Plant Physiology, **136**: 4256–4264 (reprinted here with permission of publisher). www.plantphysiol.org.

ABSTRACT

UDP-GlcA 4-epimerase (UGlcAE) catalyzes the epimerization of UDP-α-D-glucuronic acid (UDP-GlcA) to UDP-α-D-galacturonic acid (UDP-GalA). UDP-GalA is a precursor for the synthesis of numerous cell-surface polysaccharides in bacteria and plants. Using a biochemical screen, a gene encoding AtUGlcAE1 in Arabidopsis (*Arabidopsis thaliana*) was identified and the recombinant enzyme was biochemically characterized. The gene belongs to a small gene family composed of six isoforms. All members of the UGlcAE gene family encode a putative type-II membrane protein and have two domains: a variable N-terminal region approximately 120 amino acids long composed of a predicted cytosolic, transmembrane, and stem domain, followed by a large conserved C-terminal catalytic region approximately 300 amino acids long composed of a highly conserved catalytic domain found in a large protein family of epimerase/dehydratases. The recombinant epimerase has a predicted molecular mass of approximately 43 kD, although size-exclusion chromatography suggests that it may exist as a dimer (approximately 88 kD). AtUGlcAE1 forms UDP-GalA with an equilibrium constant value of approximately 1.9 and has an apparent Km value of 720 μM for UDP-GlcA. The enzyme has maximum activity at pH 7.5 and is active between 20^{o}C and 55^{o}C. Arabidopsis AtUGlcAE1 is not inhibited by UDP-Glc, UDP-Gal, or UMP. However, the enzyme is inhibited by UDP-Xyl and UDP-Ara, suggesting that these nucleotide sugars have a role in regulating the synthesis of pectin. The cloning of the AtUGlcAE1 gene will increase our ability to investigate the molecular factors that regulate pectin biosynthesis in plants. The availability of a functional recombinant UDP-GlcA 4-epimerase will be of considerable value for the facile generation of UDP-D-GalA in the amounts required for detailed studies of pectin biosynthesis.

INTRODUCTION

GalA is a major sugar residue of plant pectic polysaccharides (Mohnen, 2002) and a minor component of some plant arabinogalactan proteins (Darvill et al., 1980; Yates et al., 1996). GalA is also found in various cell-surface polysaccharides of different Gram-negative bacteria, including human pathogenic bacteria *Klebsiella pneumoniae* (for review, see Holst et al., 1998), *Shigella* spp. (Feng et al., 2004), and *Vibrio cholera* (Adeyeye et al., 2003); plant pathogenic bacteria *Agrobacterium larrymoorei* (Molinaro et al., 2003) and *Erwinia chrysanthemi* spp. (Gray et al., 2000); plant symbiotic rhizobacteria (Forsberg and Carlson, 1998); aerobic bacteria from the deep sea (Raguenes et al., 2003); and the hyperthermophilic bacterium *Aquifex pyrophilus* (Plotz et al., 2000). The existence of GalA is not restricted to Gram-negative bacteria being present in the capsular polysaccharide of the Gram-positive human pathogen type-I *Streptococcus pneumoniae* (Garcia et al., 1997; Stroop et al., 2002; for review, see Jann and Westphal, 1975), as well as in certain polysaccharides of the *Cyanobacterium* spp. (Gloaguen et al., 1995).

Genetic and biochemical evidence demonstrates that the synthesis of GalA-containing polymers requires UDP-α-D-galacturonic acid (UDP-GalA). A recent review by Mohnen (2002) predicted that synthesis of the major pectin polymers homogalacturonan, rhamnogalacturonan I, and rhamnogalacturonan II and the minor polymers xylogalacturonan and apiose-galacturonan requires more than 10 distinct enzymes (galacturonosyltransferases), each incorporating GalA from its activated nucleotide-sugar form, UDP-GalA, to a unique polymer. In 1958, Neufeld et al. reported the first isolation of a specific nucleotide-sugar 4-epimerase activity that converts UDP-α-D-glucuronic acid (UDP-GlcA) to UDP-GalA (Neufeld et al., 1958). Subsequently, the activity was isolated from *Cyanobacterium anabaena flos-aquae* (Ankel and Tischer, 1969; Gaunt et al.,

1974) and from different plants (Feingold et al., 1960; Dalessandro and Northcote, 1977a, 1977b; Liljebjelke et al., 1995; Orellana and Mohnen, 1999). A genetic screen of a polysaccharide-deficient *S. pneumoniae* strain identified the cap1J gene (Munoz et al., 1997) as encoding the active UDP-GlcA epimerase (Munoz et al., 1999). We have developed methods to identify the function of various cDNAs encoding putative nucleotide-sugar epimerases. Here, we report the isolation of a cDNA from Arabidopsis (*Arabidopsis thaliana*) that encodes a protein sharing high sequence similarity to the cap1J protein, and provide the biochemical proof for the identification of a plant UDP-GlcA epimerase that we named AtUGlcAE1 (for Arabidopsis UDP-GlcA 4-epimerase 1).

RESULT

Cloning and Characterization of AtUGlcAE from Arabidopsis

With the functional cloning of the prokaryotic, human, and plant UDP-Gal epimerases (Lemaire and Muller-Hill, 1986; Daude et al., 1995; Dormann and Benning, 1996) and the putative UDP-GlcA epimerase (Garcia et al., 1993; Munoz et al., 1997), it became apparent that most, if not all, nucleotide-sugar 4-epimerases belong to a large family of proteins that have conserved catalytic motifs (Thoden et al., 1997). These motifs include the GxxGxxG (x = any amino acid) sequence within the N-terminal domain that is likely involved in the binding of β-NAD^+, and a catalytic triad consisting of S/T and YxxxK (Weirenga et al., 1986). These conserved domains are common in nucleotide-sugar 4-epimerases from eukaryotes and prokaryotes. Given the size of the gene family among all organisms, it remains a challenge to predict the substrate specificity of a particular gene product based on the gene/protein sequence alone. As a result, we have developed an HPLC-based biochemical screen to identify the

Table 2.1. Molecular data for the three isoforms of the Arabidopsis AtUGlcAE gene family

[a]The amino acid sequence identity/similarity comparison (Altschul et al., 1997) was performed using amino acids spanning from 97 to 437 of AtUGlcAE1. [b]The type is defined arbitrarily to indicate that the 3 isoforms described have a variable amino acid region spanning from amino acid 1 to approximately 100.

AtUGlcAE	Amino Acid	Molecular Mass (kD)	Loci	%[a] Similarity/ Identity	Type[b]
AtUGlcAE1	437	48.5	At2g45310	100/100	A
AtUGlcAE2	460	47.4	At3g23820	78/89	B
AtUGlcAE3	429	50.5	At4g30440	78/88	C

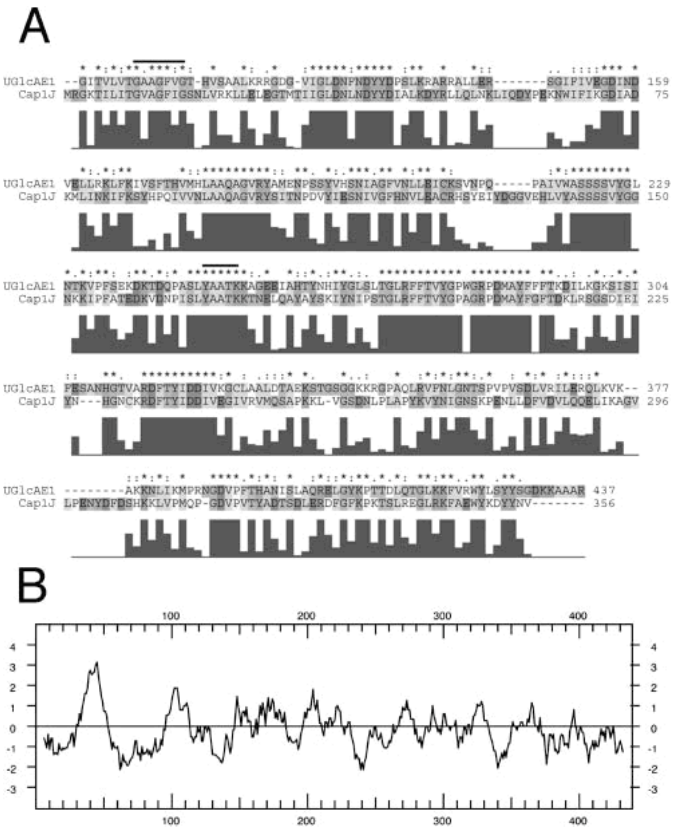


Figure 2.1. Amino acid sequence alignment and characterization of Arabidopsis AtUGlcAE1.

A, Comparison of the conserved amino acid sequence between AtUGlcAE1 (GenBank accession no. AAT06796) and bacterial cap1J (accession no. 1870164) protein. Protein sequence alignment was performed using ClustalX (version 1.83; Jeanmougin et al., 1998). The putative conserved GXXGXXG motif and the putative catalytic triad YXXXK (Weirenga et al., 1986) are marked with a line. The histograms beneath the sequence indicate the degree of amino acid identity/similarity between the two proteins as determined by ClustalX.

B, Hydropathy plot of AtUGlcAE1.

function of these putative epimerases. The amino acid sequence of the cap1J (Munoz et al., 1997) was useful to initially identify candidate gene homologs from the various Arabidopsis expressed

sequence tag cDNA database (dbEST) projects and available genomic data. Using sequence data and biochemical reports that UDP-GlcA epimerases from several plants are membrane associated (Feingold et al., 1960), we cloned several gene candidates from an Arabidopsis cDNA library (Table 2.1). We identified three distinct putative UGlcAE cDNAs each encoding a protein having two regions: (1) an N-terminal region (approximately 120 amino acids long) that shares no sequence similarity with known proteins and contains a hydrophobic domain (Fig. 2.1B) predicted to have a type- II membrane orientation; and (2) a large region (approximately 300 amino acids long) that is evolutionarily conserved, sharing more than 64% amino acid sequence similarity to the cap1J prokaryote UDP-GlcA epimerase (Munoz et al., 1999), and contains the catalytic epimerase domains (Fig. 2.1A). To determine the function of the cap1J-like Arabidopsis homologs, we overexpressed one of them in *Escherichia coli* and studied the properties of the recombinant protein.

AtUGlcAE1 Encodes Active Recombinant UGlcAE

Recombinant cap1J UDP-GlcA epimerase was shown to be difficult to express and purify, and only partial purification was achieved due to loss of activity (Munoz et al., 1999). We were unable to recover active recombinant enzymes when full-length AtUGlcAE1 and AtUGlcAE2 proteins were expressed in *E. coli*, presumably due to the hydrophobic domain that results in un- or misfolded protein in inclusion bodies within *E. coli*. We have experienced similar problems when trying to actively express other recombinant plant proteins that contain a putative trans-membrane domain (Harper and Bar-Peled, 2002). Recombinant AtUGlcAE1$_{\Delta 1-64}$, lacking the putative transmembrane domain, on the other hand, could be expressed in *E. coli* (Fig. 2.2A, lane 2) and was active (Fig. 2.2B, subsection 2) and relatively stable when stored as a crude extract. *E. coli* cells expressing the empty control vector (Fig. 2.2A, lane 3) had no UDP-GlcA epimerase

activity, indicating that the observed epimerase activity was exclusively due to the expression of the recombinant plant gene (Fig. 2.2B, subsection 3). Recombinant Arabidopsis AtUGlcAE2, lacking its putative membrane domain (amino acids 1–69), was also expressed in *E. coli*, but the activity was too low to permit further characterization.

The active 43-kD AtUGlcAE1$_{\Delta 1–64}$, expressed in *E. coli*, was partially purified by anion-exchange chromatography (Fig. 2.2, A, lane 4, and B, subsection 5). Purification of the active recombinant protein on a Ni column could not be used because the imidazole used to elute such a column completely inhibits AtUGlcAE1$_{\Delta 1–64}$ activity. AtUGlcAE1$_{\Delta 1–64}$, readily converts UDP-GlcA to a product that elutes from HPLC column at a retention time of 22.8 min (Fig. 2.3, subsection 2). This conversion did not require exogenous NAD^+ (Fig. 2.3, subsection 3). To determine the identity of the product produced by the recombinant protein, the product eluting from the HPLC column was collected and analyzed by proton ^{1}H-NMR. The 500-mHz ^{1}H-NMR spectrum (Fig. 2.4, A and B) confirmed that the enzymatic product is UDP-GalA. The spectrum contains a doublet-of-doublet signal at 5.6 ppm with a coupling constant, $^3J_{1",P'}$, of 7 Hz, diagnostic for the H-1" proton of an α-D-GalA residue that is linked to the phosphate group of uridine. The coupling constant $^3J_{1",2"}$ equal to 3.7 Hz value is diagnostic for coupling between protons H-1" and H-2" in UDP-GalA. Other specific chemical shifts and coupling constant values (Table 2.2) diagnostic for GalA are the large 10.2-Hz coupling constant between the trans-configuration of protons 2" and 3" and the short 3.2-Hz signal corresponding to coupling constants between the *cis*-configuration of protons 3" and 4". The uridine proton assignments in the ^{1}H-NMR spectra are characteristic of those for uridine in other UDP sugars (Harper and

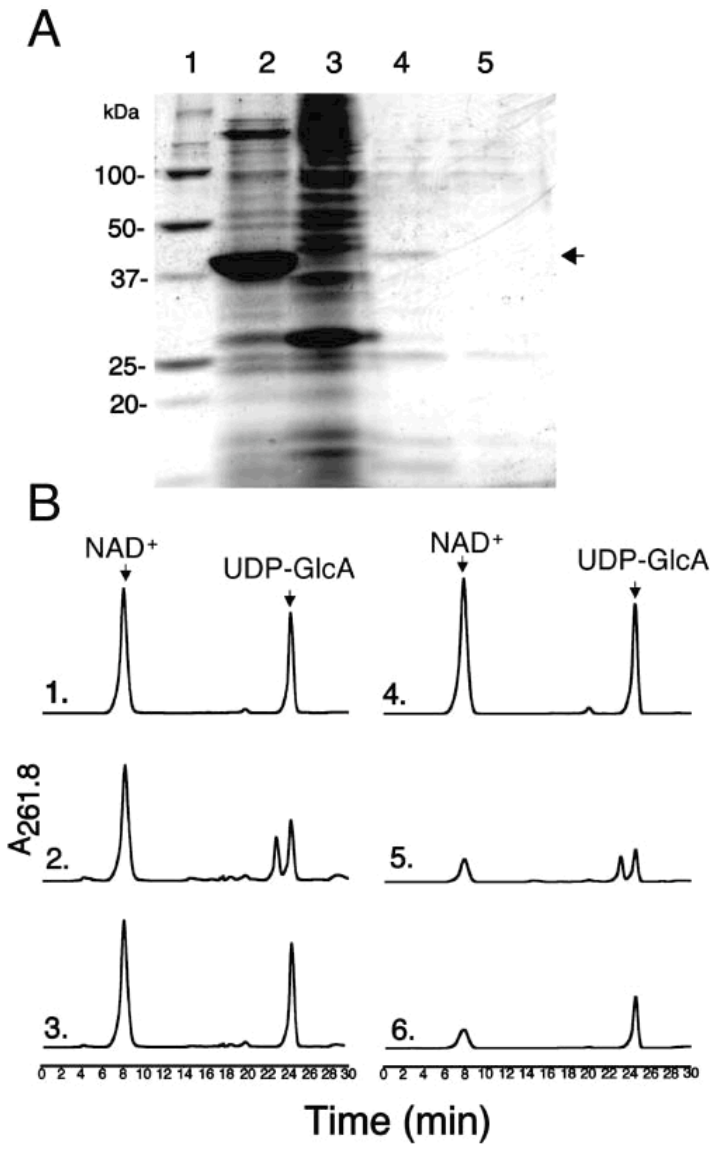


Figure 2.2. Isolation, chromatography, and initial characterization of recombinant AtUGlcAE1$_{\Delta 1–64}$.

A, SDS-PAGE of total protein isolated from *E. coli* cells expressing AtUGlcAE1$_{\Delta 1–64}$ (lane 2) or vector control (lane 3). Proteins from cells expressing recombinant AtUGlcAE1$_{\Delta 1–64}$ were separated by chromatography (see ‘‘Materials and Methods’’), and the active fractions eluted from the SourceQ column (fractions 60–65) were combined, desalted, and separated on SDS-PAGE (lane 4). As a control, total proteins from an *E. coli* cell expressing vector alone were

separated under the same chromatographic conditions and the combined desalted fractions (60–65) eluting from SourceQ column were separated on SDS-PAGE as well (lane 5). The arrow indicates the position of the partially purified recombinant AtUGlcAE1$_{\Delta 1–64}$.

B, Aliquot of each protein sample (total or from SourceQ column) was incubated with β-NAD$^+$ and UDP-GlcA, and the enzymatic reactions were separated on ion-exchange HPLC column. The arrows in subsections 1 and 4 indicate migration of standards β-NAD$^+$ and UDP-GlcA. Activity of total protein isolated from cell expressing recombinant AtUGlcAE1$_{\Delta 1–64}$ (subsection 2) or vector alone (subsection 3) is shown. The activity of partially purified recombinant AtUGlcAE1$_{\Delta 1–64}$ (subsection 5) or control vector alone (subsection 6), fractionated on SourceQ column, was determined with 0.2 mM NAD and 0.5 mM UDP-GlcA.

Bar-Peled, 2002; Watt et al., 2004), including UDP-Xyl and UDP-rhamnose. Having confirmed that the enzyme product of AtUGlcAE1$_{\Delta 1-64}$ is UDP-GalA, we tested the ability of the recombinant enzyme to convert UDP-GalA to UDP-GlcA. The data shown in Figure 2.3, subsection 4, demonstrates that AtUGlcAE1$_{\Delta 1-64}$ is in fact a reversible 4-epimerase capable of converting UDP-GalA to UDP-GlcA. This conversion also did not require exogenous NAD^+ (Fig. 2.3, subsection 5). To determine the equilibrium constant for the interconversion, the enzyme was incubated with UDP-GlcA for various times, and the UDP-GalA/UDP-GlcA ratio was determined. The data demonstrate that recombinant AtUGlcAE1$_{\Delta 1-64}$ favors the formation of UDP-GalA over UDP-GlcA with an equilibrium constant of approximately 1.9 (Table 2.4).

Further testing demonstrated that the AtUGlcAE1$_{\Delta 1-64}$ is specific for UDP-uronic acids, and none of the other nucleotide sugars examined (UDP-Glc, UDP-Gal, UDP-Ara, UDP-Xyl, CDP-Glc, and GDP-Man) are substrates for AtUGlcAE1$_{\Delta 1-64}$. Based on these product analyses and substrate-specificity studies, we have functionally cloned a plant gene whose product is involved in 4-epimerization of UDP-GlcA. We accordingly suggest naming the gene AtUGlcAE1, for Arabidopsis UDP-GlcA epimerase 1.

Enzymatic Characterization of AtUGlcAE1$_{\Delta 1-64}$

Recombinant AtUGlcAE1$_{\Delta 1-64}$ does not require metal ions for activity, and the enzyme maintains maximum activity in the presence of EDTA (Table 2.3). Glycerol, at 20% (v/v), was found to increase activity significantly, while dimethyl sulfoxide has no effect. High concentration of monovalent salts (KCl and NaCl, 300 mM), on the other hand, inhibit the activity and cause complete inactivation of the enzyme when stored at -20^{o}C for 2 weeks. However, storing the enzyme with glycerol sustains full AtUGlcAE1$_{\Delta 1-64}$ activity. The enzyme is

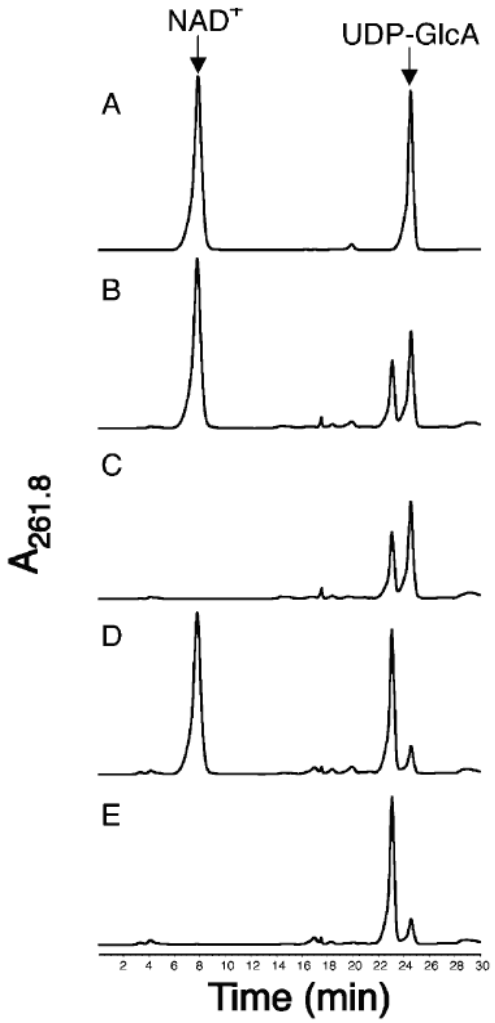


Figure 2.3. Recombinant AtUGlcAE1$_{\Delta 1–64}$ is an active UDP-GlcA/UDP-GalA 4 epimerase. The HPLC separation of standards (NAD^+ and UDP-GlcA, indicated by arrow in A) and of UDP-GlcA epimerase enzymatic products (B–E) is shown. Recombinant AtUGlcAE1$_{\Delta 1–64}$ was incubated with UDP-GlcA and 1 mM NAD^+ (B), with UDP-GlcA but without NAD^+ (C), with UDP-GalA and 1 mM NAD^+ (D), or with UDP-GalA but without NAD^+ (E).

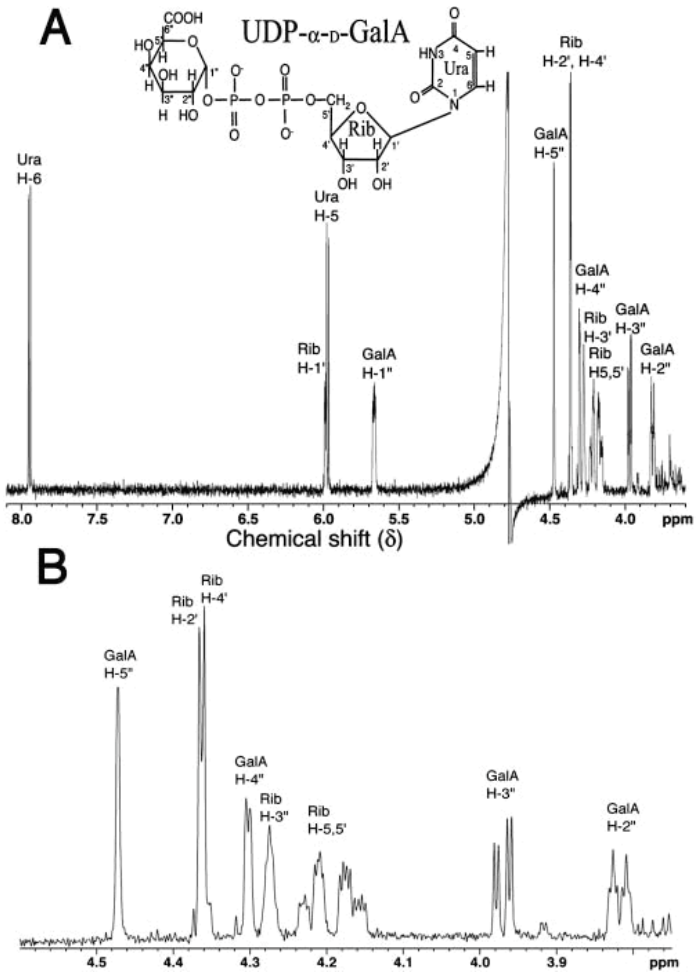


Figure 2.4. The ^{1}H-NMR spectrum of the product generated by reacting UDP-GlcA with recombinant AtUGlcAE1$_{\Delta 1-64}$.

A, One-dimensional 500-mHz NMR spectrum of AtUGlcAE1$_{\Delta 1-64}$ product (UDP-a-D-GalA) is shown between 3.5 to 8 ppm. Location for each proton residue on the spectrum is indicated: the position of protons on the uracil (Ura) ring are indicated by H, the ribose (Rib) protons by H', and the GalA protons are in H".

B, An expanded region between 3.8 and 4.6 ppm is shown.

Table 2.2. Proton chemical shifts and coupling constants of UDP-GalA formed from UDP-GlcA by recombinant AtUGlcAE1$_{\Delta 1-64}$.

[a]Chemical shifts are in ppm relative to internal acetone signal set at 2.218 ppm. GalA proton-proton coupling constants in hertz are indicated as well as the $^{3}J_{1",P}$ coupling values between phosphate and the H-1" proton of GalA. The chemical shift values for uracil protons (H) and ribose protons (H') are the same as for other UDP-sugars.

Proton	H1	H2	H3	H4	H5	H6
GalA						
Chemical shifts, δ (ppm[a])	5.664	3.818	3.970	4.302	4.472	-
J coupling constants (Hz)	$J_{1",P}$ 7 $J_{1",2"}$ 3.5	$J_{2",3"}$ 10.2 $J_{2",P}$ 3	$J_{3",4"}$ 3.2	$J_{4",5"}$ <1	-	-
Rib						
Chemical shifts, δ (ppm[a])	5.984	4.36	4.272	4.36	4.266–4.22	–
J coupling constants (Hz)	$J_{1',2'}$ 3.6	-			$J_{5,5'}$ 12	-
Uracil						
Chemical shifts, δ (ppm[a])	-	-	-	-	5.972	7.948
J coupling constants (Hz)	-	-	-	-	$J_{5,6}$ 8.1	-

active over a relatively broad pH range with activity observed between pH 6.8 and 8.2, and has maximal activity at a pH of 7.4 to 7.6 (Table 2.4) in phosphate buffer. The activity is completely abolished at pH values lower than 4 and above 9.5 in phosphate buffer (data not shown). AtUGlcAE1$_{\Delta 1-64}$ is active between 25 ^{o}C and 55 ^{o}C, with maximal activity between 30 ^{o}C and 42^{o}C (Table 2.4).

Under the assay conditions described, the reaction rate (calculated as the amount of UDP-GalA produced from UDP-GlcA) is linear with time for up to 50 min. Kinetic studies were carried out for 15 min, and the apparent Km value for UDP-GlcA was 0.72 mM. The K_{cat} of the recombinant enzyme was 24 (s^{-1}) with a catalytic efficiency [K_{cat}/Km] value of 3.2 X 10^4 ($M^{-1}s^{-1}$; Table 2.4).

To further compare the AtUGlcAE1$_{\Delta 1-64}$ to UDP-GlcA epimerases isolated from other species, the activity of AtUGlcAE1$_{\Delta 1-64}$ was tested with various nucleotides and nucleotide sugars (Fig. 2.5). Unlike the strong inhibition observed with UDP-Glc and UDP-Gal for the *Algal* (Gaunt et al., 1974) and the Streptococcal (Munoz et al., 1999) UDP-GlcA epimerases, these nucleotide sugars had no effect on the activity of the Arabidopsis enzyme (Fig. 2.5). However, UDP-Xyl and UDP both strongly inhibited AtUGlcAE1$_{\Delta 1-64}$ by 60% (Fig. 2.5), while UMP had negligible effect on the activity of the enzyme. The 4-epimer of UDP-Xyl, UDP-Ara, also inhibited the activity of AtUGlcAE1$_{\Delta 1-64}$.

The Recombinant AtUGlcAE1$_{\Delta 1-64}$ May Function as a Dimer

Gaunt et al. (1974) estimated the M_r of the native *Algal* UDP-GlcA epimerase to be 54,000. The product of the recombinant Streptococcal cap1J gene (encoding a polypeptide of approximately 40 kD) was estimated by gel filtration to have a M_r of 80,000 (Munoz et al., 1999),

Table 2.3. Effect of various additives on AtUGlcAE1$_{\Delta 1-64}$ activity.

AtUGlcAE1$_{\Delta 1-64}$ was separately mixed with each additive for 30 min on ice. One millimolar UDP-GlcA was added and the reactions were incubated for 30 min at 30^{o}C. The amounts of UDP-GalA were determined by HPLC. Data are the average relative amounts of UDP-GalA produced compared to the control water (no additives). Each value is the mean of duplicate reactions, and the values varied by no more than 5%.

Additive and Concentration in the Assay	Relative AtUGlcAE1$_{\Delta 1-64}$ Activity (%)
Water	100
KCl (200 mM)	64
NaCl (300 mM)	36
EDTA (1 mM)	98
DTT (2 mM)	91
MgCl2 (2 mM)	96
Glycerol (20%, v/v)	146
Dimethyl sulfoxide (4%)	100

Table 2.4. Enzymatic properties and characterization of recombinant AtUGlcAE1$_{\Delta 1–64}$.

[a]At the pH value indicated, the relative activity in phosphate buffer is 100%; in Tris-HCl, 80%; in Bis-Tris-HCl, 75%. [b]Optimal temperature assays were conducted in phosphate buffer. [c]The reciprocal initial velocity was plotted against UDP-GlcA concentration according to Lineweaver and Burk to calculate Km value. [d]The conversion ratio between UDP-GalA to UDP-GlcA was determined over a period of 2 h. [e]The molecular mass of the active AtUGlcAE1$_{\Delta 1–64}$ eluted from Superdex 75 gel-filtration column (17.8 min) was estimated based on extrapolation ($R^2 = 0.996$) from relative time of standard protein marker.

Optimal pH[a]	Optimal Temperature[b] (^oC)	Km[c] (mM)	K_{cat} (S^{-1})	Catalytic Efficiency (M^{-1} S^{-1})	Equilibrium Constant[d]	Mass of Active Protein[e] (Denatured) (kD)
7.4–7.6	30–42	0.72	24	3.2 X 10^4	1.9	88 (43.5)

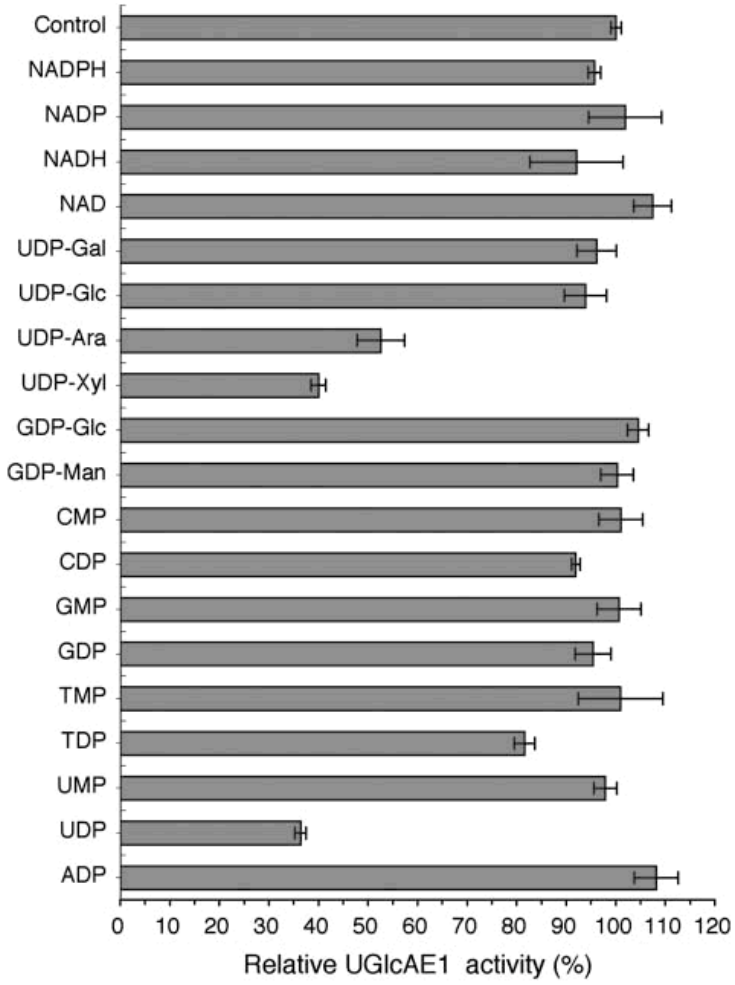


Figure 2.5. The effects of nucleotides and nucleotide sugars on the activity of the recombinant AtUGlcAE1$_{\Delta 1–64}$.

UDP-GlcA (1 mM) was reacted for 30 min with recombinant AtUGlcAE1D1–64 in the absence or presence of different nucleotides and nucleotide sugars (2 mM). Note that 100% activity corresponds to 30 nmol of UDP-GalA produced. The data are the average relative amount of UDP-GalA produced compared to the control from two experiments, and each value is the mean of triplicate reactions.

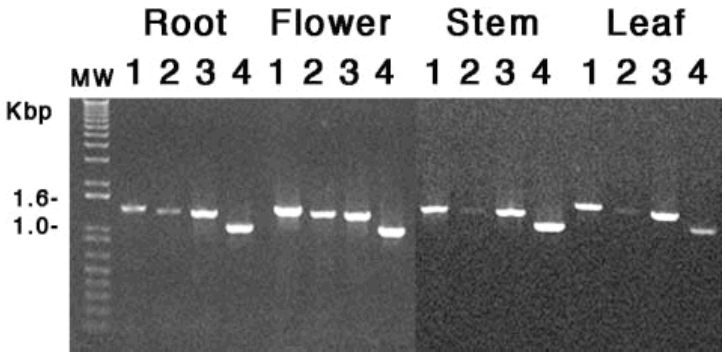


Figure 2.6. Gene expression of three AtUGlcAE isoforms in Arabidopsis.

Total RNA was extracted from root, flower, stem, and leaf and used for a RT-PCR reaction with gene-specific primers. The lane labeled 1 in each section is the AtUGlcAE3 (At4g30440) transcript, lane 2 is the AtUGlcAE1 (At2g45310) transcript, and lane 3 indicates AtUGlcAE2 (At3g23820). UXS3 (Harper and Bar-Peled, 2002) was used as a control (labeled as lane 4).

suggesting that the prokaryotic enzyme is a dimer. The recombinant Arabidopsis AtUGlcAE1$_{\Delta 1-64}$ has a M_r of 43,500. To determine the M_r of the active AtUGlcAE1$_{\Delta 1-64}$, the recombinant enzyme was separated by size-exclusion chromatography using a Superdex 75 gel-filtration column. The elution position of the AtUGlcAE1$_{\Delta 1-64}$ activity peak was calculated to correspond to a mass of about 88 kD (Table 2.4), suggesting that the Arabidopsis enzyme, like the prokaryote cap1J, is active as a dimer.

All Three Isoforms of AtUGlcAE Are Transcribed in Arabidopsis

The expression of gene-specific AtUGlcAE transcript was determined by reverse transcription (RT)-PCR. Transcripts of AtUGlcAE1, 2, and 3 were observed in all Arabidopsis tissues examined (Fig. 2.6). However, the relative amount of transcript encoding AtUGlcAE1 was lower in leaf and stem tissues when compared to AtUGlcAE2 and 3.

DISCUSSION

GalA accounts for between 20% and 35% of the glycosyl residues present in plant primary cell walls, principally as a major component of pectic polysaccharides. Thus, enzymes involved in the formation of the activated form of GalA are likely to have important roles in regulating pectin biosynthesis (Orellana and Mohnen, 1999; Mohnen, 2002).

Our study reports the functional identification of a distinct Arabidopsis gene (At2g45310) that encodes a 4-epimerase (AtUGlcAE1) that interconverts UDP-GlcA and UDP-GalA, as determined by HPLC and NMR (Fig. 2.3; Table 2.2). We have shown that recombinant AtUGlcAE1 favors the formation of UDP-GalA because at equilibrium AtUGlcAE1 has a conversion ratio of 1 to approximately 2 (UDP-GlcA:UDP-GalA; Table 2.4). We suggest that UDP-GalA is less susceptible than UDP-GlcA to epimerization at C-4 because a hydrogen bond

may form between the axial hydroxyl at C-4" and the C-6" carboxyl group of GalA. This hydrogen bond will reduce the net negative charge of UDP-GalA and will therefore stabilize the molecule when compared with UDP-GlcA. Support for the net charge differences between UDP-GalA and the more ionic UDP-GlcA is evident by the fact that the UDP-GlcA is retained longer on an anion-exchange column, when compared with UDP-GalA (Fig. 2.3).

The hydropathy plot of the N-terminal region of AtUGlcAE1 (Fig. 2.1B) predicts that the protein has a single trans-membrane domain. In addition to AtUGlcAE1, we isolated two additional cDNAs (AtUGlcAE2 and 3) that share sequence similarity with AtUGlcAE1. All three cloned cDNAs (AtUGlcAE1, 2, and 3; At2g45310, At3g23820, and At4g30440 respectively; Table 2.1) encode proteins that have distinct N-terminal amino acid regions (amino acids approximately 1–100), and each includes a putative transmembrane domain, which supports the biochemical data for the existence of multiple UGlcAE isoforms in plants. We further speculate that the different UDP-GlcA epimerases in Arabidopsis may correspond to the digitonin-soluble and digitonin-insoluble isoforms reported in other plants (Feingold et al., 1960).

A total of six genes with similarity to the *Streptococcal* cap1J gene exist in the Arabidopsis genome (Reiter and Vanzin, 2001). While this work was being revised for publication, the activities of two Arabidopsis UDP-GlcA epimerases distinct from the epimerase described here were reported, although detailed chemical analysis of the product was not performed (Molhoj et al., 2004; Usadel et al., 2004). The NMR data provided in this report clearly and without ambiguity indicate that AtUGlcAE1 is a true reversible UDP-GlcA epimerase that favors the formation of UDP-GalA. Furthermore, as AtUGlcAE1 (At2g45310) shares a high amino acid sequence similarity to At4g30440 (GAE1; Molhoj et al., 2004) and At4g12250 (GAE6; Usadel et al., 2004), as well as to the other three uncharacterized member of

this gene family, it is highly likely that all six isoforms are UDP-GlcA/UDP-GalA 4-epimerases. However, these isoforms have distinct enzymatic properties: AtUGlcAE1 (At2g45310) described in this report, for example, is inhibited by UDP, UDP-Xyl, and UDP-Ara, while the At4g30440 protein (GAE1) was reported to be inhibited by only UDP-Xyl. Recombinant AtUGlcAE1 (At2g45310) has a Km of 0.72 mM and UDP-GlcA to UDP-GalA conversion ratio of 1.9, whereas recombinant At4g30440 protein (GAE1) and At4g12250 (GAE6) protein were reported to have lower Km values (0.19 mM and 0.23 mM, respectively) and a lower conversion ratio (1.3) of UDP-GlcA to UDP-GalA (Molhoj et al., 2004). These significant biochemical differences could indicate that the distinct isoforms of UDP-GlcAE in Arabidopsis have different biological roles.

It has been suggested that control of the flux of nucleotide sugars could play a role in the regulation of polysaccharide synthesis (Dalessandro and Northcote, 1977a, 1997b). The mechanisms that control flux have yet to be determined. Recent studies provide evidence that allosteric inhibition of nucleotide-sugar interconverting enzymes could be one such control mechanism. For example, UDP-Xyl strongly inhibits UDP-Glc dehydrogenase (Hinterberg et al., 2002; Bar-Peled et al., 2004) and several isoforms of UDP-GlcA decarboxylase (Harper and Bar-Peled, 2002; Suzuki et al., 2003). We have shown here that UDP-Xyl and UDP-Ara inhibit the recombinant AtUGlcAE1 (Fig. 2.5), which suggests that these nucleotide sugars could regulate the amount of UDP-GalA that is available for pectin biosynthesis. Whether activity of all of the AtUGlcAE isoforms is affected by these (or other) nucleotide sugars remains to be determined.

MATERIALS AND METHODS

Cloning and Expression Analysis ofArabidopsis UGlcAE

Arabidopsis (*Arabidopsis thaliana*) dbEST were searched to identify cDNAs that encode amino acid sequences with similarity to the *Streptococcus pneumoniae* UGlcAE (cap1J). Several ESTs (r90663, t20825, t21711, h76650, and ai100746) and genomic clones (F6n15.16, att4ca, f17i23, and aab82632) that showed similarity to the cap1J bacterial gene products were identified and used to isolate cDNA clones and to design primers to obtain the corresponding full-length Arabidopsis cDNA by RT-PCR. Total RNA from Arabidopsis ecotype Columbia plants was isolated using Trizol reagent (Gibco-BRL, Gaithersburg, MD; Chomczynski, 1993). RNA was reverse transcribed into cDNA at 42^{o}C for 60 min using 1 mM oligo(dT) primer (5#TTCTAGAATTCAGCGGCCGCTT$_{15}$-TTV) in a 20-µL total reaction consisting of 50 mM Tris-HCl, pH 8.4, 75 mM KCl, 3 mM MgCl2, 0.2 mM of each deoxynucleotide triphosphate (dNTP), 10 mM dithiothreitol (DTT), and 200 units SuperScript-II- RNaseH^{-} reverse transcriptase (Gibco-BRL). Following the reaction, 2 units of RNaseH (Gibco-BRL) were added. An aliquot (2 µL) of the resulting reverse-transcribed products was used as a template for PCR using 1 unit of high- fidelity Platinum Taq DNA polymerase mixed (Gibco-BRL) with GB-D proofreading DNA polymerase and Platinum antibody, 0.2 mM of the sense primer (120#1S1-26 5#CATATGTCACGTCTTGACGACATACCTTC), 0.2 mM of the antisense primer (120#2AS/1294-1317 5#GCGGCCGCTCATCTAGCGGCGGCTTTTTTGTCG), and 0.2 mM of each dNTP (Roche, Basel) in buffer containing 60mM Tris-SO4, pH8.9, 18mM ammonium-sulfate, and 1mM MgSO4. The RT-PCR reaction product (1,325 bp) was separated by agarose gel electrophoresis, purified, and cloned into pCR2.1-TOPO plasmid (Invitrogen, Carlsbad, CA). The cloned RT-PCR product in vector pCR2.1:121.1 was sequenced and the nucleotide sequence submitted to GenBank (accession no. AY594693, AtUGlcAE1). For expression in *Escherichia coli*, PCR reaction was carried for 15 cycles with sense 122#3S/196-221

5#CATATGTCTCGCCGTTCCTCCGAACAACAC and antisense 120#2AS/1294-1317 primers. The 1,217-bp PCR product consisting of a truncated coding region lacking the putative trans-membrane region (Δ1–64) was subcloned into pCR2.1 to generate pCR2.1:122.2 and sequenced. The NdeI-NotI fragment from pCR2.1:122.2, consisting of the coding region spanning from amino acid 65 to 438 of AtUGlcAE1, was subcloned into pET28b *E. coli* expression vector (Novagen, Madison, WI). The resulting clone, pET28b:122.2#3, was constructed to give an in-frame N-terminal 6His-AtUGlcAE1 gene fusion.

For AtUGlcAE1, 2, and 3 expression studies in Arabidopsis Columbia, total RNA was isolated from flowers, fully expanded rosette leaves, and stems of 6-week-old plants and from roots of 4-week-old plants grown in liquid media as described (Bar-Peled and Raikhel, 1997). The RNA was reverse transcribed into cDNA in 20-μL reactions using 200 units of SuperScript II reverse transcriptase (Invitrogen) and 1 mM oligo(dT) primer using the manufacturer's recommended buffer, 10 mM DTT, and 0.2 mM of each dNTP. One-twentieth of each of the reverse-transcribed products was used as a template for PCR reactions using 0.5 units Taq DNA polymerase (Roche), the manufacturer's buffer, 0.2 mM dNTP, 1.5 mM MgCl2, and 0.2 mM AtUGlcAE1 gene-specific sense and antisense primers (see above). The transcripts of AtUXS3 (UDP-GlcA decarboxylase) and AtUGlcAE3 described by Harper and Bar-Peled (2002) and AtUGlcAE2 were used as internal RT-PCR controls. PCR conditions were: 1 cycle at 95^{o}C for 2 min; and 30 cycles (95 ^{o}C, 30 s; 54 ^{o}C, 30 s; 70 ^{o}C, 1.5 min) and a final extension at 70 ^{o}C for 5 min. One-tenth of each sample and the DNA MW marker (1 KB plus; Invitrogen) were resolved on a Tris-acetate EDTA-1% agarose gel and visualized by staining with ethidium bromide.

Protein Expression, Purification, and Mass Determination

Fifteen milliliters of an overnight culture of *E. coli* carrying the pET28b:122.2#3 (AtUGlcAE1) or control vector were used to inoculate 0.5 L of Luria-Bertani liquid broth supplemented with 50 μg/mL kanamycin and 30 μg/mL chloramphenicol. Cells were grown at 37 ^{o}C while shaking (200 rpm) until a cell density of A_{600} reached 0.6 to 0.8. Gene expression was induced by the addition of isopropylthio-β-galactoside to a final concentration of 0.5 mM. After 3 h of growth at approximately 25 ^{o}C, the cells were collected by centrifugation (6,000g at 4 ^{o}C for 10 min). The cells were washed with cold water, resuspended in 20mL of extraction buffer (20mM Tris-HCl, pH 7.6, 10% glycerol, 1 mM EDTA), and supplemented with fresh 1 mM DTT and 0.5 mM phenylmethylsulfonyl fluoride. The cells were cooled in an iced-water bath and ruptured by 24 sonication intervals (10-s pulse followed by 20-s rest) using a microtip probe and a Fisher 550 sonicator (Fisher Scientific, Loughborough, Leicestershire, UK) set at a power of 4.5. The suspension was centrifuged (20,000g, 30min, 4 ^{o}C), and 1mL of the supernatant was injected at a flow rate of 1 mL min^{-1} onto an anion-exchange column (0.5 cm X 5 cm, packed with Source Q15; Pharmacia, Piscataway, NJ) that was preequilibrated with cold buffer A (50 mM sodium phosphate, pH 7.6). After the supernatant was loaded, the column was washed with buffer A until the UV (A_{280}) baseline was stabilized (30 min). The bound proteins were then eluted using a linear 30-min salt gradient from 0 to 0.5 M NaCl in the same buffer, at a flow rate of 0.5 mL min^{-1}, followed by a 5-min gradient from 0.5 to 0.8 M NaCl at the same flow rate. Proteins were detected by monitoring the A_{280} of the column effluent. The active fractions were desalted and concentrated using a YM-10 Centricon centrifuge filter (Amicon, Bedford, MA) and stored at -20 ^{o}C. Gelfiltration chromatography was used to estimate the M_r of the native recombinant AtUGlcAE1$_{\Delta 1-64}$. AtUGlcAE1$_{\Delta 1-64}$ (0.25 mL) was loaded at a flow rate of 0.5 mL min^{-1} onto a Superdex75 column (HR10X30; Pharmacia) previously equilibrated with 0.1 M

sodium phosphate, pH 7.6, 4% (v/v) glycerol. Proteins were eluted with the same buffer at a flow rate of 0.5 mL min^{-1}. Fractions (0.25, 0.5 mL) were collected for analysis. The gel-filtration column was calibrated using protein markers (aldolase, 158 kD; bovine serum albumin, 67 kD; ovalbumin, 43 kD; chymotrypsinogen A, 25 kD; myoglobin, 17.6 kD; ribonuclease A, 13.7 kD [Amersham Biosciences, Piscataway, NJ]). The A_{210} of the eluant was monitored to determine the elution times of each protein marker. Chromatography steps were performed at room temperature, and fractions containing AtUGlcAE1$_{\Delta 1-64}$ activity were analyzed and stored at -20^{o}C. Proteins were separated by 0.1% SDS-12% PAGE (Bar-Peled et al., 1991) alongside M_r markers (Bio-Rad, Hercules, CA) and visualized by staining with Coomassie blue or SimplyBlue (Invitrogen).

Enzyme Assay, HPLC, and NMR Product Analysis

UDP-GlcA epimerase activity was initiated by adding UDP-GlcA (1mM) to a reaction mix (50 µL) consisting of 0.1 M sodium phosphate, pH 7.6, 20% (v/v) glycerol, 0.2 mM β-NAD^{+}, 1 µg/mL AtUGlcAE1$_{\Delta 1-64}$. The assay was carried out for 30 min at 37^{o}C (unless otherwise mentioned), and the reaction was terminated by the addition of 50 µL of chloroform. The mixture was vortexed and centrifuged at 16,000g for 5 min at room temperature. The aqueous phase was retained, and the organic phase extracted with 80 µL of water. The aqueous phases were combined and injected via HPLC at 1 mL min^{-1} onto an anion-exchange column (Source15Q [Pharmacia] packed in 250- X 4.6-mm column) previously equilibrated with 2 mM ammonium format. After a 5-min wash, the sample was eluted with a linear 2 to 600 mM ammonium-formate gradient formed over 25 min (Bar-Peled et al., 2001). Nucleotides and nucleotide sugars were detected by their UV absorbance using a Waters (Milford, MA) photodiode array detector. The maximum absorbance for uridine, UDP-GlcA, and UDP-GalA

was 261.8 nm in 0.5 to 0.6 M ammonium formate. The column was calibrated with authentic nucleotides and nucleotide sugars (Sigma, St. Louis). For NMR analysis, UV-absorbing peaks eluting from the HPLC column were collected, diluted five times with water, and lyophilized to remove the ammonium formate. The residue was dissolved in water, relyophilized twice, and exchanged twice with 99.96% deuterium oxide. Proton NMR spectroscopy was performed at 25^{o}C on a Varian (Palo Alto, CA) Inova spectrometer operating at 500 MHz. The chemical shifts (δ) are reported in ppm relative to external acetone (2.224 ppm).

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession number AAT06796.

CHAPTER 3

ENZYMATIC CHARACTERIZATION AND COMPARISON OF VARIOUS POACEAE UDP-GLCA 4-EPIMERASES [2]

[2]Xiaogang Gu, Kathryn E. Davis, Paul J. Guyett and Maor Bar-Peled, submitted to Journal of Biochemistry

ABSTRACT

UDP-α-D galacturonic acid (UDP-GalA) is a key precursor for the synthesis of various bacterial and plant polysaccharides. UDP-glucuronic acid 4-epimerase (UGlcAE) catalyzes the reversible conversion of UDP-α-D-glucuronic acid to UDP-GalA. UGlcAEs in two bacteria species and in one dicot species, Arabidopsis, have been identified and functionally characterized. However, little is known about the specificity of UGlcAE in Poaceae species. Therefore, we cloned and expressed several putative maize and rice UGlcAE genes, and compared their enzymatic properties with dicot homologs in Arabidopsis. All tested plant UGlcAEs, unlike their homolog in *Klebsiella pneumoniae*, can only use UDP-GlcA or UDP-GalA as substrate. Our data show that UGlcAE isoforms in different plant species have different enzymatic properties. For example, the maize UGlcAE displays a much higher inhibition by UDP-arabinose than rice and Arabidopsis homologs. However, the epimerases from different plant species share similarities as well. All of them are inhibited by UDP-xylose and UDP; but not by UMP, UDP-glucose or UDP-galactose. This study demonstrates that although UGlcAE isoforms are highly conserved in their amino acid sequences and have identical enzymatic activity *in vitro*, the activity of specific plant isoform(s) can be regulated differently by specific nucleotide-sugars.

INTRODUCTION

Galacturonic acid (GalA) is a sugar residue that exists in different bacterial polysaccharides and in various plant glycans. GalA is a component of capsular polysaccharides and lipopolysaccharides (LPS) of several bacterial species (Forsberg and Carlson, 1998; Munoz et al., 1999; Frirdich and Whitfield, 2005). In plants, all pectic polymers: homogalacturonan, rhamnogalacturonan I, rhamnogalacturonan II, xylogalacturonan and apiogalacturonan, contain GalA residue in their backbones (Mohnen 2002). GalA is also present in the reducing end-sequence of spruce and Arabidopsis xylans (Andersson et al., 1983; Pena et al., 2007) and in arabinogalactan-protein (AGP) (Pellerin et al., 1995). UDP-galacturonic acid (UDP-GalA), the activated nucleotide-sugar form of GalA, has been demonstrated to be the precursor for the synthesis of several GalA containing polysaccharides in bacteria (Munoz et al., 1999; Frirdich et al., 2005) and for the synthesis of pectin in plants (Mohnen, 2002). It has been suggested that at least four different galacturonosyltransferases (GalATs) are required to transfer the GalA from UDP-GalA during the synthesis of the plant pectic polymers (Mohnen, 2002; Sterling et al., 2006).

In bacteria, genes encoding UDP-GlcA 4-epimerase (UGlcAE) have been isolated from type I *Streptococcus. pneumoniae* (Cap1J, Munoz et al., 1999) and *Klebsiella pneumoniae* (Gla_{KP}, Frirdich and Whitfield, 2005) and biochemical data indicate that these two different bacterial UGlcAEs have distinct biochemical properties. For example, in addition to the reversible conversion of UDP-GlcA and UDP-GalA, Gla_{KP} is capable of reversibly interconverting UDP-glucose / UDP-galactose and UDP-*N* acetylglucosamine / UDP-*N*-acetylgalactosamine (Frirdich and Whitfield, 2005). Cap1J on the other hand is specific for UDP-uronic acids (Munoz et al., 1999). In plants, a UGlcAE gene family containing six isoforms

has been identified in Arabidopsis (Gu and Bar-Peled, 2004); each isoform is predicted to have a type-II membrane protein topology and some have been functionally characterized (Molhoj et al. 2004, Usadel et al. 2004, Gu and Bar-Peled 2004). However, little is known about the specificity of UGlcAEs in the Poaceae (e.g. maize and rice) and no Poaceae UGlcAE homologs have been identified or characterized so far. Therefore, to expand our knowledge regarding the UGlcAE isoforms in the Poaceae, we identified, cloned and characterized UGlcAE homologs in maize and rice, and subsequently compared their enzymatic properties with Arabidopsis homologs.

RESULT

Isolating and Cloning UGlcAE from Maize

The amino acid sequence of AtUGlcAE1 (Gu and Bar-Peled, 2004) was used to identify candidate gene homologs from maize expressed sequence tag database (dbEST) projects and available genomic data. Based on the sequence data, we cloned one putative maize UGlcAE (ZmUGlcAE) cDNA encoding a protein that shares 66% amino acid sequence identity to AtUGlcAE1. The putative maize UGlcAE has the conserved nucleotide-sugar 4-epimerase catalytic motifs, which include the GxxGxxG (x = any amino acid) motif within the C-terminal domain that is likely to be involved in the binding of β-NAD^+, and a catalytic triad consisting of S/T and YxxxK (Weirenga et al., 1986; Gu and Bar-Peled, 2004). Reverse transcription (RT)-PCR was performed to determine the tissue-specific expression of the maize cDNA. Transcripts were detected in root, seedling, ear and mature leaf (Fig. 3.1).

To determine if the maize homolog functions as a UDP-GlcA 4-epimerase, we expressed the gene in *Escherichia coli* and studied its biochemical properties.

ZmUGlcAE Encodes Active UGlcAE

Since we were unable to express recombinant proteins that contain trans-membrane domains in *E. coli* (Harper and Bar-Peled, 2002; Gu and Bar-Peled, 2004), a recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$, lacking the putative transmembrane domain, was used for expression in *E. coli* (Fig. 3.2 a, lane 2,). The recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ converted UDP-GlcA to UDP-GalA (Fig. 3.2 b, panel 2) while control *E. coli* cells containing an empty vector (Fig. 3.2 a, lane 3) had no UGlcAE activity (Fig. 3.2 b, panel 3). ZmUGlcAE$_{\Delta 1\text{-}47}$ was stable (> 6 months) when stored as a crude extract at -20^{o}C, similar to the AtUGlcAE1 (Gu and Bar-Peled, 2004), which is still fully active after 1 year in storage.

To further characterize the protein, the recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ was purified by chromatography (Fig. 3.2 a, lane 4). The purified 42.9 kDa ZmUGlcAE$_{\Delta 1\text{-}47}$ converts UDP-GlcA to a product that elutes from the HPLC column at a retention time of 22.9 min (Fig. 3.2 b, panel 5), at the same elution time as the UDP-GalA standard (Fig. 3.3, panel 1). To characterize the enzymatic product, the peak eluting from the column at 22.9 min was collected and subsequently analyzed by proton ^{1}H-NMR. The proton chemical shifts and coupling constants data obtained from the 500-mHz ^{1}H-NMR spectrum of the enzymatic product (data not shown) are identical to those obtained from authentic UDP-GalA (Gu and Bar-Peled, 2004) confirming that ZmUGlcAE$_{\Delta 1\text{-}47}$ converts UDP-GlcA to UDP-GalA. The data shown in Fig. 3.3 also demonstrate that ZmUGlcAE$_{\Delta 1\text{-}47}$ is capable of converting UDP-GalA to UDP-GlcA. Like the Arabidopsis isoform, the maize UGlcAE does not require exogenous NAD^{+} (Fig. 3.3, panel 3 and 6). Further testing demonstrated that the ZmUGlcAE$_{\Delta 1\text{-}47}$ is specific for UDP-uronic acids and cannot use other nucleotide sugars such as UDP-Glc, UDP-Gal, UDP-Xyl, UDP-Ara, as substrates (data not shown).

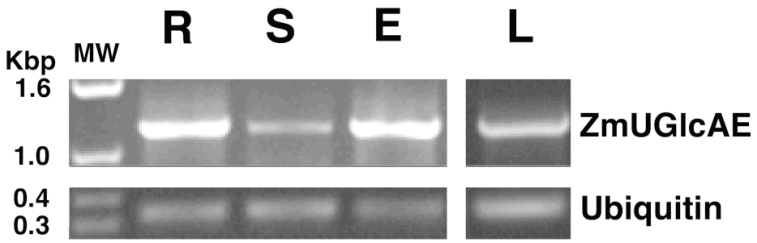


Figure 3.1. ZmUGlcAE transcript is expressed in different maize tissues.

Total RNA isolated from the maize root (R), seedling (S), ear (E), and leaf (L) was used for RT-PCR reaction with gene-specific primers. RT-PCR with ubiquitin specific primers (Livak and Schmittgen, 2001) was carried out as an internal control.

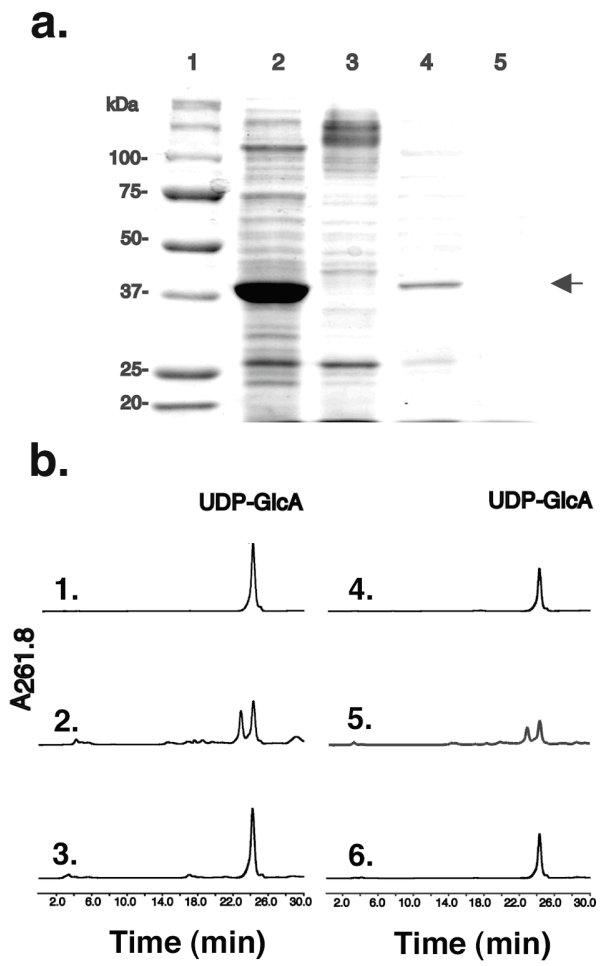


Figure 3.2. Isolation, chromatography and initial characterization of recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$

a. SDS-PAGE of total protein isolated from *E. coli* cells expressing ZmUGlcAE$_{\Delta 1\text{-}47}$ (lane 2) or vector control (lane 3). Proteins from cells expressing recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ were separated by successive chromatographies (see "Material and Methods") and the active fraction eluted from a SourceQ column was separated by SDS-PAGE (lane 4). As a control, total proteins from *E.coli* cells expressing vector alone were separated under the same chromatographic

conditions and the corresponding elution fraction from the SourceQ column were separated by SDS-PAGE as well (lane 5). The arrow indicates the position of recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$.

b. An aliquot of each protein sample (total protein or protein fraction eluted from SourceQ column) was incubated with UDP-GlcA and the enzymatic reactions were separated on ion exchange HPLC column. The migration of UDP-GlcA standard is indicated in panels 1 and 4. Activity of total protein isolated from cells expressing recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ (panel 2) or vector alone (panel 3) was determined using 1 mM UDP-GlcA as substrate. The activity of column-purified protein isolated from cells expressing recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ (panel 5), or control vector alone that was fractionated on the same SourceQ column (panel 6), was determined using 0.5 mM UDP-GlcA as substrate.

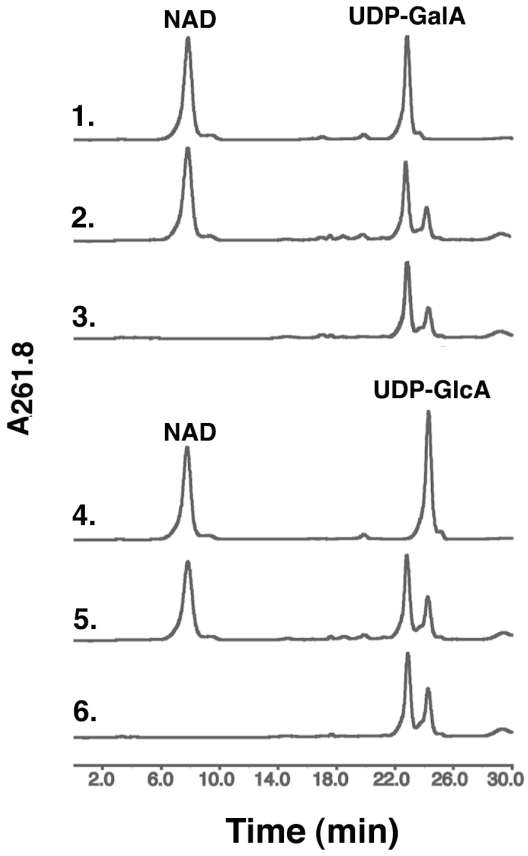


Figure 3.3. Recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ is a reversible UDP-GlcA/UDP-GalA-4 epimerase, and does not require exogenous NAD$^+$

HPLC separation of NAD$^+$, UDP-GalA and UDP-GlcA standards (panel 1 and 4) and of UDP-GlcA epimerase enzymatic products (panel 2, 3 5 and 6). Recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ was incubated for 2 h with 1 mM UDP-GalA and 1 mM NAD$^+$ (panel 2), with 1 mM UDP-GalA but without NAD$^+$ (panel 3), 1 mM UDP-GlcA and 1 mM NAD$^+$ (panel 5), or with 1 mM UDP-GlcA but without NAD$^+$ (panel 6).

Table 3.1. Enzymatic characterization of recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$

[a] Assays were conducted for 30 min in either 0.1 M sodium phosphate at various pH values, ranging from 2 to 11.5, or 0.1 M Tris-HCl (pH 7.2 to 8). The relative activity at the pH value indicated in phosphate buffer is 100%; while in Tris-HCl buffer, the relative activity is 84%. [b] Optimal temperature assays were conducted in phosphate buffer (pH 7.6) for 30 min at temperature ranging from 4^{o}C to 65^{o}C. [c] For kinetic studies, assays at different concentrations of UDP-GlcA (0.1mM to 4mM) were performed for 5 min. The reciprocal initial velocity was plotted against UDP-GlcA concentration according to Lineweaver and Burk to calculate *K*m value. [d] Assays were conducted in phosphate buffer (pH 7.6) for different time (0, 10, 20, 30, 90, 300 and 480 min). [e] The molecular mass of the active ZmUGlcAE$_{\Delta 1\text{-}47}$ eluted from a Superdex75 gel filtration column (24 min) was estimated based on extrapolation (R^2=0.993) of the elution time of standard protein markers.

Optimal pH[a]	Optimal Temperature[b] (^{o}C)	*K*m[c] (mM)	*k*cat (S^{-1})	Catalytic efficiency (M^{-1} S^{-1})	Equilibrium constant[d]	Mass of active protein[e] (Denatured) (kDa)
7.2-7.8	30-37	0.60	52	$8.67x10^4$	1.9	81.4 (42.9)

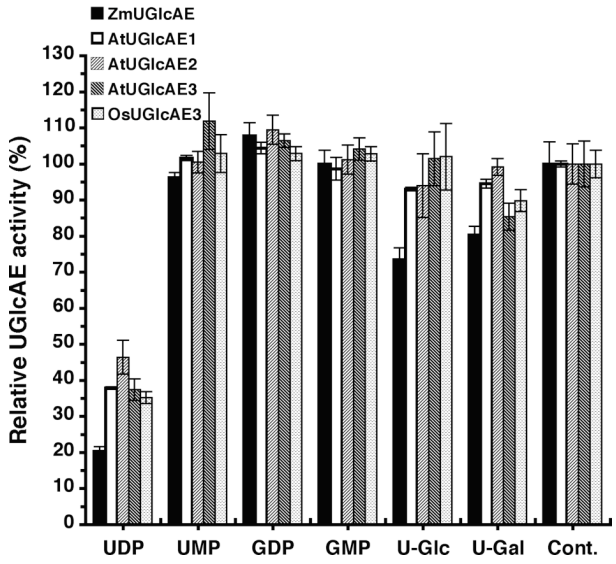


Figure 3.4. The effects of nucleotides and nucleotide sugars on the activity of the recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$, AtUGlcAE1$_{\Delta 1\text{-}64}$, AtUGlcAE2$_{\Delta 1\text{-}68}$, AtUGlcAE3$_{\Delta 1\text{-}53}$ and OsUGlcAE3$_{\Delta 1\text{-}53}$

Recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$, AtUGlcAE1$_{\Delta 1\text{-}64}$, AtUGlcAE2$_{\Delta 1\text{-}68}$, AtUGlcAE3$_{\Delta 1\text{-}53}$ and OsUGlcAE3$_{\Delta 1\text{-}53}$ were first incubated in the absence (control) or presence of different nucleotides and nucleotide sugars (2 mM) for 15 min. Assays were then initiated by adding 1 mM UDP-GlcA and terminated after 30 min. 100% activity corresponds to 700 pmol/min of UDP-GalA produced for ZmUGlcAE$_{\Delta 1\text{-}47}$, 970 pmol/min of UDP-GalA produced for AtUGlcAE1$_{\Delta 1\text{-}64}$, 463 pmol/min of UDP-GalA produced for AtUGlcAE2$_{\Delta 1\text{-}68}$, 624 pmol/min of UDP-GalA produced for AtUGlcAE3$_{\Delta 1\text{-}53}$ and 775 pmol/min of UDP-GalA produced for OsUGlcAE3$_{\Delta 1\text{-}53}$. The data presented are the average relative amounts of UDP-GalA produced compared to the control from two experiments. Each value is the mean of triplicate reactions.

Enzymatic Characterization of ZmUGlcAE$_{\Delta 1\text{-}47}$

The optimal pH for the activity of recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ is 7.2 to 7.8 in phosphate buffer, and the optimal temperature is 37^{o}C (Table 3.1). The activity of ZmUGlcAE$_{\Delta 1\text{-}47}$ is completely abolished when assays were conducted at pH value lower than 4 or higher than 10, or when the assay temperature was above 65^{o}C.

The equilibrium constant (UDP-GalA/UDP-GlcA) is 1.9 (Table 3.1). Further kinetic analyses indicate that the ZmUGlcAE$_{\Delta 1\text{-}47}$ has an apparent *K*m value of 0.6 mM for UDP-GlcA and a turnover number (k_{cat}) of 52 (s^{-1}). The catalytic efficiency [k_{cat}/Km] value of the maize UGlcAE is 8.67 x 10^4 ($M^{-1}s^{-1}$; Table 3.1). Thus the enzymatic characteristics of ZmUGlcAE$_{\Delta 1\text{-}47}$ are comparable to those of Arabidopsis AtUGlcAE1$_{\Delta 1\text{-}64}$ (Gu and Bar-Peled, 2004)

To determine the apparent molecular weight (M_r) of the active ZmUGlcAE$_{\Delta 1\text{-}47}$, the recombinant enzyme was analyzed by size-exclusion chromatography on a Superdex75 column. ZmUGlcAE$_{\Delta 1\text{-}47}$ activity was detected at an elution volume corresponding to a mass of about 82 kD (Table 3.1), suggesting that the maize enzyme is active as a dimer *in vitro* assays, similar to the prokaryotic cap1J (Munoz et al., 1999) and Arabidopsis AtUGlcAE1 (Gu and Bar-Peled, 2004) enzymes.

UDP-Ara and UDP-Xyl Strongly Inhibit the Activity of ZmUGlcAE$_{\Delta 1\text{-}47}$

The activity of ZmUGlcAE$_{\Delta 1\text{-}47}$ was tested in the presence of different nucleotides and nucleotide-sugars and compared with that of Arabidopsis isoform AtUGlcAE1 (Fig. 3.4). The activities of ZmUGlcAE$_{\Delta 1\text{-}47}$ and AtUGlcAE1$_{\Delta 1\text{-}64}$ were inhibited by UDP but not by UMP, GDP,

or GMP (Fig. 3.4). Similar to AtUGlcAE1$_{\Delta 1\text{-}64}$, ZmUGlcAE$_{\Delta 1\text{-}47}$ is also not significantly inhibited by UDP-Glc or UDP-Gal.

Our previous analysis of AtUGlcAE1 suggested that UDP-Xyl and UDP-Ara could regulate the amount of UDP-GalA produced, and thus, the amount of UDP-GalA that is available for pectin biosynthesis (Gu and Bar-Peled, 2004). We therefore compared the activity of ZmUGlcAE$_{\Delta 1\text{-}47}$ and recombinant AtUGlcAE1$_{\Delta 1\text{-}64}$ in the presence of different concentrations of UDP-Xyl and UDP-Ara. As shown in Fig. 3.5, the activity of both enzymes is strongly inhibited by UDP-Xyl. However, while no significant inhibition was observed when AtUGlcAE1 was incubated with up to 1 mM UDP-Ara, the maize UGlcAE was strongly inhibited by UDP-Ara (Fig. 3.5).

To further determine if the inhibition of UGlcAE by UDP-Ara is maize-specific, we cloned, characterized and compared other UGlcAE enzymes in rice and Arabidopsis as described below.

Comparing the Enzymatic Properties of UGlcAEs in Rice, Maize and Arabidopsis and the Effects of UDP-Xyl, UDP-Ara on These Isozymes

Rice contains five UGlcAE homologs based on the genomic data and rice dbEST project. Phylogenetic analysis was used to compare the amino-acid sequences of UGlcAE isoforms from Arabidopsis, rice and maize. This analysis separates the UGlcAEs into three evolutionary clades: Type A, B, and C (Fig. 3.6). The maize UGlcAE protein is in the Type C clade and therefore it was named as ZmUGlcAE3 to maintain consistence with the other members of the C clade. Based on the phylogenic data, it became clear that to truly evaluate if UDP-Ara inhibition of UGlcAE is specific to the maize isoform or specific to the clade C epimerases, it would be

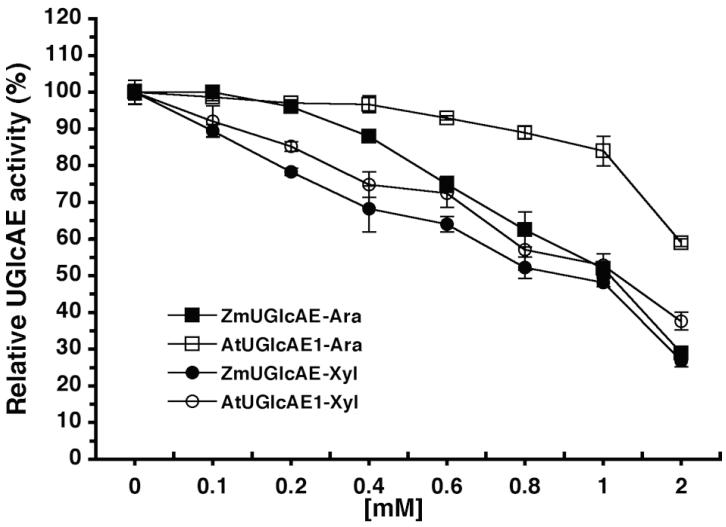


Figure 3.5. The effects of different concentrations of UDP-Xyl and UDP-Ara on the activity of the recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ and AtUGlcAE1$_{\Delta 1\text{-}64}$

Recombinant ZmUGlcAE$_{\Delta 1\text{-}47}$ (solid symbols) and AtUGlcAE1$_{\Delta 1\text{-}64}$ (open symbols) were first incubated in the absence (control) or presence of various concentrations of UDP-Xyl (circles) and UDP-Ara (squares) for 15 min. Assays were then initiated by adding 1 mM UDP-GlcA and after 30 min, reactions were terminated. 100% activity corresponds to 730 pmol/min of UDP-GalA produced for ZmUGlcAE$_{\Delta 1\text{-}47}$, 930 pmol/min of UDP-GalA produced for AtUGlcAE1$_{\Delta 1\text{-}64}$. The data are the average relative amounts of UDP-GalA produced compared to the control from two experiments. Each value is the mean of triplicate reactions.

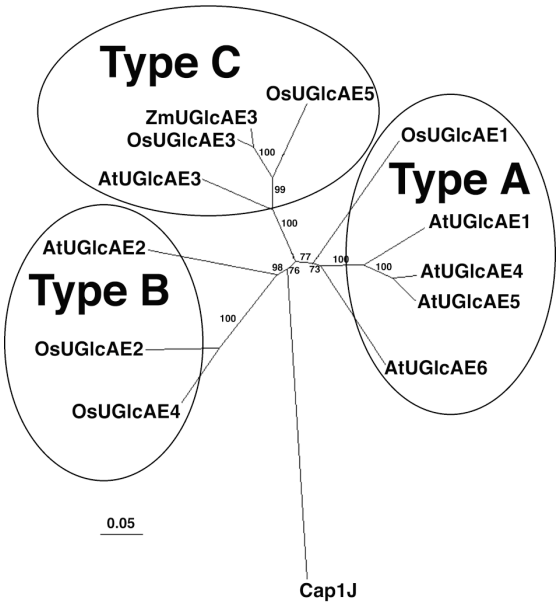


Figure 3.6. Neighbor-joining tree of the Maize, Rice and Arabidopsis UGlcAE proteins.

UGlcAE protein sequences were aligned with ClustalX software. Alignments were analyzed with PAUP software (Swofford, 1998) to generate an unrooted tree. Percentage bootstrap values of 1,000 replicates are given at each branch point. Branch lengths are shown to scale. The Genbank accession number or the loci names are in parentheses: AtUGlcAE1 (At2g45310), AtUGlcAE2 (At3g23820), AtUGlcAE3 (At4g30440), AtUGlcAE4 (At4g00110), AtUGlcAE5 (At1g02000), AtUGlcAE6 (At4g12250), ZmUGlcAE(DQ247999), OsUGlcAE1 (DQ333338), OsUGlcAE2 (DQ333337), OsUGlcAE3 (DQ333336), OsUGlcAE4 (Os09g32670), OsUGlcAE5 (Os06g08810), Cap1J (Z83335).

Table 3.2. Enzymatic characterization of recombinant OsUGlcAE3$_{\Delta 1\text{-}53}$, AtUGlcAE3$_{\Delta 1\text{-}53}$, and AtUGlcAE2$_{\Delta 1\text{-}68}$

[a] Assays were conducted for 30 min in 0.1M sodium phosphate buffer at various pH values, ranging from 2 to 11.5. [b] Optimal temperature assays were conducted in phosphate buffer (pH 7.6) for 30 min at temperature ranging from 4^{o}C to 65^{o}C. [c] For kinetic studies, assays at different concentrations of UDP-GlcA (0.1 mM to 4 mM) were conducted for 5 min. The reciprocal initial velocity was plotted against UDP-GlcA concentration according to Lineweaver and Burk to calculate *K*m values. [d] The molecular masses of the active OsUGlcAE3$_{\Delta 1\text{-}53}$, AtUGlcAE3$_{\Delta 1\text{-}53}$, and AtUGlcAE2$_{\Delta 1\text{-}68}$ eluted from Superdex75 gel filtration column (24.2 min, 24.6 min, 24.6 min, respectively) were estimated based on extrapolation (R^2=0.9997) of the elution times of standard protein markers.

	Optimal pH[a]	Optimal Temperature[b] (^{o}C)	*K*m[c] (mM)	Mass of active protein[d] (Denatured) (kDa)
OsUGlcAE3$_{\Delta 1\text{-}53}$	7.6	37	0.53	92.2 (44.3)
AtUGlcAE3$_{\Delta 1\text{-}53}$	7.6	37	0.52	88.5 (44.1)
AtUGlcAE2$_{\Delta 1\text{-}68}$	7.6	37	0.31	88.5 (45.5)

necessary to characterize and compare the rice and Arabidopsis Type C UGlcAE isoforms. Therefore, we cloned and expressed recombinant forms of several UGlcAEs (type B, AtUGlcAE2$_{\Delta 1\text{-}68}$; type C, AtUGlcAE3$_{\Delta 1\text{-}53}$ and OsUGlcAE3$_{\Delta 1\text{-}53}$) in *E. coli*, and characterized the expressed enzymes. All recombinant UGlcAEs are active under similar pH and temperature conditions as summarized in Table 3.2. In addition, based on size-exclusion chromatography, all UGlcAE isoforms analyzed are active as dimers with the following native molecular masses: AtUGlcAE2$_{\Delta 1\text{-}68}$ (88.5kDa), AtUGlcAE3$_{\Delta 1\text{-}53}$ (88.5kDa) and OsUGlcAE3$_{\Delta 1\text{-}53}$ (92.2kDa). These recombinant UGlcAEs have similar *K*m values (Table 3.2). All five recombinant enzymes are inhibited by UDP, but not by UMP, GDP, GMP, UDP-Glc or UDP-Gal (Fig. 3.4). To further compare the catalytic properties of these recombinant enzymes, the effects of UDP-Xyl and UDP-Ara on the activities of OsUGlcAE3$_{\Delta 1\text{-}53}$, AtUGlcAE2$_{\Delta 1\text{-}68}$ and AtUGlcAE3$_{\Delta 1\text{-}53}$ were tested (Fig. 3.7 a and b) and compared with ZmUGlcAE$_{\Delta 1\text{-}47}$ and AtUGlcAE1$_{\Delta 1\text{-}64}$. All five recombinant enzymes were inhibited by UDP-Xyl (Fig. 3.5, Fig. 3.7A), although the levels of inhibition varied among the different enzymes. For example, at 1 mM UDP-Xyl, AtUGlcAE2$_{\Delta 1\text{-}68}$ was inhibited by ~25% while the activities of AtUGlcAE3$_{\Delta 1\text{-}53}$ and OsUGlcAE3$_{\Delta 1\text{-}53}$ were inhibited by approximately 50 and 60%, respectively (Fig. 3.7a). Conversely, inhibition of the activity of the three enzymes (AtUGlcAE2$_{\Delta 1\text{-}68}$, AtUGlcAE3$_{\Delta 1\text{-}53}$ and OsUGlcAE3$_{\Delta 1\text{-}53}$) by UDP-Ara was only observed at a concentration of 2 mM. There was no significant inhibition when the concentration of UDP-Ara was below 1 mM (Fig. 3.7b). All the inhibition assays mentioned above were repeated four to five times and the result consistently shows that only the maize ZmUGlcAE3 is inhibited by low concentrations of UDP-Ara, while other recombinant proteins tested were not.

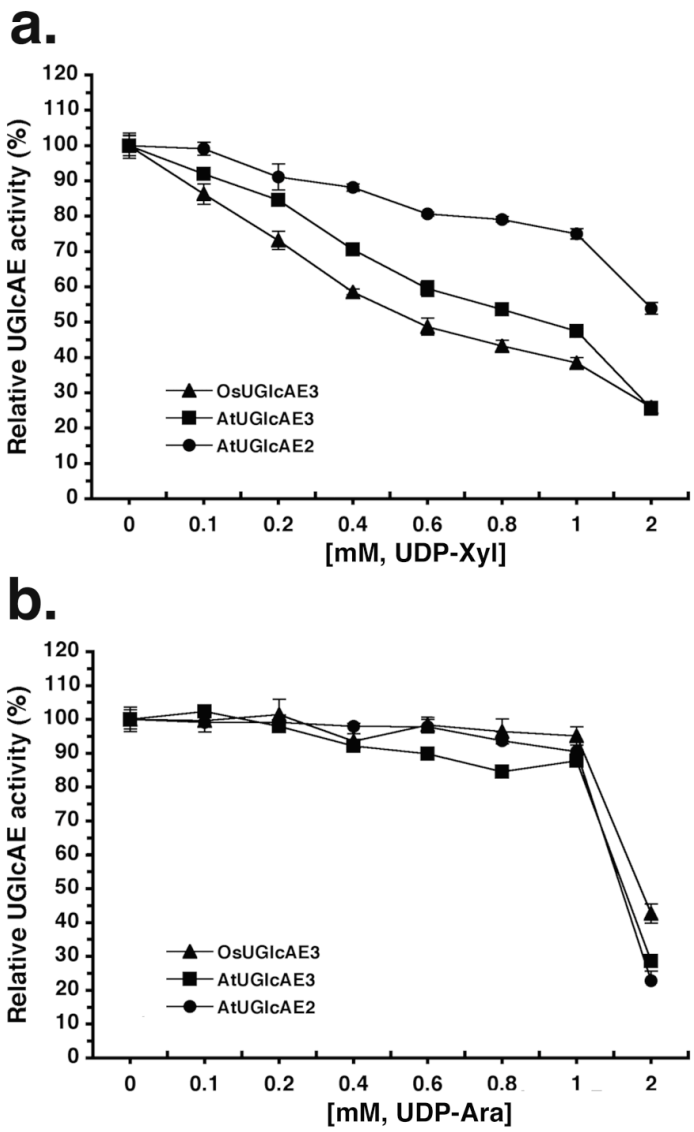


Figure 3.7. The effects of different concentrations of UDP-Xyl and UDP-Ara on the activity of recombinant AtUGlcAE2$_{\Delta 1\text{-}68}$, AtUGlcAE3$_{\Delta 1\text{-}53}$ and OsUGlcAE3$_{\Delta 1\text{-}53}$

Recombinant AtUGlcAE2$_{\Delta 1\text{-}68}$ (●) , AtUGlcAE3$_{\Delta 1\text{-}53}$ (■) and OsUGlcAE3$_{\Delta 1\text{-}53}$ (▲) were first incubated in the absence (control) or presence of various concentrations of UDP-Xyl (panel a) or UDP-Ara (panel b) for 15 min. UDP-GlcA (1 mM) was then added, and the reactions were incubated for 30 min. 100% activity corresponds to 670 pmol/min of UDP-GalA produced for

AtUGlcAE2$_{\Delta 1\text{-}68}$, 830 pmol/min of UDP-GalA produced for AtUGlcAE3$_{\Delta 1\text{-}53}$ or 860 pmol/min of UDP-GalA produced per minute for OsUGlcAE3$_{\Delta 1\text{-}53}$. The data are the average relative amounts of UDP-GalA produced compared to the control. Each value is the mean of triplicate reactions.

DISCUSSION

Our study reports the functional characterization of UGlcAEs in maize and rice. Although we cloned and characterized only one maize UGlcAE, several maize UGlcAE isoforms exist as indicated by maize dbEST (data not shown). Therefore, both dicot and monocot plant species contain multiple isoforms of UGlcAE. Furthermore, the collective biochemical data of all the UGlcAEs characterized so far indicate that both the bacterial and plant UGlcAEs are active as dimers and their activity is inhibited by specific nucleotides or nucleotide-sugars (e.g. UDP and UDP-Xyl). However, the UGlcAEs in these two different kingdoms also have distinct enzymatic properties as well. For example, the bacterial Cap1J is strongly inhibited by UDP-Gal (Munoz et al., 1999) while plant UGlcAEs are not.

Both Poaceae and dicot plants contain pectin as a major component in their cell walls. However, the amount of pectin in Poaceae and dicots is significantly different: <10% in grasses vs. ~30% in dicots (O'Neill and York, 2003), and the biological reason for this difference still remains unclear. It has been suggested that the biosynthesis of cell wall polysaccharides is controlled by both the glycosyl transferases responsible for assembling the polymers and the level of activated nucleotide-sugar precursors that supply those synthases (Freshour et al. 2003; Seifert, 2004; Lerouxel et al., 2006). Therefore, one possible explanation for the different pectin levels between Poaceae and dicots could be that their UGlcAEs, the enzymes that produce UDP-GalA for the pectin synthesis, may have distinct enzymatic properties. The strong inhibition (about 50%) of the maize ZmUGlcAE3 by UDP-Ara, as reported in this paper, could raise a hypothesis that inhibition of the activity of maize UGlcAE by these nucleotide-sugars could lead to the reduction of the *in-vivo* UDP-GalA level, resulting in less pectic polymer production in maize. Whether this inhibition occurs *in vivo*, however, remains unclear and it seems that this

hypothesis may not be applied to all Poaceae species because such inhibition was not observed in the rice homolog tested here.

A fundamental question in plant biology is why so many isoforms exist for the synthesis of particular metabolites. Several explanations can be proposed: 1) Some isoforms are inhibited by specific metabolite(s). For example, Barber et al. (2006) who studied isoforms involved in synthesis of UDP-Gal (UDP-glucose 4-epimerase, UGE) recently demonstrated that UGE1, but not UGE4, is sensitive to certain nucleotide inhibitors. This may be the case among maize, rice and Arabidopsis UGlcAE3s, which show different sensitivities to UDP-Ara. However, within one species, such as Arabidopsis, there is no significant difference among the tested AtUGlcAE isoforms toward the same inhibitors, which is also observed in Arabidopsis UDP-GlcA decarboxylase (UXS) isoforms (Harper and Bar-Peled, 2002). 2) Different isoforms evolved to be expressed in a tissue-specific manner. While this may be a case for tobacco sucrose-phosphate synthase (SPS) isoforms (Chen et al., 2005), the available expression data for the various Arabidopsis UGlcAE, UXS and UGE gene family members (Harper and Bar-Peled, 2002; Zimmermann et al., 2004; Usadel et al., 2004; Gu and Bar-Peled, 2004) suggest that expression of those isoforms is not tissue specific, although the expression level in specific cell types may be different (Birnbaum et al., 2003). 3) Some isoforms may be targeted to different locations within a cell or within an organelle. This has been demonstrated for several isozymes. For example, of the six UXS isoforms, Uxs1 is localized in Golgi whereas Uxs3 is cytosolic (Pattathil et al., 2005). Variation of targeting within an organelle was also described for the H_2O_2 scavenging isoenzymes of ascorbate peroxidase (APX). Two distinct types of chloroplastic APXs exist; one is localized in the stroma (sAPX), the other is in the thylakoid (tAPX) (Panchuk et al., 2005). Clearly, a sub-cellular specific targeting domain is required to target a specific

isoform to its final cellular destination. These specific 'targeting domains' likely evolved by the early insertion and subsequent mutation of nucleotide sequences in the different isoforms independently. Although the plant UGlcAE isoforms identified so far have a highly conserved catalytic domain, sequence variation does exist, particularly within their N-terminal (approximately first 100 amino acid) and C-terminal (approximately last 25 amino acid) regions. Similar sequence variation at the C and N-termini was observed in the membrane bound UXS isoenzymes (Harper and Bar-Peled, 2002). We suggest that these variations may localize those isoforms to different endomembranes. Current work in our laboratory is underway to address the latter possibility.

MATERIALS AND METHODS

Cloning and Expression Analysis of Arabidopsis, Maize and Rice UGlcAEs

Maize (*Zea mays*) expressed sequence tag cDNA databases (dbEST) were searched to identify cDNAs that encode amino acid sequences similar to the *Arabidopsis thaliana* UDP-GlcA 4-epimerase1 (AtUGlcAE1). Several expressed sequence tags (CO519975, CA830134, BQ668208, BU050741) that showed similarity to the AtUGlcAEl gene products were identified and used to isolate cDNA clones. To clone the coding region of maize UGlcAE, total RNA from seedlings of *Z. mays* ecotype B73 was isolated using Trizol reagent (Gibco-BRL, Gaithersburg, MD; Chomczynski, 1993). RNA was reverse transcribed into cDNA as described (Gu and Bar-Peled, 2004). An aliquot (4 μl) of the resulting reverse-transcribed products was used as a template for PCR using one unit of high fidelity Platinum Taq DNA polymerase (Gibco-BRL) mixed with GB-D proofreading DNA polymerase and Platinum antibody, 0.2 μM of the sense primer [ZmCO519975Bsp1S#1 5'GTCATGAGGGTGTTGGAGGAGGACCTCTTC], 0.2 μM of

the antisense primer [ZmAY105703Xho1AS#1 5' CTCGAGTCATTGTCGCGAGTTCTTGGATCC], and 0.2 mM of each dNTP (Roche, Basel) in buffer containing 60 mM Tris-SO_4 (pH 8.9), 18 mM ammonium-sulfate 1 mM $MgSO_4$, 5% dimethyl sulfoxide and 1M betaine. The RT-PCR reaction product (1332 bp) was separated by agarose-gel electrophoresis, purified, and cloned into pCR2.1-TOPO plasmid (Invitrogen, Carlsbad, CA). The RT-PCR product in vector pCR2.1:Zm#1 was sequenced and the nucleotide sequence was submitted to GenBank (accession no. DQ247999, ZmUGlcAE3). For expression in *E. coli*, the PCR reaction was carried out for 20 cycles with sense ZmCO519975Nde1S#2 5'CATATGACGGCGTCCTACCTCAGCTTCCAGTC and antisense ZmAY105703HindIII AS#2 5'AAGCTTTCATTGTCGCGAGTTCTTGGATCC primers using cDNA as template as above. The 1191 bp PCR product, consisting of a truncated coding region lacking the putative trans-membrane region (Δ1-47), was subcloned into pCR4-TOPO plasmid (Invitrogen) to generate pCR4:Zmt#14 and subsequently verified by sequencing. The Nde I- Hind III fragment from pCR4:Zmt#14, consisting of the coding region spanning from amino acid 48 to 440 of ZmUGlcAE3 [i.e. $ZmUGlcAE3_{\Delta 1\text{-}47}$], was subcloned into the Nde I/Hind III cloning sites of pET30a *E. coli* expression vector (Novagen, Madison, WI) resulting in clone pET30a:Zmt#14.2.

For the cloning of the Arabidopsis and rice UGlcAE homologs, total RNA was extracted from *Arabidopsis thaliana* ecotype Columbia plants or from seedlings of *Oryza sativa japonica* cultivar-group Sariceltik and the corresponding homologs were obtained by RT-PCR as described above. For the expression of the Arabidopsis homologs (AtUGlcAE2, At3g23820; AtUGlcAE3, At4g30440) and rice UGlcAE homologs (OsUGlcAE1, GenBank DQ333338; OsUGlcAE2, GenBank DQ333337; OsUGlcAE3, GenBank DQ333336) in *E. coli*, the truncated coding regions lacking the putative trans membrane regions (Δ1-68 for AtUGlcAE2, Δ1-53 for

AtUGlcAE3, Δ1-72 for OsUGlcAE1, Δ1-68 for OsUGlcAE2 and Δ1-53 for OsUGlcAE3) were amplified by PCR reactions using the corresponding cDNAs as templates and with the following primers: sense BAB03000#3S 5'CATATGCACCACCTCCACCGTCGTAGCTTCTTATC and antisense BAB03000#2AS 5'GCGGCCGCTTAAGCGGAATCTTCGGCGTGAG for AtUGlcAE2; sense 188BspH1trunS#4 5'CTCATGAGCTTCGTCGATTCCGGTAG and antisense 188kpn1bamh1AS#3 5'GGATCCGGTACCATATGTACAAGCTTGGCTTTAG for AtUGlcAE3; sense R20Nco1S#2 5'CCATGGCCCCGCGGGCGGCGCCCGACTC and antisense R20BamH1AS#1 5'GGATCCCTTGCACTGCTTGTGTTGCCACTCCG for OsUGlcAE1; sense R10BspH1S#2 5'TCATGAGCCACTCCTTCCACCTCTCCTCCG and antisense R10BamH1AS#1 5'GGATCCGACGCGGCGGACATGGCCATGG for OsUGlcAE2; sense R30BspH1S#2 5'CTCATGAGCTTCCAGTCCTTCGTCGACACCTC and antisense R30BamH1AS#1 5'GGATCCAAGTTCTTGGAGCCCCTGGTGTAGCC for OsUGlcAE3. The corresponding PCR products were subcloned into pCR2.1-TOPO plasmid or pCR4-TOPO plasmid to generate pCR2.1:63, pCR4:188.2, pCR4:R20tc.3, pCR4:R10t.1 and pCR4:R30t.3, respectively. Upon confirmation of each plasmid construct by DNA sequencing, the Nde I – Not I fragment from pCR2.1:63, the BspH I – BamH I fragment from pCR4:188.2, the BspH I – BamH 1 fragment from pCR4:R10t.1, the Nco I – BamH I fragment from pCR4:R20tc.3 and the BspH I – BamH I fragment from pCR4:R30t.3 were subcloned into Nde I/Not I or Nco I/BamH I cloning sites of the pET28b *E. coli* expression vector (Novagen, Madison, WI) resulting in clones pET28b:63.5AC, pET28b:188.1, pET28b:R10t.3, pET28b:R20t.1 and pET28b:R30t.9, respectively.

For ZmUGlcAE3 expression studies, total RNA was isolated from 8 cm long ears, the tenth leaf (from the bottom), or from the roots and whole seedling of 1-week-old maize plants

(ecotype w23). RNA (3 μg) was reverse transcribed as described (Gu and Bar-Peled 2004), and one-twentieth of each of the reverse-transcribed products was used as a template for PCR reactions using 0.5 units Taq DNA polymerase (Roche), the manufacturer's buffer, 0.2 mM dNTP, 1.5 mM $MgCl_2$, 5% dimethyl sulfoxide, 1M betaine and 0.2 mM ZmUGlcAE3 gene-specific sense and antisense primers (see above). The transcript of ubiquitin (Livak and Schmittgen 2001) was used as internal RT-PCR control. PCR conditions were: 1 cycle at 94°C for 3 min; and 30 cycles (94°C, 30 s; 62°C, 30 s; 70°C, 2.5 min) and a final extension at 70°C for 5 min. 20 more cycles were performed for leaf samples to yield enough products to check the ZmUGlcAE3 / ubiquitin expressions. One-tenth of each sample and the DNA MW marker (1 KB plus; Invitrogen) were resolved on a Tris-acetate EDTA-1% agarose gel and visualized by staining with ethidium bromide.

Protein Expression, Purification, and Mass Determination

Cultures of *E. coli* carrying the pET30a:Zmt#14.2 ($ZmUGlcAE3_{\Delta 1-47}$), pET28b:63.5AC ($AtUGlcAE2_{\Delta 1-68}$), pET28b:188.1 ($AtUGlcAE3_{\Delta 1-53}$), pET28b:R20t.1 ($OsUGlcAE1_{\Delta 1-72}$), pET28b:R10t.3 ($OsUGlcAE2_{\Delta 1-68}$), pET28b:R30t.9 ($OsUGlcAE3_{\Delta 1-53}$) or control empty vector, were used to inoculate 0.5 L Luria-Bertani liquid broth supplemented with 50 μg/ml kanamycin and 30 μg/ml chloramphenicol. Cell growth, induction, and protein isolation were as described (Gu and Bar-Peled, 2004). For unknown reasons, the activities of recombinant $OsUGlcAE1_{\Delta 1-72}$ and $OsUGlcAE2_{\Delta 1-68}$ were too low to be used for enzymatic characterization. Estimation of the molecular weight of the native recombinant proteins was carried out on a gel-filtration column [1cm x 90cm, packed with Superdex75 prep grade; Amersham Bioscience, Piscataway, NJ] as described (Gu and Bar-Peled, 2004). The active fractions of the recombinant proteins were further fractionated on an anion-exchange column [0.5cm x 5cm, packed with Source Q15;

Pharmacia, Piscataway NJ] (Gu and Bar-Peled, 2004). Proteins were separated by 0.1% SDS-12% PAGE (Bar-Peled et al., 1991) alongside molecular weight markers (BioRad) and visualized by Coomassie blue staining (Invitrogen).

Enzyme Assay, HPLC and NMR Product Analysis

UDP-GlcA epimerase activity was initiated by adding UDP-GlcA (1 mM) to a reaction mix (a total of 50 μl) consisting of 0.1M sodium phosphate, pH 7.6, 20% (v/v) glycerol and 1 μg/ml recombinant epimerase. The assay was carried out for 30 min at 37°C (unless otherwise indicated), and the reaction was terminated by the addition of 50 μl of chloroform. Assay products were determined by HPLC as described previously (Gu and Bar-Peled, 2004). Proton NMR spectroscopy of reaction product isolated by HPLC was performed at 25°C on a Varian (Palo Alto, CA) Inova spectrometer operating at 500 MHz. The chemical shifts (δ) are reported in ppm relative to external acetone (2.224 ppm).

CHAPTER 4

DISTINCT TYPES (A, B AND C) OF PLANT UDP-GLCA 4-EPIMERASE ISOFORMS ARE ALL TYPE II MEMBRANE PROTEINS LOCALIZED IN THE GOLGI[3]

[3]Xiaogang Gu and Maor Bar-Peled, to be submitted to Plant Physiol or Planta

ABSTRACT

GalA containing glycans, including pectic polysaccharides and xylan, play important roles in plant growth and development. In plants, the biosynthesis of these glycans requires UDP-galacturonic acid (UDP-GalA) as an activated sugar donor. UDP-GlcA 4-epimerase (UGlcAE) converts UDP-glucuronic acid (UDP-GlcA) to UDP-GalA. Recent work on UGlcAE isoforms among different plant species reveals that each plant has multiple UGlcAE isoforms, which can be divided into three evolutionary clades: type A, B, and C. All plant UGlcAE proteins have two domains: a putative N-terminal "targeting domain" and a large C-terminal catalytic domain. To understand the reason for the existence of multiple UGlcAE isoforms, representatives from the three types of Arabidopsis UGlcAE (AtUGlcAE) isoforms were biochemically characterized and their sub-cellular localizations were further studied. All isoforms (type A, B and C) are integral membrane proteins, and appear to be N-glycosylated. Protease protection assays confirm that the large catalytic domain is facing the organelle lumen and confocal microscopy analyses further demonstrate that EYFP tagged AtUGlcAE isoforms are localized to the Golgi apparatus.

INTRODUCTION

In plants, various pectic polymers, such as homogalacturonan (HGA), rhamnogalacturonan I (RG-I), rhamnogalacturonan II (RG-II), xylogalacturonan (XGA) and apiogalacturonan (AGA), contain galacturonic acid (GalA) sugar residues. GalA residues are also found in other plant glycans such as the "Sequence 1 Structure domain" [β-D-Xyl-(1,3)-α-L-Rha-(1,2)-α-D-GalA-(1,4)- β -D-Xyl] of xylan (Andersson et al., 1983; Pena et al., 2007) and arabinogalactan-proteins (AGPs) (Pellerin et al., 1995).

UDP-α-D-GalA (UDP-GalA) is a key precursor for the synthesis of pectins (Sterling et al., 2007) and is proposed to be the donor for the addition of the GalA-residue to the Sequence 1 Structure domain of xylan (Pena et al., 2007; Mohnen, 2008). In plants, the synthesis of UDP-GalA is catalyzed by numerous isoforms of UDP-GlcA 4-epimerase (UGlcAE) (Gu and Bar-Peled, 2004). The identification and biochemical characterization of multiple plant UGlcAE isoforms (Molhoj et al., 2004; Usadel et al., 2004; Gu and Bar-Peled, 2004; Gu Ph.D. dissertation Chaper 3), however, did not provide information about where the conversion from UDP-GlcA to UDP-GalA occurs *in vivo*. Although the activity of UGlcAE was reported to be co-fractionated with the Golgi on sucrose-gradients (Sterling et al., 2001), it did not confirm whether the catalytic activity of UGlcAE resides in the cytosol, or inside the Golgi. Here we report that all UGlcAE isoforms (type A, B, and C) (Gu Ph.D. dissertation Chaper 3) are integral membrane proteins with their catalytic domains in the Golgi lumen.

RESULT

EYFP tagged AtUGlcAEs are membrane proteins

ARAMEMNON is a program (Schwache et al. 2003) used to predict sub-cellular localizations of plant proteins. The program also predicts if a given protein is a membrane protein and if so, what its topology is. The output of ARAMEMNON relies on several different computer programs, which provides predictive information to generate hypotheses related to the function of the studied protein.

In plants, the three types of UGlcAE isoforms (A, B and C; Gu Ph.D. dissertation Chaper 3; Gu Ph.D. dissertation Appendix A) contain two distinct regions: a variable N-terminus with a putative transmembrane domain and a large conserved C-terminal catalytic domain (Gu and Bar-Peled, 2004). Based on the membrane topology prediction (ARAMEMNON version 5.2), three out of ten programs (Table 4.1) predict that the catalytic domain of AtUGlcAE3 (At4g30440, a type C UGlcAE isoform) resides in the cytoplasm and seven other programs predict the catalytic domain of the same protein to reside in the organelle lumen (Table 4.1). Similar uncertainty for membrane topology prediction is also obtained for the type A UGlcAE isoform, (AtUGlcAE1, At2g45310); two programs predict the C-terminus region of the protein to face the cytosol, while eight other programs predict the N-terminus region to face the cytosol (Table 4.1).

The ARAMEMNON sub-cellular localization prediction of UGlcAE isoform is also ambiguous. Based on the seventeen programs within the ARAMEMNON that are used for sub-cellular localization prediction, AtUGlcAE1 is predicted to reside either in the chloroplast or in the secretory pathway (Table 4.2). AtUGlcAE3 is predicted to reside in the secretory pathway and AtUGlcAE2 (At3g23820, a type B UGlcAE isoform) is suggested to localize in the

Table 4.1 ARAMEMNON (Ver 5.2) prediction for the membrane topology of three types Arabidopsis UGlcAE isoforms

ARAMEMNON programs	AtUGlcAE1 (Type A)	AtUGlcAE2 (Type B)	AtUGlcAE3 (Type C)
Alom_v2	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus
HmmTop_v2	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus
MemSat_v3	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Cytoplasmic C-terminus
PHDhtm	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Cytoplasmic C-terminus
Phobius	N/A	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus
S-Tmhmm_v0.9	Cytoplasmic C-terminus	Cytoplasmic C-terminus	Cytoplasmic C-terminus
THUMBUP_v1	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus
THUMM_v2	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus
TMMOD	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus	Non-cytoplasmic C-terminus
ConPrep_v2	Cytoplasmic C-terminus	Cytoplasmic C-terminus	Non-cytoplasmic C-terminus

Table 4.2 Prediction for the sub-cellular localization of three types of Arabidopsis UGlcAE isoforms using ARAMEMNON

[a]The consensus prediction confidence value was calculated by ARAMEMNON. A higher value indicates a stronger confidence.

		Consensus prediction value[a]
AtUGlcAE1		
(type A)	Chloroplast localization	5.6
	Mitochondrion localization	0
	Secretory pathway localization	3.4
AtUGlcAE2		
(type B)	Chloroplast localization	2.7
	Mitochondrion localization	1.3
	Secretory pathway localization	4.7
AtUGlcAE3		
(type C)	Chloroplast localization	0.0
	Mitochondrion localization	0.6
	Secretory pathway localization	5.3

mitochondria, chloroplast or the secretory pathway (Table 4.2). As the secretory system contains many distinct organelles (i.e. ER, Golgi, vacuole and plasma-membrane), prediction that a protein resides in the "secretory system" is also ambiguous. As summarized in Table 4.1 and 4.2, the outcome of these predictions suggests that different UGlcAE isoforms (type A, B or C) exist because each may be localized to a different sub-cellular organelle and may have different membrane topology (i.e. the catalytic domain of one UGlcAE may reside in the cytosol while in another isoform, the catalytic domain is in the lumen of an organelle). We decided to evaluate these predictions experimentally since such information is critical for a better understanding of UDP-GalA synthesis and it is also beneficial to evaluate how and where pectic polysaccharides are made within a cell.

In order to validate the membrane topology and the sub-cellular localization of plant UGlcAEs, different types of Arabidopsis isoforms (AtUGlcAE1, AtUGlcAE2 and AtUGlcAE3) were expressed with Enhanced Yellow Fluorescent Protein (EYFP) fused to their C-terminal regions. The nature of the protein (i.e. membrane or soluble), the topology (i.e. cytoplasmic C-terminus or lumenal C-terminus) and the sub-cellular localization (i.e. ER, Golgi or chloroplast) of AtUGlcAE-EYFP were subsequently determined by examining these expressed recombinant proteins in stably transformed Arabidopsis or transiently transformed *N. bentamiana* plants .

Preliminary analyses of five randomly selected stably transformed Arabidopsis plants for each isoform revealed that all plants expressing the same recombinant isoform show no abnormal phenotypes when they are compared to the wild type plants. Their seeds were further collected and used in the studies described in this report. To determine if an isoform is cytosolic or membrane associated, seedlings of transgenic Arabidopsis lines expressing each AtUGlcAE-EYFP were harvested and used for preparation of soluble and microsomal fractions (Pattathil et

al., 2005). Both fractions were subsequently analyzed for the presence of the recombinant isoform by western blot using anti-GFP antiserum and by HPLC for the enzymatic activity. As shown in Fig 4.1A, all three types of recombinant isoforms were detected in the microsomal fraction, but not in the soluble fraction. Similarly, UGlcAE activity was only detected in microsomes (Fig. 4.1B). In addition, all transgenic plants expressed higher UGlcAE specific activity when compared with wild-type plants (Fig. 4. 1B). Therefore, based on western blot and enzymatic activity data, we suggest that all three types of UGlcAE isoforms are microsomal proteins. These biochemical data however, do not indicate whether the studied UGlcAE isoform is a peripheral membrane protein, an integral membrane protein, or a soluble protein within the organelle lumen.

AtUGlcAEs are integral membrane proteins

To elucidate whether AtUGlcAE isoforms are integral membrane proteins, various treatments to solubilize recombinant isoforms out of the microsomes were examined. For these experiments, microsomes were isolated from transgenic Arabidopsis lines expressing a tagged soluble ER-lumen marker (CFP-HDEL, Nelson et al., 2007) or recombinant AtUGlcAE isoforms. Equivalent amounts of microsomes were re-suspended either in 0.1 M sodium carbonate, extraction buffer alone, or in buffers containing one of the following additives: the nonionic detergent Triton X-100 (2%, v/v), the ionic detergent sarkosyl (1%, w/v), 2 M urea, or 1 M NaCl. After incubation, the treated samples were centrifuged and the presence of AtUGlcAE isoforms or ER marker in the resulting supernatants was determined by western blot analyses with anti-GFP antiserum (Fig. 4.2). Salt treatment, which is known to release cytosolic proteins that ionically interact with the membrane, did not release AtUGlcAE-EYFP or the ER-marker, indicating that recombinant AtUGlcAE isoforms are not peripheral membrane proteins (Fig. 4.2).

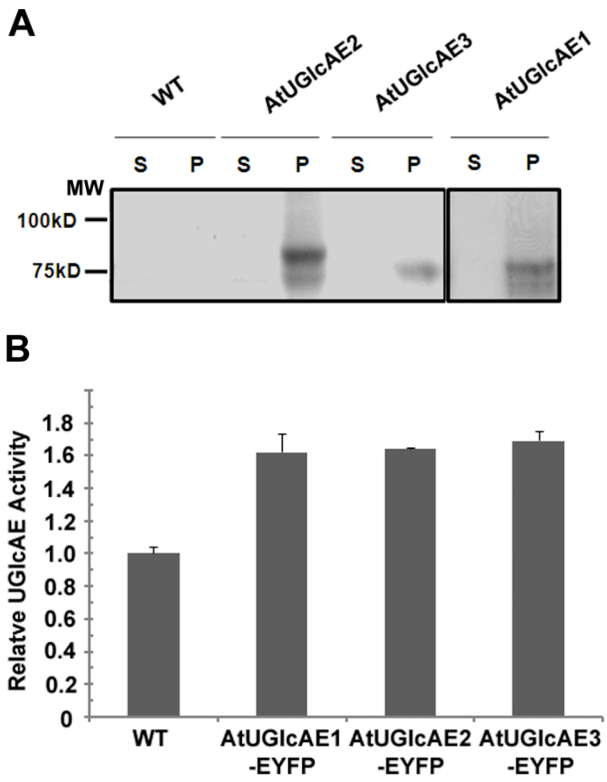


Figure 4.1. AtUGlcAEs are membrane associated proteins.

A**,** Western blot analysis of soluble (S) and microsomal (P) proteins isolated from wild type (WT) or transgenic Arabidopsis seedlings expressing AtUGlcAE1-EYFP, AtUGlcAE2-EYFP or AtUGlcAE3-EYFP. Proteins (15µg) of each fraction were separated by SDS-PAGE and western blots analysis was performed with anti-GFP antiserum.

B, UGlcAE activity of equivalent microsomal protein preparations isolated from five-day old WT or transgenic AtUGlcAE-EYFP Arabidopsis seedlings. The data are the relative amount of UDP-GalA produced per hour compared with control (WT plant). These data are the means of triplicate reactions.

The soluble ER-lumen marker (CFP-HDEL) was extracted by carbonate and urea (Fig 4.2), which are expected to solubilize soluble lumen proteins and peripheral membrane proteins (Bar-Peled and Raikhel 1997; Rancour et al., 2002; Pattathil et al., 2005). Urea and alkali treatments, however, did not release recombinant AtUGlcAE isoforms. Only detergents (Fig 4.2) solubilized AtUGlcAEs indicating they are integrated microsomal membrane proteins.

The catalytic domains of AtUGlcAEs are in the lumen of organelle

As mentioned above, membrane topology prediction programs arrive at no consensus to address if the catalytic domain of AtUGlcAE protein is located in the cytoplasm or in the organelle lumen.

To determine the membrane topology of AtUGlcAEs, protease protection assays were carried out. Microsomes extracted from transgenic plants were treated with proteinase K (PK) in the presence or absence of detergent and analyzed by immunoblotting with anti-GFP antiserum (Fig 4.3). In the presence of detergent, AtUGlcAEs-EYFP were degraded by PK (Fig 4.3), however, in the absence of detergent, the EYFP domain of recombinant isoforms were "protected" by the endomembrane from PK dgradation. Since the EYFP domain is fused to the C-terminal domain (the catalytic domain) of the AtUGlcAE isoform, the PK analysis suggests that all types of recombinant isoforms (A, B or C) are membrane proteins with their catalytic domains in the organelle lumen.

While analyzing the amino acid sequences of AtUGlAEs, we noticed several putative N-glycosylation sites within the C-terminal domains of AtUGlcAE isoforms (i.e. N^{190}PS, N^{357}TS and N^{397}IS for AtUGlcAE1; N^{372}TS and N^{412}VS for AtUGlcAE2; N^{348}TS and N^{388}IS for AtUGlcAE3), but not within their short N-terminal domains. It is assumed that if a protein

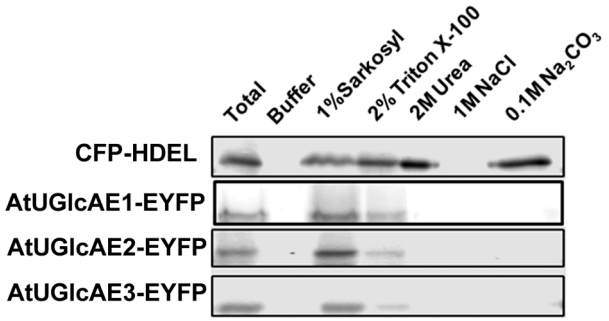


Figure 4.2. AtUGlcAEs are integral membrane proteins.

Microsomes were prepared from five-day old Arabidopsis seedlings. Membrane (15 mg) was resuspended in 100 µl 0.1 M sodium carbonate (pH 11), or in 100µl extraction buffer containing 1% (w/v) Sarkosyl, 2% (v/v) Triton X-100 , 2 M urea, 1 M NaCl or water (buffer alone). After 30 min incubation on ice, the solubilized membranes were centrifuged (100,000 g) and aliquot (20µl) of each supernatant was analyzed by western blot analysis with anti-GFP antiserum. The experiment has been repeated more than three times and representative data are presented.

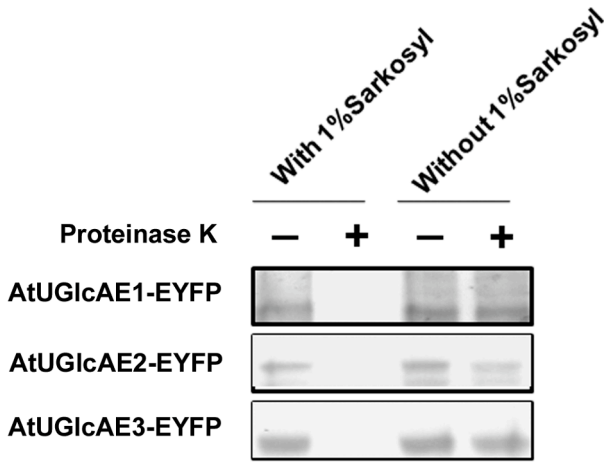


Figure 4.3. UGlcAEs are integral membrane proteins with catalytic domains located in the lumen of endomembranes.

Microsomes (15 mg) isolated from five-day old transgenic Arabidopsis seedlings were treated with proteinase K (PK) in the presence or absence of detergent (Sarkosyl). After PK digestion, proteins were analyzed by western blot using anti-GFP antiserum.

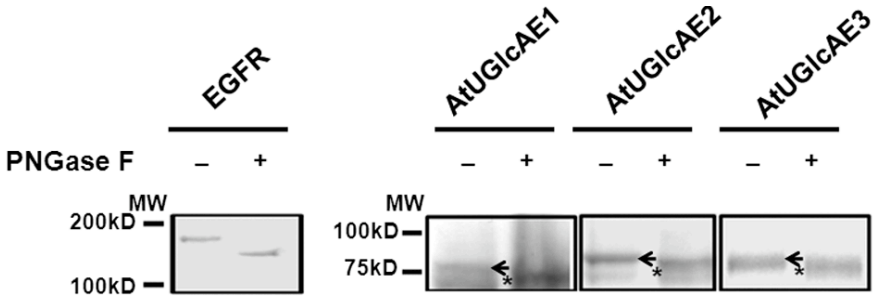


Figure 4.4. The catalytic domains of AtUGlcAEs are N-glycosylated.

15 µg microsomal proteins isolated from five-day old transgenic seedlings were treated with either 500 unit PNGase F or water at 37^{o}C for 1.5 hours. 10 µg epidermal growth factor receptor (EGFR) was used as a positive N-glycosylated protein control. Treated proteins were separated by SDS-PAGE and immunoblotted with anti-GFP antiserum for recombinant isoforms, or anti-EGFR antiserum for EGFR. The glycosylated form is indicated by an arrowhead and the de-glycosylated form is indicated by a star.

is N-glycosylated, treatment with Peptide: N-Glycosidases F (PNGF) (Maley et al., 1989) will result in its unglycosylated-form with a reduced molecular weight (MW) compared to its glycoslyated-form. In addition, if the isoforms are N-glycosylated (as determine by PNGF assay) it may provide an additional proof for membrane topology because the addition of *N*-linked glycans to proteins is known to occur inside the endomembrane lumen (van Geestand Lolkema, 2000).

Microsomes isolated from transgenic plants were treated with PNGF (Fig 4.4), and the treated proteins were separated by SDS-PAGE and analyzed by immunoblots. A clear shift in the MW of recombinant AtUGlcAE protein treated by PNGF was observed when compared with untreated control protein (Fig. 4.4), indicating that the C-terminal domains of AtUGlcAE isoforms are N-glycosylated and thus, providing additional evidence that the catalytic domains are in the organelle lumen.

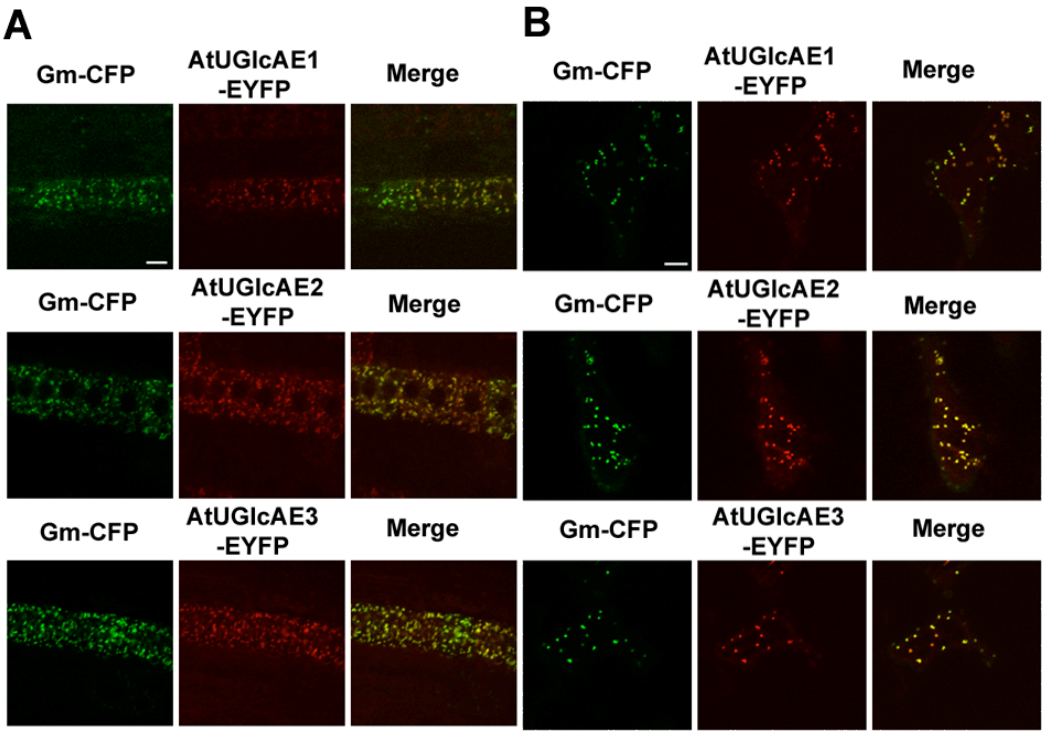
Taken together, the data from protease protection and deglycosylation assays let us to conclude that all three types of membrane bound AtUGlcAEs have their catalytic domains in the lumen of an organelle. While ARARMON predicts other 'TM" regions (data not shown), this is unlikely since truncated UGlcAE isoforms (AtUGlcAE1$_{\Delta 1\text{-}64}$, AtUGlcAE2$_{\Delta 1\text{-}68}$ and AtUGlcAE3$_{\Delta 1\text{-}53}$, lacking the N-terminal region) were expressed as soluble protein in *E. coli* and active (Gu and Bar-Peled, 2004; Gu's Ph.D. dissertation chapter 3). Therefore, we propose that all types of plant UGlcAEs are type II transmembrane proteins.

AtUGlcAEs are located in the plant Golgi

To determine the sub-cellular localization of different types of UGlcAE isoforms, stably transformed Arabidopsis plants expressing type-specific AtUGlcAE-EYFP were further transformed with either an ER marker (CFP-HDEL), or a Golgi marker (Gm-CFP) (Saint-Jore-

Dupas et al., 2006). Confocal microscopy analyses (Fig 4.5) were used to determine the localization of AtUGlcAE isoforms and the markers. Arabidopsis cells co-expressing AtUGlcAE-EYFP and Gm-CFP displayed a punctate fluorescence pattern (Fig 4.5A), which is also observed for other Golgi-localized proteins such as Uxs2 (Pattathil et al., 2005), Sed5 (Sanderfoot et al., 2000), and Memb11-YFP (Chatre et al., 2005). Overlaying (merging) (Fig 4.5A) the Gm fluorescent signals with the signals of AtUGlcAE-EYFP, showed that those fluorescent dots are well matched, suggesting that both Gm and AtUGlcAE proteins are co-localized (Fig 4.5A). Similar fluorescence data were obtained when experiments were conducted using transient expression in tobacco leaves (Fig 4.5B). Stably transformed Arabidopsis cells co-expressing the recombinant isoforms with ER marker (CFP-HDEL), however, displayed different and non-overlapping fluorescent signal patterns (Fig 4.5C). While AtUGlcAE-EYFP had a punctate pattern signal, the fluorescent signal derived from the ER marker (CFP-HDEL) had a distinct network-like pattern similar to that observed for other ER proteins, such as CmERS1 (Ma et al., 2006) and AtERdj3A (Yamamoto et al., 2008). Thus, we proposed that recombinant AtUGlcAE isoforms are co-localized with Golgi marker (Gm-CFP) but not with ER marker (CFP-HDEL).

Brefeldin A (BFA) has been reported to block the secetory pathway and to inhibit the formation of Golgi apparatus (Ritzenthaler et al., 2002; Yuasa et al., 2005). Therefore, as reported by various laboratories (Ritzenthaler et al., 2002; Yuasa et al., 2005; Chatre et al., 2005; Matheson et al., 2007), upon BFA treatment, Golgi-resident proteins would accumulate in the ER while an ER-resident protein would not change its pattern with BFA. To evaluate the response of AtUGlcAE isoforms to BFA treatments, transgenic Arabidopsis seedlings expressing either recombinant AtUGlcAE isoforms or different markers, were examined (Fig 4.6). ER-


A
Gm-CFP
AtUGlcAE1 -EYFP
Merge
Gm-CFP
AtUGlcAE2 -EYFP
Merge
Gm-CFP
AtUGlcAE3 -EYFP
Merge
B
Gm-CFP
AtUGlcAE1 -EYFP
Merge
Gm-CFP
AtUGlcAE2 -EYFP
Merge
Gm-CFP
AtUGlcAE3 -EYFP
Merge

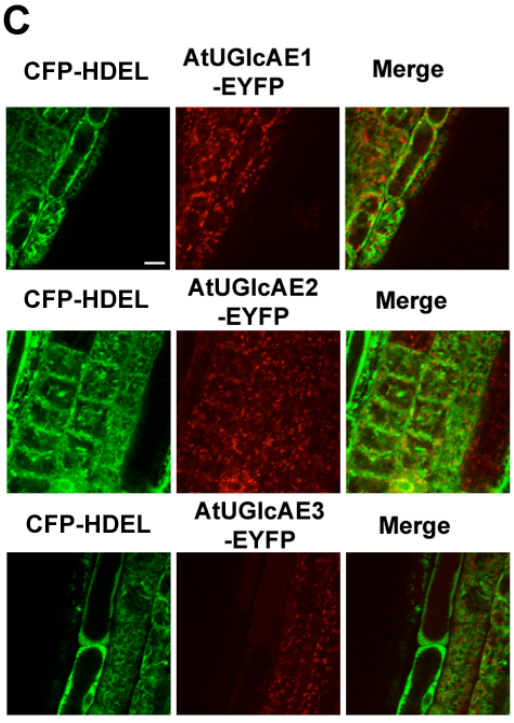

C
CFP-HDEL
AtUGlcAE1 -EYFP
Merge
CFP-HDEL
AtUGlcAE2 -EYFP
Merge
CFP-HDEL
AtUGlcAE3 -EYFP
Merge

Figure 4.5. AtUGlcAEs are co-localized with plant Golgi marker, but not with the ER marker.

A, Co-localization of AtUGlcAEs-EYFP with Gm-CFP in Arabidopsis root cells. The root cells of five-day old transgenic Arabidopsis seedlings co-expressing AtUGlcAE-EYFP and Gm-CFP were analyzed by the confocal microscopy. Left, CFP fluorescence; Center, EYFP fluorescence; right, merged left and center images.

B, Co-localization of AtUGlcAEs-EYFP and Gm-CFP in tobacco leaves. Recombinant isoform and Gm-CFP were transiently expressed in the lower epidermis cells of tobacco leaves and analyzed by the confocal microscopy. Left, CFP fluorescence; Center, EYFP fluorescence; right, merged left and center images.

C, Co-localization of AtUGlcAEs-EYFP with CFP-HDEL in Arabidopsis root cells. The root cells of five-day old transgenic Arabidopsis seedlings co-expressing AtUGlcAE-EYFP and CFP-HDEL were analyzed by the confocal microscopy. Left, CFP fluorescence; Center, EYFP fluorescence; right, merged left and center images. Bar, 10 μm.

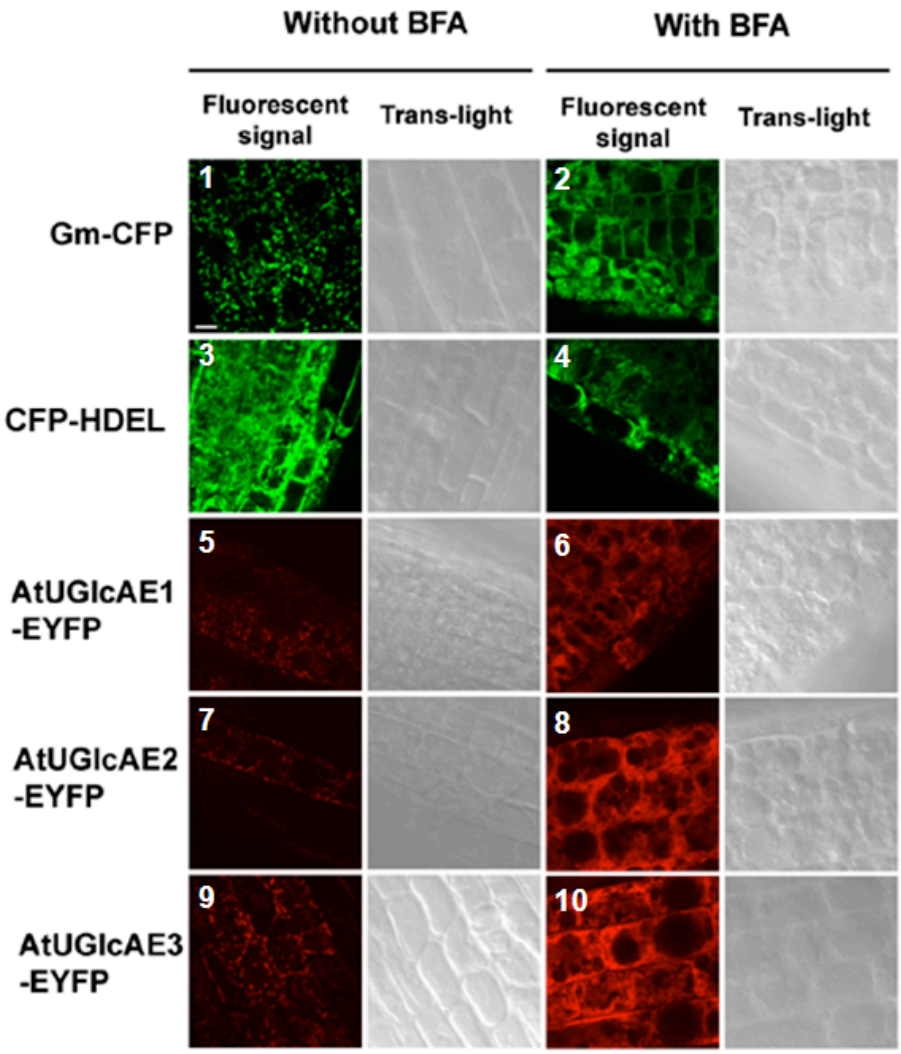


Figure 4.6. AtUGlcAEs are localized in the plant Golgi. With or without 2 hours BFA treatment, five-day old seedlings' root cells of the transgenic Arabidopsis lines expressing either recombinant isoform, Gm-CFP or HDEL-CFP, were analyzed by the confocal microscopy. The trans-light images indicate the root cells that are observed. Bar, 10 μm.

marker protein (CFP-HDEL) had similar fluorescent signal pattern (i.e. diffused network-like) in plants treated with or without BFA (Fig. 4.6 panel 3 and 4). On the other hand, different fluorescent signal patterns were observed for plants expressing AtUGlcAE-EYPF in the presence or absence of BFA (Fig. 4.6). Note the punctate pattern (Fig. 4.6 panel 5, 7, 9) in plant samples expressing AtUGlcAE-EYFP without the BFA treatment versus the diffusive pattern (Fig. 4.6 panel 6, 8, 10) observed after the samples were treated by BFA. This diffusive fluorescent pattern is somewhat similar to the network-like pattern observed in the plant samples expressing CFP-HDEL (Fig. 4.6 panel 4). Comparing the fluorescent pattern of AtUGlcAE-EYFP with plant samples expressing Golgi marker (Gm-GFP) (Fig. 4.6 panel 1 and 2) further confirms our conclusions that AtUGlcAE is a Golgi resident protein. Taken together, comparative confocal microscopy analyses of ER marker (CFP-HDEL), Golgi marker (Gm-CFP) and recombinant AtUGlcAE isoforms led us to conclude that all recombinant isoforms (type A, B and C) are localized in the plant Golgi.

DISCUSSION

Based on the data reported here, we propose that plant UGlcAE isoforms (type A, B, and C) are type II membrane proteins with their catalytic domains within the Golgi lumen. This catalytic orientation indicates that UDP-GlcA imported into the Golgi membrane lumen is converted to UDP-GalA, which then can be used by glycosyltransferases for the syntheses of GalA containing glycans.

The PNGF analyses indicate that all three types of UGlcAE isoforms are N-glycosylated. In some glycoproteins, N-glycosylation was critical for their enzymatic activitie. For example, N-glycosylation was required for the catatlytic activity of GD3 synthase (Martina et al., 1998) and the enzymatic activity of galactosylceramide sulphotransferase (CST) was significantly

reduced in the absence of its N-glycans (Echhardt et al., 2002). In addition, *N*-glycosylation was reported to be crucial for the plasma membrane localization of human nephrin (Yan et al., 2002) and deleting the N-glycosylation of sialyltransferase II (ST-II) was reported to shift the ST-II subcellular localization from Golgi to ER (Bieberich et al., 2000). The functions of N-glycosylation for AtUGlcAEs remain unclear. The enzyme (when expressed as recombinant protein in *E.coli*) is active without being N-glycosylated . Whether the N-glycosylation is important for the Golgi localization of those isoforms will be addressed in the future by specific site directed mutagenesis.

The ARAMEMNON predictions (see Table 4.1and Table 4.2) clearly required experimental validation. Relying on the predictions alone could draw wrong conclusions for the existence of various UGlcAE isoforms. Although this study indicates that all plant UGlcAE isoforms are in the Golgi, we did not address, however, if the protein is localized to the *cis*-, *medial*- or *trans*-Golgi compartments. Current work in our laboratory is underway to address the sub-cellular localization using immuno-electron microscopy along with studies aimed at identifying amino acids that are likely required for Golgi localization.

MATERIALS AND METHODS

Molecular cloning and plant transformation

The Arabidopsis cDNA prepared in Chapter 3 was used to amplify the full length AtUGlcAE1, AtUGlcAE2 and AtUGlcAE3 with the following primers: sense 120#1S 5’ catatgtcacgtcttgacgacataccttc and antisense 120 AS#8 5’ ggtaccgctctagcggcggctttttttgtcgcc for AtUGlcAE1; sense BAB03000#1S 5' catatgcccctgtcggcgacggcggatac and antisense #60Kpn1/AS#7 5’ tggtaccgcggaatcttcggcgtgagaag for AtUGlcAE2; sense cab79762BspHI/S#1 5’ atcatgatgccttcaatagaagatgagctgtttc and antisense Kpn/AS#1 5’

tggtaccatatgtacaagcttggctttagtattg for AtUGlcAE3 as described by Gu and Bar-Peled (2004). The corresponding amplified products were subcloned into pCR4-TOPO plasmid (Invitrogen) or pCR2.1-TOPO plasmid (Invitrogen) to generate pCR2.1: 127#1, pCR4:67#4 and pCR2.1:183#2, respectively. After sequencing verification, the Nco1-Kpn1 fragment from pCR2.1: 127#1, BspH1-Kpn1 fragment from pCR4:67#4 and BspH1-Kpn1 fragment from pCR2.1 183#2 were subcloned into Nco1/Kpn1 cloning site of pCam35stl-EYFPs2#4 (Pattathil et al., 2005) to generate pCam35stl-127-EYFPS2#4.1, pCam35stl-67-EYFPS2#4.1 and pCam35stl-187-EYFPS2#4.1, respectively.

Plant transformation

The following constructs were used to generate transgenic plants expressing the chimeric fluorescent protein: pCam35stl-127-EYFPS2#4.1 for AtUGlcAE1-EYFP, pCam35stl-67-EYFPS2#4.1 for AtUGlcAE2-EYFP, pCam35stl-187-EYFPS2#4 for AtUGlcAE3-EYFP, PAN61 for Golgi marker Gm-CFP and PAN103 for ER marker CFP-HDEL (The last two plasmids are kindly provided by Dr. Andreas Nebenführ in the University of Tennessee). The constructs were transformed individually to *Agrobacterium tumefaciens* GV3101/pM90 strain (Pattathil et al., 2005), and subsequently transformed to Arabidopsis using the method described by Bechtold et al. (1993). Numerous Arabidopsis plants stably expressing individual fluorescent tagged AtUGlcAE isoform were selected by hygromycin (15 μg/ml final concentration). Arabidopsis plants stably expressing Gm-CFP or CFP-HDEL were selected by kanamycin (30 μg/ml final concentration). Several independent transgenic lines expressing individual fluorescent tagged AtUGlcAE were further transformed with PAN61 or PAN103 plasmid to generate transgenic plants co-expressing each AtUGlcAE-EYFP with either CFP-HDEL or Gm-CFP. The co-expressing transgenic plants were selected by hygromycine (15 μg/ml) and

kanamycin (30 µg/ml). Transformation of tobacco (*Nicotiana tabacum*) cv Petit Havana to transiently co-express specific AtUGlcAE-EYFP with the Golgi marker (Gm-CFP) was carried out as described by Chatre et al., (2005).

Microsome preparation

Transgenic Arabidopsis plants expressing fluorescent tagged proteins were first germinated on 1% (w/v) agar plate containing Murashige-Shoog basal salt media (sigma) supplemented with 1% (w/v) sucrose at pH 5.7. Transgenic seedlings were harvested 5 days after germination; frozen in liquid nitrogen and 1 g tissue was ground to a fine powder with a mortar and pestle. Cold buffer (3 ml, of 50 mM sodium phosphate at pH 7.6, 1 mM EDTA–NaOH, 400 mM sucrose) supplemented with 1 mM DTT and 0.1 M PMSF was used for microsome extraction. After 5 min grinding, the homogenate was centrifuged (1,000 g for 10 min at 4^{o}C) and the top 2.5 ml supernatant layer was re-centrifuged at 100,000 g for 60 min at 4^{o}C. The resulting supernatant (termed S) and the pellet (microsome fraction, termed P) were carefully separated and used for analysis.

To determine if UGlcAEs are integral membrane proteins, 15 mg of microscomes were re-suspended with 100 µl extraction buffer containing either 2% Triton X-100, 1% sarkosyl, 2 M urea, 1 M NaCl, water, or resuspended in 100 µl of 0.1 M sodium carbonate (pH 11). After 30 min on ice, the homogenized microsomes were centrifuged (100,000 g at 4^{o}C for 15 min) to remove undissolved matter, and 20 µl of the supernatant was used for western blot analysis.

Protease protection assay: microsomes (15 mg) were first re-suspended with 50 µl extraction buffer supplemented with or without 1% sarcosyl. After 30 min incubation on ice, 1µg proteinase K (PK) or water (control) was added to the samples to initiate the PK digestion.

After10 min at 30^{o}C, PMSF was added to a final concentration of 2mM to terminate the digestion.

PNGase N-glycosylation assay: microscomes (15 mg) were first resuspended with 50 µl extraction buffer supplemented with 0.5% (w/v) SDS and 40 mM DTT. After 30 min solubilization on ice, samples were boiled (100 ^{o}C for 10 min) and subsequently centrifuged (100,000 g at 4^{o}C for 15 min). The resulting supernatant (15 µl) was used for N-glycosylation assay. The PNGF assays in a final volume of 20 µl consist of 50 mM sodium phosphate (pH 7.6), 1% (w/v) NP-40, 15 µl supernatant sample, supplemented with 500 unit Peptide: N-Glycosidase F (PNGF) (Bio-labs), or water (negative control). After 1.5 hours incubation at 37 ^{o}C, samples were analyzed by western blot.

For Western blot analyses, proteins were separated on 0.1% SDS-10% PAGE, blotted onto 0.45 µm PVDF (polyvinylidene fluoride) membrane in a Hoffer wet blot apparatus using a solution composed of 10 mM Tris,10 mM Glycine, 0.025% SDS, and 20% methanol. Following electroblotting (35 V for 4 hours), membranes were treated for 1 hour with blocking reagent [5% (w/v) dry milk in TBST (20 mM Tris–HCl, pH 7.6, 150 mM NaCl, 0.1% Tween 20)]. Membranes (blots) were washed with TBST and subsequently reacted with rabbit anti-GFP (Abcam, MA, USA) serum diluted 1:5,000 in blocking reagent for 1 h. Blots were then washed with TBST; reacted for 1 h with goat-anti rabbit secondary antibodies linked to alkaline phosphatase (Sigma) at 1:5,000 dilution in blocking reagent. Blots were incubated up to 5 min with alkaline phosphatase color reagents (10 ml , nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP; Sigma) as described (Pattathil et al., 2005). After color was developed, blots were washed with water and scanned.

UDP-GlcA 4-epimerase (UGlcAE) activity

Soluble (S) and microsomal (P) fractions were prepared from wild type and transgenic Arabidopsis plants expressing AtUGlcAE1-EYFP, AtUGlcAE2-EYFP or AtUGlcAE3-EYFP. Microsomes were further re-suspended with extraction buffer containing 2% Triton X-100. The amount of proteins in either the solubilized microsomal fraction (P) or the soluble fraction (S) was quantified with Bio-Rad protein assay kit, using BSA as a standard. UGlcAE activity assays were initiated by adding UDP-GlcA (1 mM final concentration) to a reaction mix (a total volume of 50 µl) consisting of 0.1M sodium phosphate buffer (pH 7.6), 20% (v/v) glycerol and 15 µg either total microsomal protein, or total soluble protein. Assays were carried out for 1 hour at 37°C and the reactions were terminated by the addition of 50 µl chloroform. Assay products were determined by HPLC as described previously (Gu and Bar-Peled 2004).

BFA treatment

Seeds of transgenic Arabidopsis plants were germinated on 1% (w/v) agar plate containing Murashige-Shoog basal salt (MS) media (sigma) supplemented with 1% (w/v) sucrose at pH 5.7. Five days after germination, the seedlings treated with 100µg/ml BFA (BFA stock solution was 5 mg/mL in dimethyl sulfoxide; Sigma) or with DMSO alone (negative control). After 2 hours incubation, the roots were analyzed by confocal microscopy.

Miscroscopy

Confocal imaging was performed using a Leica Confocal laser scanning microscope system [Leica TCS SP2 Spectral Confocal Microscope with Coherent Ti: sapphire multiphoton laser (Mira Optima 900-F)]. For imaging tissues treated with BFA or DMSO alone, the argon ion laser excitation was set at 514 nm for EYFP samples, or 458 nm for CFP samples. Fluorescence was detected using a 475 to 525 nm bandpass filter for CFP and a 530 to 590 nm bandpass filter for EYFP.

For the co-localization studies, excitation lines of an argon ion laser of 458 nm for CFP and 514 nm for EYFP were used alternately with line switching using the line-sequencing scanning facility of the microscope. Fluorescence was detected using a 475 to 525 nm bandpass filter for CFP and a 530 to 590 nm bandpass filter for EYFP.

CHAPTER 5

CONCLUSIONS

Plant species contain multiple UGlcAE isoforms.

Using the amino acid sequence of bacterial Cap1J that encodes a functional UDP-glucuronic acid 4-epimerase, we have identified several candidates of UGlcAE homologs in Arabidopsis (Gu and Bar-Peled, 2004), rice and maize (Gu Ph. D. dissertation Chapter 3; Appendix A). We functionally determined these plant genes to encode active UDP-GlcA 4-epimerases. Analyses of genomic and EST database indicates that multiple isoforms are common in other plant species, for example, six isoforms were identified in *Vitis vinifera*, five isoforms in *Sorghum bicolor*, six isoforms in *Arabidopsis,* seven isoforms in *Populus trichocarpa*, five isoforms in rice and at least three isoforms were identified in maize EST and partial genomic database (Gu Ph.D. dissertation Appendix A). Interstingly, the existence of multiple isoforms of UGlcAE is common in lower plants as well, for example nine isoforms are encoded by the *Physcomitrella* genome and three are found in *Selaginella* (Gu Ph.D. dissertation Appendix A). Phylogenetic analysis of UGlcAE isoforms (Gu Ph.D. dissertation Appendix A), at the amino-acid sequence level, indicates that members of UGlcAE in flowering plants can be divided into three different clades: A, B and C. In Arabidopis, all three distinct types of AtUGlcAE isoforms are transcribed in each individual root cell (Gu Ph.D. dissertation Appendix A), suggesting that diffrent members of this small gene family are required within a single cell.

We concluded that all flowering plant UGlcAEs have two distinct domains: a targeting domain (located at the N-terminal portion of the protein and is about 100 to 120 amino acids

long) followed by a conserved catalytic domain (about 300 amino acids long) (Gu and Bar-Peled, 2004; Gu Ph.D. dissertation Chapter 3). The Arabidopsis UGlcAE isoforms are shown to be N-glycosylated (Gu Ph. D. dissertation Chapter 4). The role of their N-glycosylations, however, is unclear and further research aiming to specifically remove or alter amino acid within the consensus N-glycosylation sites is required.

Plant UGlcAE isozymes have similar enzymatic properties.

The plant UGlcAEs characterized so far, unlike their homolog in *Klebsiella pneumoniae*, are very specific and can only 4-epimerize UDP-GlcA and UDP-GalA (Gu and Bar-Peled, 2004; Gu Ph. D. dissertation Chapter 3). The tested plant UGlcAEs, type A, B or C, are active as dimers and have similar K_m and V_{max} values. These enzymes are inhibited by UDP-xylose and UDP, but not by UMP, UDP-glucose or UDP-galactose (Gu Ph. D. dissertation Chapter 3). However, UGlcAE isoforms in different plant species do have different enzymatic properties (Gu Ph. D. dissertation Chapter 3). For example, the maize UGlcAE (type C) displays a much higher inhibition by UDP-arabinose than type C rice and all three types of Arabidopsis homologs, suggesting that the activity of specific plant isoform(s) can be regulated differently in different plants by specific nucleotide-sugars.

Plant UGlcAEs are type II membrane proteins located in the plant Golgi.

At the time that I started my research, the conversion from UDP-GlcA to UDP-GalA was predicted to occur either inside the lumen of the organelles, or in the cytosol (Orellana and Mohnen, 1999). The data in Chapter 4 demonstrate that plant UGlcAEs are type II membrane proteins with their catalytic domains within the lumen. In addition, EYFP fused AtUGlcAEs were shown to be localized to the Golgi (Gu Ph. D. dissertation Chapter 4). Therefore, my

research clarified that the conversion from UDP-GlcA to UDP-GalA occurs inside the Golgi lumen.

REFERENCE

Adeyeye J, Azurmendi HF, Stroop CJ, Sozhamannan S, Williams AL, Adetumbi AM, Johnson JA, Bush CA (2003) Conformation of the hexasaccharide repeating subunit from the Vibrio cholerae O139 capsular polysaccharide. Biochemistry **42**: 3979–3988

Ahola V, Aittokallio T, Vihinen M, Uusipaikka E (2006) A statistical score for assessing the quality of multiple sequence alignments. BMC Bioinformatics **7:** 484

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**: 3389–3402

An J, O'Neill M, Albersheim P, Darvill A. (1994). Isolation and structural characterization of b-D-glucosyluronic acid and 4-O-methyl b-D-glucosyluronic acid-containing oligosaccharides from the cell wall pectic polysaccharide rhamnogalacturonan I. Carbohydr. Res. **252**: 235-243.

Andersson SI and Samuelson O (1983) Structure of the reduced end-groups in spruce xylan. Carbohydrate Research **111**: 283-288

Ankel H, Feingold DS (1966) Biosynthesis of uridine diphosphate d-xylose: II. Uridine diphosphate d-glucuronate carboxy-lyase of *Cryptococcus laurentii*. Biochemistry 5: 182–189

Ankel H, Tischer RG (1969) UDP-D-glucuronate 4-epimerase in blue-green algae. Biochim Biophys Acta **178**: 415–419

Ahn JW, Verma R, Kim M, Lee JY, Kim YK, Bang JW, Reiter WD, Pai HS (2006) Depletion of UDP-D-apiose/UDP-D-xylose synthases results in rhamnogalacturonan-II deficiency, cell wall thickening, and cell death in higher plants. J Biol Chem. **281**:13708-13716

Barber C, Rösti J, Rawat A, Findlay K, Roberts K, Seifert G. 2006. Distinct Properties of the Five UDP-D-glucose/UDP-D-galactose 4-Epimerase Isoforms of *Arabidopsis thaliana*. J Biol Chem. **281**: 17276 – 17285.

Bar-Peled L (2005) A Novel Pathway for polysaccharide Precursor synthesis. University of Georgia, Center for Undergraduate Research Opportunities.

Bar-Peled M (2005) Nucleotide sugars. In M.E. Himmel ed, Biomass Recalcitrance. Blackwell Publishing, Oxford, pp. 134-159

Bar-Peled M, Griffith CL, Doering TL (2001) Functional cloning and characterization of a UDP-glucuronic acid decarboxylase: The pathogenic fungus Cryptococcus neoformans elucidates UDP-xylose synthesis. Proc Natl Acad Sci USA **98**: 12003–12008

Bar-Peled M, Griffith CL, Ory JJ, Doering TL (2004) Biosynthesis of UDP-GlcA, a key metabolite for capsular polysaccharide synthesis in the pathogenic fungus *Cryptococcus neoformans*. Biochem J **381**: 131–138

Bar-Peled M, Lewinsohn E, Fluhr R, Gressel J (1991) UDP-rhamnose: flavanone-7-O-glucoside-2”-O-rhamnosyltransferase. Purification and characterization of an enzyme catalyzing the production of bitter compounds in citrus. J Biol Chem **266**: 20953–20959

Bar-Peled M, Raikhel NV (1997) An efficient method for cloning in-frame fusion protein genes. Anal Biochem **250**: 262–264

Bechtold, N., Ellis, J., and Pelletier, G. (1993). *In planta Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris, Life Sciences **316**:1194-1199.

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A Gene Expression Map of the *Arabidopsis* Root, Science **302**: 1956-1960

Bouton S, Leboeuf E, Mouille G, Leydecker MT, Talbotec J, Granier F, Lahaye M, Hofte H and Truong HN (2002) *QUASIMODO1* Encodes a Putative Membrane-Bound Glycosyltransferase Required for Normal Pectin Synthesis and Cell Adhension in Arabidopsis. Plant Cell. **14**: 2577-2590

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

Burget EG, Verma R, Mølhøj M, Reiter WD. (2003). The biosynthesis of L-arabinose in plants: molecular cloning and characterization of a Golgi-localized UDP-D-xylose 4-epimerase encoded by the MUR4 gene of Arabidopsis. Plant Cell. 15(2):523-531

Chatre L, Brandizzi F, Hocquellet A, Hawes C, and Moreau P (2005) Sec22 and Memb11 Are v-SNAREs of the Anterograde Endoplasmic Reticulum-Golgi Pathway in Tobacco Leaf Epidermal Cells. Plant Physiol **139**: 1244-1254

Chen S, Hajirezaei M, Bornke F. (2005). Differential Expression of Sucrose-Phosphate Synthase Isoenzymes in Tobacco Reflects Their Functional Specialization during Dark-Governed Starch Mobilization in Source Leaves. Plant Physiol. **139**: 1163-1174.

Chomczynski P (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques **15**: 536–537

Dalessandro G, Northcote DH (1977a) Possible control sites of polysaccharides synthesis during cell growth and wall expansion of pea seedlings (*Pisum sativum* L.). Planta **134:** 39-44

Dalessandro G, Northcote, DH (1977b) Changes in enzymic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in sycamore and poplar. Biochem J **162:** 267-279

Darvill AG, Albersheim P, Delmer DP (1980) The primary cell walls of flowering plants. In NE Tobert, ed, The Biochemistry of Plants: A Comprehensive Treatise, Vol 1. Academic Press, New York, pp 92–162

Daude N, Gallaher TK, Zeschnigk M, Starzinski-Powitz A, Petry KG, Haworth IS, Reichardt JK (1995) Molecular cloning, characterization, and mapping of a full-length cDNA encoding human UDP-galactose 4'-epimerase. Biochem Mol Med **56**: 1–7

Davis KR, Hahlbrock K (1987) Induction of Defense Responses in Cultured Parsley Cells by Plant Cell Wall Fragments. Plant Physiol. **84**:1286-1290

Dormann P, Benning C (1996) Functional expression of uridine 5'-diphospho 4-epimerase (E.C. 5.1.3.2) from Arabidopsis thaliana in *Saccharomyces cerevisiae* and *Escherichia coli*. Arch Biochem Biophys **327**: 27–34

Driouich A, Faye L, Staehelin LA. (1993) The plant Golgi apparatus: a factory for complex polysaccharides and glycoproteins. Trends Biochem Sci. **18(6):**210-4

Egelund J, Peterson BL, Motawia MS, Damager I, Faik A, Olsen CE, Ishii T, Ulvskov P and Geshi N (2006) Arabidopsis thaliana RGXT1 and RGXT2 Encode Golgi-Localized (1,3)-α-D-Xylosyltransferases Involved in the Synthesis of Pectic Rhamnogalacturonan-II. Plant Cell. **18**: 2593-2607

Field R. and Naismith J. (2003) Structural and Mechanistic Basis of Bacterial Sugar Nucleotide-Modifying Enzymes. Biochemistry. **42**: 7637-7647

Feingold DS, Neufeld EF, Hassid WZ (1960) The 4-epimerization and decarboxylation of uridine diphosphate D-glucuronic acid by extracts from Phaseolus aureus seedlings. J Biol Chem 235: 910-913

Feingold DS, Avigad G (1980) Sugar nucleotide transformations in plants. *In* PK Stumpf, EE Conn, eds, The Biochemistry of Plants, Vol 3. Academic Press, New York, pp 101–170

Feng L, Tao J, Guo H, Xu J, Li Y, Rezwan F, Reeves P, Wang L (2004) Structure of the *Shigella dysenteriae* 7 O antigen gene cluster and identification of its antigen specific genes. Microb Pathog **36**: 109–111

Forsberg LS, Carlson RW (1998) The structures of the lipopolysaccharides from *Rhizobium etli* strains CE358 and CE359. The complete structure of the core region of R. etli lipopolysaccharides. J Biol Chem **273**: 2747–2757

Freshour G, Bonin CP, Reiter WD, Albersheim P, Darvill AG, Hahn MG (2003) Distribution of Fucose-Containing Xyloglucans in Cell Walls of the *mur1* Mutant of *Arabidopsis*. Plant Physiology **131**: 1602-1612

Frirdich E and Whitfield C. (2005). Characterization of Gla_{KP}, a UDP-galacturonic acid C4-epimerase from *Klebsiella pneumoniae* with extended substrate specificity. J Bacteriol. **187**: 4104-4115.

Frirdich E, Bouwman C, Vinogradov E and Whitfield C. (2005). The role of galacturonic acid in outer membrane stability in *klebsiella pneumoniae*. J Biol Chem. 280: 27604-27612.

Garcia E, Garcia P, Lopez R (1993) Cloning and sequencing of a gene involved in the synthesis of the capsular polysaccharide of *Streptococcus pneumoniae* type 3. Mol Gen Genet **239**: 188–195

Garcia E, Arrecubieta C, Munoz R, Mollerach M, Lopez R (1997) A functional analysis of the Streptococcus pneumoniae genes involved in the synthesis of type 1 and type 3 capsular polysaccharides. Microb Drug Resist **3**: 73–87

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

Gaunt MA,Maitra US, Ankel H (1974) Uridine diphosphate galacturonate 4-epimerase from the blue-green alga Anabaena flos-aquae. J Biol Chem **249**: 2366–2372

Gloaguen V, Wieruszeski JM, Strecker G, Hoffmann L, Morvan H (1995) Identification by NMR spectroscopy of oligosaccharides obtained by acidolysis of the capsular polysaccharides of a thermal biomass. Int J Biol Macromol **17**: 387–393

Gray JS, Yang BY, Montgomery R (2000) Extracellular polysaccharide of *Erwinia chrysanthemi* A350 and ribotyping of *Erwinia chrysanthemi* spp. Carbohydr Res **324**: 255–267

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 Physiology **136**: 4256-4264

Harper AD, Bar-Peled M (2002) Biosynthesis of UDP-xylose. Cloning and characterization of a novel Arabidopsis gene family, UXS, encoding soluble and putative membrane-bound UDP-glucuronic acid decarboxylase isoforms. Plant Physiol **130**: 2188–2198

Harholt J, Jensen JK, Sorensen SO, Orifila C, Pauly M and Scheller HV (2006) Arabinan deficient 1 is a putative arabinosyltransferases involved in biosynthesis of pectic arabinan in Arabidopsis. Plant Physiol. **140**: 49-58

Holst O, Susskind M, Grimmecke D, Brade L, Brade H (1998) Core structures of enterobacterial lipopolysaccharides. Prog Clin Biol Res **397**: 23–35

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

Iwai H, Masaoka N, Ishii T and Satoh S (2002) A pectin glucoronyltransferase gene is essential for intercellular attachment in the plant meristem. PNAS. **99**: 16319-16324

Iwai H, Hokura A, Oishi M, Chida H, Ishii T, Sakai S and Satoh S (2006) The gene responsible for borate cross-linking of pectin Rhamnogalacturonan-II required for plant reproductive tissue development and fertilization. PNAS. **103**: 16592-16597

Jann K, Westphal O (1975) Microbial polysaccharides. In M Sela, ed, The Antigens, Vol 3. Academic Press, New York, pp 1–125

Jackson C, Dreaden T, Theobald L et al., (2007) Pectin induces apoptosis in human prostate cancer cells: correlation of apoptotic function with pectin structure Glycobiology **17**: 805–819

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal X. Trends Biochem Sci **23**: 403–405

Lemaire HG, Muller-Hill B (1986) Nucleotide sequences of the gal E gene and the gal T gene of E. coli. Nucleic Acids Res **14**: 7705–7711

Jensen JK, Sørensen SO, Harholt J, et al. (2008) Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in Arabidopsis. Plant Cell **20**:1289-1302

Jiang L, Yang SL, Xie LF, Puah CS, Zhang XQ, Yang WC, Sundaresan V and Ye D (2006) *VANGUARD1* Encodes a Pectin Methylesterase That Enhance Pollen Tube Growth in the Arabidopsis Style and Transmitting Tract. Plant Cell. **17**: 584-596

Livak KJ, Schmittgen TD. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Method. 4: 402-408.

Johnson K, Jones B, Schultz C and Bacic A (2003) Non-enzymic cell wall (glycol) proteins. In J. K.C. Rose, ed, The Plant Cell Wall: Annual Plant Review, Vol 8. Blackwell Publishing and CRC Press, Oxford, pp. 111-143

Jones L, Milne JL, Ashford D, McQueen-Mason SJ (2003) Cell wall arabinan is essential for guard cell function. PNAS **100:** 11783-11788

Jones L, Milne JL, Ashford D, McCann MC, McQueen-Mason SJ (2005) A conserved functional role of pectic polymers in stomatal guard cells from a range of plant species. Planta. **221**:255-264

Karkonen A. and Fry S. (2006) Novel characteristics of UDP-glucose dehydrogenase activities in maize: non-involvement of alcohol dehydrogenases in cell wall polysaccharide biosynthesis. Planta 223: 858–870

Kessler G, Neufeld E, Feingold D, Hassid WZ (1961) Metabolism of D-Gglucuronic Acid and D-Galacturonic Acid by *Phaseolus aureus* Seedlings. J Biol Chem **236:** 308-312

Khotimchenko MY, **Kolenchenko EA**. (2006) Efficiency of low-esterified pectin in toxic damage to the liver inflicted by lead treatment. Bull Exp Biol Med. **144**(1):60-62

Konishi T, **Takeda T**, **Miyazaki Y**, **Ohnishi-Kameyama M**, **Hayashi T**, **O'Neill MA**, **Ishii T**. (2006) A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. Glycobiology. 17(3):345-54

Kotake T, **Hojo S**, **Yamaguchi D**, **Aohara T**, **Konishi T**, **Tsumuraya Y**. (2007) Properties and physiological functions of UDP-sugar pyrophosphorylase in Arabidopsis. Biosci Biotechnol Biochem. **71**: 761-771

Lerouxel O, Cavalier MD, Liepman HA and Keegstra K (2006) Biosynthesis of plant cell wall polysaccharides – a complex process. Current Opinion in Plant Biology **9**:621-630

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

Ma B, Cui ML, Sun HJ, Takada K, Mori H, Kamada H, and Ezura H (2006) Subcellular Localization and Membrane Topology of the Melon Ethylene Receptor CmERS1. Plant Physiol **141**: 587–597

Maley F, **Trimble RB**, **Tarentino AL**, **Plummer TH Jr**. (1989) Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. Anal Biochem.**180**:195-204

Miyazawa R, Tomomasa T, Kaneko T, Arakawa N, Shimizu N, Morikawa A (2008) Effects of pectin liquid on gastroesophageal reflux disease in children with cerebral palsy. BMC Gastroenterology **8**:11

Mohnen D (2002) Biosynthesis of pectins. In G.B. Seymour, JP Knox, eds, Pectins and their Manipulation. Blackwell Publishing and CRC Press, Oxford, pp. 52-98

Mohnen M (2008) Pectin structure and biosynthesis. Current Opinion in Plant Biology **11**: 266-277

Mølhøj M, Verma R, Reiter WD. (2003) The biosynthesis of the branched-chain sugar d-apiose in plants: functional cloning and characterization of a UDP-d-apiose/UDP-d-xylose synthase from Arabidopsis. Plant J. 35(6):693-703

Molhoj M, Verma R, Reiter WD (2004) The biosynthesis of D-Galaturonate in plants, functional cloning and characterization of a membrane-anchored UDP-D-Glucuronate 4-epimerase from *Arabidopsis*. Plant Physiol. **135**: 1221-1230

Molinaro A, De Castro C, Lanzetta R, Parrilli M, Raio A, Zoina A (2003) Structural elucidation of a novel core oligosaccharide backbone of the lipopolysaccharide from the new bacterial species Agrobacterium larrymoorei. Carbohydr Res **338**: 2721–2730

Mort A, **Zheng Y**, **Qiu F**, **Nimtz M**, **Bell-Eunice G**. (2008) Structure of xylogalacturonan fragments from watermelon cell-wall pectin. Endopolygalacturonase can accommodate a xylosyl residue on the galacturonic acid just following the hydrolysis site. Carbohydr Res. 343:1212-1221

Mouille G, Ralet M-C, Cavelier C, et al. (2007) Homogalacturonan synthesis in Arabidopsis thaliana requires a Golgi-localized protein with a putative methyltransferase domain. Plant J **50**:605-614.

Munoz R, Lopez R, de Frutos M, Garcia E (1999) First molecular characterization of a uridine diphosphate galacturonate 4-epimerase: an enzyme required for capsular biosynthesis in *Streptococcus pneumoniae* type 1. Mol Microbiol. **31**: 703-13.

Munoz R, Mollerach M, Lopez R, Garcia E (1997) Molecular organization of the genes required for the synthesis of type 1 capsular polysaccharide of Streptococcus

pneumoniae: formation of binary encapsulated pneumococci and identification of cryptic dTDP-rhamnose biosynthesis genes. Mol Microbiol **25**: 79–92

Naran R, Chen G, Carpita N (2008) Novel Rhamnogalacturonan I and Arabinoxylan Polysaccharides of Flax Seed Mucilage. Plant Physiol **142:** 132-148

Nelson B, Cai X and Nebenfuhr A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. The Plant Journal **51**, 1126–1136

Neufeld EF, Feingold DS, Ilves SM, Kessler G, Hassid WZ (1961) Phosphorylation of D-galacturonic aid by extracts from germinating seeds of *Phaseolus aureus*. J Biol Chem **236:** 3102-3105

Nguema-Ona E, Andeme-Onzighi C, Aboughe-Angone S, Bardor M, Ishii T, Lerouge P and Driouich A (2006) The *reb1-1* mutation of Arabidopsis. Effect on the structure and localization of Galactose-containing cell wall polysaccharides. Plant Physiol **140:** 1406-1417

Norman C, Vidal S, Palva ET (1999) Oligogalacturonide-mediated induction of a gene involved in jasmonic acid synthesis in response to the cell-wall-degrading enzymes of the plant pathogen *Erwinia carotovora*. Mol Plant Microbe Interact **12**:640-644.

Nuin PA, Wang Z, Tillier ER (2006) The accuracy of several multiple sequence alignment programs for proteins. BMC Bioinformatics **7:** 471

Ohashi T, Cramer N, Ishimizu T, Hase S (2006) Preparation of UDP-galacturonic acid using UDP-sugar pyrophosphorylase. Analytical Biochemistry **352:** 182-187

O'Neill M, 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 M, 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 M and York W (2003) The composition and structure of plant primary cell walls. In J. K.C. Rose, ed, The Plant Cell Wall: Annual Plant Review, Vol 8. Blackwell Publishing and CRC Press, Oxford, pp. 1-54

Orellana A, Mohnen D (1999) Enzymatic synthesis and purification of [^{3}H] uridine diphosphate galacturonic acid for use in studying Golgi-localized transporters, Anal. Biochem. **272**: 224-231

Orfila C, Sørensen 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(4):613-22.

Panchuk II, Zentgraf U, Volkov RA. (2005) Expression of the Apx gene family during leaf senescence of *Arabidopsis thaliana*. Planta. **222**: 926-932.

Parre E, Geitmann A (2005) Pectin and the role of the physical properties of the cell wall in pollen tube growth of Solanum chacoense. Planta. **220**:582-92.

Pattathil S, Harper AD, Bar-Peled M. (2005) Biosynthesis of UDP-xylose: characterization of membrane-bound At*Uxs2*. Planta **211**: 538-548.

Pellerin P, Williams SVP and Brillouet JM (1995) Characterization of five type II arabinogalactan-protein fractions from red wine of increasing uronic acid content. Carbohydrate Research **277:** 135-143

Pena M, Zhong R, Zhou G-K, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye Z-H (2007) *Arabidopsis irregular xylem8* and *irregular xylem9*: Implications for the Complexity of Glucuronoxylan Biosynthesis. The Plant Cell **19:** 549-563

Persson S, Caffall KH, Freshour G, Hilley MT, Stefan Bauer a, 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. The Plant Cell **19:** 237-255

Plotz BM, Lindner B, Stetter KO, Holst O (2000) Characterization of a novel lipid A containing D-galacturonic acid that replaces phosphate residues. The structure of the lipid a of the lipopolysaccharide from the hyperthermophilic bacterium Aquifex pyrophilus. J Biol Chem **275**: 11222–11228

Raguenes G, Cambon-Bonavita MA, Lohier JF, Boisset C, Guezennec J (2003) A novel, highly viscous polysaccharide excreted by an alteromonas isolated from a deep-sea hydrothermal vent shrimp. Curr Microbiol **46**: 448–452

Reiter WD, Vanzin GF (2001) Molecular genetics of nucleotide sugar interconversion pathways in plants. Plant Mol Biol **47**: 95–113

Renard C, Crepeau MJ and Thibault J (1999) Glucuronic acid directly linked to the galacturonic acid in the rhamnogalacturonan backbone of beet backbone. Eur. J. Biochem. **266**: 566-574

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligalacturonide-related signaling. Phytochemistry **57**: 929-967

Ritzenthaler, C., Nebenfuhr, A., Movafeghi, A., Stussi-Garaud, C., Behnia, L., Pimpl, P., Staehelin, L.A. and Robinson, D.G. (2002) Reevaluation of the effects of brefeldin A on plant cells using tobacco Bright Yellow 2 cells expressing Golgi-targeted green fluorescent protein and COPI antisera. Plant Cell, **14**, 237–261.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:** 1572-1574

Saint-Jore-Dupas C, Nebenführ A, Boulaflous A, Follet-Gueye ML, Plasson C, Hawes C, Driouich A, Faye L and Gomord V (2006) Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. The Plant Cell **18**:3182-3200

Santiago-Doménech N, **Jiménez-Bemúdez S**, **Matas AJ**, **Rose JK**, **Muñoz-Blanco J**, **Mercado JA**, **Quesada MA**. (2008) Antisense inhibition of a pectate lyase gene supports a role for pectin depolymerization in strawberry fruit softening. Journal of Experimental Botany **59**: 2769–2779

Schols HA, **Vierhuis E**, **Bakx EJ**, **Voragen AG**.(1995) Different populations of pectic hairy regions occur in apple cell walls. Carbohydr Res. **275**:343-360

Seifert GJ (2004) Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. Curr Opin Plant Biol. **7**: 277-284

Simpson SD, Ashford DA, Harvey DJ, Bowles DJ (1998) Short chain oligogalacturonides induce ethylene production and expression of the gene encoding aminocyclopropane 1-carboxylic acid oxidase in tomato plants. Glycobiology **8**:579-583.

Stafstrom, J.P., and Staehelin, L.A. (1986). The role of carbohydrate in maintaining extensin in an extended conformation. Plant Physiol. 81: 242–246.

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. PNAS **103**: 5236-5241

Sterling JD, **Quigley HF**, **Orellana A**, **Mohnen D** (2001) The catalytic site of the pectin biosynthetic enzyme alpha-1,4-galacturonosyltransferase is located in the lumen of the Golgi. Plant Physiol **127**:360-371

Stroop CJ, Xu Q, Retzlaff M, Abeygunawardana C, Bush CA (2002) Structural analysis and chemical depolymerization of the capsular polysaccharide of Streptococcus pneumoniae type 1. Carbohydr Res **337**: 335–344

Suzuki K, Suzuki Y, Kitamura S (2003) Cloning and expression of a UDPglucuronic acid decarboxylase gene in rice. J Exp Bot **54**: 1997–1999

Suzuki K, Watanabe K, Masumura T, Kitamura S. (2004) Characterization of soluble and putative membrane-bound UDP-glucuronic acid decarboxylase (OsUXS) isoforms in rice. Arch Biochem Biophys. 431(2):169-77

Swofford DL. (1998) PAUP. Phylogenetic Analysis Using Parsimony and Other Methods. Version 4. Sinauer Associates, Sunderland, MA.

Tenhaken R and Thulke O. (1996) Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-glucose dehydrogenase. Plant Physiol. 112(3):1127-1134.

Thoden JB, Hegeman AD, Wesenberg G, Chapeau MC, Frey PA, Holden HM (1997) Structural analysis of UDP-sugar binding to UDP-galactose 4-epimerase from Escherichia coli. Biochemistry **36**: 6294–6304

Usadel B, Schluter U, Molhoj M, Gipmans M, Verma R, Kossmann J, Reiter WD, Pauly M (2004) Identification and characterization of a UDP-D-glucuronate 4-epimerase in Arabidopsis. FEBS Lett **569**: 327–331

Watt G, Leoff C, Harper AD, Bar-Peled M (2004) A bifunctional 3,5-epimerase/4-keto reductase for nucleotide-rhamnose synthesis in Arabidopsis. Plant Physiol **134**: 1337–1346

Weirenga RK, Terpstra P, Hol WGJ (1986) Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint. J Mol Biol **187**: 101-107

Willats W, McCartney L, Mackie W, Knox JP (2001) Pectin: Cell biology and prospects for functional analysis. Plant Mol. Bio. **47**: 9-27

Yates EA, Valdor JF, Haslam SM, Morris HR, Dell A, Mackie W, Knox JP (1996) Characterization of carbohydrate structural features recognized by anti-arabinogalactan-protein monoclonal antibodies. Glycobiology **6**: 131–139

York W and O'Neill M (2008) Biochemical control of xylan biosynthesis — which end is up? Current Opinion in Plant Biology **11**:258–265

Yuasa K, Toyooka K, Fukuda H and Matsuoka K (2005) Membrane-anchored prolyl hydroxylase with an export signal from the endoplasmic reticulum. The Plant Journal **41:** 81–94

Zandleven J, **Sørensen SO**, **Harholt J**, **Beldman G**, **Schols HA**, **Scheller HV**, **Voragen AJ**. (2007) Xylogalacturonan exists in cell walls from various tissues of Arabidopsis thaliana. Phytochemistry 68:1219-26

Zhang Q, Hrmova M, Shirley NJ, Lahnstein J, Fincher GB. (2006) Gene expression patterns and catalytic properties of UDP-D-glucose 4-epimerases from barley (Hordeum vulgare L.). Biochem J. 394:115-124

Zhang Q, Shirley N, Lahnstein J, Fincher GB. (2005) Characterization and expression patterns of UDP-D-glucuronate decarboxylase genes in barley. Plant Physiol. 138(1):131-141

Zhao ZY, **Liang L**, **Fan X**, **Yu Z**, **Hotchkiss AT**, **Wilk BJ**, **Eliaz I**.(2008) The role of modified citrus pectin as an effective chelator of lead in children hospitalized with toxic lead levels**.** Altern Ther Health Med. 14:34-38

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. *Arabidopsis* Microarray Database and Analysis Toolbox. Plant Physiol. **136**: 2621-2632

APPENDIX A

PHYLOGENETIC ANALYSES OF PLANT UGLCAE ISOFORMS

Phylogenetic analyses indicate the existence of three evolutionary clades among dicot and Poaceae UGlcAE isoforms (Gu and Bar-Peled, 2004; Gu's Ph.D. dissertation Chapter 3). In order to further investigate the phylogenetic relationship among plant UGlcAE homologs, the amino-acid sequences corresponding to the conserved catalytic domain of maize, rice and Arabidopsis UGlcAEs (Gu's PhD dissertation Chapter 3) were used to identify UGlcAE gene homologs from three other flowering plant species (*Populus trichocarpa*, *Sorghum bicolor and Vitis vinifera*), two moss species (*Physcomitrella patens subsp. Patens,* and *Selaginella moellendorffii*), two green algal species (*Ostreococcus lucimarinus* and *Chlamydomonas reinhardtii*) and three bacterial species (*Azoarcus sp. BH72, Dechloromonas aromatic* and *Streptococcus pneumoniae*). The phylogenetic analyses of UGlcAEs were then performed in collaboration with Dr. Ying Xu's lab at the University of Georgia.

A total of fifty-one UGlcAE homologs were collected (Table A.1), all of which show more than 45% sequence identity (> 65% positives) at the protein level compared to AtUGlcAE1. All UGlcAE homologs contain highly conserved catalytic motifs, which include the GxxGxxG (x = any amino acid) motif that is reqiured for NAD^+ binding and a catalytic triad consisting of S/T and YxxxK (Weirenga et al., 1986; Gu and Bar-Peled, 2004). Multiple UGlcAE isoforms were identified in flowering plants: six isoforms in *Vitis vinifera*, five isoforms in *Sorghum bicolor*, six isoforms in Arabidopsis*,* seven isoforms in *Populus trichocarpa*, five isoforms in

rice and at least three isoforms in maize (Table A.1). Based on our membrane topology studies with Arabidopsis UGlcAE homologs, it is likely that the UGlcAE isoforms in flowering plant and the moss species are type II transmembrane proteins, as they share very similar amino acid sequences. The green algal and bacterial UGlcAE isoforms, on the other hand, are likely to be cytosolic proteins, as they a) lack the N-terminal extension found in other plant UGlcAE; b) have no predicted TM domains; c) have the GXXGXXG motif close to the N-terminal region. Interestingly, among the three isoforms identified in *Selaginella*, two isoforms contain a transmembrane domain, while isoform 3 does not and we predict that this isoform resides in the cytoplasm.

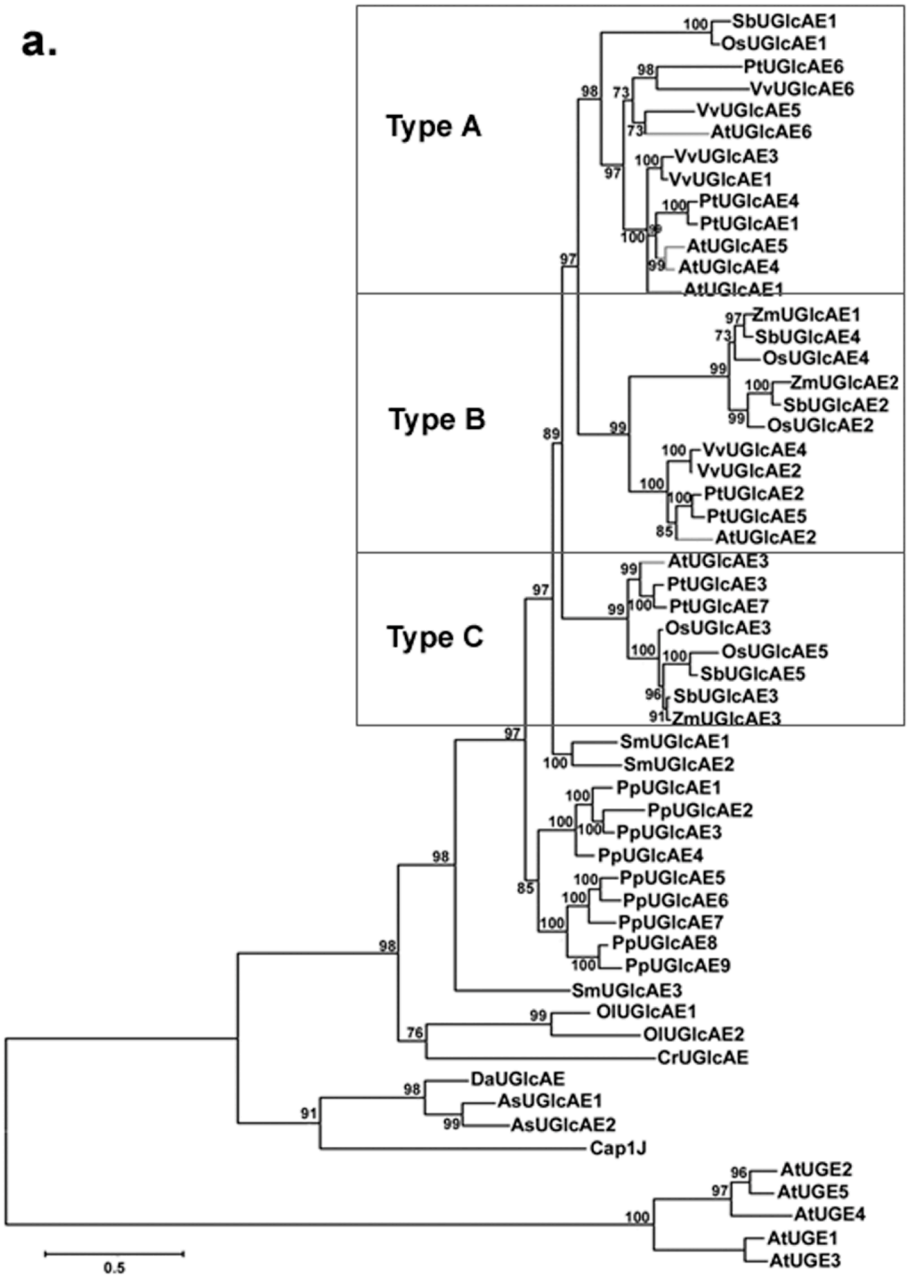
Both the full-length protein sequences and the sequences encompassing only the catalytic domain (also named as "domain sequence") of all the identified UGlcAEs were subjected to the phylogenetic analyses using MrBayes method (Ronquist and Huelsenbeck, 2003). Based on the convension that the phylogenetic relationship among different UGlcAE homologs is believed to be closer than the relationship between UGlcAE homologs and other nucleotide sugar synthases, five Arabidopsis UDP-Glucose 4-epimerase (AtUGE) (Barber et al., 2006) proteins were used as an out group. The two Bayesian phylogenetic trees (Fig A.1 a and b) strongly indicate that there are three major UGlcAE clades (type A, B and C) in flowering plants, with clade C being more ancestral than clades A and B (Fig A.1 a and b). Sequences from *Physcomitrella* and *Selaginella* form additional branches in the tree and the three sequences from green algae represent the ancestral UGlcAE homologs in aquatic plants before the generation of the three major UGlcAE clades (Fig A.1 a and b). The UGlcAE phylogeny also suggests that it is likely the bacterial UGlcAE orthologs have contributed to the origin of plant UGlcAE genes because three bacterial UGlcAE proteins are the most ancestral UGlcAE orthologs in both trees (Fig A.1 a and b).

The current phylogenetic analyses, however, are not sufficient to determine when the diversification of the three major UGlcAE clades took place because the sequences of plant specis rather than the moss and the flowering plants are still missing. Therefore, it is also uncertain how and when the soluble UGlcAE evolved into the membrane bound 4-epimerases. Further identification and subsequent phylogenetic analyses of UGlcAE isoforms from more fully sequenced plant and algal genomes may help us answer these questions.

Table A.1 Identified plant and bacterial UGlcAE isoforms.

Specie	UGlcAE isoforms	Locus or accession number
Arabidopsis thaliana	AtUGlcAE1	At2g45310
	AtUGlcAE2	At3g23820
	AtUGlcAE3	At4g30440
	AtUGlcAE4	At4g00110
	AtUGlcAE5	At1g02000
	AtUGlcAE6	At4g12250
Oryza sativa	OsUGlcAE1	DQ333338
	OsUGlcAE2	DQ333337
	OsUGlcAE3	DQ333336
	OsUGlcAE4	Os09g32670
	OsUGlcAE5	Os06g08810
Zea mays	ZmUGlcAE1	ACG26819
	ZmUGlcAE2	ACG41039
	ZmUGlcAE3	DQ247999
Populus trichocarpa	PtUGlcAE1	grail3.0035002801
	PtUGlcAE2	ABK93061
	PtUGlcAE3	eugene3.00180906
	PtUGlcAE4	grail3.0033029301
	PtUGlcAE5	eugene3.00880019
	PtUGlcAE6	grail3.0079014001
	PtUGlcAE7	eugene3.00061339
Sorghum bicolor	SbUGlcAE1	Sb01g041030
	SbUGlcAE2	Sb07g026520
	SbUGlcAE3	Sb04g035630
	SbUGlcAE4	Sb02g029130
	SbUGlcAE5	Sb10g005920
Vitis vinifera	VvUGlcAE1	CAN73016

	VvUGlcAE2	CAO43514
	VvUGlcAE3	CAO24537
	VvUGlcAE4	CAN83418
	VvUGlcAE5	CAN60968
	VvUGlcAE6	CAO49712
Selaginella moellendorffii	SmUGlcAE1	Selmo1/scaffold_1:5084261-5085613
	SmUGlcAE2	Selmo1/scaffold_6:822306-823676
	SmUGlcAE3	Selmo1/scaffold_10:1110336-1111949
Physcomitrella patens subsp. patens	PpUGlcAE1	EDQ73088
	PpUGlcAE2	EDQ53936
	PpUGlcAE3	EDQ67973
	PpUGlcAE4	EDQ65739
	PpUGlcAE5	EDQ81323
	PpUGlcAE6	EDQ83426
	PpUGlcAE7	EDQ75605
	PpUGlcAE8	EDQ74153
	PpUGlcAE9	EDQ57868
Ostreococcus lucimarinus	OlUGlcAE1	ABO94794
	OlUGlcAE2	ABP00628
Chlamydomonas reinhardtii	CrUGlcAE	EDO99605
Azoarcus sp. BH72	AsUGlcAE1	YP_932356
	AsUGlcAE2	YP_157918
Dechloromonas aromatic	DaUGlcAE	AAZ48658
Streptococcus pneumoniae	Cap1J	Z83335


a.
Type A
Type B
Type C
SbUGlcAE1
OsUGlcAE1
PtUGlcAE6
VvUGlcAE6
VvUGlcAE5
AtUGlcAE6
VvUGlcAE3
VvUGlcAE1
PtUGlcAE4
PtUGlcAE1
AtUGlcAE5
AtUGlcAE4
AtUGlcAE1
ZmUGlcAE1
SbUGlcAE4
OsUGlcAE4
ZmUGlcAE2
SbUGlcAE2
OsUGlcAE2
VvUGlcAE4
VvUGlcAE2
PtUGlcAE2
PtUGlcAE5
AtUGlcAE2
AtUGlcAE3
PtUGlcAE3
PtUGlcAE7
OsUGlcAE3
OsUGlcAE5
SbUGlcAE5
SbUGlcAE3
ZmUGlcAE3
SmUGlcAE1
SmUGlcAE2
PpUGlcAE1
PpUGlcAE2
PpUGlcAE3
PpUGlcAE4
PpUGlcAE5
PpUGlcAE6
PpUGlcAE7
PpUGlcAE8
PpUGlcAE9
SmUGlcAE3
OlUGlcAE1
OlUGlcAE2
CrUGlcAE
DaUGlcAE
AsUGlcAE1
AsUGlcAE2
Cap1J
AtUGE2
AtUGE5
AtUGE4
AtUGE1
AtUGE3
0.5

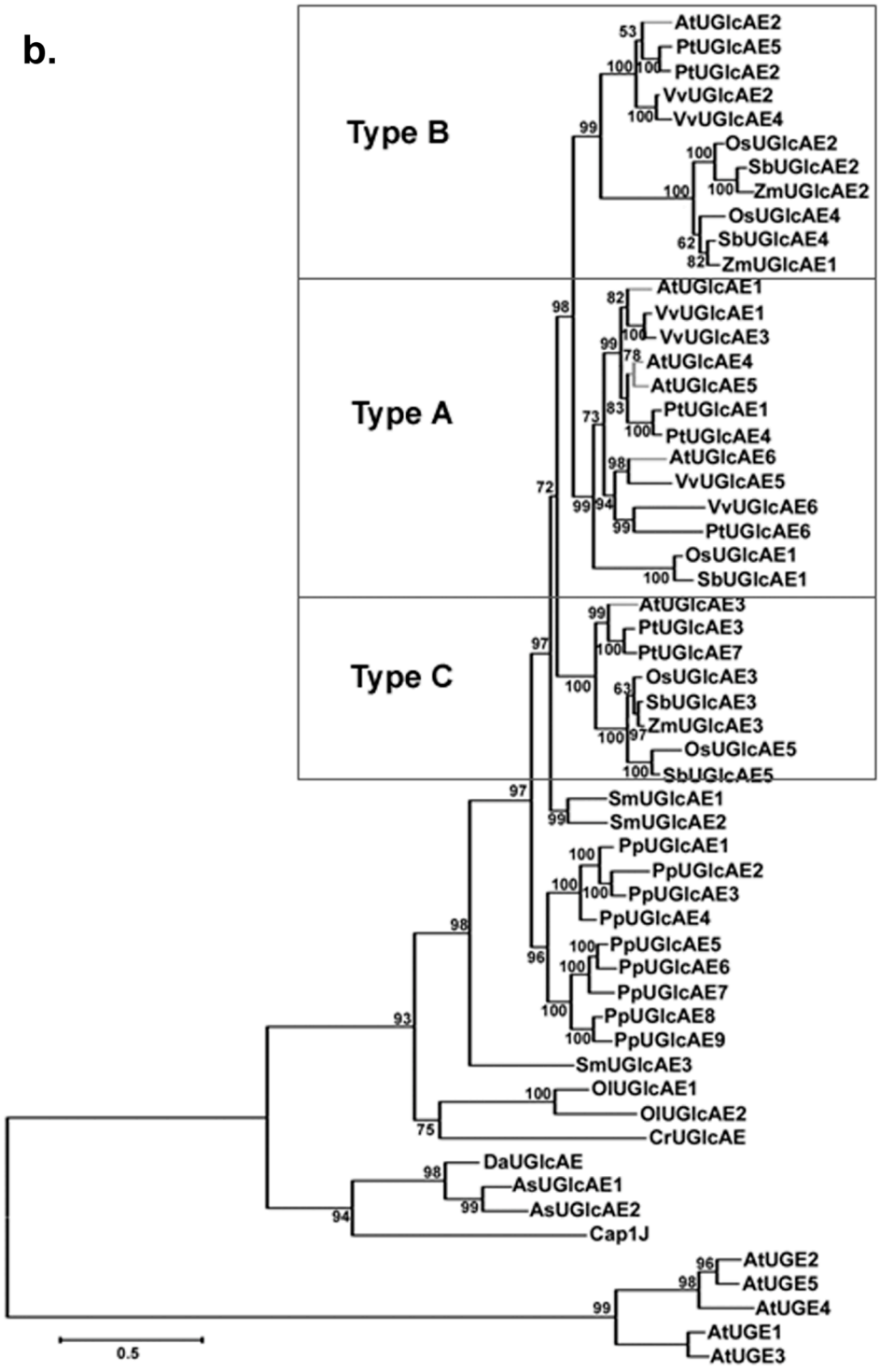

b.
Type B
Type A
Type C
AtUGlcAE2
PtUGlcAE5
PtUGlcAE2
VvUGlcAE2
VvUGlcAE4
OsUGlcAE2
SbUGlcAE2
ZmUGlcAE2
OsUGlcAE4
SbUGlcAE4
ZmUGlcAE1
AtUGlcAE1
VvUGlcAE1
VvUGlcAE3
AtUGlcAE4
AtUGlcAE5
PtUGlcAE1
PtUGlcAE4
AtUGlcAE6
VvUGlcAE5
VvUGlcAE6
PtUGlcAE6
OsUGlcAE1
SbUGlcAE1
AtUGlcAE3
PtUGlcAE3
PtUGlcAE7
OsUGlcAE3
SbUGlcAE3
ZmUGlcAE3
OsUGlcAE5
SbUGlcAE5
SmUGlcAE1
SmUGlcAE2
PpUGlcAE1
PpUGlcAE2
PpUGlcAE3
PpUGlcAE4
PpUGlcAE5
PpUGlcAE6
PpUGlcAE7
PpUGlcAE8
PpUGlcAE9
SmUGlcAE3
OlUGlcAE1
OlUGlcAE2
CrUGlcAE
DaUGlcAE
AsUGlcAE1
AsUGlcAE2
Cap1J
AtUGE2
AtUGE5
AtUGE4
AtUGE1
AtUGE3
0.5

Figure A.1 Phylogenetic analyses of UGlcAE

Fifty-one UGlcAE proteins (see Table A.1) including both the membrane-bound and the cytosolic homologs, along with five Arabidopsis UGE proteins as out-group were selected for phylogenetic analyses. The full-length protein sequences (Panel a), or the domain sequences (Panel b) of those UGlcAE homologs were first aligned with the multiple sequence alignment method, L-INS-I implemented by the MAFFT program (Katoh et al., 2005) and their phylogeny was subsequently analyzed by MrBayes methods, respectively. The Bayesian phylogenetic analyses were conducted with mixed amino acid model estimated in the run, estimated proportion of invariable sites, estimated gamma distribution parameter and 1,000,000 generations. Branch lengths are shown to scale.

APPENDIX B

TRANSCRIPTION ANALYSES OF ARABIDOPSIS UGLCAE ISOFORMS

Within a gene family, different isoforms might evolve to be expressed in a tissue-specific manner (Chen et al., 2005). To determine whether distinct type of UGlcAE isoforms may have a tissue-specific expression pattern, the transcription data of different Arabidopsis UGlcAE homologs in different plant tissues were obtained from Genevestigator (https://www.genevestigator.ethz.ch) and their transcriptions were subsequently analyzed.

Genevestigator is a public microarray database allowing researchers to study the expression and regulation of genes in a broad variety of contexts (Zimmermann et al., 2004). In Genevestigator, raw experimental data from a large diversity of experiments covering different tissues, ages, and treatments, were initially processed using Affymetrix MAS 5.0 software to a target value (TGT) of 1,000; and the transcription levels of genes in each hybridized Affymetrix GeneChip array were converted and quantified into corresponding signal intensity values for easy interpreataion (Zimmermann et al., 2004).

Based on the data from Genevestigator website (Fig B.1), it is clear that 1) all three types of UGlcAE isofroms are expressed in any given tissue; 2) the relative expression level of particular isoforms may be different in specific tissues. For example, AtUGlcAE2 (type B) appears to be expressed in all tissues at similar levels, while AtUGlcAE1 (type A) and AtUGlcAE3 (type C) have a higher expression in pollen (Fig A.2). It is also clear that each plant tissue contains multiple types of AtUGlcAEs that are expressed at very similar levels. For

example, both type B AtGlcAE2 and type C AtUGlcAE3 are equivalently expressed in plant tissues including lateral root, hypocotyl, petiole, rosette leaves, cauline leave, shoot apex, stem, silique, sepal, petal, flower, inflorescence, redicle and cotyledon. In pollen and stamen, type C AtUGlcAE3 and type A AtUGlcAE1 are expressed at very similar levels. Therefore, the expression of different types of AtUGlcAEs is not tissue specific.

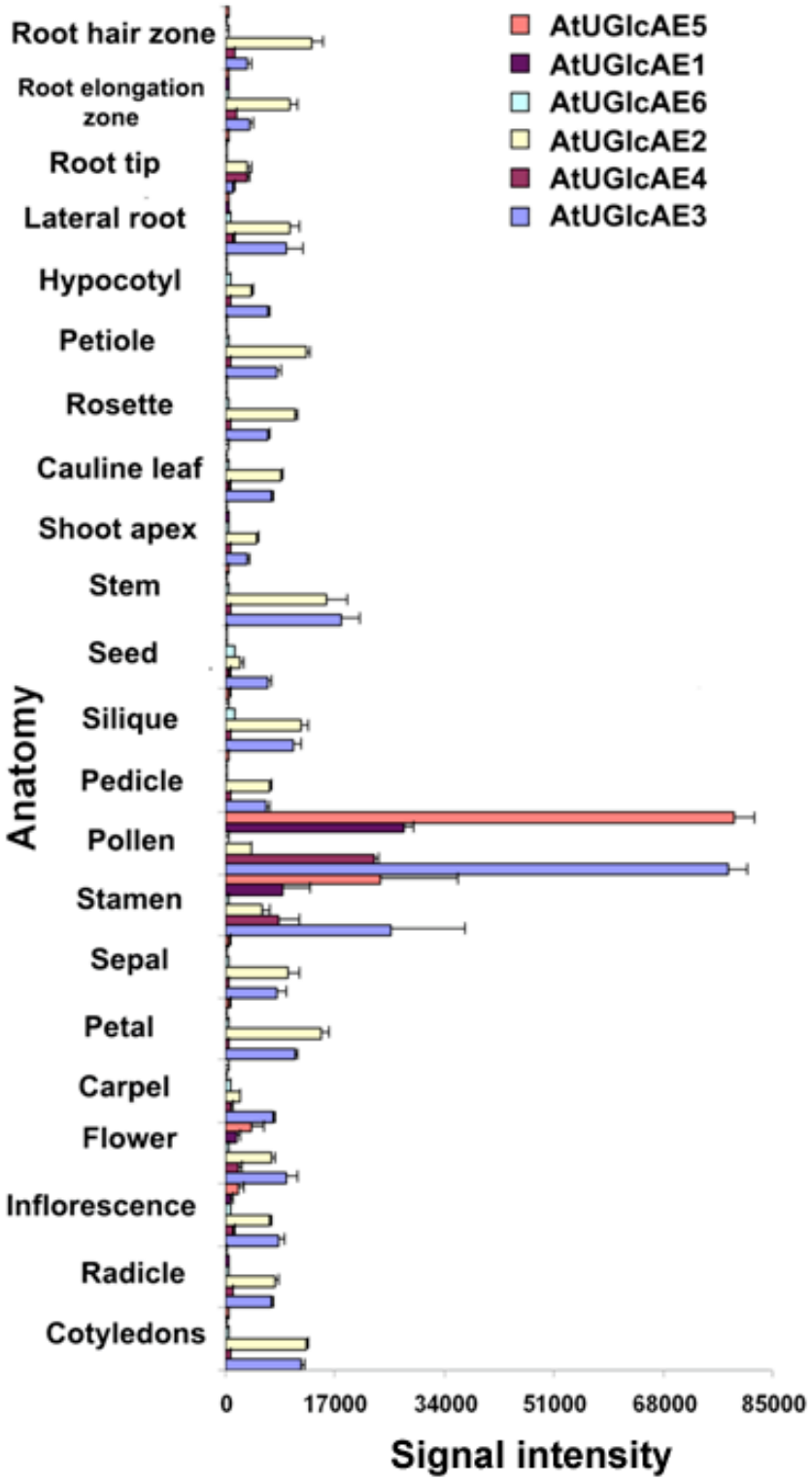

AtUGlcAE5
AtUGlcAE1
AtUGlcAE6
AtUGlcAE2
AtUGlcAE4
AtUGlcAE3
Root hair zone
Root elongation zone
Root tip
Lateral root
Hypocotyl
Petiole
Rosette
Cauline leaf
Shoot apex
Stem
Seed
Silique
Pedicle
Pollen
Stamen
Sepal
Petal
Carpel
Flower
Inflorescence
Radicle
Cotyledons
Anatomy
0
17000
34000
51000
68000
85000
Signal intensity

Figure B.1. The expression of AtUGlcAE isoforms in different Arabidopsis tissues.

The signal intensity value for each isoform in specific tissues was obtained from Genevestigator database by selecting the array type of ATH1 22K and chosing the option of high quality arrays only (total 4076 arrays were picked up) as recommended (Grenan, 2006). Higher intensity value indicates higher transcription level of the gene in that particular tissue (Zimmermann et al., 2004).

APPENDIX C

EXPERIMENTS TO TEST THE BINDING BETWEEN NAD^+ AND UGLCAE

The conversion between UDP-GlcA and UDP-GalA is believed to proceed via a transient 4-keto intermediate, which is generated by a redox reaction involving an enzyme-bound pyridine nucleotide cofactor (i.e. NAD^+ or $NADP^+$) (Feingold and Avigad, 1980). Recombinant AtUGlcAE2 and AtUGlcAE3 expressed in *P. pastoris* do not require exogenous NAD^+ for their UGlcAE activity (Molhoj et al., 2004; Usadel et al., 2004); and exogenous NAD^+ is also not required for activities of AtUGlcAE1, AtUGlcAE2, AtUGlcAE3, ZmUGlcAE3, OsUGlcAE3 and Cap1J when they are heterologously expressed in *E.coli* (Munoz et al., 1999; Gu and Bar-Peled, 2004; Gu's PhD dissertation Chapter 3). Munoz et al., (1999) suggested that Cap1J may contain an enzyme-linked NAD^+ molecule(s), since the fluorescence spectrum of the purified protein resembled that of NADH, with an excitation maximum at about 350nm and an emission maximum close to 435 nm. So far, no unambiguous evidence is available to confirm the binding between NAD^+ and the 4-epimerases and several treatments were tried to separate the co-factor from the protein: 1) the partially purified recombinant protein was denatured by boiling at 100^{o}C for 10 min; 2) the partially purified recombinant protein was denatured with chloroform; 3) the partially purified protein was precipitated with ethanol at a final concentration of 80% (v/v). After the treatments, the samples were centrifuged and the soluble fractions were subjected to HPLC analysis to test the presence of free NAD^+. However, in the HPLC assays, the existence of too many unidentifiable co-elutants made it impossible to confirm the presence of bound NAD^+

or NADH. Therefore, in the future, a new method must be developed to increase the purity of the enzyme so that the level of contaminants will be reduced, thus allowing elucidation of whether the NAD^+ is bound to the enzyme. We are also considering using ^{13}C-DNP-NMR for future work to evaluate these processes.