Structure and function of Xyloglucan Xylosyltransferases

by

Alan Thomas Culbertson

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Biochemistry

Program of Study Committee:
Olga Zabotina, Major Professor
Adam Barb
Thomas Bobik
Reuben Peters
Vincenzo Venditti

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

Table of Contents

ABSTRACT

The plant cell wall is a complex network composed mainly of polysaccharides, the most abundant biopolymers on earth and a rich source of biorenewable materials. Biosynthesis of these plant polysaccharides is poorly understood, largely due to difficulties in the structural characterization of glycosyltransferases and lack of suitable substrates for *in vitro* analysis. Xyloglucan Xylosyltransferases (XXTs) initiate side-chain extensions from a linear glucan polymer by transferring the xylosyl group from UDP-xylose during xyloglucan biosynthesis. Here, we optimized protein expression and enzymatic activity conditions of XXTs through numerous N- and C-terminal truncations, various *E. coli* strains, solubility tags, and storage conditions. This procedure was used for protein expression of three XXTs (XXT1, XXT2, and XXT5) and we show that XXT5 is catalytically active *in vitro*, though at a significantly slower rate compared to XXT1 or XXT2. As no structural information was available for any of the XXTs, we built a homology model of XXT2. This model was used to predict amino acids involved in UDP-xylose binding that were verified through mutagenesis. We subsequently solved the crystal structure of XXT1 without ligands and in complexes with UDP and cellohexaose. XXT1, a homodimer and member of the GT-A fold family of glycosyltransferases, binds UDP analogously to other GT-A fold enzymes. The structures detailed here combined with the properties of mutant XXT1s are consistent with a SN_i-like catalytic mechanism. Distinct from other systems is the recognition of cellohexaose by way of an extended cleft. The crystal structure of XXT1 demonstrates that XXT1 alone cannot produce

xylosylation patterns observed for native xyloglucans because of steric constraints imposed within the acceptor binding cleft. Homology modeling of XXT2 and XXT5, using the crystal structure of XXT1 as template, reveals a structurally altered cleft in XXT5 that could accommodate a partially xylosylated glucan chain produced by XXT1 and/or XXT2. This suggests that XXT1 and XXT2 xylosylate a growing glucan chain to produce the GXXG repeat, which is then utilized by XXT5 to produce the biologically observed XXXG repeat of native xyloglucan present in most of the plants. These results allowed us to propose a model of sequential xylosylation of glucan chain synthesized by glucan synthase and support the synthesis of xyloglucan via multiprotein complex localized in plant Golgi as proposed previously.

The work described in this dissertation was supported by NSF-MCB grant #1121163 and partially by the seed-grant from Roy J. Carver Trust.

CHAPTER 1. INTRODUCTION

Plant Cell Wall

The plant cell wall is a complex and dynamic assembly of polysaccharides, lignin, and glycoproteins (Cosgrove, 2005; Burton et al., 2010; Keegstra, 2010). The cell wall strengthens the plant body but also plays key roles in plant growth, cell differentiation, intercellular communication, water movement, and defense against biotic or abiotic attack (Keegstra, 2010). The plant cell wall is an attractive resource due to the fact that two of the three most abundant biopolymers on earth, cellulose and lignin, are major components of the cell wall. This demonstrates its enormous potential for the use of the plant cell wall for industrial applications including biofuels and biomaterials (Boerjan et al., 2003; Carroll and Somerville, 2009; Pauly and Keegstra, 2010; McFarlane et al., 2014). However, degradation of the plant cell wall is energy intensive and often very expensive due to the recalcitrant nature of cell walls. It is generally considered that the production of cellulases is the major bottleneck of complete degradation of the plant cell wall, and a large effort to increase cellulase production efficiency has yielded little success (Wilson, 2009). An alternative approach is to engineer the plant cell wall that can be more easily degraded with no detrimental effects on plant fitness (McCann and Carpita, 2008; Carpita, 2012; Ciesielski et al., 2014). Engineering a plant cell wall for human applications requires a complete understanding of the biosynthesis of the plant cell wall polysaccharides.

The composition of the plant cell wall is dependent on tissue, species, cell type, and environment. Depending on these factors, there is varying abundance of cellulose, hemicellulose, pectin, glycoproteins, and lignin. Generally, there are two types of plant cell wall: primary cell walls that surround growing cells and contain mostly cellulose, hemicellulose, and pectins, and secondary cell walls that are abundant in cellulose and lignin and play more of structural and defensive roles. The composition of the cell wall is likely more of a spectrum of these types, with the primary and secondary walls being the two extremes of the spectrum (Keegstra, 2010).

Cellulose is the most abundant biopolymer on earth. It consists of a β-(1,4) linked glucan backbone with numerous inter- and intra-hydrogen bonds to form an insoluble crystalline microfibril. Cellulose serves as the load-bearing structure in the cell wall. Cellulose biosynthesis is performed by cellulose synthases (CesA)(Pear et al., 1996). Arabidopsis has ten CesA genes consisting of two groups; CesA1, -2, -3, -5, -6, and -9, which are involved in primary cell wall biosynthesis, and CesA4, -7, and -8, which are involved in secondary cell wall biosynthesis (McFarlane et al., 2014). CesAs belong to GT family 2 and contain eight transmembrane domains that channel the cellulose chain through the plasma membrane (Morgan et al., 2013). CesAs form a rosette multi-protein structure with six-fold symmetry and are the functional unit of cellulose synthesis (Mueller and Brown, 1980). These complexes typically align with microtubules, suggesting that microtubules guide cellulose synthesis (McFarlane et al., 2014).

Recently, a crystal structure of cellulose synthase from bacteria (BcsA-BcsB) was revealed (Morgan et al., 2013). This structure demonstrates that cellulose synthase contains eight transmembrane domains, six of which directly interact with the glucan chain (Morgan et al., 2013). The catalytic domain is a GT-A fold and contains the D, D, D, Q(Q/R)XRW motif that is conserved among GT family 2 members (Morgan et al., 2013). Further, a homology model of CesA from *Gossypium hirsutum* was reported and revealed the plant-specific subdomains (Sethaphong et al., 2013). These plant-specific subdomains are predicted to be important for CesA oligomerization to form the rosette complex (Sethaphong et al., 2013).

Hemicellulose polysaccharides are loosely defined as those that require a strong alkali for solubilization and directly interact with cellulose. These polysaccharides can cross-link or act as spacers between cellulose microfibrils and influence cellulose crystallinity. The main types of hemicellulose are xyloglucan, xylan, mannan, and mixed linked glucans. The first three polysaccharides contain a β-(1,4) linked glycan backbone chain of either glucan, mannan, or xylan, whereas the fourth contains a β-(1,3:1,4) linked glucan backbone chain. These polysaccharides form the "gel-like" layer of the plant cell wall and consist of roughly a third of the dry mass of the plant primary cell wall (Pauly et al., 2013).

Mannans are mainly found in green algae (*Charophyceae*) or in the seeds of plants. The backbone is either a linear mannan chain or alternating mannose and glucose backbone. The mannan backbone is biosynthesized by CSLCA enzymes. Mannans are an essential part of seed development in Arabidopsis seeds (Goubet 2003). Mannans are grouped into four classes based on backbone and side-chain

composition: mannan, glucomannan, galactomannan, and galactoglucomannan (Scheller and Ulvskov, 2010).

Xylans are most abundant hemicellulose in secondary cell walls of monocots. They consist of a β-(1,4)-linked xylose backbone which is substituted with a wide variety of other sugars depending on species and tissue (Pauly et al., 2013). The main types of xylan are glucoronoxylan or arabinoxylan, which are substituted with α-(1,2) glucoronic acid or arabinose, respectively (Scheller and Ulvskov, 2010; Pauly et al., 2013). Mixed linked glucans are found only in grasses and not in dicot species. They consist of a β-(1,3) and β-(1,4) linkages and are a non-branched homopolymer.

Pectins are a major component in the primary cell wall and consists of three major polysaccharides: homogalacturonan (HG), rhamnogalactoronan I (RG-I), and rhamnogalacturonan II (RG-II), with HG being the most abundant (Atmodjo et al., 2013). The diversity and complexity of these polysaccharides make them difficult to study. Additionally, isolation from cell wall requires chemical or enzymatic treatments to release the pectic polysaccharides that can alter their structure (Atmodjo et al., 2013)

Glycosyltransferases

The numerous oligo- and polysaccharides found in nature entail an enormous complexity with specific linkages (Rademacher et al., 1988). The formation of these highly specific linkages requires the action of glycosyltransferases (GTs) (Lairson et al., 2008). GTs catalyze the formation of glycosidic bonds by transferring a sugar moiety from an activated donor to a variety of acceptors (Lairson et al., 2008). The activated donor is typically a nucleotide

sugar (e.g. UDP-glucose, UDP-xylose, GDP-mannose), but lipid phosphates are also used. The acceptor substrates can be any other molecule with the most common being other sugars, but can also be lipids, nucleic acids, proteins, or small molecules.

GTs are classified into 105 families based on their amino acid sequence in the Carbohydrate Active Enzyme Database (Campbell et al., 1997; Coutinho et al., 2003). Most GTs adopt one of two folds, GT-A or GT-B. GT-A fold has two Rossman-like domains that form a central β-sheet that is surrounded by α-helices. These are metal dependent enzymes that require a DxD motif for metal coordination (Busch et al., 1998; Wiggins and Munro, 1998). GT-B fold glycosyltransferases also have two Rossman-like domains, but they are less tightly associated with the active site between these two domains.

In addition to their classification by overall fold, GTs are also classified by their inversion or retention in stereochemistry in product compared to that of donor substrate. Inverting GTs follow a single displacement mechanism, a mechanism similar to the inverting glycosyl hydrolases (Davies and Henrissat, 1995; Lairson et al., 2008). This entails a protein side chain acting as catalytic base to deprotonate the acceptor hydroxyl, which then facilitates a S_N2 displacement of the anomeric carbon of the activated sugar (Lairson et al., 2008). The catalytic mechanism of retaining GTs is still under heavy debate. It was first proposed that these enzymes utilized a double displacement mechanism, similar to the retaining glycosyl hydrolases. This mechanism involves a protein side chain that performs a nucleophilic attack on the anomeric carbon of the activated sugar to form a glycosyl-enzyme intermediate with inverted stereochemistry. The acceptor hydroxyl would

then nucleophilically attack this glycosyl-enzyme intermediate, inverting the stereochemistry again, resulting in a net retention of anomeric carbon stereochemistry. This mechanism has fallen into disfavor due to the lack of a suitably placed catalytic nucleophile for the first SN_2 reaction and the failure to trap a glycosyl-enzyme intermediate. Instead, retaining GTs may utilize a S_Ni-like mechanism (Persson et al., 2001), which has gained support through kinetic isotope studies (Lee et al., 2011) and crystallographic snapshots along the pathway (Yu et al., 2015). S_Ni-like reactions are similar to S_N1 reactions but the bond collapses back to the original stereochemistry in the product. This mechanism is typically proposed due to the lack of supporting evidence for other mechanisms.

Xyloglucan Biosynthesis

Xyloglucan (XyG) is a ubiquitous plant polysaccharide that is found in all land plants (Pauly and Keegstra, 2016). XyG encompasses roughly 20-25% of the hemicellulose in the primary cell wall of dicots, while only encompassing 2-10% in grasses (Scheller and Ulvskov, 2010). XyG is believed to associate with cellulose through hydrogen bonds (Valent and Albersheim, 1974). Due to this interaction, it is believed that XyG may cross-links cellulose microfibrils or it may act as a spacer to prevent the cellulose aggregation (Somerville et al., 2004).

XyG contains a β-(1,4) linked glucan backbone that is substituted with numerous glycosyl residues (Pauly and Keegstra, 2016). The type and order of glycosyl substitution is dependent on numerous factors including species, tissue, cell type, and developmental state. Additionally, non-glycosyl substitutions can also be

present, such as acetyl groups. A nomenclature has been developed to describe the glucan backbone substitution (Fry et al., 1993). For example, G represents unsubstituted glucosyl residue, X represents glucosyl residue substituted at O-6 with α- D-xylose, L represents the xylose substituted with α-(1,2) linked D-Gal*p*, and F represents the xylose substituted with α-(1,2) linked L-Fuc*p*-(1,2)-β-D-Gal*p* (Fry et al., 1993). The composition of XyG in *Arabidopsis* primarily is composed of XXXG, XXGF, and XLFG subunits (Vanzin et al., 2002; Madson, 2003). Dicots typically contain XXXG-type xyloglucan, whereas grasses typically contain XXGG-type xyloglucan (Figure 1; Vincken et al., 1997).

XyG biosynthesis occurs in the Golgi apparatus (Chevalier et al., 2010) by action of multiple highly specific glycosyltransferases (decsribed in details below in the chapter 2). In this process, a glycan synthase synthesizes the glucan chain, transporting it into the Golgi, which is then decorated by numerous GTs (Pauly and Keegstra, 2016). Following biosynthesis, XyG is transported out of the cell via secretory vesicles. Cellulose synthase-like C4 (CSLC4) is the primary synthase responsible for XyG backbone synthesis (Cocuron et al., 2007). Interestingly, co-expression of CLSC4 with xyloglucan xylosyltransferase 1 (XXT1) increased glucan production. Additionally, it was shown that N-terminus, C-terminus, and active site loop are localized on the cytosolic side of the Golgi membrane (Davis et al., 2010).

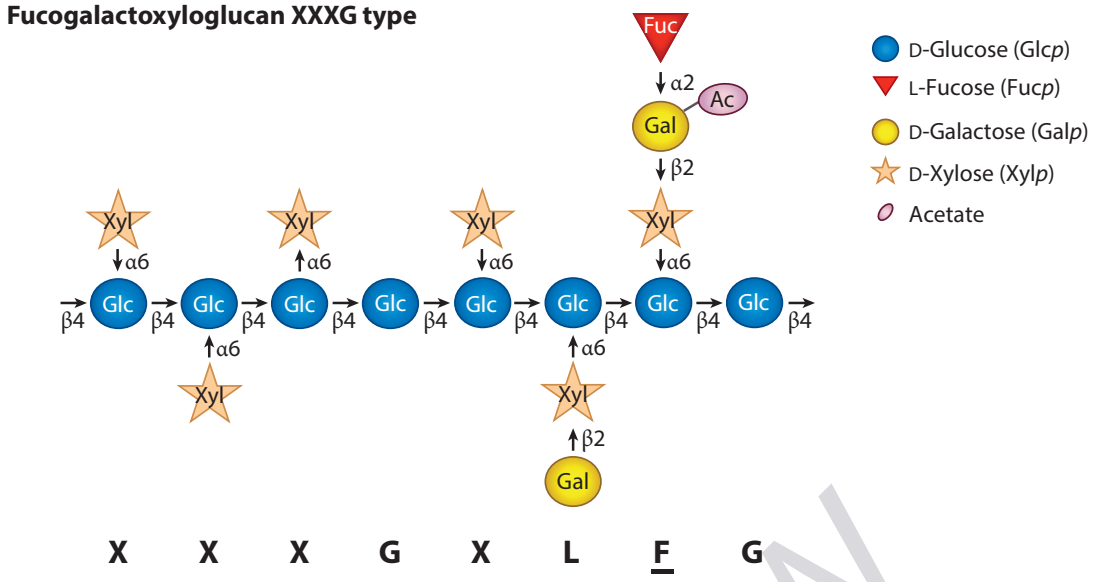

Figure 1. Schematic of the xyloglucan structure. Figure from Pauly and Keegstra, 2016.

To identify the enzymes involved in the xylosylation of the glucan backbone, Fiak *et al.* first tested enzymes that were in the same GT family as a characterized mannan GT (Faik et al., 2002). This is because the mannan GTs and XXTs both glycosylate a linear β-(1,4) glycan backbone at the O-6 position, thus it is likely that they would have high sequence similarity. This study identified XXT1 as being an enzyme that xylosylates a linear glucan backbone (Faik et al., 2002). Later, product characterization of XXT1 and XXT2 demonstrated that they both primarily xylosylate the fourth glucose from the reducing end of cellohexaose (Cavalier and Keegstra, 2006). Numerous reverse genetic studies have been investigated on the XXTs and it was found that a single knockout of either *xxt1* or *xxt2* had no change in XyG content, whereas a double knock-out of *xxt1xxt2* had no detectable XyG (Cavalier et al., 2008; Zabotina et al., 2012). This led to the proposal that XXT1 and XXT2 are at least partially redundant. The *xxt1xxt2* plants have minor morphological

phenotype changes, primarily in the root hairs (Cavalier et al., 2008; Zabotina et al., 2012). Later, XXT5 was identified to be another XXT with some distinct differences compared to XXT1 or XXT2. One difference was the fact that a single knock-out of *xxt5* resulted in a 50% reduction in xyloglucan (Zabotina et al., 2008; Zabotina et al., 2012). Additionally, attempts to demonstrate activity of XXT5 *in vitro* were all unsuccessful (Faik et al., 2002; Zabotina et al., 2008; Vuttipongchaikij et al., 2012), raising questions about the function of XXT5.

Galactosyltransferases were identified and were shown to have high specificity for the xylose they galactosylate (Madson, 2003; Jensen et al., 2012). XLT2 was shown to galactosylate the second xylose, and MUR3 galactosylates the third xylose in the XXXG subunit repeat (Madson, 2003; Jensen et al., 2012). A fucosyltransferase, FUT1, fucosylates the second galactose residue in the XLLG repeat, yielding an XLFG repeat (Perrin, 1999). All of these xyloglucan biosynthetic proteins were shown to form protein-protein interactions with each other and have been proposed to form a multi-enzyme protein complex (Chou et al., 2012; Chou et al., 2015).

Following its biosynthesis, XyG is transported to the plasma membrane via vesicles. Following incorporation in the cell wall, the structure remains highly dynamic as a result of numerous apoplastic enzymes (Frankova and Fry, 2013) which is highly correlated with cell elongation and plant growth (Pauly and Keegstra, 2016). Examples of these apoplastic enzymes include XyG endohydrolases, that hydrolyze the xyloglucan backbone releasing the XyG oligosaccharides into the

wall, and endotransglycosylases, which cleave and re-ligate the XyG to another XyG chain (Nishitani and Tominaga, 1992).

References

Atmodjo MA, Hao Z, Mohnen D (2013) Evolving views of pectin biosynthesis. Annu Rev Plant Biol **64:** 747-779

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol **54:** 519-546

Burton RA, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. Nat Chem Biol **6:** 724-732

Busch C, Hofmann F, Selzer J, Munro S, Jeckel D, Aktories K (1998) A Common Motif of Eukaryotic Glycosyltransferases Is Essential for the Enzyme Activity of Large Clostridial Cytotoxins. Journal of Biological Chemistry **273:** 19566-19572

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

Carpita NC (2012) Progress in the biological synthesis of the plant cell wall: new ideas for improving biomass for bioenergy. Curr Opin Biotechnol **23:** 330-337

Carroll A, Somerville C (2009) Cellulosic biofuels. Annu Rev Plant Biol **60:** 165-182

Cavalier DM, Keegstra K (2006) Two xyloglucan xylosyltransferases catalyze the addition of multiple xylosyl residues to cellohexaose. J Biol Chem **281:** 34197-34207

Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, Zabotina OA, Hahn MG, Burgert I, Pauly M, Raikhel NV, Keegstra K (2008) Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. Plant Cell **20:** 1519-1537

Chevalier L, Bernard S, Ramdani Y, Lamour R, Bardor M, Lerouge P, Follet-Gueye ML, Driouich A (2010) Subcompartment localization of the side chain xyloglucan-synthesizing enzymes within Golgi stacks of tobacco suspension-cultured cells. Plant J **64:** 977-989

Chou Y, Pogorelko G, Young ZT, Zabotina OA (2015) Protein-Protein Interactions Among Xyloglucan-Synthesizing Enzymes and Formation of Golgi-Localized Multiprotein Complexes. Plant Cell Physiol

Chou YH, Pogorelko G, Zabotina OA (2012) Xyloglucan xylosyltransferases XXT1, XXT2, and XXT5 and the glucan synthase CSLC4 form Golgi-localized multiprotein complexes. Plant Physiol **159:** 1355-1366

Ciesielski PN, Resch MG, Hewetson B, Killgore JP, Curtin A, Anderson N, Chiaramonti AN, Hurley DC, Sanders A, Himmel ME, Chapple C, Mosier N, Donohoe BS (2014) Engineering plant cell walls: tuning lignin monomer composition for deconstructable biofuel feedstocks or resilient biomaterials. Green Chemistry **16:** 2627-2635

Cocuron JC, Lerouxel O, Drakakaki G, Alonso AP, Liepman AH, Keegstra K, Raikhel N, Wilkerson CG (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:** 8550-8555

Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol **6:** 850-861

Coutinho PM, Deleury E, Davies GJ, Henrissat B (2003) An Evolving Hierarchical Family Classification for Glycosyltransferases. Journal of Molecular Biology **328:** 307-317

Davies G, Henrissat B (1995) Structures and Mechanisms of Glycosyl Hydrolases. Structure **3:** 853-859

Davis J, Brandizzi F, Liepman AH, Keegstra K (2010) Arabidopsis mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane. Plant J **64:** 1028-1037

Faik A, Price NJ, Raikhel NV, Keegstra K (2002) An Arabidopsis gene encoding an alpha-xylosyltransferase involved in xyloglucan biosynthesis. Proc Natl Acad Sci U S A **99:** 7797-7802

Frankova L, Fry SC (2013) Biochemistry and physiological roles of enzymes that 'cut and paste' plant cell-wall polysaccharides. J Exp Bot **64:** 3519-3550

Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau JP, Kato Y, Lorences EP, Maclachlan GA, McNeil M (1993) An unambiguous nomenclature for xyloglucan-derived oligosaccharides. Physiologia Plantarum **89:** 1-3

Jensen JK, Schultink A, Keegstra K, Wilkerson CG, Pauly M (2012) RNA-Seq analysis of developing nasturtium seeds (Tropaeolum majus): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. Mol Plant **5:** 984-992

Keegstra K (2010) Plant Cell Walls. Plant Physiology **154:** 483-486

Lairson LL, Henrissat B, Davies GJ, Withers SG (2008) Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem **77:** 521-555

Lee SS, Hong SY, Errey JC, Izumi A, Davies GJ, Davis BG (2011) Mechanistic evidence for a front-side, SNi-type reaction in a retaining glycosyltransferase. Nat Chem Biol **7:** 631-638

Madson M (2003) The MUR3 Gene of Arabidopsis Encodes a Xyloglucan Galactosyltransferase That Is Evolutionarily Related to Animal Exostosins. The Plant Cell Online **15:** 1662-1670

McCann MC, Carpita NC (2008) Designing the deconstruction of plant cell walls. Curr Opin Plant Biol **11:** 314-320

McFarlane HE, Doring A, Persson S (2014) The cell biology of cellulose synthesis. Annu Rev Plant Biol **65:** 69-94

Morgan JL, Strumillo J, Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature **493:** 181-186

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

Nishitani K, Tominaga R (1992) Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. J Biol Chem **267:** 21058-21064

Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G (2013) Hemicellulose biosynthesis. Planta **238:** 627-642

Pauly M, Keegstra K (2010) Plant cell wall polymers as precursors for biofuels. Curr Opin Plant Biol **13:** 305-312

Pauly M, Keegstra K (2016) Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan. Annu Rev Plant Biol

Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci U S A **93:** 12637-12642

Perrin RM (1999) Xyloglucan Fucosyltransferase, an Enzyme Involved in Plant Cell Wall Biosynthesis. Science **284:** 1976-1979

Persson K, Ly HD, Dieckelmann M, Wakarchuk WW, Withers SG, Strynadka NC (2001) Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. Nat Struct Biol **8:** 166-175

Rademacher TW, Parekh RB, Dwek RA (1988) Glycobiology. Annu Rev Biochem **57:** 785-838

Scheller HV, Ulvskov P (2010) Hemicelluloses. Annu Rev Plant Biol **61:** 263-289

Sethaphong L, Haigler CH, Kubicki JD, Zimmer J, Bonetta D, DeBolt S, Yingling YG (2013) Tertiary model of a plant cellulose synthase. Proc Natl Acad Sci U S A **110:** 7512-7517

Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S, Raab T, Vorwerk S, Youngs H (2004) Toward a systems approach to understanding plant cell walls. Science **306:** 2206-2211

Valent BS, Albersheim P (1974) The structure of plant cell walls: v. On the binding of xyloglucan to cellulose fibers. Plant Physiol **54:** 105-108

Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter WD (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:** 3340-3345

Vincken JP, York WS, Beldman G, Voragen AG (1997) Two general branching patterns of xyloglucan, XXXG and XXGG. Plant Physiol **114:** 9-13

Vuttipongchaikij S, Brocklehurst D, Steele-King C, Ashford DA, Gomez LD, McQueen-Mason SJ (2012) Arabidopsis GT34 family contains five xyloglucan alpha-1,6-xylosyltransferases. New Phytol **195:** 585-595

Wiggins CA, Munro S (1998) Activity of the yeast MNN1 α-1, 3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. Proceedings of the National Academy of Sciences **95:** 7945-7950

Wilson DB (2009) Cellulases and biofuels. Curr Opin Biotechnol **20:** 295-299

Yu H, Takeuchi M, LeBarron J, Kantharia J, London E, Bakker H, Haltiwanger RS, Li H, Takeuchi H (2015) Notch-modifying xylosyltransferase structures support an Si-like retaining mechanism. Nat Chem Biol

Zabotina OA, Avci U, Cavalier D, Pattathil S, Chou YH, Eberhard S, Danhof L, Keegstra K, Hahn MG (2012) Mutations in multiple XXT genes of Arabidopsis reveal the complexity of xyloglucan biosynthesis. Plant Physiol **159:** 1367-1384

Zabotina OA, van de Ven WT, Freshour G, Drakakaki G, Cavalier D, Mouille G, Hahn MG, Keegstra K, Raikhel NV (2008) Arabidopsis XXT5 gene encodes a putative alpha-1,6-xylosyltransferase that is involved in xyloglucan biosynthesis. Plant J **56:** 101-115

CHAPTER 2. THE GLYCOSYLTRANSFERASES INVOLVED IN SYNTHESIS OF PLANT CELL WALL POLYSACCHARIDES: PRESENT AND FUTURE

Modified from a paper published in *JSM Enzymology and Protein Science*

Alan T. Culbertson and Olga A. Zabotina*

Abstract

Glycosyltransferases are enzymes which transfer an activated sugar to an acceptor substrate such as polysaccharides, peptide, lipid or various small molecules. In the past 10-15 years, substantial progress has been made in the identification and cloning of genes that encode polysaccharide synthesizing glycosyltransferases. However, majority of these enzymes remain structurally and mechanistically uncharacterized. This short review will focus on the questions in biochemistry of polysaccharide synthesizing glycosyltransferases to be answered in coming years.

Introduction

Plant cell walls have been proposed to be a source of renewable energy in the form of lignocellulosic liquid biofuels. In the past 10-15 years, significant progress has been made in understanding of cell wall polysaccharide biosynthesis, particularly, in identifying and characterizing the numerous genes involved in this complex process. With respect to the challenges of revealing the genes required, molecular biology, reverse-genetics, and genomics has provided many powerful tools and significantly

advanced our understanding of plant cell wall formation. A significant body of recent reviews describes the advances and the current state in understanding of plant polysaccharide biosynthesis [1-5]. However, the progress in biochemical characterization of the gene products, glycosyltransferases, is being much slower and currently falls behind the successful genetic studies. In part, this is due to the low solubility of these enzymes and the lack of suitable enzyme assays. For example, the small stereochemical differences between sugar moieties and the multiple ways these moieties can be linked to each other, which were used by nature to achieve the wide diversity of oligo- and polysaccharide present in different types of plant cell walls, limits selection of suitable substrates and complicates characterization of products. Despite current limitations, the structural characterization and mechanisms of catalysis of plant glycosyltransferases will certainly be a subject of intensive research in the coming years due their essential function in plant cell wall biosynthesis. We present here the brief overview on the long-standing unanswered questions and directions we believe the field of cell wall polysaccharide synthesizing glycosyltransferases is headed.

Synthesis of the branched and heterogeneous polysaccharide structures requires action of multiple specific glycosyltransferases and synthases which transfer a donor sugar substrate to oligosaccharide acceptor. The donor substrates are typically activated sugars such as nucleotide sugars (UDP or GDP-bound) or, more rarely, phosphorylated sugars [6]. The structures solved for GTs from other organisms, the majority of which are involved in glycosylation of small lipophilic molecules, showed that the catalytic domains of most GTs have two types of fold,

GT-A and GT-B. However, the existence of the different type, GT-C was also proposed [6]. GT-A contains two β/α/β Rossmann-like folds tightly associated forming a continuous β-sheet and while GT-B also has two β/α/β Rossmann-like folds, they are not tightly intertwined but face each other with the active site residing between them. GT-A folds are metal dependent, which is coordinated by the well documented DxD motif [7], while GT-B are metal independent. In addition to these two structural folds, GTs are further characterized whether the chemical bond formed is an inversion or retention of stereochemistry with respect to the donor substrate. The most common donor substrate is nucleotide sugars where the sugar is linked via alpha bond. If a GT catalyzes the formation of the glycosidic bond attaching the sugar to the acceptor molecule via beta bond, the stereochemistry is inverted, while if the new glycosidic bond formed is alpha the stereochemistry is retained. The catalytic mechanism of inverting glycosyltransferases has been demonstrated to be a direct displacement S_N2-like reaction in which an active site residue acts as a general base to deprotonate the acceptor which performs a nucleophilic attack on the donor anomeric carbon [6]. The mechanisms of retaining glycosyltransferases has yet to be elucidated but numerous possibilities have been proposed such as a double displacement mechanism or a front-side single displacement (S_Ni) mechansism [8,9,10]. The plant cell wall biosynthetic enzymes are found in at least three of these classifications.

Glycosylsynthases

Cellulose Synthase (Ces) and Cellulose Synthase-Like (CSL) are integral, membrane proteins with multiple transmembrane domains (TD) which span the Golgi membrane (or plasma membrane in the case of Ces) multiple times and belong to CAZy family GT2 [11]. The first solved structure for a protein from GT2 family was the structure of the catalytic domain of polysaccharide synthesizing protein, SpsA, from B. subtilis [12], which demonstrated that catalytic domains of GT2 proteins adopt GT-A fold. The 3D structure of SpsA allowed prediction of the active site amino acids important for substrate binding and catalysis [13] and served as a prototype for the organization of other family GT2 synthases. More recently, the structure for *R. sphaeroides* Ces domains BcsA and BcsB demonstrated that TDs form a pore through which the synthesized glucan chain is translocated across the plasma membrane [14]. Another study resulted in a 3D computational model of the predicted cytosolic domain of cotton CESA (GhCESA1) [15], which showed good structural agreement between BcsA and GhCESA1. In both structures, the catalytic residues within GT-A fold included the matching motifs DDG, DCD and TED. The DDG and DCD motifs coordinate UDP and divalent cation and the D of the TED motif, most likely, acts as the catalytic base [14, 16]. The motif QRGRW in BcsA is positioned near the plasma membrane and was shown to interact with cellulose acceptor substrate, whereas in GhCESA1 a QVLRW motif was proposed to have similar function and similar positioning near plasma membrane [16]. These solved structures dismissed the long standing speculation about two active sites possibly present within the same peptide [17, 18] to explain cellulose synthesis, presenting