Biosynthesis of Xyloglucan in Grasses and Identification of Endosperm Developmental Phases in Wheat Seeds

A thesis presented to
the faculty of
the College of Arts and Sciences of Ohio University

In partial fulfillment
of the requirements for the degree
Master of Science

Ramya Nadella
November 2007
This thesis titled
Biosynthesis of Xyloglucan in Grasses and Identification of Endosperm Developmental Phases in Wheat Seeds

by

RAMYA NADELLA

has been approved for
the Department of Environmental and Plant Biology
and the College of Arts and Sciences by

____________________________
Ahmed Faik
Assistant Professor of Environmental and Plant Biology

____________________________
Benjamin M. Ogles
Dean, College of Arts and Sciences
Abstract

NADELLA, RAMYA, M.S., November 2007, Environmental and Plant Biology

Biosynthesis of Xyloglucan in Grasses and Identification of Endosperm Developmental Phases in Wheat Seeds (124 pp.)

Director of Thesis: Ahmed Faik

Xyloglucans are the major hemicellulosic polysaccharides in the primary cell walls of dicots and non-graminaceous monocots but small amounts are also seen in the walls of grasses. They serve as cross-linkers of cellulose microfibrils and maintain the structural integrity of the cell. Xyloglucan biosynthesis mechanism is well understood and many biosynthetic genes have been identified in dicots. However this process is not known in grasses where structural differences in xyloglucan exist. Thus, this study aimed to investigate xyloglucan biosynthesis in grasses by functionally characterizing some xyloglucan biosynthetic genes, particularly the xylosyltransferases from wheat and rice. Wheat (TaGT34-7) and rice (OsGT34-3) homologs of Arabidopsis xyloglucan-xylosyltransferase (AtXT1) gene were previously identified using a bioinformatic script developed by Faik et al. (2006). The full-length cDNAs encoding TaGT34-7 and OsGT34-3 proteins were expressed in Pichia pastoris and Drosophila S2 expression systems, but the proteins did not show xylosyltransferase activity in Pichia pastoris and a low activity in Drosophila S2 cells. Further optimization is required to fully identify the function of these two proteins.

In another study, the endosperm developmental phases were identified in developing wheat seeds using sectioning and light microscopy. Cell walls of wheat endosperm have ~70% (w/w) of arabinoxylan (AX), 20% (w/w) of mixed linked glucan
(MLG), and only 5% (w/w) of cellulose. MLG and AX are deposited during cellularization and differentiation phases of the endosperm development, respectively. To identify the putative glycosyltransferase genes involved in the biosynthesis of these polymers (AX, MLG), it was necessary to first identify the endosperm developmental phases namely cellularization and differentiation phases at which they are deposited. Fresh wheat seeds were collected and their endosperm developmental phases were investigated using sectioning and light microscopy methods. Also, starch deposition was monitored using polarized light microscopy.

Approved: 

Ahmed Faik
Assistant Professor of Environmental and Plant Biology
Acknowledgments

I thank Dr. Ahmed Faik, my advisor, for the continuous and constant support. I am very grateful for the help, guidance and encouragement he has provided during my stay.

I also thank my committee members Dr. Allan Showalter, Dr. Sarah Wyatt and Dr. Marcia Kieliszewski for their guidance and support.

I thank Chris Havran for teaching me histological techniques, Harold Blazier and Aaron Mather for their help in the greenhouse, Matthew Shipp for his help in the expression work, and my cousin, Vijay Nadella for his support and encouragement throughout my masters program.
# Table of Contents

Abstract ............................................................................................................................... 3  
Acknowledgments ............................................................................................................... 5  
Table of Contents ................................................................................................................ 6  
List of Tables ...................................................................................................................... 8  
List of Figures ..................................................................................................................... 9  
List of Abbreviations ........................................................................................................ 11  
Chapter 1: GENERAL INTRODUCTION: IMPORTANCE, STRUCTURE, AND BIOSYNTHESIS OF PLANT CELL WALL POLYSACCHARIDES ......................................................... 14  
  Importance of plant cell walls .......................................................................................... 15  
  Importance in plant development ................................................................................... 15  
  Uses and Commercial applications ............................................................................... 15  
  Structure of plant cell walls ......................................................................................... 16  
  Primary plant cell wall types ....................................................................................... 17  
  Structures of primary cell wall polysaccharides ............................................................ 20  
    Cellulose microfibrils ................................................................................................. 20  
    Cross-linking glycans (also called hemicelluloses) ................................................... 20  
    Xylans .................................................................................................................... 20  
    Mixed linked glucans (MLGs) ................................................................................. 21  
    Xyloglucan (XyGs) .................................................................................................. 22  
    Pectins .................................................................................................................. 25  
  Structural proteins in primary cell walls ...................................................................... 27  
    Hydroxyproline-rich glycoproteins (HRGPs) ............................................................ 27  
    Extensins (EXTs) .................................................................................................... 27  
    Arabinogalactan proteins (AGPs) ............................................................................ 28  
    Proline-rich proteins (PRPs) .................................................................................. 28  
    Glycine-rich proteins (GRPs) ................................................................................ 29  
  Biosynthesis of plant cell wall polysaccharides .............................................................. 29  
    Cellulose biosynthesis ............................................................................................. 29  
    Xylan biosynthesis .................................................................................................. 31  
    Mixed linked glucan (MLGs) biosynthesis ............................................................... 33  
    Xyloglucan (XyGs) biosynthesis .............................................................................. 33  
    Pectin biosynthesis ................................................................................................ 38  
  Specific aims of the research ......................................................................................... 40  
Chapter 2: BIOSYNTHESIS OF XYLOGLUCAN IN GRASSES ........................................... 41  
  Introduction .................................................................................................................. 42  
  Previous results in the lab ............................................................................................. 45  
    Identification of the putative xyloglucan xylosyltransferases in wheat and rice using bioinformatic approach ................................................................. 45  
  Materials and Methods ............................................................................................... 48  
  Clones, chemicals, and enzymes ............................................................................... 48
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning of his-tagged version of <em>OsGT34-3</em> and <em>TaGT34-7</em> genes</td>
<td>49</td>
</tr>
<tr>
<td>Expression of <em>OsGT34-3</em> and <em>TaGT34-7</em></td>
<td>50</td>
</tr>
<tr>
<td>Expression in <em>Pichia pastoris</em></td>
<td>50</td>
</tr>
<tr>
<td>Expression in <em>Drosophila</em> S2 cells</td>
<td>53</td>
</tr>
<tr>
<td>Bradford protein assay</td>
<td>53</td>
</tr>
<tr>
<td>SDS-PAGE and Western blot</td>
<td>54</td>
</tr>
<tr>
<td>Xylosyltransferase assay</td>
<td>55</td>
</tr>
<tr>
<td>Results</td>
<td>56</td>
</tr>
<tr>
<td>Bioinformatic analysis of the putative xyloglucan xylosyltransferases from wheat and rice</td>
<td>56</td>
</tr>
<tr>
<td>Sub-cloning of the putative wheat and rice xylosyltransferases into <em>Pichia pastoris</em> and <em>Drosophila</em> S2 expression vectors</td>
<td>60</td>
</tr>
<tr>
<td>Optimization of the expression of wheat and rice putative xylosyltransferases in <em>Pichia pastoris</em> and <em>Drosophila</em> S2 expression systems</td>
<td>64</td>
</tr>
<tr>
<td>Xylosyltransferase activity in <em>Pichia</em> and <em>Drosophila</em> extracts</td>
<td>66</td>
</tr>
<tr>
<td>Discussion</td>
<td>69</td>
</tr>
<tr>
<td>Chapter 3: IDENTIFICATION OF ENDOSPERM DEVELOPMENTAL PHASES IN WHEAT SEEDS</td>
<td>74</td>
</tr>
<tr>
<td>Introduction</td>
<td>75</td>
</tr>
<tr>
<td>Endosperm development in cereals</td>
<td>75</td>
</tr>
<tr>
<td>Cell wall components of the wheat endosperm</td>
<td>79</td>
</tr>
<tr>
<td>Deposition of cell wall polysaccharides during endosperm development in wheat</td>
<td>81</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>82</td>
</tr>
<tr>
<td>Materials</td>
<td>82</td>
</tr>
<tr>
<td>Growing of wheat plants</td>
<td>82</td>
</tr>
<tr>
<td>Grouping of wheat seeds and preparation of seeds for light microscopy</td>
<td>83</td>
</tr>
<tr>
<td>Sectioning and Microscopy</td>
<td>83</td>
</tr>
<tr>
<td>Starch visualization</td>
<td>84</td>
</tr>
<tr>
<td>Results</td>
<td>86</td>
</tr>
<tr>
<td>Endosperm development in wheat grains</td>
<td>86</td>
</tr>
<tr>
<td>Discussion</td>
<td>94</td>
</tr>
<tr>
<td>Chapter 4: GENERAL CONCLUSIONS AND FUTURE WORK</td>
<td>97</td>
</tr>
<tr>
<td>General Conclusions</td>
<td>98</td>
</tr>
<tr>
<td>Xyloglucan Biosynthesis in Grasses</td>
<td>98</td>
</tr>
<tr>
<td>Identification of Endosperm developmental phases in wheat seeds</td>
<td>99</td>
</tr>
<tr>
<td>Future work</td>
<td>101</td>
</tr>
<tr>
<td>Optimization of protein expression in <em>Drosophila</em> S2 cells</td>
<td>101</td>
</tr>
<tr>
<td>Optimization of the biochemical assay</td>
<td>101</td>
</tr>
<tr>
<td>References</td>
<td>103</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1: Nomenclature of xyloglucan derived oligosaccharides, using the one-letter code, depending on the side chain present on the glucosyl residue</td>
<td>24</td>
</tr>
<tr>
<td>3.1: Different developmental stages of wheat seeds grouped according to the fresh weight of the seed (mg/seed)</td>
<td>85</td>
</tr>
<tr>
<td>4.1: Assigning an approximate time scale to the wheat seeds with the endosperm at different developmental phases</td>
<td>100</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1: A schematic model of the cell walls of <em>Arabidopsis</em> and Rice</td>
<td>19</td>
</tr>
<tr>
<td>1.2: Schematic diagram showing the structures of three pectic polysaccharides</td>
<td>26</td>
</tr>
<tr>
<td>2.1: Structure of the basic unit of xyloglucan along with the glycosyltransferases involved in its biosynthesis</td>
<td>44</td>
</tr>
<tr>
<td>2.2: Phylogenetic tree of the putative xylosyltransferases</td>
<td>47</td>
</tr>
<tr>
<td>2.3: The hydropathic plot of TaGT34-7 and OsGT34-3 according to the method of Kyte and Doolittle</td>
<td>58</td>
</tr>
<tr>
<td>2.4: Sequence alignment of the putative wheat and rice xylosyltransferases with AtXT1 and AtXT2</td>
<td>59</td>
</tr>
<tr>
<td>2.5: Strategies used for cloning <em>TaGT34-7</em> and <em>OsGT34-3</em> into <em>Pichia pastoris</em> and <em>Drosophila</em> S2 expression vectors</td>
<td>61</td>
</tr>
<tr>
<td>2.6: <em>Pichia</em> expression vector constructs used for the expression of <em>TaGT34-7</em> and <em>OsGT34-3</em></td>
<td>62</td>
</tr>
<tr>
<td>2.7: <em>Drosophila</em> expression vector constructs used for the expression of <em>TaGT34-7</em> and <em>OsGT34-3</em></td>
<td>63</td>
</tr>
<tr>
<td>2.8: Heterologous expression of <em>TaGT34-7</em>, <em>OsGT34-3</em> in <em>Pichia pastoris</em> and <em>Drosophila</em> S2 cells</td>
<td>65</td>
</tr>
<tr>
<td>2.9: XyG-xylosyltransferase activities in detergent-soluble extract from <em>Pichia pastoris</em> expressing <em>TaGT34-7</em> and <em>OsGT34-3</em></td>
<td>67</td>
</tr>
<tr>
<td>2.10: XyG-xylosyltransferase activity in detergent-soluble extracts from <em>Drosophila</em></td>
<td>67</td>
</tr>
</tbody>
</table>
S2 cells expressing TaGT34-7 and OsGT34-3

3.1: Time course of endosperm developmental phases in the wheat grain after anthesis

3.2: Diagrammatic representation of the endosperm cellularization process

3.3: Structure of arabinoxylan and the glycosyltransferase activities involved in its biosynthesis

3.4: Transverse and longitudinal sections of wheat seeds at stage 0

3.5: Transverse and longitudinal sections of wheat seeds at stage 1

3.6: Transverse and longitudinal sections of wheat seeds at stage 2

3.7: Transverse and longitudinal sections of wheat seeds at stage 3

3.8: Transverse and longitudinal sections of wheat seeds at stage 4

3.9: Transverse section of a wheat seed at stage 5

3.10: Summary figure showing the transverse sections of the seeds, from all the stages
List of Abbreviations

AGP: arabinogalactan protein
Ara: arabinose
Asn: asparagine
AX: arabinoxylan
BMGY: buffered glycerol-complex medium
BMMY: buffered methanol-complex medium
BSA: bovine serum albumin
CAP: contig assembly program
CAZy: carbohydrate active enzymes
cDNA: complementary DNA
CesA: cellulose synthase subunit
CSL: cellulose-synthase like
DTT: dithiothreitol
EB: extraction buffer
EDTA: ethylene diamine tetra acetic acid
EMS: ethyl methanesulfonate
ER: endoplasmic reticulum
EST: expressed sequence tag
Gal: galactose
GalA: galacturonic acid
GAX: glucuronoxarabinoxylan
Glc: Glucose
GlcNAc: N-acetylglucosamine
GPI: glycosylphosphotidylinositol
GRAVY: grand average of hydropathicity
GRP: glycine-rich protein
GT: glycosyltransferase
HGA: homogalacturonan
His: histidine
HRGP: hydroxyproline-rich glycoprotein
Hyp: hydroxyproline
Man: mannose
MLG: mixed linked glucan
MW: molecular weight
NDP: nucleotide diphosphate
No: number
PCR: polymerase chain reaction
PBS: phosphate buffered saline
PM: plasma membrane
PMSF: phenylmethanesulphonylfluoride
PRP: proline-rich protein
Pro: proline
PVDF: polyvinylidene difluoride membrane
QTL: quantitative trait loci
Rf: retention factor
RG-I: rhamnogalacturonan I
RG-II: rhamnogalacturonan II
RGRC: rice genome research center
Rha: rhamnose
RT-PCR: reverse transcriptase-polymerase chain reaction
SDS: sodium dodecyl sulphate
SDS-PAGE: sodium dodecyl sulphate polyacrylamide gel electrophoresis
Ser: serine
T-DNA: transfer DNA
Thr: threonine
THRGP: threonine hydroxyproline rich glycoprotein
TMD: transmembrane domain
UDP: uridine diphosphate
XET: Xyloglucan endotransglycosylase
XTH: xyloglucan endotransglycosylase/hydrolase
Xyl: xylose
XG-I: xylogalacturonan I
YPD: yeast extract peptone dextrose medium
Chapter 1: GENERAL INTRODUCTION: IMPORTANCE, STRUCTURE, AND BIOSYNTHESIS OF PLANT CELL WALL POLYSACCHARIDES
Importance of plant cell walls

Importance in plant development

Plant cells are surrounded by an envelope-like structure called the cell wall. The cell wall not only sets boundaries for a cell, but also sets its characteristic shape, provides mechanical support, and protects the cell from fungal and bacterial pathogens (Bowles, 1990; Carpita and Gibeaut, 1993). The structural and functional properties of the cell walls can be ascribed to the polysaccharides and proteins that make up the cell wall. Plant cell walls control the rate and direction of plant growth, which ultimately determines the shape of the plant cell (Sakurai, 1991). Oligosaccharides originated from the plant cell wall by enzymatic action can serve as signaling molecules in the plant defense, growth and development (Mohnen and Hahn, 1993). Oligosaccharides of homogalacturonan serve as elicitors and induce the synthesis of antibiotics called phytoalexins, which protects the plants from pathogen attack (Mohnen and Hahn, 1993; McNeil et al., 1984). Breakdown of cell wall polysaccharides results in the softening and ripening of fruits, such as tomatoes, and causes damages during shipping and handling, which reduces the market rate (Vicente et al., 2007). As the cell wall plays an important role in all these aspects of plant growth and development, research has been directed towards understanding the biogenesis, assembly and disassembly of plant cell walls.

Uses and Commercial applications

Plant cell walls are beneficial to humans in many ways. Polysaccharides and proteoglycans of the plant cell walls are the important raw materials in wood, paper,
textile, fuel and food industries (Farrokhi et al., 2006). Cell wall carbohydrates are the main source of dietary fiber (Selvendran, 1984). Pectins are used as gelling and thickening agents in the manufacture of fruit jams (Harris and Smith, 2006). Carbohydrates in the endosperm cell walls of wheat are the contributing factors to bread quality (Izydorczyk and Biliaderis, 1995). Plant cell wall polysaccharides can be broken down into simple sugars by the action of hydrolases and glycanases, and the released sugars can be converted to bioethanol by microbial fermentation (Gray et al., 2006). Modifying the cell wall to answer the needs of humans requires a thorough understanding of the biogenesis, deposition and metabolism of the cell wall components (Farrokhi et al., 2006).

Therefore, research in our lab mainly aims to identify the genes involved in the biosynthesis of two cell wall polysaccharides of grasses, particularly arabinoxylan (AX) and xyloglucan (XyG). The present study specifically focuses on the biosynthesis of XyG in grasses to understand the XyG biosynthetic mechanisms in dicots and grasses.

**Structure of plant cell walls**

The cell wall is a complex structure, which is divided into three zones namely, middle lamella, primary and secondary wall. Middle lamella, which is made up mostly of pectic polysaccharides, is shared by adjacent cells and holds them together. Primary cell wall is deposited by the growing cells. Depending on the composition, two types of primary cell walls can be distinguished (See “primary plant cell wall types” section). Secondary cell wall is deposited only by mature, specialized cells such as tracheids,
fibers and vessels, after the growth has ceased. In addition to the cellulose and hemicellulose, secondary cell walls contain additional substances such as lignin, cutin and suberin, which contribute to their mechanical strength (Carpita and McCann, 2000).

**Primary plant cell wall types**

Two different types of primary cell walls exist in flowering plants: type I and type II walls (Carpita and Gibeaut, 1993). Type I walls, characteristic of dicots and non-commelinoid monocots, consist of cellulose microfibrils interconnected by XyG polymers, which together form a network (Fig. 1.1) (Carpita and Gibeaut, 1993; Yokoyama and Nishitani, 2004). The cellulose-XyG network is embedded in an independent pectic network (Fig. 1.1). The pectic network mainly consists of homogalacturonan (HGA), rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II) (Carpita and Gibeaut, 1993; Carpita and McCann, 2000). Although these primary cell walls are largely made up of carbohydrates, structural proteins are also present up to 10% (McNeil et al., 1984; Showalter, 1993; Showalter, 2001). The percentage composition of different components, on a dry weight basis, of type I primary cell wall is cellulose 30%; XyG 25%; pectin 35%; and structural proteins 10% (Fry, 1988; Cosgrove, 1997).

Type II walls are special cell walls found in commelinoid monocotyledons, which include cereals such as rice, wheat, oats and barley (Carpita and Gibeaut, 1993). They have lower amounts of XyG and pectin compared to type I walls (Carpita and Gibeaut, 1993). The major hemicelluloses in type II walls are glucuronoarabinoxylans (GAX) and
mixed linked glucans (MLG) (Fig. 1.1) (Carpita and Gibeaut, 1993). Small amounts of XyG are also seen in the walls of grasses but the structure is different from that of type I cell walls (Fig. 1.1) (Carpita, 1996). The percentage composition of different components, on a dry weight basis, of type II primary cell wall is cellulose ~30%, GAX ~30%, MLG ~30%, pectin ~5%, XyG ~4%, structural proteins ~0.5% (Fry, 1988).
Figure 1.1 Schematic models of the cell walls of Arabidopsis and rice, which represents the type I and type II walls, respectively. Both type I and type II walls have cellulose microfibrils that are cross-linked by XyG in type I walls and by glucuronoarabinoxylan in type II walls (Carpita and Gibeaut, 1993). Cellulose-hemicellulose network is embedded in a pectic network, consisting mainly of homogalacturonan, rhamnogalacturonan I and rhamnogalacturonan II, in type I walls whereas type II walls have low amounts of pectin (Carpita and Gibeaut, 1993). Type I walls have structural proteins such as extensins and arabinogalactan proteins whereas type II walls have low amounts of protein but they have extensive amounts of hydrocinamic acids and phenylpropanoids (Carpita and Gibeaut, 1993). Small amounts of XyG are also seen in type II walls (Carpita and Gibeaut, 1993; Carpita 1996). The figure was adapted from Yokoyama and Nishitani, 2004.
Structures of primary cell wall polysaccharides

Cellulose microfibrils

Cellulose is the most abundant bio-polymer on earth, and it plays an important role in maintaining integrity of the cell walls, the size and shape of the cell and in determining the direction of plant growth (Duchesne and Larson, 1989). Primary cell walls contain approximately equal amounts of cellulose, hemicellulose and pectins whereas secondary cell walls are highly rich in cellulose constituting about 50% to 80% of the total wall (Cooper, 2000). Cellulose is synthesized at the plasma membrane by an enzyme complex termed a “rosette” (Delmer, 1999). Each rosette uses the UDP-glucose as a substrate to form 36 $\beta(1\rightarrow4)$ glucan chains, which coalesce to form a crystalline cellulose microfibril (Delmer, 1999). The deposition and orientation of these newly synthesized cellulose microfibrils in the cell wall are guided by cortical microtubules (Saxena and Brown, 2005).

Cross-linking glycans (also called hemicelluloses)

Xylans

Xylans are large family of heteropolymers found in primary and secondary cell walls (Ebringerova and Heinze, 2000). They are the principal cross-linking glycans of cellulose microfibrils in type II walls (Carpita and Gibeaut, 1993). Xylans mainly consist of $\beta(1\rightarrow4)$ linked xylose backbone, which is substituted with arabinofuranosyl residues at the C-2 and/or C-3 position and with glucuronosyl residues at C-2 position (Carpita and
The structure and composition of xylans depend on the cell type (tissues) and on the species (Ebringerova and Heinze, 2000; Porchia and Scheller, 2000). In the primary cell walls of graminaceous species, xylans are mainly substituted with arabinofuranosyl residues (thus called arabinoxylan), whereas xylans with more of glucuronic acid residues (glucuronoxylan) are seen in secondary walls of dicots (Porchia and Scheller, 2000). Xylans substituted with both arabinose and glucuronic acid (thus called glucuronoorabinoxylan) is seen in the vegetative tissues of grasses (Wong, 2006). The AXs are interconnected by esterified and etherified hydroxycinnamates and the principal hydroxycinnamate compound being the ferulic acid, which is esterified to C-5 of the arabinosyl units (Carpita and Gibeaut., 1993). Highly substituted AXs are initially synthesized in the Golgi and deposited in the cell walls, where the arabinofuranosyl residues are hydrolyzed (Gibeaut and Carpita, 1991) to facilitate the hydrogen bonding of AXs with cellulose microfibrils (Carpita and Gibeaut, 1993). AXs constitute about 70% (w/w) of the total cell wall in the wheat endosperm (See chapter 3).

**Mixed linked glucans (MLGs)**

MLGs are unbranched chains containing glucosyl residues linked by β(1→3) and (1→4) linkages in a ratio of 1:2 to 1:3 (Carpita and Gibeaut, 1993; Carpita, 1996). MLGs constitute about 30% of the primary cell walls of grasses, on a dry weight basis (Fry, 1988). MLGs deposition is specifically seen in the cell walls of elongating tissues of grasses and these polymers are degraded after the growth has stopped (Carpita and
Gibeaut, 1993; Buckeridge et al., 2004). MLGs also accumulate in the endosperm cell walls and maternal tissues of cereal grains (Buckeridge et al., 2004).

Xyloglucan (XyGs)

In dicots, XyGs comprise about 20-25% of the primary cell walls (Hayashi, 1989). They consist of a $\beta(1\rightarrow4)$ linked glucose backbone and 70% of the glucose residues are substituted with $\alpha$-xylosyl residues at the C-6 position. The xylosyl residues are substituted with galactosyl residues at the C-2 position by a $\beta$ linkage (Kooiman, 1961; Hayashi, 1989; Carpita and Gibeaut, 1993). Fucose residues are attached to galactosyl side chains at the C-2 position by an $\alpha$ linkage (Bauer et al., 1973). Most of the dicots and non-commelinoid monocots contain the XyGs with a trisaccharide side chain, which contains fucosyl, galactosyl and xylosyl residues, but some exceptions exist. For example, XyGs in solanaceous species, have arabinofuranosyl residues attached to xylosyl side chains at the C-2 position (Eda and Kato, 1978; Ring and Selvendran, 1981; Akiyama and Kato, 1982; Eda et al., 1983; Hayashi, 1989).

In grasses, XyGs have a backbone made of $\beta(1\rightarrow4)$ linked glucosyl residues of which only 30-40% are substituted with xylosyl residues (Kato et al., 1982). Galactosyl side chains are rarely seen in grass XyGs and the terminal fucosyl residues are totally absent (Hayashi, 1989; Carpita, 1996). However, fucosylated XyG was seen in the suspension-cultured cells of Festuca arundinacea, a deep-rooted perennial grass (McDougall, 1994).
**Xyloglucan nomenclature**

To simplify the writing of XyG structure, Fry et al. (1993) proposed a unique nomenclature based on a set of one-letter codes for XyG derived oligosaccharides depending on the terminal side chain present on each glucosyl residue in the backbone of XyG. For example, a glucosyl residue with no further substitution is indicated by “G”; a glucosyl residue substituted with a xylosyl residue is assigned a “X”, a glucosyl residue with a terminal galactose on the side chain is indicated by “L” and other variations as indicated in Table 1.1. XyG derived oligosaccharides will be named, depending on the code letters, from the non-reducing to reducing terminus (Fry et al., 1993). For example, the structure of XyG in Fig. 2.1 (page 37) can be represented as XXFG by using the one-letter nomenclature, where ‘X’ represents α-D-Xyl-(1→6)-β-D-Glc; ‘F’ represents α-L-Fuc-(1→2)-β-D-Gal-(1→2)-α-D-Xyl-(1→6)- β-D-Glc segment; and ‘G’ represents an unbranched β-D-Glc residue.
Table 1.1 Nomenclature of xyloglucan derived oligosaccharides, using the one-letter code, depending on the side chain present on the glucosyl residue (Fry et al., 1993).

<table>
<thead>
<tr>
<th>Code letter</th>
<th>Structure represented</th>
<th>Mnemonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>β-D-Glc*</td>
<td>Glucose</td>
</tr>
<tr>
<td>X</td>
<td>α-d-Xylp-(1→6)-β-D-Glc*</td>
<td>Xylose</td>
</tr>
<tr>
<td>L</td>
<td>β-D-Galp-(1→2)-α-D-Xylp-(1→6)-β-D-Glc*</td>
<td>galactose</td>
</tr>
<tr>
<td>F</td>
<td>α-1-Fucp-(1→2)-β-D-Galp-(1→2)-α-D-Xylp-(1→6)-β-D-Glc*</td>
<td>Fucose</td>
</tr>
<tr>
<td>A</td>
<td>α-1-Araf-(1→2)-β-D-Glc*</td>
<td>Arabinose</td>
</tr>
<tr>
<td></td>
<td>α-D-Xylp-(1→6)-β-D-Glc*</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>β-D-Xylp-(1→2)-β-D-Glc*</td>
<td>Beta-xylose</td>
</tr>
<tr>
<td></td>
<td>α-D-Xylp-(1→6)-β-D-Glc*</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>α-1-Araf-(1→3)-β-D-Xylp-(1→2)-β-D-Glc*</td>
<td>follows A and B</td>
</tr>
<tr>
<td></td>
<td>α-D-Xylp-(1→6)-β-D-Glc*</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>α-1-Araf-(1→2)-α-D-Xylp-(1→6)-β-D-Glc*</td>
<td>Solanaceae</td>
</tr>
</tbody>
</table>

For example, cellohexaose, an exogenous acceptor, has six β(1→4) linked glucose residues. Using the one letter codes given in the table 1.1, cellohexaose can be represented as GGGGGG. If the acceptor is substituted with xylosyl residue at the third glucosyl residue from the reducing end, it can be represented as GGGXGG. The table was adapted from Fry et al., 1993.
Pectins

Pectins are a group of acidic polysaccharides rich in galacturonic acid residues. Three different pectic polysaccharides can be found in the primary walls of all flowering plants, namely HGA, RG-I and RG-II (Fig. 1.2) (Carpita and McCann, 2000). HGA is a linear polymer made up of $\alpha(1\rightarrow4)$ linked galacturonic acid residues (Carpita and McCann, 2000). HGA is synthesized in the Golgi apparatus as a highly methyl esterified polymer but the methyl ester groups are removed after it is deposited in the wall, which allows the cross-linking of the polymer with itself through calcium cations (Mohnen, 1999). Apart from methyl esterification, other biosynthetic modifications to HGA include $O$-acetylation at C-2 or C-3 positions (Mohnen, 1999), addition of xylosyl residues at C-3 position (Kikuchi et al., 1996; Le Goff et al., 2001) or addition of apiosyl residues on C-2 or C-3 positions (Longland et al., 1989).

RG-I is made up of several repeats of the disaccharide unit: $(1\rightarrow2)-\alpha-L$-rhamnose-$(1\rightarrow4)-\alpha-D$-galacturonic acid (Fig. 1.2) (Jarvis, 1984). The rhamnosyl residues in RG-I are substituted on C-4 position with side chains such as galactans, arabinans and highly branched arabinogalactans (Mohnen, 1999). HGA and RG-I are also seen in grasses but in lower amounts (Shibuya and Nakane, 1984; Carpita, 1989).

RG-II is highly branched pectic polysaccharide with a HGA backbone and the side chains are composed of eleven different sugars (O’Neill et al., 1996). RG-II can dimerize through borate ester linkages through apiosyl side chains (O’Neill et al., 1996).
Figure 1.2 Schematic diagram showing the structures of three pectic polysaccharides. RG-I and RG-II are covalently attached to HGA. Ara, arabinose. Gal, galactose. GalA, galacturonic acid. HGA, homogalacturonan. RG-I, rhamnogalacturonan I. RG-II, rhamnogalacturonan II. Rha, rhamnose. The figure was adapted and modified from Willats et al., 2001.
Structural proteins in primary cell walls

Hydroxyproline-rich glycoproteins (HRGPs)

This group of cell wall glycoproteins includes extensins (EXTs), arabinogalactan proteins (AGPs), proline-rich proteins (PRPs) and glycine-rich proteins (GRPs). As the name indicates HRGPs are rich in hydroxyproline (Hyp) residues and are glycosylated on Hyp residues according to Hyp-contiguity hypothesis, which dictates that contiguous Hyp residues are glycosylated with short arabinino-oligosaccharides whereas non-contiguous Hyp residues are decorated with large complex hetero-polysaccharide units (Kieliszewski and Lamport, 1994).

Extensins (EXTs)

EXTs are a class of HRGPs present in the cell walls of higher plants (Showalter, 1993). Dicot EXTs are rich in Hyp and serine (Ser) residues along with other amino acids such as valine, tyrosine, lysine and histidine (Showalter, 1993). They contain a characteristic repeating unit of a pentapeptide motif [Ser-Hyp₄] in which Hyp residues are decorated with arabinosyl residues and serine is decorated with a single galactosyl residue (Kieliszewski et al., 1990; Showalter, 1993). Monocot EXTs, observed in maize, were designated as threonine-hydroxyproline rich glycoproteins (Kieliszewski and Lamport, 1987) as they contain threonine and proline (Pro) amino acids in addition to Hyp, Ser and lysine; and they are lightly or not at all glycosylated (Kieliszewski et al., 1990).
Arabinogalactan proteins (AGPs)

AGPs are a class of highly glycosylated HRGPs, which are broadly classified into classical and non-classical AGPs (Showalter, 2001; Gaspar et al., 2001). AGPs contain arabinogalactan polymers and arabinoside oligomers as their carbohydrate components (90%), which are attached to Hyp residues on the protein backbone (10%) (Showalter, 2001). Classical AGPs contain Hyp, alanine, threonine, serine and glycine as their major amino acid constituents and typically contain an N-terminal secretion sequence, which is removed from the mature protein, and a C-terminal hydrophobic region (Showalter, 2001; Gaspar et al., 2001). Prolyl-hydroxylation, glycosylation, addition and processing of glycosylphosphatidylinositol-anchor (GPI-anchor) are the major post-translational modifications occurring to AGPs (Gaspar et al., 2001). In addition to these post-translational modifications, AGPs are further processed at the plasma membrane (PM) by phospholipase C or D, which releases the proteins from PM to cell wall or extracellular destinations (Showalter, 2001). Non-classical AGPs contain other domains rich in cysteine (cysteine-rich AGPs), asparagine (asparagine-rich AGPs), and have lesser amounts of Hyp compared to classical AGPs (Showalter, 2001).

Proline-rich proteins (PRPs)

PRPs are a class of lightly glycosylated HRGPs, which contain equal amounts of Pro and Hyp and repeated units of Pro-Pro residues along with other large repeat units (Datta et al., 1989; Showalter, 1993). Monocot PRP, from maize consists of an N-
terminal proline rich domain and a C-terminal proline poor domain containing cysteine residues (Jose-Estanyol et al., 1992).

**Glycine-rich proteins (GRPs)**

As the name indicates, glycine-rich proteins have glycine as their main amino acid, up to 70%, in their primary structure (Showalter, 1993). Monocot GRPs have somewhat less glycine content when compared to dicot GRPs (Showalter, 1993). Most of the GRPs contain an amino terminal signal peptide but some of them lack the signal peptide and contain an RNA binding sequence, which suggests that not all GRPs are secreted to the cell wall but some of them are localized in the cytoplasm (Showalter, 1993).

**Biosynthesis of plant cell wall polysaccharides**

**Cellulose biosynthesis**

Cellulose is synthesized at the plasma membrane by cellulose synthase complexes called rosettes. Each rosette consists of six subunits and each subunit is made of six cellulose synthase catalytic subunits (CesAs) (Scheible et al., 2001; Doblin et al., 2002). Each CesA polypeptide is involved in the synthesis of one $\beta(1\rightarrow4)$ glucan chain, and thus each rosette is synthesizing 36 $\beta(1\rightarrow4)$ glucan chains, which coalesce to form a cellulose microfibril (Doblin et al., 2002). All CesAs identified to date have some structural features in common such as eight transmembrane domains (TMD), two at the N-terminus
and six at the C-terminus (Richmond, 2000). The region between the N-terminal and C-TMDs is called the globular domain, which contain several motifs such as a DXD motif and QXXRW motifs that are responsible in substrate binding and the formation of a catalytic site respectively (Richmond, 2000). The gene encoding the CesA catalytic subunit was first identified in *Acetobacter xylinum* (Saxena et al., 1990) and the entire cellulase synthase operon containing four genes was identified by Wong et al. (1990). After the discovery of the cellulase synthase operon, many groups tried to find the plant homologues by screening the cDNA libraries, but none of them were successful (Delmer, 1999). Pear et al. (1996) used molecular approaches and identified the first plant CesA gene from cotton fibers. They identified two genes, *celA1* and *celA2*, that were highly expressed during the time at which the secondary cell wall (containing almost pure cellulose), is synthesized in the cotton fibers. A DNA segment encoding the central region of the celA1 protein was over expressed in *E.coli* and was shown to bind UDP-glucose in the presence of Mg$^{2+}$, similar to bacterial CesAs, which gave further evidence that *CelA1* may encode the plant CesA (Pear et al., 1996). Later on ~12 CesA genes were identified in rice and 10 in *Arabidopsis* (Farrokhi et al., 2006).

Using bioinformatic approach, Richmond and Somerville (2000) identified additional 41 genes in *Arabidopsis* that have sequence similarity to CesA genes and named them as *Cellulose Synthase-Like (CSL)* genes. These CSL proteins are classified as GT-2 family according to CAZy (http://www.cazy.org/) and grouped under eight families namely CslA, CslB, CslC, CslD, CslE, CslF, CslG and CslH in addition to CesA group (Richmond and Somerville, 2000; Hazen et al., 2002). Richmond and Somerville (2000) hypothesized that the CSL genes might encode the enzymes responsible for the
formation of the hemicellulosic β(1,4)-linked backbones as they present similarity to
β(1,4)glucan backbone of cellulose. Consistent with this hypothesis, several CSL genes
encoding the enzymes synthesizing the hemicellulosic backbones were linked to the
biosynthesis of mannans, MLG, and XyG (Dhugga et al., 2004; Liepman et al., 2005;
Burton et al., 2006; Cocuron et al., 2007). However, the enzyme(s) that is (are) involved
in xylan backbone synthesis is still not known.

**Xylan biosynthesis**

*Arabidopsis fragile fiber8 (fra8)* is a secondary cell wall mutant that has reduction
in the secondary cell wall fiber thickness and a decrease in the amount of cellulose
(Zhong et al., 2005). Monosaccharide composition analysis suggests a defect in the
addition of glucuronic acid residues onto xylans. *FRA8* (At2g28110) gene was cloned
using a positional cloning approach and transformation of the mutant plants with the wild
type gene completely rescued the mutant phenotype (Zhong et al., 2005). The FRA8
protein has glycosyltransferase (GT) characteristics and belongs to GT-47 family
according to CAZy classification. Therefore the authors speculate that FRA8 is a
putative glucuronyltransferase that is involved in adding the glucuronic acid residues
onto xylans (Zhong et al., 2005). However, this hypothesis still needs biochemical
confirmation.

Another series of secondary cell wall mutants having defects in the xylem
development are the *irregular xylem (irx)* mutants (Turner and Somerville, 1997). The
first three *irx* mutants (*irx1, irx2, irx3*) were identified by screening for altered xylem cell
morphology of the hand-cut sections, of an *Arabidopsis* ethyl methanesulfonate (EMS)-mutagenized populations (Turner and Somerville, 1997). Later on, microarray data mining and coexpression analyses resulted in the identification of additional genes involved in the secondary cell wall deposition (Brown et al., 2005; Persson et al., 2005). Mutant analysis of *irx8* (At5g54690; GT-8 family) and *irx9* (At2g37090; GT-43 family) revealed that IRX8 is responsible for the synthesis of a specific glycosyl sequence at the reducing end of the glucuronoxylan and IRX9 is responsible for the elongation of the glucuronoxylan chains (Pena et al., 2007). However, the biochemical function of these genes is needed. More recently, Brown et al. (2007) identified two secondary cell wall mutants, *parvus-3* (At1g19300; GT-8 family) and *irx14* (At4g36890; GT-43 family), both of which have defects in xylan biosynthesis. Studies on *parvus-3* and *irx14* in comparison with *irx7*, *irx8* and *irx9* revealed that PARVUS-3, IRX7 and IRX8 are responsible for synthesizing the specific glycosyl sequence at the reducing end of the glucuronoxylan, which acts as a primer, and IRX9 and IRX14 are responsible for the elongation of glucuronoxylan chains (Brown et al., 2007).

Baydoun et al. (1989) have shown that the membrane fractions prepared from pea epicotyls have a β(1→4)xylosyltransferase activity as evidenced by the incorporation of $[^{14}C]$xylose from UDP-$[^{14}C]$Xyl into a product containing β(1→4) xylans. The same study also has shown that the xylosyltransferase and glucuronyltransferase activities are tightly coupled, which was evident by the incorporation of glucuronic acid only when UDP-$[^{14}C]$glucuronic acid and UDP-$[^{14}C]$Xyl are incubated together (Baydoun et al., 1989). Porchia and Scheller (2000) showed that microsomal membranes prepared from wheat seedlings also incorporated xylose from UDP-$[^{14}C]$Xyl into a product containing
β(1→4) xylans. However, the structures of the product formed were not determined in any of these studies.

**Mixed linked glucan (MLGs) biosynthesis**

Comparative genomics strategy allowed the identification of the genes encoding the MLG synthase (Burton et al., 2006). By comparing the quantitative trait loci (QTL) for MLG of barley to that of rice, Burton et al. (2006) identified the *CslF* gene family as the putative genes that might encode the MLG synthase. To confirm their function, the rice *CslF* genes were expressed in *Arabidopsis*, which lack MLGs in their cell walls, thus providing a gain-of-function evidence for the role of *CslF* genes in MLG biosynthesis (Burton et al., 2006). Immunoelectron microscopy using an antibody against MLG confirmed its presence in *Arabidopsis* cell walls (Burton et al., 2006).

**Xyloglucan (XyGs) biosynthesis**

XyGs are the only hemicellulosic polysaccharides for which most of the genes involved in their biosynthesis have been identified in dicots, specifically in *Arabidopsis*. XyG biosynthesis occurs in the Golgi apparatus by the co-operative action of glucan synthase and xylosyltransferase (Ray, 1980; Hayashi and Matsuda, 1981; Brummel et al., 1990; Faik et al., 2002). The other sugars on the side chains such as galactosyl and fucosyl residues will be added independently by galactosyltransferase and fucosyltransferases respectively (Brummel et al., 1990; Faik et al., 1997). The four biosynthetic activities involved in the elaboration of XyGs are indicated in Fig. 2.1.
XyG $\beta(1\rightarrow4)$ glucan synthase (Fig. 2.1) is the enzyme that catalyzes the formation of the glucose backbone of XyG. The gene encoding the $\beta(1\rightarrow4)$ glucan synthase was recently identified by Cocuron et al. (2007). The expression profiles of XyG biosynthetic genes, $MUR3$ and $AtXT1$, were determined in the developing nasturtium seeds and the results revealed that their nasturtium homologs were highly expressed at 22 days post anthesis, which corresponds to maximum deposition of XyG in the nasturtium cotyledons. The data also revealed that one member of $CslC$ gene family was also highly expressed at that stage (Cocuron et al., 2007). The expression of $CslC$ genes at the time (22 DPA) during which the XyG is accumulated in the nasturtium seeds served as a good foundation for postulating that this member of $CslC$ gene family might encode the enzymes that make the backbone of XyG (Cocuron et al., 2007). Therefore *Pichia* cells were transformed with *Tropaeolum majus* $CslC$ (*TmCslC*) and its closest homolog from *Arabidopsis*, $AtCslC4$; and the products obtained from the over-expressing yeast cells had the glucan chains, which are $\beta(1\rightarrow4)$ linked (Cocuron et al., 2007). But the $\beta(1\rightarrow4)$ linkage is seen in both cellulose and XyG. If $AtCslC4$ is involved in the biosynthesis of cellulose, it should be localized at the site of synthesis of cellulose, i.e., plasma membrane whereas if $AtCslC4$ encodes the XyG $\beta(1\rightarrow4)$ glucan synthase, the protein should be localized in the Golgi. Localization studies revealed the expression of $AtCslC4$ in the Golgi, which confirmed that $AtCslC4$ is the gene that encodes the XyG $\beta(1\rightarrow4)$ glucan synthase (Cocuron et al., 2007).

XyG $\alpha(1\rightarrow6)$ xylosyltransferase (Fig. 2.1) is the enzyme that catalyzes the addition of xylosyl residues to the glucan backbone of XyG. Faik et al. (2002) demonstrated that pea microsomal membranes catalyze the addition of xylosyl residues to
the exogenous cello-oligosaccharide acceptors (Degree of polymerization > 4), when supplied UDP-[\(^{14}\)C]xylose. Interestingly, the biochemical characterization of the \([^{14}\)C]radiolabeled products indicated that pea xylosyltransferase activity has similar enzyme characteristics to the galactomannan galactosyltransferase from fenugreek (Edwards et al., 1999; Faik et al., 2002). Due to the similarity in the linkages of galactose in galactomannan and xylose in XyG, it was hypothesized that galactomannan galactosyltransferase and XyG-xylosyltransferase might have similar catalytic mechanisms to add the galactosyl or xylosyl residues on the glycan backbones of galactomannan or XyG, respectively (Faik et al., 2002). Therefore, galactomannan galactosyltransferase gene (TfGalT), which was previously cloned by Edwards et al. (1999), was used as query to search the Arabidopsis genome. This search led to the identification of seven Arabidopsis genes having significant sequence similarities to TfGalT (Faik et al., 2002). The seven genes, named Arabidopsis thaliana glycosyltransferase1-7 (AtGT1-7; GT-34 family), were expressed in the Pichia pastoris expression system and among the seven genes, the product of one gene AtGT1 showed the xylosyltransferase activity (Faik et al., 2002). Biochemical characterization of AtGT1 revealed that the enzyme added the xylosyl residues specifically to the penultimate glucose residue of cellopentaose (annotated GGGGG according to the one-letter code nomenclature) to form GXGGG (Faik et al., 2002). Therefore, AtGT1 gene was named Arabidopsis thaliana xylosyltransferase 1 (AtXT1) (Faik et al., 2002). More recently, Cavalier and Keegstra (2006) showed that AtXT2 also encoded an enzyme with xylosyltransferase activity by expressing it in Drosophila S2 expression system. The same study have shown that both AtXT1 and AtXT2 add the xylosyl residues to the
fourth glucose residue from the reducing end of the cellopentaose and cellohexaose acceptors to produce GXGGG and GGXGGG, respectively (Cavalier and Keegstra, 2006). In addition, both the AtXT1 and AtXT2 could add multiple xylosyl residues upon prolonged incubation to produce GGXXGG and minute amounts of GXXXGG (Cavalier and Keegstra, 2006).

XyG $\beta(1\rightarrow2)$ galactosyltransferase (Fig. 2.1) is the enzyme that catalyzes the addition of galactosyl residues to the xylosyl residues on the side chains of XyG. A gene encoding the XyG galactosyltransferase was identified, using reverse genetic approaches, by Madson et al. (2003). In this study, EMS-mutagenized Arabidopsis population was screened for alterations in the cell wall monosaccharide composition, which resulted in the identification of mur3, a mutant having $>90\%$ reduction in the fucose content and a significant reduction in the galactose content in XyGs (Madson et al., 2003). Further studies on the mur3 mutant revealed that MUR3 (GT-47 family) encoded a XyG-galactosyltransferase, which specifically transfers galactose residues to the third xylosyl residue from the non-reducing end in the XXXG subunit of XyG (Madson et al., 2003). Therefore, positional cloning was used to identify the MUR3 gene, and the protein encoded by this gene had all the characteristics of the Golgi-localized GTs (Madson et al., 2003). MUR3 was expressed in the Pichia pastoris expression system, and the product of this gene catalyzed the transfer of galactose residues to the xylosyl side chains of mur3–derived XyG, which further confirmed that MUR3 gene encodes the XyG galactosyltransferase (Madson et al., 2003).

XyG $\alpha(1\rightarrow2)$ fucosyltransferase (Fig. 2.1) is the first enzyme for which the glycosyltransferase gene had been cloned (Perrin et al., 1999) and this enzyme catalyzes
the addition of terminal fucosyl residues to the galactosyl residues on the side chains of XyG. An α-fucosyltransferase was purified from pea microsomal membranes by using a biochemical assay that was developed in Dr. Gordon Maclachlan’s lab (Faik et al., 1997), along with traditional biochemical purification methods (Perrin et al., 1999; Faik et al., 2000). Microsomal membranes prepared from pea catalyzed the addition of fucosyl residues to galactosyl residues when guanosine diphosphate (GDP)-fucose was used as a donor and tamarind seed XyG (TXyG) as an acceptor, which lacks the fucose residues \textit{in vivo} (Faik et al., 1997; Faik et al., 2000). A ~63 KDa polypeptide was purified from the pea microsomal membranes and amino acid sequencing was done on trypsin-digested fragments (Perrin et al., 1999; Faik et al., 2000). The information from the amino acid sequencing was used to isolate an \textit{Arabidopsis} expressed sequence tag (EST), which was further used to screen an \textit{Arabidopsis} cDNA library (Perrin et al., 1999). Thus, an \textit{Arabidopsis} full-length cDNA clone corresponding to the fucosyltransferase was identified and named \textit{Arabidopsis thaliana} fucosyltransferase 1 (\textit{AtFT1}; GT-37 family) (Perrin et al., 1999). The identity of this cDNA as a fucosyltransferase was confirmed by immunoprecipitation of the fucosyltransferase activity from the \textit{Arabidopsis} cell suspension cultures, using an antibody raised against a portion of AtFT1 protein expressed in \textit{E.coli} (Perrin et al., 1999). The immunoprecipitated proteins obtained from the \textit{Arabidopsis} cell suspension cultures showed XyG fucosyltransferase activity \textit{in vitro} (Perrin et al., 1999). Also, \textit{AtFT1} was expressed in a mammalian expression system, COS-7 cells, and the fucosyltransferase activity was confirmed (Perrin et al., 1999). The immunoprecipitation experiments together with the fucosyltransferase activity in the
COS-7 cell lines confirmed that AtFT1 encoded a protein with XyG fucosyltransferase activity.

Peptide sequence information of the protein purified from pea microsomal membranes was used to isolate a cDNA clone from pea cDNA library by molecular cloning experiments (Faik et al., 2000). The cDNA clone was named *Pisum sativum* fucosyltransferase 1 (*PsFT1*; GT-37 family), and the protein corresponding to *PsFT1* had all the short peptides obtained from amino acid sequencing (Faik et al., 2000). The same study has also demonstrated that the *PsFT1* fucosyltransferase specifically transfers fucose residues on to XyG and not to RG-I and RG-II, both of which also contain terminal fucosyl residues (Faik et al., 2000).

**Pectin biosynthesis**

Pectic polysaccharides are biosynthesized in the Golgi apparatus and at least 54 biosynthetic enzyme activities are probably required to synthesize the highly complex pectic polysaccharides (Mohnen, 1999). A gene encoding the galacturonosyltransferase involved in the biosynthesis of HGA was identified by Sterling et al. (2006) and was named as *galacturonosyltransferase 1, GAUT1*. Heterologous expression of *GAUT1* in human embryonic kidney cells confirmed the galacturonosyltransferase activity for GAUT1 (Sterling et al., 2006).

Screening of *Arabidopsis* T-DNA knock-out mutants in GT-47 family led to the identification of ARAD1, a putative arabinosyltransferase involved in adding the arabinans side chains to RG-I (Harholt et al., 2006). The RG-I extracted from mutant
plants has only 30% of the arabinose compared to wild type, confirming that the deficiency in arabinose is associated with RG-I (Harholt et al., 2006). The mutant phenotype of \textit{arad1} was complemented by transformation of plants with \textit{ARAD1} under the control of \textit{CaMV35S} promoter (Harholt et al., 2006). However, evidence for the biochemical activity of the arabinosyltransferase activity is still needed.

\textit{Xylogalacturonan deficient 1 (xgd1)} is another mutant from GT-47 family. In the \textit{xgd1} mutant, xylogalacturonan (XGA) is completely absent and complementation of the mutant phenotype restored the cell wall composition to wild type suggesting that XGD1 gene encodes a xylosyltransferase that attaches xylosyl residues to a HGA backbone (Scheller et al., 2007).

\textit{Nicotiana plumbaginifolia nolac-H18} (\textit{Nicotiana plumbaginifolia} non-organogenic callus with loosely attached cells) is another cell wall mutant, which has impaired cell attachment in callus culture, low or complete lack of glucuronic acid content and inability of RG-II to form dimers (Iwai et al., 2002). The gene responsible for the mutant phenotype in \textit{nolac-H18} is a glucuronosyltransferase, \textit{NpGUT1}. Complementation of the mutant plants rescued the mutant phenotype but the enzyme activity has not been established (Iwai et al., 2002).

Two xylosyltransferase genes, rhamnogalacturonan xylosyltransferase 1 (\textit{RGXTI}) and rhamnogalacturonan xylosyltransferase II (\textit{RGXTII}) were recently identified (Egelund et al., 2004). Both RGXTI and RGXTII catalyze the addition of xylosyl residues to fucosyl residues on the side chains of RG-II via an $\alpha(1,3)$ linkage (Egelund et al., 2006).
Specific aims of the research

As indicated earlier in this chapter, XyG is the major hemicellulosic polysaccharide in type I walls but small amounts of this polysaccharide are also found in type II walls of grasses. During the cell expansion, cell wall modifying enzymes such as XyG endotransglycosylase/hydrolases (XTH) re-organize the cellulose-XyG network in type I walls and aid in the wall loosening (Yokoyama and Nishitani, 2001). Arabidopsis and rice have XTH gene families consisting of 33 and 29 members, respectively (Yokoyama and Nishitani, 2001; Yokoyama et al., 2004). The presence of a large XTH gene family in rice, having the type II walls with only 5% XyG, was particularly surprising and suggests that XyG may also have important structural function in type II cell walls, though present in lower amounts. Therefore, to understand the importance of XyG in type II walls, it is necessary to study the biosynthesis of this polysaccharide in grasses and compare its biosynthetic mechanism to dicots. The putative rice and wheat XyG xylosyltransferase genes, OsGT34-3 and TaGT34-7, were previously cloned in our lab (Faik, unpublished data). The specific aims of this project are to

1. Study the biosynthesis of XyG in grasses, specifically in wheat and rice.
   (a) Sub-clone these putative xylosyltransferases into Pichia pastoris and Drosophila S2 expression vectors.
   (b) Optimize the conditions for the expression of these putative xylosyltransferases in Pichia pastoris and Drosophila S2 expression systems and test the enzyme activity using exogenous acceptors.

2. Identification of endosperm developmental phases in wheat seeds, which would provide a strong foundation for future work on AX and MLG biosynthesis.
Chapter 2: BIOSYNTHESIS OF XYLOGLUCAN IN GRASSES
Introduction

XyG in type I walls consists of $\beta(1\rightarrow4)$ linked glucose backbone and 70% of the glucosyl residues in the backbone are substituted with xylosyl residues at the C-6 position (Kooiman, 1961; Hayashi, 1989). The xylosyl residues can be substituted by galactosyl residues and the galactosyl residues in turn can be substituted by fucosyl residues (Bauer et al., 1973; Hayashi, 1989) (Fig. 2.1). The biosynthesis of XyG in type I walls, requires the activities of at least four biosynthetic enzymes, namely $\alpha(1,2)$fucosyltransferase, $\beta(1,2)$galactosyltransferase, $\alpha(1,6)$xylosyltransferase and $\beta(1,4)$glucan synthase (Fig. 2.1). Most of the genes encoding the enzymes involved in the XyG biosynthesis have been identified in dicots (see the biosynthesis of XyG section in chapter 1).

However, XyG in grasses consists of $\beta(1\rightarrow4)$ linked glucose backbone as in dicots, but only 30-40% of the glucosyl residues in the backbone are substituted with xylosyl residues at C-6 position (Carpita and Gibeaut, 1993). Galactosyl residues are rarely seen in the XyG of grasses and the terminal fucose is totally absent (Carpita and Gibeaut, 1993). Therefore, the biosynthesis of XyG in grasses may require only three enzyme activities namely, $\beta(1,2)$galactosyltransferase, $\alpha(1,6)$xylosyltransferase and $\beta(1,4)$glucan synthase. But none of the genes encoding these enzymes have been identified in grasses. Carbohydrate active enzymes (CAZy) database have listed 21 rice genes in GT-37 family in which the previously identified *Arabidopsis (AtFT1)* and pea (*PsFT1*) XyG fucosyltransferase genes are present. But as fucose residues are absent in the XyG of grasses, the wheat and rice genes listed in GT-37 family might not be involved in the biosynthesis of XyG but may be involved in the fucosylation of other cell
wall components (i.e., AGPs, pectins) or may not be even fucosyltransferases. CAZY database also listed 29 rice genes and 2 wheat genes in GT-47 family, which includes the previously identified XyG galactosyltransferase gene, *MUR3*. As the galactose residues are rarely seen in the XyG of grasses, the wheat and rice genes in GT-47 could be candidate GT genes involved in the biosynthesis of XyG. The main sugar that forms the side chain of the XyG in grasses is xylose. Therefore, it is logical to start the study on XyG biosynthesis in grasses, by identifying the genes encoding the XyG xylosyltransferases. Previous work in our lab led to the identification of the putative XyG xylosyltransferase genes from wheat and rice. The present work focused on the cloning, expression and functional characterization of two xylosyltransferases from wheat and rice.
Figure 2.1 Structure of the basic repeating unit of xyloglucan along with the glycosyltransferases involved in its biosynthesis. β(1→4)glucose residues, forming the backbone of XyG, are linked together by β(1→4)glucan synthase. Xylosyl residues are added to the glucose backbone at the C-6 position, by α(1→6)xylosyltransferase. Galactose residues are added to xylose side chains at the C-2 position by β(1→2)galactosyltransferase. The terminal fucose residues are added to galactose side chains at the C-2 position by α(1→2)fucosyltransferase. The letters below the glucose residues (XXFG) represent the one-letter codes as proposed by Fry et al. (1993). The repetition of the XyG unit in the figure is represented by ‘n’.
Previous results in the lab

Identification of the putative xyloglucan xylosyltransferases in wheat and rice using bioinformatic approach

The putative XyG xylosyltransferases were identified by Dr. Ahmed Faik using an *in-house* java script, previously developed in our laboratory (Faik et al., 2006). The script performs several bioinformatic steps sequentially and gives a final list of hits. The first step performs a tBlastn (NCBI) search, which compares the protein query sequence against the translated nucleotide sequence databases. *Arabidopsis* xylosyltransferase 1 and 2 protein sequences were used as queries and TIGR wheat EST databases and rice genome databases were used for the search. The java script collects all the hits obtained from the blast search and removes the redundant hits. Then the script makes contigs with the unique hits, using the CAP3 assembly program (Huang and Madan, 1999). Low quality nucleotide sequences are removed while performing these steps.

For the phylogenetic analysis, another in-house script developed by Faik et al. (2006) was used. The script combines PHYLIP 3.65 (PHYlogeny Inference Package) package, PHYML (Phylogeny by Maximum Likelihood) algorithm and Tree View 1.6.6 for making the alignments and creating a phylogenetic tree with the given set of sequences. The script first makes PHYLIP alignments with the given set of protein sequences. The alignments were then submitted to Jones-Taylor-Thornton model (JTT) (Jones et al., 1992) or Dayhoff (Dayhoff et al., 1978) substitution frequency matrix as implemented in PHYLM algorithm. Then the Tree View software was used to view the phylogenetic tree. The Phylogenetic tree was created by Dr. Ahmed Faik (Fig 2.2). The
accession numbers for the putative xylosyltransferases that were included in this study are At1g74380: NM_106098; At1g18690: NM_101727; At2g22900: NM_127855; At4g37690: NM_148407; At5g07720: AJ245573; AtXT1: AF424587; AtXT2: AF069298; NbGT34-3: AJ864709; OsGT34-1: AK100984; OsGT34-2: AK110842; OsGT34-3: AK098859; OsGT34-4: AK105101; OsGT34-5: AK108140; OsGT34-6: NM_001074598; StGT34-2: AJ864713; TfGalTase: AJ245478; VvGT34: AJ864705; ZmGT34: AJ864711. All the wheat genes were cloned in the lab and annotated as TaGT34-1 to TaGT34-7 and no GenBank accession number is assigned to each of them because they are not published yet (Faik, Unpublished data).
Figure 2.2 Phylogenetic tree of the putative xylosyltransferases, members of the glycosyltransferase 34 family. The phylogenetic tree was created by Dr. Ahmed Faik, using a bioinformatic script (Faik et al., 2006). OsGT34-3 (*Oryza sativa* GT-34-3) and TaGT34-7 (*Triticum aestivum* GT-34-7) are the two putative xylosyltransferases having 80% identity to AtXT1 at the amino acid level. TfGalTase: *Trigonella foenum-graecum* galactosyltransferase; α(1,2)GalTase; α(1,2)galactosyltransferase from *Schizosaccharomyces pombe*; StGT34-2: *Solanum tuberosum* glycosyltransferase 34-2; VvGT34: *Vitis vinefera* glycosyltransferase 34; ZmGT34: *Zea mays* glycosyltransferase 34; NbGT34-3: *Nicotiana benthamiana* glycosyltransferase 34-3; At: *Arabidopsis thaliana*; Os: *Oryza sativa*. Boot strap values (%) from 100 trials were indicated at branch points. The scale indicates the branch length.
Materials and Methods

Clones, chemicals, and enzymes

The full-length cDNA clones for OsGT34-3 and TaGT34-7 were obtained from the Rice Genome Research Center (RGRC), Japan and from the Wheat EST collection of the Genome Canada/Genome Prairie/Genome Quebec program. Oligonucleotide primers were synthesized by Invitrogen (Carlsbad, CA). Platinum Pfx DNA polymerase, Escherichia coli TOP10 competent cells, Pichia pastoris expression system and expression vectors, Gateway cloning and expression vectors, LR clonase, BP clonase, Drosophila S2 expression system were from Invitrogen (Carlsbad, CA). Taq DNA polymerase was from New England Biolabs (Ipswich, MA). QIAPREP Spin miniprep kit and QIAquick gel extraction kit were from Qiagen (Valencia, CA). Whatman FTA cards and FTA purification reagent were from Fisher Scientific (Hampton, NH). The SuperSignal West Femto chemiluminescence system and Zeba spin columns were from Pierce (Rockford, IL). Complete protease inhibitor tablets lacking ethylene diamine tetraacetic acid (EDTA) were from Roche Applied Science (Indianapolis, IN). Immobilin-P (polyvinylidene difluoride) membranes were from Millipore (Billerica, MA). Celloligosaccharides, Dowex 1-X8 Cl ion-exchange resin, bradford reagent, and glass beads were from Sigma-Aldrich (St. Louis, MO). UDP-[\textsuperscript{14}C]xylose (9.76 GBq/mmol) was from PerkinElmer Life Sciences (Waltham, MA). Low range pre-stained protein standard was from Bio-Rad (Hercules, CA).
Cloning of his-tagged version of *OsGT34-3* and *TaGT34-7* genes

*OsGT34-3* and *TaGT34-7* were PCR-amplified from full-length cDNA clones with gene-specific primers having the His tag before the start codon. The primer sequences are as follows: *OsGT34-3*, 5′-ATGCATCATCATCATCATCACATGTGGGTGGCGGAGCGG-3′ (forward) and 5′-TCACGTAGAAGTAGTCTTCATG-3′ (reverse); and *TaGT34-7*, 5′-ATGCATCATCATCATCATCACATGTGGGTGGCGGAGCG-3′ (forward) and 5′-TCATGTAGTAGTCTTGAGGGC-3′ (reverse). PCR reactions were performed in 50µl final reaction volume containing 1X Pfx amplification buffer, 0.3mM dNTP's, 1mM MgSO₄, 0.3µM of each primer, 50-100ng of DNA template, 1.5 units of Pfx DNA polymerase. PCR was performed in a PxE thermal cycler (Thermo Electron Corporation, Milford, MA). PCR amplification conditions were, 94°C for 2 min; 30 cycles of 94°C for 15 sec, 52°C for 30 sec, and 67°C for 30 sec; final extension at 68°C for 3 min followed by 4°C storage. PCR products were cloned into Zero Blunt TOPO vector, and transformed into *E.coli* TOP10 competent cells. Bacterial colonies were grown overnight in 5ml LB media with 50mg/L of kanamycin at 37°C, in a shaker at 225 rpm. Plasmid DNA was then extracted from these cultures using the Qiagen miniprep kit according to the manufacturer’s protocol (Qiagen, Valencia, CA). The cloned PCR products were fully sequenced to verify the absence of PCR-induced point mutations. The *OsGT34-3* and *TaGT34-7* genes were released from TOPO vectors with EcoRI and ligated into EcoRI-linearized *Pichia* expression vector, pPICZ versions B or A, both of which are same except for a single restriction site variation between NotI and the myc epitope (Invitrogen Inc, California).
For cloning of OsGT34-3 and TaGT34-7 genes into Drosophila expression vectors, two different strategies were employed: for OsGT34-3, the DNA segment was released from the Zero Blunt TOPO vector by digestion with EcoRI and ligated into EcoRI-linearized pENTR vector. For TaGT34-7, which was in pCMV-SPORT6, a gateway recombination reaction was performed using BP clonase to transfer TaGT34-7 from pCMV-SPORT6 to pDONR vector. Finally, OsGT34-3 and TaGT34-7, both in gateway cloning vectors now, were recombined into the gateway expression vector, pMT-DEST48, by overnight LR clonase reactions according to manufacturer’s protocol (Invitrogen, Carlsbad, CA).

Expression of OsGT34-3 and TaGT34-7

Expression in Pichia pastoris

X-33 Pichia cells were prepared for transformation by starting a 5ml starter culture (from the glycerol stock), on the first day, in yeast extract peptone dextrose medium (YPD, 1% yeast extract, 2% peptone, 2% dextrose) in 50ml conical tube at 30°C overnight. On the second day, 250ml of fresh YPD media was inoculated with 0.5ml of X-33 overnight culture and grown overnight at 30°C. On the third day, X-33 cells were harvested by centrifugation at 1500 × g for 5 mins at 4°C. The cell pellets were re-suspended in 250ml of ice-cold sterile water and centrifuged again at 1500 × g for 5 mins at 4°C, and this step was repeated again by re-suspending the cell pellet in 125ml of ice-cold sterile water. In the next step, the cell pellet was re-suspended in 20ml of 1M
sorbitol, centrifuged, and finally the cell pellet was re-suspended in 0.5ml of 1M sorbitol. The cells were kept on ice and used on the same day for electroporation.

X-33 cells were electroporated with linearized TaGT34-7-pPICZ and OsGT34-3-pPICZ constructs as per the manufacturer’s directions. (Invitrogen, Carlsbad, CA). Briefly, 70µl of X-33 cells were mixed with 20µl (10µg DNA) of either TaGT34-7-pPICZ or OsGT34-3-pPICZ constructs in the ice-cold 0.2cm electroporation cuvette (Fisher Scientific, Hampton, NH). Electroporation conditions for pulsing the cells were 1.8 kilo volts, 200 ohms, 25 micropores. After electroporation 1ml of ice-cold 1M sorbitol was added to the cuvette and the cells were incubated at 30°C without shaking for 2 hours. The cells were then plated on yeast extract peptone dextrose medium plates (YPDS) [1% w/v yeast extract, 2% w/v peptone, 2% w/v dextrose, 1M sorbitol, 2% w/v agar] containing 100µg/ml of zeocin. The transformants were tested for the presence of the gene, by PCR reactions with gene-specific primers. Briefly, the colonies were grown in 5ml of buffered glycerol-complex medium (BMGY) [1% w/v yeast extract, 2% w/v peptone, 100mM potassium phosphate (pH 6.0), 1.34% w/v yeast nitrogen base, 4x10^-5% w/v biotin, 1% v/v glycerol] at 30°C overnight. Then 50-100µl of the overnight culture was applied to Whatman FTA cards, allowed the sample to dry, and punched a disc out of the sample on the FTA card. The disc was cleaned several times with FTA purification reagent, as per the manufacturer’s protocol (Fisher Scientific, Hampton, NH), and was used directly in the PCR reaction. PCR reactions were performed in 20µl final reaction volume containing 2µl of 10 X standard Taq buffer, 200µM each dNTP, 0.3µM of each primer, disc from FTA card, 1.0 units of Taq DNA polymerase. PCR amplification
conditions were, 95°C for 30 sec; 25 cycles of 95°C for 30 sec, 52°C for 1 min. 30 sec, and 72°C for 1 min. 45 sec; final extension at 72°C for 10 min followed by 4°C storage.

For the induction of the expression, three positive transformants were selected for each gene construct and grown in 15ml of BMGY at 30°C until OD$_{600}$ reaches was between 2-6. To induce the gene expression, *Pichia* cells were harvested by centrifugation (3000 × g, 5 min) and re-suspended in buffered methanol-complex medium (BMMY) [1% w/v yeast extract, 2% w/v peptone, 100mM potassium phosphate (pH 6.0), 1.34% w/v yeast nitrogen base, 4x10$^{-5}$% w/v biotin, 0.5% v/v methanol] to an OD$_{600}$ = 1 (~40ml) in a 250-ml baffled flask. The induction was initiated by adding 0.5% (v/v) methanol; and during the induction period methanol was also added every 24 h to a final concentration of 0.5% (v/v). To test the efficiency of induction, 1ml of samples were taken after second and fourth day of induction for the detection of proteins in western blot. After four days of protein induction, the enzymes were extracted as described by Faik et al. (2002). Briefly, the enzyme fractions were prepared by vortexing *Pichia* cells for seven times (30 sec each) in the extraction buffer (EB) [100mM Heps-KOH pH 7.0, 0.4M sucrose, 0.1% bovine serum albumin (BSA), 1mM EDTA, 1mM dithiothreitol (DTT), 5mM MgCl$_2$, 5mM MnCl$_2$, 1mM phenylmethanesulphonylfluoride (PMSF), one complete protease inhibitor cocktail tablet/50ml of EB] containing 1.5ml of acid-washed glass beads (425-600µm), already re-suspended at 50% (v/v) in distilled water. The cells were kept on ice between each vortex to prevent the degradation of the protein. After breaking the cells by vortexing, the debris was collected by centrifugation (14000 × g, 10 min, 4°C). The debris pellets were again re-suspended in EB containing
the glass beads and 1% Triton X-100, vortexed for four times (30 sec each), centrifuged and the resulting supernatant was used immediately for enzyme assays or stored at -20°C.

**Expression in *Drosophila S2* cells**

*Drosophila S2* (Schneider 2) cells were cotransfected with the pCoBlast vector (1µg) and pMT-DEST48 (19µg) containing *OsGT34-3* and *TaGT34-7*. Selection of cell lines stably expressing *OsGT34-3* and *TaGT34-7* was done with blasticidin following the supplier’s instructions (Invitrogen, Carlsbad, CA). Copper sulfate was added to a concentration of 500µM to induce protein expression by transformed cell lines. Twenty-four hours after the addition of copper sulfate, cells were harvested by centrifugation (250 × g, 10 min, 4°C) and used for enzyme extraction or stored at -80°C. The cell pellets were suspended in ice-cold extraction buffer with 1% Triton X-100. The cells were disrupted by sonication, 3 × 15 seconds, with storage on ice between each treatment. Cells were then centrifuged (16060 × g, 5 min, 4°C), the supernatant was collected and used immediately for enzyme assays or stored at -20°C.

**Bradford protein assay**

Protein concentration in the enzyme extracts was quantified using the Bradford method (Bradford, 1976). Briefly, 0.1ml of protein sample was incubated with 3ml of bradford reagent (Sigma-Aldrich, St.Louis, MO) for 30 min and the absorbance values were measured at 595nm. To generate a standard curve, the absorbance measurements of
the BSA protein standards ranging from 0.1-4mg/ml were also taken. The protein concentration of the unknowns was calculated from the standard curve.

**SDS-PAGE and Western blot**

The enzyme extracts obtained from the *Pichia* and *Drosophila* S2 cell cultures were separated by SDS-PAGE, which was performed as described by Laemmli (1970), and Sambrook and Russel, (2001) with minor modifications. To the enzyme extracts, equal amounts of laemmli buffer (50mM Tris-HCl pH6.8, 100mM DTT, 2% w/v SDS, 0.1% bromophenol blue, 10% v/v glycerol) was added and boiled in hot water bath (100°C) for 10 mins to denature the proteins. 20µl (2-5µg of protein) of boiled samples were directly loaded into 10% SDS-PAGE gels. Electrophoresis and transfer of proteins were performed using the Mini-PROTEAN 3 Electrophoresis Cell (Bio-Rad, Hercules, CA). Proteins were separated in the electrophoresis cell, filled with running buffer (1.92M glycine, 0.25M tris base, 1% w/v SDS), at constant volts (120V) for 1 hour. After separation, proteins were transferred to polyvinylidene difluoride membranes (PVDF) in transfer buffer (25mM tris, 192mM glycine, 20% v/v methanol, pH 8.3) for 1 hour at 200mA with stirring. PVDF membrane was wet in 100% acetone before use. Membranes were blocked with a 5% milk solution and probed with (1:5000 dilution) 6×His monoclonal antibody (Clontech laboratories, Mountain View, CA). After three cycles of washing (15 min each wash) with 0.1% tween-20 in phosphate buffered saline (PBS) [136mM NaCl, 1.5mM KH₂PO₄, 2.7mM KCl, 1mM Na₂HPO₄, pH 7.5], membranes were probed with horseradish peroxidase conjugated goat anti-mouse
antibodies and developed using the chemiluminescence system (Pierce, Rockford, IL). Protein bands were compared to low-range pre-stained SDS-PAGE standards (Bio-Rad, Hercules, CA).

**Xylosyltransferase assay**

The xylosyltransferase enzyme assay was performed as described earlier (Faik et al., 2002). The enzyme fractions were supplied with UDP-[\(^{14}\)C]Xyl and a cellohexaose oligosaccharide as acceptor, which has six \(\beta(1\rightarrow4)\) linked glucose residues (GGGGGG), thus resembling the backbone of XyG. The reaction mixture contains 40\(\mu\)l (~8\(\mu\)g of protein) of the enzyme extracts, 10\(\mu\)M UDP-[\(^{14}\)C]xylose (65,000 cpm), and 10mM cellohexaose acceptor (100\(\mu\)g). The reactions were incubated at room temperature for 1 hour and terminated by adding 500\(\mu\)l of Dowex 1X8-100 Ion Exchange resin (1:1 resin/water) to complex the unused UDP-[\(^{14}\)C]xylose. The reaction slurries was loaded onto Zeba spin columns and spun in a microcentrifuge at 16060 \(\times\) g for 30 seconds. The column effluent was mixed with the scintillation fluid and the radioactivity was measured on a Beckman Coulter LS-6500 Multipurpose scintillation counter. Control reactions did not contain the cellohexaose acceptor.
Results

Bioinformatic analysis of the putative xyloglucan xylosyltransferases from wheat and rice

In order to have more information on the putative rice and wheat xylosyltransferase genes, *OsGT34-3* and *TaGT34-7*, their predicted amino acid sequences were examined for the presence of characteristics specific to the GT-34 family. Both the wheat and rice xylosyltransferases were predicted, by TMHMM (Transmembrane Hidden Markov Model, Krogh et al., 2001), to have a single TMD in the amino terminal region, consistent with the observation that Golgi-localized GTs are type II integral membrane proteins (Fig. 2.3 and 2.4). The protein sequences of putative rice and wheat xylosyltransferases have 80% and 79% identity to the *Arabidopsis* AtXT1 protein, respectively. They both have a DXD motif (where D refers to aspartic acid and X represents any amino acid), which is necessary for substrate donor (NDP-sugar) binding (Wiggins and Munro, 1998), and they also have GT-34 specific motif (Fig. 2.4). An uncleavable signal anchor, at the N-terminus, was predicted for both the wheat and rice proteins by signalP program (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al., 2004) and DGPI algorithm, which can identify signal peptide and GPI anchor addition sequences (http://129.194.185.165/dgpi/DGPI_demo_en.html). The signal anchor covers most of the sequence that forms the TMD in both wheat and rice proteins. Also, both proteins were predicted to have a single N-glycosylation site at the C-terminus, at amino acid position 421 for OsGT34-3 and at 426 for TaGT34-7, according to motif scan algorithm (http://myhits.isb-sib.ch/cgi-bin/motif_scan). Therefore, both the wheat and
rice proteins are targeted to the secretory pathway and enter the endoplasmic reticulum (ER) where they are N-glycosylated.

For sub-cellular localization predictions, WoLF PSORT (http://wolfpsort.org/) (Horton et al., 2007), TargetP (http://www.cbs.dtu.dk/services/TargetP/) (Emanuelsson et al., 2007), BaCelLo (http://gpcr.biocomp.unibo.it/bacello/) (Pierleoni et al., 2006), LOCtree (http://cubic.bioc.columbia.edu/cgi-bin/loc/nair/loctree/query) (Nair and Rost, 2005), and PA-SUB (http://www.cs.ualberta.ca/~bioinfo/PA/Sub/) (Lu et al., 2004) algorithms were used. None of these algorithms predicted the localization of rice and wheat proteins as Golgi residents. However, it is known that all the GTs are type II membrane proteins and should be localized in the Golgi. Therefore, AtXT1, which is a known XyG xylosyltransferase, was run through all the algorithms but similar results were obtained. This shows that none of the sub-cellular localization prediction algorithms were able to provide a reliable prediction of the localization of cell wall GTs, which was also observed in the work done by Egelund et al. (2004). However, LOCtree predicted the Golgi localization for both the wheat and rice proteins but this algorithm was designed specifically for non-membrane proteins while the proteins analyzed, TaGT34-7 and OsGT34-3, are membrane proteins. Therefore, the LOCtree prediction for sub-cellular localization is also not reliable. But previous biochemical and electron microscopy studies provide sufficient evidence for the Golgi retention of the GTs (Moore et al., 1991). Munro’s (1991) work showed that the TMD and the peptide regions flanking it were responsible for targeting the GTs to the Golgi.
Figure 2.3 The hydropathic plot of TaGT34-7 (A) and OsGT34-3 (B) according to the method of Kyte and Doolittle (1982). The predicted transmembrane domain in both the proteins is indicated by TMD. The red line indicates the grand average of hydropathicity (GRAVY) value, which is the average hydropathy score, of the proteins. The peaks above the red line indicate the transmembrane regions and below the red line indicate the surface regions. The figure was generated using the software available in the website: http://gcat.davidson.edu/rakarnik/kyte-doolittle.htm
Figure 2.4 Amino acid sequence alignment, using T-Coffee algorithm (Notredame et al., 2000), of the putative wheat and rice xylosyltransferases with AtXT1 and AtXT2. The alignment was imported into GeneDoc version 2.6 for shading (Nicholas et al., 1997). Black shading represents the conserved amino acids found in the *Arabidopsis*, wheat and rice xylosyltransferases. The glycosyltransferase 34 specific motif is represented by a red line and the DSD motif by a green line. The amino acid residues within the blue box form the transmembrane domain in the *Arabidopsis*, wheat and rice xylosyltransferases.
Sub-cloning of the putative wheat and rice xylosyltransferases into *Pichia pastoris* and *Drosophila* S2 expression vectors

The full length rice and wheat cDNA clones were PCR amplified, from the original clones obtained from RGRC and Genome Canada program respectively, using gene specific primers carrying the nucleotide sequence for His tag before the start codon. The PCR products were cloned into Blunt TOPO vector and then transferred to *Pichia pastoris* expression vector, pPICZ, by EcoRI digestion (Fig. 2.5a). For cloning into *Drosophila* expression vector, two different cloning strategies were employed for *TaGT34-7* and *OsGT34-3*.

For *TaGT34-7*, which was in pCMV-SPORT6, the gene was transferred to pDONR vector by a gateway recombination reaction catalyzed by BP clonase enzyme mix (Fig. 2.5b). For *OsGT34-3*, an EcoRI digestion facilitated the transferring of gene fragment from the Blunt TOPO vector to the linearized pENTR vector (Fig. 2.5b). Finally, the wheat and rice genes in pDONR and pENTR were transferred to pMT-DEST48 by a gateway recombination reaction catalyzed by LR clonase enzyme mix (Fig. 2.5b). The final wheat and rice expression vector constructs were shown in figures 2.6 and 2.7.
Figure 2.5 Strategies used for cloning *TaGT34-7* and *OsGT34-3* into (A) *Pichia pastoris* and (B) *Drosophila* S2 expression vectors. (A) Wheat and rice cDNAs were PCR amplified to add the His tag before the start codon. The PCR products were introduced into Blunt TOPO vector and then transferred to pPICZ by EcoRI digestions. (B) *TaGT34-7* was transferred from pCMV-SPORT6 to pDONR, to create a donor construct, by using BP clonase enzyme mix that catalyzes site-specific recombination, developed by Hartley et al., 2000, between attB1-attP1 and attB2-attP2 sites. *OsGT34-3* was released from Blunt TOPO vector by an EcoRI digestion and ligated with the EcoRI linearized pENTR to create a donor constructs. The genes in the donor constructs were transferred to the expression vector, pMT-DEST 48, by a gateway recombination reaction using LR clonase enzyme mix, which catalyzes site-specific recombination (Hartley et al., 2000) between attL1-attR1 and attL2-attR2 sites. AOX, Alcohol oxidase. MT, metallothionein.
Figure 2.6 Pichia expression vector constructs used for the expression of TaGT34-7 (A) (Triticum aestivum xylosyltransferase) and OsGT34-3 (B) (Oryza sativa xylosyltransferase) in Pichia pastoris. TaGT34-7 and OsGT34-3 were PCR amplified and introduced into Blunt TOPO vector. The genes were dropped from the Blunt TOPO plasmids by EcoRI digestion and ligated into the EcoRI linearized Pichia expression vectors to produce the expression constructs. Both the genes, TaGT34-7 and OsGT34-3, will be driven under the control of AOX1 (alcohol oxidase 1) promoter, which will be induced by methanol for protein expression. The restriction site and primer sites used for deriving the genes were indicated in the figure. The figure was generated by using vector NTI software (Invitrogen Inc, Carlsbad, CA).
Figure 2.7 Drosophila expression vector constructs used for the expression of TaGT34-7 (A) (Triticum aestivum xylosyltransferase) and OsGT34-3 (B) (Oryza sativa xylosyltransferase) in Drosophila S2 cells under the control of pMT, metallothionein promoter, which is induced by copper sulfate. Primer sites used for deriving the genes were indicated in the figure. attB1 and attB2 are the site specific recombination sites in pMT-DEST48 vector. TaGT34-7 was transferred from pDONR to pMT-DEST48 (A) and OsGT34-3 was transferred from pENTR to pMT-DEST48 (see Fig. 2.3b). Both the reactions were site-specific recombination reactions catalyzed by LR clonase enzyme mix. The figure was generated by using vector NTI software (Invitrogen Inc, Carlsbad, CA).
Optimization of the expression of wheat and rice putative xylosyltransferases in

*Pichia pastoris* and *Drosophila S2* expression systems

The wheat and rice putative xylosyltransferases were integrated into the genomes of *Pichia* and *Drosophila* cells by genetic recombination (See material and methods for details). The putative xylosyltransferases were expressed under the control of alcohol oxidase promoter in *Pichia*; and metallothionein promoter in *Drosophila S2* cells. The integration and expression of the genes was verified by western blot analyses using anti-6×His antibody (Fig. 2.8). One milliliter of induced *Pichia* cells was collected after 2 days and 4 days of induction and used to visualize protein production by western blot. TaGT34-7 and AtXT1 showed the same level of expression after 2 and 4 days of induction. But the expression of OsGT34-3 was higher after 2 days of induction and decreased after 4 days of induction. In *Pichia* cells, both of the proteins including the control, AtXT1, showed the expected band size of 54 kD (Fig. 2.8a), whereas OsGT34-3 expressed in *Drosophila S2* cells showed somewhat larger size i.e., 60.2 kD (Fig 2.8b), which was calculated from the molecular mass standard curve (Fig. 2.8c).
**Figure 2.8** Heterologous expression of TaGT34-7, OsGT34-3 and AtXT1 in (A) *Pichia pastoris* and (B) *Drosophila* S2 cells. (A) Western blot analysis of Triton X-100 solubilized extracts from three *Pichia* cell colonies, expressing TaGT34-7, OsGT34-3 and AtXT1, collected after 2 days (d2) and 4 days (d4) of induction with methanol. AtXT1, TaGT34-7 and OsGT34-3 proteins were detected using a 6×His monoclonal antibody. *Pichia* cells transformed with AtXT1 were used as a positive control. All the three xylosyltransferases, AtXT1, TaGT34-7 and OsGT34-3, showed the expected protein of size, 54 kD. (B) and (D) Western blot analysis of Triton X-100-solubilized S2 cell extracts, expressing AtXT1 and OsGT34-3 respectively, collected after 24 hours of induction with copper sulfate. Sizes of the protein markers are labeled on the left. (C) and (E) Molecular mass standard curves, for the figures 2.8b and 2.8d respectively, was prepared by plotting the retention factor (Rf) values against the log of molecular weights for the standard proteins. The molecular masses of the proteins, AtXT1 and OsGT34-3, were calculated from the standard curve. AtXT1 is 53.3 kD and OsGT34-3 is 60.2 kD.
**Xylosyltransferase activity in *Pichia* and *Drosophila* extracts**

The enzymes extracted from *Pichia* and *Drosophila* S2 cells were used for xylosyltransferase enzyme assay as developed by Faik et al. (2002). The enzyme fractions (~2-5µg protein) were supplied with the UDP-[\(^{14}\text{C}\)]Xyl substrate and cellohexaose acceptor (GGGGGG), which resembles the glucose backbone in XyG. The reaction mixture was incubated at room temperature and if the enzyme has the xylosyltransferase activity, the radioactivity associated with the xylose will be transferred to the cellohexaose acceptor as shown in the reaction below.

\[
\text{Cellohexaose + UDP-}^{[14}\text{C}]\text{xylose} \xrightarrow{\text{Xylosyltransferase activity}} \text{Cellohexaose-}[^{14}\text{C}]\text{xylose + UDP + Excess UDP-}\quad^{[14}\text{C}]\text{xylose}
\]

The excess of radioactive UDP-Xyl precursor in the reaction mixture was removed by adding, Dowex 1X8-100, an ion-exchange resin at the end of the incubation period. The resin associates with the UDP portion of the UDP-Xyl that is not used in the reaction. The radioactivity incorporated onto cellohexaose was measured by using a scintillation counter. TaGT34-7 and OsGT34-3 did not show the xylosyltransferase activity in *Pichia* expression system (Fig. 2.9), which could be due most likely to the differences in the post-translational modifications between plants and yeasts (see discussion). Both the wheat and rice proteins and AtXT1 showed a very low activity in *Drosophila* S2 cells (Fig. 2.10) but the enzyme activity can be improved by optimizing the conditions of expression and the enzyme assay.
Figure 2.9 XyG-xylosyltransferase activities in detergent-soluble extracts from Pichia pastoris expressing TaGT34-7 and OsGT34-3. Screening for activity in three colonies for each clone A) TaGT34-7 and B) OsGT34-3. The cells were grown under induction conditions for four days before solubilization in extraction buffer containing 1% Triton X-100 and then the soluble fraction (~2-5µg of protein) was used as the source of the enzyme. Triton X-100 solubilized Pichia cell extracts expressing either TaGT34-7 or OsGT34-3 were assayed using cellohexaose, which has six β(1→4) linked glucose residues, as the acceptor. All experimental reactions (■) were incubated for 1.5 hours in the presence of 65,000 cpm UDP-[14C]Xyl, 10mM cellohexaose acceptor substrate. Control reactions (□) lacked the acceptor substrate. The values are the average of two assays, and the bars show SE.
Figure 2.10 XyG-xylosyltransferase activity in detergent-soluble extracts from *Drosophila* S2 cells expressing TaGT34-7, OsGT34-3 and AtXT1. The transfected S2 cells expressing the xylosyltransferases were induced with copper sulfate, solubilized in extraction buffer containing 1% Triton X-100 and the soluble fraction (~2-5µg protein) was used as the source of the enzyme. Triton X-100 solubilized *Drosophila* S2 cell extracts expressing AtXT1 or TaGT34-7 or OsGT34-3 were assayed using cellohexaose, which has six β(1→4) linked glucose residues, as the acceptor. All experimental reactions (■) were incubated for 1.5 hours in the presence of 65000 cpm UDP-[\(^{14}\)C]Xyl, 10mM cellohexaose acceptor substrate. Control reactions (□) lacked the acceptor substrate. The values are the average of two assays, and the bars show SE.
Discussion

XyG is the major hemicellulosic polysaccharide in the cell walls of dicots and non-graminaceous monocots. Graminaceous monocots, which include wheat, rice, oat and barley, contain only ~5% of XyG in their cell walls. The genes encoding all the enzymes involved in the XyG biosynthesis in dicots have been identified, but none of these studies focused on the XyG biosynthesis in monocots. The comparison between the biosynthetic mechanism in dicots and monocots would be very informative and would allow crop improvement or substitution of polymers in cereal cell walls. Therefore, the specific aim of this project was to clone and express the XyG biosynthetic genes, particularly the XyG xylosyltransferase genes from wheat and rice, in eukaryotic expression systems such as *Pichia pastoris* and *Drosophila* S2 expression systems.

The putative wheat and rice XyG xylosyltransferases were predicted to have an uncleavable signal anchor at the N-terminus, which targets the proteins to the secretory pathway. Therefore, both the proteins are predicted to enter the endoplasmic reticulum and N-glycosylated at target asparagines. The consensus sequence for N-glycosylation, asparagine-glutamic acid-threonine (Asn-Glx-Thr), is present at amino acid position 421-423 and 426-428 in rice and wheat proteins respectively. According to the biochemistry and electron microscopy studies, all the GTs identified to date are Golgi resident proteins; however none of the sub-cellular localization algorithms currently available (WoLF PSORT, TargetP, PA-SUB, BaCelLo) was able to predict Golgi localization for both the wheat and rice proteins. Cytoplasmic or mitochondrial localization were predicted for both the wheat and rice proteins and the same result was obtained when the control, AtXT1, was run through all the algorithms. Therefore, the sub-cellular prediction
algorithms did not give a reliable prediction of the localization of the putative xylosyltransferases. The same observation was reported by Egelund et al. (2004), who used TargetP algorithm for predicting the sub-cellular localization of six known GTs, including AtXT1. However, both the wheat and rice proteins have the GT-34 specific motif, which confirms that they are GTs. Amino acid sequence alignment of the putative wheat and rice xylosyltransferases with AtXT1 and AtXT2 showed ~80% identity between all of them. Bioinformatic analyses revealed that the putative xylosyltransferases have a single TMD at the amino terminal region, consistent with the fact that all the GTs involved in the cell wall polysaccharide biosynthesis are type II integral membrane proteins. Moreover, they have a DXD motif, which is required for substrate donor (NDP-sugar) binding.

To confirm our hypothesis that the putative wheat and rice genes, TaGT34-7 and OsGT34-3, might encode the XyG xylosyltransferase enzymes, we used a biochemical approach that should give a direct evidence of their biochemical function. Thus, cDNAs encoding the putative XyG xylosyltransferases from wheat (TaGT34-7) and rice (OsGT34-3) were expressed in Pichia pastoris and Drosophila S2 expression systems. Both proteins expressed in Pichia pastoris did not show the xylosyltransferase activity, although the proteins were present in the extract as demonstrated by western blot (Fig. 2.9). However, when expressed in Drosophila S2 cells, both enzymes showed very low activity (Fig. 2.10). Two possibilities could explain the loss of activity in Pichia cells. First, the differences in the post-translational modifications in higher and lower eukaryotes, particularly the differences in the N-glycosylation patterns, might have affected the functionality of the heterologous proteins produced in Pichia. Both in higher
and lower eukaryotes, N-glycosylation begin in the endoplasmic reticulum by the transfer of a lipid-linked oligosaccharide unit, Glc$_3$Man$_9$GlcNAc$_2$ (Glc, Glucose; Man, Mannose; GlcNAc, N-acetylglucosamine), to asparagine in the recognition sequence, Asn-X-Ser/Thr (Asn, Asparagine; X represents any amino acid; Ser, serine; Thr, Threonine) (Patrick et al., 2005). In higher eukaryotes, the oligosaccharide unit is trimmed in ER and Golgi generating the oligosaccharides of high mannose type (Man$_{5-6}$GlcNAc$_2$), or a mixture of different sugars (complex type) or a combination of both (Patrick et al., 2005). But the most common oligosaccharides produced by Pichia pastoris is comprised of Man$_8$GlcNAc$_2$ or Man$_9$GlcNAc$_2$ (Montesino et al., 1998). Though Pichia pastoris does not hyperglycosylate the foreign proteins as is the case with Saccharomyces cerevisiae, it does add a variety of oligosaccharides and larger polymers different from that of higher eukaryotes (Daly and Hearn, 2005). The longer oligosaccharides attached during N-glycosylation in Pichia might have interfered with folding and/or catalytic/binding sites of the heterologous proteins.

Second, the amount of the proteins produced might not be enough to show the enzyme activity, which can be avoided by scaling up the expression in larger volumes. Alternatively, multicopy recombinants of Pichia cells (more than one gene copy in a single cell) can be selected for protein expression. Previous studies have shown that increasing the gene dosage can significantly increase the protein expression (Clare et al., 1991a; Clare et al., 1991b).

Faik et al. (2002) noted that AtXT2, which is 89% identical to AtXT1, also did not show the enzyme activity in Pichia cells but it showed the enzyme activity when Cavalier and Keegstra (2006) expressed it in Drosophila S2 expression system.
Therefore, the putative wheat and rice xylosyltransferases were also expressed in *Drosophila* S2 expression system. The preliminary assays with the proteins expressed in *Drosophila* S2 cells showed a very low cellohexaose-dependant xylosyltransferase activity (Fig. 2.9) but more work is needed to confirm the enzyme activity of the putative xylosyltransferases. Future work can be directed towards optimizing the protein expression in *Drosophila* S2 cells and enzyme assays for testing xylosyltransferase activities of TaGT34-7 and OsGT34-3. Though the results from the present work did not provide sufficient information about the function of TaGT34-7 and OsGT34-3, further work might provide a definite evidence for their role in XyG biosynthesis. One possibility is that XyG xylosyltransferases from grasses may have different requirements or strictly dependant on glucan synthase activity.

Cell wall modifying enzymes such as XyG endotransglycosylase/hydrolases (XTH) are involved in the re-organization of the cellulose-XyG network by catalyzing the cleavage and religation of XyG chains, which helps in the cell wall loosening (Yokoyama and Nishitani, 2001). The *XTH* gene family in *Arabidopsis* has 33 members, which is not so surprising because of the abundance of XyG in type I walls (Yokoyama and Nishitani, 2001). Yokoyama et al. (2004) compared the *XTH* gene family in *Arabidopsis* and rice representative of type I and type II walls respectively. Surprisingly, rice also has an *XTH* gene family consisting of 29 members, most of which showed growth stage specific and organ-specific expression patterns (Yokoyama et al., 2004). A wheat gene encoding the XyG endotransglycosylase (XET) was also previously identified (Okazawa et al., 1993). The presence of an *XTH* gene family in rice, which is similar to the size of *Arabidopsis* *XTH* gene family, and an *XET* gene in wheat suggests that XyG
has an important structural function despite its low content in grasses (Yokoyama et al., 2004). Therefore, to unravel the function of XyG in grasses, it is important to first understand its biosynthesis in grasses.
Chapter 3: IDENTIFICATION OF ENDOSPERM DEVELOPMENTAL PHASES IN WHEAT SEEDS
Introduction

Cell walls of grasses, including commercially important cereals such as rice, wheat and barley, are rich in AXs and (1→3),(1→4)-β-glucan, also called mixed linked glucans (MLG), in addition to cellulose microfibrils (Carpita and Gibeaut, 1993). AXs are important for plant development as they cross-link the cellulose microfibrils forming a network, which maintains the structural integrity of the cell (Carpita and Gibeaut, 1993). Also, hydrolysis of AXs is an important prerequisite for the utilization of cereal hemicelluloses in the ethanol fermentation industry (Gray et al., 2006). Despite the functional and economical importance of AXs, none of the genes involved in their biosynthesis have been identified. Thus, research in our lab is mainly directed towards identifying the genes involved in AX biosynthesis. Endosperm cell walls of wheat have ~70% (w/w) AXs (Izydorczyk and Biliaderis, 1995) making the developing endosperm a good model system to study AX biosynthesis. It has been shown that AX is deposited during the differentiation phase of the endosperm development in wheat (Philippe et al., 2006b). The goal of this project was to identify wheat endosperm developmental phases namely syncitial, cellularization, differentiation and maturation phases, which would help in focusing further research on specific stages of wheat grain at which the genes involved in the AX may be highly expressed.

Endosperm development in cereals

The life of an endosperm starts with double fertilization, which is a unique characteristic of higher plants (Chaudhury et al., 1998). Double fertilization was
discovered almost a century ago during which one sperm nucleus fuses with the egg to form a zygote whereas the other fuses with two polar nuclei in the central cell to form a triploid nucleus called the primary endosperm nucleus (Olsen, 2001). Endosperm in cereals serves as a major storage organ and it follows nuclear development (Olsen, 2001; Wilson et al., 2006), which involves four major developmental phases: syncitial, cellularization, differentiation and maturation phases (Fig. 3.1) (Rogers and Quatrano, 1983; Lopes and Larkins, 1993; Berger, 1999). During syncitial phase, the triploid endosperm nucleus undergoes several rounds of mitosis without cytokinesis resulting in the formation of a structure called the multinuclear coenocyte, which contains numerous nuclei surrounding the central cell (Fig. 3.2a and 3.2b) (Olsen, 2001). During the cellularization phase, formation of the endosperm cell walls is initiated by the radial microtubular structures arising from the nuclear membranes of the coenocyte (Fig. 3.2c) (Olsen, 2001). The radial microtubular structures arising from the adjacent nuclei will attach to form a phragmoplast, which later results in the formation of cell plate and finally the cell wall (Fig. 3.2d) (Olsen, 2001). The nuclei then divide periclinally to form multiple layers of cell files closing the central vacuole (Fig. 3.2e and 3.2f) (Olsen, 2001). During the differentiation phase, each endosperm cell grows in volume due to the accumulation of starch and water; and the protein content of the cells increases during this phase (Olsen, 2001). During the maturation phase, cell expansion stops, starch and protein accumulation ceases, and the development of the endosperm is said to be complete (Olsen, 2001).
Figure 3.1 Time course of endosperm developmental phases in the wheat grain after anthesis. The four endosperm developmental phases (syncitial, cellularization, differentiation and maturation) along with their main characteristic features are indicated. °D represents the temperature/energy cumulated per day by the grain after anthesis. Days after anthesis are indicated on the top. The figure was adapted from Philippe et al., 2006a.
Figure 3.2 Diagrammatic representation of the endosperm cellularization process (A) Triploid endosperm nucleus (orange) in the central cell. (B) Multinuclear coenocyte formed by the several mitotic divisions of the endosperm nucleus without the formation of cell walls. (C) Formation of radial microtubular structures arising from the endosperm nuclei. (D) Interaction of the radial microtubular structures from each nucleus facilitated the formation of cell walls. (E) and (F) Periclinal divisions of the endosperm nuclei along with wall formation resulting in the closure of the central vacuole, thus completing the cellularization process. The figure was adapted from Olsen, 2001.
Cell wall components of the wheat endosperm

Endosperm cell walls in wheat consist of AX 70% (w/w), MLG 20% (w/w), and cellulose 2-4% (w/w) (Izydorczyk and Biliaderis, 1995; Bacic and Stone, 1981). AX in the endosperm cell walls of wheat consists of β-(1→4)linked xylose backbone substituted with α-L-arabinofuranose residues at C-2 and/or C-3 position (Fig. 3.3), and the arabinofuranosyl residues can be esterified with ferulic acid at C-5 position (Izydorczyk and Biliaderis, 1995). Structural changes of AX, as indicated by the difference in arabinose-to-xylose ratios, are seen in different parts of the grain (Andersson et al., 1994; Delcour et al., 1999).
Figure 3.3 Schematic representation of the putative structure of arabinoxylan along with the glycosyltransferase activities involved in its biosynthesis. The β(1,4)xylan backbone is synthesized by the xylan synthase, and the arabinosyl side chains addition is catalyzed by α(1,2)- and/or α(1,3)-arabinofuranosyltransferase activities.
Deposition of cell wall polysaccharides during endosperm development in wheat

Temporal and spatial deposition of cell wall polysaccharides is seen during endosperm development in wheat (Guillon et al., 2004; Philippe et al., 2006b). Immunogold labeling experiments showed that (1→3)-β-D-glucans were deposited during the initial stages of endosperm cellularization, and the (1→3)-β-D-glucan epitope disappeared after the endosperm was cellularized (Philippe et al., 2006b). At the final stages of endosperm cellularization, MLG and arabinogalactan proteins were seen in the endosperm cell walls (Philippe et al., 2006a; Philippe et al., 2006b). During the differentiation phase, the MLG epitope was mostly seen in the aleurone layer surrounding the endosperm, and at the maturation stage, MLG deposition in the starchy endosperm cell walls depended on the cell type with the most deposition in the central cells of the endosperm (Philippe et al., 2006b). AX, which constitutes about 70% (w/w) in the endosperm cell walls of wheat, was also deposited during the differentiation phase (Philippe et al., 2006b). According to these studies, AX was deposited in a highly substituted form at the differentiation phase and as the grain matures, the degree of substitution decreased (Philippe et al., 2006a; Philippe et al., 2006c), probably by the action of hydrolases, which facilitates interactions of AX with other cell wall polysaccharides (Izydorczyk and McGregor, 2000).

Identification of the endosperm developmental phases and knowing the specific time of the deposition of cell wall polysaccharides is an important key step towards identification of the genes involved in cell wall polysaccharide biosynthesis in the endosperm. Therefore, the goal of the present study was to use light microscopy and
sectioning to identify the endosperm developmental phases during wheat grain development.

**Materials and Methods**

**Materials**

Safefix II (catalog no. 23-042-600), xylene (catalog no. X3P), embedding rings (catalog no. 22-038197), paraffin (catalog no. T555), and toluidine blue O (catalog no. T161) were from Fisher Scientific (Hampton, NH). Microscopic glass slides and cover slips were from Corning (Corning, NY). Microtome knives were from Extremus, CL. Sturkey, Inc (Lebanon, PA). Histomount was from National Diagnostics (catalog no. HS-103) (Atlanta, GA). The pots used for transplantation were from Nursery supplies co (Chambersburg, PA). *Triticum aestivum*, spring wheat, was used for this study.

**Growing of wheat plants**

Wheat plants were grown in the greenhouse where the temperature was maintained at 20°C during the night and the day temperatures varied (~22-26°C) according to the outside temperature conditions. The seeds were sown in small pots initially, and were allowed to germinate and grow up to three weeks. The seedlings were then transplanted to larger pots and supplied with a continuous release of fertilizer, Osmocote classic, containing 14-14-14 of N: P: K. Water was supplied to the plants.
through a drip irrigation system. Under these conditions the wheat seedlings took ~ 2.5 - 3 months to grow to full size and produce the spikelets required for experimentation.

**Grouping of wheat seeds and preparation of seeds for light microscopy**

Developing wheat seeds were collected and sorted by weight (mg/seed) into different stages, as shown in table 3.1. Spikelets were maintained at 4°C while sorting the seeds into different stages. For fixation, fresh seeds were submerged in Safefix II (15-20 seeds/50ml of Safefix II) for one day at 4°C and then dehydrated in graded ethanol series (30%, 50%, 60%, 70%, 80%, 90%, 95%, and 100%, for 2 hours each). After dehydration the seeds were infiltrated in xylene:paraffin solution (1:1 ratio) at 65°C for 1 hour, followed by three baths of 100% paraffin, maintained at 65°C, for 5-7 days. The infiltrated seeds were then embedded individually in blocks filled with paraffin.

**Sectioning and Microscopy**

Individual seed in paraffin blocks were transversely and longitudinally cut (20µm thickness) using a rotary microtome (American Optical Co, Buffalo, NY). The sections were collected onto glass slides and dried on a hot plate, maintained at 55°C, overnight. The dried sections were dehydrated in graded ethanol series (100%, 90%, 70%, 40%, and 30%), for 3 min each, followed by rinsing in distilled water and stained with 0.05% (w/v) toluidine blue solution for 3 min, re-hydrated and then mounted in a mounting media, histomount. The mounted sections were viewed under Nikon EFD-3 microscope. Unless otherwise specified all the sections were viewed under 2X magnification. Sections were
viewed using SPOT Advanced software (Diagnostic instruments, Sterling heights, MI) and processed using Adobe Photoshop (version 7.0).

**Starch visualization**

Starch accumulation in endosperm was monitored using polarized light, in which all the light waves vibrate in a single plane. A polarizing filter that converts unpolarized light to polarized light was used; and transmits the light waves oriented in a single plane. When starch grains are viewed under cross polarized light, they exhibit a property called bi-refringence, which means the polarized light is refracted, and thus starch granules appear as bright bodies against a dark background (Rock, 1981).
Table 3.1 Different developmental stages of wheat seeds grouped according to the fresh weight of the seed (mg/seed).

<table>
<thead>
<tr>
<th>Stages of wheat seeds</th>
<th>Fresh weight (mg/seed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0</td>
<td>1-7</td>
</tr>
<tr>
<td>Stage 1</td>
<td>8-15</td>
</tr>
<tr>
<td>Stage 2</td>
<td>16-20</td>
</tr>
<tr>
<td>Stage 3</td>
<td>21-25</td>
</tr>
<tr>
<td>Stage 4</td>
<td>26-35</td>
</tr>
<tr>
<td>Stage 5</td>
<td>36-45</td>
</tr>
</tbody>
</table>
Results

Endosperm development in wheat grains

Developing wheat grains were grouped into different stages according to their seed weight. Stage 0: 1-7 mg/seed, stage 1: 8-15 mg/seed, stage 2: 16-20 mg/seed, stage 3: 21-25 mg/seed, stage 4: 26-35 mg/seed, and stage 5: 36-45 mg/seed. Sections were prepared from each stage, stained and observed under light microscopy. The following paragraphs describe each stage:

Stage 0: Endosperm in stage 0 seeds was at the syncitial phase of endosperm development, which is characterized by the open central vacuole (CV) surrounded by many layers of maternal tissues (Fig. 3.4b). Antipodal cells were prominently seen on the ventral side of the embryo sac (Fig. 3.4a and 3.4c). Antipodal cells helps in the continuous nuclear divisions of the primary endosperm nucleus, which leads to the formation of multinuclear coenocyte (http://www.wheatbp.net/). Embryo was not seen at this stage.

Stage 1 and Stage 2: Seeds from these stages have endosperms at the cellularization phase, which is characterized by the beginning of the formation of cell walls between nuclei (Fig. 3.5a, 3.5b, 3.6a, and 3.6b). Embryo was first seen in stage 1 seeds, and it has a characteristic globular shape (Fig. 3.5c and 3.5d). More prominent development of the embryo was seen in stage 2 seeds with a scutellum, which helps in nourishing the embryo by transferring the nutrients from the endosperm (Fig. 3.6d). Shoot and root poles were also clearly distinguishable in the longitudinal sections from seeds in stage 2 (Fig. 3.6c and 3.6d).
Figure 3.4 Transverse and longitudinal sections of wheat seeds at stage 0. Wheat seeds were fixed in Safefix II, dehydrated in different concentrations of ethanol, embedded in paraffin, sectioned at 20µm thickness, mounted on glass slides and stained with toluidine blue. (A) and (B) Transverse sections of stage 0 seeds (1-7 mg/seed) showing the empty central vacuole surrounded by maternal tissue. Antipodal cells were prominently seen at this stage. (C) Longitudinal section of a stage 0 seed. Embryo was not seen at this stage. The blue double-headed arrow in (C) represents an approximate location from where the transverse section (A) was taken. Scale bars = 1mm. AC, antipodal cells. CV, central vacuole. ESW, embryo sac wall. M, maternal tissue. N, nucellus.
Figure 3.5 Transverse and longitudinal sections of wheat seeds at stage 1. Wheat seeds were fixed in Safefix II, dehydrated in different concentrations of ethanol, embedded in paraffin, sectioned at 20µm thickness, mounted on glass slides and stained with toluidine blue. (A) and (B) Transverse sections of stage 1 seeds (8-15 mg/seed) showing the endosperm at cellularization phase. Endosperm was first recognizable in stage 1 seeds. (C) Magnified version (10X) of a part of (D), showing the embryo. (D) Longitudinal section of stage 1 seed. Embryo with a characteristic globular shape was first seen at this stage. The blue double-headed arrow in (D) represents an approximate location from where the transverse section (B) was taken. Scale bars 1mm in (A) and (B), 10µm in (C) and 20µm in (D). E, embryo. En, endosperm. M, Maternal tissue.
Figure 3.6 Transverse and longitudinal sections of wheat seeds at stage 2. Wheat seeds were fixed in Safefix II, dehydrated in different concentrations of ethanol, embedded in paraffin, sectioned at 20µm thickness, mounted on glass slides and stained with toluidine blue. (A) and (B) Transverse sections of stage 2 seeds (8-15 mg/seed) showing the endosperm at cellularization phase. (C) Longitudinal section of stage 2 seed showing the embryo. The blue double-headed arrow in (C) represents an approximate location from where the transverse section (B) was taken. (D) Magnified version (10X) of a part of (C), showing the embryo with a clearly distinguishable scutellum, root and shoot poles. Scale bars 1mm in (A) and (B), 10µm in (C) and 20µm in (D). E, embryo. En, endosperm, M maternal tissue. RP, root pole. Sc, scutellum. SP, shoot pole.
Stage 3 and stage 4: Endosperm in these stages was at differentiation phase, in which the cell first increases in size due to cell expansion by accumulating water, starch and protein content (Fig. 3.7a, 3.7b, 3.8a, and 3.8b). Starch was visualized by using the polarized light. Larger lens shaped starch granules, which are type A, and smaller spherical starch granules, which are type B, are sequentially deposited during the wheat endosperm development (Bechtel et al., 1990; Peng et al., 2000). Consistent with the previous studies, larger lens shaped type A starch granules were first visible in stage 3 seeds (Fig. 3.7b). Embryo is developing rapidly; shoot apex and coleoptile were clearly distinguishable in the longitudinal section of the stage 3 seed (Fig. 3.7d). Scutellum has also increased in length. The developing embryo uses the contents of the endosperm near the scutellum thus forming a crushed cell layer (Fig. 3.8d).

Stage 5: Endosperm in seeds at stage 5 seeds is at the maturation phase of the endosperm, which is characterized by the cessation of cell division and accumulation of starch and protein (Fig. 3.9a) (Berger, 1999; Philippe et al., 2006b). When viewed under higher magnification, both type A and type B starch granules were visible in stage 5 wheat seeds (Fig. 3.9b).
Figure 3.7 Transverse and longitudinal sections of wheat seeds at stage 3. Wheat seeds were fixed in Safefix II, dehydrated in different concentrations of ethanol, embedded in paraffin, sectioned at 20µm thickness, mounted on glass slides and stained with toluidine blue. (A) Transverse section of a stage 3 seed (21-25 mg/seed) showing the endosperm at cell expansion phase. (B) Magnified version (10X) of a part of (A) showing starch deposition. (C) Longitudinal section of stage 3 seed showing the developing embryo. The blue double-headed arrow in (C) represents an approximate location from where the transverse section (A) was taken. (D) Magnified version (10X) of a part of (C), showing the embryo with a clearly distinguishable shoot apex, coleoptile and scutellum. Scale bars 1mm in (A), 10µm in (C) and 20µm in (B) and (D). C, coleoptile. DE, developing endosperm. E, embryo. S, starch. SA, shoot apex. Sc, scutellum.
Figure 3.8 Transverse and longitudinal sections of wheat seeds at stage 4. Wheat seeds were fixed in Safefix II, dehydrated in different concentrations of ethanol, embedded in paraffin, sectioned at 20μm thickness, mounted on glass slides and stained with toluidine blue. (A) Transverse section of a stage 4 seed (26-35 mg/seed) showing the endosperm at cell expansion phase. (B) Magnified version (10X) of a part of (A) showing starch deposition. (C) Longitudinal section of stage 4 seed showing the developing embryo. The blue double-headed arrow in (C) represents an approximate location from where the transverse section (A) was taken. (D) Magnified version (10X) of a part of (C), showing the embryo with a clearly distinguishable shoot apex, coleoptile and scutellum. Scale bars 1mm in (A), 10μm in (C) and 20μm in (B) and (D). C, coleoptile. CCL, crushed cell layer. DE, developing endosperm. E, embryo. S, starch. SA, shoot apex. Sc, scutellum.
Figure 3.9 Transverse section of a wheat seed at stage 5. Wheat seeds were fixed in Safefix II, dehydrated in different concentrations of ethanol, embedded in paraffin, sectioned at 20µm thickness, mounted on glass slides and stained with toluidine blue. (A) Transverse section of a stage 5 seed (36-45 mg/seed) showing the endosperm at maturation phase, which is characterized by the starch-filled endosperm. (B) Magnified version (10X) of a part of (A) showing starch deposition. Scale bars 1mm in (A), and 20µm in (B). ME, mature endosperm. S, starch.
**Discussion**

Nuclear endosperm development involves four major phases namely syncitial, cellularization, differentiation, and maturation phases. Philippe et al. (2006a; 2006b; 2006c) showed that different cell wall polysaccharides are deposited at different endosperm developmental phases. Knowledge of the specific time of deposition of cell wall polysaccharides is particularly important as it helps to focus on the specific stages at which the genes involved in particular polysaccharide (i.e., AX, MLG, or callose) biosynthesis are expressed.

Mares et al. (1975) collected the wheat seeds at various times after anthesis and dissected the early stages of endosperm development i.e., syncitial and cellularization phases. As it was difficult to determine the accurate time at which anthesis occurred, Mares et al. (1975) proposed an approximate time scale at which the earlier endosperm developmental phases occurred. In the present work, wheat seeds were grouped taking into consideration the weight of individual seeds, which eliminates the inaccuracies associated with the identification of the exact timing of the anthesis. All the four endosperm developmental phases were identified. A summary of the results from the sectioning and light microscopy of all the stages of wheat seeds is represented in Figure 3.10 and showed that we could identify most of the developmental phases of wheat endosperm. Stage 0 seeds had the multinucleate coenocyte and an empty endosperm, which are the characteristics of the syncitial phase of endosperm development (Fig. 3.4a and 3.4b). In stage 1 and 2, endosperm cells were rapidly dividing (Fig. 3.5a, 3.5b, 3.6a and 3.6b) and forming new cell walls, which are the characteristics of the cellularization phase. According to Philippe et al. (2006b), (1→3)-β-D-glucans and MLG are deposited
at the beginning and end of the cellularization phase respectively. Therefore, the genes involved in the biosynthesis of these two polymers would be highly expressed in stage 0 and stage 1. Endosperm in stage 3 and 4 of seed development had the characteristics of the differentiation phase of endosperm development (Fig. 3.7a and 3.8a). During this developmental phase AX is deposited (Philippe et al., 2006a; Philippe et al., 2006c), thus, the genes involved in the biosynthesis of this polymer would be expected to be highly expressed in stage 1 and stage 2. Stage 5 seeds correspond to the maturation phase of endosperm development (Fig. 3.9a). Thus, these results indicate that it is possible to identify various developmental stages of wheat endosperm using our strategy.

We have now established growth conditions and the exact timing to easily obtain wheat seeds at a desired developmental stage of the endosperm. These data will be the foundation for the future work aiming in identifying the genes involved in AX and MLG biosynthesis using in situ hybridization and RT-PCR strategies.

Furthermore, RT-PCR experiments also support the assigned developmental stages identified by the sectioning and microscopy experiments. The RT-PCR experiments, performed by another graduate student (Wei Zeng) in our lab, monitored the expression of starch synthase and storage protein genes in all the identified stages of wheat seeds. Starch synthase gene was highly expressed in stages 2, 3, and 4, which suggests that starch deposition is initiated in stage 3, consistent with the results from polarized light microscopy experiments showing the beginning of starch deposition at stage 3 (Fig. 3.7b).
Figure 3.10 Summary figure showing the transverse sections of the seeds from all the stages (stage 0-5). Sections were observed under 2X magnification. Scale bars in seeds of stage 1-5 = 1mm. DAA, days after anthesis. Time scale (DAA) for the endosperm developmental phases was adapted from Philippe et al. (2006b). CV, central vacuole.
Chapter 4: GENERAL CONCLUSIONS AND FUTURE WORK
General Conclusions

Xyloglucan Biosynthesis in Grasses

XyG is the major hemicellulosic polysaccharide in type I walls but small amounts of this polysaccharide are also seen in the type II walls of grasses. Due to the functional importance of XyG in type I walls, many of the genes encoding the enzymes involved in XyG biosynthesis have been identified. The specific aim of the research project was to understand the XyG biosynthesis in grasses by cloning and expression of the genes encoding the XyG biosynthetic enzymes.

The wheat and rice genes, TaGT34-7 and OsGT34-3 that were clustered with the two known XyG-xylosyltransferase proteins from *Arabidopsis* (AtXT1 and AtXT2) have 79% and 80% amino acid identity with AtXT1 respectively. Protein sequences of TaGT34-7 and OsGT34-3 were analyzed using various bioinformatic programs available publicly and several motifs characteristic to GT-34 family (CAZy classification) were also identified in both wheat and rice putative XyG-xylosyltransferases. Also, both proteins were predicted to be N-glycosylated with an estimated MW of 52kD (TaGT34-7) and 51kD (OsGT34-3).

For functional analysis, the full-length cDNAs encoding TaGT34-7 and OsGT34-3 were expressed in *Pichia* and *Drosophila* cells, and the expression of the proteins was monitored by western blot. However, the expressed proteins did not show the XyG-xylosyltransferase activity in detergent-extracts from *Pichia pastoris* cells, possibly because of the differences in the N-glycosylation patterns between *Pichia* and plant cells. Whereas, the proteins expressed in *Drosophila* S2 cells showed very low activity, which
was not strong enough to unambiguously confirm their biochemical identity as XyG-xylosyltransferases.

**Identification of Endosperm developmental phases in wheat seeds**

By transverse sectioning and light microscopy methods, different endosperm developmental phases in the wheat grain were identified. Identification of the endosperm developmental phases is particularly important because different cell wall polysaccharides are deposited sequentially in the cell walls of wheat endosperm. AX, the major polysaccharide in the endosperm cell walls (70%w/w), is deposited during differentiation phase of the wheat endosperm and MLG (20%w/w) is deposited during cellularization phase of the endosperm development in wheat (Philippe et al., 2006b). According to the data, an approximate time scale to was assigned to each group of the wheat seeds as indicated in table 4.1.

The endosperm developmental phases identified using the sectioning and microscopy experiments were confirmed by RT-PCR experiments (performed by another graduate student in our lab), in which the expression of starch synthase and storage protein genes were used to monitor in all the stages of developing wheat seeds.
Table 4.1 Assigning an approximate time scale to the wheat seeds with the endosperm at different developmental phases.

<table>
<thead>
<tr>
<th>Stages of wheat seeds</th>
<th>Fresh weight (mg/seed)</th>
<th>Endosperm developmental phases</th>
<th>Days after anthesis (estimated from Fig. 3.10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0</td>
<td>1-7</td>
<td>Syncitial</td>
<td>2</td>
</tr>
<tr>
<td>Stage 1</td>
<td>8-15</td>
<td>Cellularization</td>
<td>3</td>
</tr>
<tr>
<td>Stage 2</td>
<td>16-20</td>
<td>Cellularization</td>
<td>6</td>
</tr>
<tr>
<td>Stage 3</td>
<td>21-25</td>
<td>Differentiation</td>
<td>10</td>
</tr>
<tr>
<td>Stage 4</td>
<td>26-35</td>
<td>Differentiation</td>
<td>13</td>
</tr>
<tr>
<td>Stage 5</td>
<td>36-45</td>
<td>Maturation</td>
<td>20</td>
</tr>
</tbody>
</table>
Future work

Optimization of protein expression in *Drosophila S2* cells

Although the preliminary analyses of the proteins expressed in *Drosophila S2* cells showed a low activity, more work is needed to confirm the XyG-xylosyltransferase activity of TaGT34-7 and OsGT34-3. Santos et al. (2007) showed that the copper sulfate induction caused a decrease in the cell growth and EGFP expression at lag and stationary phases but did not effect the exponential phase of cell growth. This suggests that the highest protein expression can be obtained by inducing the cells at a specific stage of cell growth. Therefore, transfected S2 cell lines of TaGT34-7 and OsGT34-3 can be grown to different cell densities to check the stage of cell growth at which maximum protein expression can be seen. Also to optimize the parameters for maximal protein expression, the transfected S2 cell lines can be induced with different concentrations of copper sulfate and for different time periods.

Optimization of the biochemical assay

In the present study, cellohexaose was used as the acceptor in the xylosyltransferase assays to test the activities of TaGT34-7 and OsGT34-3. This is because earlier study done has shown that cellohexaose acts as a better acceptor for AtXT1 and AtXT2 (Cavalier and Keegstra, 2006). But the enzyme activity of TaGT34-7 and OsGT34-3 can be tested using different acceptors such as cellobiose, cellotriose, celloteraose and cellopentaoses. Also, pH and temperature plays an important role in the enzyme activity. Therefore, the optimum pH and temperature of the putative
xylosyltransferases can be identified by setting up the enzyme assays at a wide range of pH’s and temperatures. Finally after optimizing the enzyme assays, the reactions can be scaled up several times for further biochemical characterization of the reaction products by using size exclusion chromatography or High Performance Liquid Chromatography (HPLC). More importantly, the enzyme from grasses may have a strict requirement for the presence of glucan synthase activity. In this case, it would be difficult to measure the activity. This would be a major difference between XyG biosynthesis in dicots and grasses.
References


Carpita NC and Gibeaut DM (1993) Structural models of the primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. The Plant J 3: 1-30


Fry SC (1988) The growing plant cell wall: Chemical and metabolic analysis. pp 4


Kooiman P (1961) The constitution of Tamarindus amyloid. Rec Trav Chim Pays-Bas 80: 849-865


Munro S (1991) Sequences within and adjacent to the transmembrane segment of α-2,6-sialyltransferase specify Golgi retention. EMBO J 10: 3577-3588


Wiggins CAR and Munro S (1998) Activity of the yeast MNN1 α-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. Proc Natl Acad Sci USA 95: 7945-7950


Yokoyama R and Nishitani K (2001) A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of *Arabidopsis*. Plant Cell Physiol 42: 1025-1033
