THE ROLE OF UDP-APIOSE IN BIOSYNTHESIS OF PLANT WALL POLYSACCHARIDES AND BACTERIAL GLYCANS

by

JAMES AMOR SMITH

(Under the Direction of Maor Bar-Peled)

ABSTRACT

Apiose is a unique, branched-chain monosaccharide. It is a conserved initial residue of side chains A and B in the vascular plant cell wall polysaccharide rhamnogalacturonan II, an abundant side chain in the seagrass and duckweed-specific pectic polysaccharide apiogalacturonan, and a glycoside decorating numerous plant secondary metabolites. Detection of aqueous methanolic-soluble apiosides in mosses, liverworts, hornworts, and green algae revealed that apiosides are not exclusive to vascular plants. Recently, our research uncovered apiosyl-residues in extracts from the marine bacteria *Geminicoccus roseus* and plant pathogen *Xanthomonas pisi*. These apiose-containing glycans are synthesized using UDP-apiose as the donor. UDP-apiose together with UDP-xylose is formed from UDP-glucuronic acid by the bifunctional UDP-apiose/UDP-xylose synthase (UAS). UASs from vascular plants, avascular plants, and bacteria (bUAS) were identified and functionally characterized to distinguish them from other short-chain dehydrogenase/reductases. The initial goals of this research were to 1) determine if UDP-apiose synthesis regulates apioside deposition in plant cell walls and 2)

identify and characterize an apiosyltransferase (ApiT) capable of utilizing UDP-apiose to synthesize wall apiosides. To address the first aim, we over-expressed duckweed UAS in the moss *Physcomitrella patens*, which led to an increase in aqueous methanolic-soluble apiose but did not result in discernible amounts of cell wall-associated apiose. Thus, bryophytes and algae likely lack the ApiT required to synthesize apiose-containing cell wall glycans but can synthesize secondary metabolite apiosides. Over-expression of duckweed UAS in *Arabidopsis* also led to increased methanolic-soluble apiose but not a significant increase in pectic apiose, suggesting the limiting factor for plant wall apiose is the glycosyltransferase machinery. The second aim was addressed when small amounts of apiose were detected in the cell pellet of *Xanthomonas pisi*. Examination of the bacterial genome revealed putative GTs (termed XpGTA and XpGTB) directly downstream of the bUAS ORF. Heterologous co-expression of UAS and XpGTB in *E. coli* demonstrated that these cells utilize UDP-apiose to generate an apioside detectable in the cell pellet fraction. Enzyme purification and assay demonstrated XpGTB is an apiosyltransferase that specifically uses UDP-apiose.

INDEX WORDS: UDP-apiose, UDP-xylose, UAS, SDR, GT, apiosyltransferase, pectin, RG-II, LPS, *Xanthomonas pisi*, XpGTA, XpGTB

THE ROLE OF UDP-APIOSE IN BIOSYNTHESIS OF PLANT WALL POLYSACCHARIDES AND BACTERIAL GLYCANS

by

JAMES AMOR SMITH

BS, The University of Georgia, Athens GA (2009)

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

© 2018

James Amor Smith

All Rights Reserved

THE ROLE OF UDP-APIOSE IN BIOSYNTHESIS OF PLANT WALL POLYSACCHARIDES AND BACTERIAL GLYCANS

by

JAMES AMOR SMITH

Major Professor:	Maor Bar-Peled
Committee:	William S. York
	Michael G.Hahn
	Michael Tiemeyer

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

DEDICATION

This work is dedicated to my wife Kelly and my two children Maximus and Athena for being continual source of support and inspiration.

ACKNOWLEDGEMENTS

Over the past six years, I have been fortunate to receive abundant support and encouragement from a multitude of colleagues and mentors. I owe these people a debt of gratitude as this work would not be possible without their influence. Primarily, I thank Dr. Maor Bar-Peled for providing continual mentorship and persistent guidance and for providing me the opportunities to conduct and publish this research. Maor imparted on me his extensive experience, leadership skills, and passion for science, which has benefited me throughout my PhD program.

I am also appreciative of all the guidance provided by my committee members: Drs. William S. York, Michael G. Hahn, and Michael Tiemeyer. Each of these prolific scientists provided unique insights to advance my project. Their thoughtful evaluation and encouragement has been invaluable to my progress through this program. I specifically would also like to recognize Malcolm O'Neill for his continual advice and supervision in developing analytical techniques and editing my proposals and papers. Special thanks are owed to Malcolm for organizing and submitting grant renewals, which provided funding for this work.

It has been my distinct honor to have worked and collaborated with many amazing and gifted scientists at the CCRC and other institutions, and I would like to thank each for their help: Dr. Maria Pena at the CCRC and Dr. Jonathan Willis at UT (currently at USDA/ARS/WRRC, Albany, CA) for their collaboration on our BESC switchgrass characterization and co-authored publication; Dr. Yiwen Yang in Dr. Hahn's lab at the CCRC (currently Department of Biostatistics, Houston, TX) for growing and transforming *Physcomitrella* for our publication; Dr.

Stefan Eberhard for providing galacturonides and greenhouse maintenance support; Dr. John Glushka for his assistance in NMR method development and analysis support; Dr. Parastoo Azadi and the analytical service lab staff for allowing me use of their equipment; Qingsong Cai for his aid in equipment maintenance and repair. Special thanks to a generous and talented scientist, Dr. Breeana Urbanowicz for all her training and advice and for all the precious reagents and protocols lent to me over the years.

Additionally, I would like to thank past and current members of the Bar-Peled lab at the CCRC: current graduate student Yang Su; past graduate students and post docs Drs. Eric Yiang, Ting Yang, Gu Xu, Euna Cho, Thiya Mukherjee, Dusty Brown, Soyoun Hwang, Zi Li; and all of the undergraduates and technicians, specifically Viviana Martinez, Omar Salas, Shahar Levy, and Toyin Adelusi for all of their help keeping the lab relatively organized and for their audience and recommendations for presentations and publications.

Finally, I want to show my appreciation to all the administrative staff in Life Science and at the CCRC; Jackie Taylor, Angie Stockton, Sheilah Dixon-Huckabee, Saeid Roshanzamir, Carol Connelly, Lynn Berryman, Karen Howard, for all their support with ordering and logistics of navigating through this program.

TABLE OF CONTENTS

Page

APPENDICES

INTRODUCTION

The distinct aim of the research and this dissertation was to identify apiosyltransferases involved in plant cell wall biosynthesis and extend our understanding of how synthesis of UDP-apiose influences rhamnogalacturonan-II (RG-II) deposition in plant secondary cell walls.

Chapter 1 begins with a description of the plant cell wall and primary wall components, specifically pectins and a description of RG-II. A summary of current knowledge of plant pectin glycosyltransferases (GTs) and a description of gram-negative bacterial wall GTs is provided. The known instances of apiose in nature are described as well as the prior work done to uncover the activated sugar donor UDP-apiose. This chapter then reviews short-chain dehydrogenase/reductase (SDR) enzymes and the mechanism of UDP-glucuronic acid decarboxylation. Chapter 1 concludes with an overview of gram-negative bacteria wall glycans and GTs involved in their synthesis.

Chapter 2 describes the identification and characterization of UAS enzymes in avascular plants and the presence and location of the sugar apiose in these ancestral plants. UAS-like homologous genes from mosses, hornworts, liverworts, and green algae were cloned for heterologous expression and their gene products revealed to be true UDP-apiose/UDP-xylose synthases. Further, apiose was detected in aqueous-methanolic fractions of avascular plant cultures and not in the cell wall fractions. An attempt to force apiose into the wall fraction of the moss *Physcomitrella patens* by overexpressing UAS demonstrated increased methanolic apiose and no wall apiose.

Chapter 3 further describes the presence of apiose in bacteria, and identification and biochemical characterization of the bUAS enzymes responsible for its synthesis. Apiose was found in methanolic fractions of cultured *Geminicoccus roseus* and *Xanthomonas pisi*. *In vitro* assay of heterologously-expressed, purified UAS-like bacterial homologs showed they were truly able to synthesize UDP-apiose and had characteristics of known UAS enzymes. The bUAS transcripts as well as UDP-apiose were detected in cultures of the bacteria *G. roseus* and *X. pisi*.

Chapter 4 continues with the discovery of an apiosyltransferase (ApiT) in the bacteria *Xanthomonas pisi* and partial characterization of the enzyme. Apiosides were detected in the non-phenolic, non-aqueous pellet fractions of *E. coli* engineered to overexpress this *X. pisi* ApiT (called XpGTB).

Finally, Chapter 5 summarizes conclusions made from the work completed in the preceding chapters and suggests future research directions.

CHAPTER 1

LITERATURE REVIEW

INTRODUCTION

Plant primary cell walls; pectins and glycosyltransferases

Plant cell walls accumulate more carbon than any other structures on the planet (Duchesne, 1989; Field et al., 1998; Poorter, 1997). This naturally abundant and renewing resource has established itself as a foundation for human construction as well as a potential source of bioenergy. Because the polysaccharide polymers that make up plant cell walls are innately recalcitrant (resistant to degradation), their use as feedstocks in bioethanol production requires stringent and costly pretreatments (Himmel et al., 2007). Study of these wall polymers' composition and biosynthesis are necessary to understand and harness their usefulness.

From single cellular algae to more complex conifers, plant cell walls consist of a complex network of polysaccharides and proteins (**Figure 1.1**). As a plant cell grows and divides, it secretes a flexible primary wall first, allowing the cell to change shape and expand. The polymers comprising the primary wall include cellulose, hemicelluloses, and pectins (Popper & Fry, 2003). The secondary cell wall is then secreted by the cell and maintains the cells rigidity and shape. This wall is also made of similar polysaccharide polymers, but is much more enriched in cellulose and hemicellulose (Popper & Fry, 2008). In addition, secondary walls of woody plants are low in water content and rich in aromatic polymers called lignin (Boudet, 2000).

FIGURE 1.1

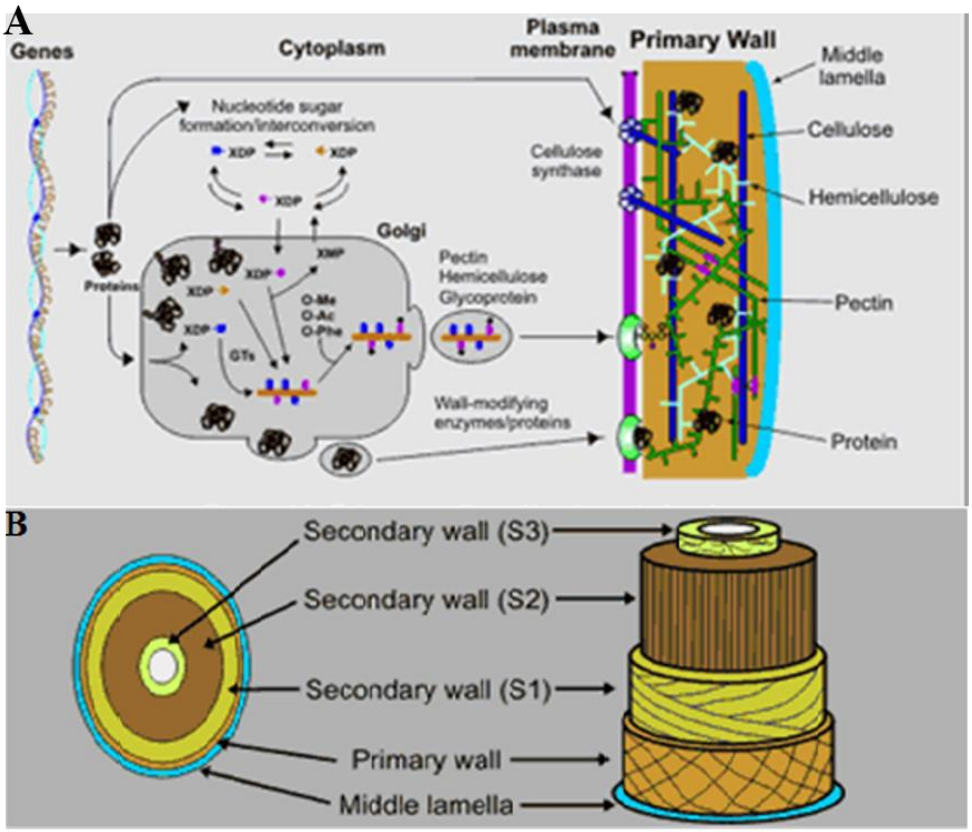


FIGURE 1.1. Schema of primary plant wall structure and biosynthesis. Adapted from Mohnen et al. (2008) (https://www.ccrc.uga.edu/~mao/intro/ouline.htm). (**A**) Primary wall with cellulose and hemicellulose fibers in a pectin matrix. Cellulose synthases secrete fibrils at plasma membrane, while hemicelluloses and pectin are synthesized in the Golgi and transported to the wall. (**B**) Simplified cross-section of a mature plant cell with cellulose-rich, lignified secondary layers S1-3.

Described as a network of cellulose fibers in a bed/gel of pectin, the wall is dynamic. It is believed the hydration of pectin forms a mutable gel in the middle lamella that allow cellulose microfibrils to adjust during growth and prevents rearrangement once growth halts (Cosgrove,

2005). In addition, this "gel" facilitates transport of intercellular signaling, and fragments of pectic domains have been found to act as bioactive signaling molecules (Bishop et al., 1984; Hahn et al., 1981).

Plant cell walls require various enzymatic activities for synthesis, modification, and reconstruction. Considered most important and intriguing for synthesis, glycosyltransferase activity is addition of glycosyl residues to polysaccharides (Albersheim, 2011). Currently, only a handful of plant cell wall glycosyltransferases (GTs) have been definitively characterized. For the pectins, these include; GAUT1 (Sterling et al., 2006) complex, RGXT1-3 (Egelund et al., 2008; Egelund et al., 2006), XGD1 (Jensen et al., 2008), and GALS1 (Liwanag et al., 2012). Unfortunately, most of the GTs involved in plant cell wall synthesis are believed to be Golgi-localized, Type II membrane-bound proteins (Oikawa et al., 2013) and difficult to express as soluble, active proteins. Additionally, these proteins may be difficult to express heterologously, since they require eukaryotic systems for post-translational modification and proper folding and may be active in protein complexes (Atmodjo et al., 2011; Harholt et al., 2012; Mutwil et al., 2008; Oikawa et al., 2013; Zeng et al., 2010). For these reasons several heterologous expression systems have been explored, including *Pichia pastoris* (Cocuron et al., 2007; Edwards et al., 1999; Egelund et al., 2008; Faik et al., 2002; X. L. Liu et al., 2011; Madson et al., 2003; Maris et al., 2009), insect cells (Egelund et al., 2006; Liepman et al., 2005), immortalized monkey kidney (COS) cells (Perrin et al., 1999), and *Nicotiana benthamiana* (Jensen et al., 2008).

Advancement of heterologous expression and characterization of plant cell wall GTs has come from development of a human embryonic kidney (HEK) cells system (Sterling et al., 2006; Urbanowicz et al., 2014; Vanzin et al., 2002). Beyond the difficulties of expression of soluble, active protein, *in vitro* activity assays require sugar donors and acceptors, which may be difficult

to acquire or may be completely unknown. Development of Promega's UDP-Glo™ Glycosyltransferase Assay (http://www.promega.com) has also progressed the field of plant wall GT research (Urbanowicz et al., 2014). Further, ability to synthesize rare or commercially unavailable nucleotide sugar donors will provided the tools to push this area of research forward (Bar-Peled et al., 2012).

Gram-negative bacteria cell walls

Bacterial cell surfaces are decorated with many different types of glycan structures (**Figure 1.2**). Gram-negative bacteria employ capsular polysaccharides (CPS), peptidoglycan, lipopolysaccharide (LPS), and some *N*- or *O*- linked glycoproteins (Logan, 2006; Messner, 2004) to interpret and respond to their environment. These cell-surface glycans are the first interactors, and are crucial for many cell processes such as motility, antibiotic resistance, and pathogenesis (Kay et al., 2010). Disturbance of bacterial wall glycans has grave consequences for the microbe; for example, improper peptidoglycan synthesis causes osmotic lysis in *Staphylococcus* (Chatterjee & Young, 1972), altered LPS structure makes *E. coli* and *Salmonella* more susceptible to antibiotics (Sanderson et al., 1974; Tamaki et al., 1971; Vaara, 1993), and defects in flagellin glycosylation attenuates *Helicobacter* colonization (Schirm et al., 2003).

FIGURE 1.2.

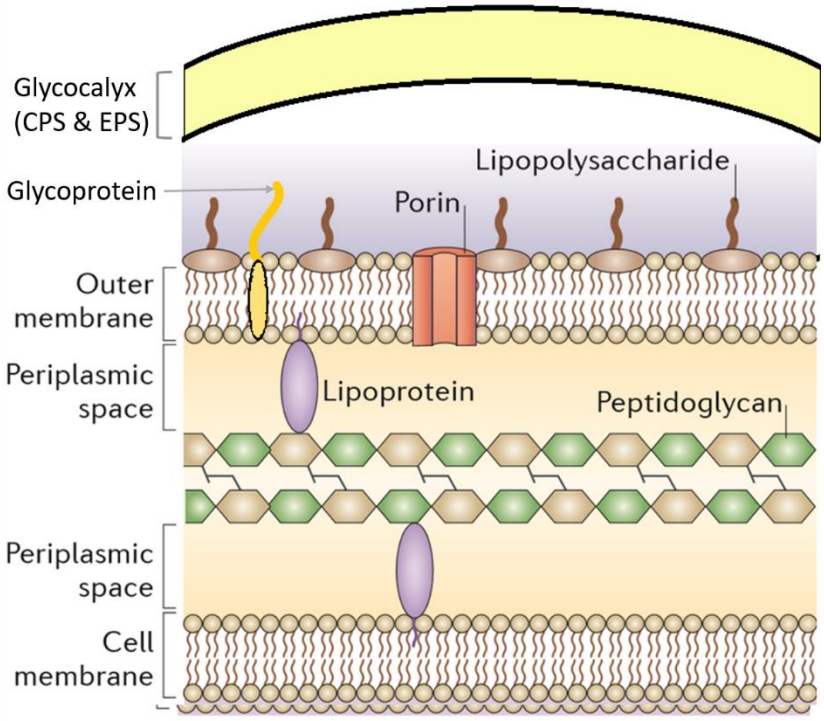


FIGURE 1.2. Adapted from Brown et al. (2015). Diagram of a gram-negative bacterial cell wall, containing a layer of peptidoglycan in the periplasmic space and LPS and a glycoprotein anchored at the cell surface. A layer of glycocalyx comprised of EPS and CPS is outermost form the cell.

Comparable to plants, synthesis of bacterial extracellular glycans utilizes mainly nucleotide sugar donors (Ardevol & Rovira, 2015), acceptor substrates, and GTs. Some GTs necessary for bacterial glycan synthesis are known, including several for the major components of gram-negative bacteria; peptidoglycan and LPS. Unlike plants, GTs associated with bacterial wall synthesis are much easier to study from the gene to the enzyme. Firstly, bacterial GTs are non-intronic and may reside in the genome in operons close to other genes required for synthesis

of the glycan structure, for example *E. coli* and *Salmonella* inner and outer core LPS GTs cluster in their respective genomes into three operons (Heinrichs et al., 1998). These GTs are typically not post-translationally modified (PTM) like plant GTs and do not localize to intracellular compartments. While some *N*- and *O*- protein glycosylation is described in certain bacteria (Szymanski et al., 1999; P. Thibault et al., 2001), few PTMs of bacterial GTs by phosphorylation have been described (Elsholz et al., 2014; Minic et al., 2007). Also, bacterial-derived genes are usually more easily heterologously expressed in *E. coli* than plant GTs (Rosano & Ceccarelli, 2014; Sahdev et al., 2008).

The CAZy database (http://www.cazy.org) has 102 GT families classified based on amino acid sequence and function (Campbell et al., 1997). Of these families, 79 have predicted bacterial members. Classification of these GTs into subfamilies is done by comparison of fold and mechanism (Coutinho et al., 2003; Osmani et al., 2009). Because eukaryotic GTs have very low sequence identity to bacterial GTs at both the nucleotide and amino acid levels, comparisons must be made at the structural level (Gloster, 2014; Jarrell et al., 2014; Liu & Mushegian, 2003). Even similarly classified GTs may have different activities, so, ultimate functional characterization of bacterial GTs requires an assay.

PECTINS

Structure and function

The pectins are plant cell wall polysaccharides rich in galacturonic acid subunits, and are divided into different domains based on composition and interaction partners. The domains include 1) the more abundant homogalacturonan (HGA, **Figure 1.3A**); 2) rhamnogalacturonan-I (RG-I, **Figure 1.3B**); 3) xylogalacturonan (XGA, (**Figure 1.3C**); 4) apiogalacturonan (ApiGalA,

(**Figure 1.3D**); and 5) rhamnogalacturonan-II (RG-II), a highly complex polysaccharide consisting of an HGA backbone with four conserved side chains A-D (**Figure 1.3E**) (Ridley et al., 2001).

FIGURE 1.3

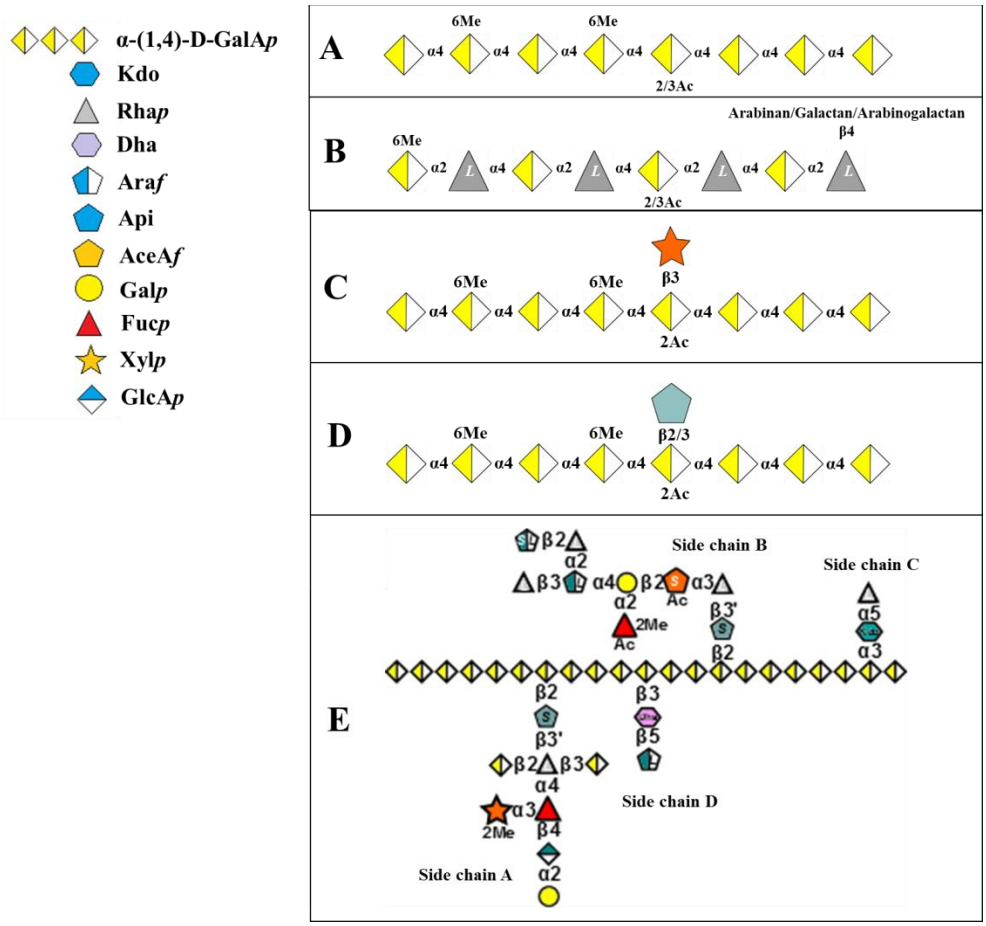


FIGURE 1.3. Structure of the pectic domains. (**A**) HGA is α-(1,4)-Gal*p*A methylated at *O*-6 carboxyl groups and acetylated at *O*-2/3. (**B**) RG-I is -[,4)-α-Gal*p*A-(1,2)- α-L-Rha*p*-(1,]- highly substituted with arabinan, galactan, or arabinogalactan branches at *O*-4 of Rha*p* residues. (**C**) XGA is an HGA backbone with Xyl*p* (or xylobiose) substitutions at *O*-3. (**D**) ApiGalA is an

HGA backbone with Api (or apibiose) substitutions at *O*-2/3. (**E**) RG-II is an HGA backbone with 4 intricate side chains A-D.

HGA is a repeating polymer of α-(1,4)-linked D-galacturonic acid (Gal*p*A) residues, which may be methylesterified at the carboxyl and *O*-2/3 acetylated depending on plant species and tissue type. HGA comprises up to 70% of total pectin (Mohnen, 2008) and has a mean degree of polymerization (DP) of 70-100 residues (J. F. Thibault et al., 1993; Yapo et al., 2007). In some species shorter and longer HGA domains have been detected, for example 4-10 Gal*p*A residues in soybean extract (Nakamura et al., 2002) and >300 Gal*p*A residues in tomato (Round et al., 2010). HGA has a dynamic roll in plant growth. Unmodified HGA is a highly charged chain that can complex metals, such as Ca^{2+} and Pb^{2+} to form more tightly packed gels and reinforce rigidity of the wall during different developmental stages and tissues (Willats et al., 2001). Conversely, HGA can be up to 80% methylesterified (Mohnen, 2008), which reduces this packing and increases the fluidity of this pectic domain in rapidly growing pollen tubes (Dardelle et al., 2010; Lehner et al., 2010).

RG-I is an L-rhamnose-containing (Rha*p*) repeating disaccharide unit of –[,4)-α-Gal*p*A-(1,2)- α-L-Rha*p*-(1,]- and makes up 20-35% of the pectins (O'Neill et al., 1990). RG-I may also be differentially modified with methyl, acetyl, and arabinan, galactan, and/or arabinogalactan side chains based on species and tissue type (Guillemin et al., 2005; Ridley et al., 2001; Willats et al., 2001). In RG-I, 20-80% of Rha*p* residues are *O*-4-substituted with side chains of 1) arabinan, a chain of α-(1,5)-L-arabinofuranose (Ara*f*) with *O*-2/3-linked Ara*f* branches; 2) galactan, a chain of β-(1,4)-D-galactopyranose (Gal*p*) with *O*-3-linked Ara*f* or arabinan branches (Nakamura et al., 2002); and 3) arabinogalactan, β-(1,4)-linked galactan with *O*-3-linked Ara*f,* β-

(1,3)-galactan, β-(1,6)- galactan, or arabinan branches (Mohnen, 2002; O'Neill & York, 2003). RG-I side chains are highly variable among different plants and have been associated with disruption of HGA cross-linking during guard cell activity of transpiration control (Jones et al., 2003).

The substituted galacturonans XGA and ApiGalA are uniquely abundant in some plant species and tissues. XGA is an α-(1,4)-Gal*p*A chain with branches of β-(1,3)-linked D-xylose (Xyl*p*) or xylobiose (Xyl*p*- β-(1,4)-Xyl*p*) and is found in stems and leaves of *Arabidopsis* as well as reproductive and storage tissues of certain plants (Zandleven et al., 2007). ApiGalA is a similar α-(1,4)-Gal*p*A chain with branches of β-(1,2/3)-linked D-apiofuranose (Api) or apiobiose (Api-β-(1,3')-Api) and is distinctive to certain duckweeds and seagrasses (Duff, 1965; Hart & Kindel, 1970; Watson & Orenstein, 1975). While these substituted galacturonans are specific to certain plants and tissues types, their function is unclear. The branching of the Gal*p*A backbone likely inhibits metal complexing and is resistant to endopolygalacturonase (EPG) hydrolysis, suggesting they may have a role in defense response to microbial pathogens (Avci et al., 2018; Jensen et al., 2008). Additionally, ApiGalA production in *Spirodela* is downregulated upon abscisic acid (ABA) treatment and dormancy induction, indicating its synthesis has a high metabolic cost and developmentally relevant function (Longland et al., 1989).

With as many as 20 different sugars in 12 different linkages, RG-II is considered the most complex polysaccharide on Earth (Perez et al., 2003). Similar to the other substituted galacturonans, RG-II consists of an α-(1,4)-Gal*p*A chain with much more complex side chains designated A-D (O'Neill et al., 2004) (**Figure 1.4**). Side chains C and D are initiated by α-(2,3)-2-keto-3-deoxy-D-manno-octulosonic acid (Kdo) and β-(2,3)-2-keto-3-deoxy-D-lyxo-heptulosaric acid (Dha), respectively and extended by α-(1,5)-Rha*p* and β-(1,5)-Ara*f*,

respectively. Side chains A and B are initiated by β-(1,2)-Api and further decorated with unique sugars in diagnostic linkages, including α-L-aceric acid (Ace*f*A), 2-*O*-methyl-α-L-fucopyranose (2MeFuc*p*), and 2-*O*-methyl-α-D-xylopyranose (2MeXyl*p*). Despite its complex composition, the glycosyl sequence of RG-II is extraordinarily conserved across diverse vascular plant species. Some minor variations from wine RG-II (**Figure 1.4**) have been reported in the substitutions at the α-L-arabinopyranose (Ara*p*) residue of side chain B, with the most notable being α-(1-2)-L-rhamnopyranose (Rha*p*) instead of α-(1-3)-Rha*p* and absence of terminal β-L-Ara*f* in *Arabidopsis* (Glushka et al., 2003). Analyses of the fine structures of wine and *Arabidopsis* RG-II indicate potential variability of side chains A and B between and within species (Buffetto et al., 2014; Pabst et al., 2013).

FIGURE 1.4

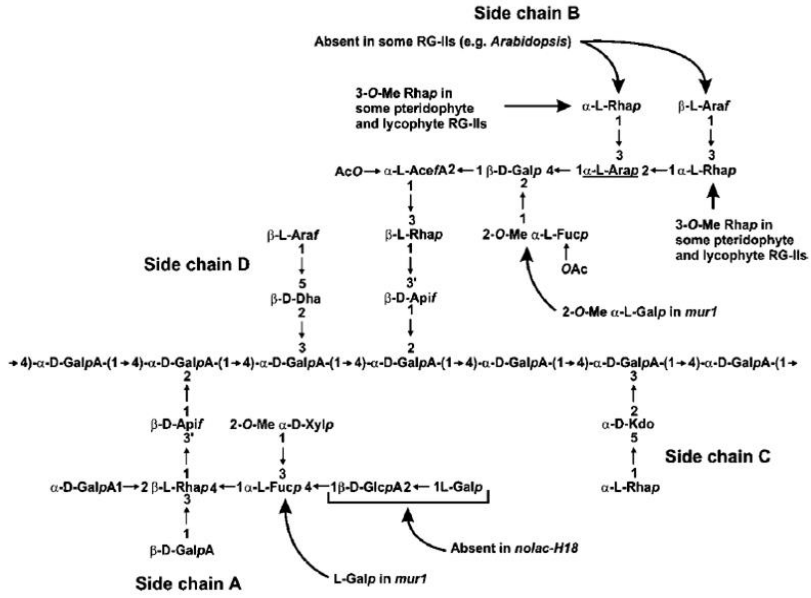


FIGURE 1.4. From O'Neill et al. (2004). Fine structure of RG-II with side chains A-D. Changes in structure from different species and mutants are indicated by arrows.

The possibility of additional tissue- or species-specific glycoforms of RG-II cannot be excluded. Though no conserved RG-II structure has been described in avascular plants, trace amounts of 2MeFuc*p* and 2MeXyl*p*, are detectable in certain avascular bryophytes suggesting RG-II may be required for specialized cell types present in bryophyte reproduction (Matsunga et al., 2003). The most primitive vascular land plants, pteridophytes and lycophytes, have similar amounts of RG-II as later-evolved angiosperms, but the avascular bryophytes (including mosses, hornworts and liverworts) appear to contain small amounts of RG-II, if any (Matsunga et al., 2003).

RG-II only makes up as much as 8 % of the cell wall (w/w), even in pectin-rich dicots (Mohnen, 2008; O'Neill et al., 2004). In vascular plants, RG-II is primarily a homodimer cross-linked by borate. The dimer forms as the two vicinal hydroxyl groups of Api in side chain A reacts with borate to form a diester (Ishii et al., 1999) (**Figure 1.5**). RG-II and its proper dimerization play an important part in higher plant growth and development. Incomplete cross-linking of RG-II induced by boron deficiency or altered RG-II structure results in stunted plants or early lethality due to weakened cell walls (Fleischer et al., 1999; O'Neill et al., 2001). Likewise, mutant *Nicotiana* and *Arabidopsis* lacking enzymes involved in the synthesis of UDP-apiose (Ahn et al., 2006; Reboul et al., 2011) and CMP-KDO (Kobayashi et al., 2011) are infertile with defective pollen tubes, indicating the importance of proper RG-II formation during plant reproduction.

FIGURE 1.5

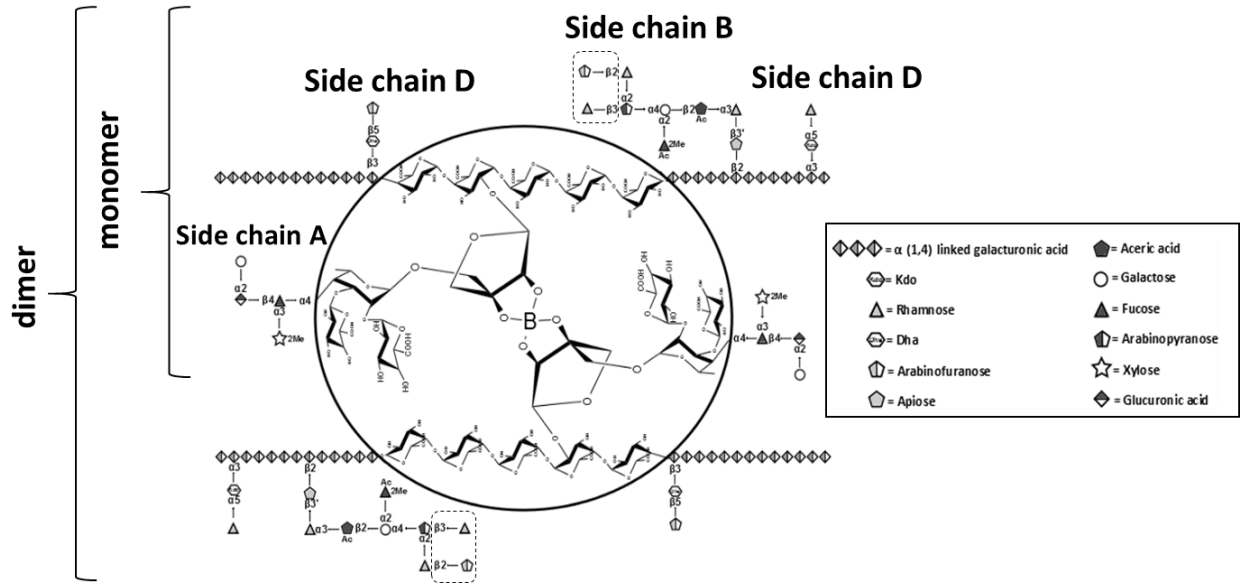


FIGURE 1.5. Structure of RG-II dimer. The borate diester through apiosyl residues of side chain A is magnified in the oval. Dotted lines represent terminal residues that are present in *Nicotinia* and wine RG-II but altered in *Arabidopsis* RG-II (O'Neill et al., 2004).

Because they have the same backbone structure and co-extract by EPG digest, it is believed that RG-II, XGA, and ApiGalA are covalently linked to HGA (Coenen et al., 2007; Ishii & Matsunaga, 2001; O'Neill et al., 1990). Reports indicate that RG-I is also linked covalently to HGA, and does not fully separate after extraction and chemical treatment (Caffall & Mohnen, 2009; G. J. Coenen et al., 2008; Nakamura et al., 2002) (**Figure 1.6**).

FIGURE 1.6

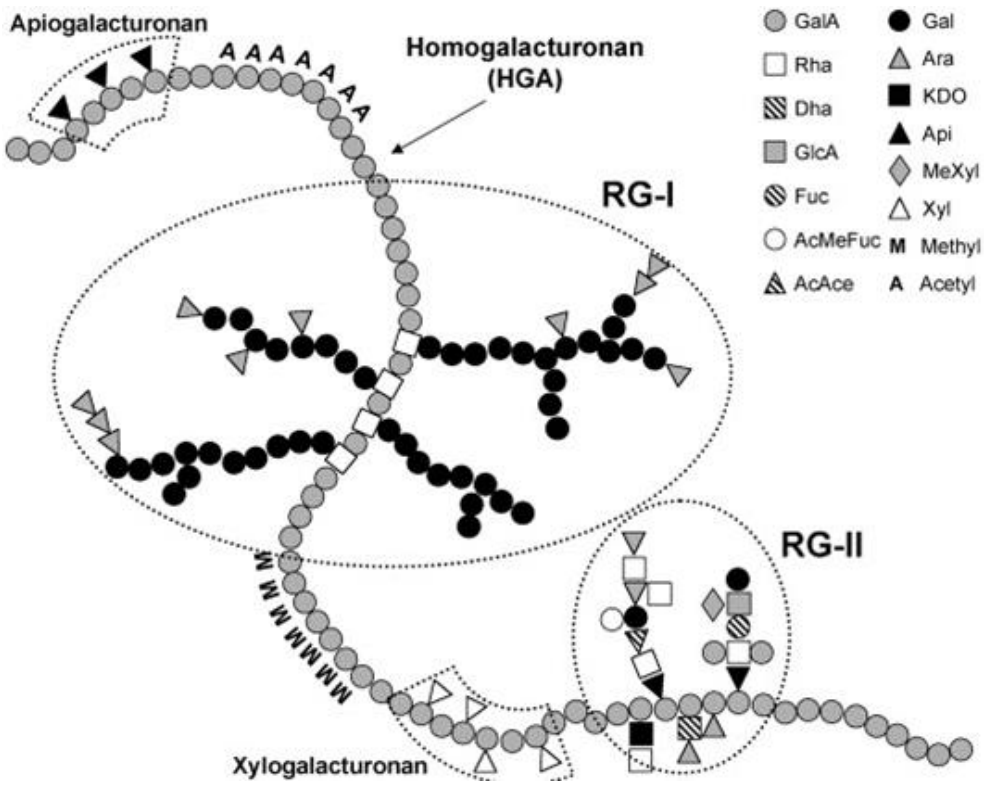


FIGURE 1.6. From Perez et al. (2003). A representation of all five pectic domains covalently linked through HGA backbone. Note: distances between domains are not to scale and are likely much greater *in vivo*.

Biosynthesis

The structures of the pectic domains have been extensively researched and characterized, but relatively little is known about their synthesis. Due to the complex glycosyl composition and modifications, at least 67 GTs, methyltransferases, and acetyltransferases are estimated to be required for biosynthesis of pectin (Caffall & Mohnen, 2009; Harholt et al., 2010; Mohnen, 2008; Ridley et al., 2001). Earlier studies using microsome preps from various plant sources demonstrated galacturonosyltransferase (GalAT) activity towards HGA elongation (Lin et al.,

1966; Villemez et al., 1965). These original assays utilized radio-labeled UDP-[^{14}C]-Gal*p*A as the donor and endogenous oligosaccharide acceptors. Later studies demonstrated exogenously supplied oligogalacturonides (OGAs) could act as acceptors (Doong et al., 1995), and more purified Golgi fractions were HGA:GalAT active (Sterling et al., 2001). Additional activities toward galactan (McNab et al., 1968) and arabinan (Odzuck & Kauss, 1972) synthesis as well as HGA methyltransferase (MT) activity (Kauss et al., 1967) were demonstrated using Golgi preps. Supporting this enzymatic evidence, immunolabeling of pectic epitopes, like HGA and RG-I demonstrated that they are located in the Golgi (Knox, 1992; Moore et al., 1991).

Plant genome and transcriptome sequencing and reverse genetic techniques have identified several putative pectin biosynthesis GTs, however characterizing the individual GTs by functional assay has been challenging. Several candidate GTs have been identified by examination of cell wall mutants. For example, QUASIMODO1 (QUA1) was identified from an *Arabidopsis* mutant with very low HGA and encodes a putative HGA:GalAT (Bouton et al., 2002; Orfila et al., 2005), and ARABINAN DEFICIENT1 (ARAD1) is a putative arabinosyltrasferase (AraT), as the *arad1* mutant has shorter arabinan side chains on RG-I (Harholt et al., 2006). Additional putative GTs involved in RG-II synthesis have been suggested by using bioinformatic approaches, but none of these have been enzymatically confirmed (Voxeur et al., 2012) (**Table 1.1**).

TABLE 1.1. *Arabidopsis* gene candidates involved in RG-II synthesis. Adapted from Voxeur et al., 2012.

Side Chain	Sugar	Linkage	Nucleotide-sugar	Mechanism	Candidates	GT family	Literature
A	β-D-Api*f*	β-1,2	UDP-α-D-Api*f*	Inverting	At3g42180, At5g25310, At4g38040; At1g04910	47; nc	Jensen et al., 2008
	*β-L-Rha*p*	α-1,3	UDP-β-L-Rha*p*	Retaining	At1g19710	4	Voxeur et al., 2012
	α-D-Gal*p*A	α-1,2	UDP-α-D-Gal*p*A	Retaining	At5g04500	64	
	α-L-Fuc*p*	α-1,4	GDP-β-L-Fuc*p*	Inverting	At5g50420	68	Voxeur et al., 2012
	α-D-Xyl*p*	α-1,3	UDP-α-D-Xyl*p*	Retaining	**At4g01220 (MGP4)**	77	Liu et al., 2011
	β-D-Gal*p*A	β-1,3	UDP-α-D-Gal*p*A	Inverting	At5g52290	31	
	β-D-Glc*p*A	β-1,4	UDP-α-D-Glc*p*A	Inverting	At3g57630	47	
	α-L-Gal*p*	α-1,2	GDP-β-L-Gal*p*	Inverting	At4g38500	Nc	Voxeur et al., 2012
B	β-D-Api*f*	β-1,2	UDP-α-D-Api*f*	Inverting	At3g42180, At5g25310, At4g38040; At1g04910	47; nc	Jensen et al., 2008
	α-L-Rha*p*	α-1,3	UDP-β-L-Rha*p*	Inverting	At1g21480	47	
	α-L-Ace*f*A	α-L	Unknown	Unknown			
	β-D-Gal*p*	β-1,2	UDP-α-D-Gal*p*	Inverting	**At5g53340 (HPGT1)**	31	Voxeur et al., 2012
	α-L-Fuc*p*	α-1,4	GDP-β-L-Fuc*p*	Inverting	At5g50420	68	Voxeur et al., 2012
	α-L-Ara*p*	α-1,4	UDP-β-L-Ara*p*	Inverting	At2g35100	47	
	α-L-Rha*p*	α-1,2	UDP-β-L-Rha*p*	Inverting	At5g25820	47	
	β-L-Ara*f*	β-1,2	UDP-β-L-Ara*f*	Retaining	At2g02060, At1g28695, At5g40900; At1g19710; At1g14590, At2g02061	77; 4	Voxeur et al., 2012
C	α-D-Kdo	α-2,3	CMP-β-Kdo	Inverting	At1g08660, At3g48820	29	Dumont et al., 2014
	α-L-Rha*p*	α-1,5	UDP-β-L-Rha*p*	Inverting	At2g03360, At2g03370	61	
D	β-D-Dha	β-2,3	Unknown	Unknown	At1G08660, At3g48820	29	Dumont et al., 2014
	β-L-Ara*f*	β-1,2	UDP-β-L-Ara*f*	Retaining	At2g02060, At1g28695, At5g40900; At1g19710; At1g14590, At2g02061	77; 4	Voxeur et al., 2012

* indicates updated configuration. Bold type indicates already characterized. NOTE: there is overlap/redundancy in candidates.

Only a select few GTs involved in pectin biosynthesis have been unambiguously defined. This short list includes GAUT1, an HGA:GalAT (Sterling et al., 2006); XGD1, a xylogalacturonan-specific β-(1,3)-xylosyltransferase (XylT) that adds Xyl*p* to HGA (Jensen et al., 2008); RGXT1, RGXT2, and RGXT3, are all α-(1,3)-XylTs that add Xyl*p* to L-Fuc*p* of RG-II side chain A (Egelund et al., 2008; Egelund et al., 2006); and GALS1, a β-(1,4)-galactosyltransferase (GalT) that adds Gal*p* to galactan (Liwanag et al., 2012).

A potential reason so few pectin biosynthesis GTs are conclusively characterized is that they are active *in vivo* in protein complexes. When GAUT1 was identified from tryptic digest of partially purified *Arabidopsis* microsomes, a homologous protein, GAUT7, was also detected by MS (Sterling et al., 2006). Heterologously expressed GAUT7 alone did not show HGA:GalAT activity, and SDS-PAGE demonstrated that GAUT7 covalently links to GAUT1 (Atmodjo et al., 2011). Bimolecular fluorescence complementation (BiFC) of transiently co-expressed GAUT1 with GAUT7 indicated GAUT7 serves to anchor GAUT1 in the Golgi (Atmodjo et al., 2011). Similarly, through fluorescence resonance energy transfer (FRET), ARAD1and its close homolog ARAD2 were shown to co-localize in the Golgi, and they homo- and hetero- dimerize, likely through disulfide bridges evidenced by non-denaturing SDS-PAGE (Harholt et al., 2012). The *arad2* mutant, however, had similar arabinan as wild type (WT) and did not complement *arad2*, suggesting ARAD1 and ARAD2 do not have redundant activities (Harholt et al., 2012).

Nucleotide sugars

Like currently characterized pectin GTs, additional GTs involved in pectin biosynthesis are predicted to use nucleotide sugars as immediate substrates (Ridley et al., 2001), and a large pool of many different activated sugar donors is necessary (Bar-Peled et al., 2012) (for RG-II,

see **Table 1.1**). Therefore, nucleotide sugar metabolic pathways are vital for plant growth and development. This is supported by studies of mutants impacting RG-II such as *mur1* in *Arabidopsis*. *Mur1* has a mutation in GDP-D-mannose-4,6-dehydratase (GMD), which catalyzes the first step in synthesis of GDP-Fuc*p* (O'Neill et al., 2001). *Mur1* plants have such severely reduced Fuc*p* that Gal*p* replaces Fuc*p* in RG-II side chains (see **Figure 1.4**), and plants are dwarfed (O'Neill et al., 2001). Additional evidence is derived from viral-induced gene silencing (VIGS) of a UDP-apiose/UDP-xylose synthase in *Nicotiana*. UAS catalyzes the synthesis of UDP-apiose (UDP-Api) and UDP-xylose (UDP-Xyl) from UDP-glucuronic acid (UDP-GlcA) (Guyett et al., 2009; Molhoj et al., 2003). Mutants had reduced levels of Api, 2MeFuc*p*, and 2MeXyl*p* and suffered wall thickening and cell death due to defective RG-II assembly (Ahn et al., 2006).

To ensure availability of required nucleotide sugars for pectin (and other glycan) synthesis, plants have evolved multiple metabolic pathways of synthesis. The “map” of plant nucleotide sugar synthesis has several nodes (**Figure1.7**) where substrates can interconvert. In addition to de-novo synthesis and interconversion, plants can salvage free monosaccharides. The process begins with phosphorylation of a monosaccharide into a sugar-1-phosphate by kinase activity. Subsequently, pyrophosphorylases catalyze nucleotidyl transfer from a nucleotide triphosphate, generating a nucleotide sugar and inorganic pyrophosphate (Kotake et al., 2004; Schnurr et al., 2006; Yang et al., 2009).

FIGURE 1.7

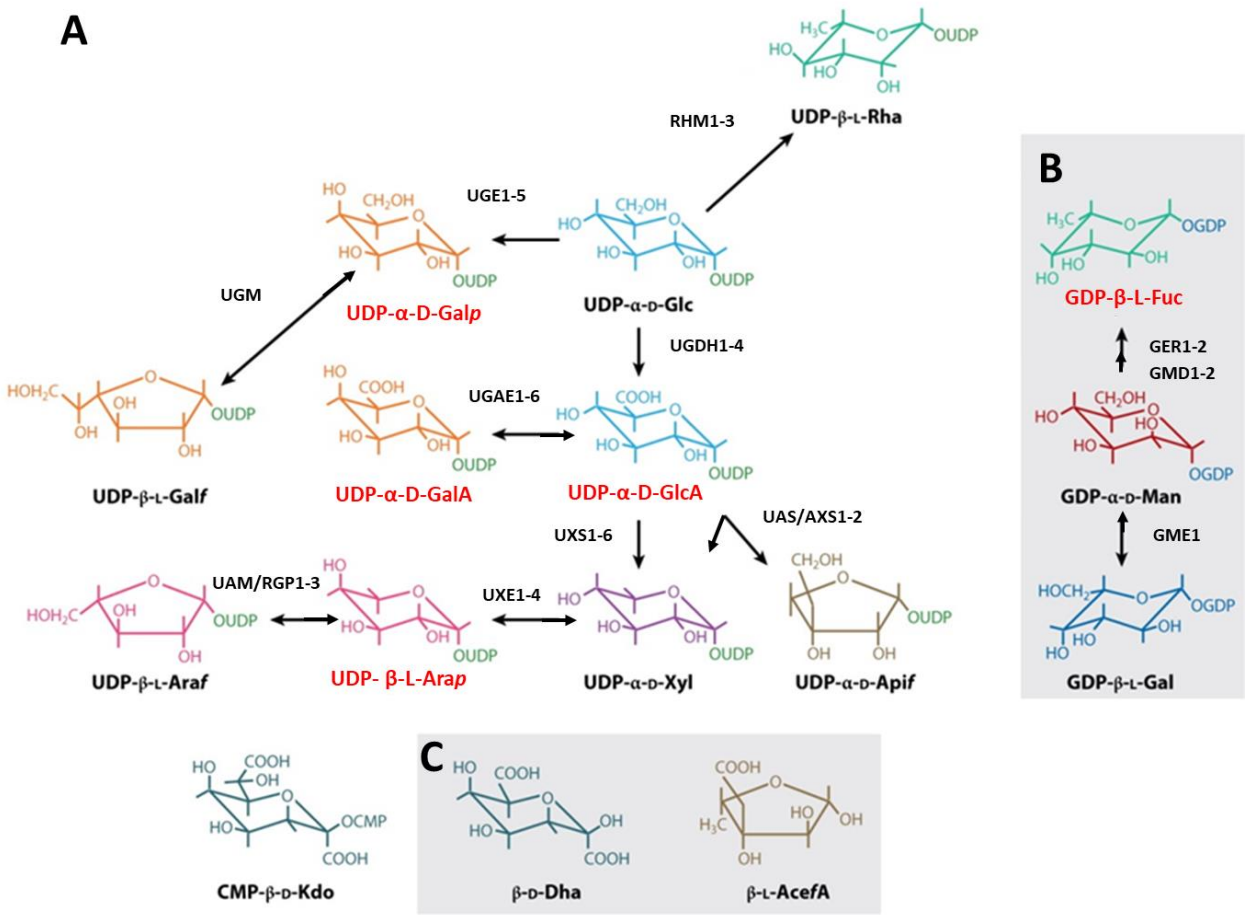


FIGURE 1.7. Adapted from Bar-Peled and O'Neill (2011). Nucleotide sugars used by plants and some interconversion pathways with known enzymes labeled. Red labels indicate a known salvage pathway. (**A**) Partial nucleotide sugar metabolic pathway involving substrates and enzymes described in text, excluding the following: UDP-galactose mutase (UGM) catalyzes interconversion of UDP-Gal*p* and UDP-galactofuranose (UDP-Gal*f*), and UDP-rhamnose synthase (RHM) catalyzes synthesis of UDP-rhamnose (UDP-Rha) from UDP-Glc. (**B**) Partial nucleotide sugar metabolic pathway. GDP-mannose epimerase (GME) catalyzes interconversion of GDP-L-galactose (GDP-L-Gal) and GDP-mannose (GDP-Man). GDP-keto-6-deoxymannose 3,5-epimerase 4-reductase (GER) catalyzes synthesis of GDP-Fuc from GDP-keto-6-

deoxymannose, the product of GMD. (**C**) No known forms of activated Dha or Ace*f*A are known.

Plants also have multiple copies of nucleotide sugar synthesis genes, further supporting the importance of nucleotide sugar substrates in proper pectin assembly. Homologs and gene families provide redundancy to ensure supply of nucleotide sugar substrates, for example there are four UDP-glucose dehydrogenase (UGDH) genes in *Arabidopsis* (Klinghammer & Tenhaken, 2007). UGDH catalyzes the synthesis of UDP-GlcA from UDP-glucose (UDP-Glc). Single mutant plants in UGD2 or UGD3 do not display growth phenotypes, but the *ugd2/3* double mutant showed thickened walls and shorter xyloglucan chains (Reboul et al., 2011). Interestingly, as UDP-GlcA is a substrate needed to synthesize other nucleotide sugars, *ugd2/3* mutants had reduced GalA, xylose, arabinose, and apiose (Reboul et al., 2011), suggesting UDP-GlcA synthesis is a limiting step in pectin synthesis.

UDP-D-glucuronate 4-epimerase (UGAE), which catalyzes interconversion between UDP-GlcA and UDP-galacturonic acid (UDP-GalA) has six members in its gene family in *Arabidopsis*, and all isoforms shows similar expression across tissue types (Gu & Bar-Peled, 2004; Molhoj et al., 2004; Usadel et al., 2004). Similarly, the UDP-xylose synthase (UXS) gene family in *Arabidopsis* also has six members (Harper & Bar-Peled, 2002; Pattathil et al., 2005). UXS catalyzes the synthesis of UDP-Xyl from UDP-GlcA. In *Arabidopsis*, UXS1, UXS2, and UXS4 are membrane-bound with active sites in the Golgi lumen, while UXS3, UXS5, and UXS6 are cytosolic. In addition, the cytosolic isoforms appear to impact xylan synthesis, as *uxs3/5/*6 mutants showed decreased xylan content, while *uxs1/2/4* mutants showed xylan content similar to wild type (Zhong et al., 2017). Mutation of one of the five UDP-glucose 4-epimerases (UGEs)

in *Arabidopsis* also had unexpected results. UGE catalyzes the conversion of UDP-Glc to UDP-galactose (UDP-Gal). Examination of the mutant demonstrated decreased xyloglucan (XyG) only in roots, with no effect on RG-I (galactan) or RG-II composition (Nguema-Ona et al., 2006). Transcription of different isoforms and partitioning of nucleotide sugar synthesis genes implies there may be levels of regulation of plant wall polysaccharide synthesis influenced by different rates of isoform activity, or isoform association with glycan-specific protein complexes.

Nucleotide sugars synthesized in the cytosol are transported via nucleoside monophosphate/nucleotide sugar antiporters into the Golgi lumen, where synthesis of many wall matrix polysaccharides occurs (Capasso & Hirschberg, 1984). In *Arabidopsis* a family of nucleotide sugar transporters (NSTs) that transport nucleotide sugars from the cytosol into the Golgi lumen has been described and several members characterized (Ebert et al., 2015; Rautengarten et al., 2016; Rautengarten et al., 2014). Transport of the nucleotide sugars UDP-arabinopranose (UDP-Ara*p*) and UDP-arabinofuranose (UDP-Ara*f*) has presented a conundrum; UDP-xylose 4-epimerase (UXE) catalyzes the interconversion of UDP-Xyl to UDP-Ara*p* in the Golgi. The enzyme UDP-arabinose mutase/reversibly glycosylated peptide (UAM/RGP) interconverts UDP-Ara*p* and UDP-Ara*f* in the cytosol. Confocal microscopy data indicates that *Arabidopsis* UAM/RGP1, UAM/RGP2, and UAM/RGP5 localize in the cytosol but highly associated with the Golgi (Rautengarten et al., 2012). In addition, immunoprecipitation of hemagglutinin (HA)-tagged protein revealed that all three UAM/RGPs were present in protein complexes (Rautengarten et al., 2012). Therefore, in order for arabinosytransferases in the Golgi to use UDP-Ara*f*, UDP-Ara*p* must be transported out of the Golgi into the cytosol, where UAM/RGP can convert it into UDP-Ara*f*, which needs to be transported back into the Golgi (**Figure 1.8**) (Oikawa et al., 2013; Rautengarten et al., 2017). For this process to work

efficiently, these proteins would have to be tightly associated if not in a complex. A similar model has been proposed for AXS1, a UDP-apiose/UDP-xylose synthase (UAS) from *Arabidopsis* (Reboul & Tenhaken, 2012). The product UDP-Api, like certain other nucleotide sugars (notably UDP-Ara*f*) is not stable in solution and quickly degrades when synthesized *in vitro*.

FIGURE 1.8

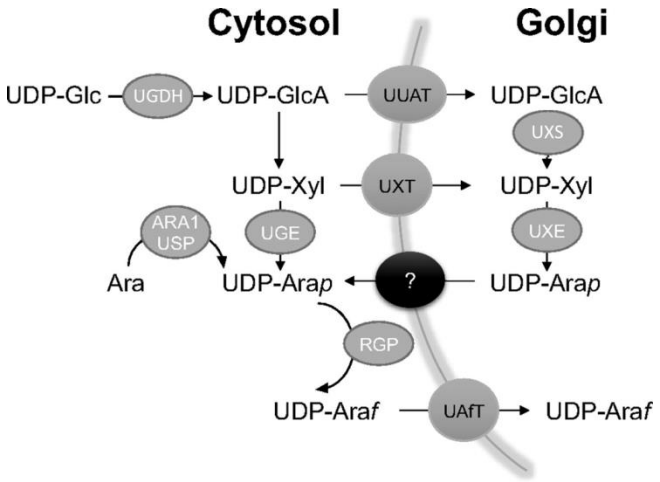


FIGURE 1.8. Adapted from Rautengarten et al. (2017). Model of UDP-Ara*f* synthesis and its transport into the Golgi lumen. Salvage pathway is represented; L-arabinokinase (ARA1) phosphorylates Ara to Ara-1-P, and UDP-sugar pyrophosphorylase synthesizes UDP-Ara*p*. Transporters listed are UDP-uronic acid transporter (UUAT), UDP-Xyl transporter (UXT), and UDP-Ara*f* transporter (UAra*f*T). UDP-Ara*p* transporter is unidentified.

Together, numerous GTs (at least 20 for RG-II, alone), several nucleotide sugar synthesis enzymes, and a handful of nucleotide sugar transporters work in a confined Golgi space to generate all non-cellulosic plant cell wall polysaccharides (Driouich et al., 2012). Considering the low amounts of RG-II in the primary wall and the fact that several of its component sugars

are unique to its structure, it is probable that the nucleotide sugars dedicated to RG-II biosynthesis are tightly controlled and limiting.

APIOSE

Natural occurrence

The pentose apiose (Api) was first described in 1901 as a sugar constituent of the flavone apiin (apiosyl-glucoside of the flavone apigenin) extracted from parsley, a member of the genus Apium (Vongerichten, 1901). Since its discovery, several additional apiosylated compounds have been described in vascular plants, including different small molecule metabolites and the structural wall polysaccharides apiogalacturonan (ApiGalA) and rhamnogalacturonan-II (RG-II) (Beck, 1967; Darvill et al., 1978; Hart & Kindel, 1970; Nepogodiev et al., 2011). To date, Api is one of 20 described branch-chain monosaccharides found in nature (Beck, 1982), many of which are synthesized by microbes (Miljković, 1990). Api with hamamelose (Haberland & Kolodziej, 1994; Ozawa, 1984; Shafizadeh, 1956) and aceric acid (Spellman, 1983; Vidal et al., 2000) are the only three branched-chain sugars found in plant walls.

Structurally, Api is unusual as it has a tertiary alcohol (**Figure 1.9**). It is most commonly found in the 3-C-(hydroxymethyl)-β-D-erythrofuranose configuration, but has been found in the 3-C-(hydroxymethyl)-α-L-threofuranose configuration in at least three instances (Ahn et al., 2004; Cheng et al., 2012; Julião et al., 2009). In 26 vascular plants and six fungi, Api is a moiety of a large variety of secondary metabolites ranging from phenolic and terpenoid to aliphatic and alkaloid (Picmanova & Moller, 2016). Recently, avascular bryophytes and algae have been added to this list of plants containing apiosylated metabolites (Smith et al., 2016). In these compounds, Api can be directly attached to the aglycone, for example in the compounds

edgeworoside B (Baba, 1990) or attached to another sugar residue as in the case of apiin (O. T. Schmidt, 1930). Api may be the terminal (non-reducing) sugar, as *O*-β-D-Api has been found linked to glucose, galactose, rhamnose, xylose, and arabinose at various positions (Hamburger et al., 1985; Malafronte et al., 2009; Maurya et al., 1996; Tsai et al., 2011; Wu et al., 2007). Additionally, apiosides may be substituted by additional sugars or modifications at Api position 2 (Julião et al., 2010), position 3' (Cornelius et al., 2010), both positions 2 and 3' (Day et al., 2000), or position 3 (Wang et al., 2007). Of note, apiosides containing a carboxyl group may cyclize with a terminal Api to form a cyclic ester, as in the case for linderofruticoside A (Song et al., 2008), similar to the apiofuranosyl-1,2-cyclic phosphate degradation product of UDP-apiose synthesis (Guyett et al., 2009).

FIGURE 1.9

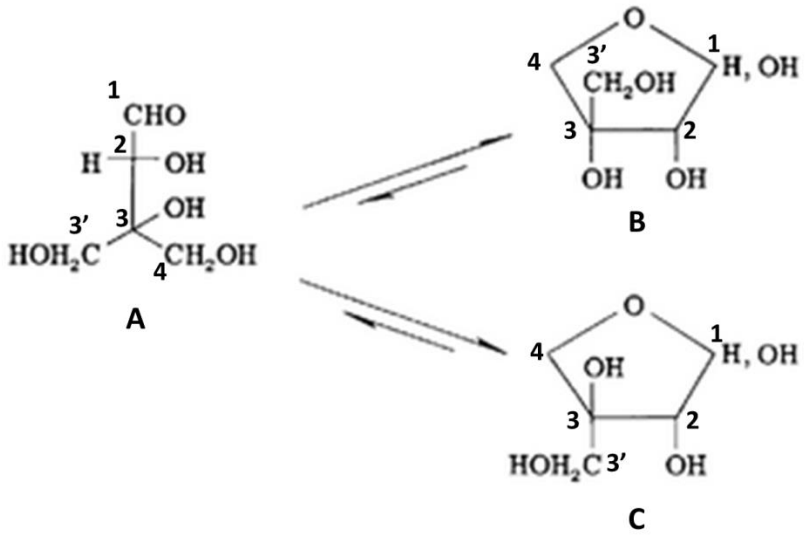


FIGURE 1.9. Adapted from Watson and Orenstein (1975). Structure of D-apiose; (**A**) open chain, (**B**) erythrofuranose, and (**C**) threofuranose configurations. Carbons are numbered.

Api is also a component of vascular plant cell walls in the pectic polysaccharides RG-II and ApiGalA. In RG-II, Api of side chain A forms the borate ester bridge to dimerize two monomers, which can occur in parallel or antiparallel alignment (Bar-Peled et al., 2012; Ishii et al., 1999; Perez et al., 2003). Modeling the RG-II dimer using computational and spectroscopic techniques suggests two "disk-like" monomers stack on top of each other to dimerize through borate (Perez et al., 2003), and that this is stabilized by Ca^{2+} (Chormova & Fry, 2016; Fleischer et al., 1999; Kobayashi et al., 1999). Since current methods of RG-II research involve EPG treatment and extraction, it is unclear which dimer orientation forms *in muro*, and our understanding of RG-II organization is limited.

Significantly large amounts of Api are found in the cell walls of the duckweeds *Lemna* and *Spirodela* and the seagrass *Zostera* in the form of ApiGalA (Hart & Kindel, 1970; Watson & Orenstein, 1975). *Spirodela* non-cellulosic cell-wall polysaccharides contain up to 26% apiose (w/w) (Longland et al., 1989), and ApiGalA comprises 11% of *Zostera* dry weight (w/w) (Gloaguen et al., 2010). It is unclear whether ApiGalA also dimerizes through borate, but recent findings that *Lemna* accumulate large amounts of boron suggest that this micronutrient is sequestered as borate esterified by ApiGalA (Avci et al., 2018).

Functions of apiosides

Since apiosides occur in a wide variety of compounds, the functional significance of naturally occurring apiose is broad. Plant secondary metabolites have been implicated in many functions, from signaling, protection against pathogens and herbivores, prevention of oxidative degradation, attracting pollinators and more. As a general property, addition of a glycose to a secondary metabolite increases its solubility of non-polar compounds and may provide storage of

energy. In the case of apiose, proposed functionality is varied and speculative. Proposed functions of apiosides are characterized by the function of the aglycone to which Api is attached or and by association to similar compounds. For example Api has been considered to protect against cold, because the branched-chain sugar alcohol hamamelitol is suggested to cryoprotect certain winter-hardy plants (Sellmair & Beck, 1968).

Flavones are natural antioxidants, protecting plant cells from oxidative stress. Because apiin biosynthesis in parsley is stimulated by ultraviolet (UV) irradiation and ozone, these apiosides may protect against oxidative insult (Eckey-Kaltenbach et al., 1993). Similarly, cell suspensions of parsley irradiated by UV demonstrate an increase in UAS transcript (Gardiner et al., 1980).

More convincing evidence for a role of Api comes from studies with apiosylated and non-apiosylated compounds. For example, gentisic acid 5-O-β-D-apiofuranosyl-(1→2)-β-D-glucopyranoside from *Orixa japonica* acted as a stronger repellant against the rutaceae-feeding butterfly *Papilio xuthus* L. than gentisic acid 5-O-β-D-glucopyranoside (Ono et al., 2004).

Though untested, the release of free Api may act as a cell signal to plants under attack, in a similar cascade response as released HGA fragments (Ferrari et al., 2013). An apiosidase from the fungus *Aspergillus niger* hydrolyzes the glycosyl bond in apiosyl-glucosides of a terpene from grape, releasing free apiose (Dupin et al. 1992; Guo et al. 1999).

Biosynthesis of apiose

Like certain other sugars, Api is first synthesized in its activated nucleotide sugar form, UDP-Api. First described by Sandermann and Grisebach in 1970, the mechanism for UDP-Api synthesis is still unsettled. Much work has been done to isolate and characterize the genes and

enzymes responsible for apiose metabolism. Partially purified extracts from parsley and Lemna cell suspensions demonstrated the generation of UDP-Api and UDP-Xyl from UDP-GlcA by an NAD^+-dependent, 86 kDa enzyme through a decarboxylation reaction with a 4-keto intermediate (Baron et al., 1972; Gebb et al., 1975; Kindel & Watson, 1973; Matern & Grisebach, 1977; Mendicino & Abou-Issa, 1974; Sandermann & Grisebach, 1970; Sandermann et al., 1968; Wellmann & Grisebach, 1971) (**Figure 1.10**). The first functionally cloned and characterized UAS, AXS1 from *Arabidopsis* (Molhoj et al., 2003) opened the possibility of finding UAS in other plant species. Real-time NMR monitoring of *Solanum* UAS showed formation of the UDP-4-ketoxylose intermediate as well as a apiofuranosyl-1,2-cyclic phosphate degradation product (Guyett et al., 2009).

UAS is classified as an extended short-chain dehydrogenase/reductase (SDR), and shares a conserved NAD^+-binding domain and catalytic site amino acid sequence with UXS, and ArnA; two other extended SDRs that use the same UDP-GlcA substrate. The proposed reaction mechanism begins similarly to UDP-xylose synthase with a decarboxylation of UDP-GlcA and formation of UDP-4-ketoxylose intermediate. Next, a unique cleavage between C-2 and C-3 opens the ring through a suggested retroaldol reaction, generating an enediol intermediate that undergoes aldol condensation and reduction by NADH to UDP-Api (Choi et al., 2012). The bi-product UDP-Xyl is generated by direct reduction of the UDP-4-ketoxylose intermediate by NADH. A recent update to the UAS reaction mechanism suggests the first reaction step is a coupled oxidation/retroaldol cleavage to form an enediol transient that is then decarboxylated, aldol ring-contracted, and reduced to form UDP-Api (Eixelsberger et al., 2017). Substrates were site-specifically isotope-labeled to observe reaction intermediates and products by NMR. This is an interesting proposal, however, absence of the UDP-4-ketoxylose intermediate observed from

other UAS studies is unexplained, and high protein homology of UAS to UXS suggests catalysis through a similar mechanism. Crystallization of UAS would provide valuable insight into the true reaction mechanism.

FIGURE 1.10

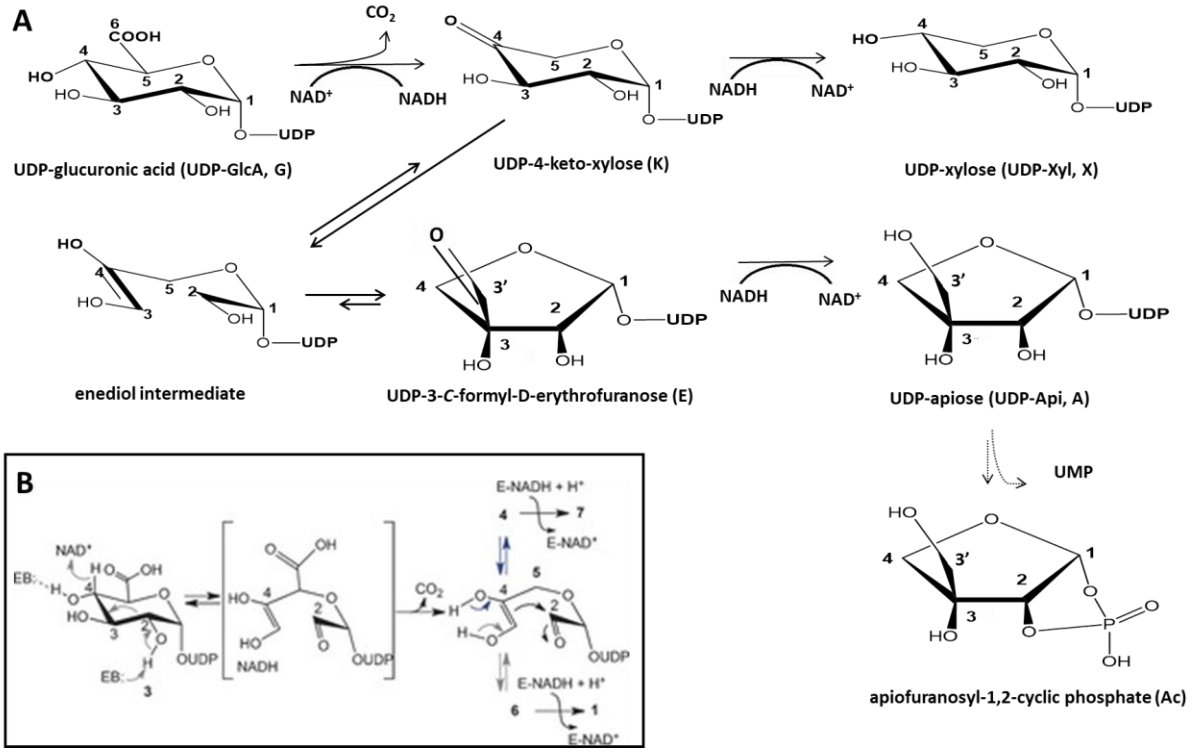


FIGURE 1.10. Adapted from Eixelsberger et al. (2017). UDP-apiose/UDP-xylose synthase reaction mechanism. (**A**) The proposed UAS reaction mechanism: UDP-GlcA decarboxylation to UDP-4-ketoxylose intermediate, retroaldol ring cleavage to an enediol intermediate, aldol condensation to UDP-3-*C*-D-erythrofuranose intermediate, and reduction to UDP-Api. UDP-Xyl is formed by direct reduction of UDP-4-ketoxylose. Degradation of UDP-Api to apiofuranosyl-1,2-cyclic phosphate and UMP occurs spontaneously. (**B**) Proposed updated UAS reaction mechanism: Coupled oxidation/retroaldol cleavage reaction generates enediol transient, followed by decarboxylation, aldol ring-contraction, and reduction in two potential products.

The only reports of apiosyltransferase activity to date are from partially purified *Lemna* and parsley cell suspensions. The *Lemna* enzyme exhibited transfer of radio-labeled [^{14}C]Api from *in vitro*-synthesized UDP-[^{14}C]-Api to endogenous acceptors to synthesize [^{14}C]ApiGalA (Mascaro & Kindel, 1977). The parsley enzyme similarly showed transfer of [^{14}C]Api to exogenously supplied apigenin-7-O-glucoside (Ortmann et al., 1972). The parsley enzyme was optimally active at neutral pH, inhibited by divalent metals, strongly inhibited by UDP and had a range of acceptors; most 7-O-glucosides of flavones (Ortmann et al., 1972).

SDR GENE FAMILY

The short-chain dehydrogenase/reductase (SDR) superfamily is vast with members in every domain of life (Jornvall et al., 1999). The members of this family are NAD(P)(H)-dependent oxidoreductases with a large range of substrates, including lipids, amino acids, and carbohydrates. For many SDRs sequence homology is not high. All share a common N-terminal glycine-rich region that binds pyrophosphate and a Rossmann fold for dinucleotide binding; a central twisted parallel β-sheet bordered by α-helices (Lesk, 1995). SDRs are sub-classed into four groups: classical, divergent, complex and extended (**Figure 1.11**).

FIGURE 1.11

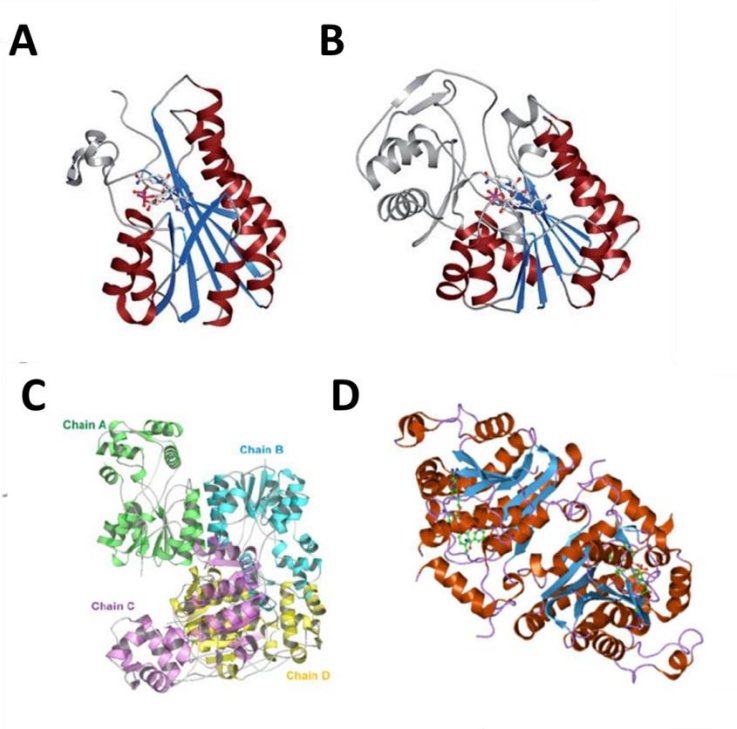


FIGURE 1.11. Adapted from Kavanagh et al. (2008), Kim et al. (2008), and Fogel et al. (2010). Ribbon diagrams of SDRs; (**A**) Classical; (**B**) Extended; (**C**) Complex; (**D**) Divergent. Rossmann folds motifs with (**A**), (**B**), (**D**) β-strands in blue and α-helices in red, additional structure grey. (**C**) multiple domains assembled into complex. (Fogel et al., 2010; Kim et al., 2008)

Another commonality of SDRs is that they perform acid-base catalysis via an ionized hydroxyl-tyrosine (Tyr, Y) that transfers a proton to/from substrate (Liu et al., 1997; Thoden et al., 2000). This activity is improved by an adjacent lysine (Lys, K) residue, which lowers the hydroxyl-tyrosine pKa (Benach et al., 1999). An active-site serine (Ser, S) polarizes the substrate carbonyl (Oppermann et al., 1997). The main-chain carbonyl of an asparagine (Asn, N) in the active site complexes water, connecting it to Tyr and assisting in proton transfer (Filling et al., 2002). Most are homodimers or homotetramers (Grimm et al., 2000).

The classical SDRs are modeled after *Drosophila* alcohol dehydrogenase (ADH). Members are ~250 a.a. in length and share a single-domain structure with some sequence variations in the cofactor binding region. The active site sequence motif is YxxxK. Reactions proceed through a mechanism wherein the coenzyme binds first and leaves last (Chang et al., 2007; Sahni-Arya et al., 2007). They perform oxidoreduction on hydroxyl- and keto groups of small molecule substrates, such as alcohols and secondary metabolites

Divergent SDRs replace the active-site Lys with a hydrophobic residue, so the active site motif is YxxMxxxK (where M is methionine). The Lys is usually located 4 residues away and is spatially arranged to still facilitate activity (Kallberg et al., 2002). As the name implies, divergent SDRs contain members that perform double-bond reduction and may not utilize the active-site Tyr for acid-base catalysis (Fillgrove & Anderson, 2001).

Complex SDRs are part of large, multidomain enzyme complexes. The Asn, Ser, Tyr, and Lys are oriented into an active-site from parts of the scaffold. (Kallberg et al., 2002). The active-site sequence motif for complex SDRs is YxxxN. The complex SDRs include fatty acid synthases.

The extended SDRs are ~350 a.a. long and have the canonical YxxxK active-site motif (Jornvall et al., 1995). These include epimerases, such as UGE, and decarboxylases, such as UXS, for which mechanisms have been extensively elucidated. In the case of the UGE reaction, NAD^+ binds tight and stays bound as its redox state changes. A proton is drawn from 4-OH with hydride transfer from C-4 to NAD^+. The 4-ketohexose intermediate rotates, presenting the opposite side to NADH for reduction.

The pocket size of epimerases influences substrate specificity; since human UGE has a larger cleft, it can accommodate catalysis of interconversion of UDP-*N*-acetyl-glucosamine

(UDP-GlcNAc) and UDP-*N*-acetyl-galactosamine (UGalNAc) (Liu et al., 1997; Thoden et al., 2001). This substrate promiscuity complicates assigning true *in vivo* function of similar epimerases. The enzyme UDP-glucuronic acid 4-epimerase (UGAE) catalyzes a similar reaction as UGE; the interconversion of UDP-GlcA and UDP-GalA. Because both UGE and UGAE operate with the same mechanism, UGAEs may be miss-annotated as UGE (Reiter, 2008).

The mechanism of extended SDR decarboxylase activity is exemplified by UXS. Again, NAD^+ cofactor binding drives oxidation of C-4-OH and decarboxylation of the C-6 carboxyl to form 4-keto-pentose. The intermediate is then reduced by NADH to generate UDP-Xyl (Bar-Peled et al., 2001). Other UDP-GlcA decarboxylases, such as bacterial ArnA and, likely, the initial steps of UAS share this mechanism. ArnA encodes a UDP-GlcA decarboxylase able to form an intermediate UDP-4-keto-L-arabinose (Breazeale et al., 2005; Gu et al., 2010) on route to the formation of UDP-arabinose-4-amino and UDP-Ara4NF.

Crystallization of human UXS (hUXS) and of bacterial ArnA demonstrates that the strength of NAD^+ cofactor binding directly correlates to rate of decarboxylase reduction of the intermediate and release of NADH (Polizzi et al., 2012). ArnA can synthesize UDP-Xyl, however, because ArnA does not tightly hold NAD^+, once the first step oxidation occurs, NADH and intermediate are allowed to leave the enzyme before UDP-Xyl can be generated through reduction by NADH. Because NAD^+ binding strength is determined by the number of H-bonds formed between enzyme and cofactor, slight differences in amino acid sequence in the NAD^+-binding domain among UDP-GlcA decarboxylases likely influences what product is formed (Smith et al., 2016; Smith & Bar-Peled, 2017) (**FIGURE 1.12)**.

FIGURE 1.12

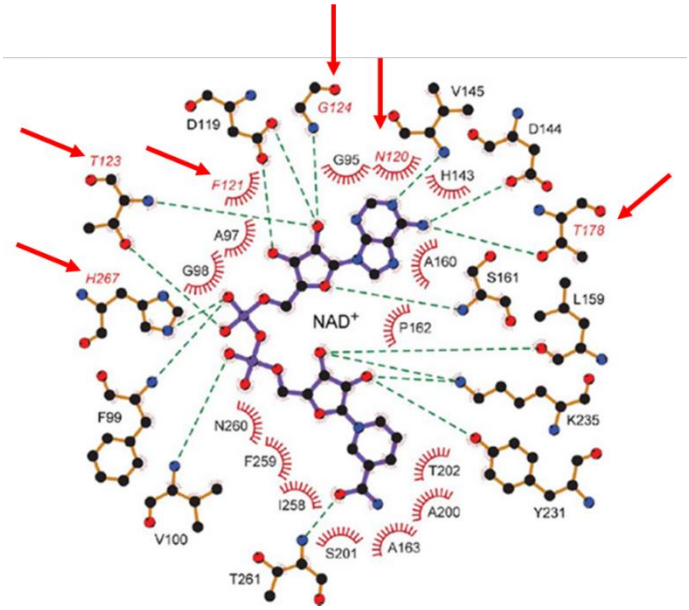


FIGURE 1.12. Adapted from Polizzi et al. (2012). Ligplot of H-bonding interactions between hUXS and NAD^+ (dotted lines) and packing interactions (feathered lines). Amino acid interactions present in hUXS not in ArnA are indicated by red arrow and in red italics.

GRAM-NEGATIVE CELL WALL GLYCANS

Structures

Bacteria are primarily classified 1) according to Gram-staining and 2) according to cell shape. Gram-negative bacteria are defined as such due to their inability to retain crystal violet, a staining method developed by the Danish scientist Hans Christian Gram. The reason these bacteria do not stain as gram-positive do is that they have a lipid-rich outer plasma membrane and a much thinner layer of peptidoglycan (Salton, 1963). Bacterial cell shape is carefully controlled and maintained to support cell motility, attachment, predation, or predator escape (Young, 2006). The shape and integrity of a bacterial cell relies on continuous reorganization and

preservation of elements of the cell wall (Holtje, 1998), including the capsule, LPS, peptidoglycan, and glycoproteins.

Many gram-negative bacteria are surrounded by a glycan layer called a capsule, which acts as a protective coat. Up to 10 μm thick, capsules are ~95% water and characterized as viscous gels (Beveridge, 1981). Capsules are made of long, typically charged polysaccharides (CPSs) with a DP of up to 200 sugars (Pelkonen & Finne, 1989). Structures are varied within and among species and serotypes. CPSs are synthesized by either the Wzy-dependent or ATP-binding cassette (ABC) transporter-dependent pathways. In the Wzy-dependent pathway, a polyprenol-linked CPS in the cytoplasm is flipped across the inner membrane (IM) by the integral membrane protein Wzx and extended by another integral membrane Wzy protein with undefined "polymerase" activity (Daniels et al., 1998). The completely assembled CPS is translocated across the OM by a channel Wza protein. In the ABC-dependent pathway, CPSs synthesized in the cytoplasm are secreted by an ABC transporter and translocated outside the cell by a protein complex coordinated through an outer membrane polysaccharide protein (OPX) (Whitfield, 2006).

At the cell surface, proteins, phospholipids, and LPS are the "barrier" layer, which must be porous enough to allow diffusion of nutrients and waste (Cowan et al., 1992). The membrane contains porins, and certain hydrophobic molecules can penetrate the barrier. Outer membranes of gram-negative bacteria are permeable to steroid probes (Plesiat & Nikaido, 1992). Surface binding proteins retain vitamins and can act as a reserve of pilins or virulence factors (Kessler & Safrin, 1988).

LPS are most widely recognized as immunogenic compounds from gram-negative bacterial pathogens, but these molecules provide structural integrity and a semi-permeable

barrier to gram-negative bacteria. The structure is divided into a conserved lipid-A in the outer membrane to which a core glycan and highly variable O-polysaccharide (O-PS) are attached (**Figure 1.13**) (Whitfield & Trent, 2014). While Lipid-A is highly conserved, the O-antigenic region is extremely inconsistent and adaptable; for example *E. coli* possess over 180 different O-antigens (Stenutz et al., 2006).

FIGURE 1.13

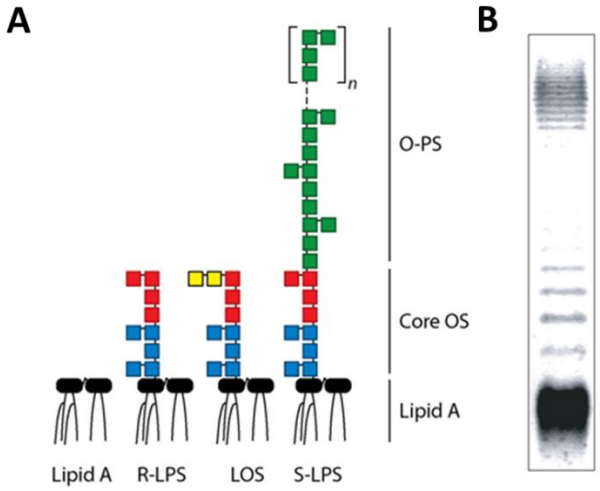


FIGURE 1.13. From Whitfield and Trent (2014). (**A**) Overview of LPS structure. (**B**) SDS-PAGE of *E. coli* LPS represents the large amount of heterogeneity within a species.

Inside the outer membrane is the periplasmic space. Here, the peptidoglycan layer provides structural support to gram-negative bacteria. Almost all prokaryotes synthesize peptidoglycan as a component of their cell walls, and the primary structure is the same. The sugar component is β-(1,4)-linked (GlcNAc) -β-(1,4)-N-acetyl-muramyl- *N*-acetylglucosamine (MurNAc-GlcNAc), which is crosslinked by tetrapeptides of L-alanine-D-isoglutamate-L-lysine (or diaminopimelic acid)-D-alanine. Peptidoglycan in gram-negative bacteria is conserved, and

consists of long glycan threads connected by flexible peptides (Matias et al., 2003). The glycan components can be up to 100 GlcNAc-β-(1,4)-MurNAc disaccharide units, averaging around 30 (Glauner et al., 1988; Harz et al., 1990). Each glycan thread is modified by a covalently linked peptide crosslinked to a peptide from another thread.

A growing number of *O*-linked and *N*-linked glycoproteins may also be present on the surface of certain gram-negative bacteria. These can include adhesins, flagella, and pilins.

Glycosyltransferases and examples in Gram-negative wall synthesis

GTs can either invert or retain the anomeric configuration of the sugar transferred, depending on their method of action (Lairson et al., 2008). Crystallization of GTs with their substrates shows that inverting type GTs act through a direct displacement mechanism, similar to an S_N2-like reaction. On the other hand, retaining GTs operate through a double displacement mechanism during which a glycosyl-enzyme intermediate occurs. Crystal structures also uncovered the two general-type folds GTs can adopt; namely GT-A and GT-B-type folds (Bourne & Henrissat, 2001). The small number of protein configurations common to all GTs studied suggests that most, if not all, have evolved from a small number of ancestral enzymes (Lairson et al., 2008). While both GT-A and GT-B have β/α/β Rossmann domains, GT-A types have an open β-sheet flanked by α-helices, and the fold is tightly associated but with separate domains of nucleotide sugar substrate and acceptor binding (Unligil & Rini, 2000) (**Figure 1.14**). Several GT-A types also have a distinct DXD motif for divalent metal cation binding (Breton et al., 2012; Wiggins & Munro, 1998). GT-B types also have two Rossman-like domains, but these are separated by a pocket (Vrielink et al., 1994).

Figure 1.14

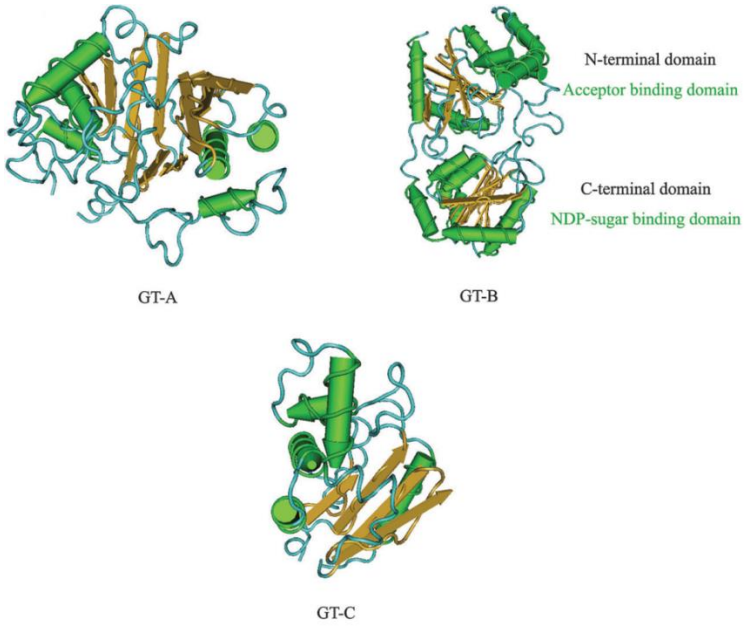


Figure 1.14. Adapted from Liang et al. (2015). Ribbon diagrams of the GT folds. GT-A has a single Rossman domain fold, while GT-B has two distinct domains. GT-C globular domain is shown.

A GT-C-type fold has also been described, and appears to be specialized to lipid-linked sugar donor substrates (Liu & Mushegian, 2003). All GT-C types described are inverting (Birch et al., 2009; Igura et al., 2008). These GTs commonly contain an N-terminal transmembrane (TM) domain and C-terminal globular domain involved in catalysis, and several resemble oligosaccharyltransferase (OST) (Liang et al., 2015; Lizak et al., 2011).

Several polysialyltransferases (PSTs) that synthesize CPSc have been characterized. PSTs have a GT-B fold, are in GT family 38 and act through an inverting mechanism. Characterization of PSTs from *E. coli* (McGowen et al., 2001) and *Neisseria meningitis* group B and C (Steenbergen & Vimr, 2003; Willis et al., 2008) demonstrate that all transfer sialic acid (NeuAc) from CMP-β-NeuAc onto a polysialic acid chain to make poly-α-(2,8/9)-NeuAc.

LPS biosynthesis of lipid-A occurs through the highly conserved Raetz pathway at the inner membrane. The process involves characterized enzymes; LpxA acylates UDP-GlcNAc to form UDP-3-*O*-acyl-GlcNAc, which is deacetylated by LpxC to make UDP-3-*O*-acyl-glucosamine (UDP-3-*O*-acyl-GlcN); LpxD acylates UDP-3-*O*-acyl-GlcN to form a UDP-2,3-diacyl-GlcN; LpxH cleaves a phosphoester to make 2,3-diacyl-GlcN-1-P (lipid-X) and UMP (Raetz & Whitfield, 2002). LpxB is the first GT in the pathway, and catalyzes condensation of one lipid-X molecule with UDP-2,3-diacyl-GlcN to form a 2,3-diacyl-GlcN-β-(1,6)- 2,3-diacyl-GlcN-P and UDP (Crowell et al., 1986). LpxB is an inverting family 19 GT with a B-type fold. The tetra-acyl disaccharide-1-phosphate is then phosphorylated by LpxK to make lipid-IV before addition of Kdo by WaaA. WaaA is an inverting GT family 30 member with a B-type fold and catalyzes addition of Kdo from CMP-Kdo α-(2,6) to lipid-IV (Schmidt et al., 2012). WaaA again transfers Kdo α-(2,4) to the lipid-IV-linked Kdo. Lipid-A synthesis is completed by two acylations through LpxL and LpxM (Raetz & Whitfield, 2002).

Lipid-A is further modified by GTs in the outer membrane to synthesize the core region. While variations are widely distributed and continue to be discovered, there is conservation in many of the best characterized systems, for example *E. coli*. One lipid-A modifying GT includes WaaC, an inverting GT in family 9 with a GT-B fold. WaaC catalyzes the addition of heptose from ADP-L-glycero-D-manno-heptose α-(1,6) to Kdo_1 of lipid-A (Grizot et al., 2006). After core region completion, O-PS is synthesized and added to the core in Wzy or ABC-transporter dependent pathways like previously described (see above), and a complete LPS molecule is translocated to the surface.

Peptidoglycan synthesis begins in the cell when MurNAc-pentapeptide is transferred from UDP- MurNAc-pentapeptide by a membrane-spanning protein MraY to form “lipid-I”

(Heydanek et al., 1969). The GT MurG then transfers GlcNAc to lipid-I from UDP-GlcNAc to form lipid-II. MurG is an inverting GT family 28 with a GT-B fold. (Hu et al., 2003). Lipid-II is then flipped into the periplasm by MurJ (Ruiz, 2008) and elongated by a membrane-bound peptidoglycan polymerase (PGP) GT, which adjoins lipid-II molecules. The GT head domain of PGP does not use a nucleotide sugar donor for catalysis and does not have a Leloir-type fold, but instead shares structural similarity to λ phage lysozyme (Lovering et al., 2007; Yuan et al., 2007).

An *N*-linked glycosylation pathway similar to that in eukaryotes has been established in the bacteria *C. jejuni*. This microbe has a protein glycosylation gene cluster (*pgl*) and transfers a heptasaccharide from undecaprenylpyrophosphate (Und-PP) to asparagine residues of target proteins with the conventional Asn-X-Ser/Thr sequon (Gupta & Brunak, 2002). PglB is the classical oligosaccharyl transferase (OST) with homology to Sttp3 of yeast (Szymanski & Wren, 2005). The diagnostic and conserved WWDYG sequence is required for activity *in vivo* (Kowarik et al., 2006a; Kowarik et al., 2006b). PglB has 10-12 predicted TM domains, and can modify completely folded proteins in the periplasm (Tabish et al., 2011). PglB is an inverting GT family 66 member with a GT-C fold (Liang et al., 2015).

O-linked protein glycosylation onto Ser or threonine (Thr) in gram-negative bacteria can occur through OST-dependent and OST-independent pathways. The OST-dependent mechanism is similar to that for N-linked glycosylation in that the Und-PP is glycosylated, flipped from the cytosol to periplasm and OST transfers the glycan to target protein (Hug & Feldman, 2011). This type of glycosylation occurs on flagellins, adhesins, and type IV pilins (Lindenthal & Elsinghorst, 1999; Logan, 2006). In *Pseudomonas aeruginosa* type IV pilins are α-(1,5)

arabinofuranosylated by a GT family 39 with an inverting mechanism and GT-C fold (Faridmoayer et al., 2007; Kus et al., 2008).

Implications for glycobiology

The presence of apiose in plant walls as the critical cross-linker to dimerize RG-II demonstrates its importance at the molecular level. The boron requirement for vascular plants (tracheophytes) proper growth and development and that 80-90% of a plant's boron is complexed in RG-II dimers reveals the biological relevance of this interesting branched-chain sugar.

It is a metabolically expensive task for an organism to make apiose. Green plants appear to have invested resources to make UDP-apiose/UDP-xylose synthase early on the evolutionary timeline, as UAS are evident in algae and ancestral land plants like mosses and liverworts. The complicated mechanism UAS used to decarboxylate UDP-GlcA and reduce the products UDP-Api and UDP-Xyl appears to be partially conserved in other extended SDRs like UXS and bacterial ArnA. Evidence of bacterial UAS suggests that its activity may have evolved from a bacterial ancestral enzyme common to UXS and ArnA.

Certain bacteria have developed glycosylation techniques that mimic eukaryotic methods. It is reasonable to hypothesize that gram-negatives that live closely with plants or animals as symbionts or pathogens use these techniques to evade host defense mechanisms. Work towards uncovering transferases that utilize unique UDP-sugars like UDP-apiose will reveal the importance of apiose in biology.

CHAPTER 2

FUNCTIONAL CHARACTERIZATION OF UDP-APIOSE SYNTHASES FROM BRYOPHYTES AND GREEN ALGAE PROVIDES INSIGHT INTO THE APPEARANCE OF APIOSE-CONTAINING GLYCANS DURING PLANT EVOLUTION[1]

[1]James Smith, Yiwen Yang, Shahar Levy, Oluwatoyin Oluwayemi Adelusi, Michael G. Hahn, Malcolm A. O'Neill, Maor Bar-Peled. 2016. Journal of Biological Chemistry. PMCID: PMC5076816

Reprinted here with permission of publisher

ABSTRACT

Apiose is a branched monosaccharide that is present in the cell wall pectic polysaccharides rhamnogalacturonan II and apiogalacturonan and in numerous plant secondary metabolites. These apiose-containing glycans are synthesized using UDP-apiose as the donor. UDP-apiose (UDP-Api) together with UDP-xylose is formed from UDP-glucuronic acid (UDP-GlcA) by UDP-Api synthase (UAS). It was hypothesized that the ability to form Api distinguishes vascular plants from the avascular plants and green algae. UAS from several dicotyledonous plants has been characterized, however it is not known if avascular plants or green algae produce this enzyme. Here we report the identification and functional characterization of UAS homologs from avascular plants (mosses, liverwort, and hornwort), from streptophyte green algae and from a monocot (duckweed). The recombinant UAS-homologs all form UDP-Api from UDP-GlcA, albeit in different amounts. Apiose was detected in aqueous methanolic extracts of these plants. Apiose was detected in duckweed cell walls but not in the walls of the avascular plants and algae. Over-expressing duckweed UAS in the moss *Physcomitrella patens* led to an increase in the amounts of aqueous methanolic-soluble apiose but did not result in discernible amounts of cell wall-associated apiose. Thus, bryophytes and algae likely lack the glycosyltransferase machinery required to synthesize apiose-containing cell wall glycans. Nevertheless, these plants may have the ability to form apiosylated secondary metabolites. Our data is the first to provide evidence that the ability to form apiose existed prior to the appearance of rhamnogalacturonan II and apiogalacturonan and provide new insights into the evolution of apiose-containing glycans.

INTRODUCTION

Apiose (3-C-[hydroxymethyl]-D-erythrofuranose, Api) is a branched-chain monosaccharide that is present in many plant secondary metabolites and in the primary cell walls of vascular plants (Picmanova & Moller, 2016). To date, only two cell wall polysaccharides, namely rhamnogalacturonan II (RG-II) and apiogalacturonan (ApiGalA), have been shown to contain Api (O'Neill et al., 1990). ApiGalAs may have a limited taxonomic distribution as they have only been detected in the cell walls of seagrasses and duckweeds (Gloaguen et al., 2010; Golovchenko et al., 2002; Hart & Kindel, 1970). By contrast, RG-II is present in the primary walls of all vascular plants examined to date (Matoh et al., 1996; O'Neill et al., 1990). Apiose links two side chains (A and B) to the galacturonan backbone of RG-II. The apiosyl residues of side chain A in two RG-II molecules are cross-linked by a borate diester to form the RG-II dimer (Ishii & Matsunaga, 1996; Kobayashi et al., 1996; Matoh et al., 1993; O'Neill et al., 1996). At least 90% of the RG-II in primary walls exists as a dimer (Ishii et al., 1999), and a reduction in the extent of RG-II cross-linking typically results in the formation of abnormal cell walls (Fleischer et al., 1999). Plants carrying mutations that affect Api metabolism as well as RG-II structure and cross-linking are dwarfed or fail to develop normally (Ahn et al., 2006; Dumont et al., 2014; O'Neill et al., 2001; Pabst et al., 2013).

Early studies of the biosynthesis of the plant flavonoid apiin (apigenin-7-[2-O-apiosylglucoside]) in parsley led Grisebach and Döbereiner (H. Grisebach & Döbereiner, 1964) to propose that UDP-apiose (UDP-Api) and UDP-xylose (UDP-Xyl) are formed from UDP-glucuronic acid (UDP-GlcA) by UDP-apiose synthase (UAS). Subsequent studies identified UDP-Api in parsley and in Lemna (Sandermann et al., 1968). It was then proposed that UDP-Api is the activated nucleotide sugar used by apiosyltransferases that catalyze the incorporation of

apiose into ApiGalA and into apiin (Hart & Kindel, 1970; Ortmann et al., 1972). No apiosyltransferase has been purified to homogeneity nor have the genes encoding this glycosyltransferase been identified.

We previously hypothesized that the enzymes responsible for synthesis of UDP-Api are present only in vascular plants (Bar-Peled et al., 2012). The recent availability of the sequenced genome of the moss *Physcomitrella patens* (Rensing et al., 2008) and publicly available transcriptomic data for other avascular land plants and for green algae produced by the 1,000 plants (1KP) project (Matasci et al., 2014) allowed us to re-examine this hypothesis.

Here we report the identification and functional characterization of UAS homologs from the monocot *Spirodela polyrhiza* (a duckweed), four mosses (*Physcomitrella patens, Dicranum scoparium, Hedwigia ciliata*, and *Sphagnum lescurii*), a liverwort (*Marchantia paleacea*), a hornwort (*Megaceros vincentianus*), and two green algae (*Mougeotia spp.* and *Netrium digitus*). Our results provide evidence that UDP-Api appeared prior to the appearance of wall-associated apiose and that bryophytes and green algae likely synthesize apiose-containing secondary metabolites but lack the biosynthetic machinery required for the synthesis of apiose-containing wall polysaccharides.

RESULTS

Identification of UDP-apiose Synthase Homologs in Avascular Plants and Green Algae

UAS-like homologs with >70% amino acid sequence identity to *Arabidopsis* AXS1/UAS1 were identified in monocots, mosses, liverworts, hornworts and streptophyte green algae (**Table 2.1**) using public ally available data from the 1,000 Plants (1KP) Project (Matasci

et al., 2014) and the Phytozome genomics portal. No UAS-like homologs were detected in the available transcriptomes of chlorophyte green algae.

An unrooted phylogenetic tree (**Figure 2.1**) was generated using the amino acid sequences of the UAS-like proteins (Fig. S1) together with other decarboxylases including *Arabidopsis* UDP-xylose synthase (UXS) and two bacterial enzymes: a bifunctional UDP-4-keto-pentose/UDP-xylose synthase (RsU4kpxs) from the plant pathogen *Ralstonia solanacearum* and the C-terminal portion of ArnA that has a UDP-glucuronic acid 4-oxidase-6-decarboxylase activity (Breazeale et al., 2005; Gu et al., 2010). The UASs cluster into a group (III, in **Figure 2.1**) comprised of several clades, which are distinct from the clades for UXS (I) and ArnA (II). These data suggest that UAS-like proteins first appeared in the streptophyte lineage of green plants. UAS, UXS, RsU4kpxs and ArnA are all decarboxylases that contain a conserved N-terminal Gly-*X*-*X*-Gly-*X*-*X*-Gly motif (Fig. S1) that is proposed to be involved in NAD^+ binding as well as the conserved Tyr-*X*-*X*-*X*-Lys motif with an upstream Ser that forms the catalytic site of the short-chain dehydrogenase/reductase (SDR) family (Harper & Bar-Peled, 2002; Kavanagh et al., 2008; Yin et al., 2011).

To date there are no reports that avascular plants or green algae produce apiose (Picmanova & Moller, 2016). Thus, we sought to ascertain if their UAS-like homologs, which have ~70% sequence identity to *Arabidopsis* AXS1/UAS1, are capable of converting UDP-GlcA to UDP-Api. To this end, UAS-like homologs from the monocot *S. polyrhiza* (SpUAS), the mosses *D. scoparium* (DsUAS), *H. ciliata* (HcUAS), *P. patens* (PpUAS) and *S. lescurii* (SlUAS), the liverwort *M. paleacea* (MpUAS), the hornwort *M. vincentianus* (MvUAS), and the green algae *Mougeotia spp.* (MougUAS) and *N. digitus* (NdUAS) were cloned and expressed in *E. coli* and then functionally characterized

In microbe Formation of UDP-apiose

The coding sequences of the selected UAS homologs were cloned into a modified pET28b *E. coli* expression vector (Yang et al., 2009). The UAS-containing plasmids or empty plasmid (negative control) were then individually transformed into *E. coli* together with a pCDFDuet plasmid containing the UDP-Glc dehydrogenase coding sequence (Broach et al., 2012) from *Bacillus thuringiensis* (BtbDH) to ensure the production of UDP-GlcA. Nucleotide sugar-containing extracts from the isopropyl β-D-thiogalactoside (IPTG)-induced *E. coli* cells were shown by hydrophilic interaction liquid chromatography with electrospray mass spectrometry (HILIC-ESI-MS/MS) to contain two product peaks eluting at 11.3 and 12.2 min (**Figure 2.2**). These peaks were not detected in the comparable extract of *E. coli* cells harboring the empty plasmid (**Figure 2.2**). The ESI mass spectra of both components contained an ion at *m/z* 535.00 (**Figure 2.2**), which corresponds to $[M-H]^-$ for a UDP-pentose. MS/MS analysis (**Figure 2.3**) of each product peak gave a fragment ion at *m/z* 323.00 that is consistent with $[UMP-H]^-$. The peak eluting at 12.2 min has the same elution time and MS fragmentation pattern as authentic UDP-Xyl. Proton NMR (^{1}H NMR) analyses confirmed that the UDP-pentose eluting at 11.3 min was UDP-Api. These data suggest that the UAS-like enzymes do synthesize UDP-Api.

Purified Recombinant UAS from S. polyrhiza, Mosses, a Liverwort, a Hornwort and Green Algae Convert UDP-GlcA to UDP-Api and UDP-Xyl

To obtain additional evidence that the monocot, avascular plant and green algal UASs form UDP-Api, the recombinant His_6-tagged proteins were solubilized from *E. coli* cells and

purified using nickel-affinity columns. Each recombinant UAS gave one major band on SDS-PAGE with a predicted mass of between 45 and 48 KDa (**Figure 2.4**). Each purified UAS was shown by HILIC-ESI-MS/MS to convert UDP-GlcA to two UDP-pentose products. MS/MS analysis (**Figure 2.5**) of these product peaks (11.3 and 12.2 min) also gave a fragment ion at *m/z* 323.00 that is consistent with $[UMP-H]^-$. Signals consistent with the presence of UDP-Api and UDP-Xyl were detected in all the ^{1}H NMR spectra when the recombinant enzyme assays were performed in deuterated buffer (**Figure 2.6**).

S. polyrhiza UAS was the most highly expressed protein and was thus selected for further characterization. Real time ^{1}H NMR spectroscopic analysis of the products formed when SpUAS reacts with UDP-GlcA (**Figure 2.7** and **Table 2.S1**) confirmed that UDP-Api is the first product formed. SpUAS produces UDP-Api and UDP-Xyl in a ratio of ~1.7: 1.0, which is similar to potato UAS (Guyett et al., 2009). Our studies with SpUAS also confirm that some of the UDP-Api is converted to the apiofuranosyl-1,2-cyclic phosphate during the *in vitro* reaction (**Figure 2.7**). No degradation of UDP-Xyl is discernible over the course of the reaction.

Real time NMR based-assays provides the opportunity to detect transient intermediates (Guyett et al., 2009). Our real time ^{1}H NMR data using recombinant SpUAS confirms that UDP-4-keto-xylose is an intermediate formed during the conversion of UDP-GlcA to UDP-Api, (**Figure 2.7**). Grisebach (H. Grisebach & Döbereiner, 1964) and subsequently Choi and colleagues (Choi et al., 2012) proposed that during UAS enzymatic catalysis a ring contraction step occurs through a retro-aldol mechanism . We detected no signals indicative of the formation of the proposed ene-diol intermediate. Nonetheless, if this intermediate is formed it may exist for such a short time or remain secured in the catalytic site of the enzyme and thus be “invisible” to ^{1}H NMR spectroscopy.

The recombinant enzyme, SpUAS, is most active in 50 mM Tris-HCl, pH 7.5 - 8.1, at temperatures between 37 and 42 °C (**Figure 2.8A-B**) and exists in solution as a dimer with a predicted size of 93 KDa (**Figure 2.8C**). SpUAS has a Km of 237 µM and Kcat/Km of 1.159 µM s^{-1}, while recombinant *Arabidopsis* AXS1/UAS1 has a reported Km of 7 µM and Kcat/Km of 0.043 µM s^{-1} (Molhoj et al., 2003). Previous studies have shown that UAS is inhibited by nucleotides and nucleotide sugars, including UDP-GalA (Molhoj et al., 2003). Under our assay conditions UDP-Xyl and NAD(P)H inhibited UAS activity by 20% and 36%, respectively, whereas UDP-GalA reduced activity by 82%. UDP-GalA is not a substrate for the recombinant UAS. Nonetheless, we cannot exclude the possibility that UDP-GalA competes with UDP-GlcA for substrate binding and thereby regulates SpUAS activity *in vivo*.

Apiose is Discernible in Aqueous Methanolic Extracts but not in the Cell Walls of Mosses, Liverworts and Green Algae

To explore the nature of apiose-containing molecules in the avascular plants and algae, living cultures of the mosses *P. patens*, *Dicranum*, *Polytrichum*, and *Sphagnum*; the liverworts *Conocephalum* and *Marchantia*; and the algae *Mougeotia* and *Netrium* were obtained (only the genera was provided by the supplier for those plants without species designation). Fresh tissue was extracted with a series of aqueous (aq.) and non-aqueous solvents (see Experimental Procedures). The residue remaining after these extractions is defined here as the cell wall. Apiose was detected after acid hydrolysis and GC-MS analyses of alditol-acetate derivatives in the aq. methanol-acetonitrile (methanolic) soluble fractions of all four mosses, both liverworts and *Mougeotia* (**Table 2.2**). No additional apiose was detected in the subsequent solvent extracts (Fractions II-IX, see Experimental Procedures) or in the cell walls generated from these plants

(**Table 2.3**). No Api was detected in any fractions from *Netrium*. By contrast, Api was abundant in the methanolic fraction and cell walls of *S. polyrhiza*. Such results are not unexpected, as the walls of this duckweed are known to contain large amounts ApiGalA (Hart & Kindel, 1970; Longland et al., 1989).

Over-expression of SpUAS in P. patens and Detection of UDP-apiose

We detected small amounts of Api in the methanolic extract of *P. patens* but no discernible amounts of Api were present in its cell wall. Moreover, no UDP-Api was detected in the nucleotide sugar-containing extracts of *P. patens* wild-type gametophyte tissue (**Figure 2.9B**) or any of the other avascular plants and green algae. Thus, we wondered if UDP-Api is being degraded or rapidly metabolized. We hypothesized that increasing the level of UDP-Api in *P. patens* gametophytes would generate a pool of activated Api that was sufficient to allow this to be incorporated into glycans of the cell wall. To this end, SpUAS was over-expressed in *P. patens*. The SpUAS transcript was detected in five independent transgenic lines (**Figure 2.9A**). None of the transformed lines had a visibly altered growth phenotype. The over-expressing lines readily formed detectable amounts of UDP-Api (**Figure 2.9B**). Increased amounts of Api were also present in the methanolic fractions from the over-expressing lines (**Figure 2.9C**). None of the lines contained discernible amounts of Api in their cell walls. Thus, we conclude that UDP-Api is not appreciably degraded and is likely not a limiting factor for the incorporation of Api residues into walls of *P. patens* gametophyte.

DISCUSSION

Our study is the first to identify functional genes encoding UDP-apiose synthase in avascular plants and green algae. The apiose in these plants was detected in aqueous methanolic extracts. However, we found no discernible amounts of apiose-containing glycans in the walls of any of the avascular plants or the algae. Thus, the Api detected in these plants is likely to be associated with a secondary metabolite. This contrasts with vascular plants, where apiose is present in the cell wall polysaccharides RG-II and ApiGalA and in secondary metabolites (Beck & Hopf, 1990; Darvill et al., 1978; Picmanova & Moller, 2016; Watson & Orenstein, 1975).

Plants, many animals, fungi, Bacteria and Archaea produce enzymes (UXS) that convert UDP-GlcA to UDP-Xyl (Gotting et al., 2000; Gu et al., 2011; Harper & Bar-Peled, 2002; Kobayashi et al., 2002; Kuhn et al., 2001), whereas UAS forms both UDP-Api and UDP-Xyl in a ratio of ~2:1. It is not known if *UXS* or *UAS* is the ancestral gene. To date, no *UAS* genes have been identified in prokaryotes, whereas *UXS* is present in many Bacteria and Archaea. Thus, it is likely that *UAS* first appeared in the plant kingdom, possibly from *UXS* (Gu et al., 2011).

The mechanism whereby UAS converts UDP-GlcA into two different UDP-sugars is not known. Plant UXS and UAS and the bacterial ArnA are all enzymes, which decarboxylate UDP-GlcA via 4,6-dehydration and a UDP-4-keto-pentose intermediate. X-ray analyses of UXS and ArnA identified domains (Fig. S1) involved in catalysis and cofactor binding (Polizzi et al., 2012; Walsh et al., 2015). Protein sequence alignment (Fig. S1), shows that residues implicated in nucleotide sugar binding and catalysis are conserved among UXSs from diverse organisms but are different in the UASs. Additionally, regions of unique amino acid insertions are present in UAS but not in UXS (Fig. S1, Regions 1-5) or ArnA (Regions 2-4). These distinct regions may facilitate or necessitate the subsequent ring cleavage and rearrangement. Site-directed

mutagenesis of these regions in combination with crystallographic and enzymatic activity studies are required to elucidate the mechanism of UAS activity and determine if it occurs via enediol intermediate.

Plants synthesize numerous secondary metabolites. In vascular plants at least 1200 of these metabolites have been reported to contain an apiose (Picmanova & Moller, 2016). These apiose-containing compounds include flavonoids, terpenoids, and cyanogenic glucosides (Beck & Hopf, 1990; H. Grisebach, 1980; Picmanova & Moller, 2016; Watson & Orenstein, 1975). Apiosylated secondary metabolites may protect a plant from pathogens and herbivores. Several secondary metabolites, including flavonoids are known to provide such defense as well as protection against UV radiation and oxidative stress (Eckey-Kaltenbach et al., 1993; Neilson et al., 2013; Treutter, 2005). These metabolites are normally water insoluble. The addition of a glycose such as apiose to these compounds would enhance their water solubility and perhaps facilitate their transport within the plant. Interestingly, *P. patens* has been reported to lack functional borate exporters (Wakuta et al., 2015), thus the possibility cannot be excluded that apiosides sequester borate, which may become toxic if accumulated to high amounts (Camacho-Cristobal et al., 2008). The ability to form apiose has broadened the library of secondary metabolites available to plants, including avascular bryophytes and algae. Mechanisms to incorporate apiose into the cell wall polysaccharides RG-II and ApiGalA must also have developed during the evolution of vascular plants from their avascular ancestors. Determining if the *P. patens* apiosyltransferases that utilize UDP-Api as a donor have common motifs with the apiosyltransferases involved in secondary metabolite and pectin biosynthesis in vascular plants will provide insight into the evolutionary origins of Api-containing glycans.

The cell walls of several green algae and bryophytes are known to contain homogalacturonan (HGA) (Baylson et al., 2001; Popper & Fry, 2003; Roberts et al., 2012). In the cell walls of vascular plants, apiose is attached to the HGA backbones of RG-II and ApiGalA. Since we could not detect wall-bound apiose, even in *P. patens* lines over-expressing SpUAS, it is probable that the apiosyltransferases required for the synthesis of RG-II and ApiGalA do not exist in bryophytes or green algae. It is notable that phylogenetic analyses of pectin-related gene families in *P. patens* suggest that this moss lacks homologs of vascular plant xylogalacturonan xylosyltransferases and rhamnogalacturonan I arabinosyltransferases (McCarthy et al., 2014). Thus, expanding the repertoire of cell wall pectin structures may have accompanied the transition from avascular to vascular plants.

The identification and functional characterization of UASs from green algae and bryophytes provides a valuable tool to study the role of apiose and apiosylated metabolites in these organisms. The use of comparative genomics and transcript analyses will reveal glycosyltransferases responsible for addition of apiose to secondary metabolites and to wall polysaccharides.

EXPERIMENTAL PROCEDURES

Plant Material and Growth Conditions

Living cultures of the mosses *Dicranum*, *Sphagnum*, *Polytrichum*, the liverworts *Marchantia* and *Conocephalum*, and the algae *Mougeotia* and *Netrium* were obtained from Carolina Biological (Burlington, NC), harvested and kept at -80 °C. *Physcomitrella patens* (Var Gransden) was maintained in liquid routine basal medium [1 L BCDAT medium = 10 ml each; medium B (0.1 mM $MgSO_4$) + medium C (1.84 mM KH_2PO_4 pH 6.5) + medium D (1 M KNO_3

and 4.5 mM $FeSO_4$) + AT (0.5 M ammonium tartrate); with 1 ml trace elements (0.22 mM $CuSO_4$, 10 mM H_3BO_3, 0.23 mM $CoCl_2$, 0.1 mM Na_2MoO_4, 0.19 mM $ZnSO_4$, 2 mM $MnCl_2$, 0.17 mM KI) and 1ml 1 M $CaCl_2$; distilled deionized H_2O added to 1 L]. Protonema were grown at 22 °C on cellophane-covered BCDAT agar plates (Nishiyama et al., 2000) in a controlled environment growth chamber (Conviron, Manitoba, Canada) with a 16/8 hour photoperiod. *Spirodela polyrhiza* was obtained from Joachim Messing (Waksman Institute of Microbiology, Rutgers University) and maintained on 0.8% agar (w/v) containing Schenk and Hildebrandt basal salts (1.6 g/L) and 0.5%sucrose (w/v), pH 5.8, in a Conviron growth chamber with a 14/10 hour photoperiod at 19 °C/15 °C. Between 6 and 8 fronds of *S. polyrhiza* were transferred to sterile liquid medium (50 ml) containing 1% (w/v) sucrose in 250 ml Erlenmeyer flasks and grown under a 14/10 hour photoperiod at 24 °C/20 °C.

Identification and cloning of DsUAS, HcUAS, MougUAS, MpUAS, MvUAS, NdUAS, PpUAS, SlUAS, and SpUAS

The moss, liverwort, and hornwort transcriptomes in the 1,000 Plants (1KP) Project and annotated monocot proteins in the Phytozome genomic portal were probed for homologs to the amino acid sequence of *Arabidopsis* AXS1/UAS1 by BLAST analyses. Analysis of the top-hits revealed proteins with high sequence identity to *Arabidopsis* AXS1/UAS1, including several green algae and moss proteins with >70% amino acid sequence identity, liverwort proteins from *Treubia lacunosa* (83% identity) and *Marchantia paleacea* (two incomplete; 65 and 82% identity), and hornwort proteins from *Nothoceros aenigmatus* and *Megaceros vincentianus* with 77 and 76% identity, respectively. The top hit for *S. polyrhiza* (locus Spipo0G0011100, annotated as “bifunctional polymyxin resistance ArnA protein”) has an amino acid sequence

identity of 85% to *Arabidopsis* AXS1/UAS1. The nucleotide sequence corresponding to the *S. polyrhiza* protein was used for primer design and cloning.

S. polyrhiza RNA was isolated from 10-day-old fronds. Fronds were collected, vacuum filtered over nylon mesh, rinsed with deionized water, blotted dry and then ground to a fine powder in liquid nitrogen using a mortar and pestle. Total RNA was extracted using a Qiagen RNeasy mini kit with an on-the-column DNAse treatment to eliminate genomic DNA contamination. RNA (0.5 µg) was then reverse transcribed with an oligo(dT) primer using SuperScript III reverse transcriptase (Life Technologies; Carlsbad, CA). A portion of the reverse transcriptase (RT) reaction (2 µl), dNTP's, 1 unit of Phusion® high-fidelity DNA polymerase (New England Biolabs; Ipswich, MA) with 0.2 µM of each forward and reverse primers (IDT; Coralville, IA; **Table 2.S2**) were used to amplify the SpUAS with the following thermal cycler conditions: one 98 °C denaturation cycle for 30 s followed by 25 cycles (each of 8-s denaturation at 98 °C; 25-s annealing at 60 °C; 30-s elongation at 72 °C), and finally termination at 4 °C. The PCR product was directly cloned into the *E. coli* expression vector pET28b (Novagen; Darmstadt, Germany) modified to contain an N-terminal His_6 tag followed by a TEV cleavage site (Yang et al., 2009).

RNA was extracted from wild-type and transformed *P. patens* using 100 mg of 2-week-old gametophyte tissue. Tissue was harvested and immediately frozen in liquid nitrogen and RNA extracted and reversed transcribed with oligo-dT. No PpUAS transcript was detected on 1% (w/v) agarose gel even after a second round of PCR amplification using 0.2 µM of each forward and reverse primer (**Table 2.S2**) and an annealing temperature of 56 °C. The recently released *P. patens* Electronic Fluorescent Pictograph browser (Ortiz-Ramirez et al., 2016; Winter et al., 2007) indicates that PpUAS (gene ID Pp1s379_19V6.1) is only expressed at substantial

levels in stage S3 sporophyte and archegonia. Thus, a synthetic ORF gene corresponding to PpUAS was obtained (GenScript; Piscataway, New Jersey, USA).

For transcript analysis of *P. patens* SpUAS-overexpressing lines, SpUAS transcript was amplified from 30 cycles using SpUAS_F and SpUAS_Mid_R primers (**Table 2.S2**). *P. patens* elongation factor 1-α (PpEf1a) was similarly amplified using forward and reverse primers (**Table 2.S2**). PCR products were run on a 1% (w/v) agarose gel.

Genes corresponding to the nucleotide sequences of *Dicranum scoparium, Hedwigia ciliata, Mougeotia spp., Marchantia paleacea, Megaceros vincentianus*, *Netrium digitus*, and *Sphagnum lescurii* UASs were obtained from GenScript, due to the lack of axenic lines. These ORF's were cloned into the modified pET28b expression vector (Yang et al., 2009), using forward and reverse primers (see **Table 2.S2**). The predicted transcript for *D. scoparium* lacked 8 amino acids at the N-terminal region based on sequence alignment with PpUAS. To obtain the entire *D. scoparium* ORF, its nucleotide sequence was extended to include the sequence corresponding to the amino acid sequence MTARVLND at the N-terminus based on the N-terminal UAS sequence of the moss *H. ciliata*. For cloning the UAS ORF of *M. paleacea*, the middle portion of the predicted sequence was constructed to include the amino acid sequence RPLDTIYSNFIDALPVVRYCTDNNKRLIHFSTCEVYGKTIGCFLPNDSPLRKD based on sequence alignment with a UAS-like homolog in the liverwort *Treubia lacunosa*.

Following cloning of the individual UAS genes, the plasmids were sequence verified (Georgia Genomics Facility; Athens, GA) and termed, pET28b-TEV-DsUAS.2, pET28b-TEV-HcUAS.3, pET28b-TEV-MougUAS.1, pET28b-TEV-MpUAS2.1, pET28b-TEV-MvUAS.6, pET28b-TEV-NdUAS.2, pET28b-TEV-PpUAS.1, pET28b-TEV-SlUAS.1, and pET28b-TEV-SpUAS.1. Their amino acid sequences were deposited in GenBank™ (accession numbers

KX344124, KX344125, KX344126, KX344127, KX344128, KX344129, KX344130, KX344131, and KX344132).

Transformation of P. patens

The ORF SpUAS was amplified by PCR using forward and reverse primers (**Table 2.S2**), directly cloned into pENTR™/SD/D-TOPO® and Gateway® cloned into pTHUbiGate, a *P. patens* expression vector that has homologous recombination sites at *P. patens* locus 108 (Vidali et al., 2007), with LR Clonase II (Life Technologies). The expression of SpUAS in pTHUbiGate is driven by a ubiquitin promoter. For plant transformation, the binary plasmid, pTHUbiGate-SpUAS (50 µg), was linearized with BsaAI (New England Biolabs) and then precipitated by the addition of ethanol. The precipitate was dissolved in sterile water (1 µg/µl) and then used to transform *P. patens* protoplasts (Nishiyama et al., 2000).

P. patens protoplasts were prepared by treating protonema suspended in aq. 8.5% (w/v) mannitol for 60 min at room temperature with 2% (w/v) Driselase (Sigma-Aldrich; St. Louis, MO) and shaking at 60 rpm. The suspension was filtered through 40 µm nylon mesh and the released protoplasts suspended in 8.5% (w/v) mannitol. The suspension was centrifuged (250*g*, 5 min, 22 °C), the supernatant discarded and the pellet re-suspended in 8.5% mannitol. The filtration and re-suspension were repeated 2 times. The number of protoplasts obtained was then determined using a hemocytometer. Protoplasts were centrifuged, and the protoplasts suspended to a density of approximately 2 x 10^6 protoplasts/ml in 0.5 M mannitol containing 15 mM $MgCl_2$ and 0.1% (w/v) MES pH 5.6 (Nishiyama et al., 2000). 0.3 ml of the protoplast suspension and 0.3 ml of PEG 8000 (Sigma-Aldrich) were added to 15 µl of BsaA1-linearized plasmid DNA, thoroughly mixed and then kept at room temperature for 10 min. The mixture was heat shocked

for 3 min at 45 °C and immediately cooled to room temperature in a water bath and kept for 10 min. The suspension was then centrifuged, the supernatant discarded, and the protoplasts suspended in 5 ml liquid protoplast regeneration medium [PRML = BCDAT plus 8% (w/v) mannitol and 10 mM $CaCl_2$]. A portion of this protoplast suspension (1.6 ml) was spread on cellophane layered over protoplast regeneration medium bottom layer [PRMB = BCDAT plus 6% (w/v) mannitol, 10 mM $CaCl_2$ and 1% (w/v) agar] (Nishiyama et al., 2000). The plates were kept for 5 days in a growth chamber at 22 °C.

The cellophane was then transferred onto BCDAT medium containing ampicillin (100 μg/ml) and kept for a further 7 days. The cellophane was then transferred onto BCDAT medium lacking antibiotic. After 7 days the cellophane was transferred to BCDAT medium with ampicillin and kept for an additional 7 days to obtain stable transformants. Transformants were verified by PCR of locus 108 using the appropriate forward and reverse primers (**Table 2.S2**) and transcript analysis.

In microbe Assays

NDP-sugars from *E. coli* harboring the expression plasmids were harvested as described (Yang et al., 2012). BL21-derived *E. coli* cells (3 ml) dually transformed with pCDFDuet-BtbDH and either pET28b-TEV-DsUAS.2, pET28b-TEV-HcUAS.3, pET28b-TEV-MougUAS.1, pET28b-TEV-MpUAS2.1, pET28b-TEV-MvUAS.6, pET28b-TEV-NdUAS.2, pET28b-TEV-PpUAS.1, pET28b-TEV-SlUAS.1, pET28b-TEV-SpUAS.1, or empty pET28b vector control were grown in LB medium [1.0% (w/v) Bacto tryptone, 0.5% (w/v) Bacto yeast extract, and 1.0% (w/v) NaCl] supplemented with chloramphenicol (35 μg/ml), kanamycin (50 μg/ml) and spectinomycin (25 μg/ml) at 37 °C and 250 rpm to an A_{600nm} of 0.6. Isopropyl β-

D-thiogalactoside (IPTG, 0.5 mM) was then added, and the cells grown at 30 °C for 4 h. Each culture was centrifuged (8,000*g*, 3 min, 4 °C); the cell pellets were washed with 4 volumes PBS (10 mM NaH_2PO_4 pH 7.5, 150 mM NaCl), transferred to Eppendorf tubes, centrifuged (18,000*g*, 1 min, 22 °C), and re-suspended in either 1 volume ice-cold deionized distilled water (DDW) for hydrophilic interaction liquid chromatography–electrospray ionization–tandem mass spectrometry (HILIC-ESI-MS/MS) analysis or deuterium oxide (D_2O) for ^{1}H NMR analysis. 10 volumes of cold chloroform/methanol (1:1, v/v) or chloroform-*d*/methanol-d_4 (Cambridge Isotopes Laboratories; Tewksbury, MA) were added, and the samples mixed for 20 min at 4 °C. The suspensions were centrifuged (18,000*g*, 5 min, 22 °C), and the upper aqueous phases collected and re-centrifuged. Portions of the aqueous phase were analyzed by HILIC-ESI-MS/MS and by ^{1}H NMR spectroscopy.

Expression and Purification

BL21-derived *E. coli* cells were transformed with pET28b-TEV-DsUAS.2, pET28b-TEV-HcUAS.3, pET28b-TEV-MougUAS.1, pET28b-TEV-MpUAS2.1, pET28b-TEV-MvUAS.6, pET28b-TEV-NdUAS.2, pET28b-TEV-PpUAS.1, pET28b-TEV-SlUAS.1, pET28b-TEV-SpUAS.1, or the empty vector control. The cells were grown at 37 °C and 250 rpm for 16 h in LB medium (20 ml) containing kanamycin (50 µg/ml) and chloramphenicol (35 µg/ml). A portion (5 ml) of the culture was transferred to fresh LB plus antibiotics (245 ml) and grown under the same conditions until its A_{600nm} was 0.8. IPTG was added to 0.5 mM to induce expression of UAS, and the culture then grown for an additional 4 h at 30 °C and 250 rpm. The induced cultures were cooled on ice and centrifuged (6,000*g*, 10 min, 4 °C). The cell pellet was suspended in lysis buffer [50 mM Tris-HCl pH 7.6, 10% (v/v) glycerol, 1 mM

ethylenediaminetetraacetic acid, 5 mM dithiothreitol 0.5 mM phenylmethylsulfonyl fluoride, 10 ml]. The cells were ruptured by sonication and proteins were then isolated after centrifugation as described (Yang et al., 2009). The final soluble protein fraction (Fraction S20) was collected and kept on ice prior to purification.

The His_6-tagged proteins were purified using fast-flow Ni-Sepharose (GE Healthcare, 2 ml resin packed in a 15 x 1 cm polypropylene column). Columns were washed and equilibrated with 50 mM sodium phosphate, pH 8.0, containing 300 mM NaCl, and then Fraction S20 was added. Bound His_6-tagged proteins were eluted with the same buffer containing increasing concentrations of imidazole (10 - 250 mM). The active enzymes were eluted in 250 mM imidazole and then dialyzed (6,000 – 8,000 molecular weight cut-off, Spectrum Laboratories, Inc) at 4 °C three times for a total of 2 h against 50 mM Tris-HCl pH 7.6, containing 0.15 M NaCl, 10% (v/v) glycerol, 1 mM DTT, and 10 µM NAD^+. The dialysates were divided into 150 µl aliquots, flash-frozen in liquid nitrogen and stored at -80 °C. Aliquots of purified protein were assayed for activity and run on SDS-PAGE.

SDS-PAGE was performed with 12% (w/w) polyacrylamide gels. Proteins were stained with 0.1% (w/v) Coomassie Brilliant Blue R-250 in aq. 20% methanol (MeOH) containing 7% (v/v) acetic acid and de-stained with aq. 20% methanol containing 7% (v/v) acetic acid. Protein concentrations were determined with the Bradford reagent (Bradford, 1976) using bovine serum albumin (BSA) as standard. The molecular mass of active recombinant SpUAS was estimated by size-exclusion chromatography eluted over a Superdex-75 100/300 GL column with dialysis buffer as eluent (Gu et al., 2010) and calibrating with proteins of known molecular mass (Bio-Rad).

Recombinant UAS Enzyme Assays

Unless otherwise indicated, reactions were performed in 50 mM Tris-HCl pH 7.9, (50 µl) containing 1 mM NAD^+, 1 mM UDP-GlcA, and 10 µg of purified protein. The mixtures were kept at 37 °C for up to 45 min and the reactions terminated by placing the tubes in boiling water for 2 min followed by the addition of an equal volume of chloroform. The suspensions were vortexed and centrifuged (12,000*g*, 5 min, 22 °C), and the aqueous phase analyzed for nucleotide sugars. ^1H NMR assays (180 µl) were performed in deuterium oxide (D_2O) using 30 µg purified protein.

Characterization of Recombinant SpUAS

SpUAS activity was assayed in different buffers, at different temperatures and with various additives and nucleotide sugars. For pH studies, purified recombinant SpUAS (10 µg) was added to standard reactions (50 µl) containing various pH buffers (100 mM) and 1 mM NAD^+ and 1 mM UDP-GlcA and kept at 37 °C for 30 min. Inhibition assays were performed by first supplementing the standard reaction mixtures with various nucleotides and nucleotide sugars, addition of purified protein and incubation. The amounts of reactants and products were determined by UV spectroscopy and used to calculate enzyme activity as follows. The products from each recombinant enzyme assay were chromatographed over a Q-15 anion exchange column (200 x 1 mm, Amersham Biosciences; Buckinghamshire, England) by elution with a linear gradient (5 mM to 0.6 M) of ammonium formate over 25 min at a flow rate of 0.25 ml/min using an Agilent (Santa Clara, CA) 1100 Series HPLC equipped with an G1313A auto-sampler, a G1315B diode array detector, and ChemStation software. Nucleotides and nucleotide sugars were detected by their A_{261nm} (for UDP-sugars) and A_{259nm} (for NAD^+). The concentrations of

reactants and products were determined by comparison of their peak areas to a calibration curve of standard UDP-GlcA (Gu et al., 2010).

Selected kinetic parameters of recombinant SpUAS (10 µg) were determined by varying the concentrations of UDP-GlcA in 50 µl reactions consisting of 1 mM NAD^+ in 50 mM Tris-HCl pH 7.9. Reactions were kept for 7 min at 37 °C quenched with an equal volume of chloroform and then vortex mixed. The reaction products in the aqueous phase were separated using a Q-15 anion exchange column as described above and reaction rates calculated from the depletion of the UDP-GlcA signal integral normalized to the NAD^+ signal integral. Values from three independent replicates were used to generate a non-linear regression plot and resultant data using GraphPad Prism Version 6.04.

HILIC-ESI-MS/MS

ESI-MS/MS analysis was performed on a Shimadzu (Kyoto, Japan) LC-MS-IT-TOF operating in the negative ion mode. Plant nucleotide sugar extracts and in microbe and recombinant enzyme assay products were mixed with 2/3 volume aq. 95% acetonitrile (ACN) containing 25 mM ammonium acetate and an aliquot (10 - 20 µl) chromatographed over an Accucore amide-HILIC column (150 x 4.6 mm; Thermo Fisher Scientific; Waltham, MA), eluted at 0.4 ml min^{-1} with a linear gradient of aq. 75% (v/v) acetonitrile containing 40 mM ammonium acetate, pH 4.4, to 50% (v/v) acetonitrile containing 40 mM ammonium acetate, pH 4.4, over 35 min using a Shimadzu LC-30AD HPLC. Mass spectra (mass range 100-2,000 *m/z*) were collected every 1.3 sec for 30 minutes. Second stage MS/MS data was collected by collision-induced dissociation (CID) with a collision energy of 35% and a nebulizing helium gas flow of 1.5 ml min^{-1} (Hwang et al., 2014).

Nucleotide sugar Extraction of Plant Tissue

Fresh tissues (100 mg wet wt) were ground in liquid nitrogen using a mortar and pestle, transferred to Eppendorf tubes and ACN:MeOH:H_2O (40:40:20, v/v/v, 1 ml) then added. The tubes were vortexed and rotated for 20 min at 4 °C. The tubes were centrifuged (18,000*g*, 5 min, 22 °C) and the supernatant transferred to a clean tube. The solutions were concentrated to ~50% of their initial volume using a stream of nitrogen gas and portions (800 µl) analyzed by HILIC-ESI-MS/MS as stated above.

Real time ^{1}H NMR Enzyme Assays

All spectra were obtained using a Varian Inova 600 MHz spectrometer equipped with a 3 mm cryogenic probe. Continuous ^{1}H NMR spectroscopic monitoring of reactions (180 µl volume) were carried out at 37 °C in a mixture of D_2O/H_2O (9:1, v/v) containing 0.83 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, internal reference), 50 mM Tris-HCl, pH 7.9, 1 mM UDP-GlcA, 1 mM NAD^+ and purified recombinant enzyme (30 µg). One-dimensional ^{1}H NMR spectra with the water resonance signal pre-saturated were collected 5 minutes post addition of enzyme in order to optimize spectrometer settings, and then spectra were continuously averaged every 2 min for up to 9 h. All chemical shifts are referenced to DSS at 0.00 ppm (Guyett et al., 2009).

Fractionation and cell wall polysaccharide extraction

Fresh plant tissue (0.1 to 1 g) was suspended in cold ACN:MeOH:H_2O (40:40:20, v/v/v, 10 vol) and ground in a mortar and pestle on ice. The suspension was transferred to a 15 ml

falcon tube and kept for 30 min at 55 °C. The suspension was centrifuged (3,000*g*, 5 min, 22 °C) and the supernatant (Fraction I, "methanolic extract") saved. Aq. 50% (v/v) EtOH (10 vol) was then added to the pellet, the suspension vortexed and kept for 30 min at 55 °C. The suspension was centrifuged and the supernatant (Fraction II) saved. The procedure was repeated using aq. 80% and 95% (v/v) EtOH to give soluble Fractions III and IV. The pellet was then suspended in water (5 vol) and kept for 30 min 55 °C. An equal volume of ethyl acetate was added and the suspension vortexed and kept for 30 min at 55 °C. The suspension was centrifuged and the top layer (Fraction V) was saved. The lower aqueous layer (Fraction VI) was transferred to a clean borosilicate tube. The pellet was suspended in $CHCl_3$:MeOH (1:1, v/v, 5 vol) and vortexed. After centrifugation, the top aqueous layer (VII) was saved. The bottom organic layer (Fraction VIII) was also saved. The pellet was suspended in acetone (5 vol) and after centrifugation the supernatant (Fraction IX) was save. The final pellet, referred to as cell wall, was allowed to air dry overnight in a fume hood.

Glycosyl Residue Composition Analysis

The aqueous and organic solvent extracts or cell wall (~1 mg) were supplemented with myo-Inositol (10 µl of 5 mM solution) as an internal standard, evaporated to dryness at room temperature using a stream of filtered air (REACTIVAP III, Thermo Fisher) and then hydrolyzed for 2 h at 120 °C with 2M TFA (1 ml). TFA was removed by evaporation under a stream of filtered air (40 °C) and the residue washed with isopropanol (3 x 500 µl). The released monosaccharides were then converted into their corresponding alditol-acetate derivatives according to York *et al* (York et al., 1986), and the final residue dissolved in acetone (100 µl).

The alditol-acetate derivatives were analyzed by gas-liquid chromatography (GLC, Agilent 7890A) equipped with a mass selective detector (EI-MS, Agilent 5975C). The sample (1 µl) was injected in the splitless mode using an Agilent 7693 auto-sampler onto a Restek RTx-2330 fused silica column (0.25 mm I.D x 30 m, 0.2 µm film thickness) with helium as gas carrier at a flow rate of 1.1 ml min^{-1}. The oven temperature was held at 80 °C for 2 min, followed by an increase of 30 °C min^{-1} to 170 °C, then at 4 °C min^{-1} to 235 °C and a hold at 235 °C for 20 min. The column was then kept at 250 °C for 7 min, cooled to 80 °C and kept at 80 °C for 1 min prior to the next injection. The injection port and the transfer line to the EI-MS were kept at 250 °C. Alditol-acetate derivatives of authentic apiose, rhamnose, fucose, ribose, arabinose, xylose, mannose, glucose, and galactose (50 µg each) were prepared under the same conditions as samples. Monosaccharides were identified based on their retention times and their EI mass-spectra. Peak areas, obtained from the total ion chromatogram, were exported to Microsoft Excel and normalized using the amount of sample and the area of the internal standard. The amounts of each monosaccharide in a sample were calculated using the response factors of the monosaccharide standards.

TABLES AND FIGURES

TABLE 2.1

UAS-like homologs identified in green algae and avascular plants.

The amino acid sequence for *Arabidopsis* AXS1/UAS1 was used as the probe to identify sequences in the 1KP and Phytozome databases. Percent sequence identity (% ID) is shown.

Organism	% ID
Green Algae	
Mesotaenium endlicherianum	75%
Roya obtusa	75%
Mougeotia spp.	73%
Netrium digitus	72%
Cylindrocystis brebissonii-M2213	72%
Cylindrocystis cushleckae	73%
Penium margaritaceum	78%
Mesostigma viride	61%
Hornworts	
Nothoceros aenigmaticus	77%
Megaceros vincentianus	76%
Liverworts	
Treubia lacunosa	83%
Marchantia paleacea[1]	65%, 82%
Mosses	
Niphotrichum elongatum	80%
Philonotis fontana	80%
Dicranum scoparium	81%
Hedwigia ciliata	78%
Sphagnum lescurii	79%
Sphagnum fallax	79%
Physcomitrella patens	74%

[1] *2 partial sequences obtained*

TABLE 2.2

Glycosyl-residue compositions of the aqueous acetonitrile-methanol soluble material from green algae, avascular plants and *S. polyrhiza*.

The glycosyl residue composition (mol %) was determined by GC-MS analysis of the alditol-acetate derivatives generated from the aq-ACN-MeOH-soluble fractions from *Conocephalum* (Cono), *Dicranum* (Dicran), *Marchantia* (March), *Mougeotia* (Moug), *Netrium* (Net), *P. patens* (Physco), *Polytrichum* (Poly), *Sphagnum* (Sphag), and *S. polyrhiza* (Spiro). Data is the average ± standard error of three independent samples. N.D. indicates not detected.

Glycose	*Moug.*	*Net*	*March*	*Cono*	*Poly*	*Physco*	*Dicran*	*Sphag*	*Spiro*
	Monosaccharide (mol %)								
Rha	0.3 ± 0.1	0.9 ± 0.1	0.8 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.3 ± 0.3	0.6 ± 0.1	0.3 ± 0.1	0.1 ± 0.1
Fuc	5.2 ± 0.1	12.7 ± 2.1	2.8 ± 0.1	0.2 ± 0.2	0.4 ± 0.1	2.0 ± 0.2	1.7 ± 0.8	1.7 ± 0. 1	3.7± 0.3
Rib	0.7 ± 0.1	4.7 ± 1.0	1.4 ± 0.1	0.2 ± 0.3	1.2 ± 0.1	1.9 ± 0.2	1.2 ± 0.1	1.9 ± 0.1	3.6 ± 0.1
Ara	5.8 ± 0.3	10.4 ± 0.6	8.8 ± 0.1	3.5 ± 0.7	2.6 ± 0.1	11.4 ± 1.1	6.0 ± 0.2	11.1 ± 1.0	9.8 ± 0.2
Xyl	3.1 ± 0.1	10.3 ± 1.7	2.9 ± 0.3	1.5 ± 0.5	0.5 ± 0.1	3.9 ± 0.3	5.9 ± 1.0	4.2 ± 0.7	3.9 ± 0.2
Api	0.8 ± 0.1	N.D.	1.1 ± 0.1	0.5 ± 0.2	0.3 ± 0.1	0.8 ± 0.1	0.9 ± 0.1	1.9 ± 0.3	3.1 ± 0.1
Man	36.7 ± 0.1	15.8 ± 2.4	41.4 ± 0.1	41.3 ± 12.5	53.1 ± 0.2	4.9 ± 0.6	56.1 ± 0.5	58.6 ± 1.2	44.0 ± 0.5
Gal	30.5 ± 0.1	39.3 ± 2.4	16.9 ± 0.1	4.0 ± 0.9	4.7 ± 0.1	54.3 ± 6.9	5.4 ± 0.6	5.0 ± 0.2	6.0 ± 0.2
Glc	16.8 ± 0.1	6.0 ± 0.7	24.0 ± 0.1	47. ± 9.7	37.2± 0.2	20.5 ± 9.3	22.5 ± 1.2	15.3 ± 1.6	26.0 ± 0.4

TABLE.2.3

Glycosyl-residue compositions of the cell walls isolated from selected green algae, avascular plants and *S. polyrhiza*.

The glycosyl residue composition (mol%) was determined by GC-MS analysis of alditol-acetate derivatives generated from the cell walls of each plant. Data is the average ± standard error of three independent samples. N.D. indicates not detected.

Glycose	*Moug*	*Net*	*March*	*Cono*	*Poly*	*Physco*	*Dicran*	*Sphag*	*Spiro*
	Monosaccharide (mol %)								
Rha	0.3± 0.1	0.4 ± 0.1	1.0 ± 0.1	0.3 ± 0.1	1.2 ± 0.1	1.2 ± 0.1	1.5 ± 0.1	1.9 ± 0.1	0.9 ± 0.1
Fuc	19.5 ± 3.1	21.4 ± 0.3	5.4 ± 0.1	1. 0± 0.1	1.1 ± 0.1	2.0 ± 0.1	4.1 ± 0.1	4.7 ± 0.1	2.8± 0.1
Rib	0.3 ± 0.1	1.5 ± 0.1	0.4 ± 0.1	0.4 ± 0.1	0.3 ± 0.1	2.1 ± 0.2	0.5 ± 0.1	0.5 ± 0.1	2.8 ± 0.1
Ara	11.7 ± 1.8	9.5 ± 0.5	24.9 ± 0.1	20.8 ± 0.9	6.7 ± 0.1	15.0 ± 0.7	11.7 ± 0.2	10.7 ± 0.1	14.8 ± 0.2
Xyl	32.9 ± 4.1	10.1 ± 0.6	11.2 ± 0.1	12.2 ± 0.3	8.2 ± 0.1	15.1 ± 1.0	12.0 ± 0.3	11.3 ± 0.1	15.8 ± 0.8
Api	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	21.4 ± 0.2
Man	1.3 ± 0.2	1.8 ± 0.1	12.3 ± 0.1	21.5 ± 0.6	58.3 ± 0.1	8.0 ± 0.5	30.5 ± 2.1	30.4 ± 0.1	2.0 ± 0.1
Gal	28.0 ± 10.2	23.8 ± 0.7	42.7 ± 0.1	40.0 ± 1.1	20.6 ± 0.1	48.5 ± 3.5	35.7 ± 2.9	36.3 ± 0.1	21.4 ± 0.6
Glc	6.0 ± 0.9	31.5 ± 0.3	2.2 ± 0.1	3.9 ± 0.1	3.8 ± 0.1	8.2 ± 3.3	4.1 ± 0.1	4.2 ± 0.1	17.9 ± 0.5

FIGURE 2.1

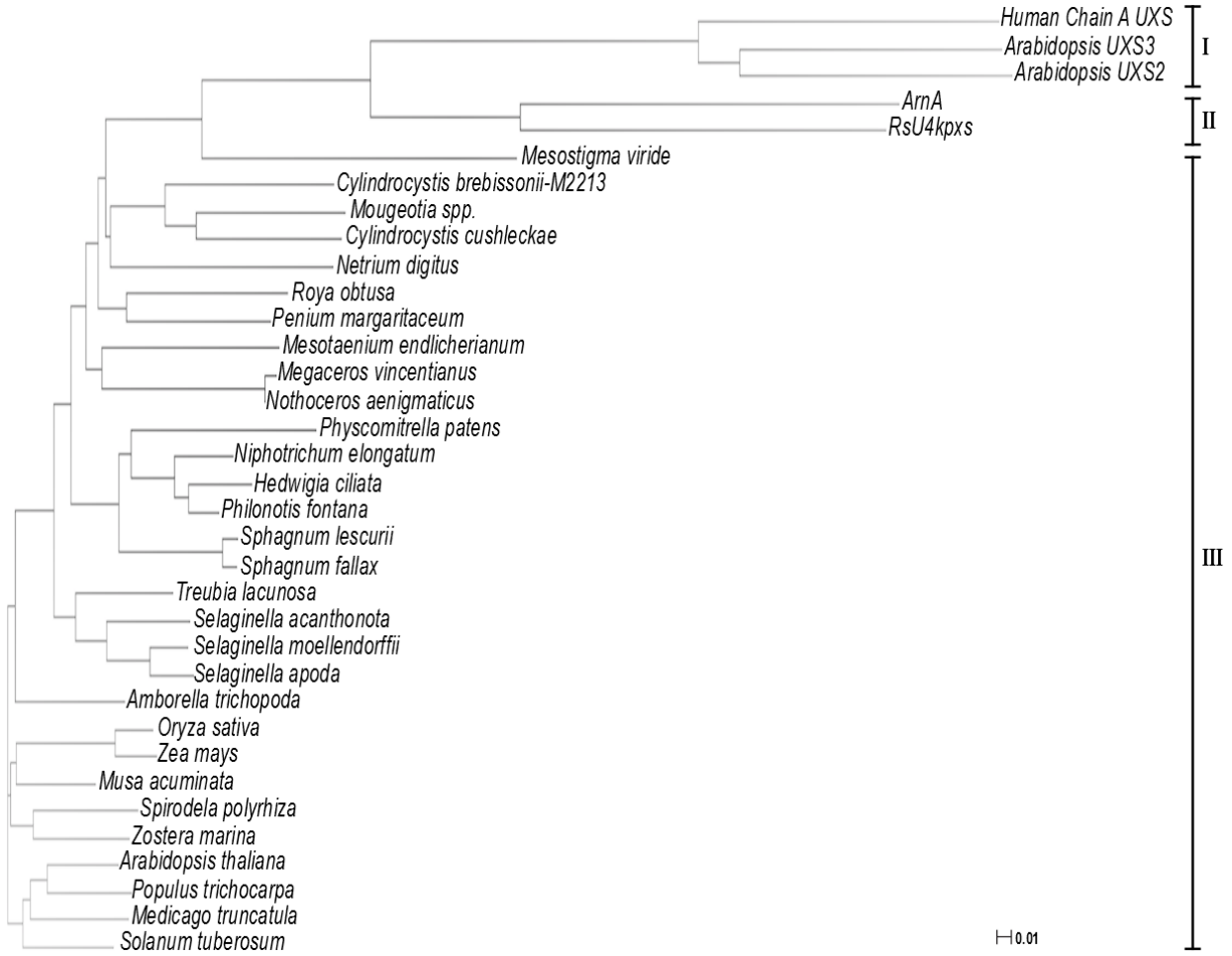


FIGURE 2.1. Unrooted phylogenetic tree of UDP-sugar decarboxylases from plants, bacteria and human. Amino acid sequences used are human UXS1 chain A (NP_079352.2), *Arabidopsis* UXS3 and UXS2 (NP_001078768.1 and NP_191842.1), *Escherichia coli* ArnA (WP_032205568.1) and *Ralstonia solanacearum* UDP-4-keto-pentose/UDP-xylose synthase (RsU4kpxs, WP_011001268.1). The UAS-like sequences used are from green algae (*Mesostigma viride*, *Cylindrocystis brebissonii*, *Mougeotia spp.*, *Cylindrocystis cushleckae*, *Netrium digitus*, *Roya obtusa*, *Penium margaritaceum* and *Mesotaenium endlicherianum*), from hornwort (*Megaceros vincentianus* and *Nothoceros aenigmaticus*), from liverwort (*Treubia lacunose*) from moss (*Physcomitrella patens*, *Niphotrichum elongatum*, *Hedwigia ciliata*,

Philonotis fontana, *Sphagnum lescurii* and *Sphagnum fallax*), from lycophytes (*Selaginella acanthonota Selaginella moellendorffii* and *Selaginella apoda*,), and from angiosperms (*Amborella trichopoda*, *Oryza sativa*, *Zea mays*, *Musa acuminata*, *Spirodela polyrhiza*, *Zostera marina*, *Arabidopsis thaliana*, *Populus trichocarpa*, *Medicago truncatula* and *Solanum tuberosum*). Alignment was made using Clustal Omega (W. Li et al., 2015; McWilliam et al., 2013; Sievers et al., 2011) and the tree generated using Dendroscope (Huson & Scornavacca, 2012).

FIGURE 2.2

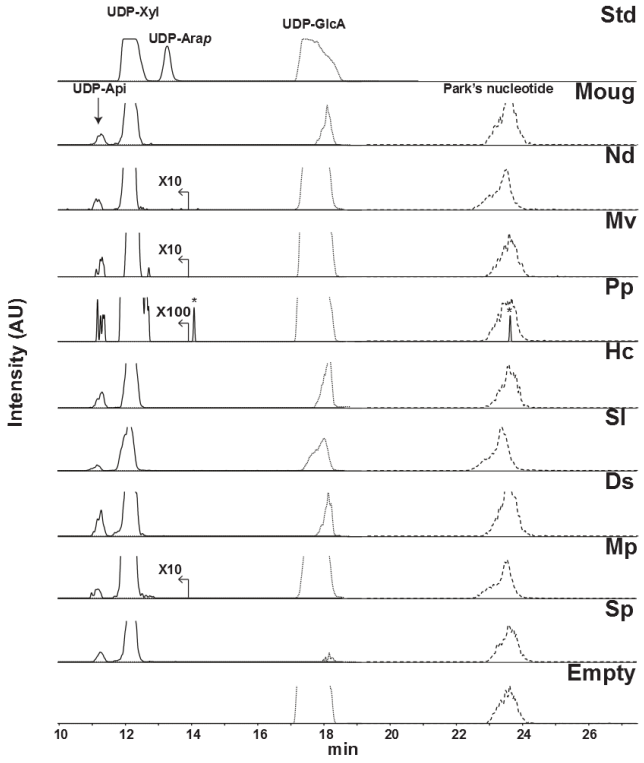


FIGURE 2.2. Extracted-ion chromatograms (XICs) of the nucleotide sugars generated in microbe using *E. coli* transformed with the UAS-like homolog from selected plants; From top to bottom, Standard (Std) UDP-GlcA, UDP-Xyl, and UDP-arabinopyranose (UDP-Ara*p*), followed by MougUAS (Moug), NdUAS (Nd), MvUAS (Mv), PpUAS (Pp), HcUAS (Hc), SlUAS (Sl),

DsUAS (Ds). MpUAS (Mp), SpUAS (Sp), and empty vector control. Negative mode $[M-H]^-$ ions diagnostic for UDP-pentose (*m/z* 535.0, solid line), UDP-uronic acid (*m/z* 579.0, dotted line) and Park's nucleotide (*m/z* 595.6, dashed line) are displayed. Park's nucleotide is a UDP-MurNAc-pentapeptide that is used as an internal standard for nucleotide sugar detection as it is abundantly made in *E. coli*. In the XICs for Nd, Mv and Mp the signal for *m/z* 535.0 is amplified by a factor of 10, and in the XIC for Pp by a factor of 100. *unknown contaminant.

FIGURE 2.3

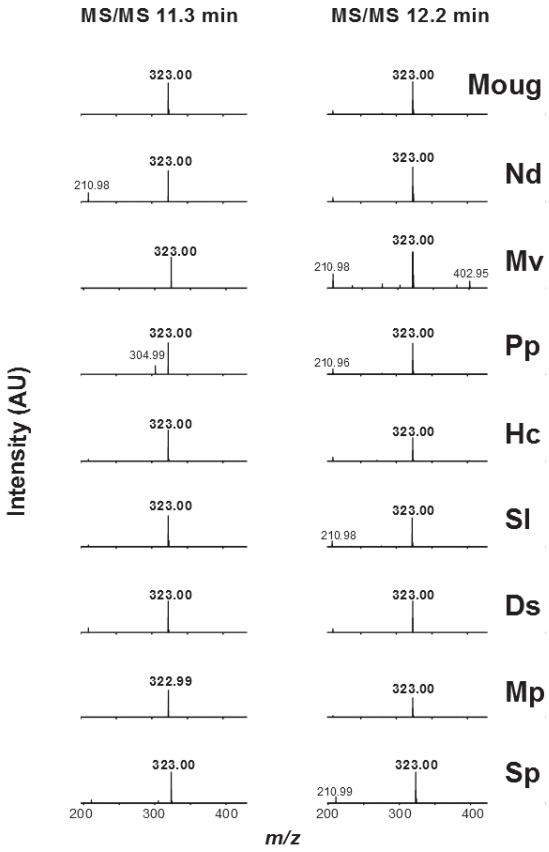


FIGURE 2.3. MS/MS spectra of the UDP-Api and UDP Xyl produced in microbe using *E. coli* transformed with UAS-like homologs from green algae and avascular plants. Second stage MS fragmentation data for the parent *m/z* 535.0 peaks at the indicated retention times; Left column

11.3 min and right column 12.2 min. From top to bottom, MougUAS, NdUAS, MvUAS, PpUAS, HcUAS, SlUAS, DsUAS. MpUAS, and SpUAS. The most abundant ion at *m/z* 323.0 is consistent with fragmentation of a UDP-sugar into $[UMP-H]^-$. Other fragment ions include *m/z* 211.0, 305.0, and 403.0; that are consistent with $[Ura-2H]^-$, $[UMP-H_2O-H]^-$, and $[UDP-H_2O-H]^-$, respectively.

FIGURE 2.4

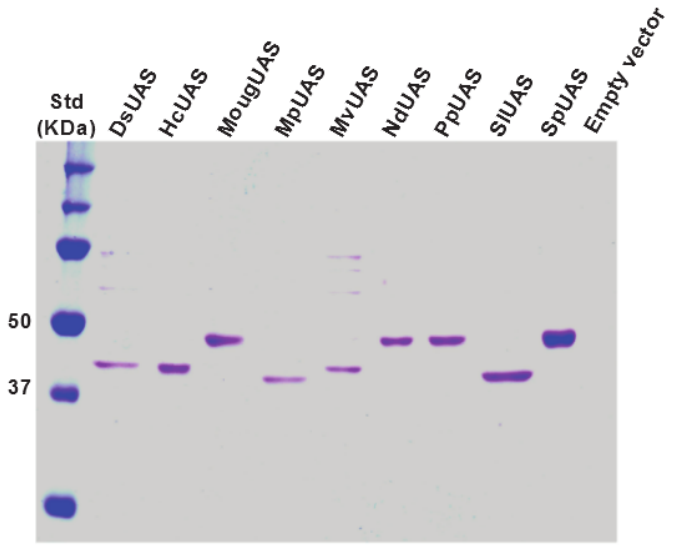


FIGURE 2.4. Expression of recombinant UAS-like proteins from green algae and avascular plants. SDS-PAGE of Ni-column-purified soluble protein extracted from *E. coli* cells induced to express DsUAS, HcUAS, MougUAS, MpUAS, MvUAS, NdUAS, PpUAS, SlUAS, and SpUAS, and empty vector control. Expected sizes are 43.0, 43.6, 45.5, 43.1, 44.8, 46.6, 45.6, 43.6, and 46.6 KDa, respectively. The rightmost lane is the empty vector control.

FIGURE 2.5

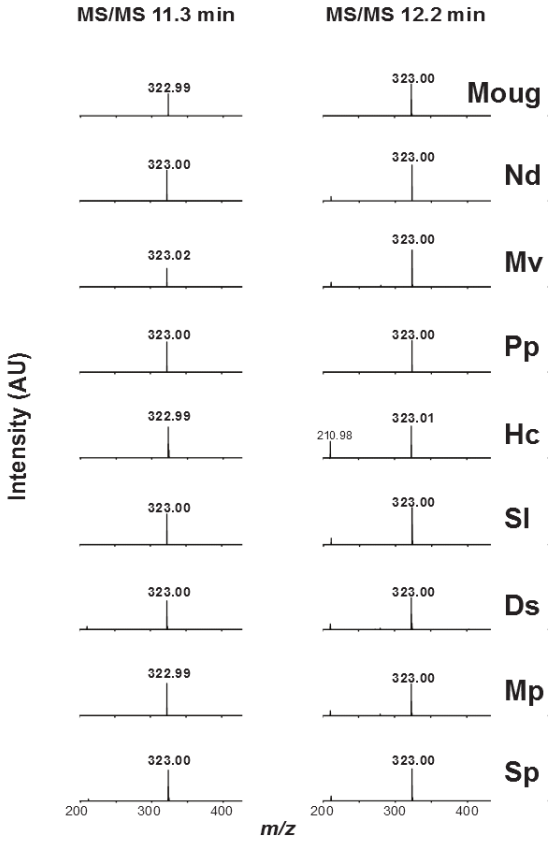


FIGURE 2.5. MS/MS spectra of UDP-Api and UDP Xyl generated by the purified recombinant UASs from green algae and avascular plants. Second stage MS fragmentation data for the parent *m/z* 535.0 peaks at the indicated retention times; Left column 11.3 min and right column 12.2 min. From top to bottom, MougUAS, NdUAS, MvUAS, PpUAS, HcUAS, SlUAS, DsUAS. MpUAS, and SpUAS. The most abundant ion at *m/z* 323.0 is indicative of UDP-sugar fragmentation into $[UMP-H]^-$. Fragments at *m/z* 211.0 are consistent with $[Ura-2H]^-$.

FIGURE 2.6

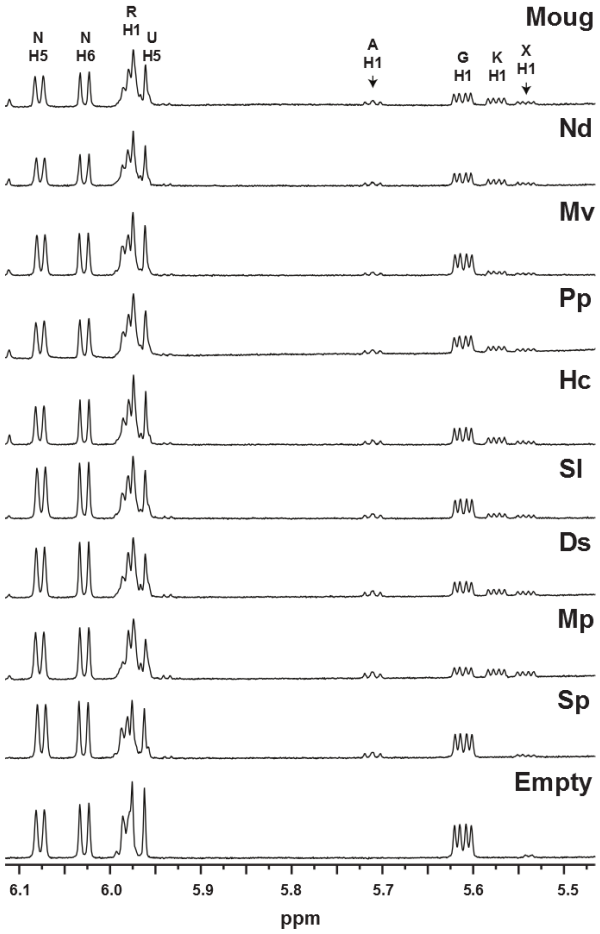


FIGURE 2.6. Selected regions of ^{1}H NMR spectra diagnostic for the products and intermediates generated by the purified recombinant UASs from green algae and avascular plants. Anomeric region between 5.50 and 5.75 ppm for the H1 protons of UDP-GlcA (G) reactant, UDP-Api (A) and UDP-Xyl (X) products and UDP-4-keto-Xyl (K) intermediate are shown (magnified by 3X for clarity). NMR region (5.95 and 6.08 ppm) diagnostic for UDP, NAD^{+} cofactor and apiofuranosyl-1,2-cyclic-phosphate (Ac) degradation product is included. NMR spectral traces from top to bottom show UAS activity of MougUAS, NdUAS, MvUAS, PpUAS, HcUAS, SlUAS, DsUAS. MpUAS, SpUAS and empty vector control. Peaks labeled N correspond to

NAD^+ H5 and H6; R corresponds to ribose H1; U corresponds to uracil H5. For additional chemical shift assignments see **Table 2.S1**.

FIGURE 2.7

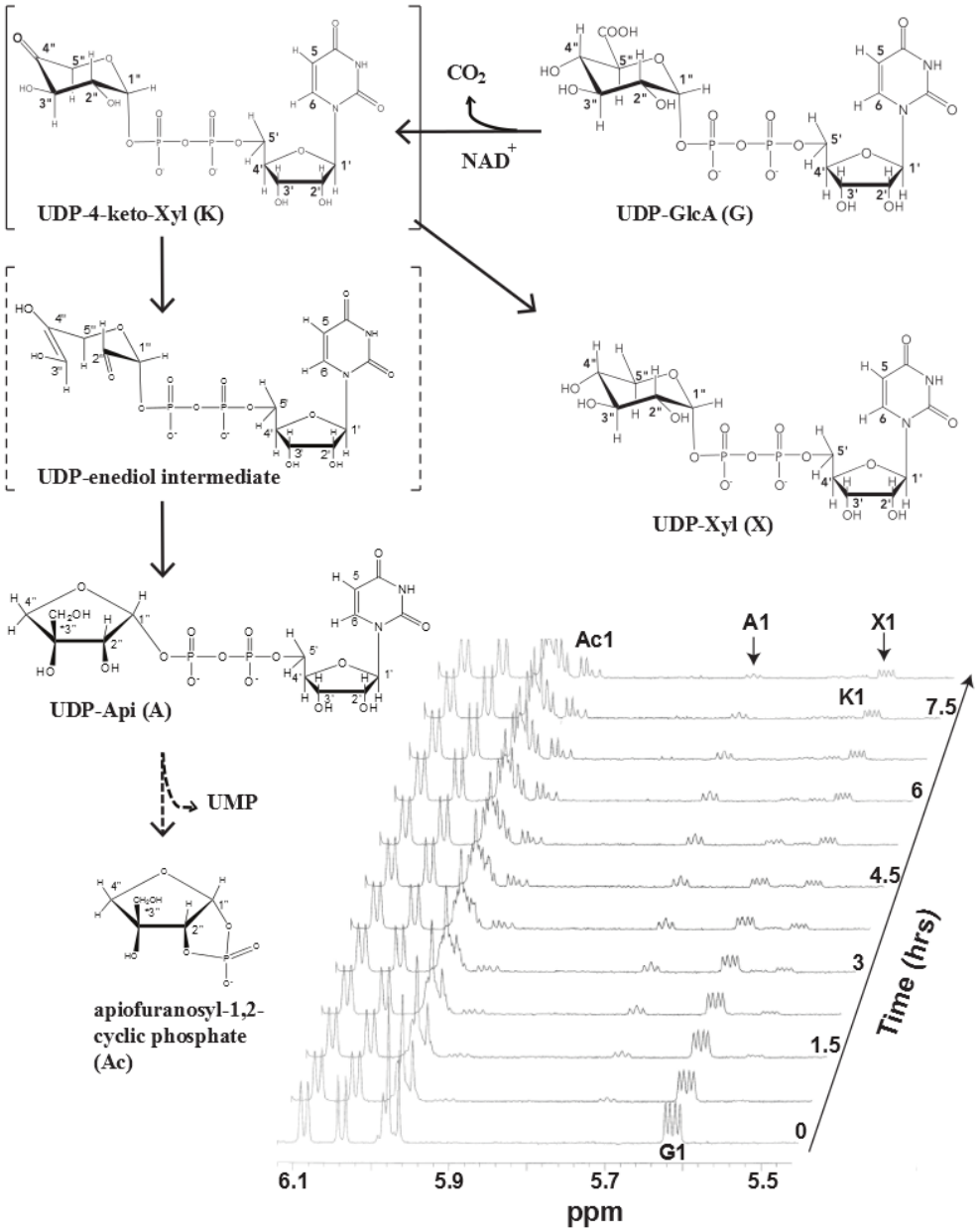


FIGURE 2.7. Proposed reaction mechanism for UAS and real time ^{1}H NMR analysis of recombinant SpUAS activity. Structures for reactant, products, detectable UDP-4-keto-Xyl intermediate (solid brackets), undetected UDP-enediol intermediate (dashed brackets) and degradation product are shown. The reaction was carried out at 37 °C. The selected anomeric region (between 5.3 and 6.1 ppm) for protons of reactant, intermediate and products is shown. Only select time-resolved spectra are displayed to prevent overcrowding of peaks.

FIGURE 2.8

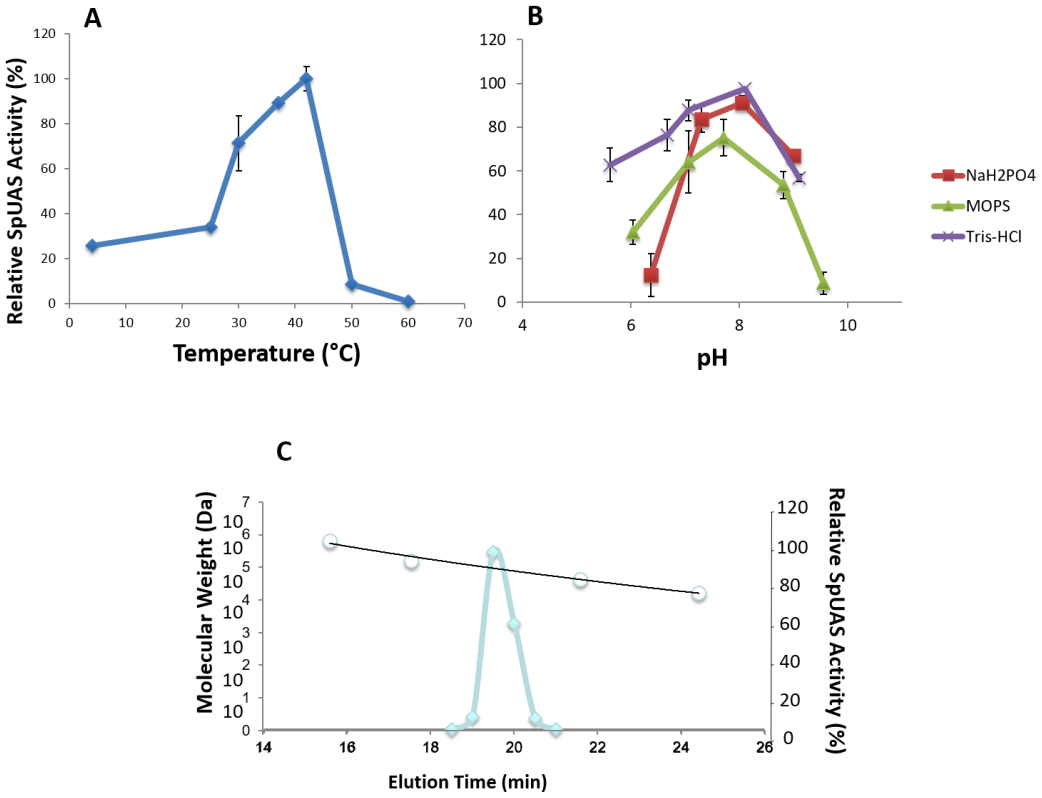


FIGURE 2.8. Characteristics of recombinant SpUAS. *A*, the effect of temperature on the relative activity of purified recombinant SpUAS. *B*, the effect of pH buffers on the relative activity of purified recombinant SpUAS. *C*, size-exclusion chromatography of recombinant SpUAS suggests that the active enzyme exists in solution as a dimer. Purified recombinant SpUAS was

fractionated on a Superdex75 gel filtration column and fractions collected (every 30 sec) and assayed for activity. The relative activity (indicated by closed diamonds) was determined by HPLC. The molecular weight of the enzyme in solution is based on the interpolation from the relative elution times of standard protein markers (indicated by open circles). Error bars represent S.E. of 3 replicates.

FIGURE 2.9

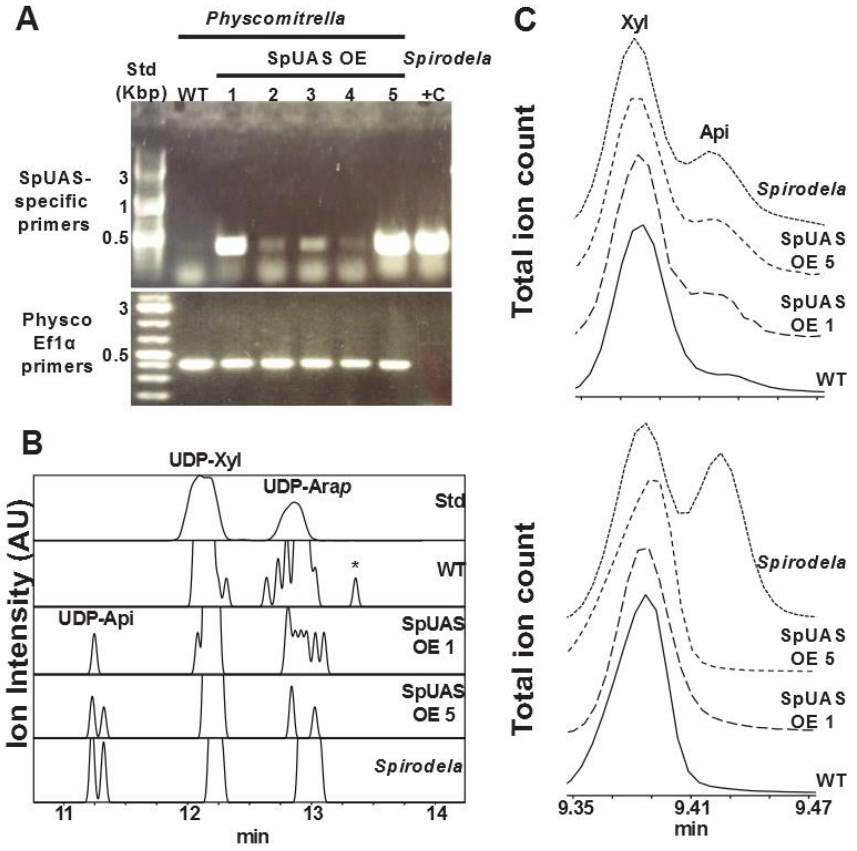


FIGURE 2.9. Overexpression of SpUAS in *P. patens* resulting in increased amounts of UDP-Api but not the formation of apiose-containing cell wall polysaccharides. *A*, transcript analysis of SpUAS. RNA extracted and cDNAs generated from wild-type *P. patens* (WT), 5 lines transformed to overexpress SpUAS (SpUAS OE 1-5), and from *S. polyrhiza*. SpUAS and

Physcomitrella Ef1α amplified by PCR and run on 1% (w/v) agarose; *B*, LC-MS analysis of ACN:MeOH:H_2O (40:40:20, v/v/v) extracts from WT *P. patens*, SpUAS OE 1, SpUAS OE 5, and *S. polyrhiza*. XICs for *m/z* 535.0, diagnostic for UDP-pentose, are shown. Std contains UDP-Xyl and UDP-Ara*p*. UDP-Api is detected in *S. polyrhiza* and in the *P. patens* SpUAS-overexpressing lines. In all but the Std trace, the *m/z* 535.0 signal has been amplified by a factor of 100; *C*, GC-MS analysis of alditol-acetate derivatives from ACN:MeOH:H_2O (40:40:20, v/v/v) extracts. The region of the total ion count for xylitol (Xyl) and apiitol (Api) is expanded, and that for cell wall fractions is below. * indicates unknown contaminant.

SUPPLEMENTAL INFORMATION

TABLE 2.S1. Chemical shifts and coupling constants for the protons of UDP-Api, UDP-Xyl and UDP-4-keto-Xyl formed from UDP-GlcA by recombinant SpUAS, and the apiofuranosyl-1,2-cyclic phosphate that is formed by spontaneous degradation of UDP-Api

	H1	**H2**	**H3*$_{a,b}$**	**H4**	**H5**
UDP-α-D-Api (A)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.71 *Quartet* $J_{1'', 2''}$ 4.5 $J_{1'', P}$ 5.6	4.01 *Quartet* $J_{1'', 2''}$ 4.5 $J_{2'', P}$ 2.2	*3.5, 3.6 $J_{3a'',3b''}$ 12.3	4.08, 4.09 $J_{4a'',4b''}$ 4.1	
UDP-α-D-Xyl (X)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.54 *Quartet* $J_{1'', 2''}$ 3.4, $J_{1'', P}$ 7.0	3.51 *Doublet* $J_{2'', 3''}$ 3.5	3.74 *Doublet* $J_{3'', 4''}$ 9.5	3.67	3.78
UDP-4-keto-Xyl (K)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.57 *Quartet* $J_{1'', 2''}$ 3.5 $J_{1'', P}$ 7.0	3.94 *Doublet* $J_{2'', 3''}$ 5.5	*3.81 *Doublet* $J_{3'', 4''}$ 9.8		3.53
Apiofuranosyl-1,2-cyclic phosphate (Ac)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.93 *Doublet of doublets* $J_{1'', 2''}$ 4.3 $J_{1'', P}$ 15.1	4.56 *Doublet* $J_{2'', P}$ 7.0	3.56, 3.6 $J_{3a'', 3b''}$ 12.3	3.86, 3.94 $J_{4a'', 4b''}$ 9.4	
Ribose (R)					
Chemical shifts, δ (ppm[a]), J coupling constants (Hz)	5.97 $J_{1', 2'}$ 3.6	4.35	4.34	4.26	4.19, 4.23 $J_{5a', 5b'}$ 12
Uracil (U)					
Chemical shifts, δ (ppm[a]) J coupling constants (Hz)					5.96 $J_{5, 6}$ 8.1

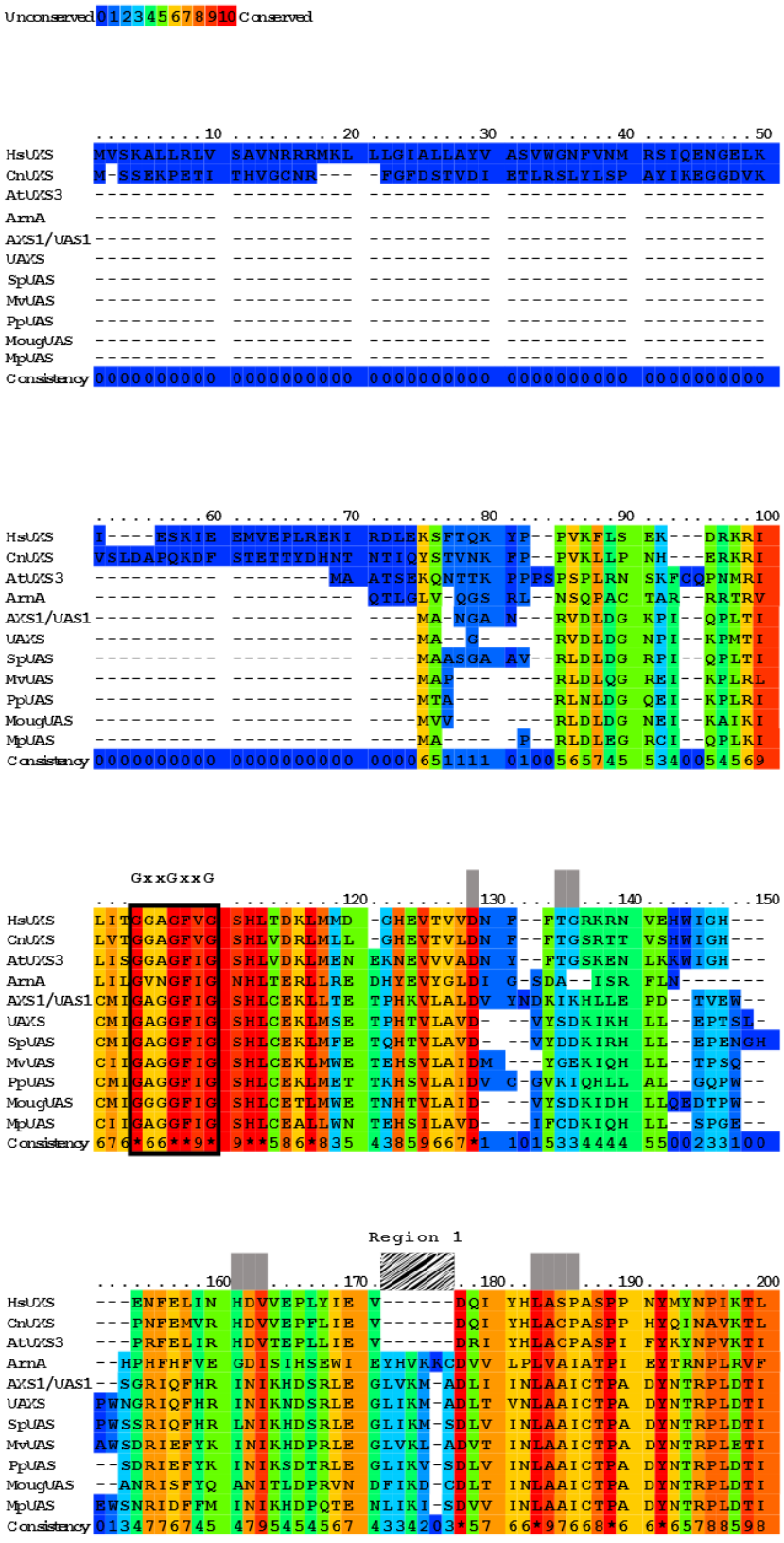
[a] Chemical shifts are in ppm relative to internal DSS signal set at 0.00 ppm. Proton-proton coupling constants in Hz are shown as well as the $J_{1'', P}$ coupling values between phosphate and the H1" proton of UDP-Api (A), UDP-Xyl (X), UDP-4-keto-Xyl (K) and apiofuranosyl-1,2-cyclic phosphate (Ac). The chemical shift values for uracil (U) and ribose (R) protons are similar

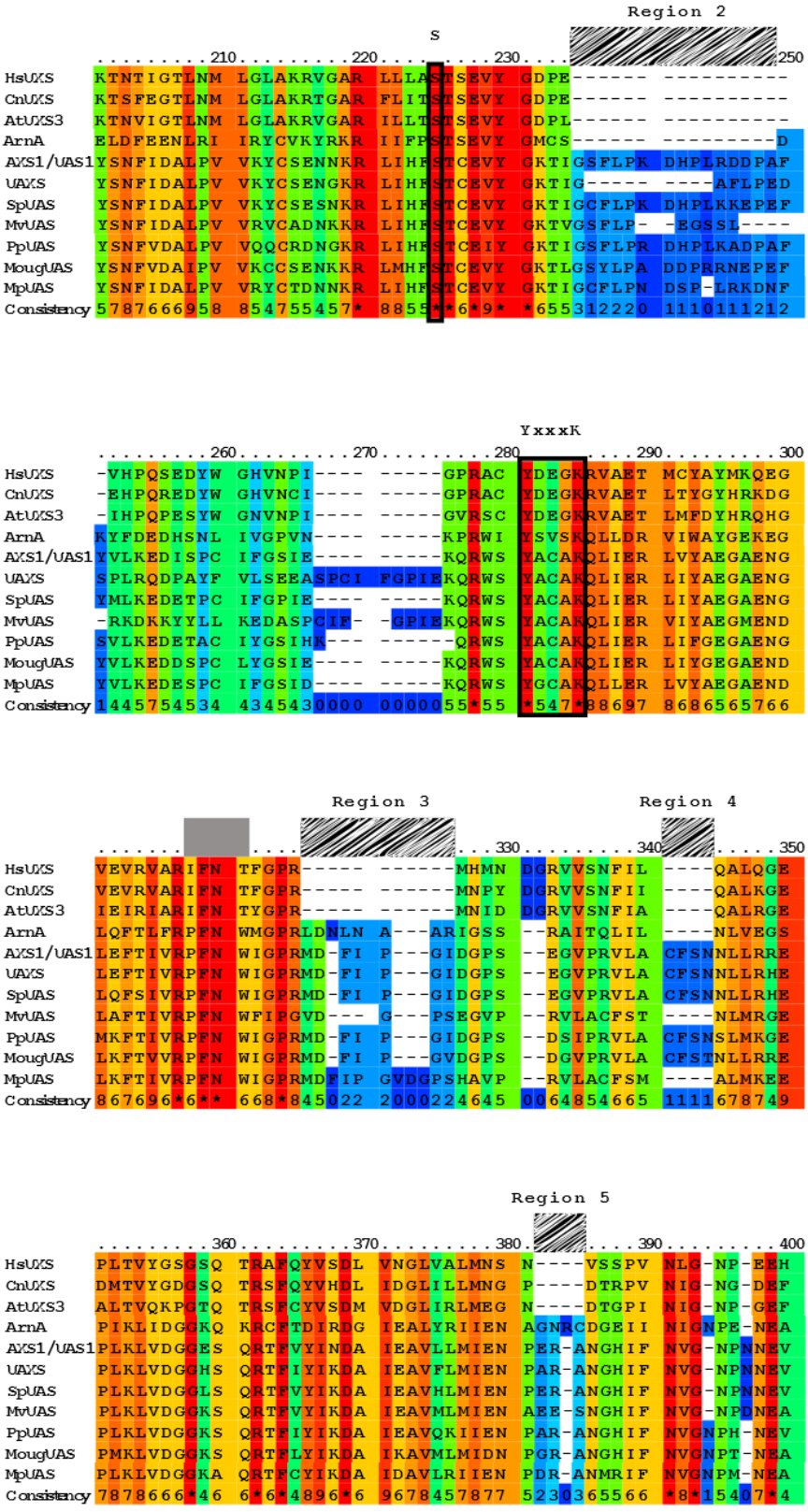
to the uracil and ribose protons of other UDP-sugars. Compounds correspond to structures in **Figure 2.7**. * refers to the proton of the exocyclic carbon in UDP-Api and apiofuranosyl-1,2-cyclic phosphate (Ac). Peak assignments were made using known values as reference (Gu et al., 2011; Guyett et al., 2009).

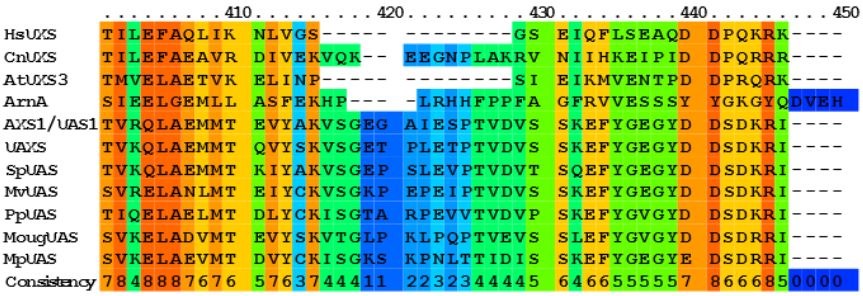
TABLE 2.S2. Primers used in plasmid generation, genotyping and transcript analysis. Obtained from IDT.

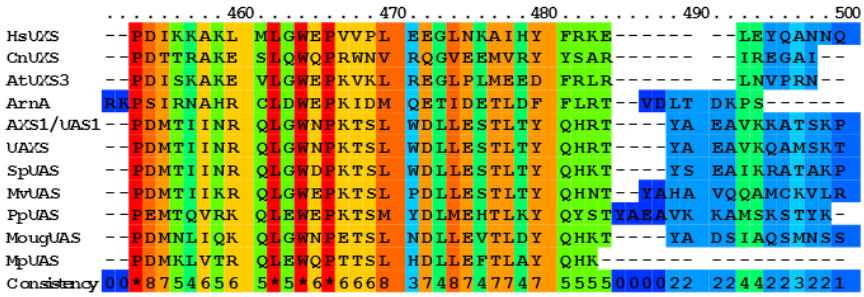
Primer	Sequence (5'-3')
DsAUS_F	GTATTTTCAGGGCGCCATGACGGCGAGGGTGCTG
DsUAS_R	GCCGGATCGAATTCACTAAGTGACAATCGCCTTCGACATTG
HcUAS_F	GTATTTTCAGGGCGCCATGGCGGGGGAGAGGGTG
HcUAS_R	GCCGGATCGAATTCACTAGTTGTTGACGATCGCCTTGGAC
MougUAS_F	GTATTTTCAGGGCGCCATGGTTGTCAGACTAGATCTGGATG
MougUAS_R	GCCGGATCGAATTCATCATGCTGCTTTAGAAGAGTTCATGG
MpUAS_F	GTATTTTCAGGGCGCCATGGCGCCGAGGCTGGAC
MpUAS_R	GCCGGATCGAATTCATTACAAGTCACTTGTCGAAAATGTCTTGGAG
MvUAS_F	GTATTTTCAGGGCGCCATGGCTCCTAGACTGGACCTG
MvUAS_R	GCCGGATCGAATTCATCAGGTTGAAGCTAGGACTTTGCAC
NdUAS_F	GTATTTTCAGGGCGCCATGGCTGGAGGCATAAGACTGG
NdUAS_R	GCCGGATCGAATTCACTAGAACTCCATCGCAACTGCCTTTAC
pENTR_SD_TOPO_F	AAGGGTGGGCGCGCCGACCCAG
pENTR_SD_TOPO_R	GGTGAAGGGCTCCTTCTTAAAGTTAAAC
pENTR_SpUAS _F	GAAGGAGCCCTTCACCATGGCGGCTTCTGGCGCGGCTG
pENTR_SpUAS _R	CGGCGCGCCCACCCTTGCTGGTTGCAACAGGTTTTGC
pET28b_TEV_F	TGAATTCGATCCGGCTGCTAACAAAGCCCG
pET28b_TEV_R	CATGGCGCCCTGAAAATACAGGTTTTC
Physco_locus108_F	CATGCTGCGTGTAGACCTGGC
Physco_locus108_R	CGCGTGTCAGGAGTACTCTTGC
PpEF1a_F	GAAGCGGCGGAGATGAAC
PpEF1a_R	ACGTCTGCCTCGCTCTAGC
PpUAS_F	GTATTTTCAGGGCGCCATGACGGCGAGATTGAACCTC
PpUAS_R	GCCGGATCGAATTCACTATTTGTAAGTAGATTTAGACATGGCCTTC
SlUAS_F	GTATTTTCAGGGCGCCATGGTAGTGACGACGAGGGTG
SlUAS_R	GCCGGATCGAATTCATCAGTTGGCAAGAGTCTTTGACATTG
SpUAS_F	GTATTTTCAGGGCGCCATGGCGGCTTCTGGCGCGGCTG
SpUAS_Mid_R	GTCGAGCGGGCGCGTGTTGTAG
SpUAS_R	GCCGGATCGAATTCAGCTGGTTGCAACAGGTTTTGC

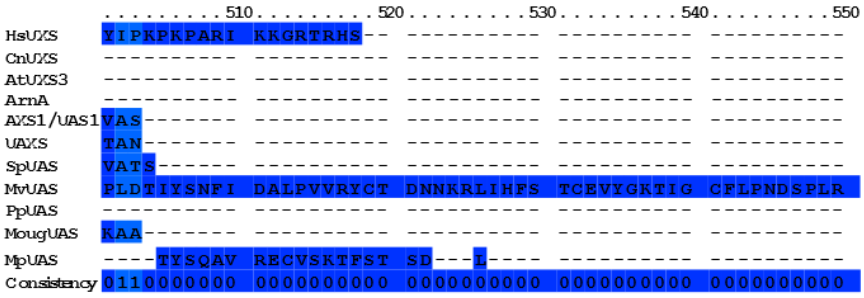
FIGURE 2.S1


Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved
HsUXS
CnUXS
AtUXS3
ArnA
AXS1/UAS1
UAXS
SpUAS
MvUAS
PpUAS
MougUAS
MpUAS
Consistency
GxxGxxG
Region 1


Region 2
S
YxxxK
Region 3
Region 4
Region 5
HsUXS
CnUXS
AtUXS3
ArnA
AXS1/UAS1
UAXS
SpUAS
MvUAS
PpUAS
MougUAS
MpUAS
Consistency


410 420 430 440 450
HsUXS TILEFAQLIK NLVGS----- --------GS EIQFLSEAQD DPQKRK----
CnUXS TILEFAEAVR DIVEKVQK-- EEGNPLAKRV NIIHKEIPID DPQRRR----
AtUXS3 TMVELAETVK ELINP----- --------SI EIKMVENTPD DPRQRK----
ArnA SIEELGEMLL ASFEKHP--- -LRHHFPPFA GFRVVESSSY YGKGYQDVEH
AXS1/UAS1 TVRQLAEMMT EVYAKVSGEG AIESPTVDVS SKEFYGEGYD DSDKRI----
UAXS TVKQLAEMMT QVYSKVSGET PLETPTVDVS SKEFYGEGYD DSDKRI----
SpUAS TVKQLAEMMT KIYAKVSGEP SLEVPTVDVT SQEFYGEGYD DSDKRI----
MvUAS SVRELANLMT EIYCKVSGKP EPEIPTVDVS SKEFYGEGYD DSDKRI----
PpUAS TIQELAELMT DLYCKISGTA RPEVVTVDVP SKEFYGVGYD DSDKRI----
MougUAS SVKELADVMT EVYSKVTGLP KLPQPTVEVS SLEFYGVGYD DSDRRI----
MpUAS SVKELAEVMT DVYCKISGKS KPNLTTIDIS SKEFYGEGYE DSDRRI----
Consistency 784888767 6 576374441 1 223234444 5 646655555 7 866685 0000


460 470 480 490 500
HsUXS --PDIKKAKL MLGWEPVVPL EEGLNKAIHY FRKE------ --LEYQANNQ
CnUXS --PDTTRAKE SLQWQPRWNV RQGVEEMVRY YSAR------ --IREGAI--
AtUXS3 --PDISKAKE VLGWEPKVKL REGLPLMEED FRLR------ --LNVPRN--
ArnA RKPSIRNAHR CLDWEPKIDM QETIDETLDF FLRT--VDLT DKPS------
AXS1/UAS1 --PDMTIINR QLGWNPKTSL WDLLESTLTY QHRT----YA EAVKKATSKP
UAXS --PDMTIINR QLGWNPKTSL WDLLESTLTY QHRT----YA EAVKQAMSKT
SpUAS --PDMTIINR QLGWDPKTSL WDLLESTLTY QHKT----YS EAIKRATAKP
MvUAS --PDMTIIKR QLGWEPKTSL PDLLESTLTY QHNT--YAHA VQQAMCKVLR
PpUAS --PEMTQVRK QLEWEPKTSM YDLMEHTLKY QYSTYAEAVK KAMSKSTYK-
MougUAS --PDMNLIQK QLGWNPETSL NDLLEVTLDY QHKT----YA DSIAQSMNSS
MpUAS --PDMKLVTR QLEWQPTTSL HDLLEFTLAY QHK------- ----------
Consistency 00*8754656 5*5*6*6668 3748747747 5555000022 2244223221


510 520 530 540 550
HsUXS VIPKPKPARI KKGRTRHS-- ---------- ---------- ----------
CnUXS ---------- ---------- ---------- ---------- ----------
AtUXS3 ---------- ---------- ---------- ---------- ----------
ArnA ---------- ---------- ---------- ---------- ----------
AXS1/UAS1 VAS------- ---------- ---------- ---------- ----------
UAXS TAN------- ---------- ---------- ---------- ----------
SpUAS VATS------ ---------- ---------- ---------- ----------
MvUAS PLDTIYSNFI DALPVVRYCT DNNKRLIHFS TCEVYGKTIG CFLPNDSPLR
PpUAS ---------- ---------- ---------- ---------- ----------
MougUAS KAA------- ---------- ---------- ---------- ----------
MpUAS ----TYSQAV RECVSKTFST SD---L---- ---------- ----------
Consistency 0110000000 0000000000 0000000000 0000000000 0000000000


HsUXS --
CnUXS --
AtUXS3 --
ArnA --
AXS1/UAS1 --
UAXS --
SpUAS --
MvUAS KD
PpUAS --
MougUAS --
MpUAS --
Consistency 00

FIGURE 2.S1. UXS, ArnA, and UAS multiple sequence alignment. Full amino acid sequence alignment of the UDP-GlcA decarboxylase domain of *E. coli* ArnA (WP_032205568.1), human UXS (HsUXS, NP_079352.2), fungal UXS from *Cryptococcus neoformans* (CnUXS, XP_572003.1), plant UAXS from *Solanum tuberosum* (NP_001275341.1), Arabidopsis AXS1/UAS1 & UXS3 (NP_180353.1 & NP_001078768.1), and the putative UAS-like enzymes of the algae *Mougeotia spp*. (MougUAS), liverwort *Marchantia paleacea* (MpUAS), hornwort *Megaceros vincentianus* (MvUAS), moss *Physcomitrella patens* (PpUAS), and the monocot *Spirodela polyrhiza* (SpUAS) were aligned with PRALINE (Simossis & Heringa, 2005) using the BLOSUM62 scoring matrix. AXS1/UAS1 was used to probe the 1KP database (Matasci et al., 2014) and the Spirodela polyrhiza genome (W. Wang et al., 2014) and aligned with other UAS and UAS-like protein sequences. Proposed catalytic and cofactor binding residues are outlined in black. Other proposed active-site residues and insertion regions are marked with grey and striped boxes, respectively in the numbering row.

CHAPTER 3

SYNTHESIS OF UDP-APIOSE IN BACTERIA: THE MARINE PHOTOTROPH *GEMINICOCCUS ROSEUS* AND THE PLANT PATHOGEN *XANTHOMONAS PISI*[1]

[1]James Smith, Maor Bar-Peled. 2017. PLOS One. PMCID: PMC5607165

Reprinted here with permission of publisher

ABSTRACT

The branched-chain sugar apiose was widely assumed to be synthesized only by plant species. In plants, apiose-containing polysaccharides are found in vascularized plant cell walls as the pectic polymers RGII and ApiGalA. Apiosylated secondary metabolites are also common in many plant species including ancestral avascular bryophytes and green algae. Apiosyl-residues have not been documented in bacteria. In a screen for new bacterial glycan structures, we detected small amounts of apiose in methanolic extracts of the aerobic phototroph *Geminicoccus roseus* and the pathogenic soil-dwelling bacteria *Xanthomonas pisi*. Apiose was also present in the cell pellet of *X. pisi*. Examination of these bacterial genomes uncovered genes with relatively low protein homology to plant UDP-apiose/UDP-xylose synthase (UAS). Phylogenetic analysis revealed that these bacterial UASs-like homologs belong in a clade distinct to UAS and separated from other nucleotide sugar biosynthetic enzymes. Recombinant expression of three bacterial UAS-like proteins demonstrates that they actively convert UDP-glucuronic acid to UDP-apiose and UDP-xylose. Both UDP-apiose and UDP-xylose were detectable in cell cultures of *G. roseus* and *X. pisi*. We could not however, definitively identify the apiosides made by these bacteria, but the detection of apiosides coupled with the *in vivo* transcription of bUAS and production of UDP-apiose clearly demonstrate that these microbes have evolved the ability to incorporate apiose into glycans during their lifecycles. While this is the first report to describe enzymes for the formation of activated apiose in bacteria, the advantage of synthesizing apiose-containing glycans in bacteria remains unknown. The characteristics of bUAS and its products are discussed.

INTRODUCTION

Apiose (3-C-[hydroxymethyl]-D-erythrofuranose, Api) is a common sugar residue of the plant pectic polymer rhamnogalacturonan-II (RG II) in vascular plant species and apiogalacturonan (O'Neill et al., 1990; Picmanova & Moller, 2016). In addition, apiose-containing small secondary metabolites were detected in lichens (Rezanka & Guschina, 2000) and in the fungus *Morchella conica* (Zheng et al., 1998).

In a screen for novel bacterial glycans, we unexpectedly found an apiosyl residue in methanolic extracts of two gram-negative proteobacteria: the alphaproteobacteria *Geminicoccus roseus* and the gammaproteobacteria *Xanthomonas pisi* (**Figure 3.1**). In addition to apiose we also detected xylose but the latter has been previously described in certain bacteria (Bystrova et al., 2006; Coyne et al., 2011; Fletcher et al., 2007; Gu et al., 2010; Gu et al., 2011; Isogai et al., 1986; Knirel, 1990).

In plants the activated nucleotide sugar donor used to synthesize apiose-containing glycan is UDP-apiose (UDP-Api). A bifunctional enzyme (UAS) converts UDP-glucuronic acid (UDP-GlcA) predominantly to UDP-Api but UDP-xylose (UDP-Xyl) is also made in a ratio close to 2:1. The UAS enzyme was characterized in several plant species and was named UDP-apiose/UDP-xylose synthase (H. Grisebach & Döbereiner, 1964; Ortmann et al., 1972; Sandermann et al., 1968). UAS belongs to the short-chain dehydrogenase/reductase (SDR) superfamily, which also includes UDP-xylose synthase (UXS) (Harper & Bar-Peled, 2002). UDP-xylose synthase (UXS) was identified in plant, human and bacteria and this enzyme has a single activity converting UDP-glucuronic acid to UDP-xylose. Functionally characterized bacterial UXSs (Coyne et al., 2011; Gu et al., 2011) belong to nitrogen fixing rhizobium species that symbiotically reside with plants. Another bacterial species that harbors UXS is the common

human microbial gut species belonging to *Bacteroides* (Coyne et al., 2011). Certain plant pathogenic bacteria like *Ralstonia solanacearum* also synthesize UDP-xylose as a byproduct of the related enzymes Rsu4kpxs (Gu et al., 2010). Lastly, the common gram-negative bacterial gene ArnA encodes a UDP-glucuronic acid decarboxylase able to form an intermediate UDP-4-keto-L-arabinose (Breazeale et al., 2005; Gu et al., 2010) on route to the formation of UDP-arabinose-4-amino and UDP-Ara4NF. None of the UXS enzymes were demonstrated to have dual activity like UAS.

In an effort to explain the existence of apiose in these bacteria [*Geminicoccus roseus* and *Xanthomonas pisi]*, we searched their genomes for potential genes encoding such activity. Using plant UAS as a probe, we identified UAS-like homologous genes from several bacterial species. We show in this report that they all are capable of synthesizing UDP-Api from UDP-GlcA. The origin of UAS in these bacteria remains unclear.

RESULTS:

Detection of Apiose in Extracts of G. roseus and X. pisi

Bacteria grown in liquid and agar media were collected and subjected to methanolic extractions. Following chemical hydrolyses, conversion of the monosaccharides to their alditol-sugar derivatives and GC-MS analyses, the methanolic extract had a peak that migrated like apiose (**Figure 3.1**). The GC chromatogram and the electron impact MS fragmentation pattern of this peak structure was not reported in Bacteria but had all the chromatographic and mass spectral features of an apiose. For example, the peaks at *m/z* 290 (**Figure 3.1**A, insert) suggest a cleavage between C2-C3 of an apiose-derivative. The peaks at *m/z* 248 and 260 and at *m/z* 247 and 187 are likely secondary ion fragments due to loss of *m/z* 42 (ketene) or 60 (acetate) from

m/z 290 and 289, respectively. The major peak at *m/z* 188 represents the deuterated form of the acetylated apiose-derivative. No apiose was detected in the organic solvent extracts. By contrast, a small amount of Api was apparent in the cell pellet of *X. pisi* but not *G. roseus* (**Figure 3.1**). In addition to apiose, these extracts also consist of other sugar residues like glucose, galactose, fucose, rhamnose, arabinose, and xylose. While both xylose and apiose have similar structures their MS fragmentation patterns differ significantly (**Figure 3.1**C).

Because mass spectra are not sufficient to discriminate apiose from its two potential epimers, we sought to further validate the nature of the apiose-like GC-MS peak by identifying genes involved in the formation of the activated sugar.

Identification and Phylogeny Analysis of Bacterial UDP-apiose/UDP-xylose Synthase-like Homologs

The BLink and BLAST programs (Altschul et al., 1997) were used to identify bacterial proteins in the NCBI non-redundant database that share sequence similarity to functional bacterial UXSs as well as plant UAS. A search in Bacteria for homologs to the amino acid sequence of *Arabidopsis* AXS/UAS1 (AEC08054.1) identified several candidates in proteobacteria: *Candidatus entotheonella* and *Geminicoccus roseus, Xanthomonas pisi* and *Yangia pacifica* with 43, 46, 48 and 49% sequence identity to *Arabidopsis* AXS/UAS1, respectively. These bacterial UAS-like gene homologs are named herein bUAS.

An unrooted phylogenetic tree (**Figure 3.2**) was generated. The analysis compared an alignment of amino acid sequences of functional plant and microbial proteins belonging to the short-chain dehydrogenase/reductase (SDR) family with the amino acid sequences of the above bacterial UAS (bUAS) proteins (Figure 3.S1). The sequence alignment included UDP-xylose

synthases (UXSs) from plant, mammal and fungi and two bacterial enzymes, a bifunctional UDP-4-keto-pentose/UDP-xylose synthase (RsU4kpxs) from the plant pathogen *Ralstonia solanacearum* and the C-terminal portion of ArnA that has a UDP-glucuronic acid 4-oxidase-6-decarboxylase activity (Breazeale et al., 2005; Gu et al., 2010). UAS, UXS, RsU4kpxs and ArnA are all decarboxylases that contain domains common to all SDRs: a conserved N-terminal Gly-*X*-*X*-Gly-*X*-*X*-Gly motif (Figure 3.S1; X= any amino acid) that is proposed to be involved in NAD^+ binding and Tyr-*X*-*X*-*X*-Lys motif with an upstream Ser that forms the catalytic site of the SDR family (Harper & Bar-Peled, 2002; Kavanagh et al., 2008; Yin et al., 2011). The bUASs cluster into a group distinct from the clades for bacterial ArnA and the UXSs (**Figure 3.2**).

According to BLAST, the *X. pisi* UAS (XpUAS) protein shares 65% sequence identity with the homolog from *G. roseus* (GrUAS), 64% with that from *Y. pacifica* and 44% with that from *C. entotheonella* (CeUAS). The closest plant homolog to XpUAS belongs to *Ornithogalum longibracteatum* (sea onion; AMM04380.1) and shares only 51% sequence identity to XpUAS. Similarly, the GrUAS is 47% identical to UAS from the sea grass *Zostera marina*, and YpUAS is 50% identical to UAS from *Vitis vinifera* (XP_002270884.1). Because no sea sponge genomes or transcriptomes are available, we cannot infer any relation of CeUAS to that of its host *Theonella swinhoei*, but its closest plant UAS homolog belongs to *Amborella trichopoda* () with 44% sequence identity. Examination of the genes flanking bUASs in each of these bacteria revealed that each is flanked by different genes, with no apparent conserved operon.

Employing the phylogeny to assess origins of the bUASs was inconclusive. While the bUASs share a branch in the phylogeny (**Figure 3.2**), they have varying sequence identities and no common theme. Additionally, the bUAS sequences are not very identical to plant UAS sequences. It is possible that bUASs originated from a bacterial ancestral gene source, however

we cannot exclude the possibility that bUASs came from plant UAS through a gene transfer event. The limited number of bUAS examples prohibits specifying their origin or the evolutionary advantage they confer. Future deposition of additional bacterial sequences will provide a better understanding of the relationship among UASs. To date there are no reports that bacteria produce apiose (Picmanova & Moller, 2016). Because we could not determine at the GC-MS level if the pentose was apiose or an apiose epimer (for example two possible epimer forms of apiose at C-2 and C-3), and because the metabolic pathway leading to formation of these apiose-like residues, was unknown, we decided to clone the genes and determine if the UAS-like homologs are capable of converting UDP-GlcA to UDP-Api, or perhaps utilizing other UDP-sugar uronates, for example UDP-GlcNAc-uronate (UDP-GlcNAcA) or UDP-galacturonic acid (UDP-GalA). To this end, XpUAS, GrUAS and CeUAS were cloned, expressed in *E. coli* and then functionally characterized.

Cloning of bUAS and in Microbe Formation of UDP-apiose

The coding sequences of the selected UAS homologs were cloned into a modified pET28b *E. coli* expression vector (Yang et al., 2009). The bUAS-containing plasmids or empty plasmid (negative control) were then individually transformed into *E. coli* together with a pCDF plasmid containing the UDP-Glc dehydrogenase (Broach et al., 2012) that provides the potential substrate for bUAS *in vivo*. Nucleotide sugar-containing extracts from the isopropyl β-D-thiogalactoside (IPTG)-induced *E. coli* cells were chromatographed by hydrophilic interaction liquid chromatography (HILIC) and analyzed by electrospray ionization mass spectrometry (ESI-MS/MS) in the negative mode. Two peaks eluting at 11.0 and 12.0 min (**Figure 3.3**) were observed in strains harboring the bUAS but not in *E. coli* cells harboring plasmid control. The

mass spectra (**Figure 3.3**) of both peaks showed an $[M-H]^-$ ion at *m/z* 535.0, that gave MS/MS ion fragments at *m/z* 403.0, 323.0 and 211.0 which are consistent with $[UDP-H_2O-H]^-$, $[UMP-H]^-$ and $[Ura-2H]^-$, respectively. The *m/z* 535.0 was not found in control *E. coli* expressing empty plasmid. Proton NMR (^{1}H NMR) analyses confirmed that the UDP-pentose eluting at 11.0 min was UDP-Api, and not UDP-apiose-epimer. These data suggest that the bUAS enzymes do synthesize UDP-Api.

Characterization of Purified Recombinant bUAS

To obtain additional evidence for the nucleotide sugar metabolism and the specific UDP-sugar uronate that the bUASs are utilizing to form UDP-Api, the recombinant His_6-tagged proteins were solubilized from *E. coli* cells and purified using nickel-affinity column. The recombinant bUASs migrated on SDS-PAGE with a predicted mass of between 43 and 45 kDa (**Figure 3.4**). Each purified UAS was shown by HILIC-ESI-MS/MS to convert UDP-GlcA to two UDP-pentose products in the presence of NAD^+. MS/MS analysis (**Figure 3.4**) of these product peaks (11.0 and 12.0 min) also gave a fragment ion at *m/z* 323.0 that is consistent with $[UMP-H]^-$. Anomeric (H-1) peaks consistent with the presence of UDP-Api and UDP-Xyl were detected in all the ^{1}H NMR spectra when the recombinant enzyme assays were performed in deuterated buffer (**Figure 3.5**).

GrUAS was the most highly expressed protein and was thus selected for further characterization. Real time ^{1}H NMR spectroscopic analysis of the products formed when GrUAS reacts with UDP-GlcA (**Figure 3.6** and **Table 3.S1**) confirmed that UDP-Api is the first product formed. GrUAS produces UDP-Api and UDP-Xyl in a ratio of ~1.7: 1.0, which is similar to characterized plant UASs (Guyett et al., 2009; Molhoj et al., 2003; Smith et al., 2016). The NMR

study with GrUAS also confirm that some of the UDP-Api is degraded and converted to the apiofuranosyl-1,2-cyclic phosphate during the *in vitro* reaction (**Figure 3.6**); this instability of UDP-apiose is a known phenomenon (Choi et al., 2012; Guyett et al., 2009; Smith et al., 2016). No degradation of UDP-Xyl is discernible over the course of the reaction.

The recombinant GrUAS is most active in 50 mM Tris-HCl, pH 8.0 - 8.5, at 37 °C (**Figure 3.7**A-B) and exists in solution as a dimer with a predicted size of 84 KDa (**Figure 3.7**C). GrUAS has a *Km* of 251 µM similar to that for *Spirodela* UAS (Smith et al., 2016), and a Kcat/Km of 60.2 nM s^{-1} (**Table 3.1**), while recombinant *Arabidopsis* AXS1/UAS1 has a reported Km of 7 µM and Kcat/Km of 43 nM s^{-1} (Molhoj et al., 2003). Previous studies have shown that UAS is inhibited by certain nucleotides and nucleotide sugars, especially UDP-Xyl and UDP-GalA (Molhoj et al., 2003; Smith et al., 2016). Under our assay conditions UDP-Xyl and UDP-GalA inhibited GrUAS activity by 9% and 77%, respectively (**Table 3.2**).

Transcript Analysis of bUAS and Detection of UDP-apiose in X. pisi and G. roseus cultures

To investigate if the bUASs are transcribed in these bacteria and if they produce UDP-Api *in vivo*, cultures of *X. pisi* and *G. roseus* were grown in liquid and agar media. Analysis of RNA for bUAS transcript was positive for both species (**Figure 3.8**) albeit at lower amounts when compared with the control *G. roseus* transcript sigma factor RpoD (WP_027134046.1), and Sig70 (EGD11293.1) as expression control for *X. pisi*. Furthermore, at culture steady state, small amounts of UDP-Api and UDP-Xyl were detected by LC-MS/MS in aqueous extracts of living culture (**Figure 3.8**). Together, these data demonstrate that genes encoding bacterial UAS are expressed and the UAS enzymes are functionally active *in vivo* leading to production of UDP-Api in the specific bacteria tested.

DISCUSSION

This report is the first to describe the sugar apiose in Bacteria, and subsequently this study led us to identify functional genes responsible for synthesizing the activated donor UDP-apiose in prokaryotes. The activity of recombinant bacterial UDP-apiose synthase (bUAS) is specific and utilizes only UDP-glucuronic acid as a substrate.

The bacterial species appearing to contain a UAS were isolated from various sources, including soil and sea. The organization of the operons harboring bUAS in these bacteria is not conserved. The genes flanking bUAS are also not conserved. In contrast, the organizations of bacterial operons that carry out synthesis of other sugar nucleotides are conserved, for example dTDP-rhamnose (Macpherson et al., 1994; Marolda & Valvano, 1995; Robertson et al., 1994). Thus, the ancestral origin of the bUAS remains somewhat elusive. Based on the current genomic database, to date only 8 bUASs exist and this number will likely increase as more marine bacteria are sequenced. Since the UAS belonging to the obligate endosymbiotic *C. entotheonella* shares less than 60% amino acid sequence identity with the other bUASs in this study, it is possible that there was no one single ancestral gene that gave rise to all of the bUASs. One possibility based on the limited sequences is that a duplication and sequence alteration of bacterial UXS or ArnA gave rise to these bUASs. However, one cannot rule out the possibility that early plant-derived UAS is the gene source for bUAS. Regardless of how or where they evolved, bUASs have the same catalytic domains and utilize the same chemistry as plant UASs (Figure 3.S1).

The bUASs are phylogenetically distinct from other short-chain decarboxylase/reductases (SDRs) and share a branch with functional plant UASs (**Figure 3.2**) All domains of life possess

enzymes (UXSs) that convert UDP-GlcA to UDP-Xyl (Gotting et al., 2000; Gu et al., 2011; Harper & Bar-Peled, 2002; Kobayashi et al., 2002; Kuhn et al., 2001), while UAS appears to be limited to plants and now Bacteria. Only a select few prokaryotic species appear to contain apiose based on the presence of bacterial UAS (bUAS), suggesting it is advantageous for these organisms' survival in their specific niche environments. Further support for this hypothesis is evidenced by the presence of UXS in the genomes of *G. roseus*, *C. entotheonella* and *Y. pacifica*, suggesting that these microbes require bUAS specifically for UDP-Api (not UDP-Xyl) synthesis.

The apiose residue observed in the methanolic extracts of *G. roseus* and *X. pisi* as well as the cell pellet fraction of *X. pisi* is likely to be incorporated as a secondary metabolite and potentially a cell wall glycan in *X. pisi*. Because no Api was detected in the chloroform-extracted fractions of culture, it is unlikely that the final Api residue is part of a glycolipid. Our previous work had shown that in green algae and basal land plants, Api residues also associated with secondary metabolites (Smith et al., 2016). In vascular plants Api is present in the cell wall polysaccharides rhamnogalacturonan-II and apiogalacturonan and in secondary metabolites (Beck & Hopf, 1990; Darvill et al., 1978; Picmanova & Moller, 2016; Watson & Orenstein, 1975). It is therefore possible that these organisms share a related family of apioside metabolites.

The identification and functional characterization of bUAS reveals a new metabolic pathway of UDP-GlcA metabolism by which select microbes have adapted to compete within their local environments. Our results provide important tools for the future study of apiose and apiosides in bacteria. Continued study of the pathways leading to apiosylated metabolites in bacteria will uncover new and interesting GT activities valuable to carbohydrate research.

EXPERIMENTAL PROCEDURES

Bacterial strains

Geminicoccus roseus ATCC BAA-1445 (*G. roseus)* and *Xanthomonas pisi* ATCC 35936 (*X. pisi*) were obtained from the American Type Culture Collection in Manassas, VA. *G. roseus* strain was isolated from a marine aquaculture system in Germany (Foesel et al., 2007), and *X. pisi* strain was isolated from *Pisum sativum* in Japan (Vauterin et al., 1995). Unless otherwise stated, *G. roseus* cultures were grown on marine agar (Difco) at 30 ^{o}C, and *X. pisi* cultures grown on nutrient agar (Difco) at 25 ^{o}C. Liquid cultures were maintained in 125 mL of either marine broth or nutrient broth, shaking at 250 rpm.

Glycosyl Residue Composition Analysis

A 5 to 7-day-old culture (30 ml) was centrifuged (10,000 g, 5 min, 4 °C) and cell pellet was suspended in 10 volumes of cold MeOH:chloroform:H_2O (40:40:20, v/v/v). The suspension was transferred to a 15 ml falcon tube and vortexed for 10 min (30 sec every 2 min, 4 °C). The suspension was centrifuged (10,000*g*, 5 min, 4 °C) and separated into a top methanolic water phase (termed methanolic), medial interphase (termed pellet) and bottom organic chloroform phase. A portion (20 µl) of the top methanolic fraction was analyzed on HILIC-ESI-MS/MS (see below) and the remainder transferred to a 13 mm borosilicate tube. The bottom chloroform fraction was transferred to a separate tube. The remaining interphase was resuspended in 2 ml DDW and samples centrifuged (10,000 g 5 min, 4 °C). Supernatant was vacuum aspirated, the pellet was again resuspended in 1 ml DDW and transferred to a new 13 mm borosilicate tube.

The methanolic and organic solvent extracts or cell pellets (~1 mg) were supplemented with myo-Inositol (10 µl of 5 mM solution) as an internal standard, evaporated to dryness at

room temperature using a stream of air (REACTIVAP III, Thermo Fisher) and then hydrolyzed for 1 h at 120 °C with 1 ml of 1 M TFA. TFA was removed by evaporation under a stream of air (40 °C) and the residue washed with isopropanol (3 x 500 µl). The released monosaccharides were then converted into their corresponding alditol-acetate derivatives according to York *et al* (York et al., 1986), and the final residue dissolved in acetone (100 µl).

A fraction (1 µl) of each of the alditol-acetate derivative samples was analyzed by gas-liquid chromatography (GLC, Agilent 7890A) equipped with a mass selective detector (EI-MS, Agilent 5975C) and separated over a Restek RTx-2330 fused silica column as previously described (Smith et al., 2016). Alditol-acetate derivatives of standard apiose, rhamnose, fucose, ribose, arabinose, xylose, mannose, glucose, and galactose (50 µg each) were prepared under the same conditions as samples. Monosaccharides were identified based on their retention times and their electron impact (EI) mass-spectra.

Identification and cloning of CeUAS, GrUAS and XpUAS

The BLASTP program (Altschul et al., 1997) and BLAST Link (Blink) tool were probed to identify bacterial proteins (taxid: 2) in the NCBI non-redundant data base that share amino acid sequence homology to the *Arabidopsis* AXS1/UAS1 (At2g27860). Analyses of hypothetical proteins belonging to the proteobacteria *Candidatus entotheonella*, *Geminicoccus roseus* and *Xanthomonas pisi* revealed that they share homology to UXS. The predicted protein for XpUAS lacked 38 amino acids at the N-terminal region based on sequence alignment with other UASs. To obtain the entire XpUAS ORF, the *X. pisi* whole genome shotgun sequence (NZ_JPLE01000032) was used to extend the nucleotide sequence to include 114 nucleotides upstream of the predicted transcript. The nucleotide sequences corresponding to the *C.*

entotheonella and *G. roseus* proteins and the extended full-length nucleotide sequence for predicted XpUAS were used for primer design and cloning.

Genomic DNA (gDNA) was isolated from 5-day-old liquid cultures (3 ml) of *G. roseus* and *X. pisi*. Cells were pelleted (14,000 rpm, 1 min) and 200 µl extraction buffer (0.2 M Tris pH 8, 0.25 M NaCl, 25 mM EDTA, 1% SDS) added, and the samples were vortexed for 2 min. Samples were spun down (12,000 rpm, 5 min), and 150 µl of supernatant was transferred to a new tube. An equal volume of isopropanol was added, and precipitated gDNA was pelleted (12,000 rpm, 5 min). Supernatant was vacuum aspirated and samples were allowed to dry under laminar flow hood for 20 min. gDNA was re-suspended in 100 µl TE (10 mM Tris pH 8, 1 mM EDTA) and stored at 4 °C.

A portion gDNA (2 µl), dNTP's, 1 unit of Phusion® high-fidelity DNA polymerase (New England Biolabs; Ipswich, MA) with 0.2 µM of each forward and reverse primers (IDT; Coralville, IA; **Table 3.S2**) were used to amplify GrUAS and XpUAS the following thermal cycler conditions: one 98 °C denaturation cycle for 30 s followed by 25 cycles (each of 8-s denaturation at 98 °C; 25-s annealing at 60 °C; 30-s elongation at 72 °C), and finally termination at 4 °C. The PCR product was directly cloned into the *E. coli* expression vector pET28b modified to contain an N-terminal His_6 tag followed by a TEV cleavage site (Yang et al., 2009).

Because no axenic monoculture of *C. entotheonella* was available, a synthetic ORF gene corresponding to the nucleotide sequence of CeUAS was obtained (GenScript; Piscataway, New Jersey, USA). The ORF was cloned into the modified pET28b expression vector (Yang et al., 2009), using forward and reverse primers (see **Table 3.S2**).

Following cloning of the individual UAS genes, the plasmids were sequence verified (Georgia Genomics Facility; Athens, GA) and termed, pET28b-TEV-CeUAS.1, pET28b-TEV-

GrUAS.1 and pET28b-TEV-XpUAS.1. Their amino acid sequences were deposited in GenBank™ (accession numbers MF191704, MF191705 and MF191706).

Analysis of nucleotide sugars produced in microbe

NDP-sugars from *E. coli* harboring the expression plasmids were harvested as described (Smith et al., 2016; Yang et al., 2012). BL21-derived *E. coli* cells (3 or 60 ml) co-transformed with pCDFDuet-BtbDH and either pET28b-TEV-CeUAS.1, pET28b-TEV-GrUAS.1, pET28b-TEV-XpUAS.1 or empty pET28b vector control were grown in LB medium [1.0% (w/v) Bacto tryptone, 0.5% (w/v) Bacto yeast extract, and 1.0% (w/v) NaCl] supplemented with 35 μg/ml chloramphenicol, 50 μg/ml kanamycin, and 25 μg/ml spectinomycin at 37 °C and 250 rpm, induced with Isopropyl β-D-thiogalactoside (IPTG, 0.5 mM) at an OD_{600nm} of 0.6 and grown at 30 °C for 4 h. In microbe nucleotide sugars were extracted and analyzed by hydrophilic interaction liquid chromatography–electrospray ionization–tandem mass spectrometry (HILIC-ESI-MS/MS) as described (Smith et al., 2016).

His_6-tagged protein expression and purification

BL21-derived *E. coli* cells were transformed with pET28b-TEV-CeUAS.1, pET28b-TEV-GrUAS.1, pET28b-TEV-XpUAS.1, or the empty vector control. The cells were grown at 37 °C and 250 rpm for 16 h in 20 ml LB media containing kanamycin (50 μg/ml) and chloramphenicol (35 μg/ml). A portion (5 ml) of the culture was transferred to fresh 245 ml LB supplemented with antibiotics and grown under the same conditions until its OD_{600nm} was 0.8. IPTG (0.5 mM) was then added to induce expression of the gene, and the culture then grown for an additional 4 h at 30 °C and 250 rpm. The induced cultures were cooled on ice and centrifuged

(6,000 g, 10 min, 4 °C). The cell pellet was suspended in 10 ml lysis buffer [50 mM Tris-HCl pH 7.6, 10% (v/v) glycerol, 1 mM ethylenediaminetetraacetic acid, 5 mM dithiothreitol 0.5 mM phenylmethylsulfonyl fluoride]. The cells were ruptured by sonication and proteins were then isolated after centrifugation as described (Yang et al., 2009). The final soluble protein fraction (Fraction S20) was collected and kept on ice prior to immediate purification.

The different His_6-tagged proteins, including control empty plasmid, were each purified using fast-flow Ni-Sepharose (GE Healthcare, 2 ml resin packed in a 15 x 1 cm polypropylene column) as previously described (Smith et al., 2016), where the purified His_6-bUAS-protein eluted in fraction E7 and had activity. The active enzymes were dialyzed (6,000 – 8,000 molecular weight cut-off, Spectrum Laboratories, Inc) at 4 °C three times for a total of 2 h against 50 mM Tris-HCl pH 7.6, containing 0.15 M NaCl, 10% (v/v) glycerol, 1 mM DTT, and 10 µM NAD^+. The dialysates were divided into 150 µl aliquots, flash-frozen in liquid nitrogen and stored at -80 °C. Aliquots of purified protein were assayed for activity and analyzed on SDS-PAGE.

SDS-PAGE was performed with 12% (w/w) polyacrylamide gels. Proteins were stained with 0.1% (w/v) Coomassie Brilliant Blue R-250 in aq. 20% methanol (MeOH) containing 7% (v/v) acetic acid and de-stained with aq. 20% methanol containing 7% (v/v) acetic acid. Protein concentrations were determined with the Bradford reagent (Bradford, 1976) using bovine serum albumin (BSA) as standard, and the molecular mass of active recombinant GrUAS was estimated by size-exclusion chromatography as previously described (Smith et al., 2016).

Recombinant bUAS Enzyme Assays

Unless otherwise indicated, the 50 µl reactions were performed in 50 mM Tris-HCl pH 7.9, containing 1 mM NAD^+, 1 mM UDP-GlcA, and up to10 µg of purified protein. The assay mixtures were incubated at 37 °C for up to 45 min and the reactions terminated by placing the tubes in boiling water for 2 min followed by the addition of an equal volume of chloroform. The suspensions were vortexed and centrifuged (12,000 g, 5 min, 22 °C), and the aqueous phase analyzed for nucleotide sugars. ^1H NMR assays were performed in deuterium oxide (D_2O) using 30 µg purified protein in final volume of 180 µl.

Characterization of Recombinant GrUAS

GrUAS activity was assayed in different buffers, at different temperatures and with various additives and nucleotide sugars. For pH studies, reactions in total volume of 50 µl consisted of purified recombinant GrUAS (10 µg), 1 mM NAD^+, 1 mM UDP-GlcA and various pH buffers (100 mM) and kept at 37 °C for 30 min. Inhibition assays were performed by first supplementing the standard reaction mixtures with 2 mM of various nucleotides and nucleotide sugars, addition of purified protein and incubation. The amounts of reactants and products were determined by UV spectroscopy and used to calculate enzyme activity as follows. The products from each recombinant enzyme assay were chromatographed over a column (200 x 1 mm) packed with 15 µm Source 15Q anion exchange resin (GE Healthcare, Pittsburgh, PA) by elution with a linear gradient (5 mM to 0.6 M) of ammonium formate over 25 min at a flow rate of 0.25 ml/min using an Agilent (Santa Clara, CA) 1100 Series HPLC equipped with an G1313A auto-sampler, a G1315B diode array detector, and ChemStation software. Nucleotides and nucleotide sugars were detected by their A_{261nm} (for UDP-sugars) and A_{259nm} (for NAD^+). The

concentrations of reactants and products were determined by comparison of their peak areas to a calibration curve of standard UDP-GlcA (Gu et al., 2010).

Selected kinetic parameters of recombinant GrUAS (10 µg) were determined by varying the concentrations of UDP-GlcA in 50 µl reactions consisting of 1 mM NAD^+ in 50 mM Tris-HCl pH 7.9. Reactions were kept for 7 min at 37 °C quenched with an equal volume of chloroform and then vortex mixed. The reaction products in the aqueous phase were separated using a Q-15 anion exchange column as described above and reaction rates calculated from the depletion of the UDP-GlcA signal integral normalized to the NAD^+ signal integral. Values from three independent replicates were used to generate a non-linear regression plot and resultant data using GraphPad Prism Version 7.

HILIC-ESI-MS/MS

ESI-MS/MS analysis was performed on a Shimadzu (Kyoto, Japan) LC-ESI-MS-IT-TOF operating in the negative ion mode. Methanolic extracts, in microbe and recombinant enzyme assay products were mixed with 2/3 volume aq. 95% acetonitrile (ACN) containing 25 mM ammonium acetate and an aliquot (10 - 20 µl) chromatographed over an Accucore amide-HILIC column (150 x 4.6 mm; Thermo), eluted at 0.4 ml min^{-1} with a linear gradient of aq. 75% (v/v) acetonitrile containing 40 mM ammonium acetate, pH 4.4, to 50% (v/v) acetonitrile containing 40 mM ammonium acetate, pH 4.4, over 35 min using a Shimadzu LC-30AD HPLC. Mass spectra (mass range 100-2,000 *m/z*) were collected every 1.3 sec for 30 minutes. Second stage MS/MS data was collected by collision-induced dissociation (CID) with a collision energy of 35% and a nebulizing nitrogen gas flow of 1.5 ml min^{-1} (59).

Real-time ^{1}H and 2-D HSQC NMR Enzyme Assays

All spectra were obtained using a Varian Inova 600 MHz spectrometer equipped with a 3 mm cryogenic probe. Continuous ^{1}H and 2-D HSQC NMR spectroscopic monitoring of reactions (180 µl volume) were carried out at 37 °C in a mixture of D_2O/H_2O (9:1, v/v) containing 0.83 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, internal reference), 50 mM Tris-HCl, pH 7.9, 1 mM UDP-GlcA or $^{13}C_6$-labeled UDP-GlcA, 1 mM NAD^+ and purified recombinant enzyme. One-dimensional ^{1}H and 2-D HSQC NMR spectra with the water resonance signal pre-saturated were collected 5 minutes post addition of enzyme in order to optimize spectrometer settings, and then spectra were continuously averaged every 2.5 min for up to 8 h. All chemical shifts are referenced to DSS at 0.00 ppm (Guyett et al., 2009).

RNA isolation and RT-PCR

RNA was extracted (Z. Li et al., 2016) from a 5-day-old culture of *G. roseus* and a 2-day-old culture of *X. pisi* grown in liquid media. A 3 ml portion of liquid culture was centrifuged (14,000 rpm, 1 min, 22 °C); supernatant discarded; and the cells were flash frozen in liquid nitrogen and stored at -80°C until extraction. Cell pellets were resuspended in 400 µl TE-lysozyme (20 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mg/ml lysozyme; Sigma L6876), vortexed at room temperature for 10 min; and extraction was carried out after addition of 40 µl of fresh 10X EB (0.3 M NaOAc, pH 5.2, 5% sarkosyl, w/v, 50 mM EDTA, 10 % β-mercaptoethanol, v/v) and incubation at 65 °C and mixing for 3 min. Subsequently, 440 µl of preheated acidic phenol was added, and samples incubated at 65 °C while mixed by vortex for 7 min. Sample was then placed on ice for 3 min and centrifuged (10,000 g 5 min, 4 °C). The top 350 µl of the aqueous phase was transferred to a new tube and an equal volume of chloroform added. Sample was vortexed for 1

min; incubated at room temperature for 7 min; and then centrifuged (14,000 rpm, 5 min, 22 °C). The top 200 µl of the aqueous phase was transferred to a new tube mixed with 200 µl of cold isopropanol and placed at -20 °C overnight. Sample was then centrifuged (10,000 g, 10 min, 4 °C) and the pellet was resuspended in 75% ethanol, again centrifuged (10,000 g, 10 min, 4 °C) and supernatant aspirated. Tubes were left open under laminar flow hood for 15 min to dry. Resulting nucleic acids were resuspended in 40 µl sterile, deionized distilled water (DDW) and RNA concentration measured with a nanodrop (Thermo Fisher Scientific; Waltham, MA). To digest remnant genomic DNA, 2 µg of RNA was DNAse treated according to manufacturer guidelines (Thermo). Following DNA digest (37 °C for 30 min) an equal volume of chloroform was added and mixed. Sample was centrifuged (14,000 rpm, 5 min, 22 °C) and top aqueous phase transferred to a new tube. RNA (0.5 µg) was then reverse transcribed with a random hexamer primer (Thermo) using SuperScript III reverse transcriptase (Life Technologies; Carlsbad, CA). A portion of the reverse transcriptase (RT) reaction (2 µl), dNTP's, 1 unit of Phusion® high-fidelity DNA polymerase (New England Biolabs; Ipswich, MA) with 0.2 µM of each forward and reverse primers (IDT; Coralville, IA; **Table 3.S2**) were used to amplify GrUAS, XpUAS, *G. roseus* sigma factor RpoD (GrRpoD) and *X. pisi* sigma factor 70 (XpSig70) with the following thermal cycler conditions: one 98 °C denaturation cycle for 30 s followed by 25 cycles (each of 8-s denaturation at 98 °C; 25-s annealing at 60 °C; 30-s elongation at 72 °C), and finally termination at 4 °C.

TABLES AND FIGURES

TABLE 3.1.

Enzymatic properties of recombinant GrUAS.

	GrUAS
Optimal pH[a]	7.7 - 8.1
Optimal Temperature (°C)[a]	37 – 42
K_m (μM)[b]	251 ± 44.6
V_{max} (nM-s^{-1})	70.4 ± 4.53
k_{cat} (s^{-1})	15.1 ± 0.97
k_{cat} /K_m (nM-s^{-1})	60.2 ± 7.05
Mass of Active Protein/Dimer (kDa)[c]	(84.3)

[a]Optimal pH was determined using different buffers (Tris-HCl, sodium phosphate, & HEPES; pH range from 5.3 to 10.3). Temperature assays were conducted in Tris-HCl buffer pH 7.9. [b]The data presented are calculated from the average of three experiments. [c]The mass of active GrUAS eluted from Superdex75 gel filtration column (10.53 min) was estimated based on extrapolation of standard protein marker.

TABLE 3.2.

Effect of nucleotide sugars and nucleotides on recombinant GrUAS activity.

Additive	Relative activity (%)[a]
Water (Control)	100 ± 4.8
UDP-glucose	89.1 ± 5.0
UDP-galactose	98.5 ± 4.2
UDP-arabinose	88.5 ± 5.6
UDP-xylose	91.4 ± 5.6
UDP-galacturonic acid	33.4 ± 9.5
UDP	72.6 ± 5.5
UMP	92.6 ± 5.0
GDP	85.2 ± 7.9
GMP	94.3 ± 3.7
CDP	85.3 ± 4.6
CMP	93.8 ± 4.5
TDP	85.0 ± 4.8
TMP	87.9 ± 4.9
NADP	79.6 ± 3.3
NADPH	82.6 ± 0.9
NADH	105.3 ± 1.9

[a]Amounts of unreacted UDP-GlcA were determined by HILIC-HPLC. The activity is calculated as the average relative amount of UDP-GlcA consumed compared to the control from three experiments.

FIGURE 3.1

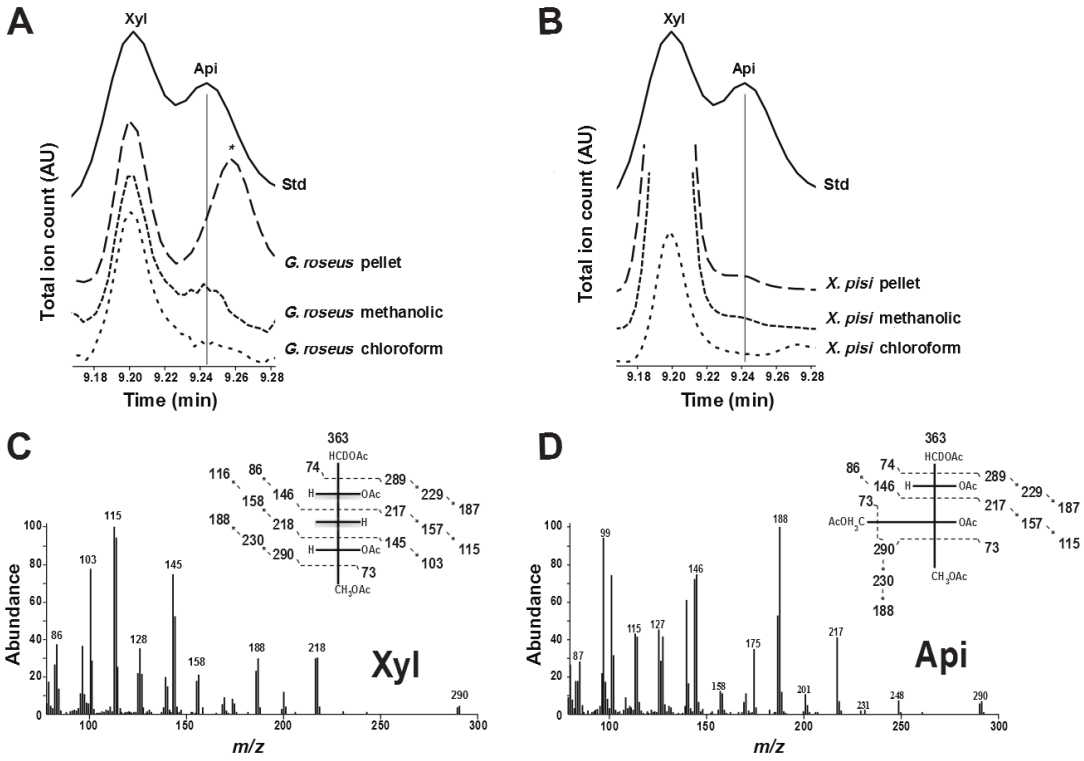


FIGURE 3.1. GC-MS analysis of alditol-acetate derivatives from methanolic, chloroform and cell pellet fractions of *G. roseus* (**Panel A**) and *X. pisi* (**Panel B**). Standard (Std) contains authentic xylose and apiose. The region of the total ion count for xylose- (Xyl) and apiose- (Api)- alditol-acetate derivatives is expanded. **Panels B** and **C** are MS fragmentation patterns for standard Xyl and Api, respectively. * indicates unidentified residue.

FIGURE 3.2

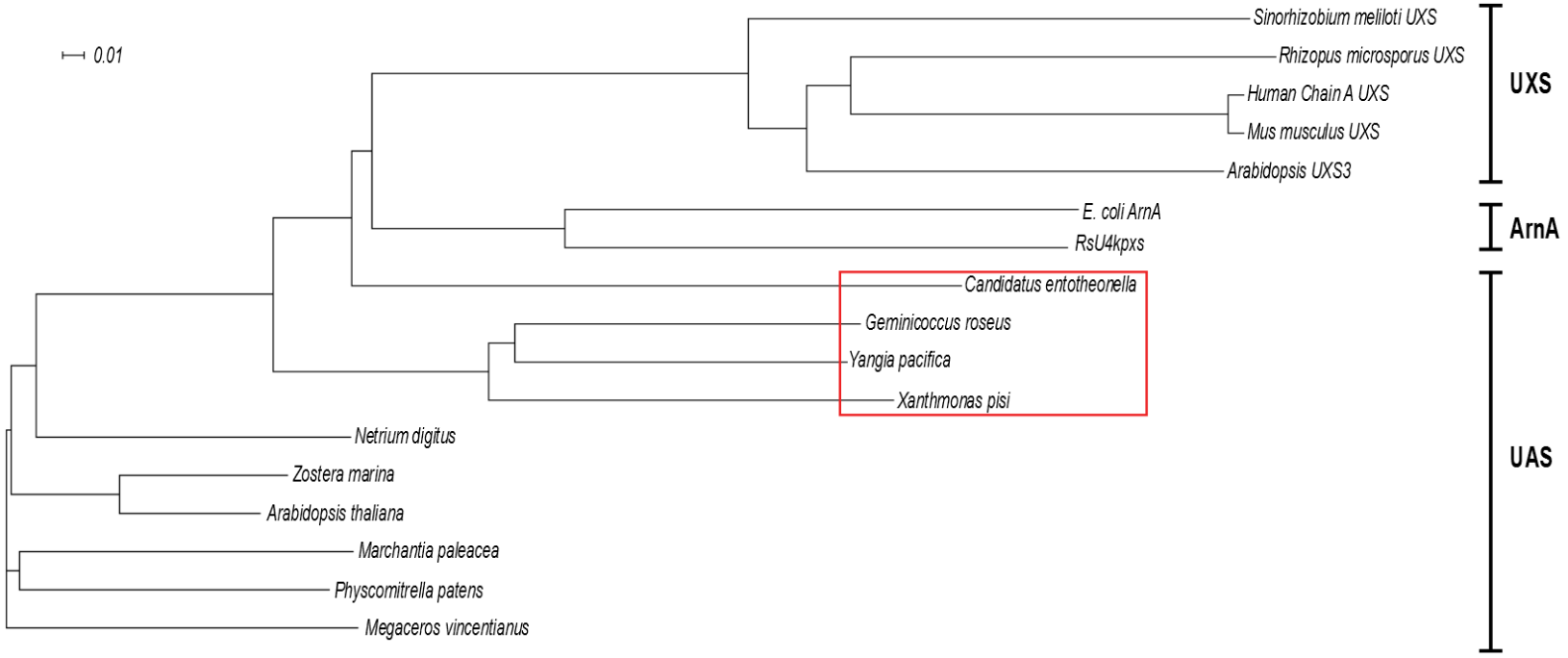


FIGURE 3.2. Phylogenetic analysis of proteins involved in the synthesis of UDP-apiose (UAS) and UDP-xylose (UXS). Amino acid sequences used are the C-terminal region of *Escherichia coli* ArnA (WP_032205568.1) that form UDP-4-keto-arabinose, *Ralstonia solanacearum* UDP-4-keto-pentose/UDP-xylose synthase (RsU4kpxs, WP_011001268.1), UXSs from bacteria (*Sinorhizobium meliloti*, ACY30251.1) mammal (human & *Mus musculus*, NP_079352.2 & NP_080706.1), fungi (*Rhizopus microspores*, CEI96046.1) and plant (*Arabidopsis* UXS3; NP_001078768.1). The bacterial UAS-like sequences used are from *Candidatus entotheonella*, *Geminicicccus roseus*, *Xanthomonas pisi* and *Yangia pacifica* (ETX00953.1, WP_084506503.1, WP_084725965.1 and

WP_066111466.1). Other UASs used are from green algae (*Netrium digitus,* AOG75413.1), from hornwort (*Megaceros vincentianus*, AOG75412.1), from liverwort (*Marchantia paleacea*, AOG75410.1) from moss (*Physcomitrella patens,* AOG75414.1), and from angiosperms (*Arabidopsis thaliana & Zostera marina*; KMZ68719.1 & NP_180353.1, respectively. Bacterial UAS are outlined by a red box. Alignment was made using Clustal Omega (W. Li et al., 2015; McWilliam et al., 2013; Sievers et al., 2011) and the tree generated using Dendroscope (Huson & Scornavacca, 2012).

FIGURE 3.3

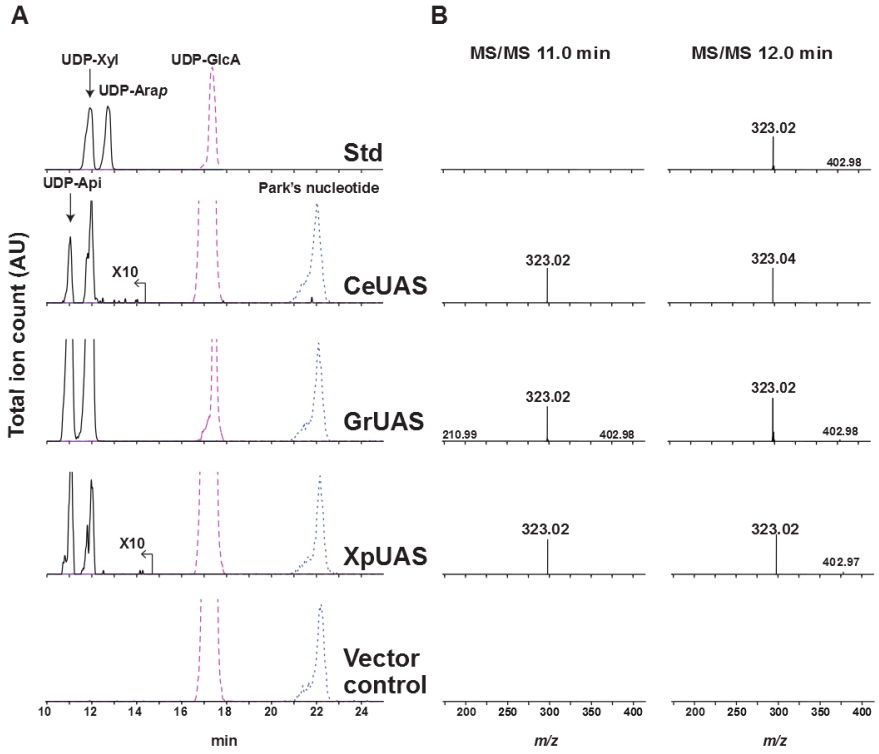


FIGURE 3.3. Activity of recombinant bacterial UDP-apiose synthase (bUAS) by in microbe assay. Analysis of in microbe nucleotide sugars by HILIC-LC-ESI-MS/MS. **Panel A,** top panel elution of standard (Std): UDP-GlcA, UDP-Xyl and UDP-arabinopyranose (UDP-Ara*p*); Nucleotide sugars were extracted from *E. coli* cells induced to express genes encoding CeUAS,

GrUAS, XpUAS or empty vector as the control (bottom panel). $[M-H]^-$ ions diagnostic for UDP-pentose (*m/z* 535.0, solid line), UDP-hexuronic acid (*m/z* 579.0, dotted line) and Park's nucleotide (*m/z* 595.6, dashed line) are displayed. Park's nucleotide is a UDP-MurNAc-pentapeptide that is used as an internal standard for nucleotide-sugar detection as it is abundantly made in *E. coli*. The *m/z* signal for CeUAS and XpUAS is amplified by a factor of 10. **Panel B**; Second stage MS fragmentation data for the parent *m/z* 535.0 peaks at the indicated retention times; Left column 11.0 min and right column 12.0 min. MS/MS ions at *m/z* 323.0, 211.0, 403.0 are consistent with predicted fragmentation of a UDP-sugar into $[UMP-H]^-$, $[Ura-2H]^-$, and $[UDP-H_2O-H]^-$, respectively.

FIGURE 3.4

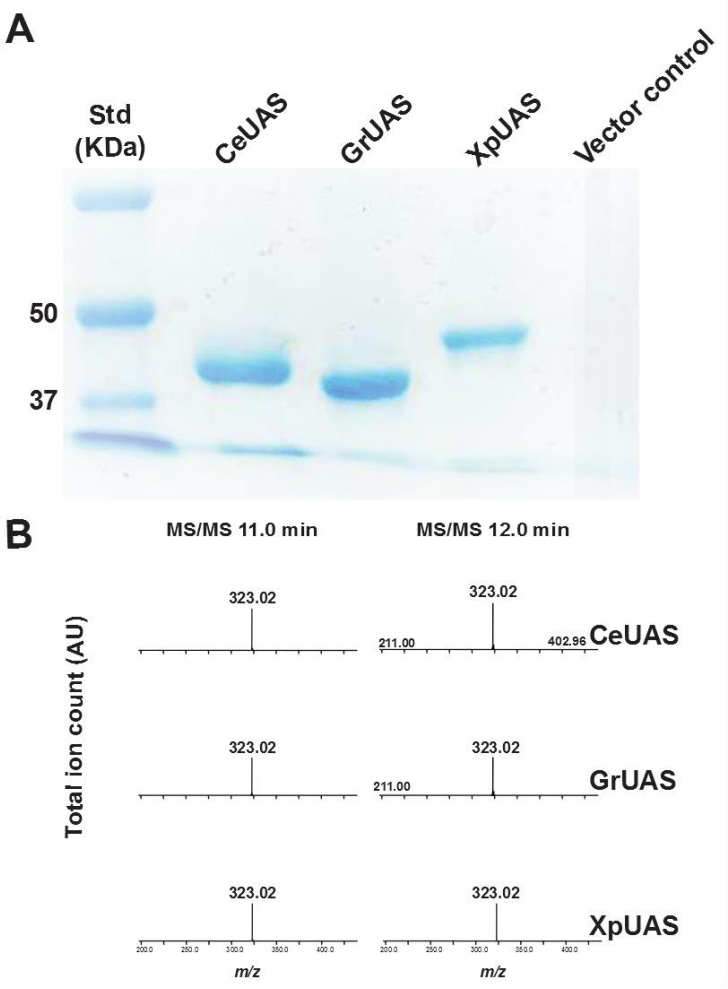


FIGURE 3.4. Activity of purified recombinant bUAS proteins. Total proteins from *E. coli* cells induced to express CeUAS, GrUAS, XpUAS and empty vector control, were isolated and the His_6-tagged-proteins were purified by Ni-column (**Panel A**) and eluted at fraction E7. The recombinant proteins were tested for UAS activity by separating the reaction mixture by HILIC column. UAS activity was determined by its ability to form UDP-apiose and UDP-xylose peaks (both at *m/z* 535.0); detection of the peaks was carried out by MS/MS (**Panel B**). **Panel A** shows the expected sizes of CeUAS, GrUAS and XpUAS: 44.2, 42.9 and 45.1 kDa, respectively. **Panel B**; shows the MS/MS for the *m/z* 535.0 peaks of UDP-Api (left column, elution time 11.0 min) and UDP-Xyl (right column, 12.0 min). MS/MS ions at *m/z* 323.0, 211.0, 403.0 are consistent

with predicted fragmentation of a UDP-sugar into $[UMP-H]^-$, $[Ura-2H]^-$, and $[UDP-H_2O-H]^-$, respectively.

FIGURE 3.5

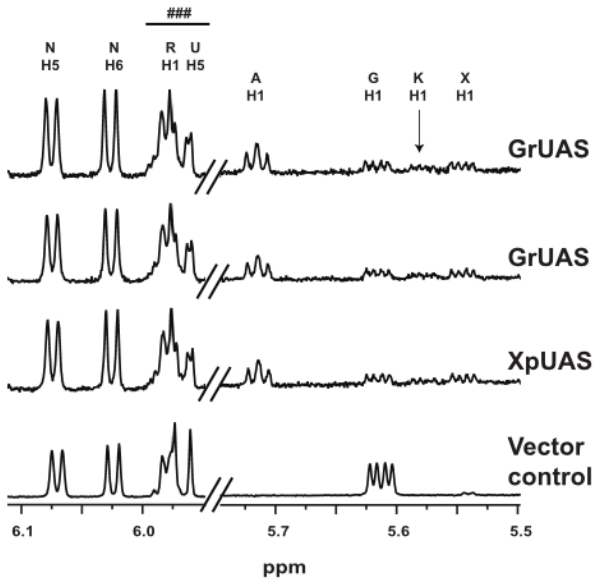


FIGURE 3.5. ^1H NMR spectra of purified recombinant bUAS reactions**.** Selected regions of ^1H NMR spectra diagnostic for the products and intermediates generated by incubation of UDP-GlcA and NAD^+ with the purified recombinant UASs from bacteria. Anomeric region between 5.50 and 5.75 ppm for the H1 protons of UDP-GlcA (G), UDP-Api (A) and UDP-Xyl (X) products and UDP-4-keto-Xyl (K) intermediate are shown. NMR region (5.95 and 6.15 ppm) diagnostic for UDP and NAD^+ cofactor is included. NMR spectral traces from top to bottom show UAS activity of CeUAS, GrUAS, XpUAS and empty vector control. Peaks labeled N correspond to H5 and H6 protons of NAD^+. ### indicates a mixture of ribose (R) and uracil (U) proton peaks of UDP from substrate and products. For additional chemical shift assignments see **Table 3.S1**.

FIGURE 3.6

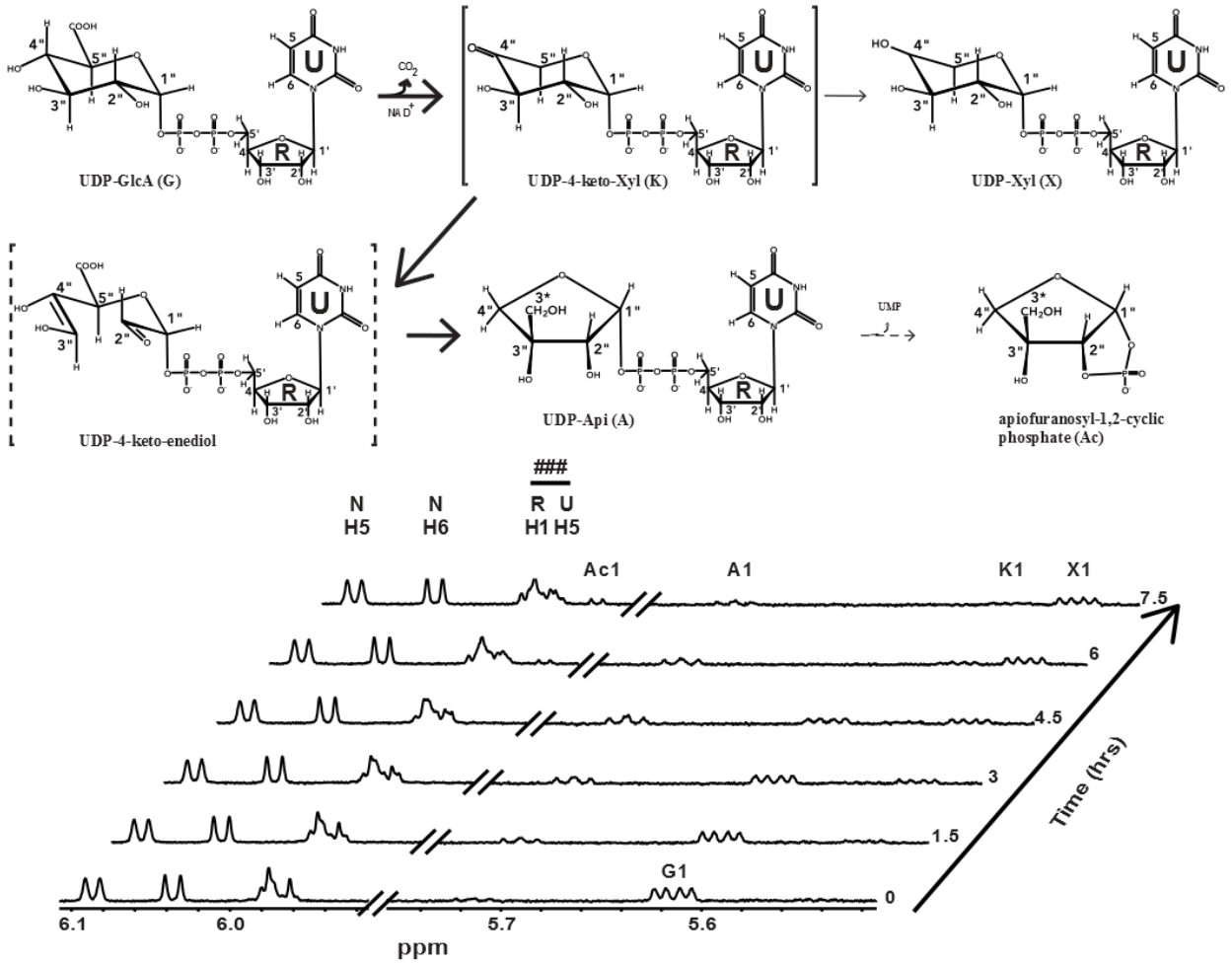


FIGURE 3.6. UAS reaction mechanism based on real time NMR analysis of recombinant GrUAS activity. NMR spectra of the UDP-apiose synthase activity at 37 ^{o}C showing conversion of the substrate (UDP-GlcA), to intermediate (UDP-4-keto-Xyl), products (UDP-Api and UDP-Xyl ~ 2:1 ratio), and degradation product (apiofuranosyl-1,2-cylcic phosphate, Ac). The proton NMR spectrum of the sugar anomeric regions (H-1s, between 5.5 and 6.1 ppm) of substrate, intermediate and products is shown. Only select time-resolved spectra are displayed to prevent overcrowding of peaks. ### indicates a mixture of ribose (R) and uracil (U) proton peaks of UDP from substrate and products. For additional chemical shift assignments see **Table 3.S1**.

FIGURE 3.7

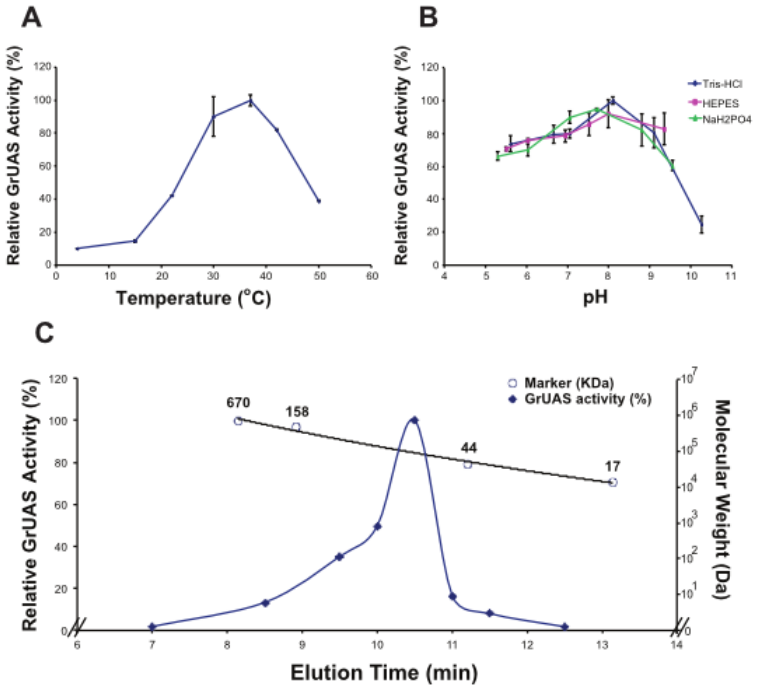


FIGURE 3.7. The effects of temperature and pH on the activity of recombinant GrUAS. **Panel A;** maximum activity of GrUAS is at 37 °C. **Panel B;** maximum activity of GrUAS is in Tris-HCl at a pH of 8.1. **Panel C;** Size-exclusion chromatography suggests recombinant GrUAS is active as dimer. GrUAS was fractionated on a Superdex-75 SEC column. Fractions were collected every minute and assayed for UAS activity. The relative activity (indicated by closed diamonds) was determined by HPLC. The molecular weight of the enzyme in solution is based on the relative elution times of standard protein markers (indicated by open circles).

FIGURE 3.8

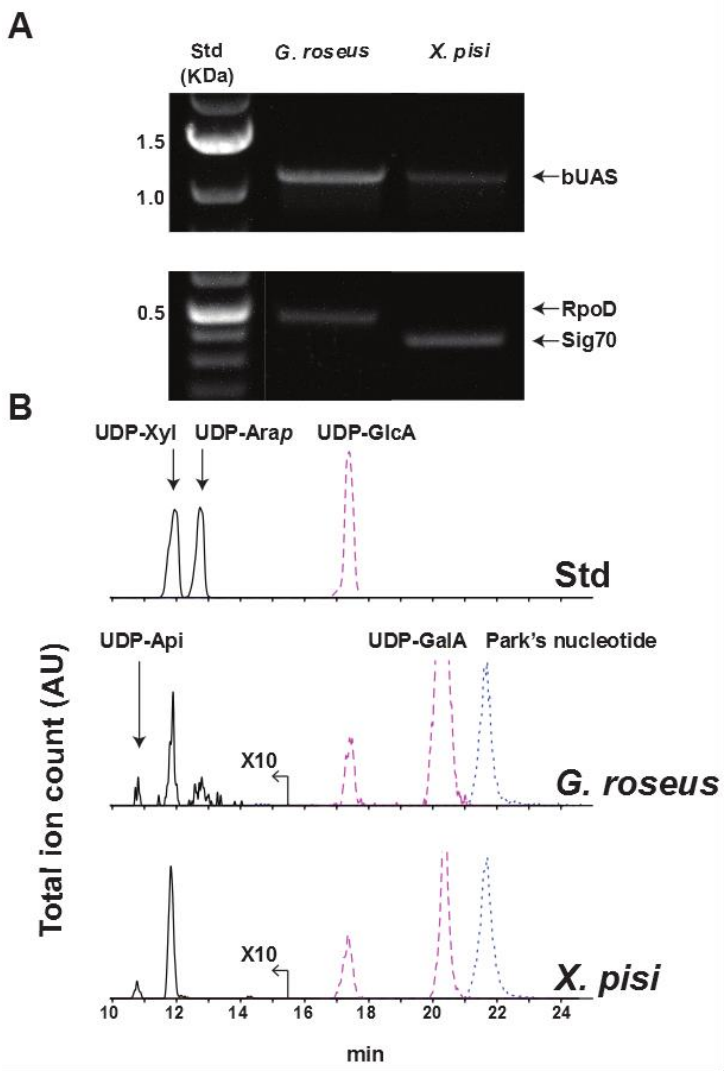


FIGURE 3.8. *In vivo* indication for the functional activity of bUAS genes and enzymes. **Panel A**; RT-PCR analyses showing bUAS transcripts of *G. roseus* and *X. pisi*. Transcripts were amplified by PCR and a portion of the products were separated by agarose and EtBr stained. **Panel B**; LC-MS analysis of aq-methanolic (MeOH:chloroform:H_2O; 40:40:20, v/v/v) extracts. Negative mode $[M-H]^-$ ions diagnostic for UDP-pentose (*m/z* 535.0, solid line; amplified by a factor of 10), UDP-hexuronic acid (*m/z* 579.0, dotted line) and Park's nucleotide (*m/z* 595.6, dashed line) are displayed.

SUPPLEMENTAL INFORMATION

TABLE 3.S1. Chemical shifts and coupling constants for the protons of UDP-Api, UDP-Xyl and UDP-4-keto-Xyl formed from UDP-GlcA by recombinant GrUAS, and the apiofuranosyl-1,2-cyclic phosphate that is formed by spontaneous degradation of UDP-Api

	H1	**H2**	**H3*$_{a,b}$**	**H4**	**H5**
UDP-α-D-Api (A)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.71 *Quartet* $J_{1",2"}$ 4.5 $J_{1",P}$ 5.6	4.01 *Quartet* $J_{1",2"}$ 4.5 $J_{2",P}$ 2.2	*3.5, 3.6 $J_{3a",3b"}$ 12.3	4.08, 4.09 $J_{4a",4b"}$ 4.1	
UDP-α-D-Xyl (X)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.54 *Quartet* $J_{1",2"}$ 3.4, $J_{1",P}$ 7.0	3.51 *Doublet* $J_{2",3"}$ 3.5	3.74 *Doublet* $J_{3",4"}$ 9.5	3.67	3.78
UDP-4-keto-Xyl (K)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.57 *Quartet* $J_{1",2"}$ 3.5 $J_{1",P}$ 7.0	3.94 *Doublet* $J_{2",3"}$ 5.5	*3.81 *Doublet* $J_{3",4"}$ 9.8		3.53
Apiofuranosyl-1,2-cyclic phosphate (Ac)					
Chemical shifts, δ (ppm[a]), *peak shape* J coupling constants (Hz)	5.93 *Doublet of doublets* $J_{1",2"}$ 4.3 $J_{1",P}$ 15.1	4.56 *Doublet* $J_{2",P}$ 7.0	3.56, 3.6 $J_{3a",3b"}$ 12.3	3.86, 3.94 $J_{4a",4b"}$ 9.4	
Ribose (R)					
Chemical shifts, δ (ppm[a]), J coupling constants (Hz)	5.97 $J_{1',2'}$ 3.6	4.35	4.34	4.26	4.19, 4.23 $J_{5a',5b'}$ 12
Uracil (U)					
Chemical shifts, δ (ppm[a]) J coupling constants (Hz)					5.96 $J_{5,6}$ 8.1

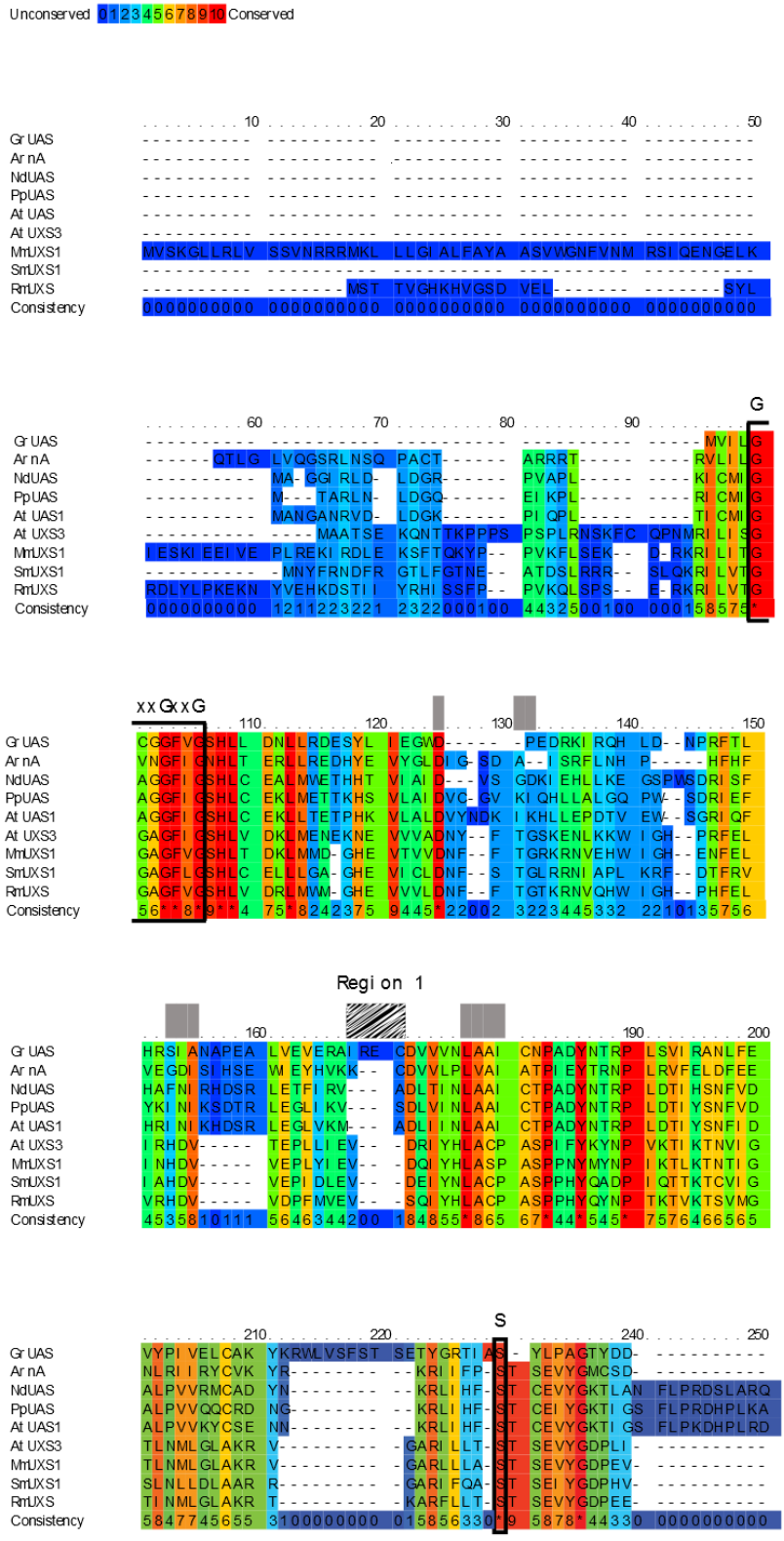
[a] Chemical shifts are in ppm relative to internal DSS signal set at 0.00 ppm. Proton-proton coupling constants in Hz are shown as well as the $J_{1",P}$ coupling values between phosphate and the H1" proton of UDP-Api (A), UDP-Xyl (X) and UDP-4-keto-Xyl (K). The chemical shift values for uracil (U) and ribose (R) protons are similar to the uracil and ribose protons of other

UDP-sugars. Compounds correspond to structures in **Figure 3.7**. * refers to the proton of the exocyclic carbon in UDP-Api and apiofuranosyl-1,2-cyclic phosphate (Ac). Peak assignments were made according to known values (Gu et al., 2011; Guyett et al., 2009).

TABLE 3.S2. Primers used in plasmid generation, genotyping and transcript analysis. Obtained from IDT.

Primer	Sequence (5'-3')
CeAUS_F	GTATTTTCAGGGCGCCATGAAAATCCTGCTGTTAGGTG
CeUAS_R	AGCCGGATCGAATTCACTACGTAAACTGTGGAATCCAAC
GrUAS_F	GTATTTTCAGGGCGCCATGCGTGTGGTCATCCTGGGTTG
GrUAS_R	AGCCGGATCGAATTCATCAGGCGGCCTTCGGCG
XpUAS_F	GTATTTTCAGGGCGCCATGCAGCGAAATCCAATTTCTCAG
XpUAS_R	AGCCGGATCGAATTCATCACTGCGTAGCTTCTTCTG
pET28b_TEV_F	TGAATTCGATCCGGCTGCTAACAAAGCCCG
pET28b_TEV_R	CATGGCGCCCTGAAAATACAGGTTTTC
GrRpoD_F	GCATCCCGGTCCACATGATC
GrRpoD_R	TCAGGTATCCAGGAAGGAGCG
XpSig70_F	GTTGTAGCGATTACCGCCCG
XpSig70_R	GCTCAGGCGCAATTTGGC

FIGURE 3.S1


Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved
GrUAS
ArnA
NdUAS
PpUAS
At UAS
At UXS3
MrUXS1
SmUXS1
RmUXS
Consistency
xxGxxG
Region 1
S

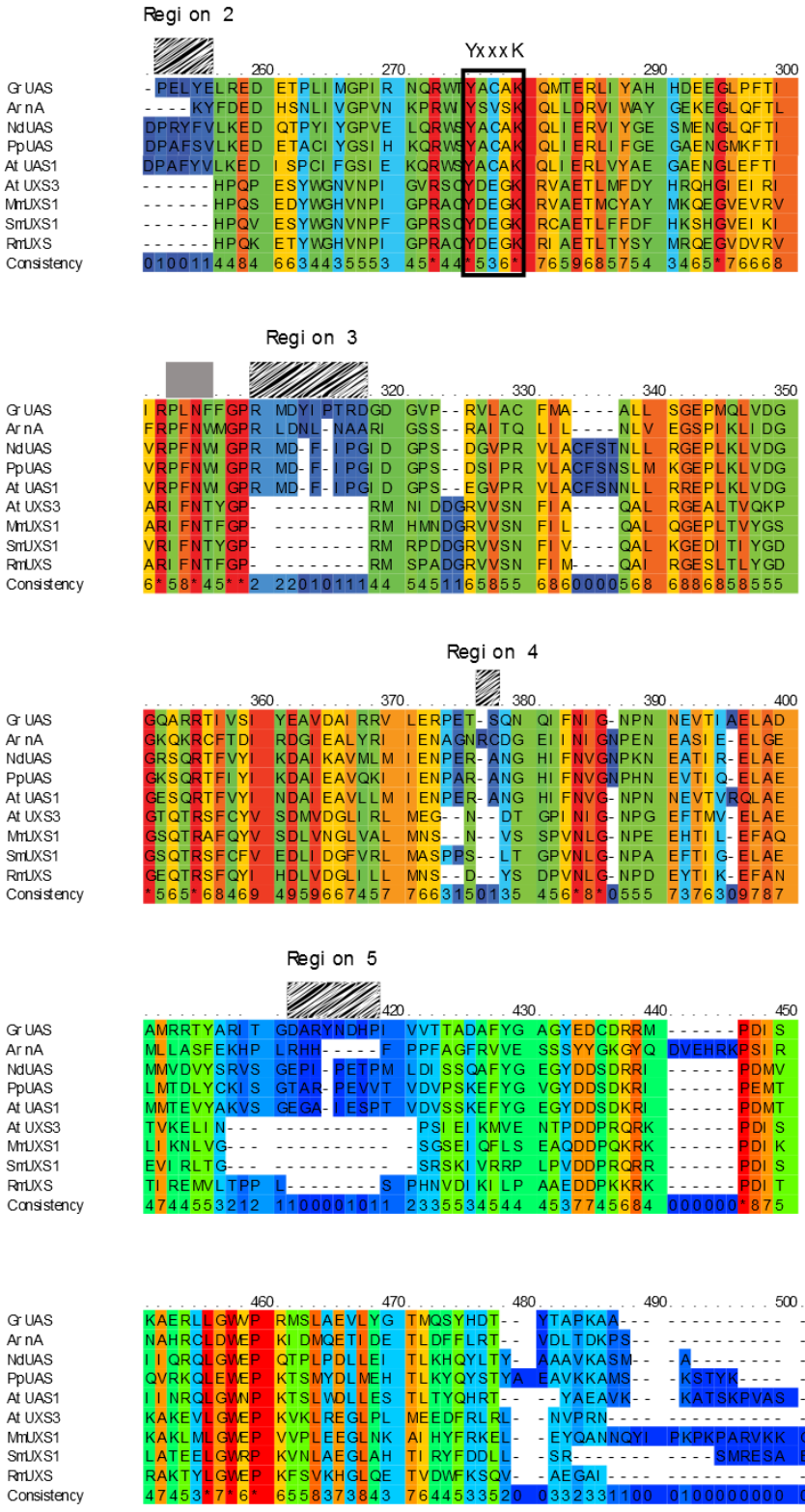

Region 2
YxxxK
Region 3
Region 4
Region 5
GrUAS
ArnA
NdUAS
PpUAS
At UAS1
At UXS3
MnUXS1
SmUXS1
RmUXS
Consistency

FIGURE 3.S1.

UXS, ArnA, and UAS multiple sequence alignment. Full amino acid sequence alignment of the UDP-GlcA decarboxylase domain of *E. coli* ArnA (WP_032205568.1), mouse UXS1 (MmUXS1, NP_080706.1), bacterial UXS1 from *Sinorhizobium meliloti* (SmUXS1, ACY30251.1), fungal UXS from *Rhizopus microsporus* (RmUXS, CEI96046.1), *Arabidopsis* UAS1 & UXS3 (AtUAS1 & AtUXS3, NP_180353.1 & NP_001078768.1), and the UAS from algae *Netrium digitus* (NdUAS, AOG75413.1), moss *Physcomitrella patens* (PpUAS, AOG75414.1), and the bacterial UAS from *Geminicoccus roseus* (GrUAS) were aligned with PRALINE (Simossis & Heringa, 2005) using the BLOSUM62 scoring matrix. Proposed catalytic and cofactor binding residues are outlined in black. Other proposed active-site residues and insertion regions are marked with grey and striped boxes, respectively in the numbering row.

CHAPTER 4

IDENTIFICATION AND PARTIAL CHARACTERIZATION OF AN APIOSYLTRANSFERASE FROM THE PLANT PATHOGEN *XANTHOMONAS PISI*[1]

[1]James Smith, Maor Bar-Peled.

To be submitted to PLOS One.

ABSTRACT

The rare branched-chain sugar apiose, once thought to only be present in the plant kingdom, was recently found in two bacterial species: *Geminicoccus roseus* and *Xanthomonas pisi*. Glycans with apiose residues were detected in aqueous methanol-soluble fractions as well as in the insoluble pellet fraction of *X. pisi*. While the genes encoding bacterial UDP-apiose synthases (bUASs) were characterized in these bacterial species, the enzyme(s) involved in the incorporation of the apiose into glycans remained unknown. In the *X. pisi* genome two genes flanking the bUAS were annotated as hypothetical glycosyltransferase (GT) proteins. The first GT (named XpGTB) has a Leloir type B fold and a conserved putative lipopolysaccharide-modifying (LPS) GT family 90 domain. The second GT (XpGTA) has a type A fold and contains a conserved prokaryotic GT family 2 domain. XpGTA and XpGTB genes were cloned and heterologously expressed. Analysis of nucleotide sugar extracts from *E. coli* expressing XpGTA or XpGTB with UAS showed XpGTA utilized UDP-xylose and XpGTB utilized UDP-apiose as substrate. Indirect activity assay revealed that XpGTB is an apiosyltransferase able to specifically use UDP-apiose. Further support for the apiosyltransferase activity was demonstrated by in microbe co-expression of UAS and XpGTB in *E. coli* showing the utilization of UDP-apiose to generate an apioside detectable in the pellet fraction. This work provides evidence that *X. pisi* evolved the ability to synthesize an apioside of indeterminate function; however the evolution of the bacterial ApiT remains to be determined.

INTRODUCTION

Bacteria produce a large array of glycan structures that are associated with the cell surface. Gram-negative bacteria produce peptidoglycan, lipopolysaccharide (LPS), capsular

polysaccharides (CPS), and some *N*- or *O*- linked glycoproteins (Logan, 2006; Messner, 2004), which have been implicated in cellular recognition of the environment and may be important for vital pathways including adherence, motility, and pathogenesis (Kay et al., 2010). Synthesis of these glycans requires sugar donors, acceptor substrates, and glycosyltransferases (GTs).

GTs are known to have specific substrate specificity and are classified based on conserved three-dimensional fold (Gloster, 2014; Jarrell et al., 2014; Liu & Mushegian, 2003) into 105 different families in the Carbohydrate-Active EnZymes (CAZy) database (http://www.cazy.org) (Campbell et al., 1997; Coutinho et al., 2003). Some GTs necessary to synthesize peptidoglycan have been identified and some crystal structures solved (Ha et al., 2000; Marrec-Fairley et al., 2000; van Heijenoort, 2001), and the steps and required machinery for synthesis of LPS have been extensively characterized in *E. coli* (C. R. Raetz, 1990, 1993; C. R. H. Raetz, 1996; Wyckoff et al., 1998). However, a large number of bacterial glycan structures and GTs that synthesize them remain a mystery. High variability and limited structural data restricts predictability of function, making identification and characterization of GTs responsible for synthesis of bacterial glycan structures challenging.

Xanthomonas is an agriculturally and industrially relevant bacterium, as it is pathogenic towards a large range of crops, and secretes an exopolysaccharide (EPS), called xanthan gum, used in the food industry as a thickening agent (Jansson et al., 1975; Melton et al., 1976). In addition, the LPS of *Xanthomonas campestris* pv. *campestris* has been shown to influence its virulence (Braun et al., 2005; Dow et al., 1995).

We previously discovered the presence of apiose (3-C-[hydroxymethyl]-D-erythrofuranose, Api) in the soil-dwelling plant pathogen *Xanthomonas pisi* (Smith & Bar-Peled, 2017) and subsequently isolated and characterize the nucleotide sugar donor UDP-apiose and the

enzyme that forms it, UDP-apiose/UDP-xylose synthase (bUAS), in this bacteria. Evidence of apiose in *X. pisi* culture suggested *X. pisi* was able to use UDP-apiose to make a cell-surface apioside. In an effort to explain this, we searched for potential genes encoding apiosyltransferase (ApiT) activity.

Here we report the identification of genes encoding two putative GTs and that one is utilizing UDP-apiose and the other UDP-xylose. No apiosyltransferase has previously been purified to homogeneity nor have the genes encoding this glycosyltransferase been identified. This is the first report to identify a gene encoding an apiosyltransferase and the first apiosyltransferase activity documented in Bacteria.

RESULTS

Identification and phylogenetic analysis of putative glycosyltransferases in apiose operon of X. pisi

Examination of *X. pisi* genomic DNA revealed two hypothetical proteins (XpiCFBP4643_06445 and XpiCFBP4643_06450) directly downstream from the bacterial UDP-apiose synthase (bUAS) (**Figure 4.1A**) one of which has 4 nucleotides overlap with the gene bUAS suggesting the genes are organized in an operon. BLAST (Altschul et al., 1997a) analysis of XpiCFBP4643_06445 revealed few proteins that align. The closes homolog is a hypothetical protein C7B62_17955 from *Pleurocapsa sp.* CCALA 161 that shares low sequence identity with an e-value of $2e^{-21}$. Interestingly, we noticed that XpiCFBP4643_06445 had very low homology e^{-19} to a protein belonging to pfam05686 classified as glycosyl transferases belonging to GT90. Based on the CAZy GT database, proteins in GT90 had a putative type B fold. Therefore, we tentatively named it XpGTB. The third gene in the operon, XpiCFBP4643_06450, was annotated

to have a domain shared by GT family 2 (pfam13704), had a putative type A fold, and was termed XpGTA.

The BLAST program (Altschul et al., 1997b) was further used to identify bacterial proteins in the NCBI non-redundant database that share sequence similarity to XpGTA and XpGTB. The BLAST hits with highest sequence identity to XpGTA were from *Synechococcus sp., Desulfonatronum thiosulfatophilum,* and *Methylobacterium sp.*. Those with highest sequence identity to XpGTB were all below 35%, but include proteins from *Mycena chlorophos*, *Pleurocapsa sp.*, and *Sphingomonadaceae*.

Unrooted phylogenetic trees were generated for representative members of GT family 2, including XpGTA and members of GT family 90, including XpGTB (**Figure 4.1B**). XpGTA clusters into a clade with the predicted procollagen galactosyltransferase 1 precursor from Norway rat *Rattus norvegicus* and predicted proteins from other eukaryotic members including a sea anemone and a marine diatom. XpGTB is in its own unique clade indicating its distinction from the closest annotated tree members; a predicted viral protein from *Gryllus bimaculatus iridovirus* and human and zebrafish KDEL1 (**Figure 4.1B**).

The analyses compared alignments of amino acid sequences of representatives from respective GT families. Since most of these representatives are uncharacterized, and because there is low sequence identity between the *X. pisi* GT representatives and their respective GT family members, these GTs were heterologously expressed in *E. coli* and functionally characterized against a screen of potential UDP-sugar donors

Initial determination of XpGTA and XpGTB glycosyltransferase activities

To investigate if these GTs use UDP-apiose and UDP-xylose as substrates, the coding sequences of XpGTA and XpGTB were cloned into modified pET28b *E. coli* expression vector (Yang et al., 2009). Since UDP-apiose has a relatively short lifespan and naturally degrades within 4 h to apiofuranosyl-1,2-cyclic-phosphate (Guyett et al., 2009; Smith et al., 2016) we set up a 'reporting system' to determine the consumption of UDP-apiose *in-vivo*. For these experiments (we named herein in microbe) pET plasmids containing XpGTA or XpGTB alone, and a pET plasmid containing both XpGTA and XpGTB together were co-transformed with a pCDF plasmid containing both a UDP-glucose dehydrogenase (UGDH) (Broach et al., 2012) and a UDP-apiose synthase (UAS) to provide the potential UDP-apiose and UDP-xylose substrates for the GTs (Smith & Bar-Peled, 2017). Cell extracts from the isopropyl β-D-thiogalactoside (IPTG)-induced *E. coli* cells were processed and examined for presence and relative amounts of UDP-Api and UDP-Xyl. Peaks corresponding to UDP-Api and UDP-Xyl were compared (**Figure 4.2**).

In cells co-expressing UAS with XpGTA, there is a noticeable decrease in the relative amount of UDP-Xyl extracted and no change in the amount of UDP-Api (**Figure 4.2**), while there is no such decrease in control samples. This suggests that XpGTA consumes UDP-xylose as a substrate, perhaps transferring it to a substrate yet to be defined. Surprisingly, cells co-expressing UAS with XpGTB show a marked decrease in UDP-Api and no change in UDP-Xyl when compared to controls (**Figure 4.2**). The decrease in UDP-Api in cells expressing XpGTB indicates XpGTB is likely specifically utilizing UDP-Api in microbe. This suggests that XpGTB consumes UDP-apiose as a substrate, perhaps transferring it to a substrate yet to be defined. There is also possibility that in the absence of 'true acceptor' (as in the *in-vivo* assays carried out in *E. coli* and not in the native environment of *X. pisi*) that the glycosyltransferase may act to

hydrolyze UDP-sugar substrate (Brockhausen, 2014; Leemhuis et al., 2003; Sheikh et al., 2017; Sindhuwinata et al., 2010). Extracts from cells expressing both XpGTA and XpGTB along with UAS have less UDP-Xyl and UDP-Api relative to the internal Park's nucleotide (UDP-MurNAc-pentapeptide) than controls (**Figure 4.2**). To further explore the specific activity of these two GTs, His-tag recombinant proteins were expressed and purified and *in vitro* assays were developed.

Purification of XpGTA and XpGTB and in vitro assay

To obtain additional evidence for the specific activities of XpGTA and XpGTB, the recombinant His_6-tagged proteins were solubilized from *E. coli* cells and purified using nickel-affinity column. The recombinant XpGTA (with theoretical mass 32.9 kDa) and XpGTB (41.4 kDa) migrated on SDS-PAGE with predicted masses (**Figure 4.3A**).

Reactions of XpGTB with UDP-Api have over ten-fold more luminescence than other UDP-sugar substrates, indicating its specificity for UDP-Api (**Figure 4.3B**). Reactions of XpGTA with UDP-sugar substrate demonstrate a 3-fold increase in luminescence with UDP-Xyl. Promega's UGlo Kit® was used to test activities of each GT. In order to develop an ApiT assay, UDP-Api substrate had to be generated quickly *in vitro*, HPLC purified and flash frozen (see methods). To determine the specific NDP-sugar donor substrate of each GT, a UDP-sugar substrate screen was initially developed, to which purified enzyme was added. Because the acceptor substrate(s) were unknown, reactions were allowed to hydrolyze UDP-sugar substrate using water as the acceptor (Sheikh et al., 2017) and reaction product analyzed by luminometer according to manufacturer's instructions.

UDP-apiose synthesis and purification

Generation of pure UDP-Api substrate is not trivial due to the fact that it spontaneously degrades into a apiofuranosyl-1,2-cyclic-phosphate in solution (Guyett et al., 2009; Smith et al., 2016). Generation of UDP-Api required *in vitro* synthesis using an active recombinant UDP-apiose synthase, UAS (Smith et al., 2016). The UAS reaction products included NAD^+, UDP-Api, and UDP-Xyl. Products were quickly separated over HILIC HPLC, and the UDP-Api peak was collected, briefly air dried, and resuspended in DDW. Small aliquots were then flash-frozen in liquid N_2 and saved at -80 °C. This methodology allowed us to control and reduce the native degradation of UDP-Api. To validate purity and relative stability of UDP-Api, a small amount of freshly prepared UDP-Api was diagnosed by LC-MS (**Figure4. 4A**). The data shown (**Figure 4.4B**) provide evidence that UDP-Api is intact and lacking UDP-Xyl product contamination. Pure UDP-Api was necessary to screen XpGTA and XpGTB for activity using the UDP-Glo™ assay. Without UDP-Api, no activity would be assigned to XpGTB.

In microbe production of apiosides

To further investigate if the XpGTA and XpGTB were actively synthesizing glycans in microbe, engineered *E. coli* cultures were fractionated and analyzed for the presence and abundance of xylose and apiose. Bacteria grown in liquid media were pelleted and subjected to aqueous-methanol:chloroform extractions to lyse cells and separate hydrophilic small molecules and hydrophobic lipid moieties from the cell surface (termed pellet) fraction. Pellets were successively washed with water and dried. Samples were chemically hydrolyzed and monosaccharides converted to their alditol-sugar derivatives for GC-MS analysis.

The pellet fractions of only the UGDH+UAS+XpGTB and UGDH+UAS+XpGTB_XpGTA strains had a peak that migrated like apiose (**Figure 4.5**). The electron impact MS fragmentation pattern of this peak's structure was similar to authentic apiose; for example, the major diagnostic peak at *m/z* 188 (Smith & Bar-Peled, 2017). Apiose was not detected in the UGDH+UAS+XpGTA, UGDH+UAS, or XpGTB_XpGTA control strains, and no xylose was detected in any of the pellet fractions. In addition to apiose, these extracts also consist of glucose, galactose, and arabinose sugar residues (**Table 4.1**).

Further fractionations of cell pellets to extract LPS using hot-phenol (Forsberg et al., 2000) were attempted; however, no apiose was detected in these extracts (data not shown).

DISCUSSION

Here we describe the first evidence of an apiosyltransferase able to specifically use UDP-apiose as substrate. We show that the *X. pisi* genome consists of an operon with genes encoding UDP-apiose synthase (bUAS), an apiosyltransferase (XpGTB), followed by another enzyme that uses UDP-xylose (XpGTA). The apiosyltransferase belongs to CAZy GT family 90 and consists of mammalian protein O-glucosyl- and xylosyltransferases and putative plant enzymes with a domain of unknown function (DUF821) that may also function in O-glucosylation.

XpGTA is classified as GT family 2 in CAZy, which contains over 120,000 bacterial proteins as well as cellulose synthase from plant and may be the largest GT family in the database. The N-terminal portion (~100 aa) of XpGTA is classified as containing β-4-glucosyltransferase (cd02511) (Breton & Imberty, 1999; Campbell et al., 1997; Coutinho et al., 2003; Gagneux & Varki, 1999; Kahler et al., 1996; Kanipes et al., 2008; Kapitonov & Yu, 1999) and WcaA (COG0463) (Stevenson et al., 1996) conserved domains by the Conserved Domain

Database (CDD) (Marchler-Bauer et al., 2017). WcaA is a GT involved in synthesis of the exopolysaccharide (EPS) colonic acid (CA) produced by *E. coli* (Stevenson et al., 1996). Domain cd02511 contains GTs important for synthesis of the lipooligosaccharide (LOS) cores of the bacteria *Neisseria meningitidis* (Kahler et al., 1996) and *Campylobacter jejuni* (Kanipes et al., 2008), and domain pfam13704 are described as putative prokaryotic glucosyltransferases (Marchler-Bauer et al., 2017). Based on these classifications, we hypothesize that XpGTA may be involved in EPS or LOS synthesis.

The C-terminal portion (~200 aa) of the apiosyltransferase, XpGTB, also consists of a CAP10 domain (smart00672) (Chang & Kwon-Chung, 1999) and GT family 90 (pfam05686) (Klutts et al., 2007) by the CDD (Marchler-Bauer et al., 2017). The CAP10 from the fungus *Cryptoccocus neoformans* is a β-1,2-xylosyltransferases involved in capsular-polysaccharide formation and virulence (Chang & Kwon-Chung, 1999). Since discovery of this domain, capsular-associated proteins of *C. neoformans* have been classified into GT family 90 and characterized as β-1,2-xylosyltransferases that modify the capsular polysaccharide as well as some glycosphingolipids (Castle et al., 2008; Klutts & Doering, 2008; Klutts et al., 2007). Proteins homologous to CAP10 have been identified in plants and animals and are described in the CDD as "putative lipopolysaccharide-modifying enzymes" (Marchler-Bauer et al., 2017). In *Drosophila melanogaster*, Rumi (AAN13920.1) was identified and characterized as a soluble, ER-localized, CAP10-domain-containing protein involved in protein O-glycosylation and O-xylosylation critical to Notch signaling (Acar et al., 2008; Takeuchi et al., 2011). Additional homologs have been characterized in human and mouse (Fernandez-Valdivia et al., 2011; Ramkumar et al., 2015; Teng et al., 2006). Because XpGTB has a conserved GT family 90 domain, we hypothesized that it may be an apiosyl- or xylosyltransferase possibly involved in

extracellular polysaccharide modification. Very few deductions could be made from the GT family assignments.

The evolution of this unique operon remains elusive. About 1500 and 2000 bp upstream of the operon a disrupted transposase coding gene is annotated. It is possible that the bUAS, XpGTB, and XpGTA encoding genes ended up in this operon by horizontal gene transfer between *X. pisi* and host plant. In fact, other *Xanthomonas* species are reported to contain evidence of cross-kingdom gene transfer (Gardiner et al., 2012). Alternatively, *X. pisi* may have independently evolved a way to use UDP-apiose to out-compete other microbes. There is evidence that *Xanthomonas* pathovars are able to rapidly diversify by genomic rearrangement and diversification from common ancestral genes (Bansal et al., 2017).

As a soil-dwelling plant pathogen, *X. pisi* is constantly struggling to evade host defenses and vie for survival in its local microbiome. One way to best the competition is to evolve unique methods of "fight or flight." The identification and characterization of XpGTB as a genuine apiosyltransferase reveals a new activity exclusive to *X. pisi*. Perhaps displaying an apioside at the surface cloaks *X. pisi* from host defense or allows it to combat rival microbes. Future studies involving knock-out strains would provide biological relevance of apiose in this microbe.

EXPERIMENTAL PROCEDURES

Bacterial strains

Xanthomonas pisi ATCC 35936 (*X. pisi*) was obtained from the American Type Culture Collection in Manassas, VA, USA. The *X pisi* strain was isolated from *Pisum sativum* in Japan (Vauterin et al., 1995). Unless otherwise stated, *X. pisi* cultures were grown on nutrient agar (BD Difco) at 25 ^{o}C. Liquid cultures were maintained in 250 mL of nutrient broth, shaking at 250

rpm. *E. coli* cells used were Bl21-derived Rosetta (DE3) cells (Novagen) modified with a pLysS plasmid (Stratagene).

Identification and cloning of XpGTA and XpGTB

The genomic DNA surrounding bUAS of *X. pisi* (10,000 bp upstream and 10,000 downstream) was examined for in-frame amino acid-coding sequences and sent into the BLASTX program (Altschul et al., 1997b). Resultant *X. pisi* predicted proteins WP_046963858.1 (XpGTB) and WP_0052764980.1 (XpGTA) were submitted to BLASTP against the NCBI non-redundant database to obtain conserved domains from the Conserved Domain Database (CDD). XpGTB had conserved CAP10 (smart00672) and GT family 90 (pfam05686) domains, and XpGTA was shown to contain conserved GT family 2 (pfam13704), Beta4Glucosyltransferase (cd02511), and WcaA (COG0463) domains.

RNA was extracted and XpGTA, XpGTB, and XpGTB_XpGTA cDNA amplified as described (Z. Li et al., 2016) from a 2-day-old culture of *X. pisi* grown in liquid media, using forward and reverse primers (IDT, Coralville, IA, USA; **Table 4.S1**). The PCR product was directly cloned into the *E. coli* expression vector pET28b modified to contain an N-terminal His_6 tag followed by a TEV cleavage site (Yang et al., 2009).

Following cloning of the individual *X. pisi* GT genes, the plasmids were sequence verified (Eurofins, LUX) and termed, pET28b-TEV-XpGTA.2, pET28b-TEV-XpGTB.1 and pET28b-TEV-XpGTB_XpGTA.7. Their amino acid sequences were deposited in GenBank™.

Analysis of nucleotide sugars produced in microbe

Nucleotide sugars from *E. coli* harboring the expression plasmids were harvested as described (Smith et al., 2016; Yang et al., 2012). *E. coli* cells co-transformed with pCDFDuet-SpUAS_BtbDH and either empty pET28b vector control (UGDH+UAS), pET28b-TEV-XpGTA.2 (UGDH+UAS+XpGTA), pET28b-TEV-XpGTB.1 (UGDH+UAS+XpGTB), pET28b-TEV-XpGTB_XpGTA.7 (UGDH+UAS+XpGTB_XpGTA) or negative control pET28b-TEV-BtbGlyT046 (UGDH+UAS+BtbGlyT046), and cells co-transformed with empty pCDFDuet vector control and either pET28b-TEV-XpGTA.2, pET28b-TEV-XpGTB.1, and pET28b-TEV-XpGTB_XpGTA.7 alone were grown in LB medium [1.0% (w/v) Bacto tryptone, 0.5% (w/v) Bacto yeast extract, and 1.0% (w/v) NaCl] supplemented with 35 μg/ml chloramphenicol, 50 μg/ml kanamycin, and 25 μg/ml spectinomycin at 37 °C and 250 rpm, induced with Isopropyl β-D-thiogalactoside (IPTG, 0.5 mM) at an OD_{600} of 0.6 and grown at 30 °C for 4 h. In microbe nucleotide sugars were extracted and analyzed by hydrophilic interaction liquid chromatography electrospray ionization tandem mass spectrometry (HILIC-LC-ESI-MS/MS) as described (Smith et al., 2016). Briefly, in microbe extracts were mixed with 2/3 volume aq. 95% acetonitrile (ACN) containing 25 mM ammonium acetate and an aliquot (10 - 20 μl) chromatographed over an Accucore amide-HILIC column (150 x 4.6 mm; Thermo), eluted at 0.4 ml min^{-1} with a linear gradient of aq. 75% (v/v) acetonitrile containing 40 mM ammonium acetate, pH 4.4, to 50% (v/v) acetonitrile containing 40 mM ammonium acetate, pH 4.4, over 35 min using a Shimadzu LC-30AD HPLC. Mass spectra (mass range 100-2,000 *m/z*) were collected every 1.3 sec for 30 minutes. Second stage MS/MS data was collected by collision-induced dissociation (CID) with a collision energy of 35% and a nebulizing nitrogen gas flow of 1.5 ml min^{-1} (Hwang et al., 2014).

His_6-tagged protein expression and purification

E. coli cells transformed with pET28b-TEV-XpGTA.2, pET28b-TEV-XpGTB.1, or the empty vector control were grown, induced, and protein was extracted and purified over Ni-column as described (Yang et al., 2009). The final soluble protein fraction (Fraction S20) was collected and kept on ice prior to immediate purification. The different His_6-tagged proteins, including control empty plasmid were separated SDS-PAGE [12% (w/w) polyacrylamide] and stained. Protein concentrations were determined with the Bradford reagent (Bradford, 1976) using bovine serum albumin (BSA) as standard.

Generation of UDP-apiose

Reactions (50 µl) of 1 mM UDP-GlcA with 10 µg of purified SpUAS and 0.5 mM NAD^+ were carried out, separated over an Accucore amide-HILIC column (150 x 4.6 mm; Thermo Fisher Scientific; Waltham, MA), as described above (Smith et al., 2016). Nucleotides sugars were detected by their A_{261nm} (for UDP-sugars) and A_{259nm} (for NAD^+). UDP-apiose eluted at ~11.2 min, was collected, and aliquots were flash-frozen in liquid N_2 and saved at -80 °C. To validate purity and relative stability of UDP-Api, a small amount of freshly prepared UDP-Api was diagnosed by on LC-MS (see above).

Recombinant XpGTA and XpGTB enzyme assays

Unless otherwise indicated, the 10 µl UDP-Glo™ (Promega, Madison, WI, USA; http://www.promega.com) reactions were performed in 50 mM KH_2PO_4 pH 7.5, containing 0.1 mM UDP-sugar 1 µg of purified protein. The assay mixtures were incubated at room temperature for up to 1.5 h. The reactions were terminated by addition of an equal volume of UDP-Glo Detection Reagent in white polystyrene 384-well assay plates (Corning, Corning, NY, USA) and

incubated for 1 h at room temp. Luminescence was measured using a multifunctional microplate reader (POLARstar OPTIMA; BMG Labtech, Ortenberg, DEU). Relative luminescence was calculated by subtracting blank, containing enzyme without UDP-sugar substrate, and dividing by background, UDP-sugar with boiled enzyme.

Glycosyl residue composition analysis

5 to 18-hr-old *E. coli* cultures (250 ml) were centrifuged (10,000 g, 5 min, 4 °C) and cell pellets suspended in 10 volumes of cold MeOH:chloroform:H_2O (40:40:20, v/v/v). The suspensions were transferred to 15 ml falcon tubes and vortexed for 10 min (30 sec every 2 min, 4 °C). The suspensions were centrifuged (10,000*g*, 5 min, 4 °C) and separated into a top methanolic-water phase (termed methanolic), medial interphase (termed pellet) and bottom organic chloroform phase. A portion (20 µl) of the top methanolic fraction was analyzed on HILIC-LC-ESI-MS/MS (see above). The bottom chloroform fraction was transferred to a separate tube. The remaining interphase was re-suspended in 2 ml DDW and samples centrifuged (10,000 g 5 min, 4 °C). Supernatant was again washed by re-suspension in 2 ml DDW and centrifugation and supernatant vacuum aspirated. The pellet was again re-suspended in 1 ml DDW and transferred to a new 13 mm borosilicate tube.

In microbe pellets, methanolic and organic solvent extracts or standards (~1 mg) were supplemented with myo-Inositol (10 µl of 5 mM solution) as an internal standard, evaporated to dryness at room temperature using a stream of air (REACTIVAP III, Thermo Fisher, Waltham, MA, USA) and then hydrolyzed for 1 h at 120 °C with 1 ml of 1 M TFA. TFA was removed by evaporation and the released monosaccharides were then converted into their corresponding alditol-acetate derivatives according to York *et al* (Rezanka et al., 2013), and the final residue

dissolved in acetone (100 µl). A fraction (1 µl) of each of the alditol-acetate derivative samples was analyzed by gas-liquid chromatography as described (Smith et al., 2016).

TABLES AND FIGURES

TABLE 4.1 Glycosyl-residue compositions of microbe pellets. The glycosyl residue composition (mol%) was determined by GC-MS analysis of alditol-acetate derivatives generated from the pellet fraction of transformed *E. coli*. Data is the average ± standard error of three independent samples. N.D. indicates not detected.

	UGDH+UAS	**UGDH+UAS +BtbGlyT046**	**UGDH+UAS + XpGTA**	**UGDH+UAS +XpGTB**	**UGDH+UAS +XpGTB_XpGTA**	**XpGTA**	**XpGTB**	**XpGTB_ XpGTA**
	Monosaccharide (mol %)							
Rib	83.2 ± 1.0	85.5 ± 3.2	90.2 ± 1.2	50.3 ± 3.5	56.6 ± 5.9	89.0 ± 1.4	86.5 ± 2.0	82.1 ± 2.2
Ara	1.0 ± 0.2	1.1 ± 0.2	0.7± 0.03	0.6 ± 0.1	0.9 ± 0.4	0.7± 0.02	0.6± 0.1	0.9± 0.01
Xyl	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Api	N.D.	N.D.	N.D.	2.78 ± 0.4	1.42 ± 0.4	N.D.	N.D.	N.D.
Man	0.4 ± 0.2	0.3 ± 0.2	0.3 ± 0.1	0.4 ± 0.04	0.4 ± 0.2	0.4 ± 0.1	0.6± 0.01	0.8± 0.02
Gal	3.9 ± 0.2	4.2 ± 0.2	4.3± 0.8	41.2 ± 4.5	35.1 ± 2.2	5.7± 0.8	7.1± .6	4.1± 0.4
Glc	2.2 ± 0.2	1.0 ± 0.1	1.7± 0.3	1.2 ± 0.03	1.0 ± 0.04	1.7± 0.2	3.1± 0.04	2.4± 0.3
GlcN	2.0 ± 0.1	1.2 ± 0.3	1.1 ± 0.1	1.0 ± 0.7	0.9 ± 0.3	0.7± 0.2	.9 ± 0.02	2.1 ± 0.2
GlcNAc	8.2 ± 0.3	7.0 ± 0.1	2.4± 0.3	2.2 ± 0.07	3.1 ± 0.13	2.0 ± 0.2	1.1 ± 0.3	7.9 ± 1.1

FIGURE 4.1

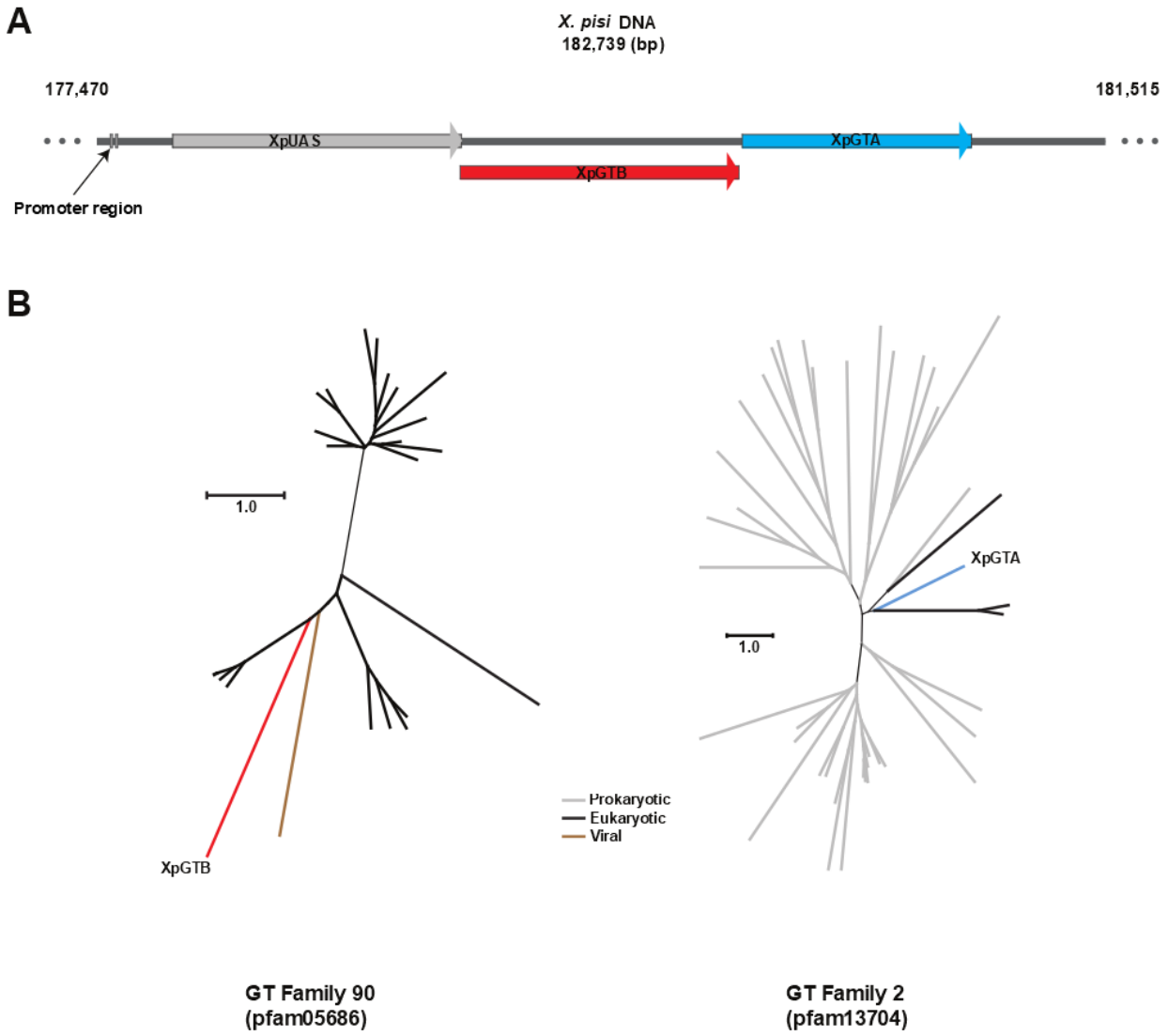


FIGURE 4.1. *X. pisi* apiose operon and phylogeny of GTs. **(**A) Gene organization of the apiose operon in *X. pisi* genome (NZ_JPLE01000032.1) from position 177,470 to 181,515 bp. -35 and -10 promoters were predicted using Bacterial Promoter Prediction website (BacPP, https://molbiol-tools.ca/Promoters.htm). Operon schema was generated using SnapGene® Viewer version 3.1.4 (GSL Biotech, Chicago, IL). (B) phylogenetic analysis of XpGTB and XpGTA. Amino acid sequences from representatives of pfam05686 and pfam13704 selected by

the CDD were used. Alignments were made using Clustal Omega (W. Li et al., 2015; McWilliam et al., 2013; Sievers et al., 2011) and the trees generated using Dendroscope (Huson & Scornavacca, 2012).

FIGURE 4.2

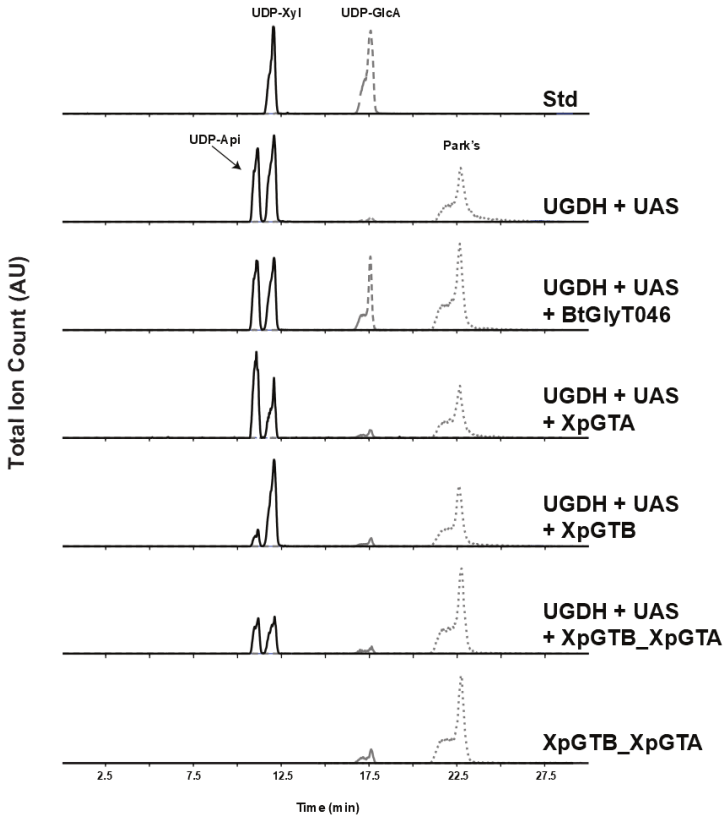


FIGURE 4.2. In microbe activity of XpGTA and XpGTB. Analysis of in microbe nucleotide sugars by HILIC-LC-ESI-MS/MS. (A) top panel elution of standard (Std): UDP-GlcA and UDP-Xyl; Nucleotide sugars were extracted from *E. coli* cells induced to express genes encoding BtbUGDH and SpUAS only; BtbUGDH and SpUAS along with a putative N-acetyl-glucosamine transferase from *Bacillus thuringiensis* (BtGlyT046), XpGTA, XpGTB, or XpGTB with XpGTA; or only XpGTB with XpGTA as control (bottom panel). Extracted ion chromatograms

(XICs) of $[M-H]^-$ ions diagnostic for UDP-pentose (*m/z* 535.0, solid line), UDP-hexuronic acid (*m/z* 579.0, dashed line) and Park's nucleotide (*m/z* 595.6, dotted line) are displayed. Park's nucleotide is a UDP-MurNAc-pentapeptide that is used as an internal standard for nucleotide-sugar detection as it is abundantly made in *E. coli*.

FIGURE 4.3

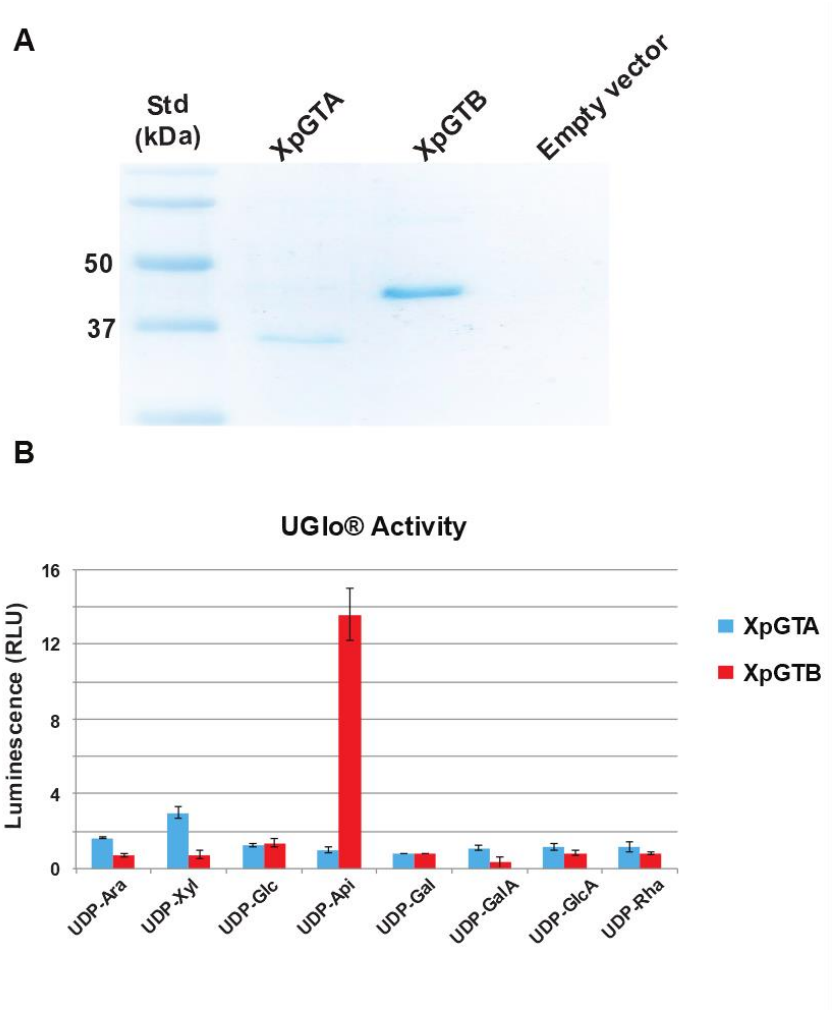


FIGURE 4.3. Activity of purified XpGTA and XpGTB proteins. (A) Nickel-purified proteins from *E. coli* cells induced to express XpGTA, XpGTb, and empty vector control with expected sizes of XpGTA and XpGTb: 32.9 and 41.4 kDa, respectively. (B) Bar graph of relative

luminescence from UDP-Glo® assays. Assays of 50 mM KH2PO4 pH 7.5, 0.1 mM UDP-sugar, 0.5 µg enzyme in 10 µl total volume reacted at room temperature for 1-1.5 hr. Equal volume of UDP-Glo® detection reagent added to stop reaction for 1 hr. Relative luminescence is fold signal over background; measured by subtracting blank (enzyme without UDP-sugar) and dividing by background (UDP-sugar with boiled enzyme).

FIGURE 4.4

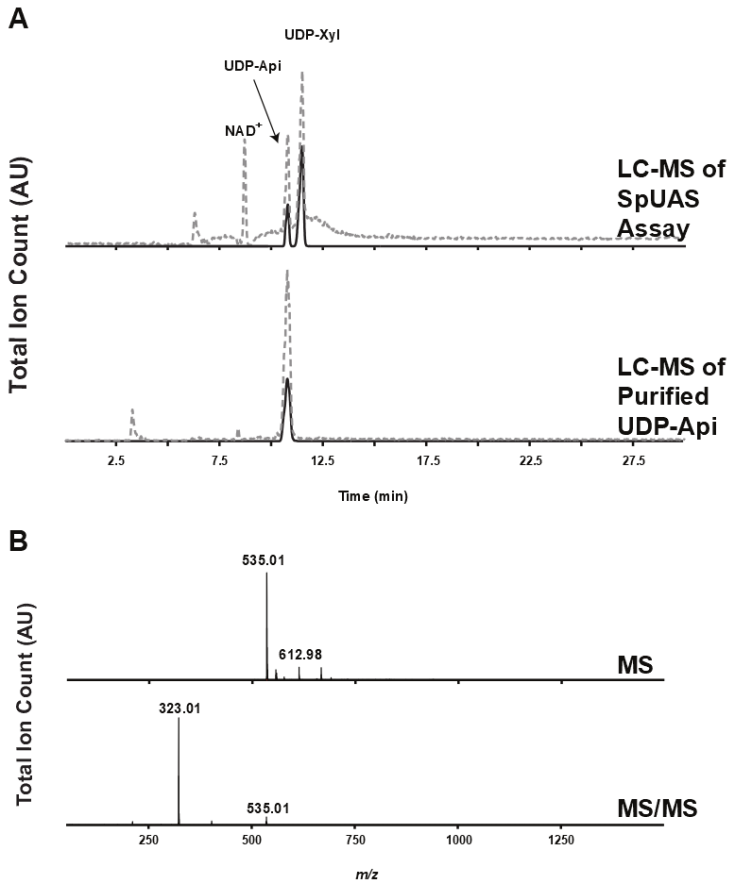


FIGURE 4.4. HPLC-purified *in vitro* synthesized UDP-apiose. (A) Total ion count (TIC, dashed line) and XIC $[M-H]^-$ ions diagnostic for UDP-pentose (*m/z* 535.0, solid line). (B) Average of all

ions at 11.0 min (MS, top panel) and average of all fragmented ions at 11.0 min (MS/MS, bottom panel).

FIGURE 4.5

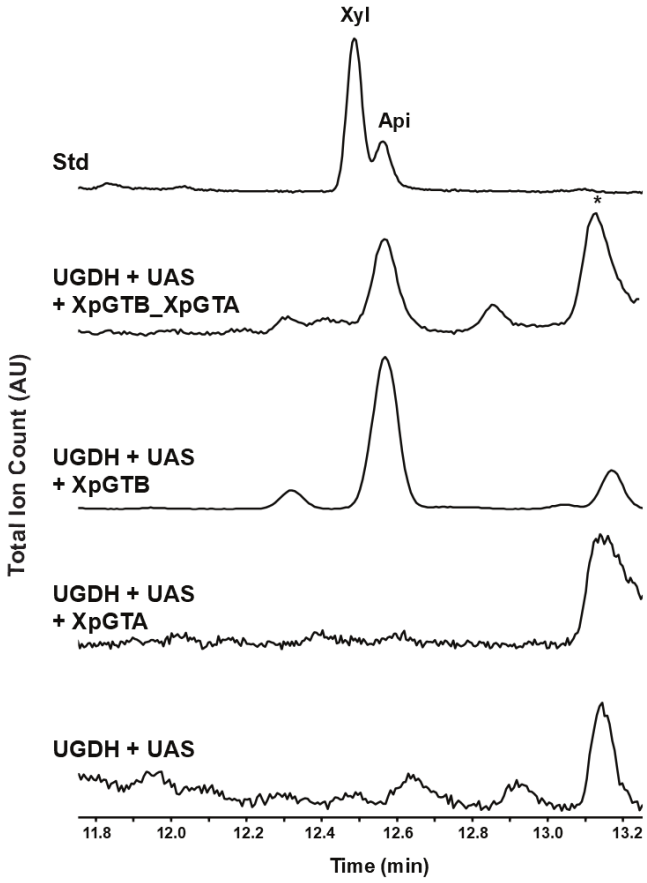


FIGURE 4.5. Detection of apiose in microbe. GC-MS analysis of alditol-acetate derivatives from cell pellet fractions of *E. coli* engineered to overexpress BtbUGDH + SpUAS with *X. pisi* GTs in microbe. Standard (Std) contains authentic xylose and apiose. The region of the total ion count for xylose- (Xyl) and apiose- (Api) alditol-acetate derivatives is expanded. * indicates unidentified residue.

SUPPLEMENTAL INFORMATION

TABLE 4.S1. Primers used in plasmid generation. Obtained from IDT.

Primer	Sequence (5'-3')
XpGTA_F	GTATTTTCAGGGCGCCATGACAGGCATGCGTCTGATAC
XpGTA_R	AGCCGGATCGAATTCACTACATGTCATTGCTGCAGGC
XpGTB _F	GTATTTTCAGGGCGCCATGACATCGAGCTCCGCCAAAG
XpGTB _R	AGCCGGATCGAATTCACCGAGGCCCCGCATGCG
BtbGlyT046_F	GTATTTTCAGGGCGCCATGTGAGTACTCCTATTATTGCTTATG
BtbGlyT04_R	AGCCGGATCGAATTCACTAGCATTCATCTGTATTCTTTTCTTCCTTTC
pET28b_TEV_F	TGAATTCGATCCGGCTGCTAACAAAGCCCG
pET28b_TEV_R	CATGGCGCCCTGAAAATACAGGTTTTC
XpSig70_F	GTTGTAGCGATTACCGCCCG
XpSig70_R	GCTCAGGCGCAATTTGGC

CHAPTER 5

CONCLUSIONS AND FUTURE PERSPECTIVES

As the cross-linking residue for RG-II dimerization through borate, the origin of Api represents an important hallmark in vascular land plant evolution. The origin of Api can be inferred from the evolution of UAS activity to synthesize UDP-Api. Both UAS and UXS decarboxylate UGlcA to a 4-keto- intermediate, and both show sequence similarity to the decarboxylase domain of the bacterial SDR ArnA. Alignment of several UDP-GlcA decarboxylase enzymes from bacteria and plant reveals that plant and bacterial UAS are more closely related to bacterial ArnA than plant and bacterial UXS. A common ancestral prokaryotic UAS/UXS probably evolved into two separate enzymes with distinct functions. ArnA may be the extant version of that ancestral enzyme. It is possible that the original enzyme did not bind NAD^+ tightly, causing release of incompletely reduced intermediates and NADH. Over time, the enzyme evolved to hold NAD^+ more tightly or in different orientations, which allowed the enzyme to perform additional activities; complete reduction of UDP-4-ketopentose to UDP-Xyl (UXS) and retroaldol ring cleavage/aldol condensation (UAS). While sequence alignment suggests domains responsible for these activities, a crystal structure of UAS or directed mutation/domain engineering is required for a full understanding of the catalytic mechanism.

The presence of UAS in avascular plants and the detection of Api in secondary metabolite extracts of bryophytes and green algae suggests that Api originated in early avascular

land plants. Because overexpression of SpUAS in *Physcomitrella* did not result in synthesis of wall apiosides, it is likely that the GTs (or nucleotide sugar transporters) responsible for wall apioside synthesis do not exist in *Physcomitrella*. This hypothesis is supported by absence of the XGD1 family of GTs from Physcomitrella genome (McCarthy et al., 2014). It would be interesting to transform Physcomitrella with XGD1 to find if it could make xylogalacturonan. Both UAS and a CMP-KDO synthase (KdsA) homolog exist in the genome of the moss *Physcomitrella*. The KdsA homolog has over 70% sequence identity to the *Arabidopsis* protein, and expressed sequence tag (EST) data substantiates that the *Physcomitrella* KdsA homolog is expressed. No Api has been detected in *Physcomitrella* (Bar-Peled et al., 2012). If, and under what conditions, any RG-II glycoform is made in *Physcomitrella* (perhaps in specialized reproductive tissues). Of note, the *Physcomitrella* UAS expression appears to be restricted to spore S3 and archegonia stages

GTs involved in pectin biosynthesis are predicted to use nucleotide-sugars as immediate donors (Mohnen, 2008), which suggests RG-II side chains are sequentially added by GTs using available nucleotide-sugar donors. These nucleotide-sugars are typically transported from the cytosol into the Golgi lumen, and the most recent data indicates that particular transporters are tightly associated or complexed with partner GTs (Oikawa et al., 2013). In some cases, the nucleotide-sugar metabolizing enzymes are also closely coupled to these Golgi-residing transporters and GTs. Such is the case for an *Arabidopsis* UDP-arabinose mutase (UAM) and is proposed for AXS1, a UDP-apiose/UDP-xylose synthase (Reboul & Tenhaken, 2012). Considering the low amounts of RG-II in the primary wall and the fact that several of its component sugars are unique to its structure, it is probable that the nucleotide-sugars dedicated to RG-II biosynthesis are tightly controlled and limiting.

Because the backbone of RG-II is HGA, it is reasonable to postulate that RG-II biosynthesis is initiated by addition of a primary branch residue (Api, KDO, or DHA) of one of the side chains to HGA. The structure of ApiGalA favors the notion that Api is the first sugar transferred to HGA to begin synthesis of side chain A or B. The substituted galacturonide ApiGalA has an HGA backbone with side chains of β-(1,2/3)-D-apiofuranose and β-(1,2/3)-apiobiose, where apiobiose is D-apiofuranose-β-(1,3)-D-apiofuranose (side chains are either attached to O-2 or O-3). Interestingly, some Xyl is also reported in extracted ApiGalA, supporting the idea that the bifunctional UDP-apiose/UDP-xylose synthase is closely coupled to a potentially promiscuous GT in the Golgi and that Api or Xyl is added to HGA as fast as it is generated (Longland et al., 1989).

The discovery of the GAUT, XGD, RGXT, and GALS enzymes provides pectin biosynthesis research with an invaluable tool whereby the genes encoding these proteins are used to probe co-expression data for additional HGA, XGA, or RG-II-specific enzymes or to validate candidates (Voxeur et al., 2012). Using these probes and excluding GTs from Physcomitrella, we constructed a candidate list worth investigating. The most recent discovery of an ApiT in bacteria may give insights and provide additional bioinformatics tools to uncover a plant-sourced apiosyltransferase involved in RG-II synthesis.

REFERENCES

Acar, M., Jafar-Nejad, H., Takeuchi, H., Rajan, A., Ibrani, D., Rana, N. A., . . . Bellen, H. J. (2008). Rumi is a CAP10 domain glycosyltransferase that modifies Notch and is required for Notch signaling. *Cell, 132*(2), 247-258. doi:10.1016/j.cell.2007.12.016

Ahn, J. W., Verma, R., Kim, M., Lee, J. Y., Kim, Y. K., Bang, J. W., . . . Pai, H. S. (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*(19), 13708-13716. doi:10.1074/jbc.M512403200

Ahn, Y. O., Mizutani, M., Saino, H., & Sakata, K. (2004). Furcatin hydrolase from Viburnum furcatum Blume is a novel disaccharide-specific acuminosidase in glycosyl hydrolase family 1. *J Biol Chem, 279*(22), 23405-23414. doi:10.1074/jbc.M311379200

Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., Staehelin, L.A. (2011). *Plant Cell Walls*. New York: Garland Science.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997a). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res., 25*(17), 3389-3402.

Ardevol, A., & Rovira, C. (2015). Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J Am Chem Soc, 137*(24), 7528-7547. doi:10.1021/jacs.5b01156

Atmodjo, M. A., Sakuragi, Y., Zhu, X., Burrell, A. J., Mohanty, S. S., Atwood, J. A., 3rd, . . . Mohnen, D. (2011). Galacturonosyltransferase (GAUT)1 and GAUT7 are the core of a plant cell wall pectin biosynthetic homogalacturonan:galacturonosyltransferase complex. *Proc Natl Acad Sci U S A, 108*(50), 20225-20230. doi:10.1073/pnas.1112816108

Avci, U., Pena, M. J., & O'Neill, M. A. (2018). Changes in the abundance of cell wall apiogalacturonan and xylogalacturonan and conservation of rhamnogalacturonan II structure during the diversification of the Lemnoideae. *Planta, 247*(4), 953-971. doi:10.1007/s00425-017-2837-y

Baba, K., Taniguti, M., Yoneda, Y., Kozawa, M. (1990). Coumarin glycosides from *Edgeworthia chrysantha*. *Phytochemistry, 29*(1), 247-249.

Bansal, K., Midha, S., Kumar, S., & Patil, P. B. (2017). Ecological and Evolutionary Insights into Xanthomonas citri Pathovar Diversity. *Appl Environ Microbiol*, 83(9). doi:10.1128/AEM.02993-16

Bar-Peled, M., Griffith, C. L., & Doering, T. L. (2001). Functional cloning and characterization of a UDP- glucuronic acid decarboxylase: the pathogenic fungus Cryptococcus neoformans elucidates UDP-xylose synthesis. *Proc Natl Acad Sci U S A, 98*(21), 12003-12008. doi:10.1073/pnas.211229198

Bar-Peled, M., Urbanowicz, B. R., & O'Neill, M. A. (2012). The synthesis and origin of the pectic polysaccharide rhamnogalacturonan II - insights from nucleotide-sugar formation and diversity. *Front. Plant Sci., 3*, 92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22639675 doi:10.3389/fpls.2012.00092

Baron, D., Wellmann, E., & Grisebach, H. (1972). Purification and properties of an enzyme from cell suspension cultures of parsley catalyzing the synthesis of UDP-apiose and UDP-D-xylose from UDP-D-glucuronic acid. *Biochim Biophys Acta, 258*(1), 310-318.

Baylson, F. A., Stevens, B. W., & Domozych, D. S. (2001). Composition and synthesis of the pectin and protein components of the cell wall of *Closterium acerosum* (Chlorophyta). *J. of Phycol., 37*, 796-809.

Beck, E. (1967). Isolation and characterization of an apiogalacturonan from the cell wall of *Lemna minor*. *Z. Pflanzenphysiol, 57*, 444-461.

Beck, E. (1982). *Branched-Chain Sugars* (Vol. 13 A). Berlin, Heidelberg, New York: Springer-Verlag

Beck, E., & Hopf, H. (1990). Branched-chain sugars and sugar alcohols. In D. D.M. (Ed.), *Methods in Plant Biochemistry* (2 ed., pp. 235–289). London: Academic Press.

Becker, A., Katzen, F., Puhler, A., & Ielpi, L. (1998). Xanthan gum biosynthesis and application: a biochemical/genetic perspective. *Appl Microbiol Biotechnol, 50*(2), 145-152.

Benach, J., Atrian, S., Gonzalez-Duarte, R., & Ladenstein, R. (1999). The catalytic reaction and inhibition mechanism of Drosophila alcohol dehydrogenase: observation of an enzyme-bound NAD-ketone adduct at 1.4 A resolution by X-ray crystallography. *J Mol Biol, 289*(2), 335-355. doi:10.1006/jmbi.1999.2765

Beveridge, T. J. (1981). Ultrastructure, chemistry, and function of the bacterial wall. *Int Rev Cytol, 72*, 229-317.

Bianco, M. I., Toum, L., Yaryura, P. M., Mielnichuk, N., Gudesblat, G. E., Roeschlin, R., . . . Vojnov, A. A. (2016). Xanthan Pyruvilation Is Essential for the Virulence of Xanthomonas campestris pv. campestris. *Mol Plant Microbe Interact, 29*(9), 688-699. doi:10.1094/MPMI-06-16-0106-R

Birch, H. L., Alderwick, L. J., Rittmann, D., Krumbach, K., Etterich, H., Grzegorzewicz, A., . . . Besra, G. S. (2009). Identification of a terminal rhamnopyranosyltransferase (RptA) involved in Corynebacterium glutamicum cell wall biosynthesis. *J Bacteriol, 191*(15), 4879-4887. doi:10.1128/JB.00296-09

Bishop, P. D., Pearce, G., Bryant, J. E., & Ryan, C. A. (1984). Isolation and characterization of the proteinase inhibitor-inducing factor from tomato leaves. Identity and activity of poly- and oligogalacturonide fragments. *J Biol Chem, 259*(21), 13172-13177.

Boudet, A. M. (2000). Lignins and lignification: Selected issues. *Plant Physiol. Biochem, 38*, 81-96.

Bourne, Y., & Henrissat, B. (2001). Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struct Biol, 11*(5), 593-600.

Bouton, S., Leboeuf, E., Mouille, G., Leydecker, M. T., Talbotec, J., Granier, F., . . . Truong, H. N. (2002). QUASIMODO1 encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in Arabidopsis. *Plant Cell, 14*(10), 2577-2590.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem., 72*, 248–254.

Braun, S. G., Meyer, A., Holst, O., Puhler, A., & Niehaus, K. (2005). Characterization of the Xanthomonas campestris pv. campestris lipopolysaccharide substructures essential for elicitation of an oxidative burst in tobacco cells. *Mol Plant Microbe Interact, 18*(7), 674-681. doi:10.1094/MPMI-18-0674

Breazeale, S. D., Ribeiro, A. A., McClerren, A. L., & Raetz, C. R. (2005). A formyltransferase required for polymyxin resistance in *Escherichia coli* and the modification of lipid A with 4-Amino-4-deoxy-L-arabinose. Identification and function oF UDP-4-deoxy-4-formamido-L-arabinose. *J. Biol. Chem., 280*(14), 14154-14167. doi:10.1074/jbc.M414265200

Breton, C., Fournel-Gigleux, S., & Palcic, M. M. (2012). Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struct Biol, 22*(5), 540-549. doi:10.1016/j.sbi.2012.06.007

Breton, C., & Imberty, A. (1999). Structure/function studies of glycosyltransferases. *Curr Opin Struct Biol, 9*(5), 563-571.

Broach, B., Gu, X., & Bar-Peled, M. (2012). Biosynthesis of UDP-glucuronic acid and UDP-galacturonic acid in *Bacillus cereus* subsp. cytotoxis NVH 391-98. *FEBS J, 279*(1), 100-112. doi:10.1111/j.1742-4658.2011.08402.x

Brockhausen, I. (2014). Crossroads between Bacterial and Mammalian Glycosyltransferases. *Front Immunol*, 5, 492. doi:10.3389/fimmu.2014.00492

Buffetto, F., Ropartz, D., Zhang, X. J., Gilbert, H. J., Guillon, F., & Ralet, M. C. (2014). Recovery and fine structure variability of RGII sub-domains in wine (Vitis vinifera Merlot). *Ann Bot, 114*(6), 1327-1337. doi:10.1093/aob/mcu097

Bystrova, O. V., Knirel, Y. A., Lindner, B., Kocharova, N. A., Kondakova, A. N., Zahringer, U., & Pier, G. B. (2006). Structures of the core oligosaccharide and O-units in the R- and SR-type lipopolysaccharides of reference strains of Pseudomonas aeruginosa O-serogroups. *FEMS Immunol Med Microbiol, 46*(1), 85-99. doi:10.1111/j.1574-695X.2005.00004.x

Caffall, K. H., & Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr Res, 344*(14), 1879-1900. doi:10.1016/j.carres.2009.05.021

Camacho-Cristobal, J. J., Rexach, J., & Gonzalez-Fontes, A. (2008). Boron in plants: deficiency and toxicity. *J. Integr. Plant Biol., 50*(10), 1247-1255. doi:10.1111/j.1744-7909.2008.00742.x

Campbell, J. A., Davies, G. J., Bulone, V., & Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J, 326 (Pt 3)*, 929-939.

Capasso, J. M., & Hirschberg, C. B. (1984). Mechanisms of glycosylation and sulfation in the Golgi apparatus: evidence for nucleotide sugar/nucleoside monophosphate and

nucleotide sulfate/nucleoside monophosphate antiports in the Golgi apparatus membrane. *Proc Natl Acad Sci U S A, 81*(22), 7051-7055.

Castle, S. A., Owuor, E. A., Thompson, S. H., Garnsey, M. R., Klutts, J. S., Doering, T. L., & Levery, S. B. (2008). Beta1,2-xylosyltransferase Cxt1p is solely responsible for xylose incorporation into Cryptococcus neoformans glycosphingolipids. *Eukaryot Cell, 7*(9), 1611-1615. doi:10.1128/EC.00458-07

Chang, Y. C., & Kwon-Chung, K. J. (1999). Isolation, characterization, and localization of a capsule-associated gene, CAP10, of Cryptococcus neoformans. *J Bacteriol, 181*(18), 5636-5643.

Chang, Y. H., Chuang, L. Y., & Hwang, C. C. (2007). Mechanism of proton transfer in the 3alpha-hydroxysteroid dehydrogenase/carbonyl reductase from Comamonas testosteroni. *J Biol Chem, 282*(47), 34306-34314. doi:10.1074/jbc.M706336200

Chatterjee, A. N., & Young, F. E. (1972). Regulation of the bacterial cell wall: isolation and characterization of peptidoglycan mutants of Staphylococcus aureus. *J Bacteriol, 111*(1), 220-230.

Cheng, W., Zhu, C. G., Xu, W. D., Fan, X. N., Yang, Y. C., Li, Y., & Shi, J. G. (2012). Phenylpropene diglycosides from the bark of Machilus wangchiana. *J Asian Nat Prod Res, 14*(11), 1046-1053. doi:10.1080/10286020.2012.702762

Choi, S. H., Mansoorabadi, S. O., Liu, Y. N., Chien, T. C., & Liu, H. W. (2012). Analysis of UDP-D-apiose/UDP-D-xylose synthase-catalyzed conversion of UDP-D-apiose phosphonate to

UDP-D-xylose phosphonate: implications for a retroaldol-aldol mechanism. *J Am Chem Soc, 134*(34), 13946-13949. doi:10.1021/ja305322x

Chormova, D., & Fry, S. C. (2016). Boron bridging of rhamnogalacturonan-II is promoted in vitro by cationic chaperones, including polyhistidine and wall glycoproteins. *New Phytol, 209*(1), 241-251. doi:10.1111/nph.13596

Cocuron, J. C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., . . . Wilkerson, C. G. (2007). A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase. *Proc Natl Acad Sci U S A, 104*(20), 8550-8555. doi:10.1073/pnas.0703133104

Coenen, G. J., Bakx, E. J., Verhoef, R. P., Schols, H. A., & Voragen, A. G. J. (2007). Identification of the connecting linkage between homo- or xylogalacturonan and rhamnogalacturonan type I. *Carbohydrate Polymers, 70*(2), 224–235.

Coenen, G. J., Kabel, M. A., Schols, H. A., & Voragen, A. G. (2008). CE-MSn of complex pectin-derived oligomers. *Electrophoresis, 29*(10), 2101-2111. doi:10.1002/elps.200700465

Cornelius, MTF, , C., MG, , d. S., TMS, . . . R. (2010). Other chemical constituents isolated from Solanum crinitum Lam. (Solanaceae)*J Braz Chem Soc, 21,* 2211–2219.

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat Rev Mol Cell Biol, 6*(11), 850-861. doi:10.1038/nrm1746

Coutinho, P. M., Deleury, E., Davies, G. J., & Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol, 328*(2), 307-317.

Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., . . . Rosenbusch, J. P. (1992). Crystal structures explain functional properties of two *E. coli* porins. *Nature, 358*(6389), 727-733. doi:10.1038/358727a0

Coyne, M. J., Fletcher, C. M., Reinap, B., & Comstock, L. E. (2011). UDP-glucuronic acid decarboxylases of Bacteroides fragilis and their prevalence in bacteria. *J Bacteriol, 193*(19), 5252-5259. doi:10.1128/JB.05337-11

Crowell, D. N., Anderson, M. S., & Raetz, C. R. (1986). Molecular cloning of the genes for lipid A disaccharide synthase and UDP-N-acetylglucosamine acyltransferase in Escherichia coli. *J Bacteriol, 168*(1), 152-159.

Daniels, C., Vindurampulle, C., & Morona, R. (1998). Overexpression and topology of the Shigella flexneri O-antigen polymerase (Rfc/Wzy). *Mol Microbiol, 28*(6), 1211-1222.

Dardelle, F., Lehner, A., Ramdani, Y., Bardor, M., Lerouge, P., Driouich, A., & Mollet, J. C. (2010). Biochemical and immunocytological characterizations of Arabidopsis pollen tube cell wall. *Plant Physiol, 153*(4), 1563-1576. doi:10.1104/pp.110.158881

Darvill, A. G., McNeil, M., & Albersheim, P. (1978). Structure of plant cell walls: VIII. A new pectic polysaccharide. *Plant Physiol., 62*(3), 418-422.

Day, S. H., Chiu, N. Y., Tsao, L. T., Wang, J. P., & Lin, C. N. (2000). New lignan glycosides with potent antiinflammatory effect, isolated from Justicia ciliata. *J Nat Prod, 63*(11), 1560-1562.

Doong, R. L., Liljebjelke, K., Fralish, G., Kumar, A., & Mohnen, D. (1995). Cell-Free Synthesis of Pectin (Identification and Partial Characterization of Polygalacturonate 4-[alpha]-

Galacturonosyltransferase and Its Products from Membrane Preparations of Tobacco Cell-Suspension Cultures). *Plant Physiol, 109*(1), 141-152.

Dow, J. M., Osbourn, A. E., Wilson, T. J., & Daniels, M. J. (1995). A locus determining pathogenicity of Xanthomonas campestris is involved in lipopolysaccharide biosynthesis. *Mol Plant Microbe Interact, 8*(5), 768-777.

Driouich, A., Follet-Gueye, M. L., Bernard, S., Kousar, S., Chevalier, L., Vicre-Gibouin, M., & Lerouxel, O. (2012). Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Front Plant Sci, 3*, 79. doi:10.3389/fpls.2012.00079

Duchesne, L. C. a. L., D. W. (1989). Cellulose and the Evolution of Plant Life: The physical and biological properties of cellulose have made it the most abundant molecule in the biosphere. *BioScience, 39*(4), 238-241.

Duff, R. B. (1965). The occurrence of apiose in *Lemna* (duckweed) and other Angiosperms. *Biochem J, 94*, 768-777.

Dumont, M., Lehner, A., Bouton, S., Kiefer-Meyer, M. C., Voxeur, A., Pelloux, J., . . . Mollet, J. C. (2014). The cell wall pectic polymer rhamnogalacturonan-II is required for proper pollen tube elongation: implications of a putative sialyltransferase-like protein. *Ann. Bot., 114*(6), 1177-1188. doi:10.1093/aob/mcu093

Ebert, B., Rautengarten, C., Guo, X., Xiong, G., Stonebloom, S., Smith-Moritz, A. M., . . . Scheller, H. V. (2015). Identification and Characterization of a Golgi-Localized UDP-Xylose Transporter Family from Arabidopsis. *Plant Cell, 27*(4), 1218-1227. doi:10.1105/tpc.114.133827

Eckey-Kaltenbach, H., Heller, W., Sonnenbichler, J., Zetl, I., Schäfer, W., Ernst, D., & Sandermann, H. (1993). Oxidative stress and plant secondary metabolism: 6”-O-malonylapiin in parsley. *Phytochemistry, 34*, 687-691.

Edwards, M. E., Dickson, C. A., Chengappa, S., Sidebottom, C., Gidley, M. J., & Reid, J. S. (1999). Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *Plant J, 19*(6), 691-697.

Egelund, J., Damager, I., Faber, K., Olsen, C. E., Ulvskov, P., & Petersen, B. L. (2008). Functional characterisation of a putative rhamnogalacturonan II specific xylosyltransferase. *FEBS Lett, 582*(21-22), 3217-3222. doi:10.1016/j.febslet.2008.08.015

Egelund, J., Petersen, B. L., Motawia, M. S., Damager, I., Faik, A., Olsen, C. E., . . . Geshi, N. (2006). Arabidopsis thaliana RGXT1 and RGXT2 encode Golgi-localized (1,3)-alpha-D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. *Plant Cell, 18*(10), 2593-2607. doi:10.1105/tpc.105.036566

Eixelsberger, T., Horvat, D., Gutmann, A., Weber, H., & Nidetzky, B. (2017). Isotope Probing of the UDP-Apiose/UDP-Xylose Synthase Reaction: Evidence of a Mechanism via a Coupled Oxidation and Aldol Cleavage. *Angew Chem Int Ed Engl, 56*(9), 2503-2507. doi:10.1002/anie.201609288

Elsholz, A. K., Wacker, S. A., & Losick, R. (2014). Self-regulation of exopolysaccharide production in Bacillus subtilis by a tyrosine kinase. *Genes Dev, 28*(15), 1710-1720. doi:10.1101/gad.246397.114

Faik, A., Price, N. J., Raikhel, N. V., & Keegstra, K. (2002). An Arabidopsis gene encoding an alpha-xylosyltransferase involved in xyloglucan biosynthesis. *Proc Natl Acad Sci U S A, 99*(11), 7797-7802. doi:10.1073/pnas.102644799

Faridmoayer, A., Fentabil, M. A., Mills, D. C., Klassen, J. S., & Feldman, M. F. (2007). Functional characterization of bacterial oligosaccharyltransferases involved in O-linked protein glycosylation. *J Bacteriol, 189*(22), 8088-8098. doi:10.1128/JB.01318-07

Fernandez-Valdivia, R., Takeuchi, H., Samarghandi, A., Lopez, M., Leonardi, J., Haltiwanger, R. S., & Jafar-Nejad, H. (2011). Regulation of mammalian Notch signaling and embryonic development by the protein O-glucosyltransferase Rumi. *Development, 138*(10), 1925-1934. doi:10.1242/dev.060020

Ferrari, S., Savatin, D. V., Sicilia, F., Gramegna, G., Cervone, F., & Lorenzo, G. D. (2013). Oligogalacturonides: plant damage-associated molecular patterns and regulators of growth and development. *Front Plant Sci, 4*, 49. doi:10.3389/fpls.2013.00049

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science, 281*(5374), 237-240.

Fillgrove, K. L., & Anderson, V. E. (2001). The mechanism of dienoyl-CoA reduction by 2,4-dienoyl-CoA reductase is stepwise: observation of a dienolate intermediate. *Biochemistry, 40*(41), 12412-12421.

Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., . . . Oppermann, U. (2002). Critical residues for structure and catalysis in short-chain

dehydrogenases/reductases. *J Biol Chem, 277*(28), 25677-25684. doi:10.1074/jbc.M202160200

Fleischer, A., O'Neill, M. A., & Ehwald, R. (1999). The pore size of non-graminaceous plant cell walls is rapidly decreased by borate ester cross-linking of the pectic polysaccharide rhamnogalacturonan II. *Plant Physiol., 121*(3), 829-838.

Fletcher, C. M., Coyne, M. J., Bentley, D. L., Villa, O. F., & Comstock, L. E. (2007). Phase-variable expression of a family of glycoproteins imparts a dynamic surface to a symbiont in its human intestinal ecosystem. *Proc Natl Acad Sci U S A, 104*(7), 2413-2418. doi:10.1073/pnas.0608797104

Foesel, B. U., Gossner, A. S., Drake, H. L., & Schramm, A. (2007). Geminicoccus roseus gen. nov., sp. nov., an aerobic phototrophic Alphaproteobacterium isolated from a marine aquaculture biofilter. *Syst Appl Microbiol, 30*(8), 581-586. doi:10.1016/j.syapm.2007.05.005

Fogel, G. B., Tran, J., Johnson, S., & Hecht, D. (2010). Machine learning approaches for customized docking scores: Modeling of inhibition of Mycobacterium tuberculosis enoyl acyl carrier protein reductase. *Symposium on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2010*, 1-6.

Forsberg, L. S., Bhat, U. R., & Carlson, R. W. (2000). Structural characterization of the O-antigenic polysaccharide of the lipopolysaccharide from *Rhizobium etli* strain CE3: a unique O-acetylated glycan of discrete size, containing 3-O-methyl-6-deoxy-L-talose and 2,3,4-tri-O-methyl-L-fucose. *J Biol Chem, 275*, 18851–18863.

Gagneux, P., & Varki, A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology, 9*(8), 747-755.

Gardiner, D. M., Upadhyaya, N. M., Stiller, J., Ellis, J. G., Dodds, P. N., Kazan, K., & Manners, J. M. (2014). Genomic analysis of Xanthomonas translucens pathogenic on wheat and barley reveals cross-kingdom gene transfer events and diverse protein delivery systems. *PLoS One*, 9(1), e84995. doi:10.1371/journal.pone.0084995

Gardiner, S. E., Schroder, J., Matern, U., Hammer, D., & Hahlbrock, K. (1980). mRNA-dependent regulation of UDP-apiose synthase activity in irradiated plant cells. *J Biol Chem, 255*(22), 10752-10757.

Gebb, C., Baron, D., & Grisebach, H. (1975). Spectroscopic evidence for the formation of a 4-keto intermediate in the UDP-apiose/UDP-xylose synthase reaction. *Eur J Biochem, 54*(2), 493-498.

Glauner, B., Holtje, J. V., & Schwarz, U. (1988). The composition of the murein of Escherichia coli. *J Biol Chem, 263*(21), 10088-10095.

Gloaguen, V., Brudieux, V., Closs, B., Barbat, A., Krausz, P., Sainte-Catherine, O., . . . Guerardel, Y. (2010). Structural characterization and cytotoxic properties of an apiose-rich pectic polysaccharide obtained from the cell wall of the marine phanerogam *Zostera marina*. *J. Nat. Prod., 73*(6), 1087-1092. doi:10.1021/np100092c

Gloster, T. M. (2014). Advances in understanding glycosyltransferases from a structural perspective. *Curr Opin Struct Biol, 28*, 131-141. doi:10.1016/j.sbi.2014.08.012

Glushka, J. N., Terrell, M., York, W. S., O'Neill, M. A., Gucwa, A., Darvill, A. G., . . . Prestegard, J. H. (2003). Primary structure of the 2-O-methyl-alpha-L-fucose-containing side chain of the pectic polysaccharide, rhamnogalacturonan II. *Carbohydr. Res., 338*(4), 341-352.

Golovchenko, V. V., Ovodova, R. G., Shashkov, A. S., & Ovodov, Y. S. (2002). Structural studies of the pectic polysaccharide from duckweed *Lemna minor* L. *Phytochemistry, 60*(1), 89-97.

Gotting, C., Kuhn, J., Zahn, R., Brinkmann, T., & Kleesiek, K. (2000). Molecular cloning and expression of human UDP-D-xylose:proteoglycan core protein beta-d-xylosyltransferase and its first isoform XT-II. *J. Mol. Biol., 304*(4), 517-528. doi:10.1006/jmbi.2000.4261

Grimm, C., Maser, E., Mobus, E., Klebe, G., Reuter, K., & Ficner, R. (2000). The crystal structure of 3alpha -hydroxysteroid dehydrogenase/carbonyl reductase from Comamonas testosteroni shows a novel oligomerization pattern within the short chain dehydrogenase/reductase family. *J Biol Chem, 275*(52), 41333-41339. doi:10.1074/jbc.M007559200

Grisebach, H. (1980). Branched-chain sugars: occurance and biosynthesis. In e. Stumpf PK & Conn EE (Ed.), *Biochemistry of Plants: A Comprehensive Treatise* (pp. 171-197). New York: Academic Press.

Grisebach, H., & Döbereiner, U. (1964). The biosynthesis of apiose in parsley. *Biochem. Biophys. Res. Commun., 17*(6), 737-741.

Grizot, S., Salem, M., Vongsouthi, V., Durand, L., Moreau, F., Dohi, H., . . . Ducruix, A. (2006). Structure of the Escherichia coli heptosyltransferase WaaC: binary complexes with ADP

and ADP-2-deoxy-2-fluoro heptose. *J Mol Biol, 363*(2), 383-394. doi:10.1016/j.jmb.2006.07.057

Gu, X., & Bar-Peled, M. (2004). The biosynthesis of UDP-galacturonic acid in plants. Functional cloning and characterization of Arabidopsis UDP-D-glucuronic acid 4-epimerase. *Plant Physiol, 136*(4), 4256-4264. doi:10.1104/pp.104.052365

Gu, X., Glushka, J., Yin, Y., Xu, Y., Denny, T., Smith, J., . . . Bar-Peled, M. (2010). Identification of a bifunctional UDP-4-keto-pentose/UDP-xylose synthase in the plant pathogenic bacterium *Ralstonia solanacearum* strain GMI1000, a distinct member of the 4,6-dehydratase and decarboxylase family. *J. Biol. Chem., 285*(12), 9030-9040. doi:10.1074/jbc.M109.066803

Gu, X., Lee, S. G., & Bar-Peled, M. (2011). Biosynthesis of UDP-xylose and UDP-arabinose in *Sinorhizobium meliloti* 1021: first characterization of a bacterial UDP-xylose synthase, and UDP-xylose 4-epimerase. *Microbiology, 157*(Pt 1), 260-269. doi:10.1099/mic.0.040758-0

Guillemin, F., Guillon, F., Bonnin, E., Devaux, M. F., Chevalier, T., Knox, J. P., . . . Thibault, J. F. (2005). Distribution of pectic epitopes in cell walls of the sugar beet root. *Planta, 222*(2), 355-371. doi:10.1007/s00425-005-1535-3

Guo, Y., Sagaram, U. S., Kim, J. S., & Wang, N. (2010). Requirement of the galU gene for polysaccharide production by and pathogenicity and growth In Planta of Xanthomonas citri subsp. citri. *Appl Environ Microbiol, 76*(7), 2234-2242. doi:10.1128/AEM.02897-09

Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*, 310-322.

Guyett, P., Glushka, J., Gu, X., & Bar-Peled, M. (2009). Real-time NMR monitoring of intermediates and labile products of the bifunctional enzyme UDP-apiose/UDP-xylose synthase. *Carbohydr. Res., 344*(9), 1072-1078. doi:10.1016/j.carres.2009.03.026

Ha, S., Walker, D., Shi, Y., & Walker, S. (2000). The 1.9 A crystal structure of Escherichia coli MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. *Protein Sci, 9*(6), 1045-1052. doi:10.1110/ps.9.6.1045

Haberland, C., & Kolodziej, H. (1994). Novel Galloylhamameloses from Hamamelis virginiana1. *Planta Med, 60*(5), 464-466. doi:10.1055/s-2006-959533

Hahn, M. G., Darvill, A. G., & Albersheim, P. (1981). Host-Pathogen Interactions : XIX. The endogenous elicitor, a fragment of a plant cell wall polysaccharide that elicits phytoalexin accumulation in soybeans. *Plant Physiol, 68*(5), 1161-1169.

Hamburger, M, , G., M, , H., & K. (1985). Flavonol glycosides from *Securidaca diversifolia*. *Phytochemistry, 24*(2689–2692).

Harholt, J., Jensen, J. K., Sorensen, S. O., Orfila, C., Pauly, M., & Scheller, H. V. (2006). ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in Arabidopsis. *Plant Physiol, 140*(1), 49-58. doi:10.1104/pp.105.072744

Harholt, J., Jensen, J. K., Verhertbruggen, Y., Sogaard, C., Bernard, S., Nafisi, M., . . . Scheller, H. V. (2012). ARAD proteins associated with pectic Arabinan biosynthesis form complexes

when transiently overexpressed in planta. *Planta, 236*(1), 115-128. doi:10.1007/s00425-012-1592-3

Harholt, J., Suttangkakul, A., & Vibe Scheller, H. (2010). Biosynthesis of pectin. *Plant Physiol, 153*(2), 384-395. doi:10.1104/pp.110.156588

Harper, A. D., & 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*(4), 2188-2198. doi:10.1104/pp.009654

Hart, D. A., & Kindel, P. K. (1970). Isolation and partial characterization of apiogalacturonans from the cell wall of *Lemna minor*. *Biochem. J., 116*(4), 569-579.

Harz, H., Burgdorf, K., & Holtje, J. V. (1990). Isolation and separation of the glycan strands from murein of Escherichia coli by reversed-phase high-performance liquid chromatography. *Anal Biochem, 190*(1), 120-128.

Heinrichs, D. E., Yethon, J. A., & Whitfield, C. (1998). Molecular basis for structural diversity in the core regions of the lipopolysaccharides of Escherichia coli and Salmonella enterica. *Mol Microbiol, 30*(2), 221-232.

Heydanek, M. G., Jr., Struve, W. G., & Neuhaus, F. C. (1969). On the initial stage in peptidoglycan synthesis. 3. Kinetics and uncoupling of phospho-N-acetylmuramyl-pentapeptide translocase (uridine 5'-phosphate). *Biochemistry, 8*(3), 1214-1221.

Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., & Foust, T. D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science, 315*(5813), 804-807. doi:10.1126/science.1137016

Holtje, J. V. (1998). Growth of the stress-bearing and shape-maintaining murein sacculus of Escherichia coli. *Microbiol Mol Biol Rev, 62*(1), 181-203.

Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., . . . Walker, S. (2003). Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc Natl Acad Sci U S A, 100*(3), 845-849. doi:10.1073/pnas.0235749100

Hug, I., & Feldman, M. F. (2011). Analogies and homologies in lipopolysaccharide and glycoprotein biosynthesis in bacteria. *Glycobiology, 21*(2), 138-151. doi:10.1093/glycob/cwq148

Huson, D. H., & Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol., 61*(6), 1061-1067. doi:10.1093/sysbio/sys062

Hwang, S., Li, Z., Bar-Peled, Y., Aronov, A., Ericson, J., & Bar-Peled, M. (2014). The biosynthesis of UDP-D-FucNAc-4N-(2)-oxoglutarate (UDP-Yelosamine) in *Bacillus cereus* ATCC 14579: Pat and Pyl, an aminotransferase and an ATP-dependent Grasp protein that ligates 2-oxoglutarate to UDP-4-amino-sugars. *J. Biol. Chem., 289*(51), 35620-35632. doi:10.1074/jbc.M114.614917

Igura, M., Maita, N., Kamishikiryo, J., Yamada, M., Obita, T., Maenaka, K., & Kohda, D. (2008). Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J, 27*(1), 234-243. doi:10.1038/sj.emboj.7601940

Ishii, T., & Matsunaga, T. (1996). Isolation and characterization of a boron-rhamnogalacturonan-II complex from cell walls of sugar beet pulp. *Carbohydr. Res., 284*, 1–9.

Ishii, T., & Matsunaga, T. (2001). Pectic polysaccharide rhamnogalacturonan II is covalently linked to homogalacturonan. *Phytochemistry, 57*(6), 969-974.

Ishii, T., Matsunaga, T., Pellerin, P., O'Neill, M. A., Darvill, A., & Albersheim, P. (1999). The plant cell wall polysaccharide rhamnogalacturonan II self-assembles into a covalently cross-linked dimer. *J. Biol. Chem., 274*(19), 13098-13104.

Isogai, E., Isogai, H., Kurebayashi, Y., & Ito, N. (1986). Biological activities of leptospiral lipopolysaccharide. *Zentralbl Bakteriol Mikrobiol Hyg A, 261*(1), 53-64.

Jansson, P. E., Kenne, L., & Lindberg, B. (1975). Structure of extracellular polysaccharide from Xanthomonas campestris. *Carbohydr Res, 45*, 275-282.

Jarrell, K. F., Ding, Y., Meyer, B. H., Albers, S. V., Kaminski, L., & Eichler, J. (2014). N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol Mol Biol Rev, 78*(2), 304-341. doi:10.1128/MMBR.00052-13

Jensen, J. K., Sorensen, S. O., Harholt, J., Geshi, N., Sakuragi, Y., Moller, I., . . . Scheller, H. V. (2008). Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in Arabidopsis. *Plant Cell, 20*(5), 1289-1302. doi:10.1105/tpc.107.050906

Jones, L., Milne, J. L., Ashford, D., & McQueen-Mason, S. J. (2003). Cell wall arabinan is essential for guard cell function. *Proc Natl Acad Sci U S A, Sep 30;100*(20), 11783-11788.

Jornvall, H., Hoog, J. O., & Persson, B. (1999). SDR and MDR: completed genome sequences show these protein families to be large, of old origin, and of complex nature. *FEBS Lett, 445*(2-3), 261-264.

Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J., & Ghosh, D. (1995). Short-chain dehydrogenases/reductases (SDR). *Biochemistry, 34*(18), 6003-6013.

Julião, LS, , L., SG, , L., C, . . . GG. (2010). Flavones and phenylpropanoids from a sedative extract of *Lantana trifolia* L*Phytochemistry71*294–300

Julião, LS, , P., AL, , M., S, . . . L. (2009). Phenylethanoid glycosides from *Lantana fucata* with in vitro anti-inflammatory activity. *J Nat Prod, 72*, 1424–1428.

Kahler, C. M., Carlson, R. W., Rahman, M. M., Martin, L. E., & Stephens, D. S. (1996). Two glycosyltransferase genes, lgtF and rfaK, constitute the lipooligosaccharide ice (inner core extension) biosynthesis operon of Neisseria meningitidis. *J Bacteriol, 178*(23), 6677-6684.

Kallberg, Y., Oppermann, U., Jornvall, H., & Persson, B. (2002). Short-chain dehydrogenases/reductases (SDRs). *Eur J Biochem, 269*(18), 4409-4417.

Kanipes, M. I., Tan, X., Akelaitis, A., Li, J., Rockabrand, D., Guerry, P., & Monteiro, M. A. (2008). Genetic analysis of lipooligosaccharide core biosynthesis in Campylobacter jejuni 81-176. *J Bacteriol, 190*(5), 1568-1574. doi:10.1128/JB.01696-07

Kapitonov, D., & Yu, R. K. (1999). Conserved domains of glycosyltransferases. *Glycobiology, 9*(10), 961-978.

Katzen, F., Ferreiro, D. U., Oddo, C. G., Ielmini, M. V., Becker, A., Puhler, A., & Ielpi, L. (1998). Xanthomonas campestris pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. *J Bacteriol, 180*(7), 1607-1617.

Kauss, H., Swanson, A. L., & Hassid, W. Z. (1967). Biosynthesis of the methyl ester groups of pectin by transmethylation from S-adenosyl-L-methionine. *Biochem Biophys Res Commun, 26*(2), 234-240.

Kavanagh, K. L., Jornvall, H., Persson, B., & Oppermann, U. (2008). Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol. Life Sci., 65*(24), 3895-3906. doi:10.1007/s00018-008-8588-y

Kay, E., Lesk, V. I., Tamaddoni-Nezhad, A., Hitchen, P. G., Dell, A., Sternberg, M. J., . . . Wren, B. W. (2010). Systems analysis of bacterial glycomes. *Biochem Soc Trans, 38*(5), 1290-1293. doi:10.1042/BST0381290

Kessler, E., & Safrin, M. (1988). Synthesis, processing, and transport of Pseudomonas aeruginosa elastase. *J Bacteriol, 170*(11), 5241-5247.

Kim, M. H., Kim, Y., Park, H. J., Lee, J. S., Kwak, S. N., Jung, W. H., . . . Oh, T. K. (2008). Structural insight into bioremediation of triphenylmethane dyes by Citrobacter sp. triphenylmethane reductase. *J Biol Chem, 283*(46), 31981-31990. doi:10.1074/jbc.M804092200

Kindel, P. K., & Watson, R. R. (1973). Synthesis, characterization and properties of uridine 5'-(-D-apio-D-furanosyl pyrophosphate). *Biochem J, 133*(2), 227-241.

Klinghammer, M., & Tenhaken, R. (2007). Genome-wide analysis of the UDP-glucose dehydrogenase gene family in Arabidopsis, a key enzyme for matrix polysaccharides in cell walls. *J Exp Bot, 58*(13), 3609-3621. doi:10.1093/jxb/erm209

Klutts, J. S., & Doering, T. L. (2008). Cryptococcal xylosyltransferase 1 (Cxt1p) from Cryptococcus neoformans plays a direct role in the synthesis of capsule polysaccharides. *J Biol Chem, 283*(21), 14327-14334. doi:10.1074/jbc.M708927200

Klutts, J. S., Levery, S. B., & Doering, T. L. (2007). A beta-1,2-xylosyltransferase from Cryptococcus neoformans defines a new family of glycosyltransferases. *J Biol Chem, 282*(24), 17890-17899. doi:10.1074/jbc.M701941200

Knirel, Y. A. (1990). Polysaccharide antigens of Pseudomonas aeruginosa. *Crit Rev Microbiol, 17*(4), 273-304. doi:10.3109/10408419009105729

Knox, J. P. (1992). Molecular probes for the plant cell surface. *Protoplasma, 167*, 1-9.

Kobayashi, M., Kouzu, N., Inami, A., Toyooka, K., Konishi, Y., Matsuoka, K., & Matoh, T. (2011). Characterization of Arabidopsis CTP:3-deoxy-D-manno-2-octulosonate cytidylyltransferase (CMP-KDO synthetase), the enzyme that activates KDO during rhamnogalacturonan II biosynthesis. *Plant Cell Physiol, 52*(10), 1832-1843. doi:10.1093/pcp/pcr120

Kobayashi, M., Matoh, T., & Azuma, J. (1996). Two chains of rhamnogalacturonan II are cross-linked by borate-diol ester bonds in higher plant cell walls. *Plant Physiol., 110*(3), 1017-1020.

Kobayashi, M., Nakagawa, H., Asaka, T., & Matoh, T. (1999). Borate-rhamnogalacturonan II bonding reinforced by Ca2+ retains pectic polysaccharides in higher-plant cell walls. *Plant Physiol, 119*(1), 199-204.

Kobayashi, M., Nakagawa, H., Suda, I., Miyagawa, I., & Matoh, T. (2002). Purification and cDNA cloning of UDP-D-glucuronate carboxy-lyase (UDP-D-xylose synthase) from pea seedlings. *Plant Cell Physiol., 43*(11), 1259-1265.

Kotake, T., Yamaguchi, D., Ohzono, H., Hojo, S., Kaneko, S., Ishida, H. K., & Tsumuraya, Y. (2004). UDP-sugar pyrophosphorylase with broad substrate specificity toward various monosaccharide 1-phosphates from pea sprouts. *J Biol Chem, 279*(44), 45728-45736. doi:10.1074/jbc.M408716200

Kowarik, M., Numao, S., Feldman, M. F., Schulz, B. L., Callewaert, N., Kiermaier, E., . . . Aebi, M. (2006a). N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science, 314*(5802), 1148-1150. doi:10.1126/science.1134351

Kowarik, M., Young, N. M., Numao, S., Schulz, B. L., Hug, I., Callewaert, N., . . . Aebi, M. (2006b). Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J, 25*(9), 1957-1966. doi:10.1038/sj.emboj.7601087

Kuhn, J., Gotting, C., Schnolzer, M., Kempf, T., Brinkmann, T., & Kleesiek, K. (2001). First isolation of human UDP-D-xylose: proteoglycan core protein beta-D-xylosyltransferase

secreted from cultured JAR choriocarcinoma cells. *J. Biol. Chem., 276*(7), 4940-4947. doi:10.1074/jbc.M005111200

Kus, J. V., Kelly, J., Tessier, L., Harvey, H., Cvitkovitch, D. G., & Burrows, L. L. (2008). Modification of Pseudomonas aeruginosa Pa5196 type IV Pilins at multiple sites with D-Araf by a novel GT-C family Arabinosyltransferase, TfpW. *J Bacteriol, 190*(22), 7464-7478. doi:10.1128/JB.01075-08

Lairson, L. L., Henrissat, B., Davies, G. J., & Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem, 77*, 521-555. doi:10.1146/annurev.biochem.76.061005.092322

Leemhuis, H., Kragh, K. M., Dijkstra, B. W., & Dijkhuizen, L. (2003). Engineering cyclodextrin glycosyltransferase into a starch hydrolase with a high exo-specificity. *J Biotechnol*, 103(3), 203-212.

Lehner, A., Dardelle, F., Soret-Morvan, O., Lerouge, P., Driouich, A., & Mollet, J. C. (2010). Pectins in the cell wall of Arabidopsis thaliana pollen tube and pistil. *Plant Signal Behav, 5*(10), 1282-1285. doi:10.4161/psb.5.10.13040

Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Curr Opin Struct Biol, 5*(6), 775-783.

Li, J., & Wang, N. (2011). Genome-wide mutagenesis of Xanthomonas axonopodis pv. citri reveals novel genetic determinants and regulation mechanisms of biofilm formation. *PLoS One, 6*(7), e21804. doi:10.1371/journal.pone.0021804

Li, J., & Wang, N. (2012). The gpsX gene encoding a glycosyltransferase is important for polysaccharide production and required for full virulence in Xanthomonas citri subsp. citri. *BMC Microbiol, 12*, 31. doi:10.1186/1471-2180-12-31

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., . . . Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res., 43*(W1), W580-584. doi:10.1093/nar/gkv279

Li, Z., Hwang, S., & Bar-Peled, M. (2016). Discovery of a Unique Extracellular Polysaccharide in Members of the Pathogenic Bacillus That Can Co-form with Spores. *J Biol Chem, 291*(36), 19051-19067. doi:10.1074/jbc.M116.724708

Liang, D. M., Liu, J. H., Wu, H., Wang, B. B., Zhu, H. J., & Qiao, J. J. (2015). Glycosyltransferases: mechanisms and applications in natural product development. *Chem Soc Rev, 44*(22), 8350-8374. doi:10.1039/c5cs00600g

Liepman, A. H., Wilkerson, C. G., & Keegstra, K. (2005). Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proc Natl Acad Sci U S A, 102*(6), 2221-2226. doi:10.1073/pnas.0409179102

Lin, T. Y., Elbein, A. D., & Su, J. C. (1966). Substrate specificity in pectin synthesis. *Biochem Biophys Res Commun, 22*(6), 650-657.

Lindenthal, C., & Elsinghorst, E. A. (1999). Identification of a glycoprotein produced by enterotoxigenic Escherichia coli. *Infect Immun, 67*(8), 4084-4091.

Liu, J., & Mushegian, A. (2003). Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci, 12*(7), 1418-1431. doi:10.1110/ps.0302103

Liu, X. L., Liu, L., Niu, Q. K., Xia, C., Yang, K. Z., Li, R., . . . Ye, D. (2011). Male gametophyte defective 4 encodes a rhamnogalacturonan II xylosyltransferase and is important for growth of pollen tubes and roots in Arabidopsis. *Plant J, 65*(4), 647-660. doi:10.1111/j.1365-313X.2010.04452.x

Liu, Y., Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Ruzicka, F. J., . . . Frey, P. A. (1997). Mechanistic roles of tyrosine 149 and serine 124 in UDP-galactose 4-epimerase from Escherichia coli. *Biochemistry, 36*(35), 10675-10684. doi:10.1021/bi970430a

Liwanag, A. J., Ebert, B., Verhertbruggen, Y., Rennie, E. A., Rautengarten, C., Oikawa, A., . . . Scheller, H. V. (2012). Pectin biosynthesis: GALS1 in Arabidopsis thaliana is a beta-1,4-galactan beta-1,4-galactosyltransferase. *Plant Cell, 24*(12), 5024-5036. doi:10.1105/tpc.112.106625

Lizak, C., Gerber, S., Numao, S., Aebi, M., & Locher, K. P. (2011). X-ray structure of a bacterial oligosaccharyltransferase. *Nature, 474*(7351), 350-355. doi:10.1038/nature10151

Logan, S. M. (2006). Flagellar glycosylation - a new component of the motility repertoire? *Microbiology, 152*(Pt 5), 1249-1262. doi:10.1099/mic.0.28735-0

Longland, J. M., Fry, S. C., & Trewavas, A. J. (1989). Developmental control of apiogalacturonan biosynthesis and UDP-apiose production in a duckweed. *Plant Physiol., 90*(3), 972-976.

Lovering, A. L., de Castro, L. H., Lim, D., & Strynadka, N. C. (2007). Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science, 315*(5817), 1402-1405. doi:10.1126/science.1136611

Macpherson, D. F., Manning, P. A., & Morona, R. (1994). Characterization of the dTDP-rhamnose biosynthetic genes encoded in the rfb locus of Shigella flexneri. *Mol Microbiol, 11*(2), 281-292.

Madson, M., Dunand, C., Li, X., Verma, R., Vanzin, G. F., Caplan, J., . . . Reiter, W. D. (2003). The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *Plant Cell, 15*(7), 1662-1670.

Malafronte, N., Pesca, M. S., Bisio, A., Morales Escobar, L., & De Tommasi, N. (2009). New flavonoid glycosides from Vernonia ferruginea. *Nat Prod Commun, 4*(12), 1639-1642.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., . . . Bryant, S. H. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res, 45*(D1), D200-D203. doi:10.1093/nar/gkw1129

Maris, A., Suslov, D., Fry, S. C., Verbelen, J. P., & Vissenberg, K. (2009). Enzymic characterization of two recombinant xyloglucan endotransglucosylase/hydrolase (XTH) proteins of Arabidopsis and their effect on root growth and cell wall extension. *J Exp Bot, 60*(13), 3959-3972. doi:10.1093/jxb/erp229

Marolda, C. L., & Valvano, M. A. (1995). Genetic analysis of the dTDP-rhamnose biosynthesis region of the Escherichia coli VW187 (O7:K1) rfb gene cluster: identification of functional homologs of rfbB and rfbA in the rff cluster and correct location of the rffE gene. *J Bacteriol, 177*(19), 5539-5546.

Marrec-Fairley, M., Piette, A., Gallet, X., Brasseur, R., Hara, H., Fraipont, C., . . . Nguyen-Disteche, M. (2000). Differential functionalities of amphiphilic peptide segments of the

cell-septation penicillin-binding protein 3 of Escherichia coli. *Mol Microbiol, 37*(5), 1019-1031.

Mascaro, L. J., Jr., & Kindel, P. K. (1977). Characterization of [14C]apiogalacturonans synthesized in a cell-free system from *Lemna minor*. *Arch. Biochem. Biophys., 183*(1), 139-148.

Matasci, N., Hung, L. H., Yan, Z., Carpenter, E. J., Wickett, N. J., Mirarab, S., . . . Wong, G. K. (2014). Data access for the 1,000 Plants (1KP) project. *Gigascience, 3*, 17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25625010 doi:10.1186/2047-217X-3-17

Matern, U., & Grisebach, H. (1977). UDP-apiose/UDP-xylose synthase. Subunit composition and binding studies. *Eur. J. Biochem., 74*(2), 303-312.

Matias, V. R., Al-Amoudi, A., Dubochet, J., & Beveridge, T. J. (2003). Cryo-transmission electron microscopy of frozen-hydrated sections of Escherichia coli and Pseudomonas aeruginosa. *J Bacteriol, 185*(20), 6112-6118.

Matoh, T., Ishigaki, K., Ohno, K., & Azuma, J. (1993). Isolation and characterization of a boron-polysaccharide complex from radish roots. *Plant Cell Physiol., 34*, 639– 642.

Matoh, T., Kawaguchi, S., & Kobayashi, M. (1996). Ubiquity of a borate rhamnogalacturonan II complex in the cell walls of higher plants. *Plant Cell Physiol., 37*, 636–640.

Matsunga, T., Ishii, T., Matsumoto, S., Higuchi, M., Darvill, A., Albersheim, P., & O'Neill, M. A. (2003). Occurrence of the Primary Cell Wall Polysaccharide Rhamnogalacturonan II in Pteridophytes, Lycophytes, and Bryophytes. Implications for the Evolution of Vascular Plants. *Plant Physiol, 134*(339-351).

Maurya, R, , W., V, , K., A, . . . RS. (1996). Cordifoliosides A and B, two new phenylpropene disaccharides from Tinospora cordifolia possessing immunostimulant activity*Nat Prod Lett8*, 7–10.

McCarthy, T. W., Der, J. P., Honaas, L. A., dePamphilis, C. W., & Anderson, C. T. (2014). Phylogenetic analysis of pectin-related gene families in *Physcomitrella patens* and nine other plant species yields evolutionary insights into cell walls. *BMC Plant Biol., 14*, 79. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24666997 doi:10.1186/1471-2229-14-79

McGowen, M. M., Vionnet, J., & Vann, W. F. (2001). Elongation of alternating alpha 2,8/2,9 polysialic acid by the Escherichia coli K92 polysialyltransferase. *Glycobiology, 11*(8), 613-620.

McNab, J. M., Villemez, C. L., & Albersheim, P. (1968). Biosynthesis of galactan by a particulate enzyme preparation from Phaseolus aureus seedlings. *Biochem J, 106*(2), 355-360.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., . . . Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res., 41*(Web Server issue), W597-600. doi:10.1093/nar/gkt376

Melton, L. D., Mindt, L., Rees, D. A., & Sanderson, G. R. (1976). Covalent structure of the extracellular polysaccharide from Xanthomonas campestris: evidence from partial hydrolysis studies. *Carbohydr Res, 46*(2), 245-257.

Mendicino, J., & Abou-Issa, H. (1974). Conversion of UDP-D-glucuronic acid to UDP-D-apiose and UDP-D-xylose by an enzyme isolated from Lemna minor. *Biochim Biophys Acta, 364*(1), 159-172.

Messner, P. (2004). Prokaryotic glycoproteins: unexplored but important. *J Bacteriol, 186*(9), 2517-2519.

Miljković, M. (1990). *Addition of nucleophiles to glycopyranosiduloses*. New York: Springer-Verlag.

Minic, Z., Marie, C., Delorme, C., Faurie, J. M., Mercier, G., Ehrlich, D., & Renault, P. (2007). Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in Streptococcus thermophilus, by EpsD tyrosine kinase. *J Bacteriol, 189*(4), 1351-1357. doi:10.1128/JB.01122-06

Mohnen, D. (2002). *Biosynthesis of pectins*. Oxford: Blackwell Publishing and CRC Press.

Mohnen, D. (2008). Pectin structure and biosynthesis. *Curr Opin Plant Biol, 11*(3), 266-277. doi:10.1016/j.pbi.2008.03.006

Molhoj, M., Verma, R., & Reiter, W. D. (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, W. D. (2004). The biosynthesis of D-Galacturonate in plants. functional cloning and characterization of a membrane-anchored UDP-D-Glucuronate 4-epimerase from Arabidopsis. *Plant Physiol, 135*(3), 1221-1230. doi:10.1104/pp.104.043745

Moore, P. J., Swords, K. M., Lynch, M. A., & Staehelin, L. A. (1991). Spatial organization of the assembly pathways of glycoproteins and complex polysaccharides in the Golgi apparatus of plants. *J Cell Biol, 112*(4), 589-602.

Mutwil, M., Debolt, S., & Persson, S. (2008). Cellulose synthesis: a complex complex. *Curr Opin Plant Biol, 11*(3), 252-257. doi:10.1016/j.pbi.2008.03.007

Nakamura, A., Furuta, H., Maeda, H., Takao, T., & Nagamatsu, Y. (2002). Structural studies by stepwise enzymatic degradation of the main backbone of soybean soluble polysaccharides consisting of galacturonan and rhamnogalacturonan. *Biosci Biotechnol Biochem, 66*(6), 1301-1313.

Neilson, E. H., Goodger, J. Q., Woodrow, I. E., & Møller, B. L. (2013). Plant chemical defense: At what cost? *Trends Plant Science, 18*, 250–258.

Nepogodiev, S. A., Fais, M., Hughes, D. L., & Field, R. A. (2011). Synthesis of apiose-containing oligosaccharide fragments of the plant cell wall: fragments of rhamnogalacturonan-II side chains A and B, and apiogalacturonan. *Org Biomol Chem, 9*(19), 6670-6684. doi:10.1039/c1ob05587a

Nguema-Ona, E., Andeme-Onzighi, C., Aboughe-Angone, S., Bardor, M., Ishii, T., Lerouge, P., & Driouich, A. (2006). The reb1-1 mutation of Arabidopsis. Effect on the structure and localization of galactose-containing cell wall polysaccharides. *Plant Physiol, 140*(4), 1406-1417. doi:10.1104/pp.105.074997

Nishiyama, T., Hiwatashi, Y., Sakakibara, I., Kato, M., & Hasebe, M. (2000). Tagged mutagenesis and gene-trap in the moss, *Physcomitrella patens* by shuttle mutagenesis. *DNA Res., 7*(1), 9-17.

O'Neill, M. A., Albersheim, P., & Darvill, A. G. (1990). The pectic polysaccharides of primary cell walls. In D. D.M. (Ed.), *Methods in Plant Biochemistry* (2 ed., pp. 415-441). London: Academic Press.

O'Neill, M. A., Eberhard, S., Albersheim, P., & Darvill, A. G. (2001). Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. *Science, 294*(5543), 846-849. doi:10.1126/science.1062319

O'Neill, M. A., Ishii, T., Albersheim, P., & Darvill, A. G. (2004). Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annu Rev Plant Biol, 55*, 109-139. doi:10.1146/annurev.arplant.55.031903.141750

O'Neill, M. A., Warrenfeltz, D., Kates, K., Pellerin, P., Doco, T., Darvill, A. G., & Albersheim, P. (1996). Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester. In vitro conditions for the formation and hydrolysis of the dimer. *J. Biol. Chem., 271*(37), 22923-22930.

O'Neill, M. A., & York, W. S. (2003). *The composition and structure of plant primary cell walls*. Ithaca, New York: Blackwell Publishing/CRC Press,.

Odzuck, W., & Kauss, H. (1972). Biosynthesis of pure araban and xylan. *Phytochemistry, 11*, 2489–2494.

Oikawa, A., Lund, C. H., Sakuragi, Y., & Scheller, H. V. (2013). Golgi-localized enzyme complexes for plant cell wall biosynthesis. *Trends Plant Sci, 18*(1), 49-58. doi:10.1016/j.tplants.2012.07.002

Ono, H., Kuwahara, Y., & Nishida, R. (2004). Hydroxybenzoic acid derivatives in a nonhost rutaceous plant, Orixajaponica, deter both oviposition and larval feeding in a rutaceae-feeding swallowtail butterfly, Papilio xuthus L. *J Chem Ecol, 30*(2), 287-301.

Oppermann, U. C., Filling, C., Berndt, K. D., Persson, B., Benach, J., Ladenstein, R., & Jornvall, H. (1997). Active site directed mutagenesis of 3 beta/17 beta-hydroxysteroid dehydrogenase establishes differential effects on short-chain dehydrogenase/reductase reactions. *Biochemistry, 36*(1), 34-40. doi:10.1021/bi961803v

Orfila, C., Sorensen, S. O., Harholt, J., Geshi, N., Crombie, H., Truong, H. N., . . . Scheller, H. V. (2005). QUASIMODO1 is expressed in vascular tissue of Arabidopsis thaliana inflorescence stems, and affects homogalacturonan and xylan biosynthesis. *Planta, 222*(4), 613-622. doi:10.1007/s00425-005-0008-z

Ortiz-Ramirez, C., Hernandez-Coronado, M., Thamm, A., Catarino, B., Wang, M., Dolan, L., . . . Becker, J. D. (2016). A transcriptome atlas of *Physcomitrella patens* provides insights into the evolution and development of land plants. *Mol. Plant, 9*(2), 205-220. doi:10.1016/j.molp.2015.12.002

Ortmann, R., Sutter, A., & Grisebach, H. (1972). Purification and properties of UDP-apiose: 7-O-(-D-glucosyl)-flavone apiosyltransferase from cell suspension cultures of parsley. *Biochem. Biophys. Acta, 289*(2), 293-302.

Ozawa, T., Kobayashi, S., Seki, R., Imagawa, H. (1984). A new gallotannin from bark of chestnut tree, *Castanea crenata* Sieb. et Zucc. *Agric Biol Chem, 48*, 1411–1416

Pabst, M., Fischl, R. M., Brecker, L., Morelle, W., Fauland, A., Kofeler, H., . . . Leonard, R. (2013). Rhamnogalacturonan II structure shows variation in the side chains monosaccharide composition and methylation status within and across different plant species. *Plant J., 76*(1), 61-72. doi:10.1111/tpj.12271

Pattathil, S., Harper, A. D., & Bar-Peled, M. (2005). Biosynthesis of UDP-xylose: characterization of membrane-bound AtUxs2. *Planta, 221*(4), 538-548. doi:10.1007/s00425-004-1471-7

Pelkonen, S., & Finne, J. (1989). Polyacrylamide gel electrophoresis of capsular polysaccharides of bacteria. *Methods Enzymol, 179*, 104-110.

Perez, S., Rodriguez-Carvajal, M. A., & Doco, T. (2003). A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. *Biochimie, 85*(1-2), 109-121.

Perrin, R. M., DeRocher, A. E., Bar-Peled, M., Zeng, W., Norambuena, L., Orellana, A., . . . Keegstra, K. (1999). Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis. *Science, 284*(5422), 1976-1979.

Picmanova, M., & Moller, B. L. (2016). Apiose: one of nature's witty games. *Glycobiology, 26*(5), 430-442. doi:10.1093/glycob/cww012

Plesiat, P., & Nikaido, H. (1992). Outer membranes of gram-negative bacteria are permeable to steroid probes. *Mol Microbiol, 6*(10), 1323-1333.

Polizzi, S. J., Walsh, R. M., Jr., Peeples, W. B., Lim, J. M., Wells, L., & Wood, Z. A. (2012). Human UDP-alpha-D-xylose synthase and *Escherichia coli* ArnA conserve a conformational shunt

that controls whether xylose or 4-keto-xylose is produced. *Biochemistry, 51*(44), 8844-8855. doi:10.1021/bi301135b

Poorter, H. a. V., R. (1997). *Plant Resource Allocation*. San Diego, CA: Academic Press.

Popper, Z. A., & Fry, S. C. (2003). Primary cell wall composition of bryophytes and charophytes. *Ann. Bot., 91*(1), 1-12.

Popper, Z. A., & Fry, S. C. (2008). Xyloglucan-pectin linkages are formed intra-protoplasmically, contribute to wall-assembly, and remain stable in the cell wall. *Planta, 227*(4), 781-794. doi:10.1007/s00425-007-0656-2

Raetz, C. R. (1990). Biochemistry of endotoxins. *Annu Rev Biochem, 59*, 129-170. doi:10.1146/annurev.bi.59.070190.001021

Raetz, C. R. (1993). Bacterial endotoxins: extraordinary lipids that activate eucaryotic signal transduction. *J Bacteriol, 175*(18), 5745-5753.

Raetz, C. R., & Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annu Rev Biochem, 71*, 635-700. doi:10.1146/annurev.biochem.71.110601.135414

Raetz, C. R. H. (1996). *Escherichia coli and Salmonella* Washington, DC: Am. Soc. Microbiol. .

Ramkumar, N., Harvey, B. M., Lee, J. D., Alcorn, H. L., Silva-Gagliardi, N. F., McGlade, C. J., . . . Anderson, K. V. (2015). Protein O-Glucosyltransferase 1 (POGLUT1) Promotes Mouse Gastrulation through Modification of the Apical Polarity Protein CRUMBS2. *PLoS Genet, 11*(10), e1005551. doi:10.1371/journal.pgen.1005551

Rautengarten, C., Birdseye, D., Pattathil, S., McFarlane, H. E., Saez-Aguayo, S., Orellana, A., . . . Ebert, B. (2017). The elaborate route for UDP-arabinose delivery into the Golgi of plants. *Proc Natl Acad Sci U S A, 114*(16), 4261-4266. doi:10.1073/pnas.1701894114

Rautengarten, C., Ebert, B., Liu, L., Stonebloom, S., Smith-Moritz, A. M., Pauly, M., . . . Heazlewood, J. L. (2016). The Arabidopsis Golgi-localized GDP-L-fucose transporter is required for plant development. *Nat Commun, 7*, 12119. doi:10.1038/ncomms12119

Rautengarten, C., Ebert, B., Moreno, I., Temple, H., Herter, T., Link, B., . . . Orellana, A. (2014). The Golgi localized bifunctional UDP-rhamnose/UDP-galactose transporter family of Arabidopsis. *Proc Natl Acad Sci U S A, 111*(31), 11563-11568. doi:10.1073/pnas.1406073111

Rautengarten, C., Ebert, B., Ouellet, M., Nafisi, M., Baidoo, E. E., Benke, P., . . . Scheller, H. V. (2012). Arabidopsis Deficient in Cutin Ferulate encodes a transferase required for feruloylation of omega-hydroxy fatty acids in cutin polyester. *Plant Physiol, 158*(2), 654-665. doi:10.1104/pp.111.187187

Reboul, R., Geserick, C., Pabst, M., Frey, B., Wittmann, D., Lutz-Meindl, U., . . . Tenhaken, R. (2011). Down-regulation of UDP-glucuronic acid biosynthesis leads to swollen plant cell walls and severe developmental defects associated with changes in pectic polysaccharides. *J Biol Chem, 286*(46), 39982-39992. doi:10.1074/jbc.M111.255695

Reboul, R., & Tenhaken, R. (2012). An emerging role of pectic rhamnogalacturonanII for cell wall integrity. *Plant Signal Behav, 7*(2), 298-299. doi:10.4161/psb.18894

Reiter, W. D. (2008). Biochemical genetics of nucleotide sugar interconversion reactions. *Curr Opin Plant Biol, 11*(3), 236-243. doi:10.1016/j.pbi.2008.03.009

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., . . . Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science, 319*(5859), 64-69. doi:10.1126/science.1150646

Rezanka, T., & Guschina, I. A. (2000). Glycosidic compounds of murolic, protoconstipatic and allo-murolic acids from lichens of Central Asia. *Phytochemistry, 54*(6), 635-645.

Rezanka, T., Nedbalova, L., Kolouchova, I., & Sigler, K. (2013). LC-MS/APCI identification of glucoside esters and diesters of astaxanthin from the snow alga Chlamydomonas nivalis including their optical stereoisomers. *Phytochemistry, 88*, 34-42. doi:10.1016/j.phytochem.2013.01.003

Ridley, B. L., O'Neill, M. A., & Mohnen, D. (2001). Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry, 57*(6), 929-967.

Roberts, A. W., Roberts, E. M., & Haigler, C. H. (2012). Moss cell walls: structure and biosynthesis. *Front. Plant Sci., 3*, 166. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22833752 doi:10.3389/fpls.2012.00166

Robertson, B. D., Frosch, M., & van Putten, J. P. (1994). The identification of cryptic rhamnose biosynthesis genes in Neisseria gonorrhoeae and their relationship to lipopolysaccharide biosynthesis. *J Bacteriol, 176*(22), 6915-6920.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol, 5*, 172. doi:10.3389/fmicb.2014.00172

Round, A. N., Rigby, N. M., MacDougall, A. J., & Morris, V. J. (2010). A new view of pectin structure revealed by acid hydrolysis and atomic force microscopy. *Carbohydr Res, 345*(4), 487-497. doi:10.1016/j.carres.2009.12.019

Ruiz, N. (2008). Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in Escherichia coli. *Proc Natl Acad Sci U S A, 105*(40), 15553-15557. doi:10.1073/pnas.0808352105

Sahdev, S., Khattar, S. K., & Saini, K. S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem, 307*(1-2), 249-264. doi:10.1007/s11010-007-9603-6

Sahni-Arya, B., Flynn, M. J., Bergeron, L., Salyan, M. E., Pedicord, D. L., Golla, R., . . . Blat, Y. (2007). Cofactor-specific modulation of 11beta-hydroxysteroid dehydrogenase 1 inhibitor potency. *Biochim Biophys Acta, 1774*(9), 1184-1191. doi:10.1016/j.bbapap.2007.07.005

Salton, M. R. (1963). The relationship between the nature of the cell wall and the Gram stain. *J Gen Microbiol, 30*, 223-235. doi:10.1099/00221287-30-2-223

Sandermann, H., Jr., & Grisebach, H. (1970). Biosynthesis of D-apiose. V. NAD+-dependent biosynthesis of UDP-apiose and UDP-xylose from UDP-D-glucuronic acid with an enzyme preparation from Lemna minor L. *Biochim Biophys Acta, 208*(2), 173-180.

Sandermann, H., Jr., Tisue, G. T., & Grisebach, H. (1968). Biosynthesis of D-apiose. IV. Formation of UDP-apiose from UDP-D-glucuronic acid in cell-free extracts of parsley (*Apium petroselinum* L.) and *Lemna minor*. *Biochem. Biophys. Acta, 165*(3), 550-552.

Sanderson, K. E., MacAlister, T., Costerton, J. W., & Cheng, K. J. (1974). Permeability of lipopolysaccharide-deficient (rough) mutants of Salmonella typhimurium to antibiotics, lysozyme, and other agents. *Can J Microbiol, 20*(8), 1135-1145.

Schirm, M., Soo, E. C., Aubry, A. J., Austin, J., Thibault, P., & Logan, S. M. (2003). Structural, genetic and functional characterization of the flagellin glycosylation process in Helicobacter pylori. *Mol Microbiol, 48*(6), 1579-1592.

Schmidt, H., Hansen, G., Singh, S., Hanuszkiewicz, A., Lindner, B., Fukase, K., . . . Mesters, J. R. (2012). Structural and mechanistic analysis of the membrane-embedded glycosyltransferase WaaA required for lipopolysaccharide synthesis. *Proc Natl Acad Sci U S A, 109*(16), 6253-6258. doi:10.1073/pnas.1119894109

Schmidt, O. T. (1930). Konstitution und Konfiguration der Apiose. Über Zucker mit verzweigter Kohlenstoffkette II. *Liebigs Ann Chem 483*115-123.

Schnurr, J. A., Storey, K. K., Jung, H. J., Somers, D. A., & Gronwald, J. W. (2006). UDP-sugar pyrophosphorylase is essential for pollen development in Arabidopsis. *Planta, 224*(3), 520-532. doi:10.1007/s00425-006-0240-1

Sellmair, J., & Beck, E. (1968). Isolierung und identifizierung von hamamelit aus *Primula clusiana* Tausch. Z. *Pfianzenphysiol, 59*, 70-90.

Shafizadeh, F. (1956). Branched-chain sugars of natural occurrence. *Adv Carbohydr Chem, 48*(11), 263-283.

Sheikh, M. O., Halmo, S. M., Patel, S., Middleton, D., Takeuchi, H., Schafer, C. M., . . . Wells, L. (2017). Rapid screening of sugar-nucleotide donor specificities of putative glycosyltransferases. *Glycobiology*, 27(3), 206-212. doi:10.1093/glycob/cww114

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., . . . Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol., 7*, 539. doi:10.1038/msb.2011.75

Simossis, V. A., & Heringa, J. (2005). PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Res., 33*(Web Server issue), W289-294. doi:10.1093/nar/gki390

Sindhuwinata, N., Munoz, E., Munoz, F. J., Palcic, M. M., Peters, H., & Peters, T. (2010). Binding of an acceptor substrate analog enhances the enzymatic activity of human blood group B galactosyltransferase. *Glycobiology*, 20(6), 718-723. doi:10.1093/glycob/cwq019

Smith, J., Yang, Y., Levy, S., Adelusi, O. O., Hahn, M. G., O'Neill, M. A., & Bar-Peled, M. (2016). Functional Characterization of UDP-apiose Synthases from Bryophytes and Green Algae Provides Insight into the Appearance of Apiose-containing Glycans during Plant Evolution. *J. Biol Chem, 291*(41), 21434-21447.

Smith, J. A., Bar-Peled, M. (2017). Synthesis of UDP-apiose in Bacteria: The marine phototroph *Geminicoccus roseus* and the plant pathogen *Xanthomonas pisi*. *PLoS One, 12*(9), e0184953.

Song, M. C., Nigussie, F., Yang, H. J., Kim, H. H., Kim, J. Y., Chung, D. K., & Baek, N. I. (2008). Phenolic glycosides from Lindera fruticosa root and their inhibitory activity on osteoclast differentiation. *Chem Pharm Bull (Tokyo), 56*(5), 707-710.

Spellman, M. W., McNeil, M., Darvill, A.G., Albersheim, P. (1983). Isolation and characterization of 3-C-carboxyl-5-deoxy-xylose, a naturally occurring, branched chain acidic monosaccharide. *Carbohydr Res, 122*(115-129).

Steenbergen, S. M., & Vimr, E. R. (2003). Functional relationships of the sialyltransferases involved in expression of the polysialic acid capsules of Escherichia coli K1 and K92 and Neisseria meningitidis groups B or C. *J Biol Chem, 278*(17), 15349-15359. doi:10.1074/jbc.M208837200

Stenutz, R., Weintraub, A., & Widmalm, G. (2006). The structures of Escherichia coli O-polysaccharide antigens. *FEMS Microbiol Rev, 30*(3), 382-403. doi:10.1111/j.1574-6976.2006.00016.x

Sterling, J. D., Atmodjo, M. A., Inwood, S. E., Kumar Kolli, V. S., Quigley, H. F., Hahn, M. G., & Mohnen, D. (2006). Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proc Natl Acad Sci U S A, 103*(13), 5236-5241. doi:10.1073/pnas.0600120103

Sterling, J. D., Quigley, H. F., 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*(1), 360-371.

Stevenson, G., Andrianopoulos, K., Hobbs, M., & Reeves, P. R. (1996). Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol, 178*(16), 4885-4893.

Szymanski, C. M., & Wren, B. W. (2005). Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol, 3*(3), 225-237. doi:10.1038/nrmicro1100

Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., & Guerry, P. (1999). Evidence for a system of general protein glycosylation in Campylobacter jejuni. *Mol Microbiol, 32*(5), 1022-1030.

Tabish, S., Raza, A., Nasir, A., Zafar, S., & Bokhari, H. (2011). Analysis of glycosylation motifs and glycosyltransferases in Bacteria and Archaea. *Bioinformation, 6*(5), 191-195.

Takeuchi, H., Fernandez-Valdivia, R. C., Caswell, D. S., Nita-Lazar, A., Rana, N. A., Garner, T. P., . . . Haltiwanger, R. S. (2011). Rumi functions as both a protein O-glucosyltransferase and a protein O-xylosyltransferase. *Proc Natl Acad Sci U S A, 108*(40), 16600-16605. doi:10.1073/pnas.1109696108

Tamaki, S., Sato, T., & Matsuhashi, M. (1971). Role of lipopolysaccharides in antibiotic resistance and bacteriophage adsorption of Escherichia coli K-12. *J Bacteriol, 105*(3), 968-975.

Teng, Y., Liu, Q., Ma, J., Liu, F., Han, Z., Wang, Y., & Wang, W. (2006). Cloning, expression and characterization of a novel human CAP10-like gene hCLP46 from CD34(+) stem/progenitor cells. *Gene, 371*(1), 7-15. doi:10.1016/j.gene.2005.08.027

Thibault, J. F., Renard, C. M. G. C., Axelos, M. A. V., Roger, P., & Crépeau, M. J. (1993). Studies of the length of homogalacturonic regions in pectins by acid hydrolysis. *Carbohydr. Res., 238*271-286.

Thibault, P., Logan, S. M., Kelly, J. F., Brisson, J. R., Ewing, C. P., Trust, T. J., & Guerry, P. (2001). Identification of the carbohydrate moieties and glycosylation motifs in Campylobacter jejuni flagellin. *J Biol Chem, 276*(37), 34862-34870. doi:10.1074/jbc.M104529200

Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., & Holden, H. M. (2000). Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase. *Biochemistry, 39*(19), 5691-5701.

Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., & Holden, H. M. (2001). Human UDP-galactose 4-epimerase. Accommodation of UDP-N-acetylglucosamine within the active site. *J Biol Chem, 276*(18), 15131-15136. doi:10.1074/jbc.M100220200

Treutter, D. (2005). Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biol. (Stuttg), 7*(6), 581-591. doi:10.1055/s-2005-873009

Tsai, SF, , L., & SS. (2011). Flavonoid composition in the leaves of twelve *Litsea* and *Neolitsea* plants. *J Chinese Chem Soc, 58*(376–383).

Unligil, U. M., & Rini, J. M. (2000). Glycosyltransferase structure and mechanism. *Curr Opin Struct Biol, 10*(5), 510-517.

Urbanowicz, B. R., Pena, M. J., Moniz, H. A., Moremen, K. W., & York, W. S. (2014). Two Arabidopsis proteins synthesize acetylated xylan in vitro. *Plant J, 80*(2), 197-206. doi:10.1111/tpj.12643

Usadel, B., Schluter, U., Molhoj, M., Gipmans, M., Verma, R., Kossmann, J., . . . Pauly, M. (2004). Identification and characterization of a UDP-D-glucuronate 4-epimerase in Arabidopsis. *FEBS Lett, 569*(1-3), 327-331. doi:10.1016/j.febslet.2004.06.005

Vaara, M. (1993). Antibiotic-supersusceptible mutants of Escherichia coli and Salmonella typhimurium. *Antimicrob Agents Chemother, 37*(11), 2255-2260.

van Heijenoort, J. (2001). Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology, 11*(3), 25R-36R.

Vanzin, G. F., Madson, M., Carpita, N. C., Raikhel, N. V., Keegstra, K., & Reiter, W. D. (2002). The mur2 mutant of Arabidopsis thaliana lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proc Natl Acad Sci U S A, 99*(5), 3340-3345. doi:10.1073/pnas.052450699

Vauterin, L., Hoste, B., Kersters, K., & Swings, J. (1995). Reclassification of *Xanthomonas*. *Int. J. Syst. Bacteriol., 45*, 472-489.

Vidal, S, , D., T, , W., P, . . . P. (2000). Structural characterization of the pectic polysaccharide rhamnogalacturonan II: Evidence for the location of the aceric acid-containing oligoglycosyl side chain*Carbohydr Res26*, 277-294.

Vidali, L., Augustine, R. C., Kleinman, K. P., & Bezanilla, M. (2007). Profilin is essential for tip growth in the moss *Physcomitrella patens*. *Plant Cell, 19*(11), 3705-3722. doi:10.1105/tpc.107.053413

Villemez, C. L., Lin, T. Y., & Hassid, W. Z. (1965). Biosynthesis of the polygalacturonic acid chain of pectin by a particulate enzyme preparation from Phaseolus aureus seedlings. *Proc Natl Acad Sci U S A, 54*(6), 1626-1632.

Vojnov, A. A., Zorreguieta, A., Dow, J. M., Daniels, M. J., & Dankert, M. A. (1998). Evidence for a role for the gumB and gumC gene products in the formation of xanthan from its pentasaccharide repeating unit by Xanthomonas campestris. *Microbiology, 144 (Pt 6)*, 1487-1493. doi:10.1099/00221287-144-6-1487

Vongerichten, E. (1901). Ueber das Apiin und Apiose. *Liebigs Ann Chem, 318*, 121-136.

Vorholter, F. J., Niehaus, K., & Puhler, A. (2001). Lipopolysaccharide biosynthesis in Xanthomonas campestris pv. campestris: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. *Mol Genet Genomics, 266*(1), 79-95.

Voxeur, A., Andre, A., Breton, C., & Lerouge, P. (2012). Identification of putative rhamnogalacturonan-II specific glycosyltransferases in Arabidopsis using a combination of bioinformatics approaches. *PLoS One, 7*(12), e51129. doi:10.1371/journal.pone.0051129

Vrielink, A., Ruger, W., Driessen, H. P., & Freemont, P. S. (1994). Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J, 13*(15), 3413-3422.

Wakuta, S., Mineta, K., Amano, T., Toyoda, A., Fujiwara, T., Naito, S., & Takano, J. (2015). Evolutionary divergence of plant borate exporters and critical amino acid residues for

the polar localization and boron-dependent vacuolar sorting of AtBOR1. *Plant Cell Physiol., 56*(5), 852-862. doi:10.1093/pcp/pcv011

Walsh, R. M., Jr., Polizzi, S. J., Kadirvelraj, R., Howard, W. W., & Wood, Z. A. (2015). Man o' war mutation in UDP-alpha-D-xylose synthase favors the abortive catalytic cycle and uncovers a latent potential for hexamer formation. *Biochemistry, 54*(3), 807-819. doi:10.1021/bi501357c

Wang, W., Haberer, G., Gundlach, H., Glasser, C., Nussbaumer, T., Luo, M. C., . . . Messing, J. (2014). The *Spirodela polyrhiza* genome reveals insights into its neotenous reduction fast growth and aquatic lifestyle. *Nat. Commun., 5*, 3311. doi:10.1038/ncomms4311

Wang, Y., Zhang, Y. J., Gao, W. Y., & Yan, L. L. (2007). [Anti-tumor constituents from Paris polyphylla var. yunnanensis]. *Zhongguo Zhong Yao Za Zhi, 32*(14), 1425-1428.

Watson, R. R., & Orenstein, N. S. (1975). Chemistry and biochemistry of apiose. *Adv. Carbohydr. Chem. Biochem., 31*, 135-184.

Wellmann, E., & Grisebach, H. (1971). Purification and properties of an enzyme preparation from Lemna minor L. catalyzing the synthesis of UDP-apiose and UDP-d-xylose from UDP-d-glucuronic acid. *Biochim. Biophys. Acta*.

Whitfield, C. (2006). Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annu Rev Biochem, 75*, 39-68. doi:10.1146/annurev.biochem.75.103004.142545

Whitfield, C., & Trent, M. S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem, 83*, 99-128. doi:10.1146/annurev-biochem-060713-035600

Wiggins, C. A., & Munro, S. (1998). Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc Natl Acad Sci U S A, 95*(14), 7945-7950.

Willats, W. G., McCartney, L., Mackie, W., & Knox, J. P. (2001). Pectin: cell biology and prospects for functional analysis. *Plant Mol Biol, 47*(1-2), 9-27.

Willis, L. M., Gilbert, M., Karwaski, M. F., Blanchard, M. C., & Wakarchuk, W. W. (2008). Characterization of the alpha-2,8-polysialyltransferase from Neisseria meningitidis with synthetic acceptors, and the development of a self-priming polysialyltransferase fusion enzyme. *Glycobiology, 18*(2), 177-186. doi:10.1093/glycob/cwm126

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One, 2*(8), e718. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17684564 doi:10.1371/journal.pone.0000718

Wu, B., Takahashi, T., Kashiwagi, T., Tebayashi, S., & Kim, C. S. (2007). New flavonoid glycosides from the leaves of Solidago altissima. *Chem Pharm Bull (Tokyo), 55*(5), 815-816.

Wyckoff, T. J., Raetz, C. R., & Jackman, J. E. (1998). Antibacterial and anti-inflammatory agents that target endotoxin. *Trends Microbiol, 6*(4), 154-159.

Yang, T., Bar-Peled, L., Gebhart, L., Lee, S. G., & Bar-Peled, M. (2009). Identification of galacturonic acid-1-phosphate kinase, a new member of the GHMP kinase superfamily in plants, and comparison with galactose-1-phosphate kinase. *J. Biol. Chem., 284*(32), 21526-21535. doi:10.1074/jbc.M109.014761

Yang, T., Bar-Peled, Y., Smith, J. A., Glushka, J., & Bar-Peled, M. (2012). *In-microbe* formation of nucleotide-sugars in engineered *Escherichia coli*. *Anal. Biochem., 421*(2), 691-698. doi:10.1016/j.ab.2011.12.028

Yapo, B. M., Lerouge, P., Thibault, J. F., & Ralet, M. C. (2007). Pectin samples from citrus peel cell walls contain homogalacturonans homogenous with respect to molar mass, rhamnogalacturonan I and rhamnogalacturonan II. *Carbohydrate Polymers, 69*, 426–435.

Yin, Y., Huang, J., Gu, X., Bar-Peled, M., & Xu, Y. (2011). Evolution of plant nucleotide-sugar interconversion enzymes. *PLoS One, 6*(11), e27995. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22125650 doi:10.1371/journal.pone.0027995

York, W., Darvill, A., McNeil, M., Stevenson, T., & Albersheim, P. (1986). Isolation and characterization of plant cell walls and cell wall components. *Methods Enzymol., 118*, 3-40.

Young, K. D. (2006). The selective value of bacterial shape. *Microbiol Mol Biol Rev, 70*(3), 660-703. doi:10.1128/MMBR.00001-06

Yuan, Y., Barrett, D., Zhang, Y., Kahne, D., Sliz, P., & Walker, S. (2007). Crystal structure of a peptidoglycan glycosyltransferase suggests a model for processive glycan chain synthesis. *Proc Natl Acad Sci U S A, 104*(13), 5348-5353. doi:10.1073/pnas.0701160104

Zandleven, J., Sorensen, S. O., Harholt, J., Beldman, G., Schols, H. A., Scheller, H. V., & Voragen, A. J. (2007). Xylogalacturonan exists in cell walls from various tissues of Arabidopsis thaliana. *Phytochemistry, 68*(8), 1219-1226. doi:10.1016/j.phytochem.2007.01.016

Zeng, W., Jiang, N., Nadella, R., Killen, T. L., Nadella, V., & Faik, A. (2010). A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. *Plant Physiol, 154*(1), 78-97. doi:10.1104/pp.110.159749

Zheng, S., Gao, L., Kang, S., Shen, X., & Wang, X. (1998). Studies on the two new stereo-saponins from *Morchella conica*. *Indian J Chem Sect B, 37*, 825–827.

Zhong, R., Teng, Q., Haghighat, M., Yuan, Y., Furey, S. T., Dasher, R. L., & Ye, Z. H. (2017). Cytosol-Localized UDP-Xylose Synthases Provide the Major Source of UDP-Xylose for the Biosynthesis of Xylan and Xyloglucan. *Plant Cell Physiol, 58*(1), 156-174. doi:10.1093/pcp/pcw179

APPENDIX A

ACRONYMS AND ABBREVIATIONS

a.a.	amino acid
AIR	alcohol insoluble residue
Api*f*	apiofuranose
ApiGalA	apiogalacturonan
ApiT	apiosyltransferase
Ara*f*	arabinofuranose
Ara*p*	arabinopyranose
AX	arabinoxylan
bUAS	bacterial UAS
CAZy	Carbohydrate Active enZYmes
CPS	capsular polysaccharide
Dha	2-keto-3-deoxy-D-lyxo-heptulosaric
EPG	endopolygalacturonase
Fuc*p*	fucopyranose
Gal*p*	galactopyranose
Gal*p*A	galacturonic acid
GAUT	galacturonosyltransferase
GAX	glucuronoarabinoxylan

GC	gas chromatography
Glc	glucopyranose
Glc*p*A	glucuronic acid
GlcAT	glucuronosyltransferase
GT	glycosyltransferase
HEK	human embryonic kidney
HGA	homogalacturonan
IM	inner membrane
Kdo	2-keto-3-deoxy-D-manno-octulosonic acid
LC-MS/MS	liquid chromatography tandem mass spectrometry
LOS	lipoooligosaccharide
LPS	lipopolysaccharide
Man*p*	mannopyranose
MWCO	molecular weight cut off
NMR	nuclear magnetic resonance
OM	outer membrane
RG-I	rhamnogalacturonan-I
RG-II	rhamnogalacturonan-II
RGP	reversibly glycosylated peptide
RGXT	RG-II:α-(1,3)-xylosyltransferase
Rha*p*	rhamnopyranose
RT	real-time
SDR	short-chain hydrogenase/reductase

SE	standard error
T-DNA	transfer DNA
UAM	UDP-arabinose mutase
UAS	UDP-apiose/UDP-xylose synthase
UXS	UDP-xylose synthase
UGDH	UDP-glucose dehydrogenase
UV	ultraviolet
XGA	xylogalacturonan
XpGTA	*X. pisi* GT A
XpGTB	*X. pisi* GT B
Xyl*p*	xylopyranose
XylT	xylosyltransferase

APPENDIX B

DOWNREGULATION OF UDP-ARABINOMUTASE GENE IN SWITCHGRASS (PANICUM VIRGATUM L.) RESULTS IN INCREASED CELL WALL LIGNIN WHILE REDUCING ARABINOSE-GLYCANS[3]

Jonathan D. Willis†, [3]James A. Smith†, Mitra Mazarei, Ji-Yi Zhang, Geoffrey B. Turner, Stephen R. Decker, Robert W. Sykes, Charleson R. Poovaiah, Holly L. Baxter, Dave G. J. Mann, Mark F. Davis, Michael K. Udvardi, Maria J. Peña, Jason Backe, Maor Bar-Peled, C. Neal Stewart, Jr. 2016. Frontiers in Plant Science. PMCID: PMC5081414

†Co-First Authors (equal contributors)

Reprinted here with permission of publisher

ABSTRACT

Switchgrass (*Panicum virgatum* L.) is a C_4 perennial prairie grass and a dedicated feedstock for lignocellulosic biofuels. Saccharification and biofuel yields are inhibited by the plant cell wall's natural recalcitrance against enzymatic degradation. Plant hemicellulose polysaccharides such as arabinoxylans structurally support and cross-link other cell wall polymers. Grasses predominately have Type II cell walls that are abundant in arabinoxylan, which comprise nearly 25% of aboveground biomass. A primary component of arabinoxylan synthesis is uridine diphosphate (UDP) linked to arabinofuranose (Ara*f*). A family of UDP-arabinopyranose mutase/reversible glycosylated polypeptides (UAM/RGPs) catalyze the interconversion between UDP-arabinopyranose (UDP-Ara*p*) and UDP-Ara*f*. The expression of a switchgrass arabinoxylan biosynthesis pathway gene, *PvUAM1*, was decreased via RNAi to investigate its role in cell wall recalcitrance in the feedstock. *PvUAM1* encodes a switchgrass homolog of UDP-arabinose mutase, which converts UDP-arabinopyranose to UDP-arabinofuranose. Southern blot analysis revealed each transgenic line contained between one to at least seven T-DNA insertions, resulting in some cases, a 95% reduction of native *PvUAM1* transcript in stem internodes. Transgenic plants had increased pigmentation in vascular tissues at nodes, but were otherwise similar in morphology to the non-transgenic control. Cell wall-associated arabinose was decreased in leaves and stems by over 50%, but there was an increase in cellulose. In addition, there was a commensurate change in arabinose side chain extension. Cell wall lignin composition was altered with a concurrent increase in lignin content and transcript abundance of lignin biosynthetic genes in mature tillers. Enzymatic saccharification efficiency was unchanged in the transgenic plants relative to the control. Plants with attenuated *PvUAM1* transcript had increased cellulose and lignin in cell walls. A decrease in cell wall-

associated arabinose was expected, which was likely caused by fewer Ara*f* residues in the arabinoxylan. The decrease in arabinoxylan may cause a compensation response to maintain cell wall integrity by increasing cellulose and lignin biosynthesis. In cases in which increased lignin is desired, e.g., feedstocks for carbon fiber production, down-regulated *UAM1* coupled with altered expression of other arabinoxylan biosynthesis genes might result in even higher production of lignin in biomass.

INTRODUCTION

Switchgrass (*Panicum virgatum*) is a perennial grass species that is considered to be a lignocellulosic bioenergy feedstock with great potential, owing to its wide adaptations to various geographies and temperate climates. Recalcitrance, which is the inherent resistance of cell wall polysaccharides to be digested into fermentable sugars, is a sizeable economic barrier to lignocellulosic biofuel production. At the center of recalcitrance is the heterogeneous composition of plant cell walls, which are made of three main types of polymers: cellulose, lignin, and hemicellulose (Dixon, 2013). Feedstock genomics and biotechnology have enabled a better understanding of cell wall recalcitrance, including that for switchgrass (Casler et al., 2011; Chen et al., 2016). Relatively few studies exist in which hemicellulose has been manipulated were carried out to determine its role in cell wall recalcitrance in biofuel crops (Vega-Sanchez and Ronald, 2010).

In plants, hemicelluloses are comprised of non-cellulose cell wall polysaccharides, and share a sugar backbone composed of 1,4-linked β-D-glycoses and include xyloglucan mixed-linkage glucan, xylan, and glucomannan. Xylan itself constitutes a sub-grouping of polysaccharides whose members are distinguished from one another by the types of

oligosaccharide side chains linked to the 1,4-linked β-D-xylopyranose (Xyl*p*) backbone (Rennie and Scheller, 2014; York and O'Neill, 2008; Scheller and Ulvskov, 2010). The backbone of xylans isolated from grasses for example, is decorated by a large number of L-arabinose residues (found only in furanose form, Ara*f*), and hence referred to as arabinoxylans. The Ara*f* residues are attached to the Xyl*p* residues in the backbone predominately at *O*-3 but occasionally at *O*-2. Some of these Ara*f* residues are linked at *O*-2 with an additional α-L-Ara*f* or a β-D-Xyl*p* residue. Grass xylans also contain small amounts of GlcA and Me-GlcA sidechains at *O*-2 (Ebringerova and Heinze, 2000). Xylan may also contain non-carbohydrate modification of *O*-acetyl esters and methyl etherified sugars, as well as feruloyl, and *p*-coumaroyl moieties (Bar-Peled and O'Neill, 2011; Faik, 2010). For example, the aromatic residues (feruloyl and *p*-coumaroyl) can be ester-linked to *O*-5 of terminal or substituted arabinose residues of xylan, whereas the acetate can be attached at *O*-2, *O*-3 or both to xylose in the backbone. In grass species, ferulic acid is ester-linked to the C5 hydroxyl of Ara*f* in arabinoxylan and in ether linkages of lignin monomers (Hartley and Ford, 1989; Scalbert et al., 1985). Although the role of feruloylation is not well understood, an increase in ferulic acid modification of arabinoxylan has been associated with cells that have stopped elongating (Carpita, 1986). Feruloylation has been hypothesized to prime polymerization of lignin thereby interconnecting a network of xylan and lignin (De O. Buanafina, 2009; Iiyama et al., 1994). Additionally, adjacent arabinoxylan chains decorated with ferulic acid can dimerize through oxidative coupling, which may condense wall polymers into a tightly packed matrix enhancing the walls stability and resistance to degradation (Hatfield et al., 1999). Disruption of these ether linkages between arabinoxylan and lignin is an inviting target for improving cell wall degradation.

The diversity in xylan structures is known, but the functional role for such diversity is largely unknown. For example, it is not understood why xylan chemotypes differ among tissues in the same plants. It was proposed that xylan interacts with cellulose and lignin, which serves to strengthen cell walls (Scheller and Ulvskov, 2010). Arabinoxylans comprise over 25% of the mass of grass cell walls (Faik, 2010; Konishi et al., 2011). The formation of arabinoxylan requires the building blocks UDP-xylose and UDP-arabinofuranose (UDP-Ara*f*). UDP-arabinopyranose mutase (UAM) converts UDP-arabinose (pyranose-form, Ara*p*) to UDP-Ara*f* (Konishi et al., 2010; Konishi et al., 2007). UAM orthologs are found in some microalgae and land plants in which they comprise a small gene family (Kotani et al., 2013). Interestingly, UAM can also reversibly glycosylate itself in the presence of UDP-sugars, such as UDP-glucose, UDP-galactose and UDP-xylose (hence the name RGP) (Dhugga et al., 1991; Konishi et al., 2010; Konishi et al., 2007; Rautengarten et al., 2011). The role of UAM as an RGP is not well understood in the context of cell wall and glycan formation. It was hypothesized that the RGP function of UAM may regulate the internal balance of UDP-sugars in the cell or compete for the formation of UDP-Ara*f*, and this hypothesis has been explored in *Arabidopsis*, algae, *Brachypodium*, and rice (Konishi et al., 2011; Kotani et al., 2013; Rancour et al., 2015; Rautengarten et al., 2011). However, UAM's potential role in recalcitrance has never been examined nor manipulated in any bioenergy feedstock.

In this study, it was hypothesized that manipulation of the level of UDP-Ara*f* in cells would alter the amount of arabinoxylan in switchgrass cell walls, and potentially alter feedstock recalcitrance. In this study, a switchgrass *UAM1* homolog (*PvUAM1*) was down-regulated in independent transgenic lines of switchgrass, wherein cell wall composition, saccharification, and plant growth were analyzed.

RESULTS

Identification of PvUAM homologs

Orthologous of functional UDP-arabinomutase (UAM1) amino acid sequences from monocot and eudicot plant species were used to identify the switchgrass PvUAM1 sequence (Fig. 1). The UAM1 has additional function or reversibly glycosylated protein (RGP1), hence forward be named UAM1/RGP1 or UAM1. PvUAM1 has 93% and 86% amino acid sequence similarity to the rice UAM1 and Arabidopsis UAM1/RGP1, respectively. Sequence relationships of UAM proteins from diverse plant species grouped into a central monocot cluster and a split eudicot grouping. PvUAM1 belongs to the monocot group. In addition to UAM1, two other UAM-homologs are known. PvUAM1 is 48% similar to PvUAM2 whereas PvUAM1 has 86% amino acid sequence similarity to PvUAM3.

Molecular and phenotypic characterization of PvUAM-RNAi transgenic plants

To study the role of PvUAM1 in switchgrass in hemicellulose metabolic pathways RNAi-transgenic plants were generated. Seven independent transgenic events regenerated from transformed callus were analyzed (Fig. 2A). Southern blot analysis showed that each transgenic line carried at least one and up to seven T-DNA inserts (Fig. S2B). One transgenic line (270-3) did not survive and was removed from subsequent analysis. The *PvUAM1* transcript abundance was less than that of the control in each of six remaining transgenic lines in both stems and leaves. For example, *PvUAM1* transcript level in stems and leaves of the RNAi plant lines, decreased by 67-95% and to 77-98% relative to the non-transgenic control, respectively (Fig. 2B). Gene expression analysis of *PvUAM* homologs (*PvUAM1, PvUAM2, and PvUAM3*) was

performed on stem internode sections (Fig. 2C). *PvUAM2* expression amongst all transgenic lines was not significantly different than the control. However, and although unintended, the *PvUAM1* RNAi target sequence was similar enough to cause significant downregulation in both *PvUAM1* and *PvUAM3* homologs for lines 270-1 and 270-2. *PvUAM3* transcript was significantly reduced in lines 270-1 and 270-2 by 90% and 74% relative to the control, respectively. An opposite effect was observed for 270-4 in which *PvUAM3* was found to be upregulated 4.6 fold over the control. *PvUAM3* transcript abundance was unchanged in 270-5, 270-6, and 270-7 compared to the non-transgenic control. Interestingly, we found no apparent correlation between number of T-DNA insert and the reduced transcript abundance of UAM1 in these transgenes.

There were several instances of altered plant growth among the transgenic switchgrass lines (Table 1). Transgenic plant lines 270-1, 270-2, 270-5 and 270-7 had equivalent number of tillers compared to the control whereas lines 270-4 and 270-6 had significantly more tillers per plant. Plant lines 270-1, 270-2, and 270-6 were shorter, whereas 270-4 and 270-5 were equivalent to control. Line 270-7 was taller when compared with control line. Tiller stem width was significantly reduced up to 22% in lines 270-1, 270-2, and 270-6, but was increased in 270-7 by up to 6%, whereas the remainder of the lines had unchanged stem width from the control. Fresh weight was significantly increased from the control by up to 102% in lines 270-2, 270-4, 270-5, and 270-7, whereas lines 270-1 and 270-6 were equivalent to control. Dry biomass results were similar to that of fresh weight except line 270-6 was also higher than control. Panicle number was significantly increased in all transgenic lines. In addition to the above mentioned phenotypic differences between PvUAM-RNAi lines and control we interestingly found that line 270-1, 270-2, and 270-4 appeared to have an increased level of red pigment in the stem nodes

when compared with the non-transgenic plants (Fig. 3A). Cross sectioning of fresh stem nodes and internodes at the E3 (elongation) developmental stage on plant line 270-4 showed a dark pigmentation that was deposited in the vascular bundles and outer tissue of the nodes (Fig. 3B).

The walls of PvUAM-RNAi transgenic plants have reduced arabinose

Following phenotypic analyses of PvUAM-RNAi transgenic plants we determined the sugar composition of polysaccharides in the cell walls of these lines. When compared with wall the control, the cell walls of leaves from transgenic PvUAM-RNAi lines had up to 51% decreased arabinose content (Fig. 4A). Similarly, the wall from transgenic stem showed up to 39% decreased arabinose content (Fig. 4B). The highest reduction in arabinose was observed in leaf and stem of line 270-1, which also exhibited the highest *PvUAM1* knockdown.

In addition to a reduced level of arabinose content in the wall, most transgenic plant lines also had an increased level of xylose (up to 16%) in leaves and stems (Fig. 4C and 4D). Line 270-5 was not significantly different in leaf or stem xylose content from the control line. The level of glucose in leaf wall was lower (up to 16% decreased) in most transgenic lines when compared with control (Fig. 4E). Lines 270-1 and 270-4 had up to 18% reduced glucose in stem, while stem glucose of the remaining lines was not significantly altered from control (Fig. 4F). The galactose content in stem cell wall was significantly lower (up to 60% decreased) in transgenic plant lines 270-1, 270-2, and 270-4 while lines 270-5, 270-6, and 270-7 were similar to the control (Table S3). In leaves, on the other hand, the amount of galactose in the wall was slightly lower (up to 17% decreased) only in line 270-1 (Table S4). The rhamnose content in leaf walls was variable among transgenic plants, with a significant increase (up to 83%) in lines 270-2, 270-4 and 270-6 (Table S4) when compared with control. In stems, most transgenic lines

exhibited reduced levels of wall rhamnose (up to 51%) (Table S3). Cell wall mannose levels in leaves and stems were similar among transgenic and control plants (Table S3 and S4). The amount of cellulose in both stem internodes and leaves was determined as well. An increase in cellulose levels was observed in stems of all PvUAM-RNAi transgenic lines (Fig. 5A), and in leaves, transgenic lines 270-1 and 270-6 showed a significantly higher amount cellulose when compared to the control (Fig. 5B). The cellulose level in leaves for lines 270-2, 270-4, 270-5, and 270-7 were equal to the control (Fig. 5B).

Because total sugar analysis is insufficient to identify gross changes in polysaccharide structure and organization, NMR analysis of extracted arabinoxylan was performed.

Arabinoxylan has altered side chains in PvUAM-RNAi mutants

Most of the Ara*f* residues in the cell walls of switchgrass are found in arabinoxylan sidechains. To determine if the synthesis of this polymer was altered in PvUAM-RNAi transgenic plants, we to analyzed arabinoxylan oligosaccharides generated by enzymatic hydrolysis of the arabinoxylan solubilized from the stem internodes and in the leaves by ^{1}H-NMR. This method for example, should distinguish in principle an arabinose in the furanose form from a pyranose form and should provide linkage anomeric configurations (α- or β-form) as well as linkage positions (xylan 1-4 and any branching 1-2, 1-3, etc.). In addition, NMR is an excellent method to identify and quantify resolved sugar signals in mixtures of polymers.

The ^{1}H-NMR analysis of material solubilized with 1 M KOH from the stem walls and hydrolyzed with a xylanase showed clearly that the most abundant component in the sample corresponded to arabinoxylo-oligosaccharides (Fig. S3). The spectra contained intense signals that corresponded to the anomeric proton of α- and β- reducing Xyl. Signals for both α-Xyl and

β-Xyl are detected by NMR, because the reducing end Xyl of the digested xylo-oligomers in solution undergoes opening that assumes both α and β closed-ring configurations. The NMR spectra contained two clearly resolved signals that corresponded to the anomeric proton of two types of arabinose residues. One signal (δ 5.39) was identified as terminal α- L-arabinosyl residues linked at *O*-3 to xylose in the backbone. This signals also corresponded to the terminal Ara in the disaccharide sidechain α-L-Ara*f*-(1→2)-α-l-Ara*f*-(1→3)-, which has the same chemical shift and cannot be differentiated with this analysis. The other signal (δ 5.55) corresponded to Ara*f* substituted at *O*-2 with a single α-Ara*f* or a β-Xyl*p* residue. The relative amount of anomeric signal for a primary branch arabinose decorating the xylan backbone (2-α-Ara) in stems was increased in lines 270-1, 270-2, and 270-4 over control and was found to be equivalent for 270-5, 270-6, and 270-7. In stems the relative anomeric signal for terminal arabinose residues (T-α-Ara) was decreased in lines 270-1 and 270-2 compared to control while remaining lines were fairly equivalent (Table 2). In stems there was a marked increase in α-4-MeGlcA anomeric signal in lines 270-1, 270-2, and 270-4 while the remaining transgenic lines were equivalent to the control. The α-4-MeGlcA signal represents side-chain decoration of xylan. Line 270-1 and 270-2 have less overall arabinose branching compared to non-transgenic primarily due to the decreased in terminal arabinose (T-α-Ara) signal (Fig. S3). Because T-α-Ara signal is representative of H1 of arabinofuranose for both monomeric and extended branches, these data suggest stem arabinoxylan in lines 270-1 and 270-2 has relatively more extended branches (2-α-Ara) than monomer. Additionally, in these lines there appears to be increased MeGlcA decoration of the stem xylan backbone.

As in stems, the leaves the α-Xyl and β-Xyl were unchanged from control (Table 3). Leaf signal for 2-α-Ara was increased in 270-1 and 270-2 compared to control and equivalent for the

other lines. Leaf signal for T-α-Ara was reduced in lines 270-1 and 270-2 compared to control. From leaves the α-4-MeGlcA signal was only markedly increased in line 270-1 (Table 3). These results largely indicate that leaf arabinoxylan structure in lines 270-1 and 270-2 has been altered similar to stems; relatively less overall branching of the xylan backbone with more extended branches than monomers.

Saccharification of PvUAM-RNAi lines is unchanged for total sugars

PvUAM-RNAi plant cell wall sugars were analyzed for polysaccharide enzymatic release from dried R1 tillers. Enzymatic sugar release is one indicator for the level of recalcitrance of the plant cell wall against enzymatic degradation. Enzymatic glucose release was increased up to 13% for 270-4, 270-6 whereas lines 270-1, 270-2, 270-5, and 270-7 were equal to control (Fig. 6A). Enzymatic release of xylose was significantly increased only in line 270-5 by 17% while the other transgenic lines were equivalent to control (Fig. 6B). When data for glucose and xylose released were added, there was no apparent change amongst transgenic plants and the control for total combined sugar release upon saccharification (Fig. 6C).

Lignin biosynthesis: gene expression, lignin content and composition in tillers

Lignin amount was up to 13% higher in transgenic plants compared with the control (Fig. 7A). Interestingly however, the relative ratio of the monolignol components syringyl and guaiacyl (also known as S/G ratio) was increased by 9% and 14%, only in transgenic plants 270-2 and 270-7, respectively (Fig. 7B) compared to control. Since py-MBMS lignin estimates do not include H lignin, thioacidolysis was performed on stems of control and representative transgenic lines 270-1 and 270-6. All samples had similar H lignin measuring between 4-6% of total lignin

(Table S5). Because the samples have similar H lignin and it does not contribute much to the spectra/total lignin content, it does not significantly change our lignin values and falls within the error rate of py-MBMS total lignin analysis.

The finding that transgenic PvUAM-RNAi lead to increase in lignin prompt us to examine selected genes involved in lignin biosynthesis. The relative amount of gene transcript was determined by qRT-PCR. There was increased expression of *PAL*, *F5H*, *4CL*, *C4H*, and *CAD* genes in PvUAM-RNAi transgenes when compared to control (Fig. 8). Expression levels of *COMT*, *C3H*, *CCR*, and *HCT* genes were unchanged compared to control.

DISCUSSION

Arabinoxylans, which comprise a relatively large portion of cell walls in grass species, likely play an important role in recalcitrance in feedstocks such as switchgrass. Arabinoxylans strengthen cell walls through cross linkages with other cell wall polysaccharides and lignin (Faik, 2010; Rennie and Scheller, 2014; Scheller and Ulvskov, 2010; Tan et al., 2013). UDP-Ara*f* has been identified as a common sugar donor for the synthesis of Ara*f*-containing side chains of the xylan backbone, which play an integral part in cross linking to other cell wall components (Anders et al., 2012). Currently the UAM class of plant proteins is the sole candidate known to convert UDP-Ara*p* to UDP-Ara*f* (Konishi et al., 2011). It has been hypothesized that decreasing the pool of available UDP-Ara*f* would, in turn, change how arabinoxylans interact with cellulose microfibrils and lignin: the reduction of numbers of cross linkages would increase the solubility of arabinoxylans (Sternemalm et al., 2008). Hence, we propose that a reduction of *PvUAM1* would reduce arabinose side chains used for cross linkages among cell wall components, ergo, reducing recalcitrance. Our data partially supports the proposition, as the decrease in Ara*f*

residues was accompanied with an increase in the total amount of lignin suggesting a compensation mechanism that has as result unchanged cell wall recalcitrance in the transgenic lines (270-1 and 270-2).

PvUAM downregulation affects plant growth

Two of the shorter plants (transgenic lines 270-1 and 270-2) had decreased expression of both *PvUAM1* and *PvUAM3*. A similar double knockdown was seen in some transgenic rice due to the close homology of the *OsUAM1* and *OsUAM3* resulting in down-regulation of both homologs (Konishi et al., 2011). Transgenic switchgrass with the double knockdown showed the significant differences in cell wall-associated arabinoxylan side chains in both stems and leaves. Downregulation of only *PvUAM1* did not result in a significant change to the side chains. The transgenic switchgrass exclusively downregulated for *PvUAM1* were taller; rice *OsUAM1*-RNAi plants were shorter (Konishi et al., 2011).

Lines 270-5, 270-6 and 270-7 did not have as strong a knock-down of PvUAM1 as events 270-1, 270-2 and 270-4. The relative lower knock-down may be due to insertional effects, but for events 270-5, 270-6 and 270-7 we also did not observe significant alteration in residual Ara or arabinoxylan structure.

Based on the qRT-PCR results we hypothesize that UAM2 may not function in arabinoxylan biosynthesis, because UAM2 appears to be expressed at or over non-transgenic wild-type in all lines without effectively recovering arabinose content in these lines. The large over expression of UAM3 in 270-4 is also silent in terms of glycosyl makeup. The largest effect seems instead to come from UAM1 expression, evidenced by a correlated decrease in residual Ara with PvUAM1 knockdown among all lines. While lines 270-4, 270-5, 270-6 and 270-7 have

reduced residual Ara in both leaf and stem, it is not as strong as the reduction in Ara in events 270-1 and 2, especially in stem tissue.

Rice UAM1 and UAM3 (*OsUAM1* and *OsUAM3*) are most homologous to *PvUAM1* and *PvUAM3* and are known to function as UDP-Ara mutases, while *OsUAM2*, which is homologous to *PvUAM2*, reportedly does not have mutase activity (Konishi et al., 2007). The reason *PvUAM2* expression fluctuates among events of line 270 is unknown, as is the fluctuation in *PvUAM3* expression (Fig. 2C).

PvUAM downregulation alters cell wall-associated sugars with no change to sugar release

The phenotype we observed of decreased cell wall-associated arabinose is congruent with prior research in rice, *Arabidopsis* and *Brachypodium* (Konishi et al., 2011; Rancour et al., 2015; Rautengarten et al., 2011). In leaves, the *PvUAM1* transcript and cell wall-associated arabinose was significantly decreased in all lines except 270-5, which had the least *PvUAM1* knockdown.

As arabinose is a component of arabinoxylan, NMR was employed to deeper characterize arabinoxylan side chain structure. While a large amount of structural information can be deduced from the NMR spectra, there are regions of overlap in the anomeric signals of the xylo-oligos (Balazs et al., 2013). Specifically, the signals for monomeric α-L-Ara*f* side chains and the terminal α-L-Ara*f* in the disaccharides side chain Ara*f*-(1,2)-α-L-Ara*f*- are unresolved as are the 2-Ara*f* signals for D-Xyl*p*-(1,2)-α-L-Ara*f*-(1,3)- and α-L-Ara*f*-(1,2)-α-L-Ara*f* disaccharide side chains. However, with NMR analysis we can detect the changes in the amount of terminal and substituted Ara*f* sidechains relative to the total amount of residues in the oligosaccharides. Overall, the data (Table 2 and 3) suggest that the switchgrass *PvUAM*-RNAi lines with knockdown of both *PvUAM1* and *PvUAM3* have an altered arabinoxylan structure, and that the

reduction of available UDP-Ara*f* causes either 1) reduced Ara*f* branching, or 2) reduced decoration of substituted side chains with a terminal Ara*f* residue (Rancour et al., 2015; Rancour et al., 2012). Evidence from 270-1 and 270-2 stems demonstrates there is a concomitant increase in glucuronate and 4-Me-glucuronate signals, which may be due to increased substitution with glucuronic acid (GlcA) to make glucuronoarabinoxylan (GAX). Arabinoxylan sidechains on the other lines with only *PvUAM1* knockdown were not significantly different from the non-transgenic control.

The majority of the transgenic lines also showed a slight change in the amount of residual glucose in the walls of both stems and leaves (Fig. 4E and 4F). The glucose observed may come from small changes in xyloglucan or mixed-linkage glucan, which is known to accumulate in developing tissues (such as seed brans) and can be detected in mature stem and leaf tissues (Carpita, 1996). Of note, glucose attributed to starch was particularly increased in all transgenic lines (Fig. S4). In transgenic rice and *Arabidopsis* in which the relevant *UAM1* homolog had decreased expression, there was no significant change to cell wall-associated glucose (Konishi et al., 2011; Rautengarten et al., 2011). Xylose was increased in the cell walls of leaves and stems of all transgenic lines except 270-5 stem (Fig. 4C and 4D). In transgenic *AtUAM1*-RNAi *Arabidopsis* xylose was increased, but in *OsUAM1*-RNAi rice in was unchanged (Konishi et al., 2011; Rautengarten et al., 2011).

Transgenic *Brachypodium* with *BdUAM1* knocked down by RNAi showed an increase of xylose in the cell walls in some lines and a decrease in others (Rancour et al., 2015).

Even though an increase in cellulose (Fig. 5A and 5B) content was detected, no significant change in enzymatic saccharification was found (Fig. 6A and 6C). Line 270-5 had increased enzymatic xylose release, which might be attributed to the increase in available stem

cell wall-associated xylose (Fig. 4D and 6B). Transgenic *Brachypodium* with *BdUAM1* knocked down by RNAi had slightly increased enzymatic glucose release from stems, but significantly lower release from leaves (Rancour et al., 2015). We did not analyze saccharification by organ, only whole tillers.

PvUAM down-regulation increases lignin content and composition

Attenuated UAM switchgrass plants had higher lignin in tillers. Furthermore, the composition of lignin in most *PvUAM1* transgenic switchgrass shifted toward more syringyl (S) lignin units, evidenced by increased S/G ratio and no significant change in H lignin content. This shift may be explained by the concomitant increase in expression of key enzymes in the lignin biosynthesis pathway. When *BdUAM1* was knocked down in *Brachypodium* lignin was increased in the leaves, but was found to be unchanged in sheath/stem portions of tillers (Rancour et al., 2015). Lignin was not analyzed in the rice and *Arabidopsis* studies involving *UAM1* knockdown (Konishi et al., 2011; Rautengarten et al., 2011). Even though there was an increase in cell wall lignin in our study, saccharification was mostly unchanged, which contrasts to the similar study in *Brachypodium*, in which saccharification increased (Fig. 7A) (Rancour et al., 2015). For line 270-5, *PvUAM1* expression was reduced to 56% of native PvUAM1 transcript, lignin content and composition was unchanged, while enzymatic xylose release increased. Further knockdown of *PvUAM1* expression (as seen in other lines) caused an increase in lignin production without reducing enzymatic sugar release. One might speculate that altered ferulated xylan formation affects lignification and in order to maintain proper plant growth and development compensation is made by increasing lignin content which has been reported in *Brachypodium* (Rancour et al., 2015) and now switchgrass. Ferulation is suggested to enable cross-linking of these

polysaccharides to each other as well as to lignin (Molinari et al., 2013). These cross-links are believed to strengthen the cell wall and in part, contribute to the enhanced rigidity of the walls.

Examination of the converse phenotype, where lignin is down-regulated shows evidence of a potential cell wall compensation mechanism as hemicellulose is increased to replace missing lignin. In the maize brown-midrib lignin mutants (*bm3)*, cell wall-associated xylose content was discovered to be equivalent or higher in certain lines while arabinose, rhamnose, and xylose-substitutions decreased (Guillaumie et al., 2008). In transgenic switchgrass using RNAi to downregulate COMT lignin biosynthetic gene, an increase in hemicellulose, xylan, and arabinan were observed (Baxter et al., 2014). In both the *bm3* maize mutants and the RNAi-COMT switchgrass lines, the disruption in lignin biosynthesis gave rise red pigmentation attributed to an increase in cinnamaldehyde (Guillaumie et al., 2008; Fu et al., 2011). The observed darkened internodes in the PvUAM-RNAi lines may also be caused by a lignin biosynthesis metabolite shift (Fig. 3). In-depth characterization of cell wall polysaccharides in cell wall mutants might reveal the interactions amongst cell wall biosynthesis pathways.

We propose a model of the interaction of hemicellulose and lignin in light of our study (Fig. 9). In this model, an increased buildup of UDP-Xyl in UDP-Ara mutase KD-lines, leads to glucose accumulation (Glc-6-phosphate, Glc-1-P, UDP-Glc, sucrose) that is shunted towards phenylpropanoid production via the shikimic acid pathway. Shikimic acid is the precursor for phenylalanine, which is at the top of the lignin biosynthesis pathway (Whetten and Sederoff, 1995). Sucrose is converted to UDP-glucose which is either up taken by cellulose synthase (Ces) complex to form cellulose or is converted to UDP-GlcA by UDP-glucose dehydrogenase (UGD). UDP-GlcA is converted by UDP-D-xylose synthase (UXS) to UDP-D-xylose. Excess UDP-D-xylose in the Golgi stack can inhibit UGD and UXS preventing further buildup of UDP-D-xylose

(Harper and Bar-Peled, 2002). UDP-D-xylose is either converted to 1,4-β-D-xylan by a xylan synthase (XS) or into the arabinose precursor UDP-L-Ara*p* (Rennie and Scheller, 2014). UDP-L-Ara*p* is converted to UPD-L-Ara*f* by UAM and then recruited into arabinose or arabinoxylan ((Konishi et al., 2011; Kotani et al., 2013; Rancour et al., 2015; Rautengarten et al., 2011; Rennie and Scheller, 2014). We propose that the reduction in available UPD-L-Ara*f* caused by *PvUAM*-RNAi results in an increase of UDP-D-xylose with a corresponding reduction in arabinoxylan branching. The possible reduction in Ara*f* side chains, which are normally ferulated by an unknown feruloyl transferase, causes an increase of feruloyl-CoA. Excess ferulic and caffeic acid accumulation is shunted to lignin biosynthesis. This model is supported by the decrease in arabinose and arabinose-furanose side chains (Fig. 6 and Tables 2 and 3) found in the *PvUAM*-RNAi lines. The increase in lignin content and S/G ratio along with upregulation of lignin biosynthetic genes (Fig. 7 and 8) supports the probability of increased synaptic acid levels being generated by increased F5H transcript. Further testing of hypotheses inferred by this model could be accomplished by making combinatorial knockdowns of genes that code enzymes in arabinoxylan and lignin biosynthetic pathways. Pleiotropic perturbations in gene expression and metabolic flux would be informative. Identification and characterization of the suspected feruloyl transferase would aid in discerning the complete molecular mechanism for how the cross linking between arabinoxylan and lignin occurs.

Interference with arabinofuranose metabolism has impacted cross-linking and lignin, with no evident influence on sugar release in switchgrass. The increase in lignin of *PvUAM*-RNAi plants might seem unfavorable for lignocellulosic ethanol production, however, the sugar release efficiency was not affected by the increase in lignin content. While *PvUAM*-RNAi might not be directly suitable for bioenergy feedstocks, they might be employed as a crossing partner with

other switchgrass low in lignin. Particular examples are *COMT* and *MYB4* transgenic lines modified for decreased lignin and increased sugar release efficiency, which might complement the increase in glucose, lignin, and biomass of *PvUAM*-RNAi transgenics in transgene stacks (Baxter et al., 2014, Baxter et al., 2015; Fu et al., 2011; Shen et al., 2012). Additional switchgrass lines that might be useful to cross with UAM1 knockdown transgenics are *MYB4* and *miR156* overexpressors, both of which were dwarfed but yielded relatively high biofuel production (Baxter et al., 2015; Fu et al., 2012; Shen et al., 2012). Additionally, any feedstock that produces inordinately high amounts of lignin might be useful for co-products, such as carbon fiber and bio-plastics (Lindsey et al., 2013; Ragauskas et al., 2014).

We have identified UAM in switchgrass and the downregulated *PvUAM*1 switchgrass plants have altered cell wall sugar content and side chains. Downregulation of *PvUAM1* produced a decrease in arabinose with concurrent increase in lignin content in the cell walls of transgenic switchgrass. We propose a model in which the decrease of available arabinoxylan caused an increase in lignin content due to excess metabolites not being used for arabinoxylan-lignin cross linking. Enzymatic saccharification was not negatively affected by the increase in lignin content possibly due to an increase in cellulose and mol % of xylose in walls of transgenic leaves and stems. Some transgenic *PvUAM1* were larger, which would be useful for commercial biomass and carbon sequestration platforms as well as a lignin feedstock.

EXPERIMENTAL PROCEDURES

PvUAM1 gene isolation, and RNAi construct

The amino acid sequence of switchgrass UAM protein was compared with UAM orthologues from eudicots and monocots: SiUAM1 (XP004982467.1), ZmUAM1

(NP001105598.1), SbUAM1 (XP002464260.1), OsUAM1 (XP006650286.1), BdUAM1 (XP003562308.1), TaUAM1 (CAA77237.1), SlUAM1 (NP001234554.1), VvUAM1 (XP002263490.1), BdUAM1 (XP003569874.1), GmUAM1 (XP003552602.1), BrUAM1 (XP009117866.1), AtUAM3 (AAM65020.1), AtUAM1 (AT3G02230.1), PtUAM1 (Potri.004G117800.1), MtUAM1 (Medtr5g046030.1), OsUAM1 (Q8H8T0.1), OsUAM3 (Q6Z4G3.1), OsUAM2 (Q7FAY6.10), AtUAM2 (NP197069.1), EgUAM1 (AGE46030.1), PdUAM1 (XP008811806.1). The sequences among UAM proteins were compared using alignment in the MUSCLE program (http://www.ebi.ac.uk/Tools/mas/muscle/) and alignment curated by Gblocks using the Phyologeny.fr software program (http://www.phylogeny.lirmm.fr) (Anisimova and Gascuel, 2006; Dereeper et al., 2008). The neighbor joining tree was generated using the MEGA 7.0 program (Tamura et al., 2013). The switchgrass *PvUAM1*, *PvUAM2*, *PvUAM3*, gene sequences were identified by BLASTN analysis of the switchgrass genome (https://phytozome.jgi.doe.gov/pz/portal.html) using the monocot UAM sequences from maize (GI: 542592), foxtail millet (GI: 101771463), and sorghum (GI: 8062976). The nucleotide coding sequence of the *PvUAM1* open reading frame was identified and a 193 bp target sequence was used to generate the RNAi plasmid construct (Fig. S1). The target sequence was amplified by PCR and was cloned into the pCR8 entry vector and then Gateway® sub-cloned into the pANIC-8A plant expression vector (Mann et al., 2012b) to yield the pANIC-8A-PvUAM1 construct (Fig. S2A).

Transgenic plant production and growth analysis

Inflorescences of the ‘Alamo’ switchgrass ‘ST1’ genotype was used to generate Type II embryogenic callus (Burris et al., 2009). *Agrobacterium tumefaciens* strain EHA105 harboring

the pANIC-8A-PvUAM1 vector was used for transformation. Transformed calli were grown on agar containing LP9 growth medium (Burris et al., 2009), supplemented with 400 mg/L Timentin (ticarcillin disodium and clavulanate potassium) and 40 mg/L hygromycin for approximately two months at 25 ° C in the dark. Subsequently the transgenic calli were transferred to regeneration medium as described by Li and Qu (2011) and was supplemented with 250 mg/L cefotaxime to stimulate regeneration (Danilova and Dolgikh, 2004). The T-DNA region of pANIC-8A-PvUAM1 plasmid also contains a cassette that constitutively expresses an orange fluorescence protein (OFP) reporter from the hard coral *Porites* (pporRFP) that is brightly fluorescent in transgenic plants (Mann et al., 2012a). Epi-fluorescence microscope having a 535/30 nm excitation filter and 600/50 nm emissions filter was used to track OFP fluorescence during transgenic callus development and to identify individual putative transgenic lines during growth on agar-plate. Regenerated transgenic plants were rooted and acclimated according to Burris et al. (2009).

T0 transgenic and non-transgenic ST1 control plants were grown in growth chambers under 16 h light/8h dark cycles at 25 °C until moved to a greenhouse with same conditions. Fertilizer (0.02% solution of Peter's soluble 20-20-20) was applied twice per month and plants were irrigated as needed. For growth analysis, each transgenic event and the non-transgenic control was vegetatively divided into three clonal replicates. The wild-type (ST1) control was derived from ST1 inflorescence callus which recovered from tissue culture and which was not exposed to *Agrobacterium tumefaciens*. Each replicate, starting from a single-tiller, was grown in a 1 L pot that was randomly sited in the greenhouse. Plants were grown until the reproductive (R1) developmental stage as defined by Hardin et al. (2013) and tiller and panicle numbers were counted. The five tallest tillers for each replicate were used to estimate total plant height. The

stem width at 10 cm from the potting surface of each of these tillers was measured with a digital caliper. Tillers were harvested and green fresh weight was recorded. Harvested tillers were placed into a drying oven at 42 °C for five days and dry weight was subsequently recorded. Hand sectioning was performed on fresh tillers and nodal sections to assess vascular phenotypes under a dissecting microscope and to depict deposition of pigment.

Southern blot analysis for T-DNA copy number

Approximately 100 mg of young (1 week-old from recently cut-back plants) fresh leaf tissue per plant was used to extract DNA (Freeling and Walbot, 1994). DNA quality was assessed using gel electrophoresis and quantified using a Nanodrop spectrofluorometer (Thermo Fisher, Wilmington, DE, USA). Twenty micrograms of DNA from each sample was digested with *NcoI*, which cuts once within the T-DNA. Digested DNA from transgenic plants and the control, was separated on a TAE-agarose gel, and transferred to a nylon membrane (Amersham Hybond™+ GE Healthcare, Pittsburgh, PA, USA). Blots were pre-hybridized with DIG easy hyb (Roche DIG kit, Nutley, NJ, USA) solution at 42 °C. The blots were then hybridized with the hygromycin DIG-PCR probe, washed, and probe was detected after the membrane was exposed to x-ray film according to manufacturer's instructions (Roche). The DNA probe (972 bp) used to detect the number of hygromycin *(hph*) gene cassette in DNA from transgene lines was amplified by PCR and labeled with digoxigenin (Roche).

RNA extraction, qRT-PCR analysis of UAM's and lignin biosynthetic gene transcript

Quantitative RT-PCR was performed to estimate transcript abundance of *PvUAM* and lignin biosynthetic genes in transgenic *PvUAM1*-RNAi and non-transgenic plants. Total RNA

was isolated from triplicate R1 tiller stem internodes and leaf cuttings using TRI Reagent following manufacturer's instructions (Sigma-Aldrich, St. Louis, MO, USA). Purified RNA was treated with DNase-1 (Promega, Madison, WI, USA) and 3 µg treated RNA was used to generate cDNA using oligo-dT and Superscript III according to manufacturer's instructions (Life Technologies, Carlsbad, CA, USA). qRT-PCR analysis was performed with Power SYBR Green PCR master mix according to manufacturer's protocols (Life Technologies) for optimization of annealing temperature, primer concentration, and cDNA concentration. Primers used for transcript analysis of *PvUAM* are listed in Table S1 and for lignin biosynthetic genes in Table S2. The optimized qRT-PCR protocol utilized a dilution of cDNA 1:100 with thermal cycling at 95 °C for 3 min, and 40 cycle repeats of (95 °C for 10 s and 60.0 °C for 30 s). The relative levels of transcripts were normalized to switchgrass ubiquitin 1 (*PvUbi1*) as a reference gene (Shen et al., 2009) using primer set: PvUBi1_F 5'- CAGCGAGGGCTCAATAATTCCA -3' and PvUbi1_R 5' - TCTGGCGGACTACAATATCCA - 3' (Xu et al., 2011). All experiments were carried out in triplicate technical replicates. The differential Ct method was used to measure transcript abundance after normalization to *PvUbi1* (Schmittgen and Livak, 2008).

Glycosyl residue composition and gas liquid chromatography (GLC) analysis

Tillers at the R1 developmental stage were collected from a single plant grown in a greenhouse for approximately 6-8 weeks. The tillers were cut and divided into stem and leaf sections. A sample section was weighed, ground in liquid nitrogen, and washed as previously reported (Martinez et al., 2012) with slight modifications. Each 1 g sample was suspended in 10 ml 80% EtOH, vortexed for 2 min, then centrifuged (6,000 x g 5 min, 25 °C). The supernatant was removed, and the resulting cell pellet was washed two times each with 10 ml 95% EtOH,

and then with 10 ml 100% EtOH. The cell pellet was resuspended in 10 ml chloroform:MeOH 1:1 (v/v) and mixed by tilting for 1 hr. Each cell pellet residue sample was filtered through Whatmann # 15 filter paper over vacuum and rinsed with acetone. Once dry, the alcohol-insoluble residue (AIR) samples were weighed and passed through a 0.5 mm mesh. AIR sample (10 mg) was suspended in 1 ml buffer (0.1 M sodium acetate, 0.01% Thimerosal, pH 5.0) and treated with an amylase mixture; Spirizyme Excel (1.2 µL) and Liquozyme SC DS (6 µL) (Novozymes, Bagsværd, DK; # NAPFM084 and AUP61163, respectively), as described by Decker (2012). Starch digestion was carried out at 55 °C overnight and subsequently the slurry sample was filtered through double filter layers (50 µm nylon mesh on top of Whatman Grade GF/A filter), and washed with buffer. The polysaccharides in each 1 mg AIR sample were hydrolyzed with 2N trifluoroacetic acid and the free monosaccharides were converted to their alditol acetate derivatives as previously described (York et al., 1986). All samples (including sugar standards) were supplemented with 50 µL 5 mM inositol as an internal standard. Alditol acetate sample or standard (1 µL) was separated on a Restek RTx-2330 fused silica column (0.25 mm I.D. x 30 m, 0.2 µm film thickness) using an Agilent 7890A GLC equipped with a flame ionizing detector (Agilent, Santa Clara, CA, USA). Relative molar percent content was calculated from the areas of sugar peaks identified by standard retention times and normalized to sample mass and internal standard.

Isolation and fractionation of wall polysaccharides and oligosaccharides

AIR sample (250 mg) was de-starched as above in 25 ml buffer treated with 30 µL Spirizyme Excel (Novozymes, Copenhagen, DNK) and 150 µL Liquozyme SC DS at 55 °C overnight, filtered and washed with buffer. To remove loosely bound pectin, the amylase-

insoluble residue was resuspended in 25 ml oxalate solvent (0.5% ammonium oxalate, 0.01% Thimerosal, pH 5.0) and shaken overnight at room temperature. The slurry was filtered through double filter layers, washed with oxalate, and the insoluble residue was reserved. Each oxalate-treated residue was resuspended in 25 ml 1 M base solution (1 M KOH, 1% $NaBH_4$) and shaken overnight at room temperature. The slurry was filtered through double filter layers. The filtrate that contained soluble polysaccharide was reserved and the insoluble residue slurry on top of double layer filter was wash with 1 M KOH and insoluble pellet was reserved. The filtrate and wash 1 M KOH-solubilized wall polymers were combined, supplemented with a drop of octanol as antifoam, neutralized to pH 7 with glacial acetic acid and later dialyzed (3500 MWCO) against deionized water for 2 to 3 days. The insoluble residue after 1 M KOH treatment was resuspended in 25 ml 4 M base solution (4 M KOH, 1% $NaBH_4$), shaken overnight at room temperature and filtered. The 4 M KOH soluble fraction was neutralized and dialyzed as described above. The dialyzed KOH fractions were centrifuged (11,000 x g for 30 min, 25 °C), concentrated by Rotovapor, lyophilized, and used to generate oligosaccharides by enzymatic digestion. The 4 M-KOH insoluble residue (enriched in cellulose) was stored at -20 °C for analysis of cellulose.

Preparation of arabinoxylooligosaccharides and NMR analysis

Between 5–20 mg of 1 M KOH soluble and dialyzed fraction (above) was dissolved in 1-5 ml of 50 mM ammonium formate, pH 5.0. One unit of endoxylanase (from *Trichoderma viride* Megazyme, Wicklow, IRL) was added and the solution incubated at 37 °C for 24 h. Hydrolase activity was terminated by boiling for 10 min in a water bath and the sample was centrifuged at 3,600 x g for 15 min at room temperature. The supernatant was transferred to a tube and

lyophilized. A portion of freeze-dried arabinoxylooligomers (1-2 mg) was dissolved in 0.5 ml deuterium oxide (99.9 %; Cambridge Isotope Laboratories, Tewksbury, MA) and supplemented with 1 µL acetone that was used as an internal chemical shift reference. One-dimensional and two-dimensional ^{1}H NMR spectra were collected on a 600 MHz Varian Inova NMR spectrometer equipped with a 3-mm cold probe and a sample temperature of 25 °C. Data were processed with MestReNova (Version 9.1; Mestrelab Research, Santiago de Compostela, ES). All chemical shifts were measured relative to internal acetone (δ H 2.225).

Cellulose quantification

Thirty milligrams of 4 M KOH insoluble residue were weighed into a conical borosilicate tube with Teflon-lined screw cap, and 3 ml solvent (acetic acid/water/nitric acid, 8/2/1, v/v/v) was added (Updegraff, 1969). The sample was vortexed, heated in a boiling water bath for 30 min with occasional mixing, cooled to room temperature, and centrifuged (2,500 x g for 3 min). The pellet was re-suspended twice in 5 ml water, centrifuged, and the supernatant discarded. The enriched cellulose pellet was treated with 2.5 ml of 72% sulfuric acid (Updegraff, 1969) and incubated at room temperature for 1 h while mixing every 10 min by vortex. Samples were then transferred to a 15 ml Falcon tube and water added to 10 ml. Ten microliters of solution was transferred to a new borosilicate tube and diluted to 400 µL with water. One milliliter of ice-cold anthrone reagent (0.2 g anthrone in 100 ml concentrated sulfuric acid (95–98%)) was added, and the mixture was heated in a boiling water bath for 15 min. Following the anthrone reaction, the amount of sugars in the cellulosic polysaccharide fraction of nitric acid-treated unfractionated cellulose was determine by measuring absorbance at 620 nm with a DU 800 series spectrophotometer (Beckman Coulter, Brea, CA, USA) using glucose from Avicel as standard.

Cell wall sugar release and lignin content and composition

Tillers were collected at the R1 developmental stage from greenhouse-grown plants and air dried for 3 weeks at room temperature before grinding to 1 mm (20 mesh) particle size. Sugar release efficiency was determined via NREL high-throughput sugar release assays on extractive- and starch-free samples using glycosyl hydrolases according to NREL protocol (Decker et al., 2012; Selig et al., 2010). Glucose and xylose release was determined by colorimetric assays with total sugar release being the sum of glucose and xylose released. Lignin analysis was performed on the same samples described above. The lignin content and composition was determined by high-throughput pyrolysis molecular beam mass spectrometry (py-MBMS) on starch-free samples (Sykes et al., 2009) at the National Renewable Energy Laboratory (NREL, Golden, CO, USA). Additionally, *p*-hydroxyphenyl (H) lignin analysis was determined by thioacidolysis according to NREL protocol (Harman-Ware et al., 2016).

Statistical analysis

Statistical analysis was carried out with biological and technical replicates using SAS® (Version 9.3; SAS Institute Inc., Cary, NC, USA) programming of mixed model ANOVA and least significant difference. This statistic analyses were performed on *PvUAM1* and *PvUAM1* homolog transcript abundance by qRT-PCR, growth analysis, cell wall-associated sugar content, NMR sugar side chain analysis, cellulose quantification, enzymatic sugar release, lignin content and composition, and lignin biosynthesis gene quantification by qRT-PCR. The standard error of the mean was calculated and displayed as error bars. *P*-values of ≤ 0.05 were considered to be statistically significant.

TABLES AND FIGURES

Table 1. Growth of down-regulated *PvUAM1* transgenic and non-transgenic (NT-ST1) switchgrass lines. Tiller number refers to the mean tally of all tillers per pot present at time of collection. Tiller height refers to the mean value of the five tallest tillers per replicate pot. Stem width of the five tallest tillers per replicate pot was measured at the potting level and averaged. Panicle number refers to the mean value of panicles present at time of collection per pot. Fresh weight refers to the mean value of fresh total biomass collected per pot. Dry weight refers to the mean value of total biomass collected and then dried for five days at 42 °C per pot. Error bars represent standard error of the mean of three whole plant replicates. Values that share the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Line	Tiller number	LSD	Tiller height (mm)	LSD	Stem width (mm)	LSD	Panicle number	LSD	Fresh weight (g)	LSD	Dry weight (g)	LSD
270-1	32.0 ± 4.9	AB	873.67 ± 17.9	D	3.61 ± 0.12	C	3.7 ± 0.8	ABC	87.70 ± 16.24	BC	22.86 ± 4.37	BC
270-2	29.0 ± 4.9	AB	923.9 ± 17.9	D	3.68 ± 0.12	C	5.3 ± 0.8	AB	165.24 ± 16.24	A	40.21 ± 4.37	A
270-4	42.3 ± 4.9	A	1089.1 ± 17.9	BC	4.12 ± 0.12	AB	5.0 ± 0.8	AB	170.35 ± 16.24	A	39.72 ± 4.37	A
270-5	33.7 ± 4.9	AB	1087.2 ± 17.9	BC	4.37 ± 0.12	A	3.3 ± 0.8	BC	135.81 ± 16.24	AB	35.64 ± 4.37	AB
270-6	40.0 ± 4.9	A	1039.9 ± 17.9	C	3.88 ± 0.12	BC	5.7 ± 0.8	A	128.15 ± 16.24	ABC	34.15 ± 4.37	AB
270-7	36.0 ± 4.9	AB	1185.9 ± 17.9	A	4.36 ± 0.12	A	4.3 ± 0.8	ABC	174.21 ± 16.24	A	41.23 ± 4.4	A
NT-	22.0 ±	B	1118.3 ±	B	3.85 ±	BC	2.7 ± 0.8	C	86.34 ±	C	18.62 ±	C

ST1	4.9		17.9		0.12				16.24		4.37	

Table 2. Glycosyl side chain analysis from stems of down-regulated *PvUAM1* transgenic and non-transgenic (NT-ST1) switchgrass lines. 2-α-Ara refers to the internal chains of arabinose present. T-α-Araf refers to the terminal residues of arabinose present. α-4-GlcA refers to the glucuronic acid present. Values represent relative % of ^{1}H-signal of the total xylan signal. Error bars represent standard error of the mean of three stem internode replicates. Values that share the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Line	2- α –Ara (% signal)	LSD	T-α-Araf (% signal)	LSD	α-4-GlcA (% signal)	LSD
270-1	3.91 ± 0.22	A	9.27 ± 0.10	ABCD	4.27 ± 0.18	A
270-2	3.68 ± 0.20	BC	9.70 ± 0.37	D	4.46 ± 0.65	ABC
270-4	3.11 ± 0.09	CD	11.4 ± 0.12	BCD	3.32 ± 0.33	BCD
270-5	2.60 ± 0.10	CD	12.1 ± 0.16	A	2.24 ± 0.38	CD
270-6	2.35 ± 0.05	CD	11.5 ± 0.16	AB	1.97 ± 0.29	CD
270-7	2.59 ± 0.05	CD	10.6 ± 0.17	AB	2.14 ± 0.28	CD
NT-ST1	2.54 ± 0.12	D	11.2 ± 0.18	CD	2.47 ± 0.28	D

Table 3 Glycosyl side chain analysis from leaves of down-regulated *PvUAM1* transgenic and non-transgenic (NT-ST1) switchgrass lines. 2-α-Ara refers to the internal chains of arabinose present. T-α-Araf refers to the terminal residues of arabinose present. α-4-GlcA refers to the glucuronic acid present. Values represent relative % of ^{1}H-signal of the total xylan signal. Error bars represent standard error of the mean of three leaf replicates. Values that share the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Line	2- α –Ara (% signal)	LSD	T-α-Araf (% signal)	LSD	α-4-GlcA (% signal)	LSD
270-1	3.72 ± 0.22	A	9.27 ± 0.07	AB	2.87± 0.66	A
270-2	4.03± 0.07	A	9.13 ± 0.06	BC	2.04 ± 0.34	AB
270-4	2.84 ± 0.04	D	9.86 ± 0.33	D	1.95 ± 0.46	B
270-5	3.11 ± 0.20	BCD	11.5 ± 0.21	A	1.40 ± 0.33	B
270-6	3.05 ± 0.15	BC	10.7 ± 0.26	AB	1.79 ± 0.28	B
270-7	3.00 ± 0.15	CD	10.9 ± 0.03	AB	1.55 ± 0.36	B
NT-ST1	2.88 ± 0.21	CD	10.9 ± 0.28	AB	2.01 ± 0.24	AB

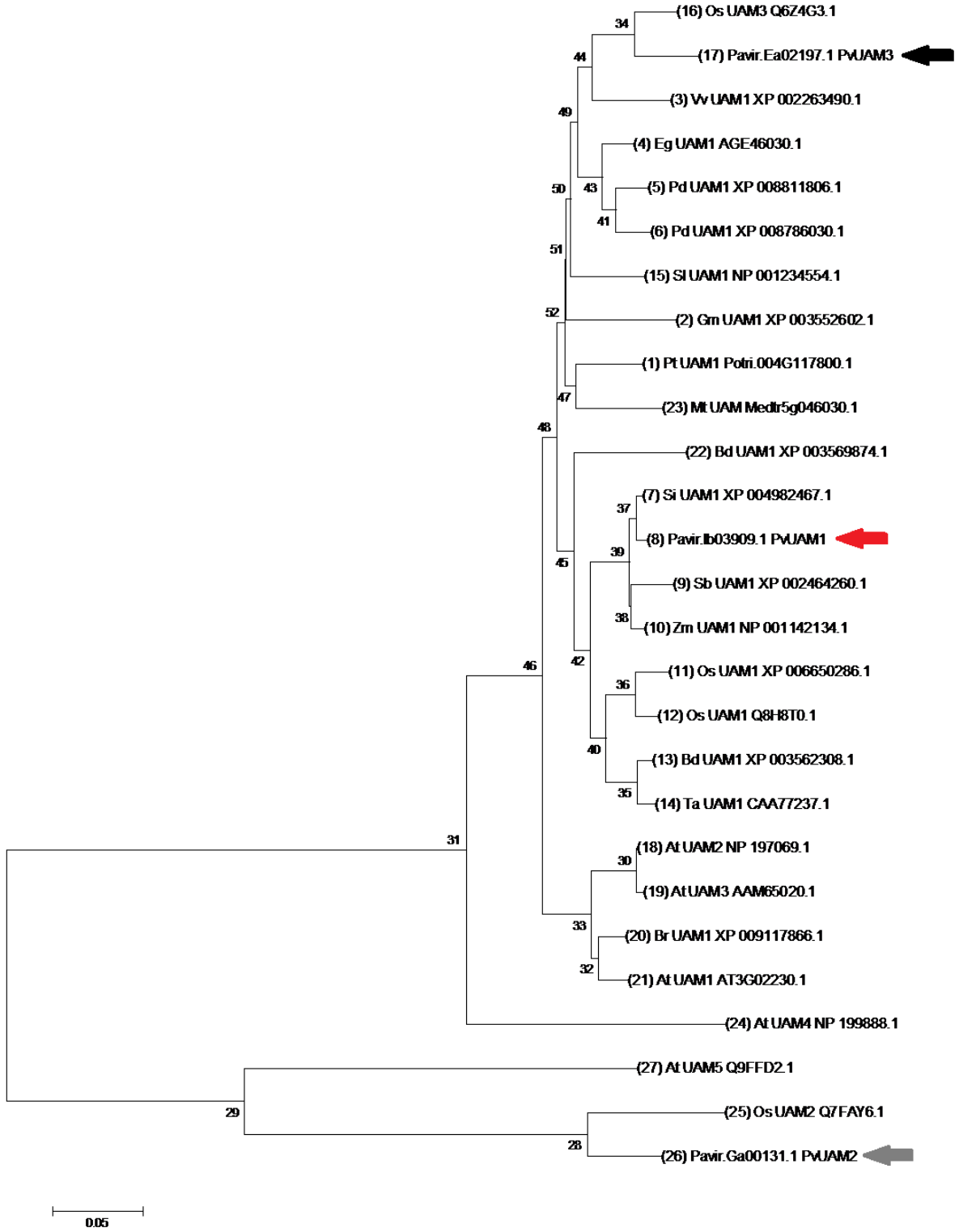


Figure 1 Neighbor-joining cluster analysis of UAM amino acids. The numerals on branches represent number of amino acid substitutions per site for known UAM protein sequences. The

scale bar shows 0.05 amino acid substitutions per site. Switchgrass UAM homologs are indicated by colored arrows, PvUAM1 (Pavirv000Ib03909; red), PvUAM2 (Pavir.GA00131.1; grey), PvUAM3 (Pavir.EA02197.1; black). Tree generated using the MEGA 7.0 program (Tamura et al., 2013) of UAM amino acid sequence alignments using Gblocks at the phylogeny.fr website (http://phylogeny.lirmm.fr). Analysis using 1000 bootstrap replicates was performed.

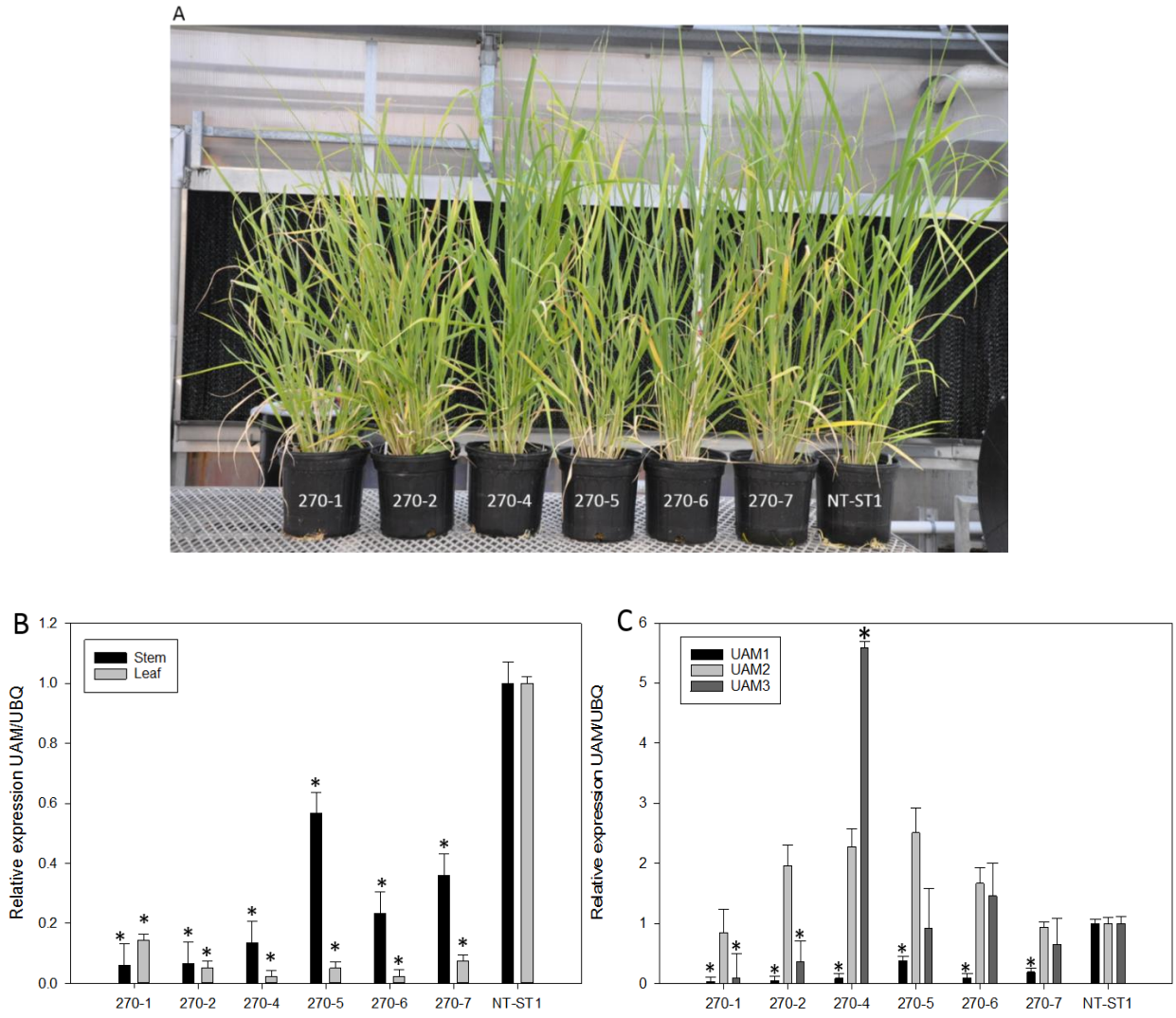


Figure 2 A) Representative down-regulated *PvUAM1* transgenic and non-transgenic (NT-ST1) switchgrass lines. B) Relative expression of *PvUAM1* in leaf and stem tissues of transgenic and non-transgenic lines. C) Relative expression of *PvUAM1*, *PvUAM2*, and *PvUAM3* in stem tissue of transgenic and non-transgenic lines. Relative expression analysis was determined by qRT-PCR and normalized to switchgrass ubiquitin 1 (*PvUbi1*). Bars represent mean values of three replicate stem internode or leaves ± standard error. Asterisks indicate significant differences from non-transgenic control plants at $p \leq 0.05$ as tested by LSD method.

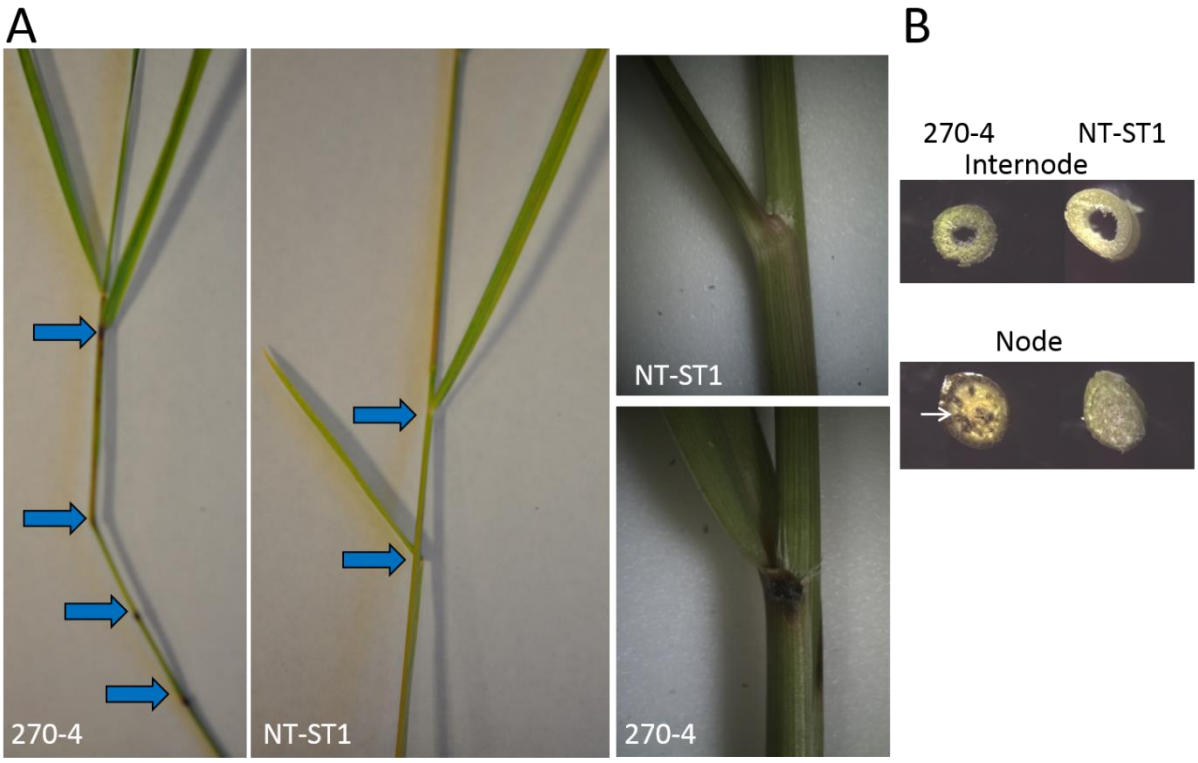


Figure 3 Stem node phenotype from fresh E3 (elongation growth stage) tillers in down-regulated *PvUAM1* transgenic switchgrass. A) Comparison of transgenic PvUAM1 (270-4) and non-transgenic (NT-ST1) nodes and internodes. Arrows indicate nodes. B) Cross-sections of vascular bundles at nodes and internodes of transgenic and non-transgenic plants. Arrow indicates darkened vascular bundle.

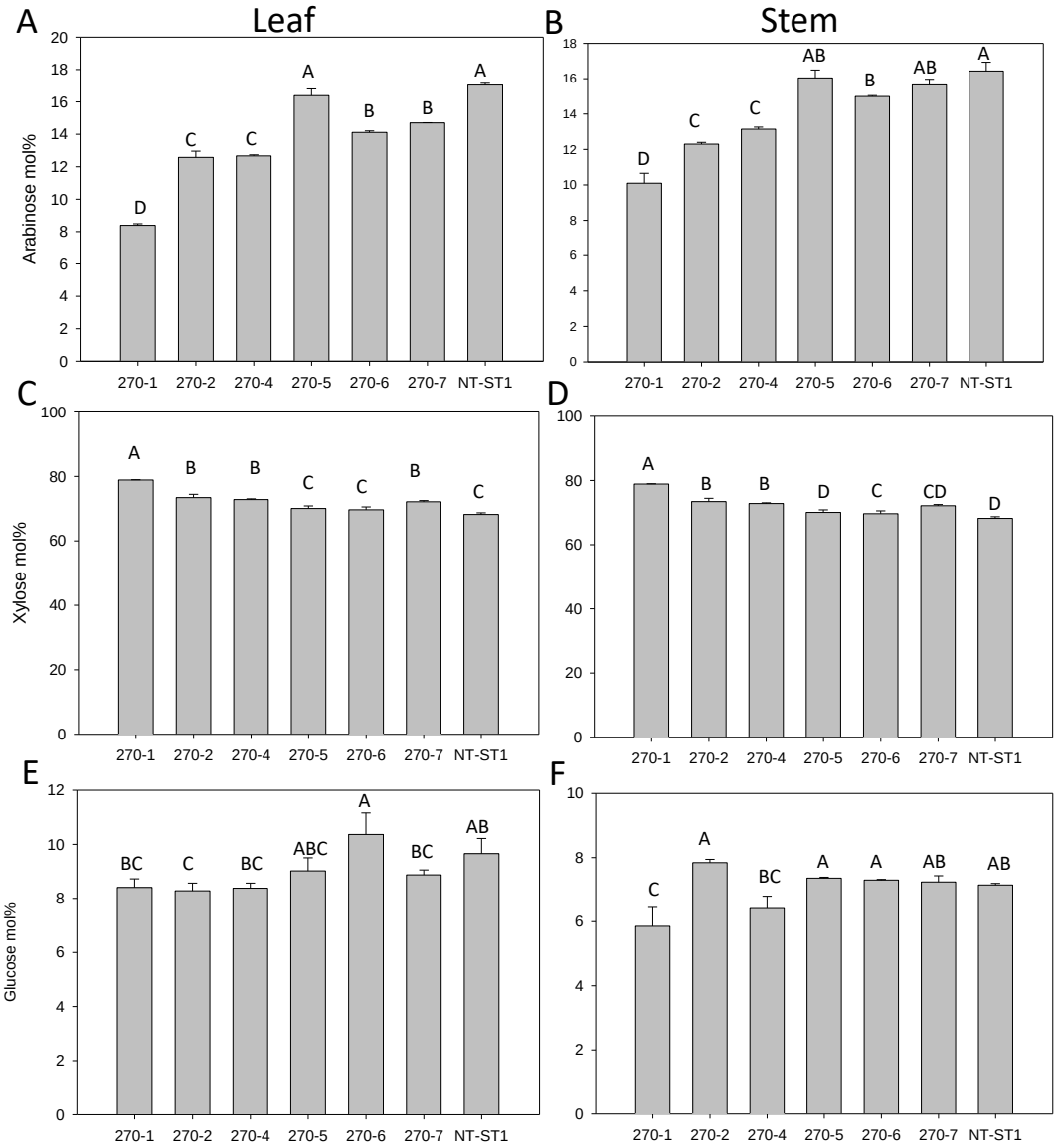


Figure 4 Arabinose (A, B), xylose (C, D), and glucose (E, F) content in leaf (A, C, E) and stem (B, D, F) of transgenic and non-transgenic (NT-ST1) lines as determined by gas chromatography. Samples were normalized to internal standard (inositol) with mol% representing the % of total cell wall-associated sugars measured. Bars represent mean values of three biological replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

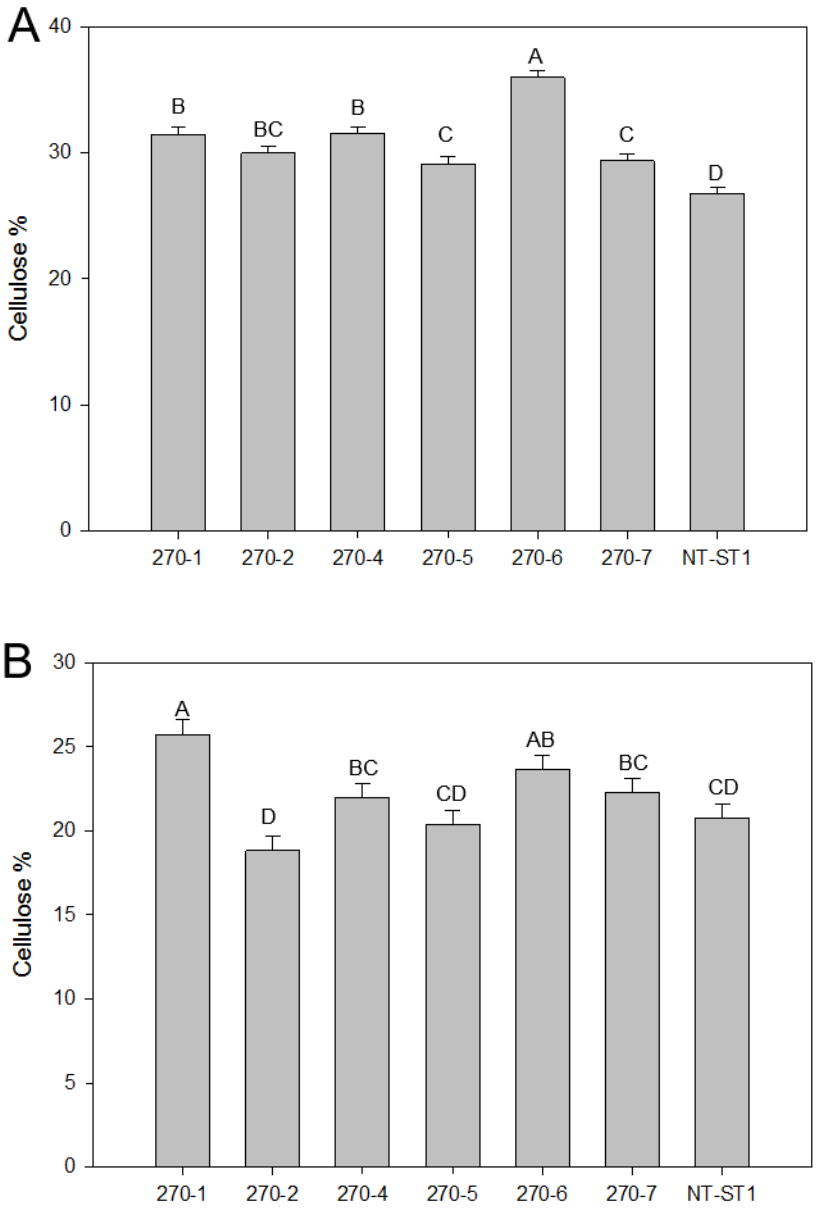


Figure 5 Cellulose content in stems (A) and in leaves (B) of transgenic and non-transgenic (NT-ST1) lines as determined by Updegraff reagent. Bars represent mean values of three biological replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

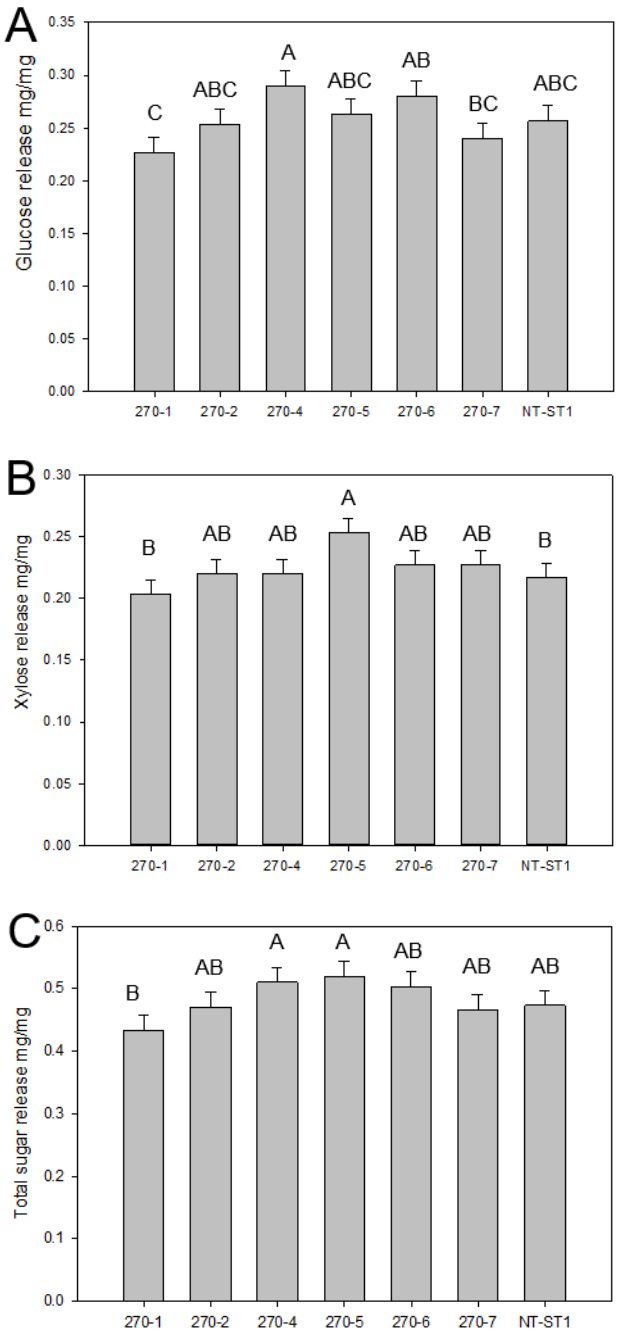


Figure 6 Glucose (A), xylose (B), and total sugar (C) release from transgenic and non-transgenic (NT-ST1) whole tiller cell wall residues as determined by enzymatic hydrolysis. Bars represent mean values of three whole plant replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

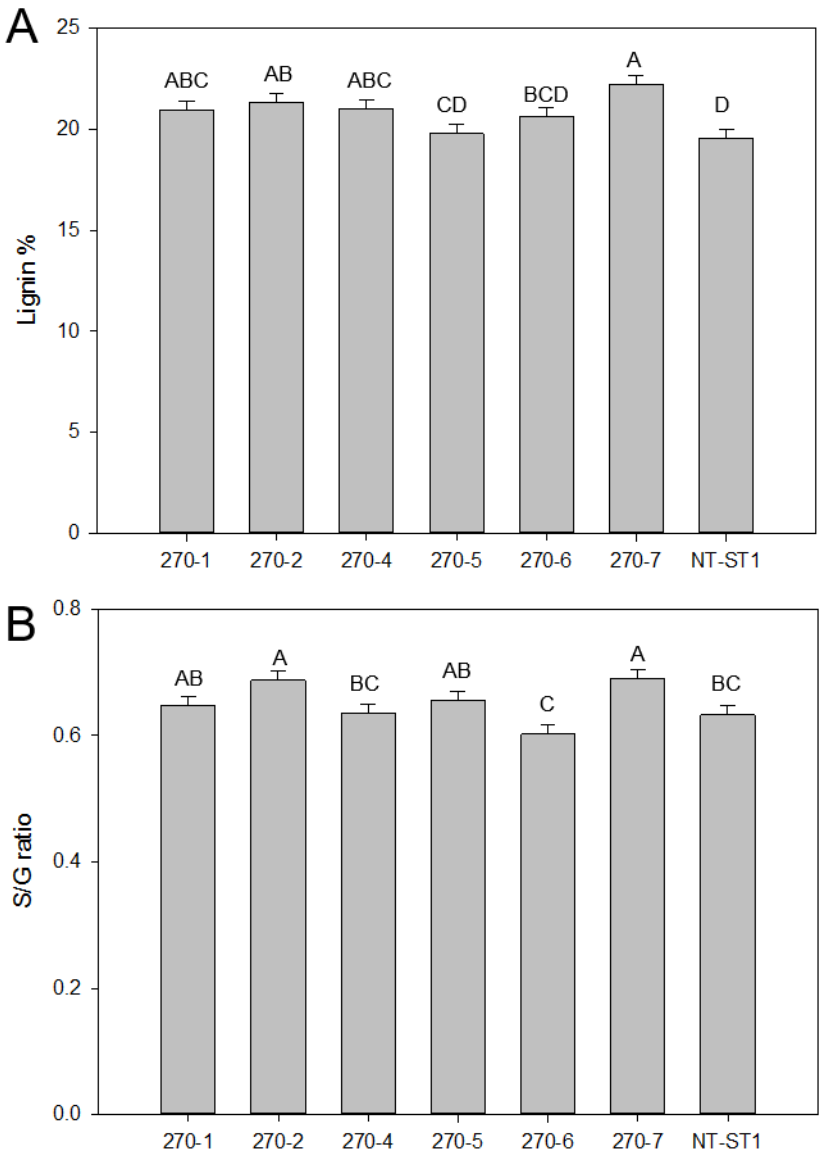


Figure 7 Lignin content (A) and S/G ratio (B) of down-regulated *PvUAM1* transgenic and non-transgenic (NT-ST1) whole tiller cell wall residues as determined by PyMBMS. Bars represent mean values of three whole plant replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

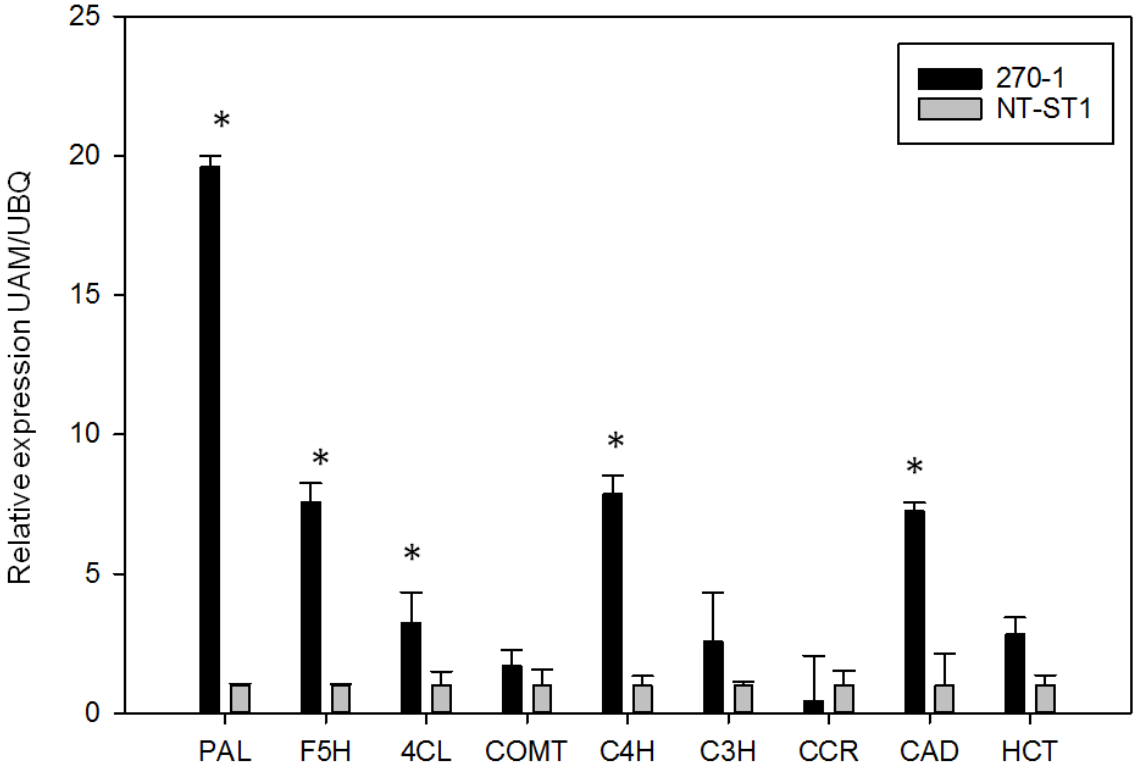


Figure 8 Relative expression of lignin biosynthetic genes in transgenic (270-1) and non-transgenic (NT-ST1) stem internodes as determined by qRT-PCR. The relative expressions were normalized to switchgrass ubiquitin 1 (*PvUbi1*; Relative Expression UBQ). Bars represent mean values of three replicate tiller internode ± standard error. Asterisks indicate significant differences from non-transgenic control plants at $p \leq 0.05$ as tested by LSD. PAL, phenylalanine ammonia-lyase; F5H, ferulate 5-hydroxylase; 4CL, 4-coumarate: CoA ligase; COMT, caffeic acid 3-Omethyltransferase; C4H, coumaroyl shikimate 4-hydroxylase; C3H, coumaroyl shikimate 3-hydroxylase; CCR, cinnamoyl CoA reductase; CAD, cinnamyl alcohol dehydrogenase; HCT, hydroxycinnamoyl.

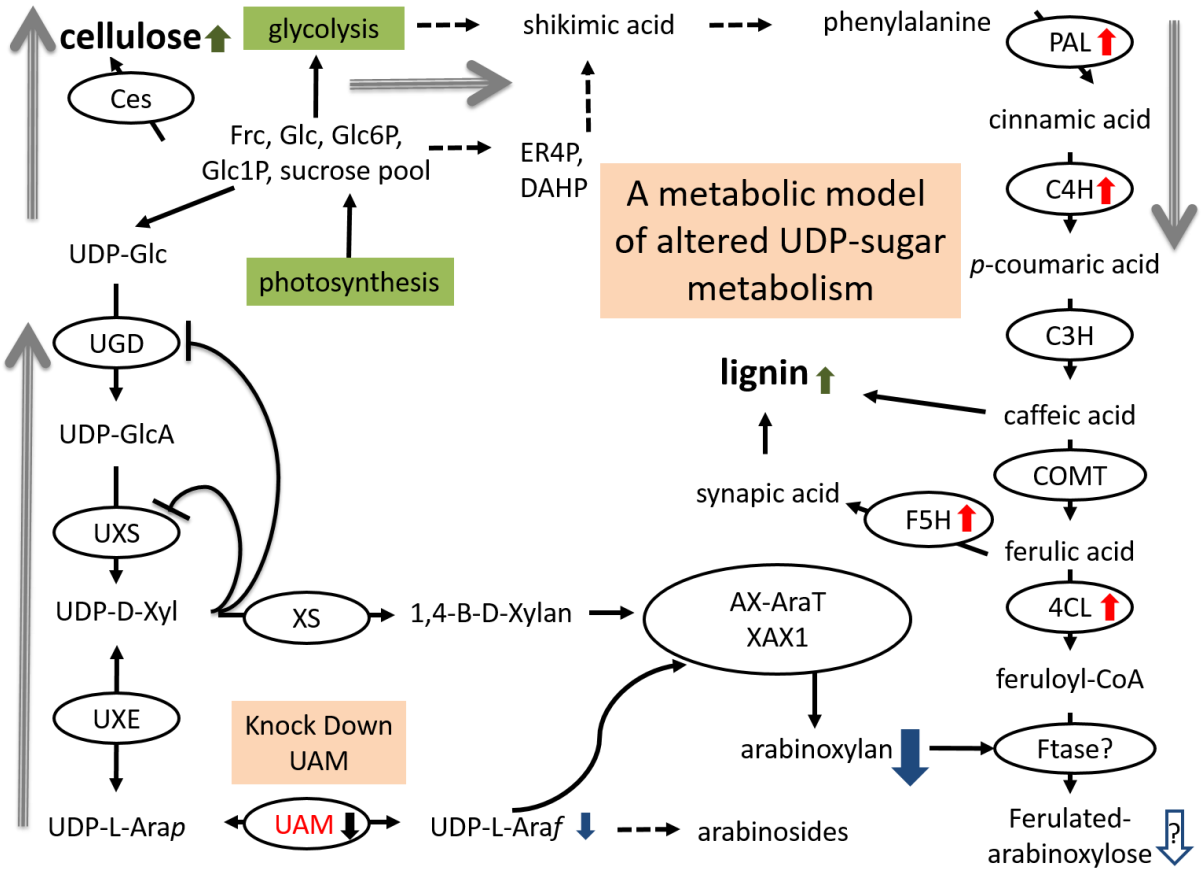


Figure 9 Proposed model arabinoxylan and lignin biosynthesis pathway interactions for down-regulated *PvUAM1* transgenic switchgrass. Proposed model arabinoxylan and lignin biosynthesis pathway interactions for down-regulated *PvUAM1* transgenic switchgrass. Biosynthesis proteins denoted in black ovals: Ces, cellulose synthase, UGD, Uridine diphosphate (UDP)-glucose dehydrogenase; UXS, UDP-D-xylose synthase, XS, xylan synthase; UXE, UDP-xylose esterase, UAM, UDP-arabinomutase (Red lettering); AX-AraT, arabinoxylan arabinosyltransferase (Porchia et al., 2002); XAX1, xylosyl-α-(1,3)-arabinosyl substitution of xylan 1 (Chiniquy et al., 2012); PAL, phenylalanine ammonia-lyase; C4H, coumaroyl shikimate 4-hydroxylase; C3H, coumaroyl shikimate 3-hydroxylase; COMT, caffeic acid 3-O-methyltransferase; F5H, ferulate 5-hydroxylase; 4CL, 4-coumarate: CoA ligase; Ftase?,

undetermined feruloyl transferase. Red arrows indicated up-regulated genes verified by qRT-PCR. Black arrow indicated down regulation of gene verified by qRT-PCR. Green arrows indicated cell wall components which have been increased. Metabolites are indicated: Frc, fructose; Glc, glucose; Glc6P, glucose-6-phosphate, Glc1P, glucose-1-phosphate; ER4P, erythrose-4-phosphate; DAHP, dihydroxyacetone phosphate; UDP-Glc, UDP-glucose; UDP-GlcA, UDP-glucuronic acid; UDP-D-Xyl, UDP-D-xylose; UDP-L-Ara*p*, UDP-L-arabinopyronose; UDP-L-Ara*f*, UDP-L-arabinofuranose. Blue arrows denote suspected reduction in metabolites resulting from down regulation of *PvUAM*. Grey arrows indicate shift in flux.

References

Anders, N., Wilkinson, M.D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T.K., Weimar, T., Mortimer, J.C., Stott, K., Baker, J.M., Defoin-Platel, M., Shewry, P.R., Dupree, P. and Mitchell, R.A. (2012) Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 989-993.

Anisimova, M. and Gascuel, O. (2006) Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst. Biol.* **55**, 539-552.

Balazs, Y.S., Lisitsin, E., Carmiel, O., Shoham, G., Shoham, Y. and Schmidt, A. (2013) Identifying critical unrecognized sugar-protein interactions in GH10 xylanases from *Geobacillus stearothermophilu*s using STD NMR. *FEBS J.* **280**, 4652-4665.

Bar-Peled, M. and O'Neill, M.A. (2011) Plant nucleotide sugar formation, interconversion, and salvage by sugar recycling. *Annu. Rev. Plant Biol.* **62**, 127-155.

Baxter, H.L., Mazarei, M., Labbe, N., Kline, L.M., Cheng, Q., Windham, M.T., Mann, D.G., Fu, C., Ziebell, A., Sykes, R.W., Rodriguez, M., Jr., Davis, M.F., Mielenz, J.R., Dixon, R.A., Wang, Z.Y. and Stewart, C.N., Jr. (2014) Two-year field analysis of reduced recalcitrance transgenic switchgrass. *Plant Biotechnol. J.* **12**, 914-924.

Baxter, H.L., Poovaiah, C.R., Yee, K.L., Mazarei, M., Rodriguez, M., Thompson, O.A., Shen, H., Turner, G.B., Decker, S.R., Sykes, R.W., Chen, F., Davis, M.F., Mielenz, J.R., Davison, B.H., Dixon, R.A. and Stewart, C.N. (2015) Field evaluation of transgenic switchgrass plants overexpressing PvMYB4 for reduced biomass recalcitrance. *Bioenerg Res.* **8**, 910-921.

Burris, J.N., Mann, D.G.J., Joyce, B.L. and Stewart, C.N. (2009) An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). *Bioenerg. Res.* **2**, 267-274.

Carpita, N.C. (1986) Incorporation of proline and aromatic amino acids into cell walls of maize coleoptiles. *Plant Physiol.* **80**, 660-666.

Carpita, N.C. (1996) Structure and Biogenesis of the Cell Walls of Grasses. Plant Physiol & Plant Mol Bio. 47, 445-476.

Casler, M.D., Tobias, C.M., Kaeppler, S.M., Buell, C.R., Wang, Z.Y., Cao, P.J., Schmutz, J., and Ronald, P. (2011). The Switchgrass genome: Tools and strategies. *Plant Genome* 4**,** 273-282.

Chen, X., Ma, Q., Rao, X.L., Tang, Y.H., Wang, Y., Li, G.Y., Zhang, C., Mao, X.Z., Dixon, R., and Xu, Y. (2016). Genome-scale identification of cell-wall-related genes in awitchgrass through comparative genomics and computational analyses of transcriptomic data. *Bioenergy Research* 9**,** 172-180.

Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E. E., Rautengarten, C., Cheng, K., Carroll, A., Ulvskov, P., Harholt, J., Keasling, J. D., Pauly, M., Scheller, H. V., Ronald, P. C. (2012) XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *PNAS* **109**:42, 17117-17122.

Danilova, S.A. and Dolgikh, Y.I. (2004) The stimulatory effect of the antibiotic cefotaxime on plant regeneration in maize tissue culture. *Russ. J. Plant. Physl.* **51**, 559-562.

De O. Buanafina, M. (2009). Feruloylation in grasses: current and future perspectives. *Mol Plant* **2,** 861-872.

Decker, S.R., Carlile, M., Selig, M.J., Doeppke, C., Davis, M., Sykes, R., Turner, G. and Ziebell, A. (2012) Reducing the effect of variable starch levels in biomass recalcitrance screening. *Methods Mol. Biol.* **908**, 181-195.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M. and Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465-469.

Dhugga, K.S., Ulvskov, P., Gallagher, S.R. and Ray, P.M. (1991) Plant polypeptides reversibly glycosylated by UDP-glucose. Possible components of Golgi beta-glucan synthase in pea cells. *J. Biol. Chem.* **266**, 21977-21984.

Dixon, R.A. (2013). Microbiology: Break down the walls. *Nature* 493**,** 36-37.

Ebringerova A. and Heinze T. (2000) Xylan and xylan derivatives–biopolymers with valuable properties, 1. Naturally occurring xylans structures, isolation procedures and properties. *Macromol Rapid Commun*. **21**, 542–556

Faik, A. (2010) Xylan biosynthesis: news from the grass. *Plant Physiol.* **153**, 396-402.

Freeling, M. and Walbot, V. (1994) *The Maize Handbook*. New York:Springer-Verlag.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Jr., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A. and Wang, Z.Y. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3803-3808.

Fu, C., Sunkar, R., Zhou, C., Shen, H., Zhang, J.Y., Matts, J., Wolf, J., Mann, D.G., Stewart, C.N., Jr., Tang, Y. and Wang, Z.Y. (2012) Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol. J.* **10**, 443-452.

Guillaumie, S., Goffner, D., Barbier, O., Martinant, J.P., Pichon, M. and Barriere, Y. (2008) Expression of cell wall related genes in basal and ear internodes of silking brown-midrib-3, caffeic acid O-methyltransferase (COMT) down-regulated, and normal maize plants. *BMC Plant Biol.* **8**, 71.

Hardin, C.F., Fu, C.X., Hisano, H., Xiao, X.R., Shen, H., Stewart, C.N., Parrott, W., Dixon, R.A. and Wang, Z.Y. (2013) Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *Bioenerg Res.* **6**, 755-762.

Hartley, R.D. and Ford, C.W. (1989) Phenolic constituents of plant-cell walls and wall biodegradability. *ACS Symp. Ser.* **399**, 137-145.

Harper, A.D., and 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.

Harman-Ware, A.E., Foster, C., Happs, R.M., Doeppke, C., Meunier, K., Gehan, J., Yue, F., Lu, F. and Davis, M.F. (2016) A thioacidolysis method tailored for higher-throughput quantitative analysis of lignin monomers. *Biotechnol J*. 10.1002/biot.201600266. [Epub ahead of print].

Hatfield, R.D., Ralph, J. and Grabber, J.H. (1999) Cell wall cross-linking by ferulates and diferulates in grasses. *J. Sci. Food. Agr.* **79**, 403-407.

Iiyama, K., Lam, T.B.T. and Stone, B.A. (1994) Covalent cross-links in the cell-wall. *Plant Physiol*. **104**, 315-320.

Konishi, T., Aohara, T., Igasaki, T., Hayashi, N., Miyazaki, Y., Takahashi, A., Hirochika, H., Iwai, H., Satoh, S. and Ishii, T. (2011) Down-regulation of UDP-arabinopyranose mutase

reduces the proportion of arabinofuranose present in rice cell walls. *Phytochemistry* **72**, 1962-1968.

Konishi, T., Ohnishi-Kameyama, M., Funane, K., Miyazaki, Y., Konishi, T. and Ishii, T. (2010) An arginyl residue in rice UDP-arabinopyranose mutase is required for catalytic activity and autoglycosylation. *Carbohydr. Res.* **345**, 787-791.

Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M.A. and Ishii, T. (2007) A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. *Glycobiology* **17**, 345-354.

Kotani, A., Tsuji, M., Azama, Y., Ishii, T., Takeda, T., Yamashita, T., Shimojima, M. and Konishi, T. (2013) Purification and characterization of UDP-arabinopyranose mutase from *Chlamydomonas reinhardtii*. *Biosci. Biotechnol. Biochem.* **77**, 1874-1878.

Lee, H.V., Hamid, S.B. and Zain, S.K. (2014) Conversion of lignocellulosic biomass to nanocellulose: structure and chemical process. *ScientificWorldJournal* **2014**, 631013.

Li, R.Y. and Qu, R.D. (2011) High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass. Bioenerg.* **35**, 1046-1054.

Lindsey, K., Johnson, A., Kim, P., Jackson, S. and Labbe, N. (2013) Monitoring switchgrass composition to optimize harvesting periods for bioenergy and value-added products. *Biomass Bioenerg.* **56**, 29-37.

Mann, D.G., Abercrombie, L.L., Rudis, M.R., Millwood, R.J., Dunlap, J.R. and Stewart, C.N., Jr. (2012a) Very bright orange fluorescent plants: endoplasmic reticulum targeting of orange fluorescent proteins as visual reporters in transgenic plants. *BMC Biotechnol.* **12**, 17.

Mann, D.G.J., LaFayette, P.R., Abercrombie, L.L., King, Z.R., Mazarei, M., Halter, M.C., Poovaiah, C.R., Baxter, H., Shen, H., Dixon, R.A., Parrott, W.A. and Stewart, C.N. (2012b) Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. *Plant Biotechnol. J.* **10**, 226-236.

Martinez, V., Ingwers, M., Smith, J., Glushka, J., Yang, T. and Bar-Peled, M. (2012) Biosynthesis of UDP-4-keto-6-deoxyglucose and UDP-rhamnose in pathogenic fungi *Magnaporthe grisea* and *Botryotinia fuckeliana*. *J. Biol. Chem.* **287**, 879-892.

Molinari, H.B., Pellny, T.K., Freeman, J., Shewry, P.R. and Mitchell, R.A. (2013) Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in *Brachypodium distachyon*. *Front. Plant Sci.* **4**, 50.

Porchia, A.C., Sørensen, S. O., Scheller, H. V. (2002) Arabinoxylan Biosynthesis in Wheat. Characterization of Arabinosyltransferase Activity in Golgi Membranes. *Plant Phys.* **130**, 432-441.

Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., Davison, B.H., Dixon, R.A., Gilna, P., Keller, M., Langan, P., Naskar, A.K., Saddler, J.N., Tschaplinski, T.J., Tuskan, G.A. and Wyman, C.E. (2014) Lignin valorization: Improving lignin processing in the biorefinery. *Science* **344**, 1246843-1-1246843-10.

Rancour, D.M., Hatfield, R.D., Marita, J.M., Rohr, N.A. and Schmitz, R.J. (2015) Cell wall composition and digestibility alterations in *Brachypodium distachyon* achieved through reduced expression of the UDP-arabinopyranose mutase. *Front. Plant Sci.* **6**, 446.

Rancour, D.M., Marita, J.M. and Hatfield, R.D. (2012) Cell wall composition throughout development for the model grass Brachypodium distachyon. *Front. Plant Sci.* **3**, 266.

Rautengarten, C., Ebert, B., Herter, T., Petzold, C.J., Ishii, T., Mukhopadhyay, A., Usadel, B. and Scheller, H.V. (2011) The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in Arabidopsis. *Plant Cell* **23**, 1373-1390.

Rennie, E.A. and Scheller, H.V. (2014) Xylan biosynthesis. *Curr. Opin. Biotechnol.* **26**, 100-107.

Scalbert, A., Monties, B., Lallemand, J.Y., Guittet, E. and Rolando, C. (1985) Ether linkage between phenolic-acids and lignin fractions from wheat straw. *Phytochemistry* **24**, 1359-1362.

Scheller, H.V. and Ulvskov, P. (2010) Hemicelluloses. *Annu. Rev. Plant Biol.* **61**, 263-289.

Schmittgen, T.D. and Livak, K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* **3**, 1101-1108.

Selig, M.J., M.P., T., Sykes, R.W., Reichel, K.L., Brunecky, R., Himmel, M.E., Davis, M.F. and Decker, S.R. (2010) Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. *Ind. Biotechnol.* **6**, 104-111.

Shen, H., Fu, C.X., Xiao, X.R., Ray, T., Tang, Y.H., Wang, Z.Y. and Chen, F. (2009) Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *Bioenerg Res.* **2**, 233-245.

Shen, H., He, X.Z., Poovaiah, C.R., Wuddineh, W.A., Ma, J.Y., Mann, D.G.J., Wang, H.Z., Jackson, L., Tang, Y.H., Stewart, C.N., Chen, F. and Dixon, R.A. (2012) Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol.* **193**, 121-136.

Shen, H., Mazarei, M., Hisano, H., Escamilla-Trevino, L., Fu, C., Pu, Y., Rudis, M.R., Tang, Y., Xiao, X., Jackson, L., Li, G., Hernandez, T., Chen, F., Ragauskas, A.J., Stewart, C.N., Jr.,

Wang, Z.Y. and Dixon, R.A. (2013) A genomics approach to deciphering lignin biosynthesis in switchgrass. *The Plant Cell* **25**, 4342-4361.

Sternemalm, E., Hoije, A. and Gatenholm, P. (2008) Effect of arabinose substitution on the material properties of arabinoxylan films. *Carbohydr. Res.* **343**, 753-757.

Sykes, R., Yung, M., Novaes, E., Kirst, M., Peter, G. and Davis, M. (2009) High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy. *Methods Mol. Biol.* **581**, 169-183.

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., Hao, Z., Zhu, X., Avci, U., Miller, J.S., Baldwin, D., Pham, C., Orlando, R., Darvill, A., Hahn, M.G., Kieliszewski, M.J. and Mohnen, D. (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* **25**, 270-287.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Updegraff, D.M. (1969) Semimicro determination of cellulose in biological materials. *Anal. Biochem.* **32**, 420-424.

Vega-Sanchez, M.E., and Ronald, P.C. (2010). Genetic and biotechnological approaches for biofuel crop improvement. *Curr Opin Biotechnol* 21**,** 218-224.

Whetten, R., and Sederoff, R. (1995). Lignin Biosynthesis. *Plant Cell* 7**,** 1001-1013..

Xu, B., Escamilla-Trevino, L.L., Sathitsuksanoh, N., Shen, Z., Shen, H., Zhang, Y.H., Dixon, R.A. and Zhao, B. (2011) Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytol.* **192**, 611-625.

York, W.S., and O' Neill, M.A., (2008) Biochemical control of xylan biosynthesis – which end is up? *Curr. Opin. Plant. Biol.* **11(3)**, 258-265.

Zhou, X., Xu, J., Wang, Z., Cheng, J.J., Li, R. and Qu, R. (2012) Dilute sulfuric acid pretreatment of transgenic switchgrass for sugar production. *Bioresour. Technol.* **104**, 823-827.

Copyright © 2016 Willis, Smith, Mazarei, Zhang, Turner, Decker, Sykes, Poovaiah, Baxter, Mann, Davis, Udvardi, Peña, Backe, Bar-Peled and Stewart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.