MEMBRANE PROTEIN TOPOLOGY AND SUBCELLULAR LOCALIZATION OF ENZYMES INVOLVED IN THE BIOSYNTHSIS OF THE HEMICELLULOSIC POLYSACCHARIDES XYLOGLUCAN AND MANNAN

By

Jonathan Kent Davis

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Plant Biology

2010

ABSTRACT

MEMBRANE PROTEIN TOPOLOGY AND SUBCELLULAR LOCALIZATION OF ENZYMES INVOLVED IN THE BIOSYNTHSIS OF THE HEMICELLULOSIC POLYSACCHARIDES XYLOGLUCAN AND MANNAN

By

Jonathan Kent Davis

Hemicellulosic polysaccharides are produced in the Golgi apparatus. The enzymes responsible for synthesizing the glycan "backbones" of the hemicellulosic polysaccharides xyloglucan and galactomannan have been identified. CSLA proteins synthesize the (gluco)mannan backbone of galacto(gluco)mannan, while CSLC proteins are the best candidates for to synthesize the glucan backbone of xyloglucan (Liepman *et al*. 2007, Cocuron *et al*. 2007). CSLA and CSLC proteins have multiple transmembrane domains, and the topology of these proteins is likely to impact how they function. For example, a CSL protein with a cytosolic active site may couple synthesis with translocase activity similarly to cellulose synthase. Alternatively, glycan synthesis by a protein with an active site in the Golgi lumen might be dependent on sugar-nucleotide transporters. Here I present fluorescent protein localization by confocal microscopy showing that CSLA and CSLC proteins reside in Golgi membranes. The colocalization of three proteins involved in the biosynthesis of xyloglucan (CSLC4, XXT1, and FUT1) within the secretory pathway was also examined. I probed the topologies of CSLA and CSLC proteins by protease protection methods designed to determine the position of the proteins' termini and their active sites. I show that AtCSLC4, a Golgi-localized xyloglucan glucan synthase, has a topology characterized by a cytosolic active site and an even number (likely 6) of transmembrane domains. Whereas the most abundantly expressed mannan synthase in *Arabidopsis*, AtCSLA9, is predicted to have 5 TMDs and has an active site facing the Golgi lumen. These topologies

have important implications for understanding the biosynthesis of cell wall polysaccharides.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my advisor, Kenneth Keegstra for his tremendous contribution to the development of my research and my growth as an individual. I would also like to thank my committee: Federica Brandizzi, for her advice and assistance that made the fluorescence microscopy work presented here possible; John Ohlrogge, for being a role-model and example of the qualities of integrity, inquisitiveness, and intellectual rigor that every successful scientist must have; and Markus Pauly, who has been an inspiration for me in his effective management style and success as an enthusiastic, young faculty member.

The members of the Keegstra lab also deserve my thanks. In particular, I want to thank Linda Danhof, David Cavalier, and John Froehlich for their assistance and advice. I would also like to thank my good friends Leron Katsir, John Withers, and Saunia Withers, who have been a constant source of inspiration and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

Introduction to cell walls and structural polysaccharides

A cell wall consisting of a complex matrix of polysaccharides, proteins, and mineral components surrounds each plant cell. In most mature plant cells, the cell wall has two compositionally and functionally distinct regions: the primary and secondary cell walls. However, some specialized cell types, such as root hairs, never deposit a secondary wall (Belford and Preston 1960). Walls of adjacent cells are separated by a pectin-rich structure called the middle lamella. The primary wall is deposited first during development, while the cell is still expanding. The primary wall of plant cells is composed mainly of polysaccharides, with smaller amounts of structural glycoproteins, phenolic esters, minerals such as calcium and boron, and enzymatic proteins. In growing cells, the primary wall controls the rate and direction of turgor-driven cell expansion and ultimately determines the size and shape of the plant cell at maturity. (Carpita and McCann 2000)

The secondary cell wall is deposited after the primary wall, when cell expansion is complete, and it is found between the primary wall and the plasma membrane. The secondary walls of certain cell types are made rigid by free radical polymerization of secreted lignin monomers, which crosslinks the matrix of polysaccharides and proteins (Bonawitz and Chapple 2010). In many cases, such as with water-conducting xylem cells, the secondary wall is deposited in a specific pattern that contributes to the function of the cell at maturity (Höfte 2010).

The polysaccharide components of both the primary and secondary cell walls can be divided into three classes: cellulose, hemicellulosic polysaccharides, and pectin. There are a number of key features of polysaccharides that must be considered to understand their biosynthesis and properties within the wall. Most monosaccharides, including all of those found

in cell wall polymers, cyclize in solution, forming 5-member furanose or 6-member pyranose rings. The position of the hydroxyl group attached to the anomeric carbon when a sugar is cyclized determines the anomeric conformation (α or β) of the sugar or glycoside. By convention, the anomeric carbon of aldoses is referred to as C1 and the anomeric carbon of ketoses is C2. (Bertozzi and Rabuka 2008)

For the D-sugars, a glycosidic bond is said to be in the β confirmation when the anomeric hydroxyl is in an equatorial position relative to the plane of the ring, and in the α confirmation if this hydroxyl is axial. This description of the positions of α and β bonds is reversed for L-sugars. In a polysaccharide, the type of linkage (α or β) determines the angle formed by adjacent residues and the relative orientation of each monomer unit (Bertozzi and Rabuka 2008). For example, the glucose monomers that make up cellulose are β(1→4)-linked, which results in a linear polymer with each glucose residue inverted with respect to its neighbor. The type of linkages in a polysaccharide has several implications for its biosynthesis. For example, if only one β(1→4) glucan synthase adds a single glucose residue at a time, how does the enzyme's active site accommodate the "flipping," or inversion of active site chemistry that must occur as each sugar is added? Also, if a β(1→4)-linked glycan is substituted with sugar side chains, as occurs in the hemicellulosic polysaccharide xyloglucan, substitutions at the same position to adjacent glucose residues would occur on opposite sides of the polymer.

Many of the unique physical properties of polysaccharides are related to the large number of hydrogen bonding sites in these molecules. Polysaccharides with higher degrees of branching or sugar side chain substitutions have more sites that can participate in hydrogen bonding with water, and therefore have increased solubility (Whistler 1973). In contrast, an unsubstituted β(1→4)-linked hexose polymer like cellulose becomes insoluble in water as the degree of polymerization passes six or seven residues (Whistler 1973). Highly substituted polysaccharides

such as xyloglucan or galactoglucomannan form stable hydrocolloids in solution, imparting properties that make these polymers useful for applications in food stabilization and other industrial processes (Dickinson 2003). The special properties of these polysaccharides also contribute to human health as the "soluble fiber" components of our diets (Edwards and Parrett 1996). Hydrogen bonding of structural polysaccharides with water contributes to the unique hydrated environment of the primary cell wall, which allows limited free diffusion of small hydrophilic molecules while providing a structure rigid enough to control cell expansion. Hydrogen bonding among polysaccharides in the cell wall is thought to be an essential component of a load-bearing network that resists uncontrolled, turgor-driven expansion and bursting (Cosgrove 2005; Hayashi 1989; Ray *et al.* 1972; Taiz 1984) . The microfibril structures typical of the cell walls of land plants are composed of cellulose polymers hydrogen bonded to each other along their length (Delmer and Amor 1995; Guerriero *et al.* 2010).

Cellulose is an unsubstituted, linear polymer of β(1→4)-linked glucose residues, and is present in the wall in both amorphous and crystalline confirmations (Carpita and McCann 2000). Crystalline cellulose is found in linear structures known as microfibrils, each a bundle of multiple individual glucan polymers in an ordered array. Cellulose microfibrils are synthesized at the plasma membrane by protein complexes containing cellulose synthase catalytic subunits (CESA) and accessory proteins (Desprez *et al.* 2007; Guerriero*, et al.* 2010). Hemicellulosic polysaccharides have greater structural complexity than cellulose, each polymer often having several different sugars as monomer units and more than one type of linkage. Hemicellulosic polysaccharides include xyloglucan, galacto(gluco)mannan, xylan, glucoronoarabinoxylan, and mixed-linkage β(1→3;1→4)-glucans, (the latter of these is found only in the grasses) (Scheller and Ulvskov 2010). Pectic polysaccharides can have even more complex structures, and are grouped into three classes: homogalaturonans (which may be substituted with xylose or apiose),

rhamnogalacturonan-I, and the highly branched, complex polymer rhamnogalacturonan-II (Mohnen 2008).

Structure and function of hemicellulosic polysaccharides xyloglucan and mannan

Xyloglucan polymers in plant cell walls often have a very regular structure characterized by a repeating pattern of substituted and unsubstituted glucose residues (Vincken *et al.* 1997). With *Arabidopsis* xyloglucan, three adjacent residues of the glucan backbone are each substituted by a single α(1→6)-linked xylose residue, and followed by one unsubstituted glucose. This pattern of substitution repeats, apparently through the length of the polymer. The second and third xylose residues from the non-reducing end of the repeating unit can be substituted by β-linked galactose on position 2. In *Arabidopsis* seedling leaves, 22.8% of the second and 52% of the third xylose from the non-reducing end of the repeating structure are galactosylated. The galactose on the third xylose from the non-reducing end is fucosylated in about 39.5% of the repeating subunits (approximately 76% of the available sites) in this tissue (Peña *et al.* 2004). The dominant structure of the repeating unit of *Arabidopsis* xyloglucan is shown in Figure 1. Although this structure is the most common repeating subunit of *Arabidopsis* xyloglucan, other structures, such as two substituted glucose residues followed by an unsubstituted glucose can also be found, albeit at much lower abundance (Obel *et al.* 2009). Xyloglucan in other plant species can have different substitution patterns. For example, the repeating subunit of xyloglucan in the Solanaceae has two glucose residues substituted with xylose, followed by two to three unsubstituted glucose residues (O'Neill and York 2003). Within a single species, the substitution pattern of xyloglucan tends to be extremely regular. Xyloglucan polymers can also be extensively acetylated, a modification that is likely to affect interactions of the polysaccharide with the solvent as well as other polysaccharides and proteins in the wall (Hayashi 1989; Kiefer *et al.* 1989; Liepman *et al.* 2007a; Scheller and Ulvskov 2010).

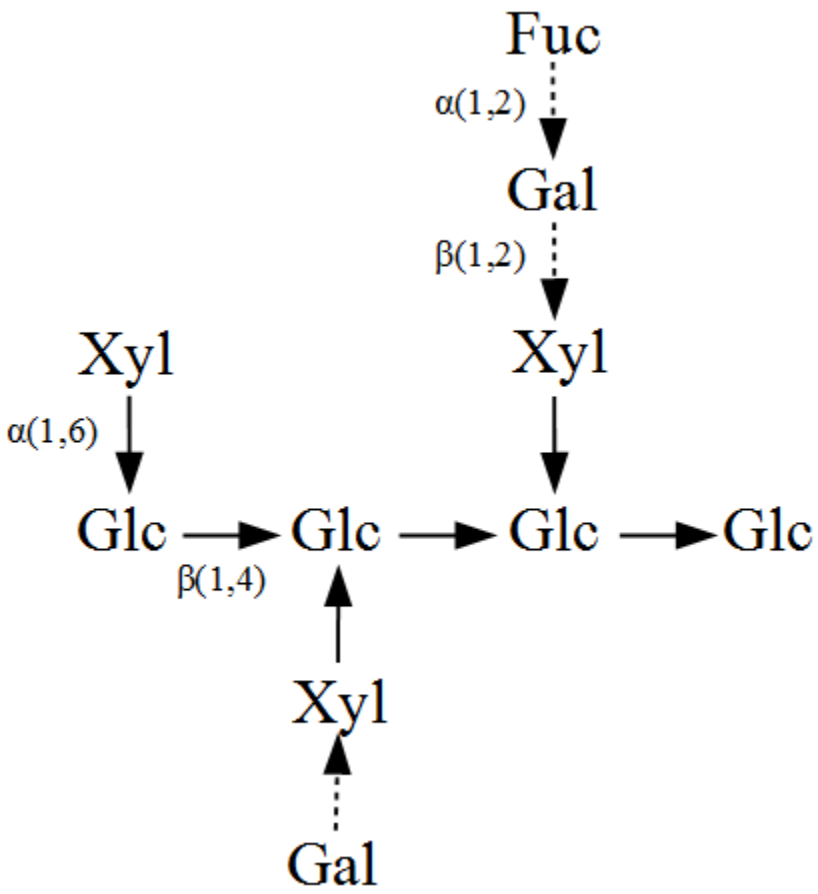

Figure 1. Glycosidic linkages of the repeating structural unit of *Arabidopsis* xyloglucan. The arrows indicate glycosidic bonds, and the types of linkage are given. Dashed arrows indicate linkages that occur in some of the repeating units. The frequency of these substitutions is described in the text. (abbreviations Fuc: fucose, Gal: galactose, Xyl: xylose, Glc: glucose)

Xyloglucan is a major hemicellulose in the primary wall of dicots and many monocots, with the exception of the poales (Scheller and Ulvskov 2010). Because xyloglucan is so abundant and widely distributed, models of primary cell wall structure and function often assign important roles to this polysaccharide (Cosgrove 2000; Cosgrove 2001; Cosgrove 2005; Keegstra *et al.* 1973). Some models propose that xyloglucan polymers crosslink cellulose microfibrils to create a load-bearing network (Somerville *et al.* 2004; Thompson 2005). In these models, parallel arrays of cellulose microfibrils are interconnected by a network of xyloglucan polymers by hydrogen bonding, such that individual xyloglucan polymers might be non-covalently bound to two microfibrils with a region of "free" xyloglucan in between. In this type of model, cell expansion can be controlled by the extent of hydrogen bonding between xyloglucan and cellulose and the length of xyloglucan polymers extending between microfibrils (Somerville*, et al.* 2004). The hypothetical "crosslinking glycan" function of xyloglucan is such an attractive model because it provides a functional explanation for many aspects of xyloglucan synthesis and post-depositional modification. The spatio-temporal pattern of deposition, modification by expansins, xyloglucan endo-transglycosidase/hydrolase (XTH) activity, acetylation, and potentially the addition or removal of sugars post-deposition all contribute to how xyloglucan might mediate cell expansion by crosslinking cellulose microfibrils.

Given the regularity of xyloglucan structure, it is likely that the substitution pattern is tightly controlled and functionally relevant. This idea also fits well with the "crosslinking glycan" hypothesis of xyloglucan function. Substitution of the glucan backbone may affect the extent and strength of interactions with cellulose microfibrils, for example. The sugar substitution pattern within "free" regions of xyloglucan may influence interactions with enzymes that target xyloglucan for cell wall remodeling during expansion. Further modifications that may occur *in muro* are also likely to play an important role in modulating the properties of xyloglucan

in the cell wall (Liepman*, et al.* 2007a; Vicente *et al.* 2007) . Despite its abundance and decades of research on this important polymer, we are only beginning to understand how xyloglucan functions (Cavalier and Keegstra 2006).

Mannan, another hemicellulosic component of the cell walls of *Arabidopsis*, is widely found in land plants (Liepman *et al.* 2007b). In the woody tissues of gymnosperms, mannans are the most abundant hemicellulosic polysaccharide (Maeda *et al.* 2000). Mannans are also present in many algal species, some of which produce no cellulose (Frei and Preston 1968; Popper and Tuohy 2010). In some of these species, mannan polymer chains form crystalline fibrils similar in appearance to the cellulose microfibrils found in the walls of higher plants. Cell wall polymers made of mannose are widely distributed through the phylogeny of photosynthetic organisms, and might be more representative of an ancestral structural polysaccharide than cellulose.

In some cases, the β(1→4)-mannan backbone is heavily substituted with galactose. Because an unsubstituted mannan polymer of mannose would be insoluble, substitution of the mannan backbone with galactose, or modifications such as O-acetylation, might contribute to allowing the polymer to be soluble during synthesis and deposition. Glucose residues may also be incorporated into the backbone of galacto(gluco)mannan (Goubet *et al.* 2009). It is not clear at this time whether the glucose composition of galactoglucomannan polymers has a functional role, or even if it is regulated beyond the ability of mannan synthases to incorporate glucose as well as mannose. For the sake of clarity, I refer to galactoglucomannan polymers of various compositions simply as "mannans" in this document.

Mannans occur mainly in secondary walls of *Arabidopsis*, and are thus an important polysaccharide component of inflorescence stems (Carpita and McCann 2000; Zhong *et al.* 2001). It is difficult to conceive a structural role for mannans in the *Arabidopsis* primary wall, given their low abundance and the lack of major growth phenotypes when mannan is eliminated

(Goubet, *et al.* 2009). Phenotypes of *Arabidopsis* plants with perturbed mannan synthesis, such as Agrobacterium rhizogenes resistance (Zhu *et al.* 2003) suggest that cell wall mannans may play a role in signaling development or defense in *Arabidopsis*. In species where mannans do accumulate to high levels in certain tissues, they could be playing a structural role similar to xyloglucan in the crosslinking model described above.

Hemicellulose biosynthesis

Unlike the synthesis of other biological macromolecules, such as proteins and nucleic acids, structural polysaccharide biosynthesis occurs without a template. Rather, the astounding complexity and regularity observed in cell wall polysaccharides is accomplished through enzyme activity alone (Perrin *et al.* 2001). The specificity of a polysaccharide biosynthetic enzyme for its acceptor substrate (polymer) and donor substrate (sugar nucleotide) ultimately determines the structure of the finished polymer product.

The glycosyltransferases involved in polysaccharide biosynthesis can be classified as either processive or non-processive, depending on the type of reaction catalyzed (Keegstra and Raikhel 2001). Enzymes that catalyze the addition of sugars to β(1→4)-linked "backbone" polymers such as cellulose, mannan, and the glucan backbone of xyloglucan, are thought to be processive enzymes. This means the acceptor substrate (the elongating polymer) remains associated with the enzyme during catalysis of sequential glycosidic linkages, and each sugar residue is added to the sugar that just preceded it.

In contrast to the polysaccharide synthases that processively add sugars to an elongating backbone chain, glycosyltransferases that decorate these backbones with sidechain substitutions add a single sugar at a time to the acceptor polymer. Enzymes in this group include all those that add the sugar side chains present on xyloglucan (XXT1, XXT2, XXT5, MUR3, FUT1) (Keegstra and Raikhel 2001; Price Nicholas *et al.* 2008; Wulff *et al.* 2000). It is interesting to note that

those enzymes identified so far to be processive glycan synthases are polytopic type III membrane proteins (belonging to the CESA superfamily), whereas the non-processive biosynthetic enzymes all appear to be type II membrane proteins (Price Nicholas*, et al.* 2008), meaning that they have a single membrane domain and the amino terminus in the cytosol.

From a mechanistic perspective, glycosyltransferases can also be classified as "retaining" or "inverting" enzymes based on whether the anomeric configuration of the sugar nucleotide substrate is conserved after sugar transfer. For example, because UDP-glucose has an α anomeric configuration, and the backbone of xyloglucan is β(1→4)-linked, CSLC4 is an inverting enzyme. In contrast, the xylosyltransferase XXT1 that adds side chains to xyloglucan is a retaining enzyme, because both the UDP-xylose and the (1→6) xylose-glucose glycosidic bonds are alpha linkages (Sinnott 1990).

The distinction between inverting and retaining glycosyltransferases is important because the two classes of enzymatic reactions are mechanistically distinct. With inverting glycosyltransferases, the orientation of the glycosidic bond of the donor substrate is inverted as the acceptor substrate performs a nucleophilic attack on the anomeric carbon of the donor sugar in a single displacement reaction. The preservation of the anomeric configuration that occurs with retaining glycosyltransferases is probably accomplished through a more complex, two-step method. The proposed mechanism thought to be used by retaining glycosyltransferases involves a donor glycosyl-enzyme intermediate that is subsequently attacked by the acceptor substrate. (Sinnott 1990)

A combination of enzymatic glycosyltransferase activities, processive and non-processive, retaining and inverting, is needed for the synthesis of cell wall polysaccharides. Biosynthesis of polysaccharides with high structural complexity, those which contain several different sugars and glycosidic bonds, is likely to be a highly coordinated process. Xyloglucan is

an example of such a polymer, as the structure of this polysaccharide (described above) is both complex and extremely regular. Given the structure of xyloglucan at deposition, it is clear that addition of sugars to the side chains must proceed in an ordered manner. For example, the fucose residue cannot be added until the galactose is present, which must be preceded by xylosylation at the third glucose of the repeating subunit.

The striking complexity and regularity of certain polysaccharides, like xyloglucan, are probably brought about through a combination of factors. First, each glycosyltransferase should be very specific for the donor substrate, acceptor substrate, and the linkage formed. Because sugars can be similar in structure, sometimes distinguished by the position of a single hydroxyl group, the low degree of substrate promiscuity among glycosyltransferases is somewhat surprising. Exceptions to the one enzyme, one substrate paradigm include the ability of CSLA proteins to use both GDP-glucose and GDP-mannose as substrates (Liepman*, et al.* 2007b; Liepman *et al.* 2005) and the incorporation of L-galatose to the xyloglucan of *mur1* mutants, apparently by FUT1 (Zablackis *et al.* 1996). Second, each enzyme must be present at the right time and place during synthesis to add the appropriate sugar in a regular manner. Third, the donor substrate must be made available at the specific location where it will be utilized. One strategy that plants may use to ensure that these circumstances are met is careful control of protein trafficking. As the biosynthesis of xyloglucan or other complex polysaccharides proceeds in the secretory pathway, protein interactions among biosynthetic enzymes, substrate transporters, and possibly other unknown actors are likely to play a key role in accomplishing the high degree of organization of biosynthetic machinery that must be required to produce polymers with such complexity and regularity.

Enzymes in hemicellulose biosynthesis

Identification of the enzymes responsible for the synthesis of cell wall polysaccharides

has proceeded more slowly than the identification and structural characterization of their polymer products. This is because these enzymes are almost always low abundance integral membrane proteins, and thus recalcitrant to traditional biochemical characterization. However, a few cell wall polysaccharide biosynthetic enzymes have been identified by traditional biochemical purification followed by protein sequencing. This was the case with galactomannan galactosyltransferase and xyloglucan fucosyltransferase, which were first isolated from fenugreek (*Trigonella foenum-graecum*) endosperm (Edwards *et al.* 1999) and pea (*Pisum sativum*) seedlings (Faik *et al.* 2000; Perrin *et al.* 1999) respectively.

More recently, advances in genomic resources and high-throughput nucleic acid sequencing technologies have allowed the identification of candidates for cell wall polysaccharide biosynthetic enzymes by sequence similarity. The identification of the cellulose synthase-like (CSL) gene family has proven to be a crucial breakthrough in understanding cell wall polysaccharide biosynthesis. The application of traditional biochemical approaches combined with genomic resources that had recently become available led to the discovery of the cellulose synthase catalytic subunits and a number of putative glycosyltransferases whose functions are still being uncovered.

Research leading to the identification of the CSL family began in the 1980s, as several groups worked to identify and clone the cellulose synthase from *Acetobacter xylinum*. Early work in this area is extensively reviewed in (Delmer 1999). In the 1990s, two groups working separately identified a consensus motif conserved among the *Acetobacter* CESA proteins and several other known glycosyltransferases that utilize UDP-glucose or UDP-glucuronic acid (Delmer and Amor 1995; Saxena *et al.* 1995). The D,D,D,QXXRW motif they identified was found by Pear *et al.* (1996) to be encoded by cDNA in a library from elongating cotton fiber, resulting in the identification of the first cellulose synthase gene from higher plants.

It soon became clear that CESA gene homologs have undergone massive duplication and expansion in higher plants. Through analysis of newly available genomic resources, Richmond and Somerville (Richmond and Somerville 2000) identified 41 CSL genes in *Arabidopsis thaliana*, and grouped them into 7 gene families based on sequence similarity (CESA, CSLA, CSLB, CSLC, CSLD, CSLE, and CSLG). As more genomes were sequenced, additional CSL homologs were identified, and the CSL superfamily has been expanded to include CSLF and CSLH, which are only present in the grasses (Hazen *et al.* 2002; Yin *et al.* 2009).

Observing the obvious similarities between cellulose and other plant cell wall polysaccharides, Richmond and Somerville (2000) posited the "CSL hypothesis." They made the prediction that the diversity of CSL protein families reflects the diversity of polymer products in the cell wall, and they supposed that each CSL clade is responsible for the synthesis of a distinct polymer. The fact that some CSL clades are present only in monocots (CSLF/H) or dicots (CSLB) also suggests functional specialization among the families, given the distinct polysaccharide composition of dicot and monocot walls. Thus far, CSL hypothesis has largely borne out to be an accurate prediction, with phylogeny recapitulating function.

However meritorious, the prediction that CSL proteins are processive glycan synthases involved in cell wall polysaccharide biosynthesis leaves many questions unanswered. Not least of the open problems suggested by the CSL hypothesis is the question of what polymer is synthesized by each CSL protein family. Some of the most successful strategies to assign function to CSL families have exploited the fact that many plants, certain legume species in particular, utilize polysaccharides that are often found in cell walls as carbon storage in seed. During seed filling in these species, the metabolism of tissues within the seed becomes specialized for the production of a single major product, with a significant portion of imported carbon destined for the storage polysaccharide (Reid and Meier 1970). In these periods of rapid

polysaccharide deposition, transcript and protein levels of enzymes involved in the production of these polysaccharides increase dramatically relative to other stages in seed development (Cocuron *et al.* 2007; Dhugga *et al.* 2004; Perrin*, et al.* 2001; Reid and Meier 1970) . Polysaccharide-accumulating seed tissues thus bypass the problem of low enzyme abundance, one of the basic obstacles to the identification and subsequent characterization of cell wall biosynthetic enzymes.

Seeds of the legume guar (*Cyamopsis tetragonoloba*) accumulate galactomannan as a storage polysaccharide. By analyzing ESTs from cDNA libraries generated before and during the peak stages of galactomannan deposition in developing seeds, Dhugga et al (2004) identified transcript encoding a CSLA homolog that increased in abundance concomitantly with mannan synthase activity. Heterologous expression of the guar CSLA protein in transgenic soybean confirmed its identity as a mannan synthase. Since this initial discovery, several research groups have provided evidence that the backbones of mannans and glucomannans are synthesized by the CSLA proteins in many, probably all plant species (Goubet*, et al.* 2009; Liepman*, et al.* 2007b; Liepman*, et al.* 2005; Suzuki *et al.* 2006). CSLA proteins may incorporate both glucose (from GDP-glucose) and mannose residues to the backbone of glucomannan. For example, the *Arabidopsis* mannan synthase CSLA9 is capable of utilizing both GDP-glucose and GDP-mannose as substrates for polymerization in vitro, though this enzyme incorporates mannose at a much higher rate (Liepman*, et al.* 2005).

Cocuron *et al.* (2007) showed that in *Arabidopsis*, and probably in other species, the glucan backbone of xyloglucan is synthesized by the CSLC4 protein. The authors of this study identified transcript of a CSLC4 homolog in developing nasturtium (Tropaeolum majus) cotyledons, which accumulate xyloglucan as a storage polysaccharide. Pichia pastoris expressing the *Arabidopsis* CSLC4 protein produce β(1→4)-linked glucan. Because *Arabidopsis*

CSLC4 localizes to the Golgi, and the β(1→4)-linked glucan synthase activity contributed by CSLC4 increases when the xyloglucan xylosyltransferase XXT1 is also expressed, the authors conclude that CSLC4 synthesizes the glucan backbone of xyloglucan.

New evidence presented by Dwivany *et al.* (2009) indicates that some CSLC homologs in barley are located at the plasma membrane. Based on this observation and the sequence relationship among CSLC proteins, they propose that this family contains two clades: one group that includes *Arabidopsis* CSLC4 and its homologs, which synthesize the backbone of xyloglucan in the Golgi, and another group of CSLC proteins that are active at the plasma membrane and are probably involved in the synthesis of another glucose-containing polysaccharide.

Unfortunately, the functions of CSL family proteins other than CSLA and CSLC have not been determined unambiguously. For example, CSLF and CSLH proteins are known to be involved in the synthesis of mixed-linkage glucan (Burton *et al.* 2010a; Burton *et al.* 2006; Doblin *et al.* 2009), but it is not clear exactly how proteins from these families contribute to the synthesis of β(1→3;1→4)-glucans (Burton *et al.* 2010b). It is also possible that some cell wall polysaccharides are not synthesized by CSL family proteins. This might be the case with xylan synthesis (reviewed in (York and Oneill 2008)) which occurs in secondary cell walls and in the primary walls of graminaceous monocots (as glucoronoarabinoxylan).

Substrates for hemicellulose biosynthesis

All enzymes that have been shown to participate in the synthesis of cell wall polysaccharides utilize sugar-nucleoside diphosphates (NDP-sugars) as substrates, the phosphoester bond providing energy for sugar transfer. The nucleotide moiety of substrates used for cell wall polysaccharide production is usually either uridine or guanine diphosphate. NDP-sugar substrates are generated from hexose phosphates by the activities of soluble nucleotide

triphosphate pyrophosphorylases, presumably in the cytosol of plant cells. Interconversion of the sugar moiety of UDP-glucose via epimerase, oxido-reductase, and decarboxylase reactions contributes to the production of the UDP derivatives of galactose, galacturonic acid, glucuronic acid, xylose, apiose, arabinose, and rhammnose (Seifert 2004). In contrast, GDP-mannose is synthesized from mannose-1-phosphate by the activity of a pyrophosphorylase following the conversion of glucose-6-phosphate to mannose-6-phosphate by a glucose-6-phosphate mutase (Seifert 2004). GDP-mannose, in turn, provides substrate for the production of GDP-fucose through a series of oxido-reductions (Seifert 2004).

The enzyme sucrose synthase (SuSy), which converts sucrose and UDP into UDP-glucose and fructose, is likely to be a source of donor substrate for glucan synthases. The abundance of SuSy protein in rapidly growing cells, and its apparent association with the plasma membrane (the site of cellulose biosynthesis and a sink location for UDP-glucose) suggests that SuSy may be a major source of substrate for at least some UDP-glucose-utilizing enzymes (Amor *et al.* 1995; Haigler *et al.* 2001; Koch 2004; Winter and Huber 2000). In fact, SuSy has been found to directly interact with a component of the cellulose synthase complex that synthesizes β(1→4)-glucan in vitro (Fujii *et al.* 2010).

Because sugar-nucleotides are too highly charged to be membrane-permeable, and they are probably synthesized in the cytosol (Seifert 2004), they must be transported across a membrane in order to be utilized in another cellular compartment. Because many glycosyltransferases involved in the synthesis of cell wall polysaccharides are type II membrane proteins with their active sites in the Golgi lumen, (Goubet and Mohnen 1999; Scheible and Pauly 2004; Sterling *et al.* 2006) it is very likely that sugar-nucleotide transporters play a central role in cell wall biosynthesis. Although some sugar-nucleotide interconverting enzymes localize to the Golgi (Burget *et al.* 2003; Reiter and Vanzin 2001), the sugar-nucleotide substrates of

these enzymes must be imported. Sugar-nucleotide transporters are membrane proteins with multiple membrane spans, and generally appear to be selective for a specific sugar-nucleotide or a narrow range of substrates (Reyes and Orellana 2008; Wulff*, et al.* 2000). Recent progress regarding the role of sugar-nucleotide transporters in cell wall biosynthesis is reviewed extensively by Reyes and Orellana (2008).

Location of synthesis: Golgi apparatus

Hemicellulosic polysaccharides are probably synthesized in the Golgi apparatus, though some synthesis in the ER cannot be completely ruled out. This has been known for many years, as antibodies specific for fucosylated xyloglucan were developed and used for transmission electron microscopy in the early 1990s (Zhang and Staehelin 1992). In that work researchers identified locations within the cell where specific forms of hemicellulose polysaccharides are present. The development of antibodies that recognize specific polysaccharide moieties has allowed the locations of xyloglucan backbone synthesis and the addition of specific sugar substitutions to be determined with some degree of resolution (Chevalier *et al.* 2010).

With the advent of fluorescent protein confocal microscopy, researchers are now able to observe the trafficking and localization of polysaccharide biosynthetic enzymes in living plant cells. Fluorescence microscopy can thus confirm the presence of the biosynthetic machinery for polymers that have previously been immunolocalized to Golgi stacks. In addition, for those enzymes that have already been identified, protein fluorescence microscopy can provide insight into how these proteins are organized within the secretory pathway. Localization of *Arabidopsis* proteins involved in the biosynthesis of mannan and xyloglucan to Golgi stacks is demonstrated in this work. The localization and organization within the Golgi of various enzymatic steps in the biosynthesis of these polymers is discussed at length in Chapter 3.

In addition to confirming the localization of known biosynthetic enzymes, fluorescent

protein microscopy can help to determine the function of proteins with unknown activity. For example, localization of CSLD proteins at the Golgi (Bernal *et al.* 2007; Bernal *et al.* 2008; Li *et al.* 2009; Zeng and Keegstra 2008) is evidence that this family of proteins probably produces a product in the Golgi, potentially narrowing the range of candidate polysaccharides that can be synthesized by these proteins. In contrast, fluorescence localization of the barley CSLC2 protein at the plasma membrane has been interpreted to mean that this particular CSLC homolog may not be involved in the production of xyloglucan (Dwivany*, et al.* 2009).

Secretory pathway: general description

The endomembrane system of plant cells is organized somewhat differently from other well-studied eukaryotes (yeast and animals). The structural organization of the plant endomembrane system is likely related to the unique architectural characteristics of the plant cell. Most plant cells have a very large lytic vacuole that surrounded by a thin cytoplasm bounded by the plasma membrane and a rigid cell wall (90% the volume of a typical parenchyma cell is occupied by the vacuole (Burgess 1985)). The cytoplasmic region of the plant cell can thus be visualized as a hollow sphere composed of a small volume (as little as 10% of the cell) occupying a relatively large area between the tonoplast and plasma membrane of each cell. In a typical plant cell, the secretory pathway is characterized by hundreds of Golgi stacks closely associated with an endoplasmic reticulum (ER) that extends from the nuclear envelope throughout the cytoplasm (Dupree and Sherrier 1998; Hanton and Brandizzi 2006). This organization allows dynamic distribution of membrane components and soluble cargo to specific locations in the cell by ensuring that transport machinery is always near target destinations.

The Golgi stacks function as central sorting apparati, packaging cargo for delivery to various locations within the cell. Of course, the Golgi is also a metabolically active organelle, being a location of some protein glycosylation and the synthesis of non-cellulosic cell wall

polysaccharides. In plant cells, Golgi bodies are closely associated with the cytosolic surface of the ER membrane. Each Golgi body consists of stacks of cisternae, which can be grouped into morphologically and functionally distinct regions called *cis*, medial, or *trans*-Golgi based on proximity to the ER. There is no direct evidence of an intermediate compartment between the ER and Golgi, and although movement of Golgi stacks is cytoskeleton-dependant, transport of cargo from the ER to the Golgi is not (Boevink *et al.* 1998; Brandizzi *et al.* 2002; Nebenführ and Staehelin 2001; Neumann *et al.* 2003; Saint-Jore *et al.* 2002). Because of their tight association and the absence of an obvious intermediate compartment, it can be useful to think of the plant ER-Golgi network as a single system for exchanging and sorting cargo.

The *trans* Golgi network (TGN) is a distinct structure that sometimes appears associated with the *trans* face of Golgi stacks (Griffiths and Simons 1986; Gu *et al.* 2001). In the plant cell, TGNs can also be found trafficking independently from the Golgi, disassociating and re-associating with Golgi bodies, and a single TGN can even be shared between two Golgi stacks (Viotti *et al.* 2010). Although direct evidence of secretory cargo passing from the Golgi through the TGN to the plasma membrane is somewhat limited, the TGN has historically been conceived of as a sorting station for cargo destined for the plasma membrane or vacuole (Gu*, et al.* 2001). The presence of fucosylated xyloglucan epitopes (Zhang and Staehelin 1992) is one of line of direct evidence for the plant TGN as a post-Golgi compartment carrying cargo destined for exocytosis.

There are significant differences in the gross organization and appearance of the endomembrane system between plant and animal cells due to differences in cell architecture, as discussed above. In addition, the plant TGN, in contrast with yeast and animal cells, also seems to function as the early endosome, directly receiving recycled plasma membrane and endocytotic cargo (Viotti*, et al.* 2010). These important differences notwithstanding, the individual Golgi

bodies of all eukaryotic cells have many features in common. The Golgi is where protein glycosylation and the synthesis of major components of the extracellular matrix occurs across kingdoms. The morphological similarity of Golgi structure among eukaryotes suggests that the biogenesis and organization of this organelle may occur through similar mechanisms. In support of this idea, several fluorescent protein constructs used as Golgi markers show the same localization across kingdoms.

Vesicle trafficking in the secretory pathway: specific components and regulation

Transport of cargo from the ER to Golgi in *Arabidopsis* occurs in a SAR1/COPII dependent manner, though no independent COPII-coated vesicles have been observed in plant cells. It is thought that, like their yeast homologs, components of the plant COPII complex (Sec23p/Sec24p, or Sec13p/Sec31p) directly or indirectly recognize soluble cargo and membrane proteins destined for the Golgi (Barlowe 2003; Jurgens 2004). ER-derived COPII vesicles, which can be observed in animal cells, fuse with the *cis* Golgi. ER-resident proteins that make it into the Golgi are recruited by H/KDEL receptors such as the protein ERD2 (ER retention-deficient), and packaged into COPI coated vesicles for ARF1-dependant retrograde transport from the Golgi to the ER.

Transport and organization within the Golgi is clearly more complex than the cotamer/GTPase mechanism that regulates ER-to-Golgi vesicle trafficking. This complexity is necessary because the Golgi functions both as a biochemically active organelle and a sorting station controlling the distribution of material to myriad locations within the cell. How the organization of the Golgi is established and maintained has long been an topic of debate. Two opposing models have classically been used to explain Golgi structure and function. One is known as the vesicular transport model (Pfeffer and Rothman 1987), where cargo and enzymatic components are packaged into (COPI coated) vesicles for anterograde transport between

cisternae. In this model, molecules enter the *cis* Golgi by vesicle fusion. Cargo destined for locations further along the secretory pathway is then collected and packaged into vesicles which bud off to fuse with the next Golgi cisterna in a *cis* to *trans* direction. In this model, each cisterna is a discrete, stable structure, passing along cargo by vesicle trafficking and retaining resident enzymatic activities and lipid components because these are not selected for export to subsequent stacks.

The other classical model for Golgi structure and function is known as the cisternal maturation model. In this model, Golgi cisternae on the *cis* face are created by the fusion of ER-to-Golgi transport vesicles and vesicles retrieving *cis* Golgi-resident cargo from subsequent stacks. At the same time, cisternae on the *trans* Golgi face are consumed as they are converted to TGN, secretory vesicles, and retrograde vesicles that serve to maintain *trans* Golgi character. Thus in the cisternal maturation model, the Golgi can be imagined as a conveyer system, where cisternae move through the stack, constantly being consumed on the *trans* face and replenished on the *cis* side. Retrograde vesicle transport maintains the proper composition and function of each cisterna (Pelham 1998).

The cisternal maturation model has become the more widely accepted of the two classical models, though there have even been attempts to reconcile the two, such as is described in the review by Pelham and Rothman (2000). However, neither the vesicular transport nor the cisternal maturation model can completely explain observed behavior of Golgi stacks and secreted material. For example, the Golgi transport vesicles thought to carry anterograde cargo in the vesicular transport model are too small to accommodate aggregated cargo, such as the collagen precursors seen in Golgi cisternae of animal cells (Bonfanti *et al.* 1998). However, cisternal maturation does not proceed fast enough to be the major route of transport for many secreted proteins that are known to pass through the Golgi (Pelham and Rothman 2000).

Recently, Patterson *et al.* (2008) proposed a third model for the creation and maintenance of the spatial organization of the Golgi. This new model conceives of Golgi organization as arising through thermodynamic two-phase partitioning of secretory cargo and Golgi-resident material. The two-phase model is particularly elegant in its simplicity and its ability to predict the behavior of molecules in the Golgi.

Employing fluorescence recovery after photobleaching (FRAP) techniques, Patterson *et al.* (2008) found that cargo does not exit from the Golgi at a linear rate following a specific delay, as would be predicted by the cisternal maturation model. Rather, proteins targeted to the plasma membrane are depleted from the Golgi with exponential kinetics. In other words, export rates are proportional to the total amount of a given cargo protein present in the Golgi. Also, newly synthesized proteins targeted for secretion are rapidly distributed throughout the Golgi in less than 5 min, rather than proceeding sequentially from the *cis* to *trans* cisternae, as both the cisternal maturation and vesicular transport models would predict (Patterson *et al.* 2008, Pelham and Rothman 2000). Furthermore, the researchers observed no preferential localization of secreted proteins to the *trans* Golgi (Patterson *et al.* 2008).

Since the observed behavior of Golgi cargo does fit the traditional models, Patterson *et al.* (2008) generated a simplified model of the Golgi apparatus based on their observations. In their initial minimal model, they draw upon the observation that Golgi resident enzymes do not perfectly colocalize with secreted proteins, though both can be found in the Golgi. They suppose that this occurs because Golgi membranes have two distinct domains: export subdomains where packaging of cargo for transport out of the Golgi occurs, and so-called "processing" domains that are responsible for maintaining the local characteristic lipid environment and enzymatic activity. Exported cargo rapidly exchanges between these domains, but Golgi-resident enzymes are mostly restricted from the export domains.

A second property of the two-phase partition model proposed in Patterson *et al.* (2008) is that the lipid composition of Golgi membrane export regions is distinct from the processing regions. Localization of membrane proteins to either of these regions is determined by free-energy preferences of the proteins' transmembrane domains (TMDs) for one lipid environment over the other. It has been known for some time that the lipid composition of Golgi membranes is not homogenous, and some lipid-modifiying enzymes localize to specific subdomains of the Golgi (van Meer and Sprong 2004). Also, Golgi membrane lipids are delivered by vesicles that originate at different locations in the cell and therefore have different compositions (Gkantiragas *et al.* 2001; van Meer and Sprong 2004). The observation that vesicles adjacent to different regions of the Golgi (*cis*, medial, or *trans*) are distinct in their size and coat components (Donohoe *et al.* 2007) could indicate that these vesicles arise from distinct sources and therefore have distinct lipid composition.

The observations of Donohoe *et al.* might be combined with those of Patterson *et al.* to connect the two-phase hypothesis of Golgi transport with the trans-to-*cis* maturation of Golgi cargo, but this idea has not been explored in the literature. Figure 2 shows a schematic representation of a model of intra-Golgi trafficking that combines the two-phase model (Patterson*, et al.* 2008) with the observation that peri-Golgi vesicles have distinct compositions depending on what regions of the Golgi are adjacent (Donohoe*, et al.* 2007). Since Golgi-resident enzymes partition to the proper cisternae in part through interactions between TMDs and the local lipid environment (Brandizzi*, et al.* 2002), this results in an adapted two-phase model that includes the *cis* to *trans* gradient of lipid composition that could drive a gradient of enzyme activities. Although it is not clear how Golgi-resident proteins are able move through the stack, this may occur through vesicle budding and fusion, though transient or stable connections between cisternae, or a combination of these. Patterson *et al*. (2008) find that secreted proteins

do not preferentially localize to the *trans*-Golgi, and thus could be packaged into vesicles that might originate at any of the cisternae. However, proteins targeted for direct trafficking to the plasma membrane could follow a different route from Golgi cargo such as cell wall polysaccharides, which may require sequential processing in the Golgi.

The model presented in Figure 2 describes one possible way to reconcile the new paradigm of Golgi organization posited by Patterson *et al*. (2008) with the classic models of *cis* to *trans* maturation of Golgi cargo. If this model is accurate, then Golgi-resident enzymes, such as those that synthesize hemicellulosic polysaccharides, localize to subdomains of the Golgi based on the local membrane environment and the cargo these enzymes modify would mature through the Golgi before being packaged into secretory vesicles primarily on the *trans* face of the Golgi. However, the model presented in Figure 2 has not been confirmed, and additional studies on the localization of cell wall biosynthetic enzymes and their products are required to reject or confirm such a model.

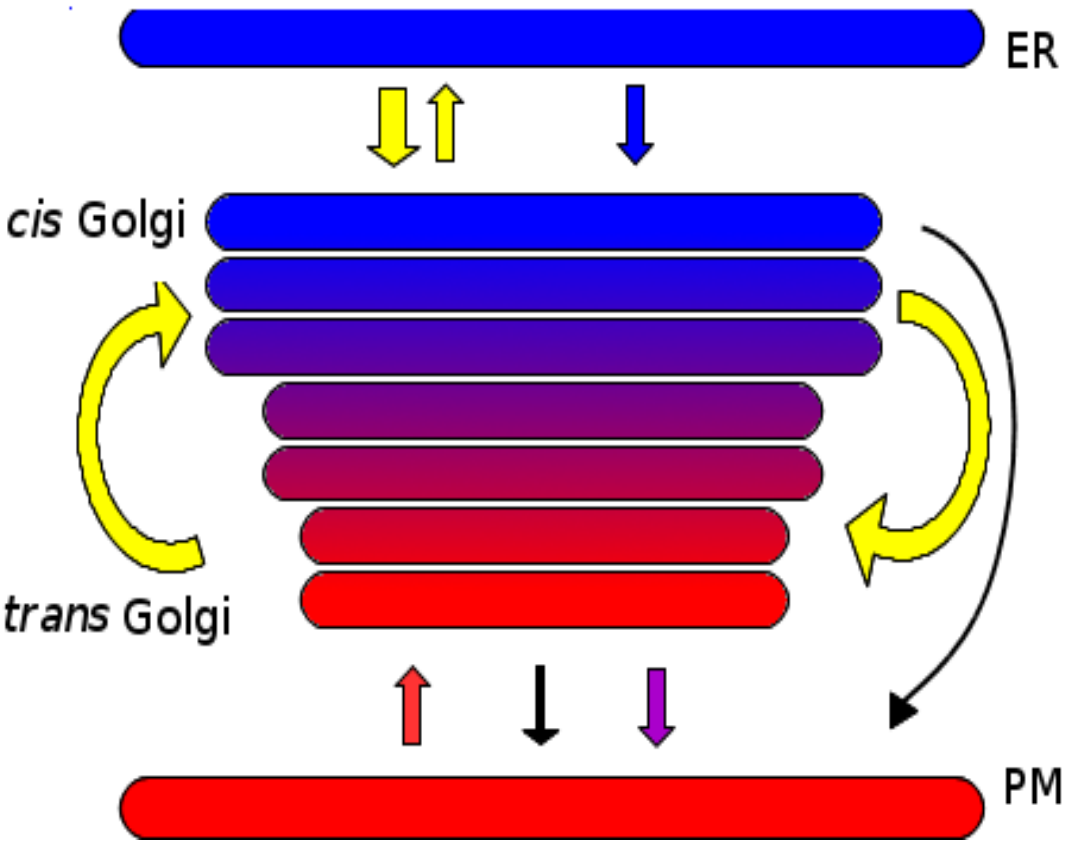

Figure 2. Model of trafficking in the Golgi
The colored ovals represent membranes of the ER, Golgi, and plasma membrane (PM). Blue shading indicates a lipid composition more similar to the ER, and red shading indicates lipid composition that is more similar to the PM. Yellow arrows show movement of Golgi-resident enzymes as they sample the local lipid environment and localize to subdomains of the Golgi where the gradient of lipids through the organelle creates the lowest free energy environment, which probably varies for each protein. Black arrows indicate possible routes for proteins targeted for secretion, and the purple arrow indicates a likely route for secreted cargo that undergoes processing in the *trans* Golgi. The red and blue arrows represent lipids entering the Golgi from the PM and ER, respectively. Note: For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

Membrane insertion of proteins in the secretory pathway

Typically, integral membrane proteins contact three distinct types of environment: at least one aqueous compartment, the hydrophobic core of the membrane, and the interfacial region of the membrane, where lipid headgroups interact with the solvent. Each of these environments has distinct properties, imposing more complex conditions for attaining a favorable, low free-energy state upon membrane proteins compared to soluble proteins. The additional structural component membrane proteins have as a result of their environment that soluble proteins lack is known as membrane protein topology.

In its most basic expression, the membrane topology of a protein describes what parts of the protein intersect a membrane and what side of the membrane each extra-membrane region faces. Thus, membrane protein topology consists of two essential components: (1) the number and position of transmembrane domains, and (2) the orientation of the protein in the membrane. Understanding membrane topology is essential for understanding the function of individual membrane proteins and the biological processes with which they are involved. Some of what is known about how organisms accomplish TMD insertion and determine membrane protein orientation is described below. Also, some recent advances in the areas of topology prediction and determination are discussed.

Golgi-resident integral membrane proteins, such as glycosyltransferases involved in the biosynthesis of cell wall polysaccharides, are inserted into the membrane of the ER during translation. This process begins when a soluble GTPase known as the signal recognition particle (SRP) recognizes and binds to a nascent polypeptide emerging from the ribosome, causing translation to arrest. The SRP then functions to recruit the ribosome to a ER-membrane-bound GTPase dubbed the SRP receptor (SR), which in turn drives association of the ribosome with the ER translocon complex. Binding of SRP to SR stimulates hydrolysis of GTP by both proteins,

which results in their disassociation from the ribosome now docked at the translocon complex, and translation proceeds.

In eukaryotes, the translocon is a trimer of Sec61 proteins (α, β and γ subunits). The α subunit of Sec61 has ten TMDs, whereas the β and γ subunits usually have a single TMD each. Much of what we know about the fine structure of the eukaryotic translocon is based on structural studies of the highly similar translocon complex protein homologs from eubacteria (SecYEG). From a perspective perpendicular to the membrane, the translocon complex resembles a horseshoe, the transmembrane domains forming a ring that is open on one side to the lipid bilayer. A "collar" of hydrophobic residues is found at the middle of the pore formed by the TMDs of the translocon complex, apparently serving to make a seal around translocating polypeptides. When the complex is in a closed confirmation, i.e. when it is not actively translocating a protein emerging from a docked ribosome, a short helix on the lumenal side of the hydrophobic collar is thought to function as a "plug" that prevents diffusion across the ER membrane. The opening between the legs of the "horseshoe" is the only conceivable location where a TMD of a newly-synthesized protein can exit the translocon and enter the ER membrane bilayer. However, the crystal structure has only been determined for the closed complex, so it is not clear what conformational changes occur during translation and TMD insertion. (Johnson and van Waes 1999; Rapoport *et al.* 2004)

Based on the structure of the translocon complex, one can imagine a mechanism by which TMDs are selected and inserted to the membrane through a free-energy partitioning process. In other words, the hydrophobic region of the ER membrane is available via the "horseshoe" opening to nascent helixes passing through the translocon, and newly-synthesized peptide helixes sample the hydrophobic lipid bilayer. A TMD will either enter the bilayer by lateral diffusion or remain within the hydrated translocon pore. This type of model for TMD

selection and insertion is supported by cross-linking studies (Hessa *et al.* 2005).

The free-energy partitioning model of TMD selection and insertion might lead to the conclusion that hydrophobicity of the peptide functional groups is the determining factor for whether a particular peptide sequence is part of a TMD or resides in an aqueous compartment. Although side-chain hydrophobicity definitely plays role in TMD selection, the interactions at the translocon that determine partitioning of nascent polypeptides to the ER membrane are clearly more complex. One obvious departure from the simple idea that more hydrophobic regions prefer insertion to the membrane is the fact that amino acid side chains do not fall into simple categories of polar or non-polar. For example, the aromatic residues tyrosine and tryptophan, though not highly charged, have a preference for the membrane-solvent interface, and the addition of these aromatic residues to the middle of a TMD significantly reduces the probability of membrane insertion. When positioned in the interfacial region at the ends of a TMD however, these residues increase the probability of insertion (Hessa *et al.* 2007). Thus, there is a clear effect of a residue's position within the TMD on how the side chain influences TMD selection.

The position effect applies to polar residues as well as aromatic or small hydrophobic functional groups. Based on modeling studies, charged amino acids buried in an inserted TMD are able to interact with lipid headgroups via salt bridges, or even force "snorkeling" of the membrane to interact with water molecules in the solvent phase. Because of this, the presence of a polar group within a TMD does not completely preclude insertion, but rather imposes a specific free-energy "cost" on the probability of membrane insertion based on the residue's position, the length and composition of the TMD, and to some extent the specific lipid environment in regard to both fatty acid and headgroup composition (Hessa*, et al.* 2005).

The simplest possible case for the selection and insertion of a TMD would be a type I

membrane protein. These proteins have a single TMD oriented with the amino-terminal side of the TMD nearest to the ER lumen, so the TMD would already be oriented properly as it passes through the translocon pore amino-terminus-first, and could simply enter the membrane through lateral diffusion as it samples the lipid environment. Translocation and membrane insertion of the polytopic protein aquaporin-4 (AQP4), described below, illustrates how complex membrane insertion becomes when considering proteins with multiple TMDs.

Sadlish *et al.* (2005) used a strategy based on the incorporation of a photo-activatable crosslinker to the polytopic protein aquaporin-4 (AQP4) to study the dynamics of membrane insertion of a protein with multiple TMDs. Sadlish *et al.* created numerous radiolabeled AQP4 variants, combining the addition of a crosslinker to three sites in each TMD (by the so-called amber codon approach) with the addition of premature stop codons introduced at a wide range of locations 3' of the crosslinker codons. In separate experiments, they translated each construct in vitro in the presence of canine pancreatic microsomes. They then performed immunoprecipitation against SEC61α following UV activation of the crosslinker and evaluated the presence of AQP4 protein. Because they used a set of serially truncated proteins, each experiment produced a steady-state pool of truncated AQP4 protein crosslinked to the translocon.

The presence of truncated AQP4 protein in SEC61α immunoprecipitations indicated that the AQP4 TMD with a crosslinker is associated with the translocon when translation halts at the introduced stop codon. This should reflect the position of that TMD in the native polypeptide at the instant of amino acid incorporation at the position substituted with a stop. Thus, the various truncations allowed the researchers to determine how long each TMD is in contact with the translocon complex.

Their results showed that each TMD interacted with the translocon complex sequentially in the order of synthesis (amino to carboxyl terminus of the nascent polypeptide), but that the

interactions were unique and sometimes dynamic. For example, the first TMD of AQP4 shows two independent cycles of reassociation and disassociation from the translocon following its initial insertion to the membrane, returning to the complex while TMD2 and TMD3 are being translocated. In contrast, the second TMD of AQP4 only appears to associate with the translocon once during synthesis and translocation.

During the insertion of AQP4, separate TMDs in a single polypeptide interact with each other during translocation (Sadlish*, et al.* 2005). These interactions may play a role in the arrangement and fold structure of helixes within the membrane as well as regions of the protein found in aqueous compartments. Indeed, some *bona fide* TMDs within polytopic proteins like AQP4 may not partition to the membrane during translocation in the absence of interactions with another, possibly distal TMD.

Prediction of transmembrane domains

A long-standing pursuit in the field of bioinformatics has been the reliable prediction of TMDs from amino acid sequences. The first transmembrane domain prediction algorithms were based solely on hydrophobicity of amino acid side chains (von Heijne 1994). These methods met with limited success because hydropathy is not the only determining factor for partitioning to the membrane. As empirical data from studies on actual transmembrane proteins accrued, TMD prediction algorithms were improved by machine learning methods through statistical sampling and analysis of biochemically characterized TMDs. Some recent efforts to improve TMD prediction (Bernsel *et al.* 2009; Hessa *et al.* 2007) have focused on attempting to understand how polypeptide composition drives TMD selection and insertion from the perspective of the physico-chemical properties of amino acid side chains.

The study by Hessa *et al.* (2007) is particularly groundbreaking in the field of TMD prediction because it was among the first to empirically measure the affect of all 20 amino acids

on the apparent ΔG of membrane insertion. Briefly, each of the naturally-occurring amino acids was scanned through an artificial transmembrane segment in linearly and symmetrically (with respect to the TMD center), and the effect on the efficiency of TMD insertion was evaluated. The functions derived from these studies provide valuable information about the contribution of individual amino acids to the apparent free energy of TMD insertion, and represent a significant advance in TMD prediction. Even if the reliability of the new algorithms is not better than existing prediction programs, they are preferable in that they do not rely on a "black box" methodology such as machine entrainment. When one predicts TMDs with utilities based on the theoretical composition-based apparent ΔG, such as predDG or SCAMPI, the results can be explained based on the fundamental properties of peptides. However, there are still improvements to be made to these methods, as predDG will sometimes predict transmembrane domains in hydrophobic regions of a globular domain that do not insert to the membrane.

The fact that TMDs of polytopic membrane proteins do not necessarily insert independently from each other is a significant obstacle in the pursuit of reliable TMD predictions. Also, protein segments that have a favorable ΔG for membrane insertion may not necessarily partition to the membrane, but could be buried in the hydrophobic core of a protein structure that nevertheless resides in a aqueous compartment. Eventually, we may have enough knowledge about the behavior of each amino acid side chains in a membrane, given their individual chemical properties and positions within a potential TMD, to accurately predict the free energy preference of a single TMD. However, it may be some time before TMD prediction algorithms reach the sophistication necessary to accurately predict the position and extent of TMDs in polytopic membrane proteins, given the enormous complexity of factors contributing to TMD membrane insertion.

Membrane Protein orientation

In addition to the number and position of TMDs, an explanation of topology of a membrane protein must include a description of the orientation of the protein. Membrane protein orientation describes the direction of each TMD embedded in the membrane and whether the hydrophilic regions of the protein are located facing the cytosol or another subcellular compartment. Each helical TMD that completely spans the membrane can adopt one of two possible orientations, with either the amino or carboxyl end of the TMD facing the cytosol. Descriptions of membrane topology of polytopic proteins, such as CSL family proteins, can be very complex. For example, the introduction or removal of a single TMD by natural or artificial means can cause complete inversion, partial inversion, or have a minor effect on the orientation, depending on where in the protein sequence the change occurs. The orientation aspect of topology is further complicated by the fact that membrane domains do not always extend across the membrane, but sometimes enter and exit the bilayer on the same side.

It is not known precisely how organisms determine the orientation of membrane proteins, but as a general trend, membrane proteins tend to follow the "positive inside rule," described by von Heijne (1989). This rule reflects the trend that positive residues occur with much higher frequency in regions of membrane proteins facing the cytosol. First observed in bacteria, the positive inside rule also applies in eukaryotes. In the case of endomembrane proteins, the lumen of secretory pathway organelles and the vacuolar space are considered to be analogous to the apoplast for the purposes of the positive inside rule, and are thus "outside" locations.

The molecular mechanism responsible for the positive inside rule are poorly understood. Several studies of the membrane topologies of *E. coli* transporter proteins with twelve membrane domains have indicated that the lipid composition of the membrane may play some role in determining protein topology (Bogdanov *et al.* 2002; Bogdanov *et al.* 2008; Zhang *et al.* 2003; Zhang *et al.* 2005). In one of these studies (Bogdanov*, et al.* 2008), the membrane

topology of a lactose permease (LacY) of *E. coli* was examined in wild-type cells and cells unable to synthesize the major *E. coli* lipid phosphatidylethanolamine (PE). The organization of LacY in wild type cells that produce PE is characterized by twelve integral membrane regions. The second membrane domain from the amino terminus forms a re-entrant loop, connecting two hydrophilic domains on the same side of the membrane. In mutant *E. coli* cells unable to synthesize PE, TMD2 is inserted as a membrane-spanning domain, and TMD7 becomes associated with the cytoplasmic face of the membrane. Surprisingly, PE synthesis by induction of a biosynthetic enzyme caused LacY protein that had previously been synthesized to revert to the wild type confirmation. This indicates that membrane topology can be a dynamic property, and is not necessarily determined solely during membrane protein insertion.

Investigations like the LacY study described above indicate that membrane protein topogenesis involves membrane lipids in addition to protein components of the translocon apparatus (reviewed in (Bogdanov *et al.* 2009)). These studies also provide the beginning of a framework for understanding the molecular mechanism of the positive inside rule. One of the conclusions drawn by Bogdanov *et al.* (2008) is that the relative charge density of anionic lipid headgroups is a key factor in membranes' contribution to topogenesis. Thus, their explanation for the role of zwitterionic PE is to dilute the negative character of the membrane interfacial region contributed by components such as phosphatidylglycerol and cardiolipin. Anionic charge density as a topological determinant also offers an explanation of the positive inside rule, as a nascent integral membrane protein has the opportunity to interact with the cytosolic face of the membrane first. Also, this model explains why retention of soluble regions with cationic character on the cytosolic side of the membrane is a stronger orientation determinant than translocation of net negatively-charged regions (Bogdanov*, et al.* 2009).

Dual-topology proteins that insert in both opposite orientations have proven to be an

invaluable tool for learning how protein sequence determines the orientation aspect of membrane protein topology. A well studied member of the small multidrug-resistance protein family of *E. coli*, EmrE, is such a protein. EmrE protein has four membrane-spanning TMDs, and it inserts to the membrane in both opposite orientations with a stoichiometry ratio of approximately 1:1 (Rapp *et al.* 2006). As would be expected for a dual topology protein, EmrE has little bias for positive charges in the hydrophilic regions. By the addition and removal of positive and negative residues on the hydrophilic regions of EmrE, researchers were able to create EmrE variants that predictably insert in one or the other of the two possible orientations (Nasie *et al.* 2010; Rapp *et al.* 2007; Schuldiner 2009). It is even possible to control the membrane orientation of EmrE with a single substitution at the carboxyl terminus (Seppälä *et al.* 2010), indicating that peptide sequence contribution to membrane protein orientation can be very subtle.

Though a reasonably reliable predictor of orientation in bacteria, the positive inside rule does not apply in every case. In eukaryotes, combined contributions of both positive and negatively-charged residues to the net charge of each loop must be considered, and there are many examples of cytoplasmic domains with net negative charge that nonetheless reside in the cytosol (von Heijne 2006). Because of the complexities inherent to membrane protein insertion and topological determination, we are unable to predict these properties based on protein sequence alone. The complex thermodynamic interactions that determine membrane protein topology, described above, and the contribution of as-yet uncharacterized factors such as protein chaperones may mean that reliable and accurate prediction methods are all but impossible, simply from the perspective of computational practicality. Ever-increasing examples of close homologs with distinct topologies (Chapter 2), dual-topology proteins (Nasie*, et al.* 2010), and proteins with dynamic topologies (Goder *et al.* 1999), indicate that at least for the near future, membrane protein topology must be experimentally determined on a case-by-case basis.

Over the years, a number of techniques have been developed to analyze protein topology. From the perspective of experimental design, it is important to attempt multiple approaches to probe protein topology. One reason for this is that most of these approaches rely on producing proteins in a non-native system by heterologous expression or in vitro translation. Because every protein has unique properties, predicting whether a particular protein will be produced in a given expression system is impossible. Even closely-related proteins can have dramatically different behaviors in the same expression system. For example, the CSLA9 and CSLC4 proteins discussed in Chapter 2 express to different levels in Pichia pastoris, and each requires different induction regimes to allow production of sufficient protein for downstream analysis. In addition to the unpredictability of heterologous systems, membrane protein topology studies almost always require that the protein sequence is altered in some way. Because minor changes to protein sequence have been shown to affect topology in some cases (Seppälä*, et al.* 2010), employing multiple approaches increases confidence in the veracity of experimental results and interpretations. Some techniques for studying membrane protein topology are briefly described below. Each of these has specific advantages and disadvantages, which underlines the need to employ multiple complementary approaches where possible.

One of the more widely-used approaches for probing membrane protein topology exploits the fact that regions of protein inserted in ER membranes that are exposed to the lumen may be glycosylated at asparagine residues (Asn-X-Ser or Asn-X-Thr, where X is not Pro). This glycosylation, which can be reversed by treatment with a specific glycosidase, results in a shift in the protein's migration rate in a separation method such as SDS-PAGE (Devoto *et al.* 1999; Maley *et al.* 1989). There are several advantages of using N-linked glycosylation to probe membrane protein topology. For example, full-length protein can be analyzed, and the simplicity of the NXS/T motif means that individual artificial glycosylation sites can be added without

large changes to protein sequence. However, for N-linked glycosylation analysis to produce unambiguous results, all native glycosylation sites should be removed by mutagenesis. Removal of every native site can amount to significant changes in protein sequence and can be very time consuming. Also, there is some evidence that glycosylation can influence topogenesis (Goder*, et al.* 1999).

A relatively new technique for probing membrane protein topology relies on a redox-sensitive form of GFP (roGFP) (Brach *et al.* 2009). roGFP variants were created by the introduction of redox-sensitive cysteine residues to surface sites near the GFP fluorophore (Hanson *et al.* 2004). The oxidation state of these residues influences the excitation and emission spectra of the fluorophore. Because the lumen of the early secretory pathway (ER and Golgi) is more oxidizing than the cytosol, analysis of the fluorescence spectra of roGFP can indicate where the fluorescent tag lies. Constructs can be created such that the roGFP is fused to a protein of interest, becoming a reporter for the orientation of the region of the protein to which it was fused. An advantage of using roGFP for topology studies is that membrane protein topology can be analyzed in living plant cells. This property would be particularly useful in cases where the topology of a protein is dynamic, in response to external stimulus for example. An obvious disadvantage of the roGFP approach to characterization of membrane topology is the size of the fluorescent protein tag. The addition of a peptide sequence approximately 30 kDa in size might have unpredictable effects on the topology of a protein of interest. Changes to topology that might result could even be restricted to regions near the tag, meaning that the functionality of a tagged protein of interest could be confirmed by a method such as mutant phenotype complementation, but the interpretation of local topology based on roGFP spectra could still be incorrect.

A classic technique to investigate membrane protein topology is protease protection. In

this approach, intact membrane vesicles containing a protein of interest are isolated and treated with low levels of protease so that sites external to the vesicles are available to the protease for cleavage, while the vesicle membrane prevents the protease from accessing sites in the vesicle lumen (Gilstring and Ljungdahl 2000; Kageyama-Yahara and Riezman 2006; Muñoz *et al.* 1996; Urbanowicz *et al.* 2004; Zeng and Keegstra 2008). Detergent can be added to permeabilize the vesicles and allow the protease to access sites on the other side of the membrane. Despite the apparent simplicity of the protease protection approach, experimental conditions must be carefully designed and confirmed. For example, high levels of protease can disrupt membrane integrity and lead to the incorrect conclusion that internal sites are on the outside of the vesicles. Also, to maintain intactness and right-sidedness of vesicles during cell fractionation, levels of salt and osmoticum must be carefully controlled. All reaction conditions, i.e. concentrations of protease, detergent, and osmoticum, and vesicle fractionation methods, must be carefully controlled and verified with multiple independent methods. Determination of the membrane protein topologies of *Arabidopsis* CSLA9 and CSLC4 by protease protection techniques is presented and discussed in the following chapter.

Biological membranes are found in every living organism, and the protein components of these membranes are essential for their function. Approximately 25% of expressed sequences in the *Arabidopsis* genome are predicted to contain at least one transmembrane domain (Schwacke *et al.* 2003). Although astounding advances have been made in our understanding of how membrane proteins are synthesized, inserted, and trafficked in the cell, we are still unable to predict which regions of a protein are in the membrane and how a protein is oriented. Since rigorous determination of topology is only possible using methods like those described above, exhaustive characterization of the position and orientation of every transmembrane domain is a daunting prospect even when considering a small number of proteins. In order to simplify this

problem when approaching a membrane protein of unknown topology, it is key to identify the most important parts of a protein to analyze. For example, the substrate-binding motif by which CSL proteins were identified resides on the largest hydrophilic region of these proteins. Knowing the orientation of this region is likely to be most useful for understanding how the topology of these proteins impacts the biosynthesis of cell wall polysaccharides.

The following chapters of this manuscript discuss the membrane protein topology and subcellular localization of enzymes involved in the synthesis of the hemicellulosic polysaccharides mannan and xyloglucan. Determination of the membrane topologies of the *Arabidopsis* mannan synthase CSLA9 and xyloglucan glucan synthase CSLC4 by protease protection methods are discussed in chapter two. The subcellular localization of these proteins, and enzymes catalyzing other steps in xyloglucan biosynthesis was examined by live cell confocal microscopy. A discussion of these data is presented in chapter 3.

CHAPTER 2 : MEMBRANE PROTEIN TOPOLGY OF CSLC4 AND CSLA9

Introduction

Considerable progress has been made recently in identifying the proteins responsible for the synthesis of the major hemicellulosic polysaccharides mannan and xyloglucan. Several research groups have provided evidence that the backbones of mannans and glucomannans are synthesized by the CSLA family of glycan synthases in many, probably all, plant species (Dhugga*, et al.* 2004; Goubet*, et al.* 2009; Liepman*, et al.* 2007a; Liepman*, et al.* 2005; Suzuki*, et al.* 2006). Cocuron *et al.* (2007) provided evidence that in *Arabidopsis*, and probably in other species, the glucan backbone of xyloglucan is synthesized by the CSLC4 protein. However, new evidence presented by Dwivany *et al.* (2009) indicates that some CSLC homologs in barley are located in the plasma membrane. Based on this observation and the sequence relationship among CSLC proteins, they propose that this family contains two clades, one group that includes *Arabidopsis* CSLC4 and its homologs, which synthesize the backbone of xyloglucan in the Golgi, and another group of CSLC proteins that are active at the plasma membrane and are probably involved in the synthesis of another glucose-containing polysaccharide.

All of the proteins in the CESA superfamily, including CSLA and CSLC proteins, are polytopic membrane proteins with multiple transmembrane domains (TMDs). Specifically, the CSLA and CSLC proteins are predicted to have 5 to 6 TMDs, with two internal hydrophilic regions. With all CESA superfamily proteins, including CSLA and CSLC proteins, the larger of these two hydrophilic loops contains the predicted active site motif (D,D,D,Q/RXXRW) associated with processive glycosyltransferases (Richmond and Somerville 2000; Saxena *et al.* 2001).

It has long been known that cell wall matrix polysaccharides, including xyloglucans and mannans, appear in Golgi vesicles before being deposited to the cell wall (Moore and Staehelin

1988; Zhang and Staehelin 1992). When the intracellular localization of recently identified proteins involved in the biosynthesis of xyloglucan and mannans has been examined, those proteins have been found in the Golgi apparatus (Madson *et al.* 2003; Wulff*, et al.* 2000).

Despite progress made in identifying the proteins that synthesize xyloglucans and mannans, many important questions remain about the biosynthesis of these polysaccharides. Understanding the properties of recently identified enzymes, such as CSLA and CSLC proteins, can provide insight into the biochemical and molecular mechanisms involved in hemicellulosic polysaccharide biosynthesis. For example, the topology of a CSL protein in the Golgi membrane profoundly impacts the mechanism of polysaccharide synthesis. A CSL glycan synthase oriented with its active site facing the Golgi lumen would most likely obtain nucleotide sugar substrates from the Golgi lumen and deposit its polymeric product there (see Figure 3, model a). Conversely, if the active site faces the cytosol, then the enzyme would use nucleotide sugars from the cytosol and transport the polymer product across the Golgi membrane (Figure 3, model b). This latter scenario is similar to the mechanism that is postulated to occur at the plasma membrane during the synthesis of the glucan chains that make up cellulose microfibrils (Delmer 1999; Guerriero*, et al.* 2010).

In this study, we investigated the topology of the CSLA and CSLC proteins by expressing epitope-tagged versions of these proteins in yeast (*Pichia pastoris*) cells. Following isolation of intact Golgi membranes, we treated the membranes with proteases to determine the sensitivity of the tagged proteins. We concluded that the topology of CSLC4, a putative xyloglucan glucan synthase, is characterized by a cytosolic active site and six transmembrane domains (Figure 3, model b); the most abundantly expressed mannan synthase in *Arabidopsis*, CSLA9, is predicted to have five TMDs and has an active site facing the Golgi lumen (Figure 3, model a). The implications of these observations for the mechanism of polysaccharide synthesis are discussed.

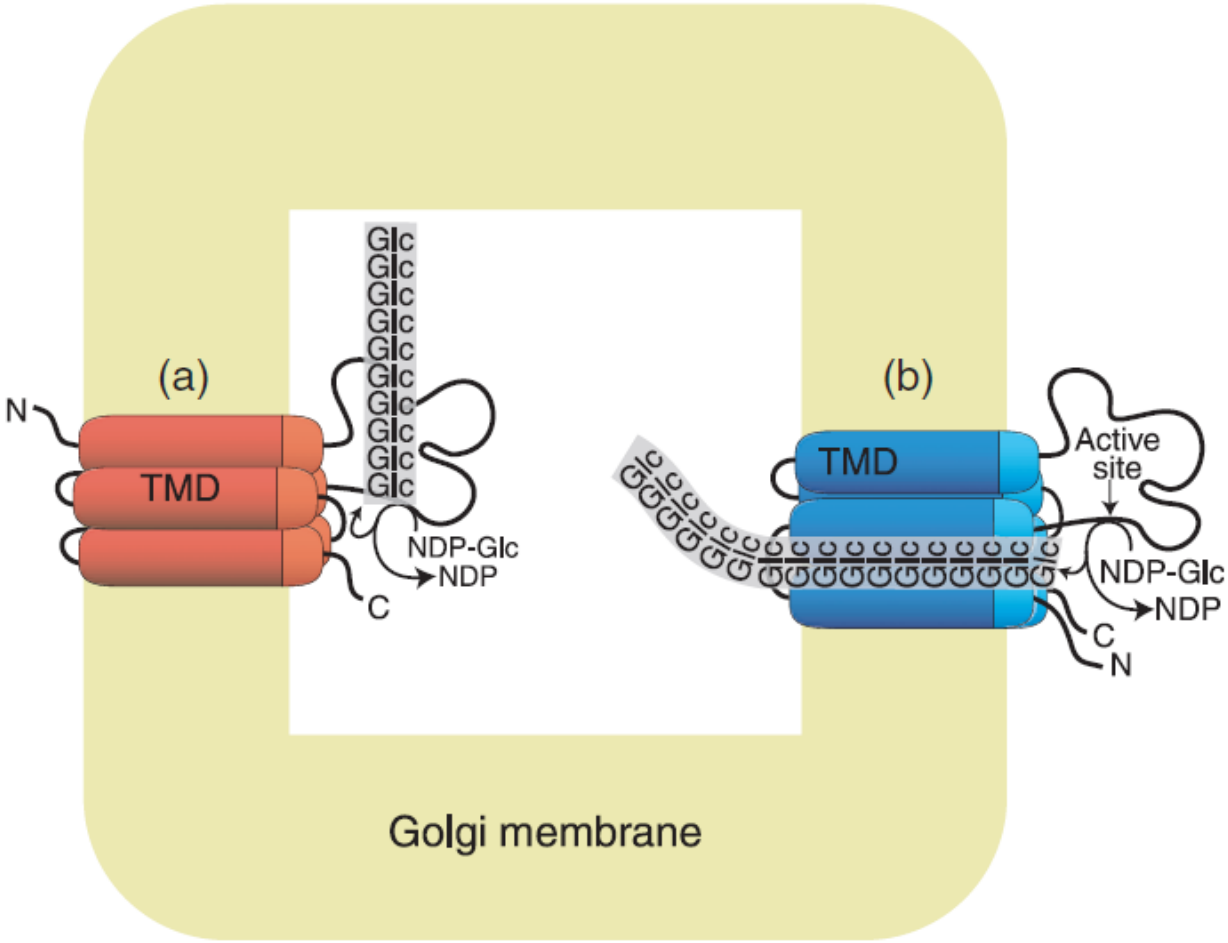

Figure 3. Schematic representation of CSL proteins.
Two CSL proteins are shown as if residing in the same Golgi vesicle (not to scale). Protein TMDs are represented by shaded cylinders, and hydrophilic regions are represented by lines. A polytopic glycosyltransferase with five transmembrane domains and an active site in the Golgi lumen, similar to CSLA9, is shown on the left, utilizing substrate and producing a product in the lumen. On the right, a polytopic glycosyltransferase with six transmembrane domains and a cytosolic active site, similar to CSLC4, is shown utilizing cytosolic substrate and depositing a polysaccharide product in the Golgi lumen.

Results

Topology of CSLA9 in the Golgi membrane

We investigated the topology of the CSLA9 protein in the membrane using protease protection experiments to probe the protease sensitivity of epitope-tagged versions of this protein expressed in *Pichia pastoris*. We made two constructs such that the expressed proteins had either one T7 epitope on the amino terminus (Figure 4a) or T7 epitopes on both the amino and carboxyl termini (Figure 4b). Cells producing the tagged proteins were disrupted, and a Golgi-enriched membrane fraction (microsomes) was prepared by differential centrifugation. Microsomes containing T7-tagged CSLA9 proteins were subjected to protease treatment in the absence or presence of a membrane-permeabilizing detergent. The presence of T7 epitopes on CSL proteins following protease treatment was evaluated using immunoblot analysis (Figure 4). To verify that the microsomal membrane vesicles isolated from *Pichia* were intact and that the protease protection conditions were appropriate, the *Arabidopsis* xyloglucan fucosyltransferase (FUT1), a type-II membrane protein with a known orientation (Wulff*, et al.* 2000), was used as a control (Figure 5).

Results from these protease protection experiments showed that in non-permeabilized microsome samples the amino terminus of CSLA9 was sensitive to protease treatment, while the carboxyl terminus was protected. Therefore the amino terminus faces the cytosol, and the carboxyl terminus is in the Golgi lumen.

Although the protease protection assays indicate the locations of the termini of the CSLA9 protein, it is also important to determine the location of the active site. TMD prediction algorithms provide a starting point for understanding the topology of a membrane protein (Figure 4c). However, the uncertainty inherent to such predictions necessitates experimental validation of protein topology. To this end, we exposed intact microsomes from cells expressing T7-

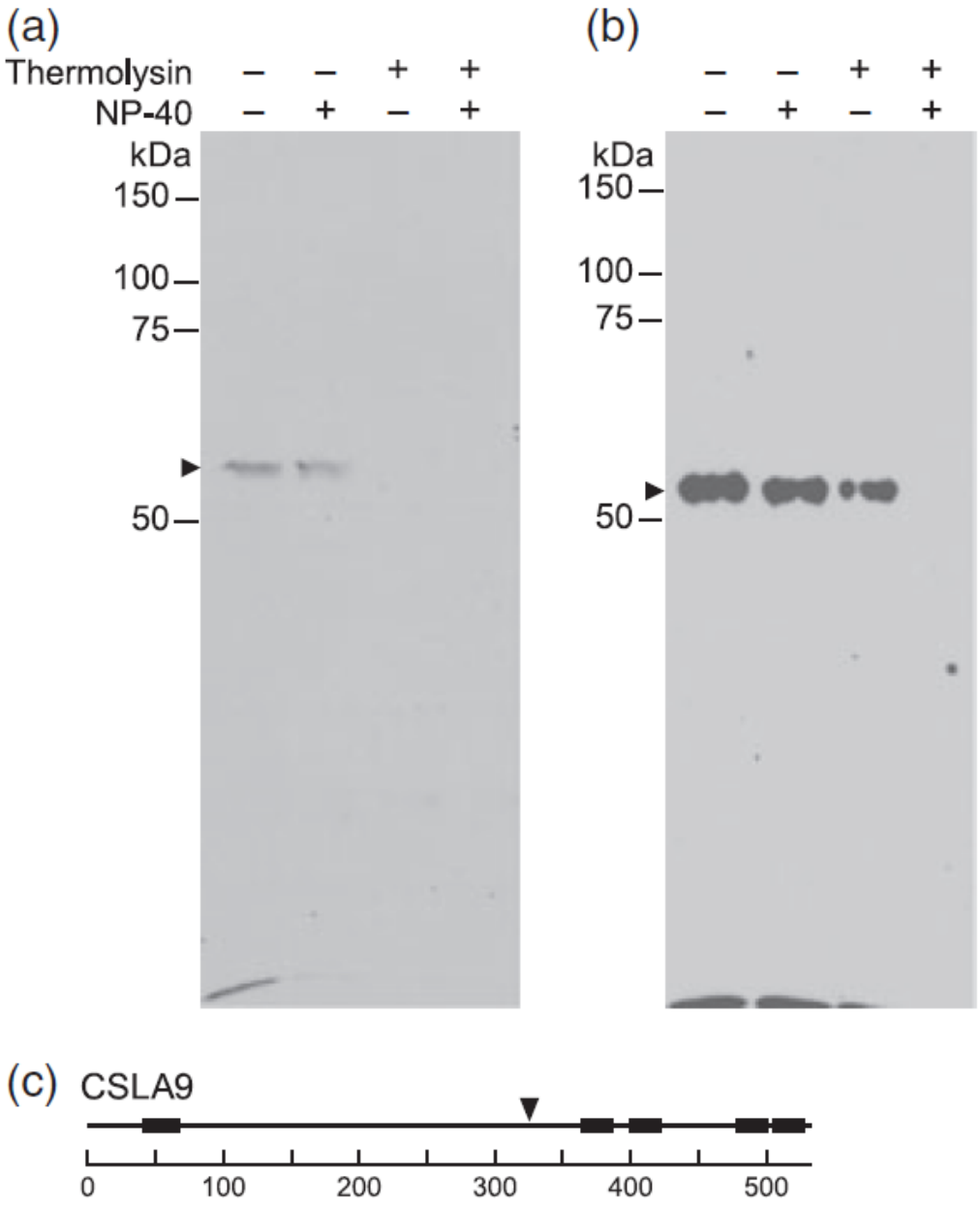

Figure 4. Protease protection of CSLA9 termini.
(a, b) CSLA9 proteins containing a T7 peptide epitope added to the N-terminus (a) or both termini (b) were expressed in Pichia cells, and microsomal membranes were isolated and subjected to protease treatment. The resulting microsomal proteins were separated by SDS–PAGE, and tagged proteins were detected by immunoblotting using antibodies against the T7 epitope. The labels above the blots indicate the protease protection assay conditions. CSLA9 migrates to approximately 55 kDa, as indicated by an arrowhead. (c) CSLA9 protein with TMDs predicted by TMMOD represented by black bars. The position of the sugar-nucleotide binding motif is indicated by an arrow.

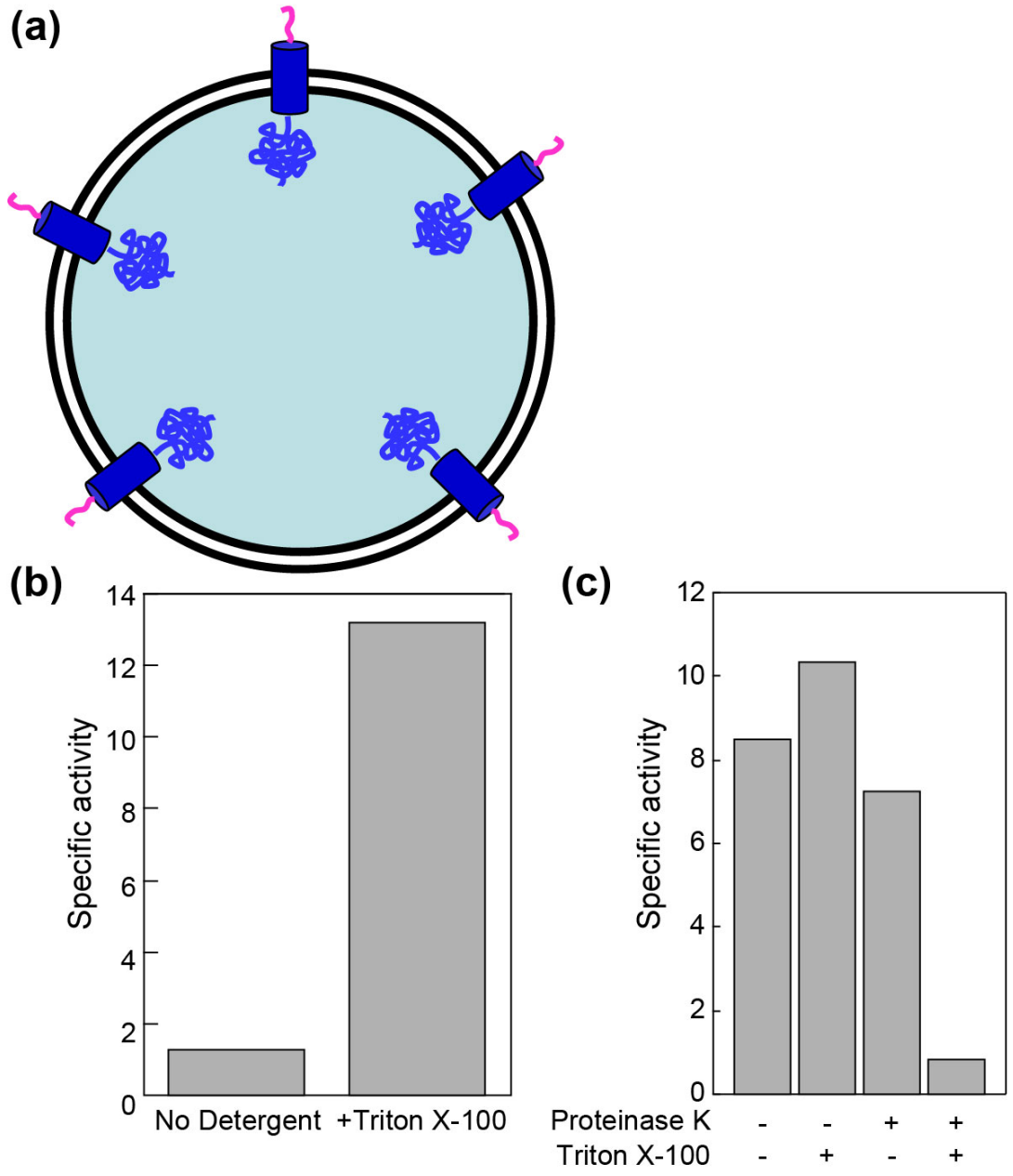

Figure 5. Latency of access to acceptor substrate shows that vesicles are intact
A schematic representation of the type II membrane protein, FUT1, is shown (a). Transfer of ^{14}C-fucose from UDP-fucose onto tamarind xyloglucan in the presence and absence of detergent (b). Permeation of the membrane with Triton X-100 is necessary for the enzyme to access the membrane-impermeable acceptor substrate (tamarind xyloglucan), indicating that Golgi vesicles are intact. FUT1 activity following protease protection is shown in (c). The Y-axis units are pmoles fucose per µg total protein per hour.

CSLA9 to protease in the absence or presence of detergent and then measured mannan synthase activity in the treated microsomes. Mannan synthase activity contributed by the CSLA9 protein is about 90% resistant to protease treatment in the absence of detergent (Figure 6), leading us to conclude that the active site is located in the Golgi lumen (as represented by model (a) in Figure 3). The decrease in mannose incorporation (approximately 10%, Figure 6, third bar) seen with protease-only treatment could be due to decreased mannan synthase activity from cleavage of external sites on CSLA9, or this decrease might represent a fraction of the vesicles that are not intact. This interpretation is consistent with the transmembrane domain predictions shown in Figure 4c, given the positions of the protein termini determined by the protease protection assays presented in Figure 4a-b.

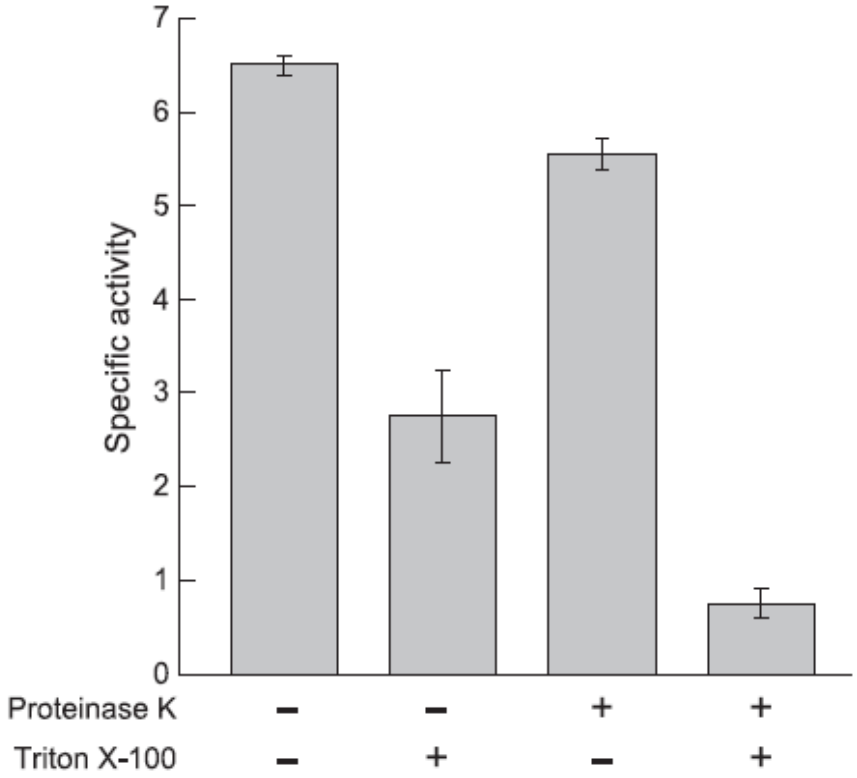

Figure 6. In vitro mannan synthase activity following protease protection.
Incorporation of mannose from GDP-14C-mannose into an ethanol-insoluble polysaccharide product by CSLA9. The x axis labels indicate the protease treatment conditions used for each sample. Proteinase K was inactivated before activity assays were performed. The y axis units are pmoles mannose per µg total microsomal protein per hour.

Topology of CSLC4 in the Golgi membrane

Constructs for heterologous expression of CSLC4 were created similarly to those made with CSLA9. Specifically, a T7 epitope tag was added to the amino terminus, the carboxyl terminus, or both ends (Figure 7) of the protein. Intact microsomal membrane vesicles were isolated from *Pichia* expressing each of these tagged CSLC4 proteins and exposed to protease, with or without detergent, in a manner similar to that described above for the CSLA9 constructs. Immunoblots following the protease protection experiments revealed that both the amino and carboxyl termini of CSLC4 were sensitive to protease treatment, even in the absence of detergent, and therefore we concluded that both termini face the cytosol (Figure 7a-c). In the transmembrane domain predictions shown in Figure 7d, the larger hydrophilic loop containing the active site of CSLC4 is separated from both termini of the protein by an even number of transmembrane domains. This predicts that the active site of CSLC4 protein is on the same side of the membrane as the termini, and thus likely to be cytosolic.

We have not identified the assay conditions required to measure glucan synthase activity of the recombinant CSLC4 protein *in vitro*. Therefore, an alternative approach was needed to probe the topology of the active loop of this protein. We introduced the four-amino-acid recognition sequence (IEGR) of the protease Factor Xa into the active-site loop of amino-terminal T7-tagged CSLC4 using site-directed mutagenesis. The relative positions of the inserted protease site, the protein's active site, and predicted TMDs are indicated in Figure 8a.

The predicted size of the amino terminal Factor Xa digestion product containing the T7 epitope is 44 kDa (Figure 8a). As shown in Figure 8b, peptides near this size were detected only upon Factor Xa digestion of the recombinant CSLC4 protein containing the corresponding protease recognition sequence. These cleavage products were generated even in the absence of detergent, under conditions that maintain intact vesicles based on FUT1 activity, as described

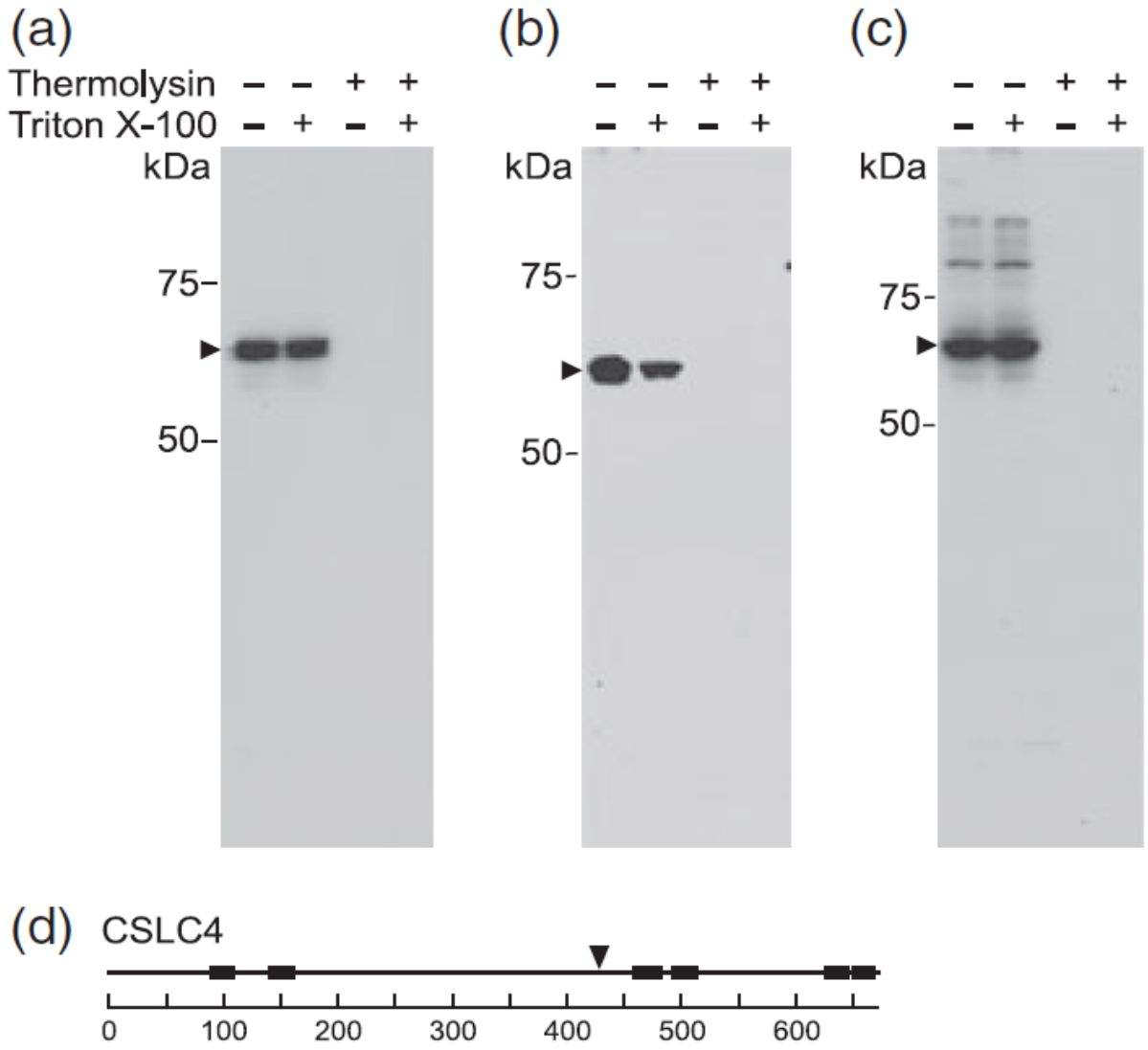

Figure 7. Protease protection of AtCSLC4 termini.
(a–c) CSLC4 proteins containing a T7 peptide epitope added to the N-terminus (a), the C-terminus (b) or both termini (c) were expressed in Pichia cells. Microsomal membranes were isolated and subjected to protease treatment. The resulting microsomal proteins were separated by SDS–PAGE, and the tagged proteins were detected by immunoblotting using antibodies against the T7 epitope. Labels above the blots indicate the protease protection assay Conditions. (d) CSLC4 protein with TMDs predicted by TMMOD represented by black bars. The position of the sugar-nucleotide binding motif is indicated by an arrow.

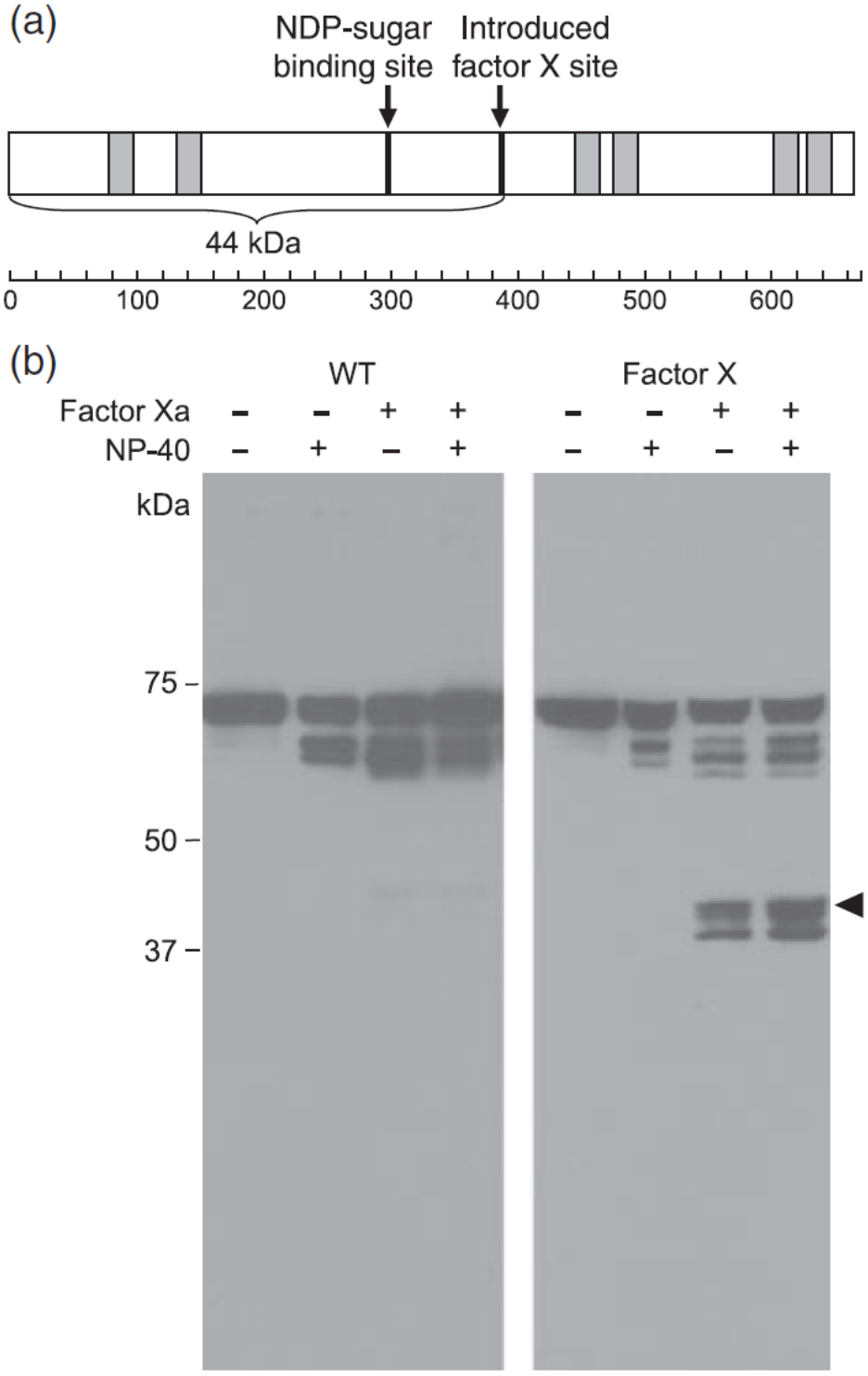

Figure 8. Specific protease cleavage of the active loop of CSLC4.
The recognition sequence for the site-specific protease factor Xa was added to the active loop of CSLC4 by site-directed mutagenesis. In the schematic representation of CSLC4 protein (a), gray bars represent TMDs, and the positions of the substrate-binding site and the introduced factor Xa site are indicated by arrows. A T7 tag on the N-terminus of CSLC4 was used for detection of full-length and cleaved CSLC4 protein. Cleavage of the mutated CSLC4 produces a 44 kDa band after treatment with factor Xa in the absence of detergent, indicating that the active loop of CSLC4 faces the cytosol (b).

above (see Figure 5). The additional bands that appear in both WT CSLC4 and the Factor Xa mutant with the addition of detergent and/or protease seem to be the products of cleavage by endogenous proteases and Factor Xa at a secondary site within CSLC4. From these results, we conclude that the protease cleavage site engineered within the active-site loop of CSLC4 faces the cytosol (Figure 3, model b). This is consistent with the conclusion reached from the protease protection experiments and TMD prediction shown in Figure 7d. Because it directly measures the location of the active-site loop, the result of the Factor Xa cleavage experiment is not dependent upon predictions of the number and location of the TMDs.

Discussion

For the protease protection experiments described in this Chapter, three epitope-tagged constructs made with CSLC4 protein are presented, but only two constructs (amino-terminus and both termini tagged) to CSLA9 are shown. In general, the expression level of CSLA9 in *Pichia* was low compared to CSLC4, and over 100 transformation events were screened to isolate a line expressing sufficient levels of T7-CSLA9 for downstream analysis. Several transformants expressing CSLA9-T7 (with the epitope on the carboxyl-terminus) were isolated, and low levels of CSLA9-T7 could be detected following a rapid microsomal fractionation used to evaluate transformants. However, following the microsomal membrane preparation method designed to yield intact vesicles, and subsequent treatment of these samples for protease protection assays, no signal from CSLA9-T7 was observed, even in mock-treated controls. It is my conclusion that low abundance of CSLA9-T7 protein expressed in *Pichia*, and instability of the protein due to endogenous proteases and other factors, resulted in undetectable levels following preparation of intact vesicles. It is not clear at this time why expression levels for CSLA9 constructs in *Pichia* were so much lower than for CSLC4 constructs. One likely explanation for this is that the active CSLA9 proteins disrupt *Pichia* metabolism through the production of mannans, or perhaps more

likely, by disrupting GDP-mannose or GDP-glucose homeostasis in the yeast.

In this study, we present evidence that two closely related Golgi-localized processive glycan synthases have opposite orientations in the Golgi membrane. Specifically, the topology of CSLC4, a putative xyloglucan glucan synthase, results in a cytosolic active site, whereas the mannan synthase CSLA9 has an active site in the Golgi lumen. The topology of these proteins is likely to impact many aspects of polymer synthesis, including substrate availability, the mechanism of chain initiation, elongation, and termination, as well as the organization of possible protein complexes. The observation that CSLC and CSLA proteins have distinct orientations raises many questions and carries important implications that are discussed below.

This conclusion predicts that the putative xyloglucan glucan synthase has a topology similar to the cellulose synthase, and may operate in a similar manner, whereas the mannan synthase has a different topology and may operate differently. Because of the similarities in their sequence, to compare and contrast the two CSL proteins studied here with the CESA proteins involved in cellulose biosynthesis could be interesting and informative. The CESA proteins that make up cellulose synthase are present in large complexes (Guerriero*, et al.* 2010), where they synthesize unsubstituted β-(1→4)-linked polymers of glucose and deposit them directly into the wall. Like CSLA and CSLC proteins, CESA proteins are integral membrane proteins with multiple transmembrane domains (eight are predicted by most algorithms). One important difference is that CESA proteins are active in the plasma membrane, although they have also been observed in internal membranes (Paredez *et al.* 2006).

The topology of CESA proteins has not been probed directly, but there is indirect evidence that the hydrophilic loop containing the substrate-binding site is in the cytosol. Two phosphorylation sites have been identified on CESA3 proteins: one site in the hydrophilic region containing the active site, and another between the amino-terminus and the first predicted

transmembrane domain (Benschop *et al.* 2007). Because protein phosphorylation takes place in the cytosol, this observation is evidence that these regions of CESA proteins are cytosolic. A second line of evidence that CESA proteins have a cytosolic active site is the availability of UDP-glucose, which is present in the cytosol, though the elongating β-glucan polymers appear in the apoplast (Guerriero*, et al.* 2010).

CSL proteins are similar in sequence and topological organization to CESA proteins, and they may also utilize substrate from one cellular compartment (cytosol) while depositing their products in another (Golgi lumen). The cytosolic location of the active site of CSLC proteins should allow them to utilize cytosolic pools of UDP-glucose. Thus, CSLC probably functions analogously to CESA proteins, coupling the addition of glucose monomers to the growing chain with translocation across the Golgi membrane of either an activated form of the sugar monomer or the elongating polymer product. In contrast, the active site of CSLA9 faces the Golgi lumen. It is therefore likely that its substrate, GDP-mannose, is drawn from a pool in the Golgi lumen during mannan synthesis. This model predicts that Golgi membranes should have a GDP-mannose transporter, which has been confirmed experimentally (Baldwin *et al.* 2001; Handford *et al.* 2004), and that GDP-mannose transporter activity would be important during mannan biosynthesis.

Prior to this study, several groups had examined the topology of polysaccharide biosynthesis in Golgi membrane fractions. Because the enzymes responsible for synthesis of hemicellulosic polysaccharides have only recently been identified, the approach taken in most of these studies was to characterize topology based on analysis of enzyme activity.

Muñoz *et al.* (1996) investigated the protease sensitivity of glucan synthase-I (GS-I) activity in intact vesicles isolated from pea (*Pisum sativum*). The authors found that GS-I activity is sensitive to protease treatment, which results in a loss of about 80% of GS-I activity

when non-permeabilized Golgi vesicles were protease treated. Because GS-I activity is reduced only about 50% when membranes are permeablized with detergent after protease treatment, the authors conclude that the active sites of enzyme(s) responsible for GS-I activity face the Golgi lumen. They propose that the protease-sensitive component responsible for the reduction in GS-I activity is a UDP-glucose transporter that provides substrate to the Golgi lumen. Although this is one possible interpretation of their data, inferring membrane protein topology based only on enzyme activity is difficult, especially when the measured activity requires the action of several proteins. Using these data to determining the membrane topology of the glucan synthase is especially difficult because the enzyme or enzymes responsible for GS-I activity have not been identified. Topological studies using protease protection assays are much easier to interpret when the sensitivity of an individual protein under investigation can be measured directly, as is the case for the CSLA9 and CSLC4 proteins reported here.

Another example where protease protection of enzymatic activity has been used to investigate membrane topology is the mixed-linkage (1→3),(1→4)-β-glucan synthase activity found in maize microsomal membrane preparations (Urbanowicz*, et al.* 2004). The authors utilized vesicles enriched in Golgi membranes and observed that protease treatment "selectively lowered the amount of cellotriosyl units produced in vitro without significant alteration in the amount of cellotetraosyl units." They interpreted this to indicate that "part of the catalytic domain is exposed to the cytoplasmic face of the Golgi membrane" (Urbanowicz*, et al.* 2004). Now that the CSLF and CSLH proteins are known to be involved in the synthesis of mixed-linkage glucan (Burton*, et al.* 2006; Doblin*, et al.* 2009), it will be possible to pursue approaches like those presented in this study to determine the membrane topology of proteins in these CSL families.

To our knowledge, the only other CSL protein topology that has been investigated

directly is that of CSLD2 (Zeng and Keegstra 2008). In that study, antibodies recognizing an amino-terminal region of CSLD2 were used to evaluate the position of the epitope relative to the Golgi membrane. Proteinase K treatment in the presence and absence of detergent indicated that the amino-terminal region of the protein is cytosolic. The transmembrane domain predictions presented by these authors led to the conclusion that the active site of CSLD2 is also located in the cytosol. Perhaps this result is not surprising, given that CSLD family proteins are so closely related to the CESA proteins and are both thought to contain 8 transmembrane domains.

The observation that the two closely related membrane proteins investigated here, CSLA9 and CSLC4, have evolved opposite membrane topologies raises many questions. For example, once a functional ancestral enzyme is established with a given topology, what structural changes would be necessary to produce a protein with an opposite topology? Once such changes have occurred, what selective advantage is conferred by the newly evolved membrane protein orientation, and if such an advantage exists, why is the ancestral orientation also maintained?

In the past ten years, our understanding of membrane protein topology has increased dramatically (reviewed in von Heijne (2006) Some proteins are known to have dual topology: they insert in the membrane in two opposite orientations. In some other cases, homologous proteins can have opposite orientations (von Heijne 2006). These situations provide precedence for how the CSLA and CSLC proteins could have evolved opposite topologies. Small changes to protein sequence, such as the addition or removal of positive residues in a hydrophilic region of the protein, or the loss, addition, or duplication of a single transmembrane domain, can cause significant changes to protein topology, including the conversion of a dual topology protein to one that favors a specific orientation (Rapp*, et al.* 2007).

The CSLA and CSLC protein families are conserved throughout higher plants, and TMD prediction programs for other members of these families not assayed here predict the same

number of TMDs in the same locations, with relatively few exceptions. It is interesting to speculate on the advantages that may be conferred by either orientation. For example, the cytosolic hydrophilic regions of CSLC4 might mediate interactions with cytosolic factors either for substrate channeling (e.g., sucrose synthase) to increase metabolic efficiency, or post-translational modification to regulate CSLC4 activity. If CSL proteins are product-inhibited, a cytosolic active site might result in faster catalysis by allowing the nucleotide product to diffuse away and separating the polymer product from the enzyme's active site. On the other hand, an active site in the Golgi lumen could confer properties to a glycan synthase that might contribute to organismal fitness in regard to appropriate synthesis of extracellular polysaccharides. CSLA proteins incorporate glucose as well as mannose into a β-(1→4)-linked polysaccharide *in vitro* (Liepman*, et al.* 2005). The lumenal active site of CSLA might serve to indirectly control glucomannan composition and/or rate of synthesis by controlling substrate identity and amounts imported to the Golgi from the cytosol. These and other hypotheses will need to be tested in future work.

As discussed above, one can now understand how some changes to protein sequence might result in an inverted topology. However, the sequence requirements for a glycosyltransferase to have membrane transporter activity are impossible to predict at this time. Although we expect that CSLA proteins require substrate to be transported to the lumen, and Golgi-localized GDP-mannose transporters have been identified (Handford*, et al.* 2004), other possibilities cannot be ruled out. It is possible that CSLA proteins may utilize cytosolic GDP-mannose directly, even though their active sites face the Golgi lumen. This scenario would require that CSLA proteins transport sugars across the membrane before rather than after catalysis. This would be a similar but distinct situation to what supposedly occurs with CESA and CSLC proteins, where the TMDs of these glucan synthases are thought to participate in the

formation of a membrane pore through which the glucan product is passed. In the case of a CSLA protein with a lumenal active site able to access cytosolic substrate pools, the membrane pore would allow part or all of the GDP-mannose substrate to pass from the cytosol to the enzyme's active site in the Golgi lumen. If transport activity occurs with CSL proteins regardless of orientation, then a newly duplicated paralog with a cytosolic active site would not necessarily need to acquire transporter functionality. Similarly, plants would not need to provide GDP-mannose to the Golgi lumen for mannan biosynthesis.

It is interesting that the mannan synthase activity of multiple CSLA homologs from several species has been measured *in vitro* (Liepman*, et al.* 2007a; Liepman*, et al.* 2005), although this has not yet been achieved with any CSLC proteins (Cocuron*, et al.* 2007; Liepman*, et al.* 2005). It is tempting to postulate that this phenomenon is related to their topologies in the Golgi membrane. For example, perhaps the six transmembrane domains of CSLC4 are insufficient to translocate the product across the membrane, a predicted requirement for a glycan synthase with a cytosolic active site. An unsubstituted β-(1→4)-linked glucan chain is between 0.5 and 1.0 nm in diameter (Kondo *et al.* 2001). Assuming standard transmembrane alpha helix dimensions, at least eight transmembrane helixes would be required to form an active pore. To form a water-filled pore large enough to accommodate both the glucan chain and water molecules, sixteen TMDs might be required to participate in the formation of a pore (Sandhu *et al.* 2009). Thus, glycan synthases with topologies similar to CSLC and CESA might require the interaction of multiple proteins to translocate the polymer across the membrane. *Pichia* cells heterologously co-expressing CSLC4 and XXT1 produce increased levels of insoluble glucan when compared to wild-type yeast (Cocuron*, et al.* 2007). These data are consistent with the idea that proteins other than CSLC4 participate in translocation of the elongating glucan chain.

Although we have some understanding of how polysaccharide elongation occurs, the

process of glycan chain initiation in mannan and xyloglucan biosynthesis remains unknown. One possibility is that polysaccharide synthesis initiates with the transfer of a sugar from a sugar nucleotide onto a "primer" molecule. Sitosterol-β-glucoside has been proposed to play a role in cellulose initiation (Peng *et al.* 2002). However, it is not clear what role this molecule might have in cell wall glycan initiation in the Golgi. Proteins such as CSLC4 and CESA, which must utilize substrate on one side of the membrane to form a product on the other side, might have specific requirements for a "primer," in that the growing chain needs to be long enough to span the membrane before "steady-state" elongation can occur. On the other hand,different constraints on primer properties might be imposed by proteins such as CSLA9 that utilize substrate on the same side of the membrane as their product. These and other important questions regarding the mechanism of polysaccharide biosynthesis can now be experimentally addressed in light of the topological constraints derived from the studies reported here.

Experimental procedures

Construction of expression plasmids and sequence analysis

T7 epitope tags (encoding amino acid sequence MASMTGGQQMG) were added to *CSL* cDNA sequences by PCR using PFU-turbo polymerase (Stratagene 600250) with the manufacturer's recommended conditions. Primers contained a DNA sequence encoding the T7 tag followed by a gene-specific region (see supplemental table for all primer sequences used). Sense primers also contained CACC for directional cloning and CATA, a yeast Kozak sequence (Kozak 1990) upstream of the T7 tag and/or gene-specific region.

T7-tagged constructs were subcloned in pENTR-D-TOPO (Invitrogen) and the resulting constructs were used to transform *Escherichia coli*. Plasmid DNA was isolated from positive transformants and sequenced to confirm the constructs before the T7-tagged ORFs were recombined using the Gateway (Invitrogen) LR clonase into the pPICZ vector for *Pichia*

expression. The attR Gateway recombination site had been previously introduced by restriction enzyme digestion and ligation at the *Sfu*I and *Eco*RI sites in the pPICZ vector.

Growth and maintenance of *Pichia pastoris* cultures

pPICZ constructs were used to transform *Pichia* (Strain X33) by electroporation with a MicroPulser™ electroporator (BioRad 165-2100) using the manufacturer's setting for *Pichia*. PCR screening confirmed the presence of transgenes. Briefly, 1-ml cultures were grown for 48 h in YPD + Zeocin (100 μg/ml) and then cells were harvested and spotted on Whatman FTA Classic cards. A 2.5-mM craft punch was used to excise a portion of the spotted cells for semiquantitative PCR analysis of genomic DNA using REDTaq PCR mix (Sigma D4309). Amplification from the transgene was compared to that of an endogenous *Pichia* gene, alcohol oxidase 2 (*AOX2*), in duplex PCR to identify transformants likely to contain multiple copies of the transgene.

For expression of CSLC4 and FUT1 protein constructs, cultures of *P. pastoris* strain X33 were grown in baffle flasks for 24 h at 28° with shaking at 200 rpm in buffered glycerol-complex medium (BMGY; 1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% Yeast Nitrogen Base [Gibco], 4 x 10^{-5} % biotin, 1% glycerol). Cells were then transferred to buffered methanol-complex medium (BMMY; 1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% Yeast Nitrogen Base [Gibco], 4 x 10^{-5} % biotin, 0.5-1% methanol) and grown for 72 h at 28° with shaking. Methanol was added to 1% every 24 h to maintain inductive conditions. For expression of CSLA9 protein constructs, X33 cells were inoculated directly to BMMY. Cells were grown in inductive conditions for 20 h at 28° in a shaking incubator, methanol was added to 1.5%, and the cultures were harvested 26 h after inoculation.

Isolation of intact microsomal vesicles from *Pichia pastoris*

Except where noted, all steps were conducted at 4° C. Cells were harvested from saturated *Pichia* cultures by centrifugation at 1,500 x g for 5 min, and washed twice with ice cold water. Spheroplasting for osmotic lysis to yield intact vesicles was done as in Graham *et al.* (1994), with the following modifications. Cells were resuspended in 20 ml SED buffer (0.95 M sorbitol, 25 mM EDTA, pH 8.0, and 50 mM DTT) per 70-ml initial culture volume and incubated at 30° with gentle shaking for 40 min. The cell suspension was then centrifuged under the same conditions, and the cells were resuspended in 20 ml SCE buffer (1 M sorbitol, 1 mM EDTA, 10 mM sodium citrate, pH 5.8). Zymolyase (MP Biomedicals) was added to a final concentration of 15 μg/ml. Cells were incubated again for 40 min with gentle agitation at 30° C. The resulting spheroplasts were harvested by centrifugation at 700 x g for 10 min and washed twice with 1 M sorbitol. The cells were then lysed osmotically by resuspension in a volume of lysis buffer (0.1 M sorbitol, 10 mM triethanolamine, pH 6.5, 1 mM EDTA, 1 EDTA-free protease inhibitor tablet [Roche] per 10 ml) of approximately 7x the volume of the *Pichia* cell pellet, followed by Dounce homogenization (25 strokes). The lysate was then centrifuged at 13,000 x g for 20 min to remove intact cells and dense organelles. The resulting supernatant was centrifuged at 120,000 x g to produce a pellet containing intact microsomes. The 120,000 x g pellet was resuspended in buffer containing 0.8 M Sorbitol, 20 mM HEPES-KCl, pH 7.0, and Roche protease inhibitor cocktail (1 tablet / 10 ml buffer). The resulting suspensions of intact microsomal vesicles contained between 10 and 20 μg microsomal protein per μl suspension.

Protease protection assays

Intact microsomal vesicles were isolated as above, diluted to 5-10 μg/μl, and treated with thermolysin (0.5-2.0 μg/ml) and/or detergent (0.1% NP-40 or Triton X-100) in 100-μl reactions for 30 min at room temperature. For thermolysin assays, $CaCl_2$ was added to the microsomal membrane resuspension buffer to a final concentration of 0.5 mM. Thermolysin digestions were

terminated by the addition of EGTA to 5 mM. Treated membranes were prepared for SDS-PAGE as in Liepman *et al.* (2005). Microsomal proteins were separated on 10% acrylamide gels before being transferred to PVDF membranes for western blotting. Monoclonal peroxidase-conjugated T7 antibody (Novagen) was used for chemiluminescent detection. TMMOD (Kahsay *et al.* 2005) was used to generate transmembrane domain predictions.

For mannan synthase and fucosyltransferase assays, intact microsomal vesicles were treated with proteinase K rather than thermolysin, because the EGTA used to stop thermolysin cleavage inhibited mannan synthase and fucosyltransferase activities. Proteinase K digestions were inactivated with PMSF (20 mM final concentration) to terminate activity before measurement. Fucosyltransferase assays were done as described (Perrin *et al.,* 1999) with some changes. Briefly, treated membrane proteins were incubated with 1.2 μM GDP-[^{14}C] -fucose, 0.2% tamarind xyloglucan, and 2 mM $MgCl_2$ in buffer containing 40 mM Pipes-KOH, pH 6.8, 0.4 mM DTT, 0.16 M sucrose, and 1% Triton (where appropriate). Mannan synthase assays were done exactly as described in Liepman *et al.* (2005). For all glycosyltransferase assays, 100-200 μg of pre-treated membrane proteins were incubated for 30 min at room temperature in a volume of 40 μl. Reactions were stopped by the addition of 1 ml 70% ethanol and polysaccharide products were precipitated overnight at -20°. Pellets were washed 3 times with 1 ml 70% ethanol containing 2 mM EDTA before being dissolved in scintillation fluid. ^{14}C incorporated to ethanol-insoluble products was detected with the 1450 MicroBeta TriLux liquid scintillation counter (LSC, Perkin Elmer).

Specific incorporation of mannose from GDP-mannose to an ethanol-insoluble product by wild-type *Pichia* microsomes was generally between 5 and 10 % of the amount incorporated by microsomes containing CSLA9 protein constructs. To confirm that the ethanol-insoluble product produced by CSLA9 protein was indeed a β-(1→4)-mannan, reaction products were

incubated with endo-β-mannanase from *Aspergillus niger* (Megazyme) as in Liepman *et al.* (2005). Undigested polysaccharide was then precipitated at -20° by the addition of 1 ml 70% ethanol. At least 70 % of the radiolabeled product from CSLA9 was susceptible to cleavage by endo-β-mannanase, whereas only 10 % of the label incorporated by wild-type *Pichia* microsomes was released with hydrolase treatment.

Site-specific protease cleavage of CSLC4

A factor Xa protease recognition site (IEGR) was added to the active loop of CSLC4 by mutating L_{384} to I, and adding a glycine residue following E_{385} using the Gene-Tailor® site-directed mutagenesis kit (Invitrogen). The mutant T7-tagged CSLC4 was expressed in *Pichia* and intact vesicles were obtained as above. Factor X cleavage assays were based on conditions reported in Gilstring and Ljungdahl (2000), and were performed on 0.5 to 1.0 mg total membrane protein in buffer containing 0.8 M sorbitol, 100 mM HEPES-KCl, pH 7.5, 50 mM NaCl, and 5 mM $CaCl_2$, with 0.2 μg factor Xa protease in the presence or absence of detergent in a volume of 100 μl. Reactions were terminated by the addition of PMSF to 10 mM.

CHAPTER 3: SUBCELLULAR LOCALIZATION OF HEMICELLULOSIC POLYSACCHARIDE BIOSYNTHETIC ENZYMES

Introduction to confocal fluorescence microscopy

It would be difficult to overstate the impact the isolation (Prendergast and Mann 1978) and cloning (Prasher *et al.* 1992) of green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* has had on the field of cell biology (Tsien 1998). The chromophore of GFP is formed by cyclization of the peptide backbone at a serine and a glycine residue located on either side of a tyrosine, which results in an imidazole ring with a 4-methylphenol (from tyrosine) at the 5 position. The carbon-carbon bond extending from the imidazole to the methyl carbon then undergoes spontaneous oxidation to form the active chromophore p-hydroxybenzylidene-imidazolidone (Heim *et al.* 1994). Since the chromophore of GFP is formed from standard amino acid constituents in the presence of molecular oxygen, fluorescent protein fusions can be created with straightforward molecular cloning techniques. Aside from some special circumstances, such as in strict obligate anaerobes, researchers can proceed rapidly from gene identification to the visualization of cellular processes in a variety of species in real time. The impact of GFP as an example of biological fluorescence as well as an applied biotechnology was recognized with the Nobel prize in chemistry in 2008.

The development of fluorescence confocal microscopy applications for biological research is a striking example of synergy in the advancement of two disparate technologies. The invention of confocal microscopy, reviewed by Amos and White (2003), permits imaging of individual cells by limiting the signal that reaches a detector to an "optical slice." It is this ability to exclude light from outside the focal plane by the confocal aperture that makes confocal microscopy superior to conventional microscopy approaches, rather than the relatively modest improvement in resolution. However, most cellular material does not have enough diffraction

contrast to be visualized with confocal microscopy alone. Thus, confocal microscopy was not very useful for studying biological systems in the absence of photoactive labeling.

Fluorescence microscopy, whether with FP fusions or immunofluorescence techniques, is technically possible with conventional light microscopes. However, since most multicellular tissues are layered, it can be difficult to resolve fine structure within individual cells with conventional microscopy because fluorescence emitted from regions outside the focal plane is transmitted to the detector. Confocal microscopy techniques overcome this problem, allowing researchers to image regions of individual cells in living tissues. Since the discovery of GFP, various spectrum-shifted variants (reviewed by Müller-Taubenberger and Anderson (2007)) have expanded the capability of fluorescent proteins as biological probes. Fluorescent probes with distinguishable spectra allow researchers to examine colocalization of multiple fluorescence-tagged proteins.

As discussed in chapter 1, immunolocalizion of xyloglucan polymer epitopes by electron microscopy indicates that these polysaccharides appear in the Golgi apparatus (Moore and Staehelin 1988; Zhang and Staehelin 1992). The enzymes responsible for the biosynthesis of these polymers have recently been identified, allowing visualization of the localization of these proteins as fluorescent protein fusions in living plant cells. In this chapter, I discuss the use of confocal fluorescence microscopy in the localization of enzymes that synthesize mannan and catalyze steps in the biosynthesis of xyloglucan.

Results

CSLA9 and CSLC4 localize to the Golgi

My first objective was to confirm that CSLA and CSLC proteins are located in the Golgi in living plant cells, as indicated by preliminary studies (Cocuron*, et al.* 2007; Sandhu*, et al.* 2009). For live cell confocal microscopy *Nicotiana tobacum* leaves were infiltrated with a

mixture of three *Agrobacterium* cultures: one culture harbored a plasmid containing a CSL protein fused to cerulean CFP, one culture contained a construct encoding a fusion of the ER/Golgi marker protein ERD2 with YFP (ERD2-YFP) (Brandizzi*, et al.* 2002), and one culture contained a construct for expression of the p19 suppressor of gene-silencing protein from tomato bushy stunt virus (Voinnet *et al.* 2003). Co-infiltration with p19 was necessary for efficient expression of the CSL fluorescent protein fusions.

CFP-CSLC4 was found in Golgi stacks, as shown by its colocalization with ERD2-YFP (Figure 9a-c). CFP-CSLC4 was also observed in other punctate structures, some of which are associated with the Golgi and some that appear to be independent from the ERD2-containing Golgi stacks (Figure 9c and inset, arrowhead). These structures may indicate localization outside the Golgi proper, possibly in the trans-Golgi network or secretory structures. CFP tagged *Arabidopsis* CSLA9 (CSLA9-CFP) also colocalizes at the Golgi with ERD2-YFP (Figure 9d-f). Similarly to CFP-CSLC4, CSLA9-CFP is present in structures that are not labeled with the Golgi marker (Figure 9f and inset, arrowhead). These data show that the CFP-tagged CSL proteins have similar subcellular distribution with an clear localization at the Golgi apparatus and at small punctate structures of unknown nature (Figure 9b-c and e-f).

The peri-Golgi bodies labeled with CFP-tagged CSLC4 and CSLA9 show dynamic association with Golgi bodies labeled with ERD2. Figure 10 shows individual frames from a time-lapse movie of the cortex of a tobacco leaf epidermis cell expressing CFP-CSLC4 and ERD2-YFP. A Golgi body labeled with both CFP-CSLC4 and ERD2-YFP (arrowhead) can be seen tracking from the bottom right to the top left of the frame. For the first 5 frames, a structure labeled with only CFP-CSLC4 (*) moves with this Golgi stack, then the two structures disassociate from each other (frames 6-8). This type of behavior is characteristic of the TGN (Viotti*, et al.* 2010), but the identity of the compartments labeled with CFP-CSLC4 should be

confirmed by colocalization with a well-characterized marker.

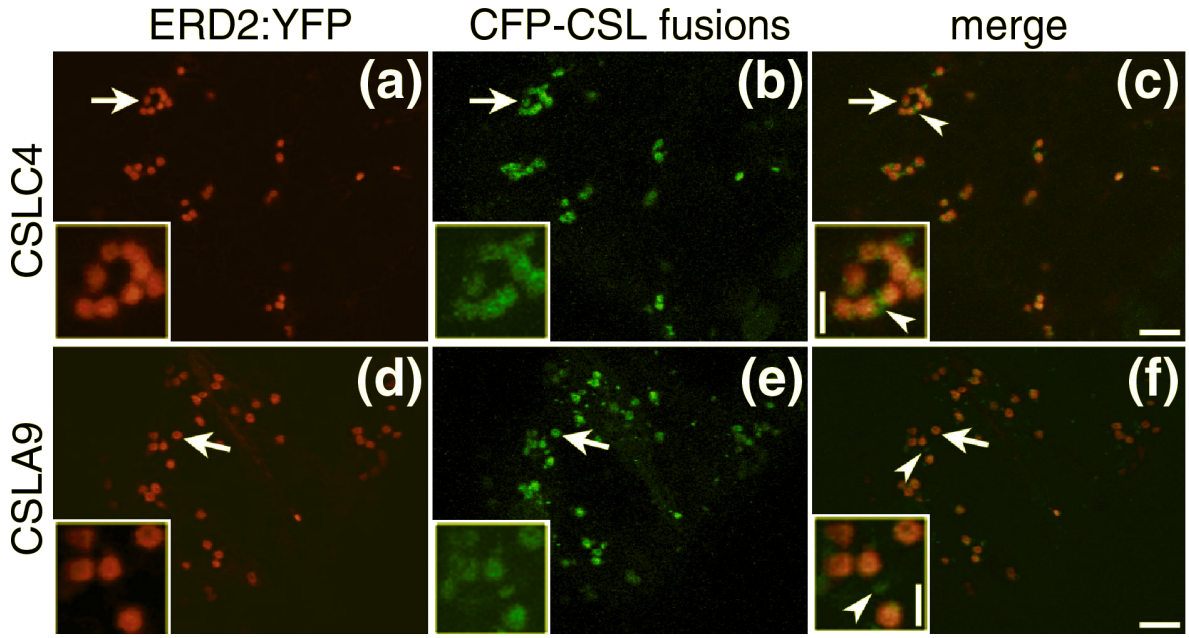

Figure 9. Subcellular localization of CSL proteins.
Laser scanning confocal microscopy of tobacco leaf epidermal cells expressing CFP-tagged CSLC4 (b, c) and CSLA9 (e, f) with the ER/Golgi marker ERD2–YFP (a, c, d, f) shows Golgi localization of CSL proteins. White arrows indicate areas of interest that are magnified in the insets. Arrowheads (c, f) indicate structures containing CSL protein but not ERD2. Scale bars = 5 μm (main panels) and 2 μm (insets).

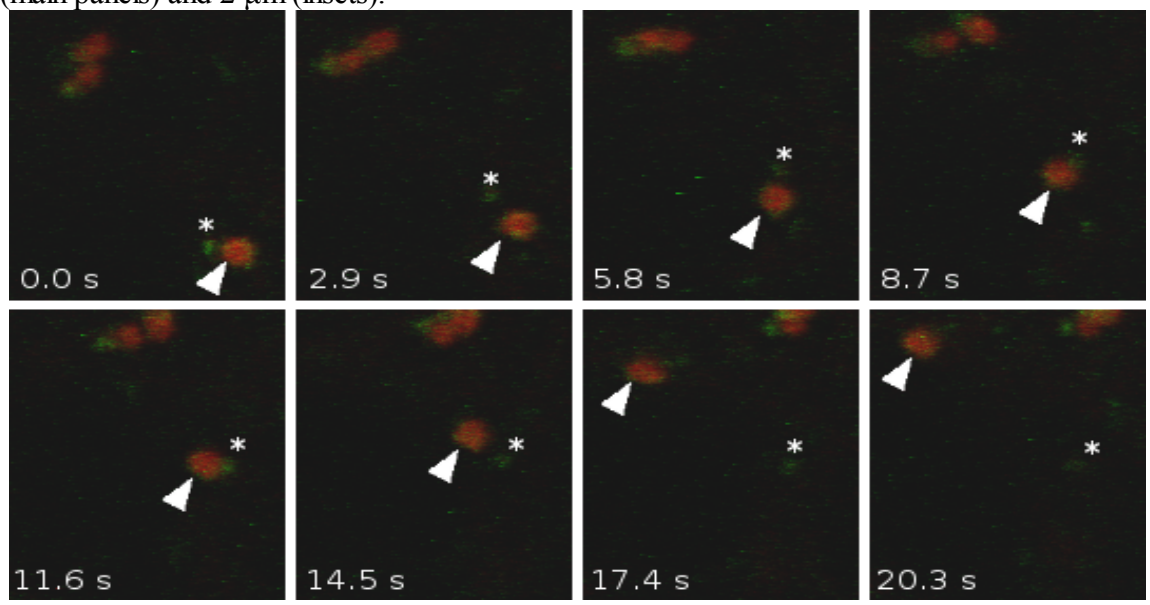

Figure 10. Structures labeled by CSLC4 show dynamic association with the Golgi apparatus.
A series of time-lapse, merged images of CFP-CSLC4 (green) and the Golgi marker ERD2-YFP (red). A small body labeled with CFP-CSLC4 (*) first traffics with a Golgi stack (arrowhead), and then the two structures disassociate from each other during the final three time points of the series. At the top of each frame, two other Golgi stacks can be seen maintaining association with CFP-CSLC4-labeled bodies during the time course. The time, in seconds, when each frame was imaged is given.

Colocalization of xyloglucan biosynthetic enzymes

As described in the introduction chapter, the hemicellulosic polysaccharide xyloglucan is synthesized by a combination of enzymatic activities. At least one enzyme is required for each type of glycosidic bond present in the mature polymer. Fucosylation of xyloglucan must occur after the addition of xylose and galactose at the appropriate residue (see xyloglucan structure, Figure 1). Since fucosylated xyloglucan can be detected in Golgi vesicles, it is logical that the enzymes responsible for fucosylation and upstream biosynthetic steps are also found in the Golgi apparatus.

Fluorescent protein fusions to the *Arabidopsis* xyloglucan fucosyltransferase (FUT1), the xyloglucan xylosyltransferase (XXT1), and the xyloglucan glucan synthase CSLC4 were created and infiltrated to *Nicotiana tobacum* leaves in various combinations. Cells in the abaxial epidermal layer were imaged 48 hours following infiltration. For analysis of the subcellular localization of xyloglucan biosynthesis, xyloglucan biosynthetic enzymes fused to cerulean CFP or venus YFP were co-infiltrated with a *trans*-Golgi marker (sialyltransferase transmembrane domain, ST) fused to monomeric red fluorescent protein (RFP). However, CFP-CSLC4 and YFP-XXT1 colocalize more completely with each other than they do with RFP-ST (ARRROWHEAD). If the smaller extra-Golgi structures labeled by CFP-CSLC4 actually correspond to the TGN, then the pattern of CFP-CSLC4/YFP-XXT1/RFP-ST localization indicates that CFP-CSLC4 and YFP-XXT1 probably localize at the *cis* or medial Golgi cisternae.

In contrast to CFP-CSLC4, CFP-FUT1 and YFP-XXT1 only localize to compartments labeled by Golgi markers. In fact, based on the data presented in Figure 11, CFP-FUT1 and YFP-XXT1 show complete localization in Golgi stacks. Taken together, these data can be interpreted to mean that three steps in xyloglucan biosynthesis (glucan synthesis, xylosylation, and fucosylation) occur in the *cis* and medial Golgi. The fact that the glucan synthase CSLC4 is

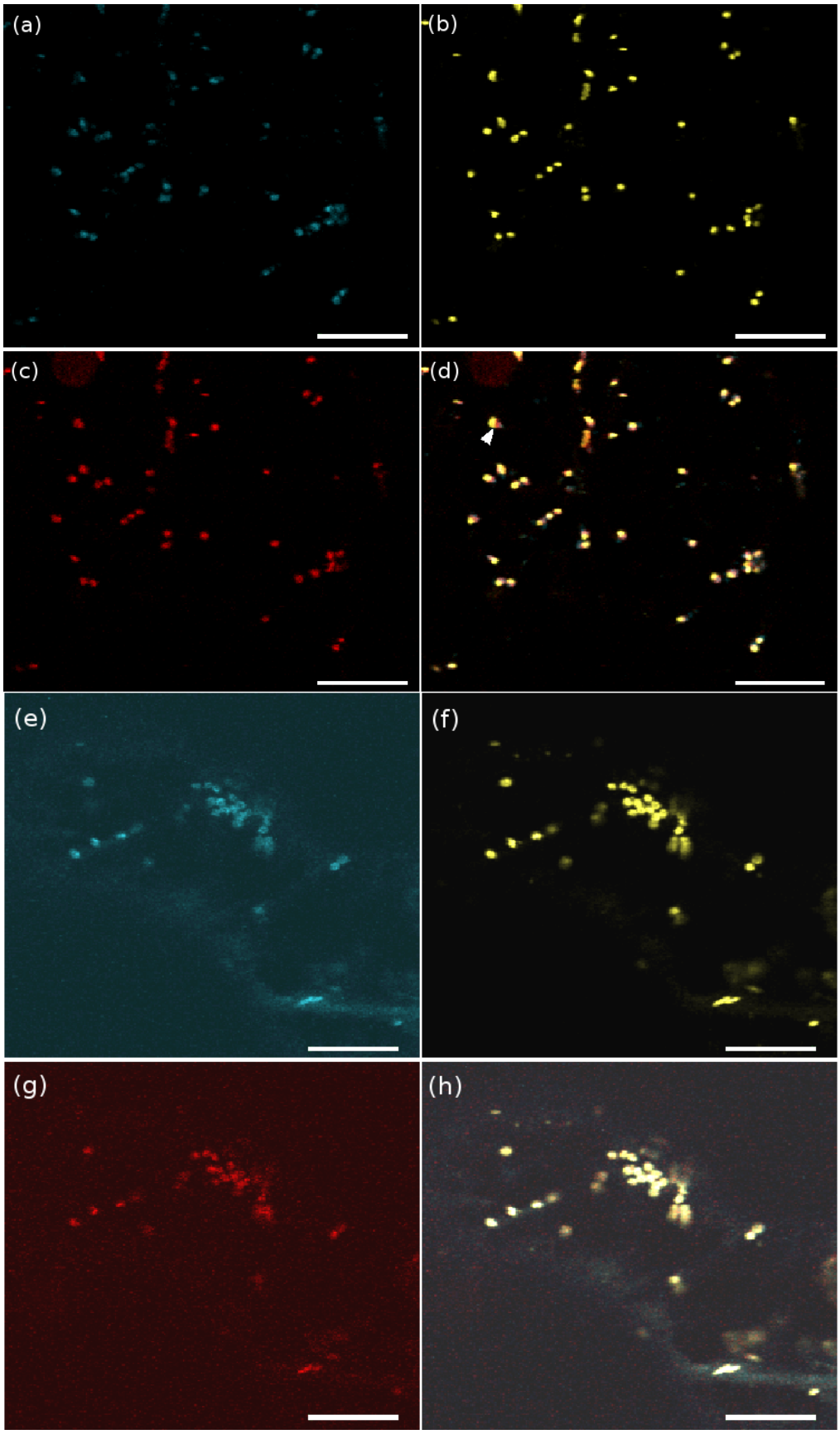
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Figure 11. Colocalization of enzymes catalyzing steps in the biosynthesis of xyloglucan. Images of a tobacco leaf epidermis cell expressing YFP-XXT1 (b and f) and ST-RFP (c and g) with CFP-CSLC4 (a) or CFP-FUT1 (e). Merge images are shown in (d and h). All fusion proteins localize to the Golgi, with CFP-CSLC4 also present in smaller structures associated with the Golgi. The arrow head in panel (d) indicates a Golgi stack oriented such that the relative distributions of CFP-CSLC4, YFP-XXT1, and ST-RFP can be seen. Scale bars = 10 μm

also found in structures that may correspond to the TGN could indicate that additional glucan synthase activity may occur in the absence of XXT1 or FUT1, downstream in the secretory pathway. This could mean that CSLC4 continues to add glucose residues to xyloglucan polymers to produce stretches of unsubstituted backbone, or CSLC4 in the TGN might produce glucan polymers (i.e. cellulose) that remain unsubstituted. However, there is no direct evidence for such glucan synthase activity from CSLC4 protein, so CSLC4 located in these extra-Golgi regions may not be active. In that case, CFP-CSLC4 signal from outside the Golgi may represent the native localization of a population of CSLC4 protein that has not been completely retained in the region of the Golgi where xyloglucan biosynthesis is most likely to occur. If this is true, the apparent TGN localization of CFP-CSLC4 might represent CSLC4 protein being recycled or turned over. Another possibility is that overexpression of CFP-CSLC4 causes "overloading" of the secretory pathway, and labeling of the small extra-Golgi structures represents a non-native localization as the plant cell attempts to maintain homeostasis in the secretory pathway.

Discussion

Confocal microscopy of fluorescence-tagged proteins affords a unique opportunity to visualize the behavior of proteins involved in various cellular processes as they occur in real time. Despite the power and elegance of this type of approach, one must always be careful interpreting the results. Perhaps the easiest mistake to make is to assume that enzymes are

always active where they are most abundant. For example, CESA proteins have been observed in vesicles in the cell cortex (Wightman and Turner 2010), but these proteins are known to be active at the plasma membrane.

Fluorescence-tagged CSLC4 and CSLA9 localize to compartments outside the Golgi, but we cannot assume that these proteins are active in these structures in the absence of additional evidence. In fact the absence of signal from XXT1 in these compartments may indicate that CSLC4 protein is not active there. As demonstrated in the study by Cocuron *et al.* (2007) that identified CSLC4 as a xyloglucan glucan synthase, CSLC4 expressed in *Pichia pastoris* produces short, soluble oligo glucans. Only when CSLC4 was coexpressed with XXT1 was the glucan synthase able to produce long glucan chains (DP>7). Interestingly, *Pichia* do not produce substrate for XXT1, which suggests that the presence of XXT1 *protein* is required for glucan synthase activity, possibly because it physically interacts with CSLC4. Given the sequence similarity among xyloglucan xylosyltransferase proteins, it would be interesting to see if any of the XXT1 homologs in *Arabidopsis* can substitute for XXT1 to increase the length of glucan chains produced by CSLC4 in *Pichia*. If another XXT has this effect, and it colocalizes to the same extra-Golgi compartments as CSLC4, then it might be more reasonable to assume xyloglucan biosynthesis occurs there.

In the course of investigating the subcellular localization of CSLA9 and CSLC4, we found that 35S-driven expression of these proteins in *N. tobaccum* leaves infiltrated with *Agrobacterium* did not result in the accumulation of detectable levels of protein. For localization studies of these CSL proteins, it was necessary to co-express the p19 suppressor of gene-silencing protein from the tomato bushy stunt virus (Tombus). RNA-mediated gene-silencing is a mechanism by which eukaryotes inactivate genes in a sequence-specific manner. In addition to regulating endogenous genes by directing chromatin modification and mRNA degradation, gene-

silencing in plants and other eukaryotes plays a role in suppression of virus replication (Meister and Tuschl 2004; Voinnet 2001). In an example of a host-pathogen evolutionary arms race, many viruses have developed mechanisms to suppress gene silencing (Lakatos *et al.* 2006). Voinnet *et al.* (2003) discovered that the expression of transgenes in *Nicotiana* leaves infiltrated with *Agrobacterium* is reduced by an endogenous gene-silencing mechanism that can be overcome by co-expression of the p19 protein.

The Tombus p19 protein bound to an siRNA target has been crystallized (Ye *et al.* 2003). The protein functions as a homodimer which specifically binds dsRNA molecules in the size range of 19-21 nucleotides. Two tryptophan residues on each of the p19 monomer units are positioned so that they "stack" with the nucleotide bases on either side of the dsRNA molecule. It is thought that these residues function in the selection of dsRNA molecules of a specific size, those that are involved in RNA-mediated gene-silencing (Vargason *et al.* 2003). P19 binding to siRNA prevents them from entering the RNA-mediated gene-silencing pathway and directing cleavage of complementary mRNA molecules.

It is not clear why the CSL proteins discussed here do not express to detectable levels in the absence of p19 protein. In fact, the abundance of CFP-tagged CSLA9 and CSLC4 proteins remains very low even when coexpressed with p19, when compared to fluorescence-tagged XXT1 or FUT1. It is tempting to speculate that an endogenous RNA-mediated silencing mechanism is involved in suppressing the 35S-driven expression of CSL protein in the absense of p19. Interestingly, a natural siRNA derived from the gene encoding CESA6 in barley (*Hordeum vulgare*) has been implicated in the regulation of CSL gene expression in that species (Held *et al.* 2008). An alternative explanation for the requirement of p19 in the transient expression of CSL protein is that the abundance of CSL transcript with 35S driven expression induces silencing without the involvement of endogenous siRNA-mediated signaling. Consistent

with either of these explanations is the observation that after 72 hours, leaves infiltrated with *Agrobacterium* harboring expression constructs for p19 and CSL protein had large regions that had undergone cell death. Specifically, leaf tissue that contained cells expressing CSL protein after 48 hours showed complete tissue death after 72 hours, while the rest of the leaf appeared normal. The cause of this cell death is not clear, but it seems to be a specific effect of CSL protein expression, as cell death was not observed with expression of the other glycosyltransferases described here.

In addition to confirming the presence of mannan and xyloglucan biosynthetic enzymes in the Golgi apparatus, fluorescent protein localization of these proteins yielded some interesting results. First, colocalization of fluoresence-tagged CSLC4 with XXT1 strengthens the hypothesis that these proteins may interact, as suggested by earlier studies (Cocuron*, et al.* 2007). Additional preliminary evidence of CSLC4, XXT1 interaction is presented in an appendix to this document. Localization of CSL protein outside of the Golgi proper was a surprising result, and may have relevance for other biological processes besides cell wall polysaccharide biosynthesis, such as RNA-mediated silencing, or protein recycling in the secretory pathway. Additional research into the regulation and trafficking of cell wall biosynthetic enzymes may help illuminate the significance of these preliminary observations in the future.

Experimental procedures

For transient expression of fluorescent proteins in tobacco, *Agrobacterium tumefaciens* strain GV3101 was transformed with plasmid pVKH18En6 (Batoko *et al.* 2000) containing CSL protein-coding sequence fused in frame to cerulean CFP (cCFP). Expression from pVKH18En6 is driven by a 35S promoter. Although some CSLD proteins with fluorescent protein fusions at their carboxyl terminus localize to the ER, amino-terminal fluorescent protein fusions localize to Golgi bodies (Bernal*, et al.* 2007; Li*, et al.* 2009). Localization to Golgi bodies is thought to

better reflect the *in vivo* localization of native CSLD proteins because independent methods have detected CSLD homologs in the Golgi (Dunkley *et al.* 2006). Also, CSL family proteins are thought to be involved in the synthesis of cell wall polysaccharides, which occurs at the plasma membrane or in the Golgi apparatus (Guerriero*, et al.* 2010; Scheller and Ulvskov 2010), but probably not in the ER. Because the CSL proteins analyzed in this study are similar in sequence to CSLD proteins, we first attempted to express them as amino-terminal CFP fusions for subcellular localization.

We detected fluorescence from CFP-CSLC4 in Golgi stacks, but no signal from CFP-CSLA9. Therefore, we fused CFP to the carboxyl-terminus of CSLA9 (CSLA9-CFP) for our localization studies. Cultures of *A. tumefaciens* (strain GV3101 for fluorescent constructs and strain C58C1 for p19 protein) were infiltrated into the abaxial leaf epidermis of four-week-old tobacco (*Nicotiana tabacum*) plants grown at 25°C (Batoko*, et al.* 2000) at a uniform bacterial optical density (OD_{600}=0.05). I used an inverted laser scanning confocal microscope (ZEISS LSM510 META) with a 63x oil immersion objective. For two-color imaging of CFP and YFP, I used the multi-track capability, focusing on coexpressed fluorescence-tagged proteins with argon laser excitation lines 458 nm for cCFP, 514 nm for YFP, and 594 nm for RFP. I detected fluorescence using a 458/514-nm dichroic beam splitter, a 465–510 nm bandpass filter (CFP), and a 520–555 nm bandpass filter (YFP). For 3 color colocalization of CFP, YFP, and RFP, I used a 405/514/594-nm dichroic beam splitter with bandpass filters at 465-510 nm (CFP), 530-600 nm (YFP), and 604-657 nm (RFP), also multitrack. 495 nm and 515 nm dichroic mirrors were employed to prevent signal from YFP being detected on the other channels. A 1 Airy unit pinhole was used for all channels.

CHAPTER 4: Impact, discussion, and future directions

Plant cell wall material constitutes a huge portion of the earth's biosphere, and has always been essential for human survival. The ability of cell walls to determine plant form and function and the usefulness of cell wall material to humans for tools, food, fuel, clothing, and shelter are due in large part to their polysaccharide components. Despite the obvious importance of plant cell wall polysaccharides, we are only beginning to understand how these polymers are made.

CSLA9 and CSLC4 proteins localize to the Golgi (Chapter 3) and produce polymers of β-(1→4)-linked mannose and glucose, respectively. Despite the fact that these proteins are similar in sequence and catalytic activities, they have opposite orientations in the Golgi membrane (Chapter 2). The active site of CSLA9 faces the Golgi lumen and the active site of CSLC4 faces the cytosol, though the products of both enzymes are deposited in the lumen. Some of the possible determinants and implications of CSL membrane protein topology are discussed at length in Chapter 2 of this document. However, many of the implications described there deserve further investigation in future studies. Some possibilities for future studies are described in this Chapter and in Appendix A.

I determined the orientations of the active sites of CSLC4 and CSLA9 by expressing epitope-tagged versions of these proteins in *Pichia pastoris* and using protease protection methods. *Pichia* was chosen as a heterologous system because this species has a developed eukaryotic endomembrane system (Rossanese *et al.* 1999), and transgenic lines are easily generated and maintained. *Pichia* has been used to heterologously express a number of plant cell wall glycosyltransferases (Cocuron*, et al.* 2007; Edwards*, et al.* 1999; Faik *et al.* 2002; Madson*, et al.* 2003), often in active forms. Because *Pichia* is a non-native system for the expression of plant proteins, it remains a formal possibility that CSLA9 or CSLC4 may adopt a different topology *in planta* than reported here. However, for the various reasons described in the

following paragraph, it is likely that the membrane topology of these proteins in plants is the same in *Pichia.*

CSLA9 protein produced in *Pichia* is active *in vitro* and *in vivo* (Chapter 2, Liepman *et al*. 2007b), and CSLC4 (Cocuron*, et al.* 2007) has *in vivo* glucan synthase activity when expressed in *Pichia*. This indicates that these proteins are folded correctly, and are thus likely to be oriented correctly. Although the fact that these proteins are active in *Pichia* does not prove that the topology in *Pichia* is identical to that found in plants, the similarities between these systems (described below) support the idea that these plant proteins have the proper orientation when expressed in *Pichia.* *Pichia* synthesizes peptides from the same amino acids as plants, and the Sec61 translocon apparatus is similar in sequence, meaning that the proteinaceous topological determinants, i.e. the transmembrane protein sequences and the structure of the translocon complex (discussed in Chapter 1) in both systems are similar. Although it is likely that the ER membrane lipid composition of yeasts is not identical to plants, one would expect to find similar properties regarding the lipid contributions to topology determination among these systems. Trends such as the positive inside rule are generally conserved among eukaryotes and bacteria (von Heijne 2006). Because the mechanism driving the positive inside rule as thought to involve the net charge contributed by lipids at the membrane-solvent interface (Bogdanov *et al.* 2009), it is likely that from the perspective of membrane protein orientation, ER lipid composition among eukaryotes may be functionally similar. In support of the fact that yeast can be useful in probing the topology of plant proteins, plant protein topology has been determined in the yeast *Saccharomyces cerevisiae* (Paul *et al.* 2007), which compared to *Pichia* has an endomembrane system more morphologically distinct from plants (Rossanese*, et al.* 1999).

Because they have opposite topologies but similar sequences, CSLC4 and CSLA9 proteins could be valuable tools to improve our understanding of how membrane topology is

determined. For example, these proteins differ in that CSLC4 probably has two transmembrane domains between the amino terminus and the substrate-binding site whereas CSLA9 has one. If the number of TMDs between the amino terminus and the active site is the primary determinant of the orientation of these proteins, then the addition of the "missing" TMD from CSLC4 to CSLA9 in the appropriate location between the first TMD and the catalytic domain of CSLA9 would cause the protein to adopt the cytosolic active site orientation observed with CSLC4. Alternatively, the factors that determine the lumenal orientation of the substrate-binding region of CSLA9 protein, such as the net charge contributed by the hydrophilic regions of this protein, might be strong enough to maintain the native orientation at the active site, resulting in an inversion of the orientation of the amino-terminal region of CSLA9. Such a result could lead to the identification and characterization of the specific factors that control membrane orientation of CSL and other proteins. A complementary experiment would be the removal of one TMD from the amino-terminal pair found in CSLC4 to investigate whether the orientation of this protein's active site is determined primarily by the number of TMDs following a amino terminus with a fixed localization in the cytosol, or by a more complex set of factors.

The observation that CSLC and CSLA proteins have opposite topologies in the Golgi membrane also raises questions regarding the topological orientations of other proteins in the CSL superfamily. As discussed in Chapter 2, CSLD2 is the only other CSL superfamily protein where the membrane topology has been probed directly (Zeng and Keegstra 2008). CSLD proteins have been implicated in the synthesis of non-crystalline cellulose (Manfield *et al.* 2004), and at least some of these proteins seem to play a role in tip growth (Bernal*, et al.* 2008), but the catalytic activity of proteins in the CSLD clade has not been unambiguously determined. If CSLD proteins are truly glucan synthases, and the topology predicted for CSLD2 by the Zeng and Keegstra (2008) study holds for other CSLD proteins, then an interesting pattern of topology

and catalytic activity begins to emerge. Specifically, all of the cell wall glucan synthases identified so far (CESA, CSLC, and possibly CSLD proteins) have their catalytic site on the opposite side of the membrane from where their products accumulate.

Despite differences in the subcellular localization of these putative glucan synthases and whether their products are unsubstituted crystalline or amorphous cellulose or the highly branched polymer xyloglucan, their similar topologies probably result in the conservation of some mechanistic similarities. For example, it is likely that transfer of their glucan product across the membrane would be required for all of these proteins, suggesting the presence of a transmembrane pore. The number of transmembrane helixes necessary to form a pore large enough to accommodate a hydrated β-(1→4)-linked glucan is discussed briefly in chapter 2. CSLD proteins and CESA proteins are predicted to have 8 transmembrane domains, the minimum number necessary to accommodate β-(1→4)-linked glucan (Kondo *et al.* 2001), though it is likely that the pore would need to be much larger to fit a hydrated glucan, possibly composed of TMDs from a dimer of CESA or CSLD proteins. CSLC4 has fewer TMDs, probably 6, so these proteins might require the participation of TMDs from other proteins to form a pore, whereas CESA and CSLD may not. If CESA, CSLD, or CSLC proteins function as dimers, a dimer of CLSC proteins would still have fewer TMDs and might still require the participation of an interacting protein. Good candidates for such an interacting protein are the XXT homologs from CAZy family 34. Further evidence for a physical interaction between CSLC4 and XXT1 is described below and in Appendix A.

Xyloglucan is a particularly interesting polysaccharide because of its complexity, regular structure, and abundance. The near-ubiquity of this polymer among higher plants and the high degree of structural regularity suggest that plants may have developed specific mechanisms to ensure this regular structure. Data presented in Chapter 3 show colocalization of an *Arabidopsis*

xyloglucan xylosyltransferase (XXT1) with the xyloglucan glucan synthase CSLC4. However, colocalization in the same compartment may not be sufficient for these enzymes to work in concert to produce the regular repeating pattern of xylose substitution of the glucan backbone that is characteristic of *Arabidopsis* xyloglucan. Evidence presented in the study by Cavalier and Keegstra (2006) supports the idea that the mere presence of XXT enzyme activity in a compartment containing acceptor substrate produced by CSLC4 would not result in the typical xyloglucan substitution pattern. In their study, Cavalier and Keegstra (2006) measured the *in vitro* activity of solubilized XXT1 and its close homolog XXT2 using cellohexaose as an acceptor substrate. The substitution pattern of xylose residues following incubation did not reflect the native substitution pattern shown in figure 1. Rather, each of these enzymes prefer to add xylose to the fourth glucose residue from the reducing end of the acceptor. Both XXT1 and XXT2 subsequently transfer a second xylose residue the next glucosyl residue in the direction of the reducing end, but at much lower efficiency. With longer incubation times and sufficient concentrations of UDP-xylose, each of these enzymes will add a third xylose residue to form a product with three adjacent glucosyl residues substituted by xylose, but this is not the preferred product formed by either enzyme.

Two important differences between assays described by Cavalier and Keegstra (2006) and xyloglucan biosynthesis in planta are the absence of CSLC4 protein and solubilization of the membranes containing XXT protein in the in vitro assays. In the study by Cocuron et al. (2007) that uncovered the catalytic activity of CSLC4, glucan synthase activity from CSLC4 was altered by co-expression of XXT1, in the absence of UDP-xylose, the substrate for XXT1. This indicates that XXT1 and CSLC4 might directly interact in a way that is important for CSLC4 function. It is logical that XXT enzyme function could also be influenced by interaction with CSLC4, such that the typical substitution pattern found in Arabidopsis xyloglucan is achieved

only when these proteins are in a complex. Affinity chromatography using antibodies to a T7 epitope fused to CSLC4 (Appendix A) indicates that XXT1 and CSLC4 proteins expressed in *Pichia* interact. Future studies using techniques such as Förster Ressonance Energy Transfer (FRET) will be able to confirm whether XXT1 and CSLC4 interact in plant cells. Hopefully, the methods I have developed for the isolation of CSLC4 in complex with XXT1 in *Pichia* can be applied to plant cell cultures to identify additional members of the putative xyloglucan biosynthetic complex. The solubilization conditions for CSLC4 – XXT1 complexes, and a method for the purification of these interacting proteins from *Pichia* are described in Appendix A. Future studies may be able to apply similar conditions with plant cells to identify other components of a putative xyloglucan synthase complex. With the structural similarity among CSL proteins, these methods could also be adapted for other members of the CSL superfamily.

Protein complexes are involved in the biosynthesis of cell wall polysaccharides other than xyloglucan. CESA proteins are active as part of large biosynthetic complexes in the plasma membrane. Several mutants that have perturbed cellulose synthesis have been isolated. Some of these mutants, for example *irregular xylem* mutants *irx3-1* and *irx5-2*, produce truncated CESA proteins that are impaired in their ability form complexes (Atanassov *et al.* 2009). Because the catalytic domains of the CESA isoforms disrupted in *irx3-1* and *irx5-2* plants remain intact, the IRX phenotypes in these lines could be indicative of the importance of proper complex formation in cellulose biosynthesis. The interaction between XXT1 and CSLC4 proteins could contribute to the structure of xyloglucan by helping to place the active sites of the xyloglucan biosynthetic enzymes in the appropriate location. If this is true, then interaction among CESA proteins may occur for similar reasons. Although each CESA catalytic unit within the complex probably produces its own glucan chain, the proper positioning in the complex of each CESA protein could be important for the formation of microfibril structures with the appropriate properties.

The activities of some proteins involved in pectin biosynthesis have not been completely purified to the exclusion of other activities involved pectic polysaccharide metabolism (Ohashi *et al.* 2007; Sterling*, et al.* 2006), suggesting that the proteins responsible for these activities may be in complex. Harholt *et al*. (2006) found that over-expression of some individual pectin biosynthetic enzymes did not result in increased accumulation of pectin, a result which could be explained by a requirement for protein complexes in pectin biosynthesis. If that is the case, over-expressing all components of the complex may be necessary to produce an over-accumulation of pectin. Complexes in pectin biosynthesis might function to ensure the proper structures or confirmation of the polysaccharide products, similarly to the putative roles described above for complexes synthesizing xyloglucan and cellulose. Proteins involved in sugar-nucleotide metabolism or transport might also be present in some or all polysaccharide biosynthetic complexes to accomplish increased efficiency through metabolite channeling. Extrapolating from various lines of evidence, it seems likely that protein complexes could be involved, if not always required, in the biosynthesis of many, and perhaps all cell wall polysaccharides.

The putative transmembrane pore that appears to be a component of a xyloglucan synthase complex could have importance for understanding cellulose biosynthesis. The co-expression and purification of CSLC4 and XXT1 in *Pichia*, which may work together to form a functional glucan synthase/translocase complex, might serve as a starting point for better understanding how cellulose synthase complexes accomplish a similar feat. The structure and organization of the pore components of cellulose synthase complexes are likely to influence how the individual cellulose strands are extruded and incorporated into the microfibrils present in the wall. At present this process is a complete mystery.

One of the major unresolved areas in cell wall biology is how the various components are deposited and incorporated into a functional cell wall. The topology and subcellular localization

studies on hemicellulose biosynthetic enzymes presented in Chapters 2 and 3 are a starting point for a better understanding of how their polysaccharide products are synthesized and trafficked. By combining what we learn from additional studies like these with analysis of the micro-scale properties of the plant cell wall during development, the field can begin to address exactly how components of the wall are deposited and incorporated.

Colocalization of hemicellulose biosynthetic enzymes, presented in Chapter 3, indicates that these proteins are most abundant in the Golgi apparatus. There is incomplete overlap between the fluorescent tagged proteins and the *trans* Golgi marker ST. Also, fluorescence from CSL proteins in compartments that appear to be TGN is located on the other side of ST signal relative to fluorescence from tagged hemicellulose biosynthetic enzymes in the Golgi proper (Figure 11d, arrowhead). Taken together, these data suggest that XXT1 and FUT1 are present in the *cis* and/or medial Golgi. CSLA9 and CSLC4 appear to localize to the *cis*/medial Golgi and the TGN.

One potential problem that can arise with fluorescent protein localization is that fluorescent tags can sometimes affect a protein's localization or activity. For example, some fluorescent protein fusions to CSLD proteins localize primarily at the ER when the fluorescent tag is on the carboxyl terminus, but localize at the Golgi when the the fluorescent tag is on the amino terminus (Bernal*, et al.* 2007; Li*, et al.* 2009). It is not likely that both of these localizations reflect the behavior of the native CSLD proteins. One way to increase confidence in our interpretations of fluorescence localization data is to confirm that the enzymatic activity of the tagged proteins is not compromised. For most of the mannan and xyloglucan biosynthetic enzymes described here (XXT1, FUT1, CSLA9), it would be possible to confirm that tagged versions of these enzymes are still active with *in vitro* activity assays, perhaps by expressing the fluorescence-tagged proteins in *Pichia*, which would have lower background for these activities

than plant cells would. However, it is still possible that a mis-targeted protein might retain its enzymatic activity, and we are currently unable to measure *in vitro* activity from CSLC4. A better approach to confirming that these fluorescence-tagged proteins are being trafficked properly would be to complement mutant phenotypes of gene knockouts. For the xyloglucan biosynthetic enzymes described here, this would require a cell wall polysaccharide composition analysis to detect whether biochemical phenotypes are complemented. For CSLA9, the *Agrobacterium* resistance phenotype (Zhu*, et al.* 2003) might be tested.

Trafficking of the proteins investigated in Chapter 3 to the appropriate positions within the Golgi apparatus could be occur through the mechanisms described in Chapter 1 (see the section "Vesicle trafficking in the secretory pathway: specific components and regulation" and figure 2). However, the only specific mechanistic determinant identified so far for localization of proteins to sub-compartments of the Golgi is an impact of TMD length on localization (Brandizzi*, et al.* 2002). It is possible that other factors may influence the free energy preference of a protein for a specific membrane environment and contribute to the observed localization pattern. Future studies might refine our understanding of intra-Golgi trafficking by combining directed modification of protein TMDs with confocal fluorescence analysis techniques that can allow researchers to extrapolate the precise position of proteins producing diffraction-limited signal (Snyder *et al.* 2004).

As our understanding of trafficking in the secretory pathway improves, researchers have begun to uncover the thermodynamic processes controlling protein and lipid sorting. We know much less about how cell wall polysaccharides produced in the secretory pathway are targeted and packaged for deposition in the wall. It could be that an unidentified protein binds or otherwise recognizes wall polysaccharides in the Golgi lumen and transmits a signal to cytosolic factors responsible for packaging the polymers into vesicles and trafficking to the plasma

membrane. Burton *et al*. (2010b) speculate that hemicellulosic polysaccharides, specifically mixed-linkage (1→3),(1→4)-β-glucan, may be covalently attached to a membrane-bound component in the lumen of the Golgi or secretory vesicles during synthesis and trafficking. An apoplastic activity would then free the polysaccharides from the membrane following vesicle fusion.

The fact that it is still unclear how structural polysaccharides are packaged, targeted, and utimately incorporated into to wall emphasizes how much we still have to learn about plant cell walls. Future advances will depend on a combination of classical biochemical techniques, such as the protease protection methods presented in this dissertation, with new technologies and techniques as they become available. It will also become more and more important to incorporate diverse techniques and perspectives, from computational biology to material physics, to further improve our understanding of all the processes that contribute to the production and properties of plant cell walls.

TABLE 1. PRIMER SEQUENCES

Primer Name	Sequence	Purpose
T7	ATGGCTAGCATGACTGGTGGACAGCAA ATGGGT	Sense sequence of the T7 epitope tag
A95’	ATGGAGCTAGGAGATACG	Amplification primers for *Pichia* expression
A93’	(TCA)ATGGTTAGGCACAATTG	
C45’	ATGGCTCCAAATTCAGTAG	
C43’	(CTA)GCTGATCTGTTCTCCG	
FT5’	ATGGATCAGAATTCGTACAGG	
FT3’	(TCA)TACTAGCTTAAGTCCCCAG	
attR	CATAGTGACTGGATATGTTGTGTTTTACA GTATTATGTAGTCTGTTTTTTATGCAAAA TCTAATTTAATATATTGATATTTATATCATT TTACGTTTCTCGTTCAGCTTTCTTGTACA AA	attR sequence for Gateway cloning
AOX2Fwd	TCGGAGTGAAAACCCCTTT	Primers for amplifying a region of the endogenous *Pichia* gene AOX2
AOX2Rev	TCGTAATCAGAAGCGGAACC	
XXT15'BamHI	GGTGCTGGATCCATGATAGAGAAGTGTA TAGGAGCGCATCG	Primers to add restriction enzymes sites for cloning into pVKH18En6 for fluorescent protein fusions
XXT13'SacI	GCGCAGGAGCTCTCACGTCGTCGTCGT ACTAAGCTTGG	
FUT15'BamHI	GGTGCTGGATCCATGGATCAGAATTCGT ACAGGAGAAGATCG	
FUT13'SacI	GCGCAGGAGCTCTCATACTAGCTTAAGT CCCCAGCTGATATCC	
CSLA95’XbaI	CAGGACGTCTAGATGGAGCTAGGAGAT ACGACGTCG	
CSLA93’SalI	CATGACCGTCGACATATGGTTAGGCACA ATTGTCCCAATTTGC	
CSLC45’BamHI	GGTGCTGGATCCATGGCTCCAAATTCAG TAGCAGTGAC	
CSLC43’SacI	GCGCAGGAGCTCTCAGCTGATCTGTTCT CCGATCAAATCCAAC	

APPENDICES

APPENDIX A: Preliminary evidence for protein:protein interactions between CSLC4 and XXT1

A previous study (Cocuron*, et al.* 2007) indicated that there may be a direct interaction between the xyloglucan glucan synthase, CSLC4 and the xyloglucan xylosyltransferase XXT1. The authors of this study found that glucan synthase activity from CSLC4 was stimulated by the presence of XXT1 protein in a system (*Pichia pastoris*) that does not produce the substrate for XXT1, UDP-xylose. In brief, *Pichia* cells expressing CSLC4 produced soluble oligo glucans that were not detected in wild-type yeast. When XXT1 was co-expressed with CSLC4, the *Pichia* cultures accumulated insoluble (longer) β(1→4)-linked glucan polymers. Thus, it is possible that XXT1 and CSLC4 proteins directly interact. Also, a protein complex of xyloglucan biosynthetic enzymes could be a mechanism by which plants ensure the regular structure of this polymer. Colocalization at the Golgi of fluorescent protein fusions to these enzymes, presented in Chapter 3 of this document, indicate that such an interaction is possible in plant cells.

Figure 12 shows the results of an experiment where an immobilized T7 antibody was used to co-purify T7 epitope-tagged CSLC4 and XXT1 from *Pichia* expressing both proteins. The positions of CSLC4 (a) and XXT1 (b) are indicated with stars. A T7 epitope tag was added to the amino terminus of CSLC4 and T7 antibodies were used to isolate T7-CSLC4 and interacting proteins. The presence of XXT1 was evaluated with a polyclonal antibody raised against this protein. I was able to detect XXT1 in the wash steps as well as in the fraction bound to anti-T7 beads (Figure 12b circle), indicating that the interaction between XXT1 and CSLC4 protein is stable enough to be preserved through several wash steps and thus, crosslinking with formaldehyde may not be necessary to observe this interaction. Interestingly, the low pH elution buffer designed to release T7-CSLC4 bound to the column did not efficiently disrupt the

antibody affinity, but did gradually deplete XXT1 from the column.

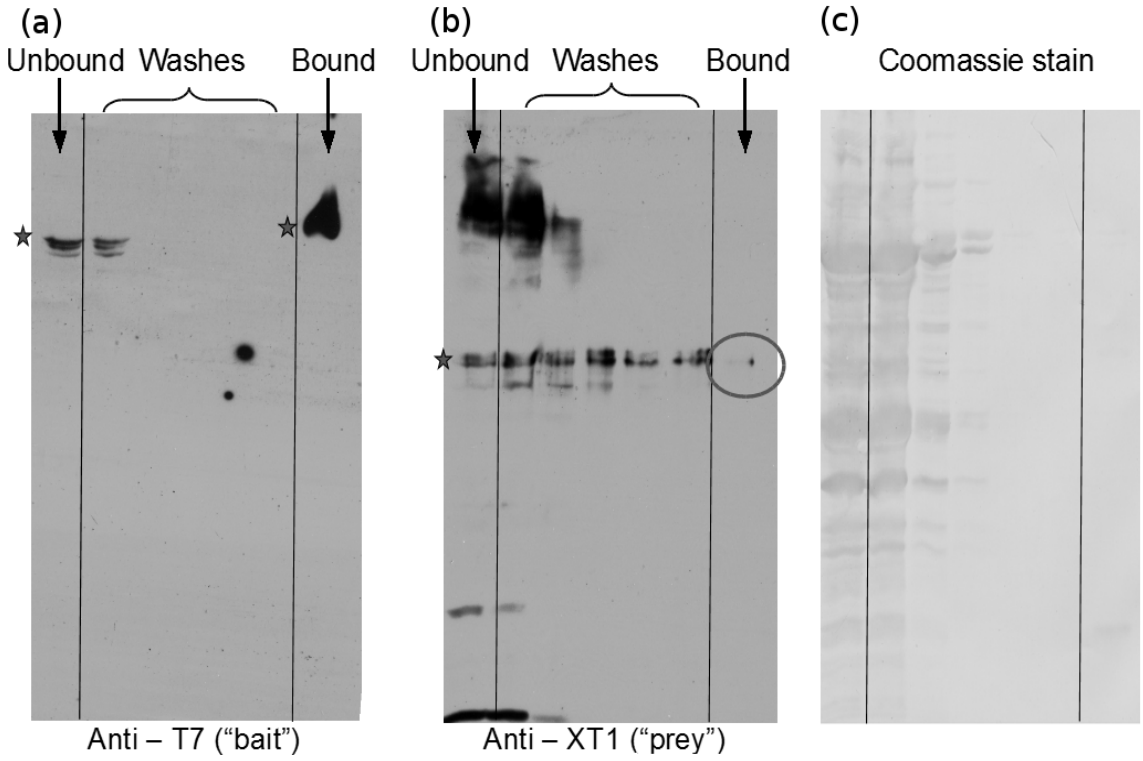

Figure 12. XXT1 interacts with T7-CSLC4 in *Pichia*.
Immunoblots using antibodies against the T7 epitope fused to CSLC4 (a) and XXT1 protein (b) coexpressed in *Pichia pastoris*. Separate fractions from the T7 antibody affinity purification described in this appendix were loaded to each lane of duplicate gels before transfer to the membrane and subsequent western blotting. The membrane used to expose the film in (b) was stained with Coomassie blue for total protein content (c). The positions of CSLC4 and XXT1 protein in (a) and (b) are indicated with a star.

These data are an important first step in determining the composition of a xyloglucan biosynthetic complex. Figure 12 shows that XXT1 and CSLC4 proteins interact in *Pichia*. The next steps to confirm this interaction and investigate additional components of a putative complex should be done in plants, if possible. To this end, I have begun to develop a method for the purification of fluorescence-tagged hemicellulose biosynthetic enzymes from plant cell cultures based on the solubilization conditions determined for CSLC4 in *Pichia* that lead to the data presented in Figure 12. In theory, any or all of the various fluorescence-tagged constructs presented in Chapter 3 could be suitable as "bait" for an antibody-epitope affinity chromatography approach similar to the one presented in Figure 12, using an antibody against GFP.

Experimental procedures

Pichia spheroplasts were prepared as described in the Experimental procedures section of Chapter 2. Following the 13,000 x g spin, formaldehyde was added to the supernatant to a final concentration of 0.2%. The lysate was allowed to incubate with the crosslinker on ice for 20 minutes. Crosslinking reactions were terminated by the addition of 2.5 M glycine to a final concentration of 0.125 M, and microsomal vesicles were pelleted by centrifugation at 120,000 x g for 60 min. Vesicles were resuspended in 1x bind/wash buffer + 0.5% Triton X-100 from a T7-tag affinity purification kit (Novagen), subjected to centrifugation at 13,000 x g for 10 minutes to pellet aggregated vesicles. The supernatant was applied to T7-affinity purification beads from Novagen and incubated at 4° C for 30 min. The beads and lysate were applied to Micro Bio-Spin chromatography columns, and the lysate was allowed to flow through (unbound fraction). The beads were then washed 3 times with one bed volume of bind/wash buffer + 0.5% Triton X-100 for each wash. The beads were then washed 2 times with one bed volume of T7 affinity elution buffer (Novagen). All fractions were run on 10% acrylamide SDS-PAGE before

transfer to PVDF membranes and blotting with the appropriate antibodies.

APPENDIX B: Unsuccessful approaches used to probe CSL membrane protein topology

Probing the membrane topology of a protein often relies on the use of non-native systems to produce material to analyze. Also, topology is a complex property, so results from a single approach can sometimes be difficult to interpret. For theses reasons, I attempted multiple approaches to address the membrane topology of CSLA9 and CSLC4 proteins. In the end, the technical difficulties of a number of these techniques proved impossible to overcome.

The first alternative technique to be tried was the N-linked glycosylation approach. This method exploits the fact that proteins entering the eukaryotic secretory pathway can be glycosylated on sites that face the ER lumen. This modification to the protein can be detected because it alters the rate of migration in separation techniques such as SDS-PAGE. Glycosylation can be reversed *in vitro* by a specific endoglycosidase such as EndoH. This means that N-linked glycosylation can be used to show if a region of a polytopic protein faces the lumen of the ER by removing all native glycosylation sites and then introducing sites at specific points in the protein sequence. For some examples of the successful use of N-linked glycosylation to probe membrane protein topology, see the following references (Devoto*, et al.* 1999; Popov *et al.* 1999; Popov *et al.* 1997).

In practice, N-linked glycosylation relies on *in vitro* translation of a protein of interest. In brief, a membrane protein is produced *in vitro* to yield radiolabeled product (Melancon and Garoff 1986), which can be incubated with microsomes (ER) allowing membrane insertion (Walter and Blobel 1983) and glycosylation of sites that face the ER lumen. For *in vitro* translation, I cloned the coding regions for CSLA and CSLC proteins into pDEST14 (Invitrogen). This vector contains a T7 RNA polymerase initiation site for *in vitro* transcription. These constructs were employed with the TNT® coupled transcription/translation system from

Promega. Production of 35S-methionine labeled CSL protein was successful, but the addition of canine pancreatic microsomes (Promega) prevented detectable accumulation of the protein for unknown reasons (figure 13a, compare to positive controls figure 13b). Plant microsomes prepared from *Arabidopsis* cell culture were tried as a substitute for canine pancreatic microsomes, but these membrane preparations did not result in import or glycosylation of control proteins (Figure 13c). These experiments were done by John Froehlich, an expert with *in vitro* translation techniques.

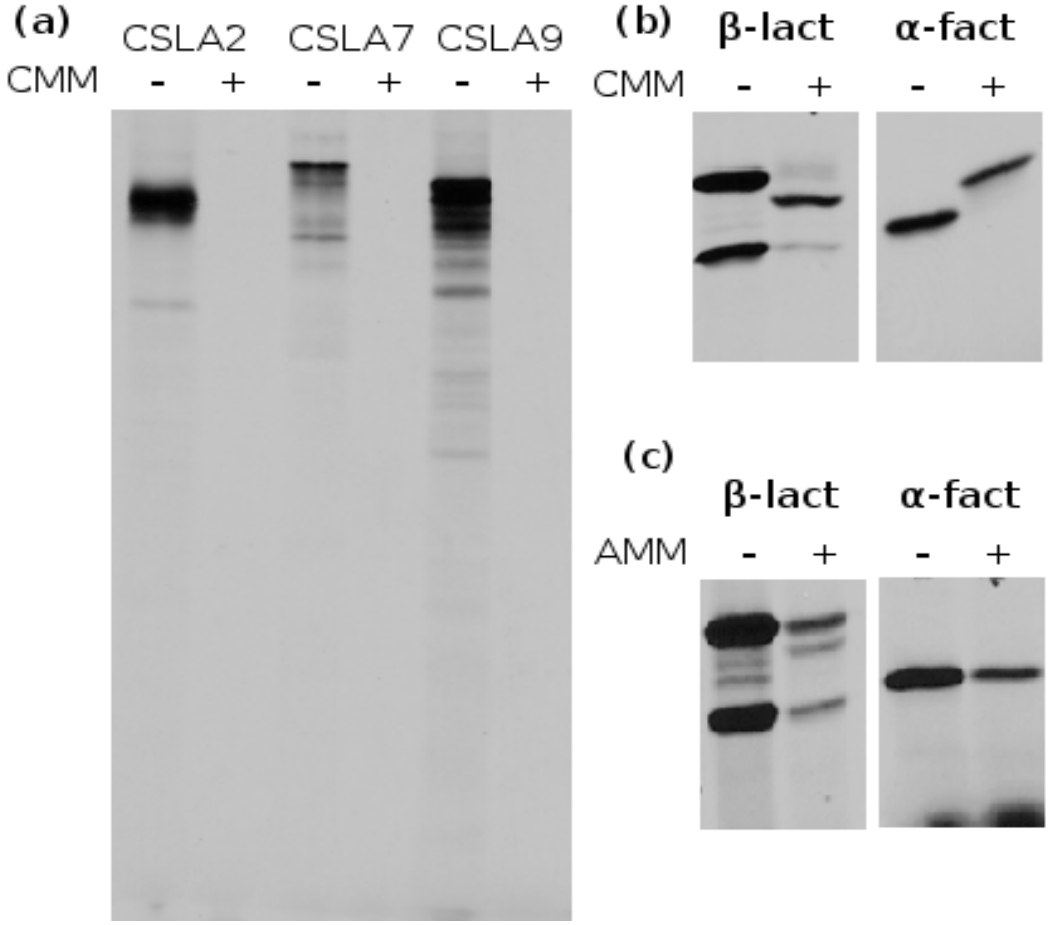

Figure 13. Addition of canine microsomal membranes inhibits *in vitro* translation of CSLA proteins.
Autoradiographs of *in vitro* translation of CSLA proteins (a) in the presence of canine microsomal membranes (CMM) resolved by SDS-PAGE. The addition of CMM to the reactions inhibits translation of CSLA proteins. Positive controls (β-lactamase, α-mating factor) are shown in (b). The size shifts of the bands corresponding to β-lactamase and α-mating factor indicate successful translocation by signal peptide cleavage and glycosylation, respectively. *Arabidopsis* microsomal membranes (AMM) were tried as an alternative to canine microsomes, but translocation of control proteins was not detected (c).

Another approach that I attempted to probing the topology of CSL proteins is the bimolecular fluorescence complementation (BiFC) fluorescence topology technique described in Zamyatnin *et al.* (2006). This technique relies on interaction between two non-fluorescent pieces of a GFP-variant protein that form a functional fluor upon their interaction in living cells. One part of the GFP-variant protein is fused to a protein of interest, and the other part is made to accumulate in a specific compartment of the cell. If both parts reside in the same compartment, the two parts of the GFP-variant protein can interact, producing fluorescence and indicating the compartment occupied by the tag on the protein of interest. Despite attempting a variety of combinations of expression patterns, times, and infiltration conditions, fluorescence was not observed from any of the CSL protein fusions. It is possible that the CSL proteins did not accumulate to high enough levels for the non-specific interaction between the fluorescent protein parts to be detected. Although additional conditions could have been attempted to promote this interaction, the approach only allows detection of the position of the *termini* of a full-length protein, data which we had already obtained from protease protection of CSL protenis produced in *Pichia* (described in chapter 2). Since the data that could be generated though the BiFC approach would not be qualitatively different from what we had learned in *Pichia*, the BiFC method was not pursued further.

REFERENCES

REFERENCES

Amor, Y., Haigler, C.H., Johnson, S., Wainscott, M. and Delmer, D.P. (1995) A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. . *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 9353-9357.

Amos, W.B. and White, J.G. (2003) How the confocal laser scanning microscope entered biological research. *Biology of the Cell*, **95**, 335-342.

Atanassov, I.I., Pittman, J.K. and Turner, S.R. (2009) Elucidating the Mechanisms of Assembly and Subunit Interaction of the Cellulose Synthase Complex of Arabidopsis Secondary Cell Walls. *Journal of Biological Chemistry*, **284**, 3833-3841.

Baldwin, T.C., Handford, M.G., Yuseff, M.-I., Orellana, A. and Dupree, P. (2001) Identification and Characterization of GONST1, a Golgi-Localized GDP-Mannose Transporter in Arabidopsis. *Plant Cell*, **13**, 2283-2295.

Barlowe, C. (2003) Signals for COPII-dependent export from the ER: what's the ticket out? *Trends in Cell Biology*, **13**, 295-300.

Batoko, H., Zheng, H.Q., Hawes, C. and Moore, I. (2000) A Rab1 GTPase is required for transport between the endoplasmic reticulum and Golgi apparatus and for normal Golgi movement in plants. *The Plant Cell Online*, **12**, 2201-2201.

Belford, D.S. and Preston, R.D. (1960) The Structure and Growth of Root Hairs. *Journal of Experimental Botany*, **12**, 157-168.

Benschop, J.J., Mohammed, S., O'Flaherty, M., Heck, A.J.R., Slijper, M. and Menke, F.L.H. (2007) Quantitative phosphoproteomics of early elicitor signaling in Arabidopsis. *Molecular & Cellular Proteomics*, **6**, 1198-1198.

Bernal, A.J., Jensen, J.K., Harholt, J., Sørensen, S., Moller, I., Blaukopf, C., Johansen, B., de Lotto, R., Pauly, M., Scheller, H.V. and Willats, W.G.T. (2007) Disruption of AtCSLD5 results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in Arabidopsis. *The Plant Journal*, **52**, 791-802.

Bernal, A.J., Yoo, C.M., Mutwil, M., Jensen, J.K., Hou, G., Blaukopf, C., Sorensen, I., Blancaflor, E.B., Scheller, H.V. and Willats, W.G.T. (2008) Functional analysis of the cellulose synthase-like genes CSLD1, CSLD2, and CSLD4 in tip-growing Arabidopsis cells. *Plant Physiology*, **148**, 1238-1238.

Bernsel, A., Viklund, H., Hennerdal, A. and Elofsson, A. (2009) TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acid Research*, **37**, W465-W468.

Bertozzi, C.R. and Rabuka, D. (2008) Structural basis of glycan diversity. In *Essentials of Glycobiology* (Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi,

C.R., Hart, G.W. and Etzler, M.E. eds). Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 23-36.

Boevink, P., Oparka, K., Santa Cruz, S., Martin, B., Betteridge, A. and Hawes, C. (1998) Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *Plant Journal*, **15**, 441-447.

Bogdanov, M., Heacock, P.N. and Dowhan, W. (2002) A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. *EMBO*, **21**, 2107-2116.

Bogdanov, M., Xie, J. and Dowhan, W. (2009) Lipid-Protein Interactions Drive Membrane Protein Topogenesis in Accordance with the Positive Inside Rule. *Journal of Biological Chemistry*, **284**, 9637-9641.

Bogdanov, M., Xie, J., Heacock, P. and Dowhan, W. (2008) To flip or not to flip: lipid–protein charge interactions are a determinant of final membrane protein topology. *The Journal of Cell Biology*, **182**, 925-925.

Bonawitz, N.D. and Chapple, C. (2010) The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. *Annual Review of Genetics*, **44**, 337-363.

Bonfanti, L., Mironov, A.A.J., Martínez-Menárguez, J.A., Martella, O., Fusella, A., Baldassarre, M., Buccione, R., Geuze, H.J., Mironov, A.A. and Luini, A. (1998) Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell*, **95**, 993-1003.

Brach, T., Soyk, S., Müller, C., Hinz, G., Hell, R., Brandizzi, F. and Meyer, A.J. (2009) Non-invasive topology analysis of membrane proteins in the secretory pathway. *The Plant Journal*, **57**, 534-541.

Brandizzi, F., Frangne, N., Marc-Martin, S., Hawes, C., Neuhaus, J.-M. and Paris, N. (2002) The Destination for Single-Pass Membrane Proteins Is Influenced Markedly by the Length of the Hydrophobic Domain. *Plant Cell*, **14**, 1077-1092.

Burgess, J. (1985) *An introduction to plant cell development* Cambridge, UK: University Press.

Burget, E.G., Verma, R., Molhoj, M. and Reiter, W.D. (2003) The Biosynthesis of L-Arabinose in Plants: Molecular Cloning and Characterization of a Golgi-Localized UDP-D-Xylose 4-Epimerase Encoded by the MUR4 Gene of Arabidopsis. *Plant Cell*, **15**, 523-531.

Burton, R.A., Collins, H.M., Kibble, N.A.J., Smith, J.A., Shirley, N.J., Jobling, S.A., Henderson, M., Singh, R.R., Pettolino, F., Wilson, S.M., Bird, A.R., Topping, D.L., Bacic, A. and Fincher, G.B. (2010a) Over-expression of specific HvCslF cellulose synthase-like genes in transgenic barley increases the levels of cell wall (1,3;1,4)-β-d-glucans and alters their fine structure. *Plant Biotechnology Journal*.

Burton, R.A., Gidley, M.J. and Fincher, G.B. (2010b) Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat Chem Biol*, **6**, 724-732.

Burton, R.A., Wilson, S.M., Hrmova, M., Harvey, A.J., Shirley, N.J., Medhurst, A., Stone, B.A., Newbigin, E.J., Bacic, A. and Fincher, G.B. (2006) Cellulose Synthase-Like CslF Genes Mediate the Synthesis of Cell Wall (1→ 3),(1→ 4)-β-D-Glucans. *Science*, **311**, 1940-1942.

Carpita, N.C. and McCann, M.C. (2000) Chapter 2 "The Cell Wall". In *Biochemistry and molecular biology of plants* (Buchanan, B.B., Gruissem, W. and Jones, R. eds). Rockville, MD: ASPB.

Cavalier, D.M. and Keegstra, K. (2006) Two Xyloglucan Xylosyltransferases Catalyze the Addition of Multiple Xylosyl Residues to Cellohexaose. *Journal of Biological Chemistry*, **281**, 34197-34207.

Chevalier, L., Bernard, S., Ramdani, Y., Lamour, R., Bardor, M., Lerouge, P., Follet-Gueye, M.-L. and Driouich, A. (2010) Subcompartment localization of the side chain xyloglucan-synthesizing enzymes within Golgi stacks of tobacco suspension-cultured cells. *The Plant Journal*, in press.

Cocuron, J.C., Lerouxel, O., Drakakaki, G., Alonso, A.P., Liepman, A.H., Keegstra, K., Raikhel, N. and Wilkerson, C.G. (2007) A gene from the cellulose synthase-like C family encodes a β-1, 4 glucan synthase. *Proceedings of the National Academy of Sciences, USA*, **104**, 8550-8555.

Cosgrove, D.J. (2000) Expansive growth of plant cell walls. *Plant physiology and Biochemistry*, **38**, 109-124.

Cosgrove, D.J. (2001) Wall Structure and Wall Loosening. A Look Backwards and Forwards. *Plant Physiology*, **125**, 131-134.

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

Delmer, D.P. (1999) Cellulose biosynthesis: exciting times for a difficult field of study. *Annual Review of Plant Biology*, **50**, 245–276.

Delmer, D.P. and Amor, Y. (1995) Cellulose biosynthesis. *The Plant Cell*, **7**, 987-987.

Desprez, T., Juraniec, M., Crowell, E.F., Jouy, H., Pochylova, Z., Parcy, F., Höfte, H., Gonneau, M. and Vernhettes, S. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 15572–15577.

Devoto, A., Piffanelli, P., Nilsson, I.M., Wallin, E., Panstruga, R., von Heijne, G. and Schulze-Lefert, P. (1999) Topology, subcellular localization, and sequence diversity of the Mlo family in plants. *Journal of Biological Chemistry*, **274**, 34993-35004.

Dhugga, K.S., Barreiro, R., Whitten, B., Stecca, K., Hazebroek, J., Randhawa, G.S., Dolan, M., Kinney, A.J., Tomes, D., Nichols, S. and Anderson, P. (2004) Guar Seed β-Mannan Synthase Is a Member of the Cellulose Synthase Super Gene Family. *Science*, **303**, 363-366.

Dickinson, E. (2003) Hydrocolloids at interfaces and the influence on the properties of dispersed systems. *Food Hydrocolloids*, **17**, 25-39.

Doblin, M.S., Pettolino, F.A., Wilson, S.M., Campbell, R., Burton, R.A., Fincher, G.B., Newbigin, E. and Bacic, A. (2009) A barley cellulose synthase-like CSLH gene mediates (1, 3; 1, 4)-β-D-glucan synthesis in transgenic *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA*, **106**, 5996-6001.

Donohoe, B.S., Kang, B.H. and Staehelin, L.A. (2007) Identification and characterization of COPIa- and COPIb-type vesicle classes associated with plant and algal Golgi. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 163-168.

Dunkley, T.P.J., Hester, S., Shadforth, I.P., Runions, J., Weimar, T., Hanton, S.L., Griffin, J.L., Bessant, C., Brandizzi, F., Hawes, C., Watson, R.B., Dupree, P. and Lilley, K.S. (2006) Mapping the Arabidopsis organelle proteome. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 6518-6523.

Dupree, P. and Sherrier, D.J. (1998) The plant Golgi apparatus. *BBA-Molecular Cell Research*, **1404**, 259–270.

Dwivany, F.M., Yulia, D., Burton, R.A., Shirley, N.J., Wilson, S.M., Fincher, G.B., Bacic, A., Newbigin, E. and Doblin, M.S. (2009) The cellulose-synthase like C (CSLC) family of barley includes members that are integral membrane proteins targeted to the plasma membrane. *Molecular Plant*, **2**, 1025-1039.

Edwards, C.A. and Parrett, A.M. (1996) Plant cell wall polysaccharides, gums and hydrocolliods, nutritional aspects. In *Carbohydrates in Food* (Eliasson, A. ed. New York, NY: Marcel Dekker, Inc, pp. 319-346.

Edwards, M.E., Dickson, C.A., Chengappa, S., Sidebottom, C., Gidley, M.J. and Reid, J.S.G. (1999) Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *Plant Journal*, **19**, 691–698.

Faik, A., Bar-Peled, M., DeRocher, A.E., Zeng, W., Perrin, R.M., Wilkerson, C., Raikhel, N.V. and Keegstra, K. (2000) Biochemical characterization and molecular cloning of an α-1, 2-fucosyltransferase that catalyzes the last step of cell wall xyloglucan biosynthesis in pea. *Journal of Biological Chemistry*, **275**, 15082-15089.

Faik, A., Price, N.J., Raikhel, N.V. and Keegstra, K. (2002) An Arabidopsis gene encoding an α-xylosyltransferase involved in xyloglucan biosynthesis. *Proceedings of the National*

Academy of Sciences of the United States of America, **99**, 7797-7802.

Frei, E. and Preston, R.D. (1968) Non-Cellulosic Structural Polysaccharides in Algal Cell Walls. III. Mannan in Siphoneous Green Algae. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **169**, 127-145.

Fujii, S., Hayashi, T. and Mizuno, K. (2010) Sucrose synthase is an integral component of the cellulose synthesis machinery. *Plant and Cell Physiology*, **51**, 294-294.

Gilstring, C.F. and Ljungdahl, P.O. (2000) A Method for Determining the in Vivo Topology of Yeast Polytopic Membrane Proteins Demonstrates That Gap1p Fully Integrates into the Membrane Independently of Shr3p. *Journal of Biological Chemistry*, **275**, 31488-31495.

Gkantiragas, I., Brugger, B., Stuven, E., Kaloyanova, D., Li, X.Y., Lohr, K., Lottspeich, F., Wieland, F.T. and Helms, J.B. (2001) Sphingomyelin-enriched microdomains at the Golgi complex. *Molecular biology of the cell*, **12**, 1819-1819.

Goder, V., Bieri, C. and Spiess, M. (1999) Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. *Journal of Cell Biology*, **147**, 257-266.

Goubet, F., Barton, C.J., Mortimer, J.C., Yu, X., Zhang, Z., Miles, G.P., Richens, J., Liepman, A.H., Seffen, K. and Dupree, P. (2009) Cell wall glucomannan in *Arabidopsis* is synthesised by CSLA glycosyltransferases, and influences the progression of embryogenesis. *The Plant Journal*, **60**, 527-538.

Goubet, F. and Mohnen, D. (1999) Subcellular localization and topology of homogalacturonan methyltransferase in suspension-cultured <i>Nicotiana tabacum</i> cells. *Planta*, **209**, 112-117-117.

Graham, T.R., Seeger, M., Payne, G.S., MacKay, V.L. and Emr, S.D. (1994) Clathrin-dependent localization of α1,3 mannosyltransferase to the Golgi complex of *Saccharomyces cerevisiae*. *Journal of Cell Biology*, **127**, 667-678.

Griffiths, G. and Simons, K. (1986) The trans Golgi network: sorting at the exit site of the Golgi complex. *Science*, **234**, 438-443.

Gu, F., Crump, C.M. and Thomas, G. (2001) Trans-Golgi network sorting. *Cellular and Molecular Life Sciences*, **58**, 1067-1084.

Guerriero, G., Fugelstad, J. and Bulone, V. (2010) What Do We Really Know about Cellulose Biosynthesis in Higher Plants? *Journal of Integrative Plant Biology*, **52**, 161-175.

Haigler, C.H., Ivanova-Datcheva, M., Hogan, P.S., Salnikov, V.V., Hwang, S., Martin, K. and Delmer, D.P. (2001) Carbon partitioning to cellulose synthesis. *Plant Molecular Biology*, **47**, 29-51.

Handford, M.G., Sicilia, F., Brandizzi, F., Chung, J.H. and Dupree, P. (2004) *Arabidopsis thaliana* expresses multiple Golgi-localised nucleotide-sugar transporters related to GONST1. *Molecular Genetics and Genomics*, **272**, 397-410.

Hanson, G.T., Aggeler, R., Oglesbee, D., Cannon, M., Capaldi, R.A., Tsien, R.Y. and Remington, S.J. (2004) Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. *Journal of Biological Chemistry*, **279**, 13044–13053.

Hanton, S.L. and Brandizzi, F. (2006) Fluorescent proteins as markers in the plant secretory pathway. *Microscopy research and technique*, **69**, 152–159.

Harholt, J., Jensen, J.K., Sorensen, S.O., Orfila, C., Pauly, M. and Scheller, H.V. (2006) ARABINAN DEFICIENT 1 Is a Putative Arabinosyltransferase Involved in Biosynthesis of Pectic Arabinan in Arabidopsis. *Plant Physiol.*, **140**, 49-58.

Hayashi, T. (1989) Xyloglucans in the primary cell wall. *Annual Review of Plant Biology*, **40**, 139–168.

Hazen, S.P., Scott-Craig, J.S. and Walton, J.D. (2002) Cellulose Synthase-Like Genes of Rice. *Plant Physiology*, **128**, 336-340.

Heim, R., Prasher, D.C. and Tsien, R.Y. (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 12501-12504.

Held, M.A., Penning, B., Brandt, A.S., Kessans, S.A., Yong, W., Scofield, S.R. and Carpita, N.C. (2008) Small-interfering RNAs from natural antisense transcripts derived from a cellulose synthase gene modulate cell wall biosynthesis in barley. *Proceedings of the National Academy of Sciences*, **105**, 20534-20539.

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S.H. and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, **433**, 377-381.

Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S.H. and von Heijne, G. (2007) Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature*, **450**, 1026-1030.

Höfte, H. (2010) Plant Cell Biology: How to Pattern a Wall. *Current Biology*, **20**, R450–R452.

Johnson, A.E. and van Waes, M.A. (1999) The Translocon: A Dynamic Gateway at the ER Membrane. *Annual Review of Cell and Developmental Biology*, **15**, 799-842.

Jurgens, G. (2004) Membrane trafficking in plants. *Annual Review of Cell and Developmental Biology*, **20**, 481-504.

Kageyama-Yahara, N. and Riezman, H. (2006) Transmembrane topology of ceramide synthase

in yeast. *Biochemical Journal*, **398**, 585-593.

Kahsay, R.Y., Gao, G. and Liao, L. (2005) An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. *Bioinformatics*, **21**, 1853-1858.

Keegstra, K. and Raikhel, N. (2001) Plant glycosyltransferases. *Current Opinion in Plant Biology*, **4**, 219-224.

Keegstra, K., Talmadge, K.W., Bauer, W.D. and Albersheim, P. (1973) The Structure of Plant Cell Walls: III. A Model of the Walls of Suspension-cultured Sycamore Cells Based on the Interconnections of the Macromolecular Components. *Plant Physiology*, **51**, 188-197.

Kiefer, L.L., York, W.S., Darvill, A.G. and Albersheim, P. (1989) Xyloglucan isolated from suspension-cultured sycamore cell walls is O-acetylated. *Phytochemistry*, **28**, 2105-2107.

Koch, K. (2004) Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. *Current Opinion in Plant Biology*, **7**, 235–246.

Kondo, T., Togawa, E. and Brown, R.M. (2001) "Nematic Ordered Cellulose": A Concept of Glucan Chain Association. *Biomacromolecules*, **2**, 1324-1330.

Lakatos, L., Csorba, T., Pantaleo, V., Chapman, E.J., Carrington, J.C., Liu, Y.-P., Dolja, V.V., Calvino, L.F., Lopez-Moya, J.J. and Burgyan, J. (2006) Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J*, **25**, 2768-2780.

Li, M., Xiong, G., Li, R., Cui, J., Tang, D., Zhang, B., Pauly, M., Cheng, Z. and Zhou, Y. (2009) Rice cellulose synthase-like D4 is essential for normal cell-wall biosynthesis and plant growth. *The Plant Journal*, **60**, 1055-1069.

Liepman, A., Cavalier, D., Lerouxel, O. and Keegstra, K. (2007a) Cell wall structure, biosynthesis and assembly. In *Plant cell separation and adhesion* (Roberts, J.A. and Gonzales-Carranza, Z. eds). Oxford, UK: Blackwell Publishing, pp. 8-39.

Liepman, A.H., Nairn, C.J., Willats, W.G.T., Sorensen, I., Roberts, A.W. and Keegstra, K. (2007b) Functional Genomic Analysis Supports Conservation of Function Among CslA Gene Family Members and Suggests Diverse Roles of Mannans in Plants. *Plant Physiology*, **143**, 1881-1893.

Liepman, A.H., Wilkerson, C.G. and Keegstra, K. (2005) Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. *Proceedings of the National Academy of Sciences*, **102**, 2221-2226.

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

Online, **15**, 1662-1670.

Maeda, Y., Awano, T., Takabe, K. and Fujita, M. (2000) Immunolocalization of glucomannans in the cell wall of differentiating tracheids in *Chamaecyparis obtusa*. *Protoplasma*, **213**, 148-156.

Maley, F., Trimble, R.B., Tarentino, A.L. and Plummer, T.H. (1989) Characterization of Glycoproteins and Their Associated Oligosaccharides through the Use of Endoglycosidases. *Analytical Biochemistry*, **180**, 195-204.

Manfield, I.W., Orfila, C., McCartney, L., Harholt, J., Bernal, A.J., Scheller, H.V., Gilmartin, P.M., Mikkelsen, J.D., Paul Knox, J. and Willats, W.G.T. (2004) Novel cell wall architecture of isoxaben-habituated Arabidopsis suspension-cultured cells: global transcript profiling and cellular analysis. *The Plant Journal*, **40**, 260-275.

Meister, G. and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343-349.

Melancon, P. and Garoff, H. (1986) Reinitiation of translocation in the Semliki Forest virus structural polyprotein: identification of the signal for the E1 glycoprotein. *The EMBO Journal*, **5**, 1551-1551.

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

Moore, P.J. and Staehelin, L.A. (1988) Immunogold localization of the cell-wall-matrix polysaccharides rhamnogalacturonan I and xyloglucan during cell expansion and cytokinesis in *Trifolium pratense L.*; implication for secretory pathways. *Planta*, **174**, 433–445.

Müller-Taubenberger, A. and Anderson, K.I. (2007) Recent advances using green and red fluorescent protein variants. *Applied Microbiology and Biotechnology*, **77**, 1-12.

Muñoz, P., Norambuena, L. and Orellana, A. (1996) Evidence for a UDP-glucose transporter in Golgi apparatus-derived vesicles from pea and its possible role in polysaccharide biosynthesis. *Plant Physiology*, **112**, 1585-1594.

Nasie, I., Steiner-Mordoch, S., Gold, A. and Schuldiner, S. (2010) Topologically random insertion of EMRE supports a pathway for evolution of enverted repeats in ion-coupled transporters. *Journal of Biological Chemistry*, **285**, 15234-15244.

Nebenführ, A. and Staehelin, L.A. (2001) Mobile factories: Golgi dynamics in plant cells. *Trends in Plant Science*, **6**, 160-167.

Neumann, U., Brandizzi, F. and Hawes, C. (2003) Protein transport in plant cells: in and out of the Golgi. *Annals of Botany*, **92**, 167-180.

O'Neill, M.A. and York, W.S. (2003). In *The plant cell wall* (Rose, J.K.C. ed. Boca Raton, FL: CRC Press, pp. 1-54.

Obel, N., Erben, V., Schwarz, T., Kühnel, S., Fodor, A. and Pauly, M. (2009) Microanalysis of Plant Cell Wall Polysaccharides. *Molecular plant*, **2**, 922-932.

Ohashi, T., Ishimizu, T., Akita, K. and Hase, S. (2007) In Vitro Stabilization and Minimum Active Component of Polygalacturonic Acid Synthase Involved in Pectin Biosynthesis. *Bioscience, Biotechnology, and Biochemistry*, **71**, 2291-2299.

Paredez, A.R., Somerville, C.R. and Ehrhardt, D.W. (2006) Visualization of Cellulose Synthase Demonstrates Functional Association with Microtubules. *Science*, **312**, 1491-1495.

Patterson, G.H., Hirschberg, K., Polishchuk, R.S., Gerlich, D., Phair, R.D. and Lippincott-Schwartz, J. (2008) Transport through the Golgi Apparatus by Rapid Partitioning within a Two-Phase Membrane System. *Cell*, **133**, 1055-1067.

Paul, S., Gable, K. and Dunn, T.M. (2007) A Six-membrane-spanning Topology for Yeast and Arabidopsis Tsc13p, the Enoyl Reductases of the Microsomal Fatty Acid Elongating System. *Journal of Biological Chemistry*, **282**, 19237-19246.

Pear, J.R., Kawagoe, Y., Schreckengost, W.E., Delmer, D.P. and Stalker, D.M. (1996) Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 12637-12637.

Pelham, H.R. and Rothman, J.E. (2000) The debate about transport in the Golgi--two sides of the same coin? *Cell*, **102**, 713-719.

Pelham, H.R.B. (1998) Getting through the Golgi complex. *Trends in Cell Biology*, **8**, 45-49.

Peña, M.J., Ryden, P., Madson, M., Smith, A.C. and Carpita, N.C. (2004) The galactose residues of xyloglucan are essential to maintain mechanical strength of the primary cell walls in *Arabidopsis* during growth. *Plant Physiology*, **134**, 443-451.

Peng, L., Kawagoe, Y., Hogan, P. and Delmer, D. (2002) Sitosterol-β-glucoside as primer for cellulose synthesis in plants. *Science*, **295**, 147-150.

Perrin, R., Wilkerson, C. and Keegstra, K. (2001) Golgi enzymes that synthesize plant cell wall polysaccharides: finding and evaluating candidates in the genomic era. *Plant Molecular Biology*, **47**, 115-130.

Perrin, R.M., DeRocher, A.E., Bar-Peled, M., Zeng, W., Norambuena, L., Orellana, A., Raikhel, N.V. and Keegstra, K. (1999) Xyloglucan Fucosyltransferase, an Enzyme Involved in Plant Cell Wall Biosynthesis. *Science*, **284**, 1976-1979.

Pfeffer, S.R. and Rothman, J.E. (1987) Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. *Annual review of biochemistry*, **56**, 829-852.

Popov, M., Li, J. and Reithmeier, R.A. (1999) Transmembrane folding of the human erythrocyte anion exchanger (AE1, Band 3) determined by scanning and insertional N-glycosylation mutagenesis. *Biochemical Journal*, **339**, 269-279.

Popov, M., Tam, L.Y., Li, J. and Reithmeier, R.A.F. (1997) Mapping the ends of transmembrane segments in a polytopic membrane protein. Scanning N-glycosylation mutagenesis of extracytosolic loops in the anion exchanger, band 3. *Journal of Biological Chemistry*, **272**, 18325-18332.

Popper, Z.A. and Tuohy, M.G. (2010) Beyond the Green: Understanding the Evolutionary Puzzle of Plant and Algal Cell Walls. *Plant Physiol*, **153**, 373-383.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. and Cormier, M.J. (1992) Primary structure of the Aequorea victoria green-fluorescent protein. *Gene*, **111**, 229-233.

Price Nicholas, J., Reiter, W.-D. and Raikhel Natasha, V. (2008) Molecular Genetics of Non-processive Glycosyltransferases. In *The Arabidopsis Book*: The American Society of Plant Biologists, pp. 1-7.

Rapoport, T., Goder, V., Heinrich, S. and Matlack, K. (2004) Membrane-protein integration and the role of the translocation channel. *Trends in Cell Biology*, **14**, 568-575.

Rapp, M., Granseth, E., Seppälä, S. and von Heijne, G. (2006) Identification and evolution of dual-topology membrane proteins. *Nature Structural & Molecular Biology*, **13**, 112-116.

Rapp, M., Seppala, S., Granseth, E. and von Heijne, G. (2007) Emulating Membrane Protein Evolution by Rational Design. *Science*, **315**, 1282-1284.

Ray, P.M., Green, P.B. and Cleland, R., 163–& (1972) Role of turgor in plant cell growth. *Nature*, **239**.

Reid, J.S.G. and Meier, H. (1970) Formation of reserve galactomannan in the seeds of Trigonella foenum-graecum. *Phytochemistry*, **9**, 513-520.

Reiter, W.D. and Vanzin, G.F. (2001) Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant Molecular Biology*, **47**, 95-113.

Reyes, F. and Orellana, A. (2008) Golgi transporters: opening the gate to cell wall polysaccharide biosynthesis. *Current Opinion in Plant Biology*, **11**, 244-251.

Richmond, T.A. and Somerville, C.R. (2000) The cellulose synthase superfamily. *Plant physiology*, **124**, 495-498.

Rossanese, O.W., Soderholm, J., Bevis, B.J., Sears, I.B., O'Connor, J., Williamson, E.K. and

Glick, B.S. (1999) Golgi Structure Correlates with Transitional Endoplasmic Reticulum Organization in Pichia pastoris and Saccharomyces cerevisiae. *The Journal of Cell Biology*, **145**, 69-81.

Sadlish, H., Pitonzo, D., Johnson, A.E. and Skach, W.R. (2005) Sequential triage of transmembrane segments by Sec61α during biogenesis of a native multispanning membrane protein. *Nature Structural & Molecular Biology*, **12**, 870-878.

Saint-Jore, C.M., Evins, J., Batoko, H., Brandizzi, F., Moore, I. and Hawes, C. (2002) Redistribution of membrane proteins between the Golgi apparatus and endoplasmic reticulum in plants is reversible and not dependent on cytoskeletal networks. *Plant Journal*, **29**, 661-678.

Sandhu, A.P.S., Randhawa, G.S. and Dhugga, K.S. (2009) Plant Cell Wall Matrix Polysaccharide Biosynthesis. *Molecular Plant*, **2**, 840-850.

Saxena, I.M., Brown Jr, R.M., Fevre, M., Geremia, R.A. and Henrissat, B. (1995) Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. *Journal of bacteriology*, **177**, 1419-1424.

Saxena, I.M., Brown, R.M. and others (2001) Structure-function characterization of cellulose synthase: relationship to other glycosyltransferases. *Phytochemistry*, **57**, 1135–1148.

Scheible, W.-R. and Pauly, M. (2004) Glycosyltransferases and cell wall biosynthesis: novel players and insights. *Current Opinion in Plant Biology*, **7**, 285-295.

Scheller, H.V. and Ulvskov, P. (2010) Hemicelluloses. *Annual Review of Plant Biology*, **61**, 263–289.

Schuldiner, S. (2009) EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics*, **1794**, 748–762.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W.B., Flügge, U.I. and Kunze, R. (2003) ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiology*, **131**, 16-26.

Seifert, G.J. (2004) Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. *Current Opinion in Plant Biology*, **7**, 277-284.

Seppälä, S., Slusky, J.S., Lloris-Garcerá, P., Rapp, M. and von Heijne, G. (2010) Control of membrane protein topology by a single C-terminal residue. *Science*, **328**, 1698-16700.

Sinnott, M.L. (1990) Catalytic mechanism of enzymic glycosyl transfer. *Chemical Reviews*, **90**, 1171–1202.

Snyder, G.E., Sakamoto, T., Hammer Iii, J.A., Sellers, J.R. and Selvin, P.R. (2004)

Nanometer localization of single green fluorescent proteins: evidence that myosin V walks hand-over-hand via telemark configuration. *Biophysical journal*, **87**, 1776-1783.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A.R., Persson, S., Raab, T., Vorwerk, S. and Youngs, H. (2004) Toward a Systems Approach to Understanding Plant Cell Walls. *Science*, **306**, 2206-2211.

Sterling, J.D., Atmodjo, M.A., Inwood, S.E., Kumar Kolli, V.S., Quigley, H.F., Hahn, M.G. and Mohnen, D. (2006) Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 5236-5241.

Suzuki, S., Li, L., Sun, Y.H. and Chiang, V.L. (2006) The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. *Plant Physiology*, **142**, 1233-1245.

Taiz, L. (1984) Plant Cell Expansion: Regulation of Cell Wall Mechanical Properties. *Annual Review of Plant Physiology*, **35**, 585-657.

Thompson, D.S. (2005) How do cell walls regulate plant growth? *Journal of experimental botany*, **56**, 2275-2285.

Tsien, R.Y. (1998) The green fluorescent protein. *Annual review of biochemistry*, **67**, 509-544.

Urbanowicz, B.R., Rayon, C. and Carpita, N.C. (2004) Topology of the Maize Mixed Linkage (1→ 3),(1→ 4)-β-D-Glucan Synthase at the Golgi Membrane. *Plant physiology*, **134**, 758-768.

van Meer, G. and Sprong, H. (2004) Membrane lipids and vesicular traffic. *Current opinion in Cell Biology*, **16**, 373-378.

Vargason, J.M., Szittya, G., Burgyán, J. and Hall, T.M.T. (2003) Size Selective Recognition of siRNA by an RNA Silencing Suppressor. *Cell*, **115**, 799-811.

Vicente, A.R., Saladié, M., Rose, J.K.C. and Labavitch, J.M. (2007) The linkage between cell wall metabolism and fruit softening: looking to the future. *Journal of the Science of Food and Agriculture*, **87**, 1435-1448.

Vincken, J.P., York, W.S., Beldman, G. and Voragen, A.G.J. (1997) Two General Branching Patterns of Xyloglucan, XXXG and XXGG. *Plant Physiology*, **114**, 9-13.

Viotti, C., Bubeck, J., Stierhof, Y.D., Krebs, M., Langhans, M., van den Berg, W., van Dongen, W., Richter, S., Geldner, N., Takano, J. and others (2010) Endocytic and secretory traffic in Arabidopsis merge in the trans-Golgi network/early endosome, an independent and highly dynamic organelle. *The Plant Cell Online*, **22**, 1344-1344.

Voinnet, O. (2001) RNA silencing as a plant immune system against viruses. *Trends in*

genetics : TIG, **17**, 449-459.

Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal*, **33**, 949–956.

von Heijne, G. (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature*, **341**, 456-458.

von Heijne, G. (1994) Membrane proteins: from sequence to structure. *Annual Review of Biophysics and Biomolecular Structure*, **23**, 167-192.

von Heijne, G. (2006) Membrane-protein topology. *Nature Reviews Molecular Cell Biology*, **7**, 909–918.

Walter, P. and Blobel, G. (1983) Preparation of microsomal membranes for cotranslational protein translocation. In *Methods in Enzymology* (Sidney, F. and Becca, F. eds): Academic Press, pp. 84-93.

Whistler, R.L. (1973) Solubility of Polysaccharides and Their Behavior in Solution. In *Carbohydrates in Solution*: AMERICAN CHEMICAL SOCIETY, pp. 242-255.

Wightman, R. and Turner, S. (2010) Trafficking of the plant cellulose synthase complex. *Plant Physiology*, **153**, 427-432.

Winter, H. and Huber, S.C. (2000) Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *CRC Crit Rev Plant Sci* **19**, 31-67.

Wulff, C., Norambuena, L. and Orellana, A. (2000) GDP-fucose uptake into the Golgi apparatus during xyloglucan biosynthesis requires the activity of a transporter-like protein other than the UDP-glucose transporter. *Plant Physiology*, **122**, 867-877.

Ye, K., Malinina, L. and Patel, D.J. (2003) Recognition of small interfering RNA by a viral suppressor of RNA silencing. *Nature*, **426**, 874-878.

Yin, Y., Huang, J. and Xu, Y. (2009) The cellulose synthase superfamily in fully sequenced plants and algae. *BMC plant biology*, **9**, 99-113.

York, W. and Oneill, M. (2008) Biochemical control of xylan biosynthesis — which end is up? *Current Opinion in Plant Biology*, **11**, 258-265.

Zablackis, E., York, W.S., Pauly, M., Hantus, S., Reiter, W.-D., Chapple, C.C.S., Albersheim, P. and Darvill, A. (1996) Substitution of L-Fucose by L-Galactose in Cell Walls of Arabidopsis mur1. *Science*, **272**, 1808-1810.

Zamyatnin, A.A., Solovyev, A.G., Bozhkov, P.V., Valkonen, J.P.T., Morozov, S.Y. and Savenkov, E.I. (2006) Assessment of the integral membrane protein topology in living

cells. *The Plant Journal*, **46**, 145-154.

Zeng, W. and Keegstra, K. (2008) AtCSLD2 is an integral Golgi membrane protein with its N-terminus facing the cytosol. *Planta*, **228**, 823-838.

Zhang, G.F. and Staehelin, L.A. (1992) Functional compartmentation of the Golgi apparatus of plant cells: immunocytochemical analysis of high-pressure frozen-and freeze-substituted sycamore maple suspension culture cells. *Plant Physiology*, **99**, 1070-1083.

Zhang, W., Bogdanov, M., Pi, J., Pittard, A.J. and Dowhan, W. (2003) Reversible topological organization within a polytopic membrane protein is governed by a change in membrane phospholipid composition. *Journal of Biological Chemistry*, **278**, 50128-50135.

Zhang, W., Campbell, H.A., King, S.C. and Dowhan, W. (2005) Phospholipids as determinants of membrane protein topology. Phosphatidylethanolamine is required for the proper topological organization of the gamma-aminobutyric acid permease (GabP) of Escherichia coli. *Journal of Biological Chemistry*, **280**, 26032-26038.

Zhong, R., Burk, D.H. and Ye, Z.H. (2001) Fibers: a model for studying cell differentiation, cell elongation, and cell wall biosynthesis. *Plant Physiology*, **126**, 477–479.

Zhu, Y., Nam, J., Carpita, N.C., Matthysse, A.G. and Gelvin, S.B. (2003) Agrobacterium-mediated root transformation is inhibited by mutation of an Arabidopsis cellulose synthase-like gene. *Plant Physiology*, **133**, 1000-1000.