Characterization of the putative catalytic domain of two xyloglucan xylosyltransferases (XXTs) involved in the biosynthesis of xyloglucan in Arabidopsis thaliana

Rachel Ann Morris

Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd

Part of the Biochemistry Commons

Recommended Citation
Morris, Rachel Ann, "Characterization of the putative catalytic domain of two xyloglucan xylosyltransferases (XXTs) involved in the biosynthesis of xyloglucan in Arabidopsis thaliana" (2013). Graduate Theses and Dissertations. 13107.
http://lib.dr.iastate.edu/etd/13107

This Thesis is brought to you for free and open access by the Graduate College at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Characterization of the putative catalytic domain of two xyloglucan xylosyltransferases (XXTs) involved in the biosynthesis of xyloglucan in Arabidopsis thaliana

by

Rachel Ann Morris

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Biochemistry

Program of Study Committee:
Olga Zabotina, Major Professor
Reuben Peters
Edward Yu

Iowa State University

Ames, Iowa

2013
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER 1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Rationale</td>
<td>2</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>3</td>
</tr>
<tr>
<td>Objectives</td>
<td>3</td>
</tr>
<tr>
<td>CHAPTER 2. LITERATURE REVIEW</td>
<td>5</td>
</tr>
<tr>
<td>The Plant Cell Wall</td>
<td>5</td>
</tr>
<tr>
<td>Xyloglucan</td>
<td>16</td>
</tr>
<tr>
<td>Glycosyltransferase Structures</td>
<td>22</td>
</tr>
<tr>
<td>CHAPTER 3. MATERIALS AND METHODS</td>
<td>27</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>27</td>
</tr>
<tr>
<td>Amino Acid Alignment and Analysis</td>
<td>27</td>
</tr>
<tr>
<td>Cloning and Construct Design</td>
<td>27</td>
</tr>
<tr>
<td>Optimization of Protein Expression</td>
<td>32</td>
</tr>
<tr>
<td>XXT Protein Isolation</td>
<td>34</td>
</tr>
<tr>
<td>XT Activity Assay</td>
<td>36</td>
</tr>
<tr>
<td>Plant Growth Conditions and Complementation Assay</td>
<td>37</td>
</tr>
<tr>
<td>Genomic DNA (gDNA) Extraction and Analysis</td>
<td>38</td>
</tr>
<tr>
<td>Total Membrane Protein Extraction from Plants</td>
<td>39</td>
</tr>
<tr>
<td>Cell Wall Analysis</td>
<td>40</td>
</tr>
<tr>
<td>CHAPTER 4. RESULTS</td>
<td>41</td>
</tr>
<tr>
<td>E. coli Protein Expression</td>
<td>42</td>
</tr>
<tr>
<td>XXTmut Protein Expression in <em>A. thaliana xxt</em> Mutant Plants</td>
<td>53</td>
</tr>
<tr>
<td>CHAPTER 5. DISCUSSION</td>
<td>59</td>
</tr>
<tr>
<td>Optimization of tXXT2-His Expression and Extraction Conditions</td>
<td>59</td>
</tr>
<tr>
<td>tXXT2 Isolated in Native or Hybrid Conditions</td>
<td>60</td>
</tr>
<tr>
<td>Demonstrates XXT Activity</td>
<td>60</td>
</tr>
<tr>
<td>Molecular Modeling of XXT2 and XXT5 Predicts Catalytic Center</td>
<td>62</td>
</tr>
<tr>
<td>Both DWD and DSD Motifs are Required for XXT2 Catalytic Function</td>
<td>65</td>
</tr>
<tr>
<td>The DSD Motif is Required for XXT5 Function</td>
<td>67</td>
</tr>
<tr>
<td>Chapter/Section</td>
<td>Page</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>CHAPTER 6. CONCLUSIONS AND FUTURE DIRECTIONS</td>
<td>71</td>
</tr>
<tr>
<td>Conclusions</td>
<td>71</td>
</tr>
<tr>
<td>Future Directions</td>
<td>72</td>
</tr>
<tr>
<td>WORKS CITED</td>
<td>74</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>90</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>91</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>92</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Model of the plant cell wall</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Schematic representation of the xyloglucan structure in <em>A. thaliana</em></td>
<td>17</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Proposed mechanisms of glycosyltransferases</td>
<td>24</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Common glycosyltransferase folds</td>
<td>26</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Construct design schematic</td>
<td>31</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Amino acid alignment of XXT2 and XXT5</td>
<td>42</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>tXXT2-His protein expression in different cell lines</td>
<td>43</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td><em>E. coli</em> XXT protein expression</td>
<td>47</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>MALDI-TOF analysis of tXXT2 activity assay</td>
<td>49</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Native and hybrid tXXT2-His isolation</td>
<td>51</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>MALDI-TOF analysis of tXXT2-His isolated fractions</td>
<td>52</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Genomic DNA analysis if XXT5mut(1, 2, 12):xxt5 T2 plants</td>
<td>54</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Root phenotype of XXT5mut(1, 2, 12):xxt5 T3 plants</td>
<td>55</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>Protein analysis of HA-XXT5mut(1, 2, 12):xxt5 T2 plants</td>
<td>56</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>Cell wall analysis of HA-XXT5mut(1, 2, 12):xxt5 T2 plants</td>
<td>58</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Molecular models of XXT2 and XXT5</td>
<td>64</td>
</tr>
<tr>
<td>Table Number</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Table 3.1</td>
<td><em>A. thaliana</em> XXTmut:xtt5 transformation table</td>
<td>38</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Optimization of tXXT2-His protein expression in BL21(DE3) codon plus <em>E. coli</em></td>
<td>45</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Optimization of tXXT2-His protein extraction and lysis conditions</td>
<td>46</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Masses of expected products of XT activity assay</td>
<td>48</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the people who aided me in the research and writing of this thesis. First and foremost, I would like to thank my thesis advisor, Dr. Olga Zabotina, for her guidance and support throughout this research project. I would like to thank Dr. Gennady Pogorelko for the construction of the DXD point mutant templates as well as the tXXT5-Myc:pET-15b plasmid. I would also like to thank Dr. Alesia Tietze for construction of the XXT2 and XXT5 models. I want to thank the other members of the Zabotina lab, both past and present, Joy Bailey, James Steadham, and Andre Salazar for their knowledge and assistance through the years. I especially want to thank Yi-Hsiang Chou, graduate student in the Zabotina lab, for his unwavering support and direction during the course of my research. I would also like to thank my program of study committee members, Dr. Reuben Peters and Dr. Edward Yu, for their time and contributions to my work. I thank Desiree Gunning for all her assistance in helping me be a teaching assistant for BBMB 102. I would like to thank the Iowa State University DNA, Hybridoma, and Protein Facilities for the instruction and use of their equipment. I would like to thank the staff, graduate students, and faculty in the Roy J. Carver Department of Biochemistry, Biophysics, and Molecular Biology for their help and support. I expressly want to thank the National Science Foundation for the funding of this research. Lastly, I would like to thank my friends and family for all their love and support of me throughout my graduate studies.
Xyloglucan is the primary hemicellulosic component of the primary cell wall in dicotyledonous plants. It is characterized by a $\beta$(1-4)-D-glucan backbone that is substituted in a regular pattern with $\alpha$(1-6)-D-xylose, and can be substituted further with $\beta$(1-2)-D-galactose or $\beta$(1-2)-D-galactose-$\alpha$(1-2)-L-fucose disaccharide. Three xyloglucan xylosyltransferases (XXTs), XXT1, XXT2, and XXT5, are responsible for the D-xylosylation of the $\beta$(1-4)-D-glucan backbone in *Arabidopsis thaliana*. They are Golgi-localized type II transmembrane proteins that have putative DXD catalytic centers. All have been shown to interact in a protein complex, XXT2 and XXT5 showing the strongest interaction, and XXT1 and XXT2 have demonstrated $\alpha$-D-xylosyltransferase (XT) activity *in vitro*. The catalytic mechanism of XXT2 and XXT1, as well as the biological function of XXT5, are still unknown.

To study the functional role of DXD motifs, DXD to AXA point mutants were made in two DXD motifs (DWD and DSD) of XXT2 and XXT5. N-terminal truncated versions of these proteins were expressed in *E. coli* to study *in vitro* catalytic activity and full-length proteins were stably transformed to *xxt* knock-out *A. thaliana* to study their roles *in vivo*. Results demonstrate that both DWD and DSD motifs are required for XXT2 catalytic activity *in vitro*. Additionally, it was shown that tXXT2-His could be isolated to >90% purity using both native and hybrid conditions and exhibit XT activity.

XXT5mut(1, 2, 12):xxt5 complemented transgenic *A. thaliana* plants were generated, genotyped, phenotyped, analyzed for protein expression, and analyzed for XyG content in their cell walls. The results show that XXT5 with mutations in the DWD motif can complement the *xxt5* phenotype while XXT5 with mutations in the DSD motif only partially complement,
demonstrating that the DSD motif is essential for full XXT5 function. Though it is still unclear whether XXT5 possesses XT activity, its putative DSD catalytic center is crucial for function. It is hypothesized that either, one, XXT5 requires interaction with XXT2 (or another protein) for catalytic function or two, XXT5 does not play a direct catalytic role but promotes the activity of XXT2 or XXT1 via protein-protein interactions or by binding and stabilizing the elongating glucan backbone of XyG during xylosylation.
CHAPTER 1. OVERVIEW

Introduction

The plant cell wall is a complex structure composed primarily of the polysaccharides cellulose, hemicelluloses, and pectin, as well as the non-polysaccharide polymer lignin, and various cell wall-associated glycoproteins (1). It plays many key roles including maintaining cell structure, aiding in pathogen defense, and cell-to-cell signaling. Recently, the cell wall has been a target for biofuel production, as the polysaccharides within are energy-rich and can be fermented to ethanol (2). It is therefore very important to elucidate the mechanisms by which the plant cell wall is biosynthesized, including transcriptional and translational regulation of the enzymes involved, as well as mechanisms of their function. Many of these questions, however, remain to be answered.

Cellulose, a β(1-4)-linked D-glucan polymer, is the most abundant component of the plant cell wall and is synthesized at the plasma membrane by CESA in a rosette-shaped protein complex (3). All other cell wall polysaccharides, hemicelluloses and pectins, are synthesized in the Golgi apparatus by various membrane-bound glycosyltransferases, and then transported to the plasma membrane and out to the cell wall in a mechanism that is not fully understood (4).

Xyloglucan (XyG) is the most abundant hemicellulosic component of the primary cell wall of dicotyledonous plants (5). Arabidopsis thaliana XyG is characterized by four-residue repeats of a β(1-4)-D-glucan backbone (G) substituted at the first, second, and third positions by α-1-6-D-xylose (X), the second and third positions additionally with β-1-2-D-galactose (L) and the third position further substituted with α-1-2-L-fucose (F), simply described as XLFG subunits (see Figure 2.1, (6)). Seven proteins in A. thaliana have been identified that function in
XyG biosynthesis; a cellulose-synthase like protein to synthesize the β-glucan backbone, CSCL4, three D-xylosyltransferases, XXT1, XXT2, and XXT5, two D-galactosyltransferases, MUR3 and XLT2, and a L-fucosyltransferase, FUT1 (5). Studies have suggested that these proteins work together in a complex to synthesize the XyG subunits (7), however since there is no structural data available for cell wall glycosyltransferases, the mechanism by which synthesis occurs is largely unclear.

**Rationale**

The search for alternative fuels is a very important issue in today’s economy. One of the primary candidates for biofuel production is the polysaccharide-rich plant cell wall. The major holdup of using this energy-rich source is that the useful polysaccharides are embedded within a non-carbohydrate network of the polyphenol polymer lignin (8). This network, known as lignocellulosic biomass, will need to be separated before it can be used for fuel, as lignin is non-hydrolyzable and non-fermentable. Understanding the mechanisms by which the different components of lignocellulosic biomass, such as XyG, are synthesized will identify targets for potentially modifying it, and thus produce cell walls that are more readily used for biofuel production. Previous studies have suggested that XXT2 and XXT5 are the two key players responsible for the D-xylosylation of the β(1-4)-D-glucan backbone (7, 9). XXT2 and XXT5 are both retaining glycosyltransferases that are Golgi-localized, display a type II transmembrane topology, and contain putative DXD catalytic motifs (10). Characterization of the role of these DXD motifs will give insight as to the mechanism of XXT activity, and in the future can assist in modification of polysaccharide synthetic enzymes.
Hypothesis

Based on previous reverse genetics studies utilizing xxt1, xxt2, xxt5, xxt1xxt2, xxt1xxt5, xxt2xxt5, and xxt1xxt2xxt5 knock-out A. thaliana transgenic lines, it is hypothesized that XXT2 mediates the biosynthesis of xyloglucan in a catalytic-dependent mechanism and XXT5 mediates synthesis in a non-catalytic, structural mechanism (9). The catalytic site of XXT2, as well as the biological role of XXT5, remains unknown. Available structural data for a few glycosyltransferases have identified various characteristics of their putative catalytic sites: a conserved DXD motif, coordination of a divalent metal cofactor (such as Mn$^{2+}$), and a Rossmann-type β/α/β fold (11). XXT2 and XX5 contain multiple DXD motifs, thus it is hypothesized that at least one of these motifs is involved in the catalytic mechanism of XXT2, and might be important for XXT5 function.

Objectives

Objective 1. Characterize DXD Motif Function in XXT2

To investigate the function of two DXD motifs in XXT2, DXD to AXA point mutations were made in D126WD128 (DWD, mut1) and/or D228SD230 (DSD, mut2). N-terminal truncation versions of these mutated proteins, lacking the transmembrane domain, were cloned into pET-15b and recombinantly expressed in E. coli. These expressed proteins were then assayed for XXT activity as well as isolated using native and hybrid methods. Full-length versions of the same XXT2 mutant proteins were also cloned into pGWB-15 and stably transformed to xxt1xxt2 and xxt1xxt2xxt5 A. thaliana plants.

Objective 2. Characterize DXD Motif Function in XXT5

Two DXD motifs in XXT5, D127WD129 (DWD, mut1) and/or D228SD230 (DSD, mut2), were point-mutated to AXA motifs to investigate their role in the function of XXT5. N-
terminal truncated versions of these mutant proteins were cloned into pET-15b and expressed in
*E. coli*. Full-length versions of XXT5 DXD to AXA mutants were cloned into pEarley-201 and
stably transformed to *xt5 A. thaliana* plants. Selected T₂ and T₃ progeny were genotyped,
phenotyped, analyzed for protein expression, and analyzed for xyloglucan content in their cell
walls.
CHAPTER 2. LITERATURE REVIEW

The Plant Cell Wall

Function and Importance

The plant cell wall is a dynamic, complex network composed of polysaccharides, non-polysaccharide polymers, and proteins (see Figure 2.1, (12)). The plant cell wall plays many essential roles including providing mechanical strength, facilitating cell growth, cell-to-cell signaling, and environment-cell signaling (13). Plant cell walls also play an important part in human civilization; wood provides shelter and fuel, cotton is used to make fiber and clothing, and pectins are used in foods and pharmaceuticals (14).

Figure 2.1. Model of the plant cell wall. Figure 2 obtained from Cosgrove et. al 2005 (15). Schematic representation of the plant cell wall and cell wall biosynthesis pathways. Cellulose (purple rods) is synthesized at the plasma membrane by the cellulose synthase complex. Pectins (red and orange thick wavy lines) and hemicelluloses (blue and grey thin wavy lines) are synthesized in the Golgi apparatus and transported out of the cell in vesicles. Glycoproteins (orange ovals) and lignin polymers surround the cell wall polysaccharides.
Plant cell walls are structurally and compositionally divided into two categories; primary cell walls and secondary cell walls. Primary cell walls are typically thin and able to change and expand as a newly formed cell matures (16). Secondary cell walls, however, are more rigid and are present only when the cell stops growing (17). It should be noted, however, that these definitions are guidelines to the types of cell walls found in plants. There is a wide array of cell wall compositions, both tissue-specific and plant-specific. For example, grasses contain 10-15% pectin material and 40-50% hemicelluloses, whereas most dicots contain 30% pectin and 20-30% hemicellulose (16, 18). While the proportions of the cell wall components differ among plants and plant tissues, the components themselves and their proposed biosynthesis remain similar.

All plant cell walls are rich in high-energy polysaccharides that are potential candidates for biofuel production (8), however they are very recalcitrant and the process of conversion is not economically favorable (19). Focus is now being put on characterizing and bioengineering plants with modified cell walls that are less recalcitrant for bioethanol production (20). For this reason, it has become extremely important to not only study plant cell wall physiology and morphology, but also the mechanisms and molecular pathways by which the cell wall is biosynthesized.

**Cellulose**

*Structure and Function*

Cellulose is the major polysaccharide found within plant cell walls and is the most abundant biopolymer on Earth (21). Though there are different crystalline forms of cellulose known to exist (22), the two predominant forms found in plants are cellulose Iα and cellulose Iβ (3), which differ only in unit cell parameters, Iα being a single-chain triclinic unit cell and Iβ being a two-chain monoclinic unit cell (23, 24). Both forms are characterized by parallel, linear, \( \beta(1-4)\)-D-glucan chains, with each successive D-glucose unit rotated 180°C, forming a flat
ribbon. These parallel chains then associate via hydrogen bonds and Van der Waals forces to form bundles, known as microfibrils (25-27). Cellulose microfibrils function by aiding in cell rigidity and maintaining cell strength (27).

**Biosynthesis**

Freeze fracture images of the green algae *Micrasterias denticulate* (28) and the higher plant *Zea mays* (29) revealed oligomer complexes associated with the biosynthesis of cellulose, suggesting that the parallel β-glucan strands are synthesized concurrently within a large complex. These complexes, later termed rosettes, are symmetrical hexamers, each of which capable of synthesizing six cellulose chains, for a total of thirty-six chains being synthesized simultaneously (30). In higher plants, cellulose synthase, named CESA, was first identified in cotton (*Gossypium hirsutum*) by using sequence similarity to the bacterial cellulose synthase gene, *celA* (31, 32). *Arabidopsis thaliana* contains ten CESA genes belonging to six different groups with non-redundant functions (33). It has been shown that the rosette complexes contain three different types of CESA proteins, and those proteins differ with primary cell wall cellulose synthesis (CESA1, CESA3, CESA2, 5, 6, or 9) and secondary cell wall synthesis (CESA4, CESA7, CESA8) (1, 34, 35). It has also been shown that a membrane-bound *endo*-β(1-4)-glucanase, KOR, associates with CESA proteins and is thought to function by either separating the cellulose microfibril from the rosette, or by cleaving non-crystalline β-D-glucan from the crystalline cellulose (36, 37).

The chemical mechanism by which UDP-D-glucose (UDP-Glc) binds to the cellulose synthase cytoplasmic active site, is attached to the growing β-D-glucan chain, and is extruded from the cell is not well understood. CESA proteins contain eight putative transmembrane domains; two near the amino terminus and six near the carboxy terminus (25). The cytoplasmic
N-terminus contains a RING-type zinc finger motif, which is thought to function in the protein-protein interactions among CESA proteins (38). The catalytic site of cellulose is predicted to be on the cytoplasmic side of the plasma membrane and is composed of a D, D, D, Q/RXXRW domain. This domain is seen in many other glycosyltransferases (21) and, based on site-directed mutagenesis studies, is thought to be both the site of UDP-D-Glc binding and catalytic transfer to the growing β-D-glucan chain (39, 40). The exact catalytic mechanism, however, remains unclear.

**Hemicelluloses**

*Structure and Function*

Hemicelluloses comprise 20-30% of the primary plant cell wall in dicots and monocots, and 40-50% in grasses (1). They are typically classified into four major subgroups; xyloglucan, xylan, mannans/glucomannans, and mixed linkage glucans. They range in composition and morphology but are similar in that they contain equatorial β(1-4)-linked backbones of D-xylose (D-Xyl), D-mannose (D-Man), or D-glucose (D-Glc), with the exception of mixed linkage glucans which have β(1-3, 1-4)-linked backbones (6). Xyloglucan (XyG) is the most predominant hemicellulosic component of the primary cell wall of dicot plants (41), and will be discussed in great detail later in this chapter (see Chapter 2: Xyloglucan).

Xylans are composed of a β(1-4)-linked D-Xyl backbone which is commonly α(1-2) substituted with D-glucuronic (D-GalA) acid or L-arabinose (L-Ara), termed glucuronoxylans and arabinoxylans, respectively (6). Glucuronoxylans are the major non-cellulose component of the secondary cell wall of dicot plants (42). Arabinoxylan and glucuronoarabinoxylan are the major non-cellulose components of the primary cell wall of grasses (43). Xylans in conifers and some dicots contain a unique reducing end containing the oligosaccharide β-D-Xyl-(1-4)-β-D-
Xyl-(1-3)-α-L-Rhamnose-(1-2)-α-D-GalA-(1-4)-D-Xyl (44, 45). Mannans and glucomannans are β(1-4) linked polysaccharides that contain either D-Man, such as mannan and galactomannan, or mixed D-Man and D-Glc backbones, such as glucomannans and galactoglucomannans (6). Mixed linkage glucans, characterized by β(1-4)-D-Glc and β(1-3)-D-Glc mixed linkage backbones, are mostly found in grasses, and none to date have been found in dicots (46).

Hemicelluloses play an important structural role in the plant cell wall and are thought to associate with cellulose via hydrogen bonds in a tethering mechanism (47, 48) Hemicelluloses’ exact role in in the primary cell wall structure is still unclear, however, due to the fact that there are plants, such as celery (Apium graveolens), which contain very little hemicellulose and yet are able to function normally (49). There is evidence that hemicelluloses play an important structural role in the secondary cell wall, for example the A. thaliana irx8 mutant, which has greatly reduced xylan content in the secondary cell wall, has a severe dwarf phenotype and reduced secondary cell wall thickness (42). Hemicelluloses also play a role in signaling. Hemicellulose breakdown products, termed oligosaccharins, are able to inhibit or promote cell expansion (50, 51). For example, a nonasaccharide product of XyG digestion, termed XXFG, is capable of inhibiting auxin-induced cell elongation in etiolated pea stems (52, 53). The reverse effect has been shown with the octasaccharide XXLG and nonasaccharide XLLG XyG digestion products, suggesting that these oligosaccharides function in a very specific way (51).

**Biosynthesis**

Hemicelluloses are synthesized by Golgi-localized glycosyltransferases which catalyze transfer of a sugar molecule from a nucleotide diphosphate donor onto a specific acceptor to create a specific glycosidic bond (4). These glycosyltransferases are predicted to share a common type II membrane topology, with a short cytosolic N-terminal, a transmembrane domain, and C-
terminal globular domain within the Golgi lumen (54). The exception is the cellulose-synthase-like proteins, which have a cellulose synthase-like topology of multiple transmembrane domains (55). Many of the glycosyltransferases involved in the biosynthesis of xyloglucan in *A. thaliana* have been identified and characterized (5), and will be discussed in detail later in this chapter (see Chapter 2: Xyloglucan).

Though the enzymes have not been catalytically characterized, xylan-deficient *A. thaliana* mutants have been investigated and four β(1-4)-D-xylan synthases have been identified; IRX9 (56, 57), IRX14 (58), IRX10/GUT2, and IRX10-LIKE/GUT1(59). It is hypothesized that these proteins work together to synthesize the xylan backbone in a protein complex, though no experimental evidence has been published (6). Three proteins, IRX7, IRX8 (58), and PARVUS (60) have been implicated in synthesis of the xylan reducing end oligosaccharide, however no catalytic activity has been demonstrated. Xylan synthesis must also include α-D-glucuronsyltransferases and α-L-arabinofuranosyltransferases, as these are the most common side chains found in xylans. Activities of these transferases have been seen *in vitro*, however no proteins have been identified (61, 62).

The first purified glycosyltransferase involved in plant cell wall biosynthesis to demonstrate *in vitro* activity was TfGalT from Fenugreek (*Trigonella foenum-graecum*). TfGalT is a D-galactosyltransferase that is able to synthesize galactomannan (63). In 2004, Dhugga et. al (64) identified a mannan synthase, CSLA/ManS, in guar (*Cyamopsis tetragonoloba*). Several other members of the CSLA family from *A. thaliana* and rice (*Oryza sativa*) were studied and shown to be able to synthesize both glucomannan and mannan using GDP-D-Glc and GDP-Man as substrates (65). Using reverse genetics techniques, CSLD2, CSLD3, and CSLD5 have also
been shown to be involved in mannan synthesis, though no *in vitro* activity has been demonstrated (66).

Mixed linked glucans are synthesized by CSLF (67) and CSLH proteins (68), though their catalytic activity has not yet been demonstrated. These corresponding gene families are absent in *A. thaliana* and present in *Brachypodium distachyon*, consistent with the cell wall structural data showing that mixed linkage glucans are only present in grasses (6). When CSLF or CSLH proteins from rice are over-expressed in *A. thaliana*, small amounts of β-(1-3, 1-4)-D-glucan can be detected (67, 68). CSLH and CSLF proteins have been shown to localize to the Golgi (68), and mixed linkage glucan synthesis occurs in the Golgi (69), however β-(1-3, 1-4)-D-glucan itself cannot be detected (70). This could be due to polymer assembly occurring outside of the Golgi, or that the β-(1-3, 1-4)-D-glucan is masked within the Golgi vesicles and thus cannot be visualized.

**Pectins**

*Structure and Function*

Pectins are a class of polysaccharides characterized by containing D-galacturonic acid (D-GalA). There are four major pectin components: homogalacturonan, xylogalacturonan, rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (71). These different polymers are present not as separate molecules, but as distinct, linked domains (72, 73). Homogalacturon (HG) is the simplest, composed of linear chains of α(1-4)-linked D-GalA. It is also the most abundant pectin, accounting for approximately 60% of pectins in the plant cell wall (72). HG can be modified at the C-6 hydroxyl group with a methyl ester, or acetylated at O-2 or O-3, in a pattern that is both species- and tissue-specific (74-76). When unesterified HG is complexed with calcium (Ca$^{2+}$), it is capable of forming a gel-type structure, referred to as the egg-box model.
(77). Common substitutions on the D-GalA backbone of HG are a D-Xyl on the C-3 position and is termed xylogalacturonan (78), or a β(1-3) or β(1-2)-linked D-apiose, termed apiogalacturonan (79). HG is thought to be covalently associated with XyG; in A. thaliana cultured cells it has been shown that xyloglucan is synthesized to on a pectin primer, then secreted into the apoplast and integrated into the cell wall (80). Additionally, it was shown that isoprimerverose (IP), a diagnostic fragment of XyG that is present as a result of treatment with Driselase, is present in a pectin-rich fragment of cell wall extraction, suggesting that XyG and pectins interact covalently (81).

Rhamnogalacturonan II (RG-II) also contains an α1-4 D-GalA backbone and is highly substituted with different moieties including L-rhamnose (L-Rha), L-arabinose (L-Ara), L-fucose (L-Fuc), D-Gal, and many others in a complex pattern of four characterized subunits; A, B, C, and D (71, 73). RG-II molecules are capable of dimerizing using a boron diester bond (82) which binds a the D-Apiose residue on the A subunit and is thought to aid in stability of RG-II (83). Rhamnogalacturonan I (RG-I) is unique in that it has an alternating backbone of D-GalA-(α1-2)-L-Rha-(α1-4) that is highly branched with D-Gal and L-Ara (84). The substitution pattern of RG-I is species-, tissue-, and development-dependent (85, 86). RG-I has been known to associate with galactan, a β(1-4)-D-Gal chain(84), Type-I arabinogalactan, β(1-4)-D-Gal chain with one α(1-5)-L-Ara (87), and Type-II arabinogalactan, β(1-3)-D-Gal chain with β(1-6)-D-Gal side chains and capped with α(1-5)-L-Ara (88).

Extracted pectins are often used as thickening agents in food, especially in the making of jams and jellies (89). In plant cells, the pectin network plays an important role in cell strength and flexibility. It has been shown that the more cross-linked HG present in the cell wall, the tighter the packing and thus a stronger cell wall (90). Additionally, RG-II dimerizes via boron,
which is essential to plant cell wall strength (91). Pectins also play a large role and cell signaling, especially in pathogen defense. Many plant pathogens, such as necrotrophic fungi, cause degradation of pectin, releasing 10-15 residue oligogalacturonides, or OGs (92, 93). These OGs can elicit a number of defense responses including induction of glucanases and chitinases (94, 95), producing reactive oxygen species (96), and accumulating phytoalexins (97). Wall-Associate Kinase (WAK)-1, was identified in \textit{A. thaliana} as a receptor on the plasma membrane responsible for OG signaling (98), however the mechanism by which this occurs and how the downstream signal is carried is very unclear (99).

\textit{Biosynthesis}

The glycosyltransferases responsible for the biosynthesis of pectins are still largely unknown, but it is hypothesized that they are Golgi-localized transmembrane proteins with topologies similar to those described in hemicellulose biosynthesis (1). Due to the vast complexity of the structure of pectins and the glycosidic bond specificity of glycosyltransferases, it is assumed that there are many different proteins predicted to be involved in pectin biosynthesis (73). A \(\beta(1-4)-D\)-GalA backbone synthase, termed GAUT1, has been identified and characterized in \textit{A. thaliana} as responsible for synthesizing the backbone of HG (100). An additional protein, GAUT7, was co-purified with GAUT1 and shows 36% sequence identity, but no \textit{in vitro} \(\beta(1-4)-D\)-GalA activity (100). It was later shown that GAUT1 and GAUT7 form a hetero-complex and GAUT7 is required for the localization of GAUT1 to the Golgi lumen (101). A putative HG-specific D-xylosyltransferase, responsible for the addition of \(\beta(1-3)-D\)-Xyl, termed XGD1 has been identified in \textit{A. thaliana} and characterized using a combination of reverse genetics and recombinant expression in \textit{Nicotiana benthamiana} (102).
Two α(1-3)-D-xylosyltransferases, RGXT1 and RGXT2, have been identified and characterized as RG-II xylosyltransferases based on reverse genetics and recombinant expression studies (103, 104). At1g56550 has been shown in vitro to be a D-xylosyltransferase, and has sequence identity with RGXT1 and RGXT2, however there is currently no in vivo evidence that it is specific for RG-II synthesis (105). At4g01220 is also proposed to also be a D-xylosyltransferase based on amino acid identity, though no in vitro activity has been shown (106). The α(1-4)-rhamnosyltransferase and α(1-2)-galacturonyltransferase for the synthesis of the RG-I backbone as well as the glycosyltransferases responsible for decorating the side chain subunits of RG-I and RG-II have not been identified. In addition to the various glycosyltransferases, pectin biosynthesis also requires methyltransferases and acetyltransferases. These proteins have yet to be identified, however their activities have been demonstrated in isolated, solubilized, membrane fractions from cell culture (107).

Lignin and Cell Wall-Associated Glycoproteins

**Lignin**

Lignin is a complex, non-polysaccharide, phenolic polymer that is embedded within the cell wall matrix, and is most commonly found in secondary cell walls (108). Lignin functions in providing mechanical support and water impermeability to the cell wall by associating with cellulose and other cell wall polysaccharides (109). The lignin support also aids in plant defense by providing a barrier to protect the cell against pathogen attack (110).

Lignin is composed of three major monomers (monolignols): p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol (111). The monolignols are synthesized from phenylalanine through the general phenylpropanoid and monolignol-specific pathways, and many of the proteins responsible for these pathways have been identified and characterized (1,
Unlike the polysaccharide components of the plant cell wall, lignin does not have a defined structure and is composed of a large, diverse set of linkages. Monolignols are polymerized via a two-step mechanism. First, the phenol group is dehydrogenated, resulting in a phenol radical. This reaction is catalyzed by peroxidases that use hydrogen peroxide as a substrate, or laccases which use oxygen, however the proteins that catalyze these reactions have not been identified (113, 114). In the second step, two phenol radicals are combined, forming a new covalent linkage (111), which is a mostly random process and depends on the availability of monolignol radicals (108).

**Cell Wall-Associated Glycoproteins**

There are many different proteins associated with the plant cell wall that aid in cell wall strength, control cell wall assembly and modification, and function in signaling cascades (1). Cell wall-associated proteins are similar in that they are typically rich in one amino acid, have highly repetitive sequences, and are either highly glycosylated or not glycosylated (115). Structural proteins aid in cell wall strength and structure but do not perform a catalytic function. There are three different classes of structural proteins seen in plant cell walls; hydroxylproline (Hyp)-rich glycoproteins (HRGPs), glycine-rich proteins, and proline-rich proteins (116). These proteins adopt specific structural domains, such as X-Hyp-Hyp-(Hyp)$_n$ forming a left-handed helix (117), that provide a framework for the cell wall polysaccharides to interact. For example, extensins are a type of HRGP that contain repeating Ser(Hyp)$_4$ sequences (118), which play a key role in self-assembly of the plant cell wall (119).
Xyloglucan

Function and Importance

Xyloglucan (XyG) is found in all terrestrial plants (120) and is the major hemicellulosic component of the primary cell wall of dicotyledonous plants, comprising 10-20% of the cell wall dry weight (41). XyG in grasses, on the other hand, constitutes only 1-5% of total cell wall (121). Xyloglucan associates with cellulose via hydrogen bonds and is thought to function by preventing the cellulose microfibrils from self-associating (122).

Structure

There is a broad array of structural morphologies of XyGs present in many different plant species (41, 121, 123). Though there are differences present, all XyGs are the same in that they contain a β(1-4)-linked D-glucan backbone that is D-xylosylated α(1-6) in a regular pattern of subunits. These D-Xyl units can be further substituted, giving a highly branched polymer. Due to the fact that there are structural similarities among XyGs, a simple nomenclature for these structures has been developed (124). The β(1-4)-linked D-glucan backbone is designated “G” and α(1-6)-linked D-Xyl additions are termed “X”. Therefore, the subunit termed “XXGG” would represent four β(1-4)-linked D-Glc with an α(1-6)-linked D-Xyl on the first and second positions, with the non-reducing end on the left side. Additionally, “L” represents a β(1-2)-D-Gal-D-Xyl disaccharide, “F” an α(1-2)-L-Fuc-β(1-2)-D-Gal-D-Xyl trisaccharide, and “A” an α(1-2)-L-Ara-D-Xyl disaccharide (124). The major xyloglucan subunits found in A. thaliana are XXFG, XXXG, and XLFG (see Figure 2.2) with minor amounts of XXLG, XLLG, and XLXG (125). Additionally, the 0-6 position of D-Gal can be acetylated.
Figure 2.2. Schematic representation of the xyloglucan structure in *A. thaliana*. A schematic representation of the XLFG xyloglucan subunit found in *A. thaliana*. Abbreviations are Glc, D-glucose (G), Xyl, D-xylose (X), Gal, D-galactose (L), Fuc, L-fucose (F). The glycosyltransferases identified from *A. thaliana* that catalyze the formation of the different subunits are listed on the right. β/α1-# indicates the type of glycosidic linkage.

**Biosynthesis**

Glycosyltransferases tend to be both highly substrate- and linkage-specific, meaning that they add a specific activated sugar donor to an acceptor, resulting in a specific linkage (4). Based on this specificity and the XyG XLFG subunits seen in *A. thaliana*, it can be assumed that at least seven different proteins are involved in its biosynthesis; a β(1,4)-D-glucan backbone synthase, three α-D-xylosyltransferases, two β-D-galactosyltransferases, and one α-L-fucosyltransferase.

**Fucosyltransferase (FUT1)**

The first XyG biosynthetic enzyme to be identified was an L-fucosyltransferase from *Pisum sativum* (Pea) epicotyl microsomes (126, 127). Using amino acid sequence identity, Perrin et. al (1999) were able to identify the gene in *A. thaliana* (*AtFUT1*). To confirm FUT1 protein function, Perrin et. al recombinantly expressed FUT1 in Cos-7 cells and assayed for fucosyltransferase activity, using tamarind XyG as an acceptor, which does not naturally contain L-Fuc, and GDP-L-Fuc as the donor (126, 128). Additionally, it was shown that *A. thaliana* with
a mutation in AtFUT1 (also termed mur2) showed a reduction in L-fucose content (129, 130). Bioinformatic analysis based on sequence similarity to AtFUT1 revealed 9 other potential L-fucosyltransferases, some of whose function was later confirmed using heterologous expression (131, 132).

**Galactosyltransferase (MUR3, XLT2)**

The next XyG glycosyltransferase to be characterized was the D-galactosyltransferase MUR3, a member of the Carbohydrate-Active enZYme (CAZy) GT47 family (132, 133). It was initially identified from the A. thaliana mur3-1 and mur3-2 mutants which showed a lack of L-fucose (129). XyG isolated from these mutant plants contained increased XXXG and XLXG fragments and lacked XXFG and XXLG fragments, indicating that MUR3 was specific for the third D-Xyl position (125, 133). Using recombinant expression in Pichia pastoris, it was confirmed that MUR3 did specifically add D-Gal to the third D-Xyl from the non-reducing end (133). Since the XLFG, XLLG, and XLXG subunits are found in A. thaliana, it can be assumed that another D-galactosyltransferase is involved. Recently, an additional D-galactosyltransferase, XYLOGLUCAN L-SIDE CHAIN GALACTOSYLTRANSFERASE (XLT2) (134) was identified using RNA-Seq analysis of Tropaeolum majus (nasturtium) seeds. T-DNA A. thaliana insertion lines, xlt2, showed no XLXG or XLFG fragments, however did contain the XXFG fragment, indicating that XLT2 specifically links D-Gal to the second D-Xyl from the non-reducing end of the XyG chain (134).

**Glucan Synthase (CSLC4)**

The XyG β-D-glucan backbone synthase CSLC4, a member of CAZy GT2 family (132), was identified from Nasturtium by producing a complementary DNA (cDNA) library from mRNA extracted during the last stages of seed development. From that library, one member of
the C subfamily of cellulose synthase-like (CSL) genes was identified. When overexpressed in *P. pastoris*, it was capable of synthesizing β-D-glucan, confirming its identity as a β-D-glucan synthase (135). Additional members of GT2 family (*CSLC4, CSLC5, CSLC6*) are also thought to be involved in XyG biosynthesis (5) however they are tissue-specific (136).

**Acetyltransferase (AXY4)**

XyG in *A. thaliana* can contain an acetyl group on the C-6-hydroxy group of D-Gal (137). The mechanism by which acetyl groups are transferred from their acceptor, most likely acetyl-CoA, to their donor is still unknown (138). An *A. thaliana* mutant, termed *rwa2*, was identified that contained reduced O-acetylation in multiple cell wall polysaccharides (139). Additionally, two XyG-specific acetyltransferases, AX4 and AXYL4, were identified and characterized in vivo. The results showed that AX4 specifically acetylates XyG, while AX4L needs further characterization (140). The role of acetylation in plant cell walls, however, is still unknown. It is predicted to function by aiding in cell wall morphology as well as prevent enzymatic breakdown (138, 141, 142). Recently, it was demonstrated that acetylation of different polysaccharides is important for signaling during plant defense reactions against fungal pathogens (142).

**Xyloglucan Xylosyltransferases (XXT1, XXT2, XXT5)**

The first XyG D-xylosyltransferase was discovered in pea microsomes, having the ability to transfer UDP-D-xylose onto a β-D-glucan chain (143). Candidate *A. thaliana* genes were then identified and recombinantly expressed in *P. pastoris*. Of the seven genes expressed, one, later to be called *XT1*, a member of the CAZy GT34 family (132), was able to catalyze D-xylosyltransferase activity when cellopentose was used as an acceptor (143). *XT2*, another CAZy GT34 family member, showed 83% identity and 91% similarity to *XT1* (144) and was thought to
also act as a D-xylosyltransferase. When both XXT1 and XXT2 were recombinantly expressed in *Drosophila* S2 (Schneider 2) and *Spodoptera frugiperda* 21 (Sf21) cells, both XXT1 and XXT2 demonstrated D-xylosyltransferase activity when either cellohexose or cellopentose was used as the acceptor (144). Additionally, it was observed that both XXT1 and XXT2 were able to xylosylate three sequential positions on the cellohexose backbone, forming GGXXXG products, though XXT2 showed a slightly higher activity (144).

*A. thaliana* lines that have a T-DNA insertion in both XXT1 (*xxt1*) and XXT2 (*xxt2*) have severe root hair phenotypes, displaying shortened, bulb-like root hairs. They also contain no detectable XyG in their cell wall, as shown by OLIMP (OLIgo Mass Profiling) and HPAEC-PAD (High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection) analysis of XyG-specific endo-glucanase (XEG) digested crude cell wall (145). The plants also had no detectable isoprimerverose (IP, a characteristic disaccharide of digested XyG) when digested with the hydrolase mixture Driselase, and no labeling when probed with XyG-specific antibodies, further confirming that no XyG is present (145). *xxt1* and *xxt2* single-mutant lines, however, do not show a dramatic phenotype and display 10.2% and 20.8% reduction in XyG, respectively (145). It is hypothesized that these two genes are partially redundant, as shown by the fact that the single mutants display different labeling when probed with XyG-specific antibodies (145).

Another putative D-xylosyltransferase, XXT5, also a GT34 family member, was also confirmed to be involved in XyG biosynthesis (146). XXT5 T-DNA insertion lines (*xxt5*) display a shortened root hair phenotype, similar to that seen in the *xxt1xxt2* double mutant, and approximately 50% reduction of total XyG in the cell wall. *xxt5* display a slightly altered XyG composition, having a decrease in XXXG, XXFG, and XXLG fragments and an increase in
XXG and GXXG fragments compared to wild type (146). However, no in vitro D-xylosyltransferase activity for XXT5 has been demonstrated to date (143, 146).

Further investigation of XXT knock-out double (xxt1xxt2, xxt1xxt5, xxt2xxt5) and a triple mutant (xxt1xxt2xxt5) has led to the hypothesis that XXT1 and XXT2 are responsible for xylosylating most of the XyG, and at least one of these enzymes must be present for all XyG subunits to be synthesized. It is also hypothesized that XXT5 is required for complete XyG synthesis, though potentially not for catalysis (9, 147). The xxt2xxt5 and xxt1xxt5 double mutants display ~50% decrease in XyG content, similar to the xxt5 single mutant, as well as a shortened root hair phenotype (9, 146). Overexpression of XXT1 can complement the xxt2, xxt5, and xxt2xxt5 mutants, XXT2 can complement xxt2 and weakly complement xxt5 and xxt2xxt5 (147). XXT5 can complement xxt2, xxt5, but not the xxt2xxt5 (147). The xxt1xxt2xxt5 triple mutant displays a shortened root phenotype and contains no detectable amount of XyG, similar to that of the xxt1xxt2 double mutant (9). Additionally, it was observed that the lack of XyG in xxt1xxt2 and xxt1xxt2xxt5 plants does not cause a compensatory change in other cell wall polysaccharides (9). Glycome profiling using carbohydrate epitope-specific monoclonal antibodies (148-150) was done in roots, hypocotyls, and stems to analyze the plant cell wall structural changes due to XXT knock-out (9). The xyloglucan epitope localization patterns of xxt1, xxt2, xxt5, xxt1xxt2, xxt1xxt5, xxt2xxt5, and xxt1xxt2xxt5 are distinctly different from each other in different tissues, suggesting that even though XXT1 and XXT2 have redundant activities and XXT5 so far displays no XT activity, they serve different biological functions and are all essential to XyG biosynthesis (9).

The other four members of the A. thaliana GT34 family (XXT3, XXT4, GT7, and GT6) have been studied and it was shown XXT4 has putative D-xylosyltransferase activity, based on recombinant expression of N-terminal putative transmembrane domain truncation mutants in
*Escherichia coli* (147). Overexpression of XXT3 in xxt2, xxt5, and xxt2xxt5 mutants fully complements phenotype, indicating that it is also a D-xylosyltransferase, however no *in vitro* or *in vivo* XT activity has been shown (147).

Using the promoter:GUS reporter assay system, the expression profile of XXT1, XXT2, XXT5, XXT3, and XXT4 was analyzed (147). XXT1, XXT2, and XXT5 are expressed in most tissues, with the strongest expression in growing tissues. The other putative XTs are only expressed in specific tissues; XXT4 is in the rosette, pedicel, and vascular tissue of the stem, and XXT3 is in the stem 2 and floral buds (147).

Once the basic components of *A. thaliana* XyG biosynthesis were identified, it was hypothesized that they work together in a protein complex. To address this question, Bimolecular Fluorescence complementation (Bi-FC) constructs of XXT1, XXT2, XXT5, and CSLC4 (151) were generated and transiently expressed in *A. thaliana* wild type protoplasts and the resulting fluorescence signal was quantified using flow cytometry (7). XXT2 and XXT5 showed the highest interaction, XXT2 and XXT5 are able to form homodimers, and it was shown that the XXT2 homodimer is associated via disulfide bonds (7). CSLC4 is also able to form homodimers and strongly interacts with XXT5 and XXT2 (7). These results indicate that a multiprotein complex comprised of XXT1, XXT2, XXT5, and CSLC4 does exist within *A. thaliana* to synthesize the XyG backbone.

**Glycosyltransferase Structures**

Glycosyltransferases play a role, direct or indirect, in almost all aspects of cell function; cell-cell interactions (152), cell growth and development (153), infection and immunity (154), signaling, and metabolism (155). Glycosyltransferases are classified in the Carbohydrate-Active enZYme (CAZy) database into 94 families based on amino acid sequences (132, 156).
Glycosyltransferases are typically highly specific for the type of glycosidic bond they generate, therefore there is a wide variety of glycosyltransferases predicted in cells, most of which have yet to be characterized (157).

Mechanisms

Glycosyltransferases catalyze the general mechanism of transferring a sugar residue from a donor to an acceptor, thereby generating a new glycosidic bond (157). Donors are typically activated nucleotide-diphosphate (NDP)-sugars and acceptors can be anything from carbohydrates or lipids, to proteins or small molecules (54).

Glycosyltransferases are mechanistically classified into two groups; inverting and retaining. The classification is based on the stereochemistry at the anomeric carbon of the sugar in the formed glycosidic bond in relation its stereochemistry on the donor molecule (11). For example, XXT2, a retaining glycosyltransferase, retains the α-conformation from the donor UDP-α-D-Xyl to the α(1-6) linkage on the β-D-glucan backbone (144). Glycosyltransferases that invert this conformation are termed inverting glycosyltransferases.

The inversion mechanism is thought to be an S$_\text{N}2$ reaction; direct attack of the acceptor onto the anomeric carbon of the donor with an oxocarbenium-ion transition state (see Figure 2.3). In this mechanism, the glycosyltransferase abstracts a proton from the acceptor, allowing it to act as a nucleophile, as well as stabilize the transition state (11). Depending on the glycosyltransferase, different amino acids such as Asp (158), His (159), and Glu (160) have been identified from X-ray structures as being the base catalyst. One notable exception is the retaining glycosyltransferase POFUT1, which is thought to use an Arg in a S$_\text{N}1$-like mechanism (161).

The mechanism of retaining glycosyltransferases, however, is highly debated. The first proposed mechanism is a double displacement; first, nucleophilic attack of the enzyme to the
anomeric sugar, generating a sugar-enzyme intermediate. Second, there is nucleophilic attack of the acceptor onto the sugar, releasing the enzyme (162, 163). The second proposed mechanism is $S_{N)i}$-like; the acceptor nucleophile attacks on the same side as the leaving nucleotide diphosphate. In this mechanism, the enzyme stabilizes the transition state by shielding the oxocarbenium ion (see Figure 2.3, (164-166)).
Structural Information

Of the approximately 89,500 structures in the Protein Data Bank (www.rcsb.org/pdb), only 2% are classified as glycosyltransferases. 20.8% of those are from *H. sapiens*, 12.4% are from *O. cuniculus*, 5.1% are from *E. coli*, and only 0.6% from *A. thaliana*. None of these known structures are involved in cell wall biosynthesis; almost all of them are involved in protein glycosylation or secondary metabolism. The cell wall-synthesizing glycosyltransferases, with the exception of cellulose synthase and cellulose synthase-like proteins, are predicted to share a similar topology of a type II transmembrane protein (4). The C-terminal catalytic globular domain is located within the Golgi lumen and is connected to a 20-30 amino acid luminal stem region. This region is connected to one transmembrane domain followed by a short, cytosolic N-terminus (4). Cellulose synthase, and it is hypothesized that cellulose synthase-like proteins are similar, have several transmembrane domains, with the catalytic site in the cytoplasm (25).

Though there is a large diversity in type and amino acid sequence of glycosyltransferases, there are only two general three-dimensional folds seen in known X-ray structures, referred to as the GT-A and GT-B folds (see Figure 2.4, (167)). It should be noted that these folds exist outside of glycosyltransferase types of proteins and therefore is not a diagnostic for characterizing proteins of unknown function (168). The GT-A fold was first described in 1999 by Charnock and Davies (169) in the crystal structure of SpsA, a glycosyltransferase involved in the synthesis of the spore coat of *Bacillus subtilis*. The GT-A fold has two Rossmann-like β/α/β domains and two distinct substrate-binding sites; a donor-site and an acceptor-site (170, 171). Many GT-A proteins also contain a conserved DXD motif in their catalytic center which coordinate a divalent metal cation, such as Mn$^{2+}$ or Mg$^{2+}$, and potentially act as the catalytic base in the chemical reaction (see Figure 2.4A, (172, 173)). The GT-B fold was first described in the
crystal structure of β-glucosyltransferase from bacteriophage T4 in 1994 by Vrielink et. al (174).

The GT-B fold is similar to GT-A in the fact that it also contains two Rossmann-like β/α/β domains, except they are located farther from each other and function as distinct donor- and acceptor-binding sites (see Figure 2.4B, (157)).

Using sequence searches and Basic Local Alignment Search (BLAST) tools, a third fold, named GT-C, was proposed, however there is currently no structural evidence (177). This fold is much different from that of GT-A and GT-B in that it contains 8-13 transmembrane domains and its active site is in a loop region (178, 179). The type of fold does not predict the mechanism of the glycosyltransferase; there are retaining and inverting enzymes which have either GT-A or GT-B folds (156).
CHAPTER 3. MATERIALS AND METHODS

Instrumentation

All PCR-amplifications were done in MyCycler Thermal Cycler (BioRad). All primers were synthesized by the Iowa State University DNA Facility using MerMade-192 synthesizer (BioAutomation). Constructs were sequenced by the Iowa State University DNA Facility by Sanger sequencing using the 3730xl DNA Analyzer (Applied Biosystems). Phenotype analysis was done using a Leica DMIRE2 microscope (www.leica.com) equipped with a Retiga 1300 camera (www.qimaging.com). Matrix-Assisted Laser Desorption/Ionization (MALDI)-Time of Flight (TOF) Mass Spectrometry (MS) was done using Perseptive Biosystems Voyager-DE PRO MS at the Iowa State University Protein Facility. High-Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) was doing using an ICS Dionex System (www.dionex.com).

Amino Acid Alignment and Analysis

Amino acid sequences of XXT2 and XXT5 were aligned using CLC Sequence Viewer 6.5.3 software (CLC Bio, www.clcbio.com). Alignment parameters were gap open cost 10, gap extension cost 1, end gap cost as any other, and very accurate alignment. N-terminal transmembrane domain of XXT2 and XXT5 was predicted using HMMTOP protein topology software (www.enzim.hu/hmmtop/).

Cloning and Construct Design

Site-Directed Mutagenesis Amplification

XXT2mut1 and XXT5mut1 were amplified from XXT2 and XXT5 A. thaliana cDNA, respectively, using the primer sets XXT-F/XXTm1-R and XXTint-F/XXT-R (see Figure 3.1A).
and Appendix C). PCR was done using Encylco DNA Polymerase (Evrogen, PK001) according to manufacturer’s instructions. Amplification was done by 95°C for 3 min, then 20 cycles of 94°C 10 sec, 56°C 20 sec, 72°C 90 sec, and final extension of 72°C for 3 min. The resulting fragments were analyzed by electrophoresis on ethidium bromide-containing 1% agarose. The bands were then cut from the gel, placed in 500µL 6M KI, and incubated at 65°C for 3 minutes, when the gel had completely dissolved. 20µL of 10% silicon dioxide (0.5-10µm, Sigma, S5631-100G) solution was added and incubated on ice for 5 min. Solution was centrifuged at 10,000xg for 30 seconds and supernatant was removed. The pellet was washed 3 times with 100µL DNA wash buffer (100mM NaCl, 10mM Tris, pH 8.0, 1mM EDTA, 50% ethanol (%v/v)). Once final wash was removed, pellet was dried at 65°C for 2 min. DNA was eluted from the pellet by washing twice with 10µL water, combined.

The purified fragments were then mixed with Encyclo DNA polymerase (Evrogen) using manufacturer’s instructions, denatured at 95°C for 3 min, allowed to self-anneal by slow reduction of the temperature (1°C/min) to 56°C, and final incubation at 72°C for 10 min. The resulting full-length fragment was then gel-purified as described previously. A final PCR-amplification of the full-length XXTmut1 was done using XXT-F and XXT-R primers using Encyclo DNA polymerase (Evrogen).

XXT2mut2 and XXT5mut2 were made in a similar fashion, except using XXT-F, XXTint-R and XXTm2-F, XXT-R during the initial amplification step (see Figure 3.1A and Appendix C). XXT2mut12 and XXT5mut12 were made the same as XXT2mut1 and XXT5mut1, except using XXT2mut2 and XXT5mut2, respectively, as the initial template. All XXT2mut(1, 2, 12) and XXT5mut(1, 2, 12) constructs were cloned into pENTR following manufacturer’s instructions (pENTR/D-TOPO Cloning Kit, Invitrogen, K243520), transformed to DH10B E. coli cells
(Invitrogen, 18297-010), and selected on 1%-agarose supplemented with lysogeny broth (LB) media and 100µg/mL ampicillin.

**Construct Design for E. coli Expression**

N-terminal truncated XXT2 and XXT5 were amplified from *A. thaliana* cDNA with the primers tXXT2-F, XXT2-R and tXXT5-Myc-F, XXT5-R (see Appendix C), respectively, using Encyclo DNA polymerase (Evrogen). Obtained PCR fragments and pET-15b vector (Novagen, 69661-3) were cut using appropriate restriction enzymes (Promega) according to manufacturer’s instructions and ligated using T4 DNA Ligase (Thermo Scientific, EL0014) to make tXXT2-His:pET-15b and tXXT5-Myc:pET-15b. These ligations were transformed to chemically competent DH10B *E. coli* using 42°C heat shock and selected on 1% LB-agarose plates containing 100µg/mL ampicillin. Transformed colonies were identified using GoTaq Green Master Mix (Promega, M712) and primers listed previously. Plasmid DNA (pDNA) was extracted from confirmed colonies (PureYield Plasmid Miniprep System, Promga, A1223) and sequenced to confirm correct insertion (Iowa State University Protein Facility). N-terminal truncated DXD mutants were cloned to pET-15b in similar way, using XXTmut:pENTR vectors described previously as initial DNA template. All final constructs were transformed to chemically competent BL21(DE3) codonplus *E. coli* (Agilent Technologies, 230245) using 42°C heat shock and selected on 100µg/mL ampicillin plus 34µg/mL chloramphenicol 1% LB-agarose plates.

**Construct Design for A. thaliana Expression**

Full-length XXT2mut(1, 2, 12):pENTR and XXT5mut(1, 2, 12):pENTR constructs generated previously were moved to pGWB-15 (180), which confers hygromycin selection, and pEarley-201 (Arabidopsis Biological Resource Center), which confers 2-amino-4-(hydroxyl
(methyl) phosphonoyl) butanoic acid (Basta) resistance, respectively, using the Gateway LR Clonase reaction (Invitrogen, 11791-019), transformed to chemically competent DH10B *E. coli* using 42°C heat shock, and selected using 50µg/mL kanamycin. Transformed colonies were identified using GoTaq Green Master Mix (BioRad) with the primers XXT2-F, XXT2-R and XXT5-F, XXT5-R, respectively (see Appendix C). Plasmid DNA (pDNA) was extracted from confirmed colonies as described previously and sequenced to confirm correct insertion (Iowa State University Protein Facility). Final XXT2mut(1, 2, 12):pGWB-15 and XXT5mut(1, 2, 12):pEarley-201 plasmids were transformed to GV3101 *Agrobacterium tumefaciens* using electroporation and selected using 50µg/mL kanamycin, 30µg/mL gentamycin, and 10µg/mL rifampicin.
Figure 3.1. **Construct design schematic.** Schematic of construct and PCR design for (A) DXD to AXA point mutations, (B) N-terminal truncation mutants utilized for *E. coli* expression, (C) over-expression in *A. thaliana*, and (D) genotyping of xxt5, xxt1xxt2, and xxt1xxt2xxt5 mutant *A. thaliana* plants. Small boxes in (B) represent the lac operator and T7 promoter sequences found in the pET-15b plasmid. Thin lines represent non-coding regions, thick boxes represent coding regions. Arrows indicate direction and relative position of PCR primers, with names listed above. TM represents transmembrane domain.
Optimization of Protein Expression

Cell Lines

tXXT2-His:pET-15b was transformed to BL21(DE3) (Invitrogen), BL21(DE3) codon plus (Agilent Technologies), C43(DE3) (Lucigen), Rosetta 2 (DE3) (EMD Millipore), and ArcticExpress(DE3) codon plus (Agilent Technologies) E. coli lines by using chemically competent cells and heat shock at 42°C. Transformed colonies were selected by 1% LB-agarose plates containing 100µg/mL ampicillin. Pilot expression was carried in 5mL culture volume. Cells were grown at 37°C until OD_{600} = 0.5. Cells were then moved to 16°C for 1 hour and induced with β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5mM. Cells were then grown at 37°C for four hours. Cells were harvested using centrifugation (14,000 rpm for 5 minutes) and the media removed. The cells were then suspended in 500µL lysis buffer (20mM HEPES, 300mM NaCl, 1mg/mL lysozyme, pH 7.4) and incubated on a rotational shaker at room temperature for 30 minutes. Lysates were frozen in liquid nitrogen and allowed to thaw in cold water for five freeze/thaw cycles. Lysates were then centrifuged at 14,000rpm for 30 min at 4°C and the crude lysate (supernatant) separated from the pellet. 20µL of both fractions were analyzed via SDS-PAGE followed by western blotting with Anti-6xHis AB. Relative band intensities of western blots were analyzed with ChemiDoc XRS+ Imager (BioRad).

Growth and Induction Conditions

tXXT2-His:pET-15b expressed in BL21(DE3) codon plus E. coli was used for growth and induction conditions pilot expression. There were three growth times: 4, 8, and 18 hours after induction, four induction concentrations of IPTG: 1mM, 0.7mM, 0.5mM, and 0.25mM, and three incubation temperatures following induction: 37°C, 28°C, and 16°C. A 5mL culture volume was used for each growth time/[IPTG]/growth temperature combination. All pilot tests were grown at
37°C until OD\textsubscript{600} = 0.5, incubated at 16°C for 1 hour, then induced with IPTG and moved to the appropriate temperature. Cells were harvested and lysed as described previously. 20µL of soluble and insoluble fractions were analyzed via SDS-PAGE followed by western blotting with Anti-6xHis AB. Relative band intensities of western blots were analyzed with ChemiDoc XRS+ Imager (BioRad)

**Lysis Conditions**

\textit{tXXT2-\textit{His:pET-15b}} expressed in BL21(DE3) codon plus \textit{E. coli} was used for the lysis condition tests. 5mL pilot cultures were grown at 37°C, moved to 16°C for 1 hour, induced with 0.5mM IPTG, and grown at 28°C for 18 hours. Cells were harvested as described previously. Lysis buffers varied by pH and buffer (200mM): MES at pH 5.0, HEPES at pH 7.4, and Tris at pH 8.8, and [NaCl]: 300mM, 500mM, 1M. Lysis conditions were performed as described previously, using all possible pH-Buffer/[NaCl] conditions. 20µL of insoluble and soluble fractions were analyzed via SDS-PAGE followed by western blotting with Anti-6xHis AB. Relative band intensities of western blots were analyzed with ChemiDoc XRS+ Imager (BioRad).

**SDS-PAGE, Coomassie Stain, and Western Blot Analysis**

100µL of each fraction was added to 33.3µL 4X denaturing protein loading buffer (50mM Tris, pH 6.8, 10% glycerol (%v/v), 0.1% bromophenol blue (%v/v), 2% SDS (%w/v)) and boiled at 100°C for 10min. 20µL was loaded onto a 1.0mm, 15-well, 10% sodium dodecylsulfate (SDS)-polyacrylamide gel and run in Tris/Glycine buffer (25mM Tris, pH 8.8, 250mM glycine) containing 1% SDS (%w/v) at 80V 30 min, 110V 2 hours. For coomassie staining analysis, gel was incubated in coomassie R-250 stain (0.25% coomassie R-250 (%w/v), 50% methanol (%v/v), 10% acetic acid (%v/v)) for 2 hours at room temperature then destained...
(10% acetic acid, 50% methanol) at room temperature for 18 hours. For Western blot analysis, protein was transferred to a nitrocellulose membrane at 100V for 1 hour in Tris/glycine buffer containing 20% methanol (%v/v). The membrane was then blocked for 1 hour at room temperature (5% milk, 25mM Tris, 150mM NaCl, 2mM KCl, pH 7.4, 0.01% Tween-20), and washed (1xTBST, 25mM Tris, 150mM NaCl, 2mM KCl, pH 7.4, 0.01% Tween-20) 3 times, 10 min each, at room temperature. The membrane was then incubated in Anti-His AB (Santa Cruz, 1:10,000 in 1xTBST, 0.5% milk) for 18 hours at 4°C. The membrane was then washed 3 times, 10 min each, at room temperature, then incubated in Anti-rabbit AB (Invitrogen, 1:10,000 in 1xTBST, 0.5% milk) for 1 hour at 4°C, and finally washed 3 times, 10 min each, at room temperature. Visualization was done on a ChemiDoc XRS+ Imager (BioRad) using HyGlo Quick Spray (Denville Scientific) according to manufacturer’s instructions. Membranes were exposed for 10 minutes total with images taken every 2 minutes.

**XXT Protein Isolation**

**Protein Expression**

Overnight cultures of BL21(DE3) codon plus *E. coli* (Agilent Technologies) were made by growing at 37°C with 200rpm shaking for 18 hours in LB media supplemented with 100µg/mL ampicillin plus 34µg/mL chloramphenicol. For protein expression, 1µL overnight culture/mL expression culture was grown at 37°C with 200rpm shaking until OD₆₀₀ = 0.5. Culture was then shaken at 16°C for hour, and induced with IPTG to a final concentration of 0.5mM. The culture was then grown at 28°C with 200rpm shaking for 18 hours.

**Cell Lysis**

Induced cells were harvested using centrifugation (14,000 rpm for 5 minutes) and the media removed. The cells were then resuspended in lysis buffer (20mM HEPES, 300mM NaCl,
1mg/mL lysozyme, pH 7.4) at a concentration of 100µL buffer/1mL initial culture volume and incubated on a rotational shaker at room temperature for 30 minutes. Lysate was then centrifuged at 14,000rpm for 30 min at 4°C and the crude lysate (supernatant) separated from the pellet.

**Native Isolation**

Obtained crude cell lysate was applied to Ni-NTA resin (ThermoScientific, 1mL lysate/250µL resin) and incubated on a rotational shaker at 4°C for 18 hours. Lysate/resin slurry was then moved to a 10mL column (BioRad) and allowed to drain by gravity flow (FT, flow-through). The column was then washed 4 times (4mL buffer/mL resin) with wash buffer (20mM HEPES, 300mM NaCl, 25mM imidazole, pH 8.0) via gravity flow at 4°C. Protein was eluted using elution buffer (20mM HEPES, 300mM NaCl, 250mM imidazole, pH 7.4) from the column via gravity flow and collected in 1mL fractions. 20µL of each sample was analyzed by SDS-PAGE followed by coomassie R-250 staining and western blotting with Anti-His AB as described previously.

**Hybrid Isolation**

Obtained pellet from cell lysate was suspended in guanidine lysis buffer (20mM HEPES, 300mM NaCl, 6M guanidine-HCl, pH 7.4, 100µL buffer/mL original culture) at room temperature for 1 hour. The lysate was then centrifuged at 14,000 rpm for 30 min at 4°C and the lysate (supernatant) was separated from the pellet. Lysate was then applied to Ni-NTA resin (ThermoScientific, 1mL lysate/250µL resin) and incubated on a rotational shaker at 4°C for 18 hours. Lysate/slurry was then moved to a column (BioRad) and allowed to drain by gravity flow (FT, flow-through). The column was then washed 4 times (4mL buffer/mL resin) with denaturing wash buffer (20mM HEPES, 300mM NaCl, 8M Urea, pH 8.0) and 4 times with native wash buffer (20mM HEPES, 300mM NaCl, 25mM imidazole, pH 8.0) all via gravity flow...
at 4°C. Protein was eluted using elution buffer (20mM HEPES, 300mM NaCl, 250mM imidazole) from the column via gravity flow and collected in 1mL fractions. 20µL of each sample was analyzed by SDS-PAGE followed by coomassie R-250 staining and western blotting with Anti-His AB as described previously.

**XT Activity Assay**

Activity Assay

Activity assays were done using crude *E. coli* cell lysates. 200µL lysate was incubated with 1µL 1% cellohexose (Seikagaku Corporation), 10µL 20mM UDP-xylose (Complex Carbohydrate Research Center), and 88µL activity assay buffer (100mM HEPES, 5mM MnCl₂, pH 7.4) for 18 hours at 28°C with 200rpm shaking. To extract product, 1.5mL of 100% ethanol was added to reaction and incubated at -20°C for 6 hours, centrifuged at 14,000rpm for 30 min at 4°C, and the supernatant completely removed. The pellet was suspended in 40µL water at room temperature for 15 min followed by centrifugation at 14,000rpm for 10 min. Resulting supernatant was used for MALDI-TOF analysis.

**MALDI-TOF Analysis**

2µL supernatant was mixed with 1µL DHB solution (1mg/mL 2,5-dihydroxybenzoic acid (DHB) in 100% acetonitrile, 1% trifluoroacetic acid (%v/v, TFA)). 1µL of solution was applied to a 100-well MALDI-TOF plate and allowed to dry. Samples were analyzed using Perseptive Biosystems Voyager-DE PRO MS. Data collection was done using reflector mode, delayed extraction, positive polarity, 20000V accelerating voltage, 93.8% grid voltage, 350nsec extraction delay time, laser intensity 4000. Three datasets of each spot were analyzed, 100 laser shots/dataset.
Plant Growth Conditions and Complementation Assay

Plant Material and Growth Conditions

*Arabidopsis thaliana* plants were grown in Sunshine-100 soil in a growth incubator set at 16-h-light/8-h-dark photoperiod at 22°C. *A. thaliana* seeds were sterilized by incubating in 70% ethanol (%v/v) for two min, then in 50% bleach, 0.1% Triton X-100 for eight minutes. Seeds were then washed four times; five minutes each wash, in sterile water. Finally, the seeds were suspended in 0.5% agarose and stored at 4°C for 18 hours. Sterilized seeds were plated onto one-half strength Murashige and Skoog medium, 0.3% Gel-Rite, containing selection if required, and grown under 16-h-light/8-h-dark photoperiod conditions in a growth incubator a 22°C.

Plant Transformation and Selection

XXT2mut(1, 2, 12):pGWB-15 and XXT5mut(1, 2, 12):pEarley-201 transformed GV101 *A. tumefaciens* was used to transform appropriate *xxt* mutant *A. thaliana* (see Table 3.2) using the floral dip method according to Zhang et. al 2006 (181). Briefly, 250mL of LB media supplemented with 50µg/mL kanamycin, 30µg/mL gentamycin, and 10µg/mL rifampicin was inoculated with XXTmut:pGWB-15/pEarley-201 transformed *A. tumefaciens* and grown 15 hours at 28°C. Cells were harvested by centrifugation at 5,000rpm for 10 min. Cells were then suspended in 400mL transformation buffer (5% sucrose (%w/v), 0.5% MES (%w/v), pH 5.7, 0.02% Silwet (%v/v)). Plants were transformed by immersion in *A. tumefaciens* solution for five minutes. XXT2mut-complemented seeds (*xxt1xxt2* and *xxt1xxt2xxt5*) were selected on one-half strength Murashinge and Skoog (MS) media, 0.3% Gel-Rite, containing 25µg/mL hygromycin-B. XXT5mut-complemented seeds (*xxt5*) were selected by spraying germinated seeds with 2-amino-4-(hydroxyl (methyl) phosphonoyl) butanoic acid (BASTA, 250mg/L) three times.
<table>
<thead>
<tr>
<th>Construct</th>
<th><em>A. thaliana</em> Line(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXT2mut1:pGWB-15</td>
<td>xxt1xxt2\n xxt1xxt2xxt5</td>
</tr>
<tr>
<td>XXT2mut2:pGWB-15</td>
<td>xxt1xxt2\n xxt1xxt2xxt5</td>
</tr>
<tr>
<td>XXT2mut12:pGWB-15</td>
<td>xxt1xxt2\n xxt1xxt2xxt5</td>
</tr>
<tr>
<td>XXT5mut1:pEarley-201</td>
<td>xxt5</td>
</tr>
<tr>
<td>XXT5mut2:pEarley-201</td>
<td>xxt5</td>
</tr>
<tr>
<td>XXT5mut12:pEarley-201</td>
<td>xxt5</td>
</tr>
</tbody>
</table>

Table 3.1 *A. thaliana* XXTmut:xxt5 transformation table. Table depicting the transformation scheme of XXT2mut(1, 2, 12):pGWB-15 and XXT5mut(1, 2, 12) to xxt knock-out *A. thaliana*.

**Root Phenotype Analysis**

XXT5mut:xxt5 T₃ seeds were sterilized as described previously. Seeds were plated on one-half strength Murashinge and Skoog (MS) media with 0.3% Gel-Rite. Once seeds had germinated, the plates were placed at a 45° angle, allowing the roots to grow into the media. Pictures of 10-day old roots were taken using a Leica DMIRE2 light microscope with a Retiga 1300 camera.

**Genomic DNA (gDNA) Extraction and Analysis**

**gDNA Extraction**

Genomic DNA (gDNA) was extracted from rosette leaves of 3-week old T₂ XXT5mut:xxt5 plants. 100mg of leaf tissue was ground in liquid nitrogen then suspended in 700µL CTAB solution (100mM Tris, pH 8.0, 20mM EDTA, 2% NaCl (%w/v), 0.5% cetyl trimethylammonium bromide (CTAB, %w/v), 0.5% polyvinylpyrrolidone (PVP, %w/v)) supplemented with 2% β-mercaptoethanol (%v/v) and incubated at 60°C for 30 minutes. 700µL 24:1 chloroform:isoamyl alcohol was added and solution was centrifuged at 14,000xg for 10 min. The aqueous layer was removed and placed in a new 1.5mL tube and 700µL isopropanol was added. After centrifugation at 14,000xg for 10 min, the supernatant was removed and 700µL
70% ethanol (%v/v) was added to wash the remaining pellet. After centrifugation at 14,000xg for 10 min, the pellet was suspended in 500µL TE buffer (10mM Tris, pH 8.0, 1mM EDTA), 50µL 3M NaCl, and 500µL isopropanol. After centrifugation, 14,000xg for 10min, pellet was washed for a final time with 100µL 70% ethanol (%v/v). The final pellet was dried at 50°C for 5 min then suspended in 40µL water. DNA concentration was determined using a NanoDrop 2000 UV-Vis Spectrophotometer (ThermoScientific).

**PCR Analysis**

For genotyping analysis of XXT5mut:xxt5 T2 plants, 100ng of extracted gDNA was used as a template. PCR-analysis was done using GoTaq Green Master Mix (Promega) according to manufacturer’s instructions. To determine insertion of HA-XXT5mut construct, 35S-F and XXT5-222-R primers were used; to determine T-DNA insertion in XXT5, p5e-F and LB-TDNA-R were used (see Figure 3.1 and Appendix C). PCR was done as described previously with 56°C annealing temperature, 30 second extension time, and 25 cycles. PCR products were analyzed on a 1% agarose containing ethidium bromide.

**Total Membrane Protein Extraction from Plants**

Total membrane fractions were extracted from XXT5mut:xxt5 10-day old T2 seedlings. 1g of plant material was ground in liquid nitrogen, suspended in 5mL of 4°C extraction buffer (40mM HEPES, 500mM sucrose, 1mM EDTA, 2mM MgCl2, 2mM KCl, pH 7.0), and homogenized using a Polytron homogenizer, three times, 10 seconds each time at 25,000rpm. The solution was then filtered through three layers of miracloth and centrifuged 30min at 10,000rpm. The resulting supernatant was then centrifuged 45 min at 38,000xg to pellet all membranes. The resulting pellet was suspended in 40µL protein loading buffer (50mM Tris, pH 6.8, 10% glycerol (%v/v), 0.1% bromophenol blue (%v/v), 2% SDS (%w/v). 30µL of each
sample was analyzed using SDS-PAGE followed by western blotting with Anti-HA AB (ThermoScientific) as described previously.

**Cell Wall Analysis**

**Cell Wall Extraction**

Cell wall alcohol-insoluble residues were extracted from XXT5mut: xxt5, xxt5, and Col-0 T2 21-day old above-ground plants, excluding the roots. Plant tissue was ground in liquid nitrogen, suspended in 10mL 100% ethanol, and homogenized with a Polytron homogenizer 5 min at 25,000rpm. Homogenized cell wall was then incubated at 80°C for 1 hour, homogenized again (5min at 25,000rpm), and centrifuged 30 min at 4100rpm. Pellet was suspended with 10mL 80% ethanol (%v/v) and incubated at 80°C for 1 hour. Cell wall solution was centrifuged (4100rpm for 30 min) and washed three times with 80% acetone (%v/v). Pellet was washed with 100% acetone then allowed to dry at 50°C for 48 hours.

**Driselase Digestion and HPAEC-PAC Detection**

Driselase was partially purified as described in Fry 1988 (182). 1 mg of previously extracted AIR was incubated with 0.5% Driselase (%w/v) prepared in 20mM ammonium formate, pH 5.0 for a total volume of 100µL. Samples were incubated at 37°C with shaking for 18 hours. Samples were then centrifuged and supernatant was analyzed with High Performance Anion Exchange Chromatography with Pulsed Pulsed Amperometric Detection (HPAEC-PAD) using an ICS Dionex System (www.dionex.com). 20µL of each sample were separated on a CarboPac PA-20 column using the following gradient: 0-20min 125mM NaOH, 20-30min increase NaOH to 300mM, 30-45min column regeneration with 125mM NaOH.
CHAPTER 4. RESULTS

To fully characterize the role of DXD motifs in the catalytic function of XXT2 and XXT5, a combination of *in vivo* and *in vitro* techniques combined with mutagenesis were employed. XXT2 and XXT5 display 69.95% sequence identity (see Figure 4.1) and they are highly identical in the C-terminal catalytic domain as compared to the N-terminal transmembrane and stem region domains (183). HMMTOP protein topology software (www.enzim.hu/hmmtop/) was used to predict the transmembrane region of XXT2 and XXT5 (see Figure 4.1). XXT2 contains three DXD motifs, D50LD52, D126WD128, and D228SD230. D50LD52 is located within the putative stem region (54), just past the transmembrane domain (see Figure 4.1). XXT5 has four DXD motifs; D99SD101, D104LD106, D127WD129, and D228SD230 with D99SD101 and D104LD106 located in the putative stem region (see Figure 4.1). The stem region is not believed to be involved in catalysis (1, 10, 54), so the DXD motifs located within were left unchanged. Both XXT2 D126W128 and XXT5 D127W129 are DWD motifs and termed here as mut1, and XXT2 or XXT5 D228SD203 are DSD motifs and termed here as mut2. In the XXTmut1 constructs, both the Asp in DWD were changed to Ala, in the XXTmut2 constructs the Asp in DSD were changed to Ala, and in XXTmut12, both sets of Asp have been changed to Ala (see Figure 4.1). All DXD to AXA point mutants were generated for XXT2 and XXT5, their N-terminal truncated versions were recombinantly expressed in *E. coli*, and full-length versions were stably transformed to xxt mutant *A. thaliana* for phenotype analysis.
Figure 4.1. Amino acid alignment of XXT2 and XXT5. Amino acid sequence alignment of XXT2 and XXT5 using CLC Sequence Viewer 6.5.3 (CLC Bio, www.clcbio.com). Numbers on top and right side designate approximate amino acid position. XXT2 and XXT5 display 69.95% sequence identity. The N-terminal transmembrane domain was predicted using HMMTOP protein topology software (www.enzim.hu/hmmtop/). Key designates identity of underlined DXD motifs and N-terminal truncation mutants.

E. coli Protein Expression

tXXT Protein E. coli Expression Construction

XXT2 and XXT5 mut1 and mut2 point mutations were made by first amplifying the mutation of interest with two overlapping fragments (see Figure 3.1) which were then denatured and self-annealed, forming the template for the final amplification of the full-length XXTmut fragment. XXT2mut12 and XXT5mut12, containing both mut1 and mut2 point mutations, were made using XXTmut1 as the template and amplified with mut2 primers (see Appendix C for primer sequences).

To aid with solubility in the E. coli recombinant protein expression system, N-terminal putative transmembrane domain truncation mutants of XXT2, XXT2mut, XXT2mut2, XXT2mut12 (ΔM1-G41), XXT5, XXT5mut1, XXT5mut2, and XXT5mut12 (ΔM1-G71) were amplified and
cloned into the *E. coli* expression vector pET-15b containing an N-terminal 6x-His or Myc tag to produce tXXT2-His and tXXT5-Myc constructs, respectively (see Figure 4.3 A).

**Optimization of tXXT2-His Protein Expression in *E. coli***

To optimize the recombinant expression conditions of tXXT2-His:pET-15b in *E. coli*, a systematic approach was taken. First, five different cell lines were tested; BL21(DE3), BL21(DE3) codon plus, C43(DE3), Rosetta 2(DE3), and ArcticExpress(DE3) codon plus (see Figure 4.2). For each trial, 5mL of each *E. coli* cell line were grown, induced, harvested, and lysed as described in *Materials and Methods*. Crude lysate and insoluble pellet fractions were analyzed by SDS-PAGE and western blotting using Anti-6xHis AB. Band intensity analysis was done using ChemiDoc XRS+ System (BioRad, 170-8265). The results show that all tested cell lines expressed tXXT2-His at a high level. BL21(DE3) codon plus and C43(DE3) had the highest amount of protein in the soluble lysate. BL21(DE3) codon plus expressed tXXT2-His at a slightly higher level overall, and was thus chosen for all future experiments.

![Figure 4.2 tXXT2-His expression level in different cell lines](image)

*Figure 4.2 tXXT2-His expression level in different cell lines.* SDS-PAGE with Anti-His AB western blot analysis of tXXT2-His crude lysates and insoluble pellets protein extraction samples from the *E. coli* cell lines BL21(DE3), BL21(DE3) codon plus, C43(DE3), Rosetta 2 (DE3), and ArcticExpress (DE3) codon plus. Samples were induced with 0.5mM IPTG and grown for 18 hours at 37°C. 20uL of each sample was loaded onto 1.0mm, 10% SDS-PAGE and run at 80V for 30min, 110V 120min. Ladder describes approximate mass in kDa (Spectra Multicolor Broad Range Protein Ladder, Pierce)

To further optimize recombinant tXXT2-His protein expression, different growth conditions were tested. Three growth times following expression: 4 hours, 8 hours, 18 hours,
three temperatures of growth after induction: 37°C, 28°C, and 16°C, and four [IPTG] for induction: 1mM, 0.7mM, 0.5mM, and 0.25mM were analyzed for optimal expression (see Table 4.1). Soluble lysates and insoluble pellet fractions of each growth time/growth temperature/[IPTG] combination were analyzed by SDS-PAGE followed by western blotting with Anti-His AB. Band intensity analysis was performed using ChemiDoc XRS+ Imager (BioRad). A scale of relative intensity was used to rank the expression level of each set of conditions; 100-80% intensity = 4, 79-50% = 3, 49-30% = 2, 29-10% = 1, 9-0% = 0 (see Table 4.1). The results show that growth at 28°C for 8 hours or longer is the highest and independent of the [IPTG] used to induce. Growth at 37°C did not show expression in the soluble fraction, at all tested [IPTG] and growth times. The highest amount of soluble lysate was detected when grown at 28°C for 18 hours and inducing with 0.5mM IPTG. Thus, for all expression experiments, these conditions were used.
### Table 4.1 Optimization of tXXT2-His protein expression in BL21(DE3) codon plus E. coli.

Table of relative expression tXXT2-His protein content in insoluble pellets and soluble lysates using three growth times following expression: 4 hours, 8 hours, 18 hours, three temperatures of growth after induction: 37°C, 28°C, and 16°C, and four [IPTG] for induction: 1mM, 0.7mM, 0.5mM, and 0.25mM. A scale of relative intensity was used to rank the expression level of each set of conditions; 100-80% intensity = 4, 79-50% = 3, 49-30% = 2, 29-10% = 1, 9-0% = 0. Relative intensity was determined using ChemiDoc XRS+ Imager (BioRad).

<table>
<thead>
<tr>
<th>Growth Time Following Induction (Hours)</th>
<th>[IPTG] (mM)</th>
<th>37</th>
<th>28</th>
<th>16</th>
<th>37</th>
<th>28</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Optimization of tXXT2-His Lysis and Extraction from E. coli**

To confirm that optimal amount of tXXT2-His was being extracted from the soluble lysate, different lysis buffer conditions were examined; [NaCl]: 1M, 500mM, 300mM, and type of buffer and pH: 20mM MES pH 5.0, 20mM HEPES pH 7.4, 20mM Tris pH 8.8. tXXT2-His:pET-15b was expressed in BL21(DE3) codon plus E. coli as described previously for optimal soluble expression. Soluble lysates and insoluble pellets were analyzed as described previously with western blotting with Anti-His AB. The same scale of relative intensity as described previously was used to rank the amount of soluble tXXT2-His protein in each set of
lysis conditions (see Table 4.2). Relative intensity was determined using ChemiDoc XRS+ Imager (BioRad). The results show that 300mM NaCl, 20mM HEPES, pH 7.4 gave the highest amount of tXXT2-His protein. The other examined conditions showed a slightly less amount of tXXT2-His protein.

<table>
<thead>
<tr>
<th>Buffer/pH</th>
<th>[NaCl]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M MES/5.0</td>
<td>3</td>
</tr>
<tr>
<td>500mM HEPES/7.4</td>
<td>3</td>
</tr>
<tr>
<td>300mM HEPES/7.4</td>
<td>4</td>
</tr>
<tr>
<td>20mM Tris/8.8</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4.2 Optimization of tXXT2-His protein extraction and lysis conditions. Table representing the relative amount of tXXT2-His protein present in BL21(DE3) codon plus E. coli crude cell lysates. Lysates were prepared using different buffer/pH: 20mM MES pH 5.0, 20mM HEPES pH 7.4, 20mM Tris pH 8.8, and different [NaCl]: 1M, 500mM, 300mM. Relative amount of protein was determined by band intensity analysis of western blots using Anti-His AB using ChemiDoc XRS+ Imager (BioRad). Scale is as follows: 100-80% intensity = 4, 79-50% = 3, 49-30% = 2, 29-10% = 1, 9-0%.

Mutated tXXT Recombinant Expression

All final tXX75, tXX75mut, tXX72, and tXXT2mut constructs were transformed to BL21(DE3) codon plus E. coli cells. Protein expression and crude extraction was performed using optimal conditions described previously. Western blotting and coomassie staining were used to determine the presence of the expressed truncated proteins as well as their relative expression level (see Figure 4.3).
Figure 4.3 *E. coli* XXT protein expression. SDS-PAGE analysis of tXXT2-His (tXXT2), tXXT2mut1-His (mut1), tXXT2mut2-His (mut2), tXXT2mut12-His (mut12), tXXT5-Myc, tXXT5mut1-Myc, tXXT5mut2-Myc, and tXXT5mut12-Myc crude lysates and insoluble pellets protein extraction samples from BL21(DE3) codon plus *E. coli* induced with 0.5mM IPTG and grown for 18 hours at 28°C. 20uL of each sample was loaded onto 1.0mm, 10% SDS-PAGE and run at 80V for 30min, 110V 120min. (A) Western blot analysis was done on nitrocellulose membrane, XXT2-His variants analyzed with Anti-His (Santa Cruz) and XXT5 variants with Anti-Myc (ThermoScientific). (B) Coomassie stain analysis. Ladder describes approximate mass in kDa (Spectra Multicolor Broad Range Protein Ladder, Pierce). tXXT2-His has a predicted molecular weight of 48kDa and tXXT5-Myc has a predicted molecular weight of 45kDa.

All tXXT2-His proteins are expressed at a high level, however only approximately 10% is present in the soluble crude lysate compared with the insoluble pellet, based on western blot band intensity analysis using ChemiDoc Imager (Bio-Rad). tXXT5-Myc is also expressed at a high level, and again only 10% is present as soluble. tXXT5mut (1, 2, 12)-Myc are not expressed at high level (approximately 10% that of the tXXT5-Myc), however 40% of that protein is soluble.

**Catalytic Activity of Truncated and Mutated Recombinant XXT2**

To test whether the *E. coli* recombinant expression and N-terminal truncation would affect the D-xylosyltransferase activity of XXT2, XT activity assays were performed using tXXT2-His crude *E. coli* cell lysates. The assay was performed according to Cavalier and
Keegstra 2006 (144) as described in Materials and Methods and product formation was analyzed using Matrix-Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS) using Applied BioSystems Voyager System 6075. BL21(DE3) codon plus crude cell lysate, was used as a negative control.

<table>
<thead>
<tr>
<th>Fraction</th>
<th>Mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGGGGG + Na+</td>
<td>1013.7</td>
</tr>
<tr>
<td>GGXGGG + Na+</td>
<td>1145.8</td>
</tr>
<tr>
<td>GGXXGG + Na+</td>
<td>1277.9</td>
</tr>
<tr>
<td>GGXXXG + Na+</td>
<td>1410.0</td>
</tr>
</tbody>
</table>

Table 4.3 Masses of expected products of XT activity assay. Masses of expected products of XT activity assay using cellohexose (GGGGGG) as an acceptor and UDP-xylose as a donor. G designates β1-4 linked D-glucose, X designates α1-6 linked D-xylose onto D-glucose.
Figure 4.4 MALDI-TOF analysis of tXXT2 activity assay. Matrix-Assisted Laser Desorption/Ionization (MALDI)-Time of Flight (TOF) Mass Spectrometry (MS) spectra for XT activity assay of (A) BL21(DE3) codon plus, tXXT2-His and (B) tXXT2mut-His soluble protein from crude *E. coli* cell lysates using cellohexose as an acceptor. Samples were mixed 2:1 with 2,5-dihydroxybenzoic acid (DHB) 0.5mg/mL matrix and applied to MALDI-TOF plate (SOURCE). Samples were collected in reflector mode, delayed extraction for 350nsec, positive polarity, 20000V accelerating voltage, and 200 shots per spectrum. Expected masses can be seen in Table 4.3.
The results show that tXXT2-His is active and able to add one D-Xyl onto the acceptor cellohexose, forming the product GGXGGG (see Figure 4.4 A, Table 4.3). The XT activity assay was also done on tXXT2mut1-His, tXXT2mut2-His, and tXXT2mut12-His E. coli crude lysates to test if the DWD (mut1) or DSD (mut2) were involved in catalysis. None of the DXD mutated variants were able to form GGXGGG (see Figure 4.4 B, Table 4.3), demonstrating that both DWD and DSD motifs are essential for XXT2 enzymatic activity.

**tXXT2 Protein Isolation using Native and Hybrid Methods**

Protein isolation of tXXT2-His was done using Ni-NTA resin (ThermoScientific) applying both native (non-denaturing) and hybrid (combination of denaturing and native conditions) protocols as described in *Materials and Methods*. The native protocol used the soluble portion of tXXT2-His and involved no denaturing steps. In the hybrid protocol, the pellet was solubilized by denaturing in guanidine-HCl. The guanidine-soluble fraction was added to the Ni-NTA resin (ThermoScientific), and then allowed to re-fold by slow dilution with non-denaturing buffer, finally being eluted in non-denaturing buffer. For both protocols, 20µL of each sample was analyzed by SDS-PAGE followed by coomassie staining (see Figure 4.5 and Appendixes 1, 2).
Figure 4.5 Native and hybrid tXXT2-His isolation. SDS-PAGE analysis of tXXT2-His isolated with the (A) Native (non-denaturing) and (B) Hybrid (combination of denaturing and native) protocols. 20µL of each sample was loaded onto 1.0mm, 10% SDS-PAGE and run at 80V for 30 min, 110V for 120min, and coomassie stained. All images were taken with ChemiDoc Imager (BioRad). Ladder describes approximate mass in kDa (Spectra Multicolor Borad Range Protein Ladder, Pierce). The expected molecular masses of tXXT2-His and tXXT5-Myc are 48kDa and 46kDa, respectively.

Isolated tXXT2-His present in elution fractions obtained from both protocols showed one band stained by coomassie R-250 staining (see Figure 4.5). The protein concentration of each sample was determined using the bicinchoninic acid assay (BCA, ThermoScientific) and percent yield was calculated. The percent yield of the hybrid protocol was much higher, 23.7% compared to 5.1%, than that of the native protocol. This is due to the fact that there is more tXXT2-His protein in the solubilized guanidine fraction than in the crude lysate, thus it had less competition for binding onto the Ni-NTA resin as compared with the crude lysate.

XT activity assays were performed on Ni-NTA purified elution fractions to demonstrate that tXXT2-His isolated from native and hybrid protocols were catalytically active. XT assays were performed on elution fractions, dialyzed against buffer not containing imidazole, as described Cavalier and Keegstra 2006 (144) and in Materials and Methods. Samples were analyzed using MALDI-TOF MS (Applied BioSystems Voyager System 6075, see Figure 4.6).
Untransformed BL21(DE3) codon plus lysate was used as a negative control and tXXT2-His was used as a positive control (see Figure 4.5).

**Figure 4.5** MALDI-TOF analysis of tXXT2-His isolated fractions. Matrix-Assisted Laser Desorption/Ionization (MALDI)-Time of Flight (TOF) Mass Spectrometry (MS) spectra for XT activity assay of tXXT2-His isolated from Native and Hybrid protocols. Samples were mixed 2:1 with 2,5-dihydroxybenzoic acid (DHB) 0.5mg/mL matrix and applied to MALDI-TOF plate (SOURCE). Samples were collected in reflector mode, delayed extraction for 350nsec, positive polarity, 20000V accelerating voltage, and 200 shots per spectrum. Expected masses can be seen in Table 4.3.

The tXXT2-His native purified sample demonstrated positive D-xylosyltransferase activity, showing that the Ni-NTA purification protocol does not have an effect on XXT2 activity. tXXT2-His hybrid purified also demonstrated positive activity, demonstrating that tXXT2-His can be denatured and solubilized from inclusion bodies and, when allowed to slowly re-fold by dilution, can restore native XXT2 catalytic activity.
XXTmut Protein Expression in *A. thaliana xxt* Mutant Plants

**Transformation and Selection**

Full-length *XXT2mut*(1, 2, 12) and *XXT5mut*(1, 2, 12) were cloned into the binary vectors pGWB-15 (180) and pEarley-201 (Arabidopsis Biological Resource Center), respectively. *XXT2mut:pGWB-15* constructs were transformed to *xxt1xxt2* and *xxt1xxt2xxt5* *A. thaliana* knock-out mutants and *XXT5mut:pEarley-201* were transformed to the *xxt5* mutant. All transformations were done via the flora dip method using *Agrobacterium tumefaciens*-mediated transformation according to Zhang et. al 2006 (181). Transformed T₀ *XXT2mut*(1, 2, 12) seeds were collected and transformants were selected using Hygromycin resistance. Transformed T₀ *XXT5mut*(1, 2, 12) seeds were collected and transformants selected using Basta resistance. Seeds from selected T₁ plants were collected and selected using Basta a second time, generating T₂ plants.

**Genomic DNA (gDNA) Analysis**

gDNA analysis of selected HA-*XXT5mut*(1, 2, 12):*xxt5* T₂ plants was done to confirm that first, the plants contain the transformed HA-*XXT5mut*(1, 2, 12) gene of interest and second, the transformed plants are in the *xxt5* knock-out mutant background. gDNA was extracted from rosette leaves as described in *Materials and Methods* and subjected to PCR-analysis (see Figure 3.1 and Figure 4.7). At least four independent lines of each construct and a minimum of four plants from each of those lines were analyzed. Three independent lines of each construct were randomly selected for further analysis.
Figure 4.7 Genomic DNA analysis of XXT5mut(1, 2, 12):xxt5 T2 plants. gDNA was extracted from rosette leaves of Col0, xxt5, XXT5mut1:xxt5, XXT5mut2:xxt5, and XXT5mut12:xxt5 plants as described previously. 100ng of gDNA was used for PCR-analysis using GoTaq Green (Promega) according to manufacturer’s instructions and (A) 35S-F, XXT5-222-R or (B) p5e-F, LB-tDNA-R primers. PCR was done as described previously for 25 cycles before analysis on ethidium bromide-containing 1% agarose gels. Three independent lines (listed below the construct name) of each construct and three plants from each line (listed below the line name) were analyzed. Depicted ladder is 1Kb plus DNA Ladder (Invitrogen).

The results show that all analyzed plants from every independent transgenic line (a minimum of four lines in each construct) of HA-XXT5mut(1, 2, 12):xxt5 are stably transformed and their gDNA contains the HA-XXT5mut(1, 2, 12) gene of interest. These plants also contain the T-DNA insertion within the native XXT5 gene, confirming that the HA-XXT5mut(1, 2, 12) is being expressed in the xxt5 mutant background.

Root Phenotype of Selected XXT5mut:xxt5 T3 Plants

XXT5mut(1, 2, 12):xxt5 T3 seeds were plated on 1/2 MS media and grown vertically for 10-days. Roots were observed using light microscopy (Leica DMI RE2 microscope with Retiga...
1300 camera) as described previously (9, 146) and in Materials and Methods. Col0 and xxt5 seeds were used as controls to compare the phenotype (see Figure 4.8)

![Image](image-url)

**Figure 4.8 Root phenotype of XXT5mut(1, 2, 12):xxt5 T3 plants.** Root pictures taken of Col0, xxt5, XXT5mut1::xxt5, XXT5mut2::xxt5, and XXT5mut12::xxt5 10-day old T3 seedlings. All seedlings were grown in same conditions on 1/2 MS media, 0.3% GelRite plates as described previously. Once seeds had germinated, plates were tilted 45° so seedlings were growing vertically. All images were taken with (Leica DMIRE2 light microscope with Retiga 1300 camera). Labels indicate the plant line show and black bar is equivalent to 200µm in all images. Black boxes in (A) depict enlarged area shown in (B). Images in (A) were taken at 10X magnification and the same image was taken at 20X magnification (B) with isolated root shown to clearly depict root hair phenotype. Three lines of each construct and a minimum of three plants from each line were analyzed for phenotype.

As described in Zabotina et. al 2008 (146), xxt5 plants display a mutant root hair phenotype characterized by shortened hairs that have bulb-like protrusions. HA-XXT5mut1::xxt5 was able to complement this phenotype, resulting in Col0-like morphology root hairs. HA-XXT5mut2::xxt5 and HA-XXT5mut12::xxt5, however, did not complement this phenotype. These results demonstrate that mutation of the second DXD motif, DSD, is essential for XXT5 function and mutation in the first DXD motif, DWD, is not.
Total Membrane Protein Analysis in Selected XXT5mut:xxt5 T2 Plants

To test whether the HA-XXT5mut(1, 2, 12) protein is being expressed in the selected transgenic plants, total membrane proteins were extracted from 0.5mg of rosette leaves from T2 plants as described in Materials and Methods. The pellet membrane proteins were resuspended in 1xSDS-PAGE loading buffer and analyzed by SDS-PAGE followed by western blotting using anti-HA AB (see Figure 4.9). HA-XXT5 complemented xxt5 was used as a positive control (146) and Col-0 was a negative control. Three independent lines of each construct were analyzed.

![Figure 4.9 Protein analysis of HA-XXT5mut(1, 2, 12):xxt5 T2 plants.](image)

The results show that HA-XXT5mut(1, 2, 12) is present in all lines, confirming that the complementation of xxt5 by XXT5mut1 is due to a presence of XXT5mut1 within the cell. The lack of complementation seen with XXT5mut2 and XXT5mut12 is not due to the absence of expressed protein in the cells, further confirming that an intact DSD motif is required for XXT to function and complement the root phenotype in xxt5.
Cell Wall Analysis of Selected XXT5mut:xxt5 T2 Plants

To further characterize the effect of XXT5mut(1, 2, 12) over-expression in xxt5 A. thaliana plants, alcohol-insoluble crude cell wall was extracted from T2 plants as described in Materials and Methods. xxt5 and Col0 A. thaliana alcohol-insoluble residues (AIRs) were also extracted and used for comparison.

Amount of XyG was determined by digesting AIRs with Driselase, an enzyme mixture that lacks an α(1-6)-D-xylosidase. A signature disaccharide, termed isoprimervverose (IP), β-D-glucose-α(1-6)-D-xylose, is released from XyG and can be quantitated using High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD). The results show that xxt5 has a significant reduction in XyG compared to Col-0 (p-value <0.05, 2-sided T-test), which is what was seen previously ((146) see Figure 4.10). XXT5mut1:xxt5 does not have a significant reduction in XyG compared to Col-0, which confirms that XXT5mut1:xxt5 fully complements the xxt5 phenotype. XXT5mut2:xxt5 and XXT5mut12:xxt5 also do not have a significant reduction in XyG, which is contradictory to what is seen with the root phenotype.
Figure 4.10 Cell wall analysis of HA-XXT5mut(1, 2, 12):xxt5 T2 plants. Analysis of xyloglucan content using Driselase digestion. Alcohol-insoluble residue crude cell wall was extracted from Col-0, xxt5, XXT5mut1:xxt5, XXT5mut2:xxt5, and XXT5mut12:xxt5 T2 plants and subjected Drislease digestion. 1mg cell wall was incubated with 0.5% Driselase and incubated at 37°C as described in Materials and Methods. XyG signature, isoprimerverose (IP) peak area in the soluble fraction following digestion was determined using high performance anion exchange chromatography with Pulsed Amperometric Detection (HPAEC-PAD) as described in Zabotina et. al 2008 (146). Stars indicate significant difference (p-value < 0.05, 2-sided T-test) from Col-0. Error bars indicated standard deviations. Three independent lines were analyzed for each construct and data from all lines was combined and averaged. Three technical replicates were performed for each sample.
CHAPTER 5: DISCUSSION

Optimization of tXXT2-His Expression and Extraction Conditions

There are many different systems currently available for the recombinant expression of proteins. The most popular, due to the fact that it is quick, relatively cheap, and well-characterized, is recombinant over-expression in *E. coli*. Though *E. coli* is capable of expressing large amounts of protein, up to 50% of total cell protein (184), in most circumstances protein is expressed as insoluble inclusion bodies; misfolded or unfolded protein aggregates (185). Many *E. coli* cell lines have been generated which are optimized for different expression conditions. The most common is the BL21(DE3) cell lines, which contain a plasmid allowing them to over-express T7 RNA polymerase (186). Modified versions of this cell line include BL21(DE3) codon plus, which contain *argU, ileY*, and *leuW* tRNA genes (186), Rosetta2(DE3), which also contains tRNA genes (187), C43(DE3), which express subunits b and c of the *E. coli* F-ATPase, an alanine-proton symporter, and the ADP/ATP/phosphate carriers from mitochondria, which assists in the expression of toxic proteins and membrane proteins (188, 189), and ArcticExpress(DE3) codon plus, which co-express Cpn60 and Cpn10 chaperones to aid in the stability of proteins at low temperatures (190).

There is a large variety of expression conditions and extraction techniques available, and it is essential to optimize for the protein of interest, as each protein will act differently in different circumstances. Factors that can influence protein expression are growth temperature, concentration of inducer, growth time after induction, pH of the growth media, and salinity of the growth media (191). Extraction conditions, such as pH, salinity, and detergents can also affect the amount of soluble protein by stabilizing proteins and preventing aggregation (192).
The XXTs have been recombinantly expressed in *Spodoptera frugiperda* 21(Sf21) insect cells (145), *Pichia pastoris* (143), and BL21(DE3) *E. coli* cells as a N-terminal truncation mutants (XXT2ΔM1-Q49, XXT5ΔM1-L72) and glutathionie-S-transferase (GST)-tag proteins (147). No work has been done, however, in the optimization or purification of the XXTs, and there are no available examples of the purification of glycosyltransferases involved in cell wall biosynthesis. It was therefore essential to establish and optimize the expression of N-terminal truncation XXTs to effectively study them *in vitro*.

In this study, an N-terminal truncation mutant of XXT2 (ΔM1-G41) was cloned into pET-15b and optimized for expression and extraction in *E. coli*. The optimal expression was determined to be the following; BL21(DE3) codon plus cell line, 0.5mM IPTG, grow at 28°C for 18 hours after induction, and use 20mM HEPES, 300mM NaCl, pH 7.4 as the protein extraction buffer. These optimal conditions were used for all *E. coli* expression experiments, and will be used for future XXT purification and structural characterization studies.

**tXXT2 Isolated in Native or Hybrid Conditions Demonstrates XXT Activity**

As stated previously, a common problem associated with the recombinant expression of proteins in *E. coli* that the protein of interest is expressed as aggregated insoluble inclusion bodies (193). Expression in inclusion bodies is not always a undesirable, however, as the expressed protein of interest comprises >50% of protein within the inclusion bodies (194), which can easily be isolated by centrifugation and solubilized with chaotrophic salts such as guanidine-HCl or urea (185). Once the protein of interest is denatured and solubilized, there are many different methods to remove the denaturing agent and re-fold it to its native state, such as dialysis or rapid dilution (194). These methods, however, can often induce protein aggregation (195), produce proteins that are not folded in a native conformation, or produce a mixture of different
protein folds (196). For example, proper folding of recombinantly expressed tissue plasminogen activator (tPA), a protein used to prevent bleeding clotting, requires the presence of the *E. coli* cysteine oxidoreductase DsbC to generate the native disulfide bonds and the correct protein fold (197). In another example, recombinant human growth hormone (r-hGH) requires a low concentration of denaturant (2M) urea, which does not disrupt all secondary structures, for correct refolding and catalytic function (198). Due to the fact that there is a wide variety of techniques available, it is essential to optimize and establish a denaturing/refolding protocol for each recombinantly expressed protein of interest (199). Additionally, once the protein of interest has been isolated and re-folded, native folding and functional assays need to be performed, including activity assays and circular dichroism to study secondary structures (200).

In this study, recombinantly expressed tXXT2-His (ΔM1-G41) protein was isolated utilizing two different methods. The first, termed the Native protocol, utilized the soluble crude lysate and purified on Ni-NTA using non-denaturing buffers. The second, termed the hybrid protocol, first denatured the insoluble *E. coli* pellet, bound it to Ni-NTA resin, and slowly re-folded it via slow dilution. To confirm that the isolated proteins were catalytically active, they were assayed with the previously described XT assay. tXXT2-His isolated from native conditions produced the GGXGGG product, demonstrating that it retains catalytic activity through the affinity chromatography purification process. tXXT2-His isolated from the hybrid protocol also demonstrated XT activity, demonstrating that the re-folded variant present in the elution fractions is able to perform the native XXT2 catalytic function (see Figure 4.5). This developed protocol will be used in future structural studies of XXT2 and XXT5.
Molecular Modeling of XXT2 and XXT5 Predicts Catalytic Center

Currently, there are no available structures for glycosyltransferases involved in the biosynthesis of the plant cell wall (Protein Databank, www.rcsb.org/pdb). It is therefore difficult to predict structural features of these proteins such as protein-protein interaction surfaces, substrate-binding sites, and catalytic pockets. Much progress has been made in generating software which will predict an unknown protein’s structure based on homology with available structural data.

Molecular models of XXT2 and XXT5 were made using YASARA modelling software (Yet Another Scientific Artificial Reality Application, www.yasara.org, (201)) by Dr. Aleisa Tietze (see Figure 5.1). The template used to generate the models was A64R (PDB ID 2P6W), a putative UDP-D-glucosyltransferase of the CAZy GT34 family from Paramecium bursaria chlorella virus-1 (PBCV-1, (202)) that has 23.95% sequence identity to XXT2. In this model, the N-terminal 110 residues (M1-E110) and C-terminal XXT2 L445-V461 and XXT5 N404-N457 residues were removed due to the fact that they are predicted to be random loops, and do not fit with the template.

The location of the studied DXD motifs, mut1 (XXT2 D126WD128, XXT5 D127WD129) and mut2 (XXT2 and XXT5 D228SD230) was first examined. As can be seen from Figure 5.1A in green, the DWD motif, mut1, is located on a loop outside of the putative globular catalytic center. The DSD motif, mut2 (red), however, is located within this putative catalytic center and is coordinated with Mn$^{2+}$ (magenta). Figure 5.1B shows a close-up of this domain, with other amino acids that are within the distance to be involved in XXT catalysis (highlighted in yellow and blue). In both XXT2 and XXT5, localization of His378 (blue) predicts its participation in catalysis as the catalytic base in the retaining glycosyltransferase mechanism.
of the XXTs (11, 144). Other charged or polar residues, D319, E200, E257, D355, Y357, C382 and N269 in XXT2 and Y204, D257, S158, and D355 in XXT5 (yellow), are within the distance of the catalytic site and might be involved in stabilizing intermediates or transition states (11).

Additionally, a putative UDP-D-Xyl binding pocket can be seen in XXT2 and XXT5 (see Figure 5.1C). For both XXT2 and XXT5, the binding pocket is cylindrical-shaped and goes through the protein (depicted with arrows in Figure 5.1C), with the catalytic center in the middle. In the XXT2 model, this binding pocket is a narrow, tunnel-shape that is small enough for the donor or acceptor to enter. In the XXT2 model, the putative catalytic center is hidden from the protein surface and resides in a small pocket within the putative binding tunnel. The XXT5 model, however, has a much larger binding pocket that is cleft-shaped. This putative catalytic center, which includes the studied DSD motif, is more accessible to the outside of the protein and could be involved in protein-protein interactions.
**Figure 5.1. Molecular models of XXT2 and XXT5.** Molecular models of XXT2Δ1-110, 445-461 and XXT5Δ1-110, 404-457 were made using YASARA software (www.yasara.org) by Dr. Alesia Tietze. Template used was A64R, a putative D-glucosyltransferase of the CAZy GT34 family from Paramecium bursaria chlorella virus-1 (PBCV-1, (202)). (A) depicts the entire model, with the DXD motifs (DWD, mut1, in green, DSD, mut2, in red), putative catalytic His378 (blue) and Mn$^{2+}$ (magenta) highlighted. (B) depicts the putative catalytic domain, with the DSD motif (red), His 378 (blue), Mn$^{2+}$ (magenta), and amino acids within range of the catalytic center to be possibly involved in the catalytic mechanism (yellow). Amino acid names and positions are labeled. (C) depicts a space-filling representation of the XXT2 and XXT5 models outside of the catalytic domain. Colors represent amino acids highlighted in (B). Arrow indicates possible substrate-binding pocket through the catalytic center. The dashed part of the arrow indicates portion of the binding pocket that is inside the protein and not accessible from the outside.
Both DWD and DSD Motifs are Required for XXT2 Catalytic Function

XXTs are predicted to be type II membrane proteins, with a small N-terminal transmembrane domain anchoring a larger C-terminal catalytic domain within the Golgi lumen (54, 183). Approximately 50% of predicted glycosyltransferases contain a conserved DXD motif (11) which acts as a catalytic site, functioning in nucleotide binding and potentially providing the catalytic base necessary for the double displacement mechanism of retaining glycosyltransferases (10, 54). The catalytic mechanism of retaining glycosyltransferases is still largely debated (162, 163), and nothing is known about the catalytic mechanisms of cell wall-synthesizing glycosyltransferases. It is therefore very important to identify the residues involved in catalysis and to characterize their roles both in vitro and in vivo.

XXT2 is a retaining glycosyltransferase (144) that contains three DXD motifs at positions D50LD52, D126WD128, and D228SD230. D50L-D52 is located within the putative stem region, which is not predicted to be involved in catalysis (4) and thus was not examined in this study. The other DXD motifs, D126WD128, here designated mut1, and D228SD230, here designated mut2, were point-mutated to AXA motifs to investigate their role in the catalytic function of XXT2. To aid with protein solubility and recombinant expression, N-terminal transmembrane domain truncation mutants (ΔM1-G41) containing mut1, mut2, both (mut12), or neither AXA mutations were cloned into pET-15b and expressed in BL21(DE3) E. coli. The catalytic activity of these mutants was examined using the XT assay as described by Cavalier et. al 2006 (144) using UDP-xylose as a donor and cellohexose as an acceptor.

It was first shown that tXXT2-His, lacking the cytosolic N-terminal and transmembrane domain, recombinantly expressed in E. coli was active and able to add one D-Xyl to a cellohexose backbone, forming the GGXGGG product (see Figure 4.3. A). Interestingly, D-
xylosylated products with two or three D-Xyl were not observed, contrary to what was seen previously in insect cells (144). This difference is most likely due to a higher XXT2 concentration in insect cells (144) than in the *E. coli* crude lysate. We cannot exclude the effect of potential glycosylation, however. Using EnsembleGly (Iowa State University, http://turing.cs.iastate.edu/EnsembleGly/predict.html) it is predicted that there are two putative N-glycosylation sites in XXT2 at N137 and N431, however N137 is followed by P, so it is highly unlikely to be glycosylated (203). If glycosylation is required for the optimal function of XXT2 in vivo, then recombinant expression in insect cells, which contain the machinery required for protein glycosylation, would produce this activity. *E. coli*, however, do not have the ability to glycosylate proteins, thus recombinantly expressed tXXT2-His could potentially not possess optimal XT activity. Additional studies will be needed in the future to investigate one, if XXT2 is glycosylated or otherwise post-translationally modified in vivo, and two, if the full XT activity of XXT2 is dependent on these potential post-translation modifications.

The same XT assay was performed on tXXT2mut1-His, tXXT2mut2-His, and tXXT2mut12-His. None of the variants demonstrated D-xylosyltransferase activity, concluding that both DXD motifs, mut1 and mut2, are essential for XXT2 catalytic activity. This is contrary to what was predicted from XXT2 models (see Figure 5.1. A), where the mut1 DXD motif (DWD) is located on a loop outside of the putative catalytic globular domain. It should be noted, however, that these loop structures do not fit well with the template, and are thus viewed as randomly oriented on the model. In vivo, it could be that this domain is located near the active site and facilitates substrate binding of either the UDP-D-Xyl donor or β-D-glucan acceptor. These obtained results will be further clarified after the characterization of HA-XXT2mut:xxt1xxt2 and HA-XXT2mut:xxt1xxt2xxt5 transgenic *A. thaliana* plants.
The DSD Motif is Required for XXT5 Function

Very little is known about the *in vivo* biological function of XXT5. XXT5 was identified to be in the same gene family as the known D-xyloslytransferases XXT1 and XXT2 (143). XXT5 is a member of the CAZy GT34 family and is predicted to have the same type II membrane topology and Golgi-localization as XXT1 and XXT2 (146). Using reverse genetics studies, it was shown that complete knock-out of *xxt5* results in a 50% reduction in XyG content of the cell wall and results in a phenotype of short, bulb-shaped root hairs (146). It is unknown what causes this phenotype: the lack of total XyG or a difference in XyG subunit composition. Immunolocalization studies using XyG-specific antibodies on *xxt5* and Col-0 cell walls demonstrated that not all XyG epitopes are reduced in the mutant compared to the wild-type, indicating that not all XyG epitopes are affected by loss of the XXT5 protein (9, 146). The combination of these results led to the hypothesis that XXT5 is a D-xylosyltransferase.

To examine the *in vitro* activity of XXT5, it was recombinantly expressed in both *P. pastoris* and Sf21 insect cells (146). A similar XT assay as the one described in this study was used to test XXT5, however no *in vitro* XT activity was shown. Many different oligosaccharide acceptors were used including GGGGG, XXGG, XXXG, XXGGG, XLGGG, XLGGXLG, GXXGGGXXGG, and others. In all of the trials, incubation of the recombinant XXT5 with UDP-D-Xyl donor and oligosaccharide acceptor did not result in XT activity. It was therefore predicted that XXT5 plays a non-catalytic, most likely structural, biological role in the biosynthesis of XyG. Since XXT5 is essential for full synthesis of XyG, fully characterizing its functional role *in vivo* has become an important area of study.

XXT5 contains four DXD motifs at positions D99SD101, D104LD106, D127WD129, and D228SD230. Two of these motifs, D99SD101 and D104LD106, are located within the
putative stem region, which is not predicted to be involved in catalysis (4), and were not examined in this study. The other two DXD motifs, D127WD129, here designated mut1, and D228SD230, here designated mut2, were point-mutated to AXA motifs to investigate their role in the function of XXT5 in vivo.

Full-length, N-terminal HA-tagged, XXT5mut(1, 2, 12) constructs were stably transformed into xxt5 knock-out A. thaliana (146) plants. Transgenic plants were selected using Basta resistance and the T2 and T3 progeny of three independent lines were characterized. Genotyping analysis confirmed that all selected transgenic plants contained the HA-XXT5mut construct and were in the xxt5 background (see Figure 4.6). When root phenotypes were compared, XXT5mut1:xxt5 fully complemented the xxt5 phenotype, indicating that this DWD motif is not essential for XXT5 functioning. XXT5mut2:xxt5 and XXT5mut12:xxt5, however, do not complement the mutant phenotype. They display shortened, bulb-like root hairs similar to that of xxt5, indicating that this DSD motif is essential for XXT5 functions. Total membrane fractions from the rosette leaves of transgenic plants were extracted and analyzed by Anti-HA western blotting (see Figure 4.8), demonstrating that HA-XXT5mut protein was being expressed. These results confirm that the described root phenotype is due to the presence of HA-XXT5mut protein.

The amount of XyG was analyzed from crude cell wall extractions based on the amount of a signature XyG disaccharide, isoprimerverose (IP), when the cell wall is digested with Driselase (see Figure 4.9). The results show that XXT5mut1:xxt5, XXT5mut2:xxt5, and XXT5mut12:xxt5 do not have a significantly different (p-value<0.05, 2-sided T-Test) amount of IP compared to Col-0. These results indicate that the lack of complementation seen in the root phenotype is not due to total reduction of XyG in these plants as it was shown before in xxt5
mutant plants (146). Most likely, the presence of XXT5 protein, though with mutated DXD motifs, is enough to restore XyG biosynthesis. The question still remains if the synthesized XyG in these complemented plants have wild type structures. Since XXT5mut2:xxt5 and XXT5mut12:xxt5 complemented plants have root phenotype similar to xxt5, it is possible that the structure/composition of XyG in these plants is not complete due to these point mutations.

Obtained results support the prediction made using XXT5 modeling, (see Figure 5.1), which show the DWD motif (green) outside of the putative catalytic center whereas the DSD motif (red) is situated within this center and is predicted to be coordinated with Mn\(^{2+}\) (magenta).

Based on the results of this study, two hypotheses of the function of XXT5 are proposed. First, XXT5 plays a catalytic role, however it requires another interacting protein for full activity. It has been shown that XXT2 and XXT5 strongly interact in the Golgi apparatus (7). In the case of the pectin synthesis GAUT1/GAUT7 complex, GAUT1 requires GAUT7 for proper localization (101). XXT2 and XXT5 could work in a similar fashion; the activity of XXT5 requires the presence of XXT2. The second hypothesis is that the DSD motif is not involved in catalytic function, but is required to stabilize the growing D-xylosylated \(\beta\)-D-glucan chain. As can be seen by the molecular models, XXT5 has a much wider substrate-binding site and more accessible putative catalytic center compared to XXT2 (see Figure 5.1C). As XXT2 and XXT1 decorate the \(\beta\)-glucan chain, XXT5, which was shown to strongly interact with XXT2 (7) binds the growing chain in its open site and aids in its stability. CSCL4 has also been shown to interact with XXT5 (7). This would support the hypothesis that XXT5 plays a stabilizing role; as CSCL4 synthesizes the growing \(\beta\)-D-glucan chain, XXT5 loosely binds it in its DSD putative catalytic domain, positioning it such that XXT2 can also bind the chain and add D-Xyl units. Future
studies investigating the activities of XXT5 in vitro with the presence of XXT2 will need to be performed to determine which of the hypotheses occur in vivo.
CHAPTER 6. CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

Objective 1. Characterize DXD Motif Function in XXT2

DXD to AXA point mutations were made in two DXD motifs in XXT2; DWD, mut1, and DSD, mut2. N-terminal truncation variants were expressed in *E. coli* and assayed for D-xylosyltransferase activity. XXT2mut constructs were also stably transformed to *xxt1xxt2* and *xxt1xxt2xxt5* *A. thaliana*. First, it was demonstrated that truncating XXT2 and recombinantly expressing it without the transmembrane domain retains XT catalytic activity. Secondly, it was shown that tXXT2-His can be isolated to >90% purity, using either native or hybrid methods, and the resulting tXXT2-His protein exhibits XT catalytic activity. Using the same XT assay, it was shown that both the DWD and DSD motifs are essential for the catalytic function of XXT2. These results were not predicted from a molecular model of XXT2, which shows the DWD motif in an unstructured loop outside of the putative catalytic domain, and the DSD motif inside. Future studies characterizing generated XXT2mut(1, 2, 12):*xxt1xxt2* and XXT2mut(1, 2, 12):*xxt1xxt2xxt5* transgenic plants will further elucidate these motif functions *in vivo*.

Objective 2. Characterize DXD Motif Function in XXT5

DXD to AXA point mutants were made in two DXD motifs in XXT5, DWD (mut1) and DWD (mut2). Full-length XXT5mut(1, 2, 12) constructs were stably transformed to *xxt5* *A. thaliana*. T2 and T3 selected transgenic progeny were genotyped to confirm presence of the XXT5mut construct and *xxt5* background, their root hairs were assayed for mutant phenotype, total membrane protein fractions were collected and analyzed for presence of HA-XXT5mut protein, and the cell wall was analyzed for XyG content. The results show that the DWD motif is
not required for XXT5 function but the DSD motif is, as XXT5mut2 and XXT5mut12 were not able to complement the xxt5 phenotype. Cell wall analysis, however, showed that there is no significant reduction in XyG in the cell wall, suggesting that there is a partial complementation. It is therefore hypothesized that the root hair phenotype is due to a difference in XyG structure and composition, not total amount. Future studies analyzing the XyG structural subunits in the XXT5mut:xxt5 plants will aid in confirming this hypothesis.

Molecular models of XXT5 suggest that it contains a catalytic center that includes the DSD motif, coordinated Mn\(^{2+}\), and an open pocket for substrate binding. The in vivo studies of complemented XXT5mut:xxt5 plants confirm this hypothesis. XXT5 with a mutated DSD motif is not able to fully complement the xxt5 phenotype, meaning that it is not acting in a native manner. The question of whether it is catalytically active, however, remains unanswered. There are two proposed hypotheses to explain the role of XXT5; first, XXT5 requires the interaction of XXT2 (or another protein) to function. Second, XXT5 could play an accessory structural role, binding the growing \(\beta\)-D-glucan chain or substrates and interacting with XXT2 and XXT1 to promote or facilitate their catalytic activity.

**Future Directions**

**Characterize XXT2mut:xxt1xxt2 and XXT2mut:xxt1xxt2xxt5 Transgenic Plants**

\(T_0\) seeds from the XXT2mut:pGWb-15 transformations will be selected on hygromycin, and selected plants will be grown for seeds. These \(T_1\) seeds will be plated on hygromycin for a second round of selection, and the seeds from the survived plants collected. The \(T_2\) seeds will be plated on regular MS media and the resulting plants will be genotyped to confirm presence of XXT2mut insertions as well as xxt2xxt1 and xxt1xxt2xxt5 backgrounds. These seeds will also be analyzed for root phenotype and protein expression using the same described protocol as
XXT5mut:xt5 plants. Finally, crude cell wall will be extracted and analyzed using Driselase and XEG digestions to determine relative amount of XyG in the cell wall.

**Characterize XyG Subunit Structures Found in XXT5mut:xt5 Plants**

To characterize the XyG subunit structures in XXT5mut:xt5 plants, crude cell walls will be digested with XyG-specific endo-glucanase (XEG). XEG cleaves between a β(1-4)-D-glucose and a substituted β(1-4)-D-glucose, giving XXG, GXXG, XXXG, XXLG, XLXG, XXFG, and XLFG subunits (204). The relative amount of these subunits can then be determined using oligosaccharide mass fingerprinting (125). Additionally, immunolocalization studies using XyG-epitope specific primary antibodies (149) can be performed to study the XyG epitope profiles in XXT5mut:xt5 plants.

**Investigate the Function of XXT5 When XXT2 (Inactive) is Present**

It has been shown previously that tXXT2-His and tXXT5-Myc interact and be pulled-down together from Ni-NTA resin (7). Utilizing this method, tXXT5-Myc will be incubated with inactive tXXT2mut2-His on Ni-NTA resin. The XT activity assay will be performed as described previously while the proteins are interacting and bound to the resin. Oligosaccharide products will be analyzed using MALDI-TOF MS as described previously. Additionally, to test this hypothesis *in vivo, XXT5mut constructs will be stably transformed to xxt1.xxt5 A. thaliana* and complementation studies will be done in a similar fashion as described with XXT5mut:xt5. The obtained results will indicate if XXT5mut can complement phenotype with the presence of XXT2, thus indicating whether it plays a structural or catalytic role.
WORKS CITED


Appendix A. **tXXT2-His protein isolation in native conditions.** Coomassie R-250 stain analysis of tXXT2-His native isolation fractions as described in *Materials and Methods*. 20μL of each sample was loaded and run on 10%, 15-well, SDS-PAGE gel. NW = native wash, E = elution, numbers indicate the fraction. Ladder describes approximate mass in kDa (Spectra Multicolor Broad Range Protein Ladder, Pierce). tXXT2-His was able to be isolated to >90% purity as seen in E9 – E16.
Appendix B. tXXT2-His protein isolation under hybrid conditions. Coomassie R-250 stain analysis of tXXT2-His hybrid isolation fractions as described in Materials and Methods. 20µL of each sample was loaded and run on 10%, 15-well, SDS-PAGE gel. DW = denaturing wash, NW = native wash, E= elution, numbers indicate the fraction. Ladder describes approximate mass in kDa (Spectra Multicolor Broad Range Protein Ladder, Pierce). tXXT2-His was able to be isolated to >90% purity using this method, as can be seen in E10-E16.
### APPENDIX C

<table>
<thead>
<tr>
<th>Function</th>
<th>Primer Name</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXT2 gene specific primers</td>
<td>XXT2-F</td>
<td>CGTCATATGATGATTGAGAGGTGT</td>
</tr>
<tr>
<td></td>
<td>δXXT2-F</td>
<td>CTAAGGTCAATCTATCGGTAGTAGAGGAGGTGT</td>
</tr>
<tr>
<td></td>
<td>XXT2-200-R</td>
<td>GAGGCTCACCTGTTTCGTCGTCGTT</td>
</tr>
<tr>
<td></td>
<td>XXT2int-F</td>
<td>CAGAGATCTGATTGGTTAGCTAAGAGGACCGGCTTCCCA</td>
</tr>
<tr>
<td></td>
<td>XXT2m1-R</td>
<td>TAGGAAGCTAGGGTTCTTAGCTAATCCACGATCTGTCGCTACGCGTTAGTTAGGAGGTGT</td>
</tr>
<tr>
<td></td>
<td>XXT2int-R</td>
<td>CACAAAGAAAACTCAATCTGAGTGGACAGACAAAGACAAGACTGATATAGAGGTGT</td>
</tr>
<tr>
<td></td>
<td>XXT2m2-F</td>
<td>AGCTTTTGTGTCACATCCCTGAGATGGTTAGGTTAGGAGGAGGTGTGACTGATGCTGACTTTTAC</td>
</tr>
<tr>
<td></td>
<td>XXT2-R</td>
<td>AAATGGATCCCTOTAACCCATTGGTGT</td>
</tr>
<tr>
<td>XXT5 gene specific primers</td>
<td>XXT5-F</td>
<td>AGCCATGATGGATGAAACAAAACACTCAATCTGAGAGAGGTGACCTGGAAGACAGGACACGCGG</td>
</tr>
<tr>
<td></td>
<td>δXXT5-Myc-F</td>
<td>AGCCATGATGGATGAAACAAAACACTCAATCTGGAAGACAGGACACGCGG</td>
</tr>
<tr>
<td></td>
<td>XXT5-R</td>
<td>ACCATATGCTAGTTCTGAGTTTGTTTTGCCAC</td>
</tr>
<tr>
<td></td>
<td>δXXT5-222-R</td>
<td>AGTGCCAGTTGAGTGAGGAGAATTG</td>
</tr>
<tr>
<td></td>
<td>XXT5int-F</td>
<td>CACAGTAAAGATTGGTGAAATCTGCTAGTTTCCCA</td>
</tr>
<tr>
<td></td>
<td>XXT5m1-R</td>
<td>TAGGAAACCTAGGATGATCTGACTGACTGACTGACTGATTAGGACTGCAAGCCTGATGCTTATAGAGGAGTTAGGAGGTGTGT</td>
</tr>
<tr>
<td></td>
<td>XXT5int-R</td>
<td>CCAACGATTGCTCAGCTCAGCTGAGACAGAACAAGGACAGGACAGAACAAGGAGGTGTGTGTGCTTATAGAGGAGTTAGGAGGTGTGT</td>
</tr>
<tr>
<td></td>
<td>XXT5m2-F</td>
<td>GTGATGTTTTTGTGCGCTTACCCAGAAGTTGGATGCTAGTTAAGGATGCTGACTGATGCTTATAGGAGGAGTTAGGAGGTGTGT</td>
</tr>
<tr>
<td>pET-15b</td>
<td>T7-F</td>
<td>TAATACGACTCACTATAGGG</td>
</tr>
<tr>
<td>Genotyping</td>
<td>T7-R</td>
<td>CTTGCAAATATGCTGTCAGG</td>
</tr>
<tr>
<td>primers</td>
<td>35S-F</td>
<td>GAAACCTCTCCGGATGATTTTATG</td>
</tr>
<tr>
<td></td>
<td>LB-TDNA-R</td>
<td>CCCATGCAATGTGGATGATGACG</td>
</tr>
<tr>
<td></td>
<td>p5-F</td>
<td>CACACGAGACGGTTCAGGAG</td>
</tr>
<tr>
<td></td>
<td>LBb1-F</td>
<td>GCCTGAGACCCGCTTCTGGAAC</td>
</tr>
<tr>
<td></td>
<td>RP2-R</td>
<td>CGCTTGTAGTCCTGAGA</td>
</tr>
<tr>
<td></td>
<td>RP1-F</td>
<td>ATCCCGCAATTGAGCAGCTGACTTGA</td>
</tr>
<tr>
<td></td>
<td>LB3-R</td>
<td>GAAATTCGAGAAGGCACTAAAGG</td>
</tr>
</tbody>
</table>

**Appendix C. Table of PCR primers.** DNA primers used for polymerase chain reaction (PCR) analyses. Sequences are written 5’-3’. All primers were obtained from the Iowa State University DNA Facility.