Analyzing the complex machinery of cell wall biosynthesis

Jaap Timmers

Thesis committee

Thesis supervisor

Prof. dr. R.G.F. Visser
Professor of Plant Breeding
Wageningen University

Thesis co-supervisor

Dr.ir. L.M. Trindade
Assistant professor at the Laboratory of Plant Breeding
Wageningen University

Dr.ir. J.-P. Vincken
Assistant professor at the Laboratory of Food Chemistry
Wageningen University

Other members

Prof.dr.ir. A.G.J. Voragen	Wageningen University
Dr. H. Höfte	INRA Centre de Versailles, France
Dr.ir. G.H. Immink	Plant Research International
Prof.dr. M.E. Janson	Wageningen University

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.

Analyzing the complex machinery of cell wall biosynthesis

Jaap Timmers

Thesis

submitted in partial fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof.dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Doctorate Board
to be defended in public
on Friday 4 December 2009
at 11 AM in the Aula.

Timmers, J.F.P

Analyzing the complex machinery of cell wall biosynthesis

Ph.D. thesis Wageningen University, the Netherlands, 2009 with summary in Dutch
ISBN 978-90-8585-516-3

Contents

General introduction

Importance of plant cell walls

One of the more obvious distinctions between animal cells and plant cells is the plant cell wall, a layer of well-structured material located beyond the cell plasma membrane, a rigid but dynamic extra cellular matrix that surrounds each cell. This cell wall has the functions of providing strength, and defining the shape and size of a plant cell, ultimately, the plant cell wall functions as the determinant of plant morphology. In addition, plant cell walls are instrumental in controlling cell interactions within tissues and recogninizing pathogens, in the development of water- and ion-binding capacity of plant tissues, and in mechanisms of seed germination, fruit maturation and leaf abscission, and in stress responses (Carpita *et al.*, 2000). Furthermore the plant cell wall polysaccharides have great practical significance in the food, feed, and chemical industry, as well as in bio-energy production. There has recently been an upsurge of interest in the use of biofuels to replace fossil oil. This has been stimulated by a very rapid increase in the price of petroleum and concerns about global climate change. The link between climate change and biofuels is related to the fact that biofuels are a renewable source of energy. Energy from sunlight is collected by the photosynthetic system of plants and used to reduce and condense atmospheric CO_2 into the polysaccharides that comprise the cell wall of plants. These polysaccharides can be fermented into bio-ethanol or biogas, and used as transport fuel, or they can be burned to produce electricity and heat. The energy resulting from oxidation is released as heat, and the CO_2 is recycled into the atmosphere. Gaining genetic control of the amount, composition, and structure of cell walls in different cell types will impact both the quantity and yield of fermentable sugars from biomass for biofuel production, and the challenge is to achieve this target without compromising plant performance. The progress towards this goal is limited by the incomplete understanding of the cell wall polysaccharide biosynthesis and the genes involved.

The plant cell wall composition

Two types of cell walls can be distinguished in a plant cell; the primary wall, which is deposited during cell growth, and the secondary cell wall, which is deposited inside the primary wall at the onset of differentiation. The primary wall needs to be both mechanically stable and sufficiently extensible to permit cell expansion while avoiding the rupture of cells under their turgor pressure. The primary cell wall consists mainly of cellulose, hemicellulose, pectins and structural proteins. Hemicellulose is defined as wall polysaccharides that are not

solublized from wall materials with buffers, hot water, or chelating agents, but only with more or less strong chaotropic agents such as alkali (O'Neill and York, 2003). Pectins are structural complex polysaccharides, to date three classes of pectic polysaccharides have been characterized: homogalacturonans, rhamnogalacturonans, and substituted galacturonans. In the primary cell wall the cellulose microfibrils are cross-linked with hemicellulose to form a cellulose-hemicellulose network, and imbedded in a pectin matrix (Fig. 1). The much thicker secondary cell wall contributes to specialized functions related to a specific cell type, such as xylem fibers. The secondary wall is further strengthened by the incorporation of lignin, polymers of aromatic alcohols that covalently bind to hemicellulose (Vanholme *et al*., 2008).

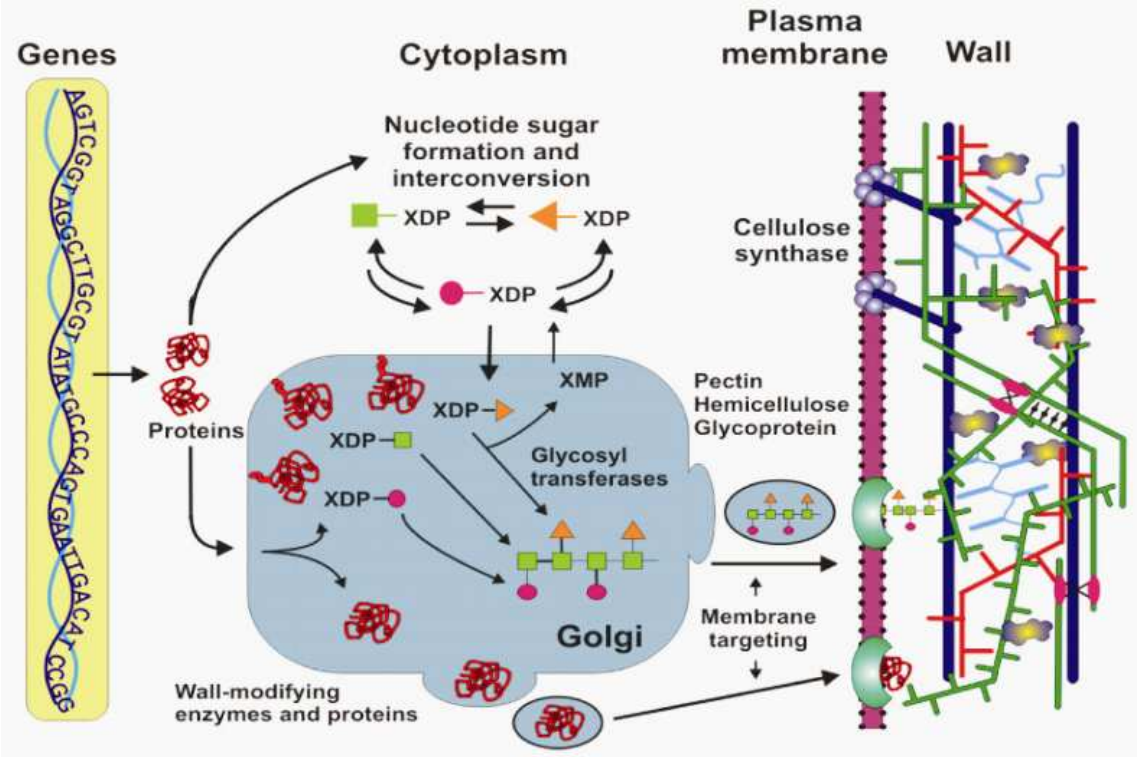


Figure 1. Model for the cell wall biosynthesis. Cellulose microfibrils (blue rods) are synthesized by large hexameric complexes in the plasma membrane, whereas hemicelluloses (red) and pectins (green), which compose the matrix polysaccharides, are synthesized in the Golgi apparatus and are deposited to the wall surface by vesicles.

Glycosyltransferases

The importance of the plant cell wall is revealed in the high number of genes likely to be involved in cell wall biogenesis, assembly, and modification. With the completion of the *Arabidopsis thaliana* genome sequence, it became clear that about 15% of the plant genes participate in cell wall formation and functioning (Carpita *et al.*, 2001; Carpita and McCann, 2002). The key enzymes that are directly engaged in polysaccharide assembly are the glycosyltransferases (GTs). GTs are enzymes that transfer the glycosyl moiety of activated

sugar residues to various acceptors and they are specific towards donor- and acceptor-substrates, as well as the bonds they form. The high specificity of these enzymes supports the abundance of the genes needed for the diverse glycoside bonds in the polysaccharides of the plant cell walls (Faik *et al*., 2002; Carpita *et al*., 2001). About 1.6% of the *Arabidopsis* genome is made up of genes that encode GTs (Egelund *et al.*, 2004). The GTs can be divided into two distinctive classes; type I and type II. Type I are processive enzymes that continuously incorporate glycosyl residues in the molecular backbones that constitute the main chains of wall polysaccharides. These polysaccharide synthases are believed to remain in close association with the nascent polysaccharide during its elongation, and are involved in the biosynthesis of homo-polysaccharides (Farrokhi *et al.*, 2006). These integral membrane proteins contain multiple transmembrane domains and have been proven difficult to purify by traditional biochemical methods (Delmer, 1999). The second class of GTs comprises enzymes that transfer glycosyl residues to a polysaccharide backbone. They have a single transmembrane domain (TMD), which functions as an anchor, an extended hydrophilic stem region, and a globular catalytic domain within the lumen of the Golgi (Perrin *et al*., 2001).

The CESA superfamily

A breakthrough in cell wall research was the identification of a type I glycosyltransferase family containing the cellulose synthases (CESA) proteins (Pear *et al*., 1996). Analysis of the *Arabidopsis* genome project indicated that this plant has ten cellulose synthases (CESAs) and thirty cellulose-like (CSL) genes (Richmond and Somerville, 2000). The catalytic region of these proteins was identified as sequence motif; D, D, D, QXXRW. The CSL genes are classified into the A, B, C, D, E, and G families and are proposed to encode processive glycosyltransferases that synthesize cell wall polysaccharides. A member of the CSLA subfamily was found to encode β-mannan synthase in guard seeds (Dhugga *et al*., 2004) whereas genes from respectively the CSLC and CSLF families encode β(1,4)-glucan and β-(1,3;1,4) glucan synthases (Cocuron *et al.*, 2007; Burton *et al*., 2006). The *Arabidopsis* CESA proteins range from 985 to 1088 amino acids in length and are integrated in the membrane by eight TMDs (Taylor, 2008). The cytosolic amino terminus contains a RING-type zinc finger though to be involved in protein-protein interaction (Saurin *et al*., 1996) (Fig. 2). The RING-finger domain of Cotton CESA protein was shown to homodimerise and interact with the RING-finger of another CESA protein (Kurek *et al*., 2002). The central cytosolic domain between the second and third TMD contains the catalytic site and is highly conserved between

all the CESA proteins, except for one region, the hyper variable region (Vergara and Carpita, 2001).

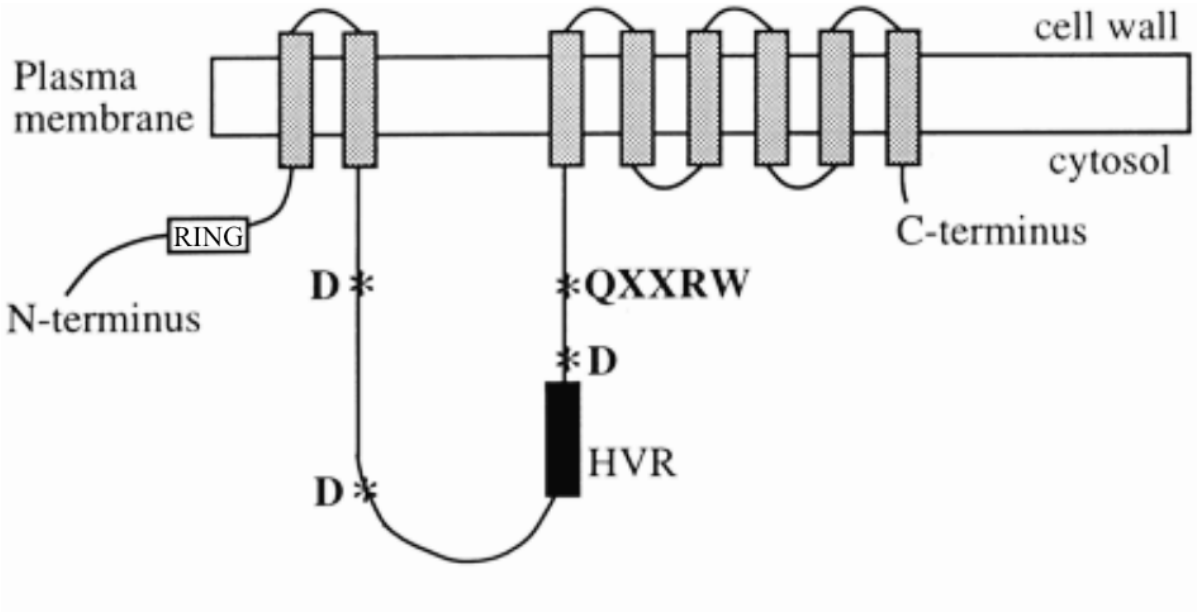


Figure 2. Model for the topology of a CESA protein. The CESA protein contains eight transmembrane domains and large central domain containing the QXXRW motif, believed to be important for substrate binding and catalysis, and the hyper variable region (HVR).

Structure of cellulose

Due to its abundance and importance, many efforts have been done to understand the structure and properties of cellulose and identify the enzymes involved in its biosynthesis. Unlike most known biopolymers, cellulose is a simple molecule of unbranched β(1,4)-glucan chains (Fig. 3A). Therefore cellulose is defined less by its primary structure, with β-1,4 linked glucosyl residues as repeating unit, but more by its secondary and higher-ordered structure.

Different forms of cellulose are defined by their crystallinic form. In crystalline cellulose, the glucan chains are arranged in a specific manner with respect to each other, no specific arrangement of the glucan chains occurs in non-crystalline or amorphous cellulose, nematic-ordered cellulose which is an intermediate form (Kondo *et al*., 2004). In plants cellulose is assembled in a structure referred to as a microfibril and the glucan chains are arranged in a parallel manner thought to be composed of thirty-six glucan chains (Taylor, 2008).

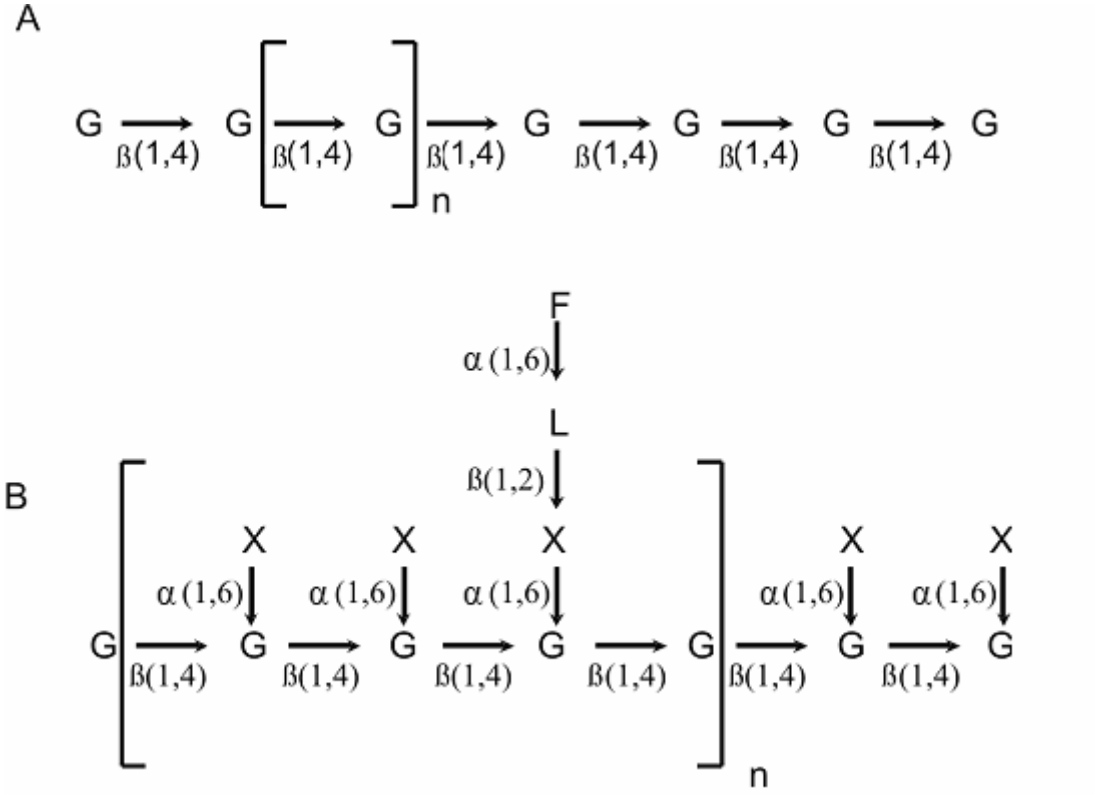


Figure 3. Chemical structure of cellulose and xyloglucan. A) The structure of cellulose, G indicates glucose, B) the structure of xyloglucan G indicates glucose, X xylose, L galactose and F fucose. Between the brackets (n) the repeating unit is indicated.

Cellulose biosynthesis

In higher plants cellulose is synthesized at the plasma membrane by rosette complexes (Fig. 1) (Somerville, 2006). The rosette structures are believed to contain thirty-six individual CESA proteins and microfibrils are believed to be composed of thirty-six glucan chains (Somerville *et al.*, 2004). The rosette complex is organized as a hexamer, as it consists of six subunits, probably resulting in synthesis of six glucan chains per subunit (Fig. 4). The rosette complexes are assumed to be assembled in the Golgi and then exported to the plasma membrane (Somerville, 2006). It was found that at least three different CESA proteins are required to form a functional rosette complex, CESA1, 3 and 6 and CESA4, 7 and 8 during, respectively, primary and secondary cell wall formation (Persson *et al*., 2007; Taylor *et al*., 2003). Mutant analysis revealed that in the primary cell wall synthesis CESA1 and CESA3 are essential, whereas CESA6 maybe functionally redundant by CESA2, CESA5 and CESA9, as null alleles of *CESA1* and *CESA3* proteins result in embryo-lethal phenotypes (Persson *et al*., 2007). Defects in the secondary cell wall shown by irregular xylem vessels that are unable to withstand the negative pressure generated during water transport through the xylem, the IRX phenotype. Mutations in each of the secondary cell wall CESAs results in such defects,

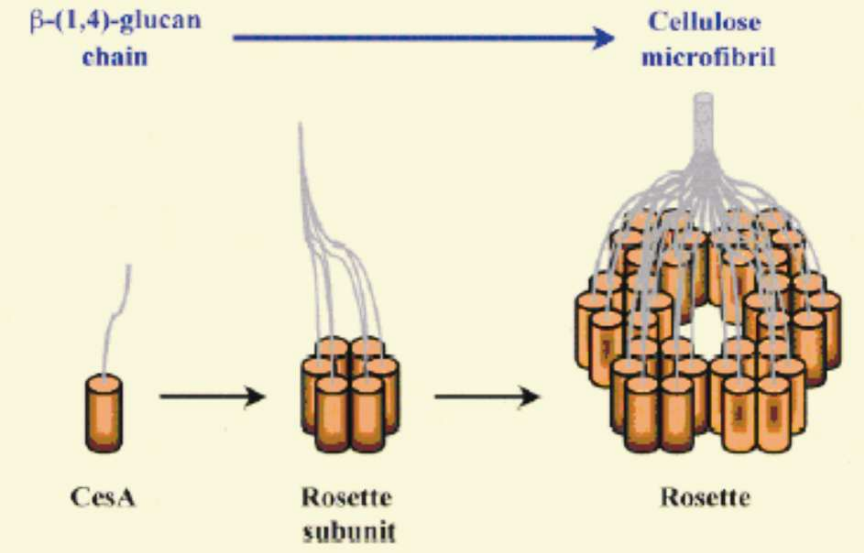


Figure 4. Model for the structure of the rosette. Six subunits, containing six CESA proteins, interact to form a rosette. Adopted from Doblin *et al*., 2002.

and it seems that all CESAs involved are equally important for cellulose deposition (Turner and Somerville, 1997). These *irx1*, *irx3* and *irx5* mutants, respectively *atcesa8*, *atcesa7*, and *atcesa4*, contain approximately 30% of the cellulose compared with wild type. Cellulose in the primary cell wall appeared unaffected in these mutants (Ha *et al*., 2002; Turner and Somerville, 1997). The identical phenotypes and similar reduction in cellulose content indicate that these genes are nonredundant, and that the enzymes are active in the same cell, at the same time (Taylor *et al*., 2003).

Apart from the CESA proteins, cellulose biosynthesis requires the participation of other proteins. Defects in the secondary cell wall characterized can be indicated by the IRX phenotype, therefore this phenotype was used to isolate *Arabidopsis* mutants defective in the biosynthesis of cellulose of the secondary cell wall (Turner and Somerville, 1997). Although this phenotype is indicative for any secondary cell wall mutations, it is not particularly suited to very large genetic screens. Mutants, which only exhibit slight distortions, are harder to discriminate and very severe wall defects may result in reduced viability, indicating that this screen may not identify all genes involved in proper secondary cell wall formation. Mutations in *KORRIGAN* result in reduced cellulose in both primary and secondary cell wall (Szyjanowicz *et al*., 2004). The *korrigan* mutant is deficient in an endo-β(1,4)-D-glucanase (Nicol *et al*., 1998) and although the exact role of the glucanase KOR1 in the synthesis of cellulose is not known, the mutant phenotype indicates that this enzyme is essential in cellulose biosynthesis (Sato *et al*., 2001). It has been suggested that KOR1 might be involved in cleavage of glucosylated sitosterol primers for cellulose synthesis, or in editing the growing

microfibrils to ensure proper packing of individual glucan chains, or in termination of the microfibril elongation (Peng *et al*., 2002; Roberts *et al*., 2004). Mutations in the *COBRA* locus affect the orientation of cell expansion and cause reduction in cellulose production (Schindelman *et al*., 2001); however; this reduction is thought to be due to the disordered deposition of cellulose microfibrils in which the *COBRA* genes are involved. Mutations in *KOBITO* result in cellulose-deficient, dwarfed mutants with a randomized cellulose microfibril orientation (Pagant *et al*., 2002). Mutations in an endo chitinase-like gene (CTL1) cause ectopic deposition of lignin and cell deformation in pith cells due to a decrease in cellulose (Zhong *et al*., 2002). The *kobito*, *knopf* and *botero1* mutants also show compromised cellulose content. Although these mutants have an effect on the cellulose synthesis, it is unknown whether they are directly linked to the rosette structure or even the cellulose machinery.

Structure of xyloglucan

The major hemicellulose found in the primary cell wall of dicots and non-graminaceous monocots is xyloglucan, where it constitutes 10 to 25% of the cell wall dry weight in dicots (McNeil *et al.*, 1984; O'Neill and York, 2003). Xyloglucans have a β(1,4) glucan backbone, regularly decorated at C6 with α-linked xylose residues, which in turn may be substituted with galactosyl and fucosyl residues (Fig. 3B). This xyloglucan substitution is highly regular and conserved in the majority of primary cell wall xyloglucans (Vincken *et al*., 1997). It might act as a spacer preventing cellulose microfibrils from aggregating or as an adapter that enables cellulose to interface with other cell wall matrix components (Thompson, 2005). Models of the plant cell wall postulate that xyloglucan cross-link adjacent cellulose microfibrils to form a cellulose-xyloglucan network which constitutes the major load-bearing structure of the primary cell wall (O'Neill and York, 2003). However it is not essential for the load-bearing capacity of the expanding primary cell wall, as lack of xyloglucan (XyG) is not fatal (Cavalier *et al*., 2008). In addition to this structural role, XyG can be a food reserve in seed or may act as signal molecules (Thompson, 2005).

Xyloglucan biosynthesis

Considerable progress has been made in the identification of glycosyltransferases involved in the biosynthesis of XyG. As glycosyltransferases are highly specific for the substrate and the types of linkages they generate, it has been suggested that a distinct enzyme is required for each linkage that is created (Keegstra and Raikhel, 2001). To date, several of the enzymes

involved in XyG production have been identified and characterized in *Arabidopsis*. The xyloglucan fucosyltransferase (AtFUT1) has been identified in *Arabidopsis* (Perrin *et al*., 1999). *AtFUT1* was found to be a member of a multi-gene family that has ten members (Sarria *et al*., 2001). *AtFUT1* is expressed at similar levels in all plant organs, whereas the other putative fucosyltransferases display a more complex expression pattern. The lack of fucosylated XyG and the expression pattern suggest that the other family members encode glycosyltransferases that are involved in the fucosylation of other polysaccharides like rhamnogalacturonan-I, rhamnogalacturonan-II, arabiongalactan proteins and N-linked glycans (Reiter, 2002).

Another fucose-deficient mutant *mur3* led to the identification of a xyloglucan galactosyltransferase that specifically adds galactose onto the third xylosyl residue. The lack of fucose is caused by the missing galactose (Madson *et al*., 2003). It was found that galactosylation rather than fucosylation is important for the mechanical strength during hypocotyl growth in *Arabidopsis* (Peña *et al*., 2004). The *MUR3* gene is expressed in similar quantities throughout the plant (Li *et al*., 2004). The xyloglucan xylosyltransferases (XXT) have been identified on the basis of their amino acid sequence homology with a fenugreek α(1,5)-galactosyltransferase (Faik *et al*., 2002). In *Arabidopsis*, the family containing these proteins consists of seven proteins with three clades. Additional research demonstrated that XXT1, XXT2, and XXT5 are involved in xyloglucan biosynthesis based on their mutant phenotype and their ability to link xylose to cellohexaose (Cavalier *et al*., 2006; Cavalier and Keegstra, 2006; Zabotina *et al*., 2008). The CSLC4 proteins is though to processively produce the glucan backbone of xyloglucan, as there is a strong correlation between the expression of *CSLC4* and *XXT1* and it has been observed that combined expression in *P. pastoris* of the two genes results in a longer β-(1,4) glucan synthesis by CSLC4 (Cocuron *et al*., 2007).

Scope and outline of the thesis

In this thesis we study the synthesis of two plant cell wall polysaccharides in *Arabidopsis thaliana* with the intention to identify and characterize the protein complexes involved in cell wall polysaccharide biosynthesis. A new approach in cell wall research was chosen to obtain new insights in the cell wall biosynthesizing machinery. This strategy can provide us with valuable information about the interactions between known constituents of the complexes involved in cell wall biosynthesis as well as providing new components of these complexes.

In Chapter 2, the advantages and limitations of several protein interaction assays are discussed and a motivation for the strategy followed in this research program is provided. The membrane-based yeast two-hybrid (MbYTH) is described, the potential pitfalls are mentioned and the details of an improved protocol for the identification cell wall protein complexes are given.

In Chapter 3, the interactions between the different CESA proteins (#4, #7 and #8), components of the complex that synthesizes the cellulose in the secondary cell wall, have been tested using the MbYTH system. A model for the composition of rosette structure of the secondary cell wall of *Arabidopsis* is proposed, and the importance of the RING-finger domain on the interaction between different CESA proteins is tested.

The interactions between different CESA proteins and the cellulase KOR1 were tested to confirm its presence in the rosette structure, which is described in Chapter 4. This interaction was studied in more detail to identify the domain responsible for the interaction.

In Chapter 5, the MbYTH system was used to identify new candidates involved in cell wall biosynthesis. The different secondary CESA proteins were used in a library screen for interactors. Several criteria were used to discriminate between the proteins found in order discuss their relevance in cellulose biosynthesis.

In Chapter 6, the interactions between proteins which are known to be involved in the biosynthesis of xyloglucan are described. In a library screen several new interactors of the xylosyltransferase XXT1 have been identified and their biological relevance is discussed.

In Chapter 7, the results obtained in the previous chapters, the importance of the MbYTH system in cell wall research, and the insights into the biosynthesis of cellulose and xyloglucan are discussed.

References

Burton R.A., Wilson S.M., Hrmova M., Harvey A.J. Shirley N.J., Medhirst 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-glucans. Science 311: 1940-1942

Carpita N., and McCann M. (2000) The cell wall, biochemistry and molecular biology of plants. Plant Physiol 52-108

Carpita N., and McCann M. (2002) The functions of cell wall polysaccharides in composition and architecture revealed through mutations. Plant and Soil 247: 71-80

Carpita N., Tierney M., and Campbell M. (2001) Molecular biology of the plant cell wall: searching for genes that define structure, architecture, an dynamics, Plant Mol Biol 47: 1-5

Cavalier D.M., Lerouxel O., Neumetzier L., Yamauchi K., Reinecke A., Freshour G., Zabotina O.A., Hahn M.G., Burgert I., Pauly M., Raikehl N.V., and Keegstra, K. (2008) Disruption of two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. Plant Cell 20: 1519-1537

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

Cocuron J.C., Lerouxel O., Drakaki G., Alonso A.P., Liepman A.H., Keegstra K., Raikhel N., and Wilkerson C.G. (2007) A gene for the cellulose synthase-like C family encodes a β-1,4 glucan synthase. Proc Natl Acad Sci USA 104: 8550-8555

Delmer D. (1999) Cellulose biosynthesis: exciting times for a difficult field of study. Annu Rev Plant Physiol Plant Mol Biol 50: 245–276

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 supergene family. Science 303: 363–366

Egelund J., Skjot M., Geshi N., Ulvskov P., and Petersen B.L. (2004) A complementary bioinformatics approach to identify potential plant cell wall glycosyltransferase-encoding genes. Plant Physiol 136: 2609–2620

Faik A., Price N.J. Raikhel N.V., and Keegstra K. (2002) An *Arabidopsis* gene encoding an α-xylosyltransferase involved in xyloglucan biosynthesis. Proc Natl Acad Sci USA 99: 7797-7802

Farrokhi N., Burton R.A., Brownfiled L., Hrmova M., Wilson S.M., Bacic A., and Gincher G.B. (2006) Plant cell wall biosynthesis: genetic, biochemical and functional genomics approaches to the identification of key genes. Plant Biotech J 4: 145-167

Ha M.-A., MacKinnon I.M., Sturcova A., Apperley D.C., McCann M.C., Turner S.R., Jarvis M.C. (2002) Structure of cellulose-deficient secondary cell walls from the *irx3* mutant of *Arabidopsis thaliana*. Phytochemistry 61**:** 7–14

Kondo T., Kasai W., and Brown R.M. Jr. (2004) Formation of nematic ordered cellulose and chitin. Cellulose 11: 463-474

Kurek I., Kawago Y., Jacob-Wilk D., Doblin M., and Delmer D. (2002). Dimerization of cotton fibre cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. Proc Natl Acad Sci USA 99: 11109-11114

Madson M., Dunand, C., Li X., Verma R., Vanzin G.F., Caplan J., Shoue D.A., Carpita, and Reiter W.D. (2003) The MUR3 gene of *Arabidopsis* encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. Plant Cell 15: 1662-1970

Nicol F, His I., Jauneau A., Vernhettes S, Canut H., and Höfte H. (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis* EMBO J 17: 5563-5576

O'Neill M.A., and York W.S. (2003.) The composition and structure of plant primary cell walls. In The Plant Cell Wall, J.K.C. Rose, ed (Boca Raton, FL: CRC Press) pp. 1–54

Pagant S, Bichet A, Sugimoto K, Lerouxel O, Desprez T, McCann M, Lerouge P, Vernhettes S., Höfte H. (2002) *Kobito1* encodes a novel plasma membrane protein necessary for normal synthesis of cellulose during cell expansion in *Arabidopsis*. Plant Cell 14: 2001-2013.

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. Proc Natl Acad Sci USA 93: 12637-12642

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 Physiol 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.M., DeRocher, A.E., Bar-Peled M., Zeng W., Norambuena, L., Orellana, A., Raikehl, N.V., and Keegstra, K., (1999) Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis. Science 284: 1976-1979

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

Persson S., Paredez A., Carroll A., Palsdottir H., Doblin M., Poindexter P., Khitrov N., Auer M., and Somerville C.R. (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis.* Proc Natl Acad Sci USA 104: 15566-15571

Reiter W.-D. (2002) Biosynthesis and properties of the plant cell wall. Curr Opin Plant Biol 5: 536-542

Robert S., Mouille G., and Höfte H. (2004) The mechanism and regulation of cellulose synthesis in primary walls: lessons from cellulose-deficient *Arabidopsis* mutants. Cellulose 11: 351-364

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

Sarria, R., Wagner, T.A., O'Neill, M.A., Faik, A., Wilkerson, C.G., Keegstra, and K., Raikehl, N.V. (2001) Characterization of a family of *Arabidopsis* genes related to xyloglucan fucosyltranferase1. Plant Physiol 127: 1595-1606

Sato S., Kato T., Kakegawa K., Ishii T., Liu Y.-G., Awano T., Takabe K., Nishiyama Y., Kuga S., Sato S., Nakamura Y., Tabata S., and Shibata D. (2001) Role of the putative membrane-bound Endo-1,4-β-glucanase KORRIGAN in cell elongation and cellulose synthesis in *Arabidopsis thaliana.* Plant Cell Physiol 42: 251-263

Saurin A.J., Borden K.L., Boddy M.N., and Freemont P.S. 1996. Does this have a familiar ring? TIBS 21: 208–214

Schindelman G., Morikami A., Jung J., Baskin T.I., Carpita N.C., Derbyshire P., McCann M.C., and Benfey P.N. (2001) COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in *Arabidopsis*. Genes Dev 15: 1115-1127

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

Somerville C. (2006) Cellulose synthesis in higher plants. Annu Rev Cell Dev Biol 22: 53-78

Szyjanowicz P.M.J., McKinnon I., Taylor N.G., Gardiner J., Jarvis M.C., Turner S.R. (2004) The irregular *xylem 2* mutant is an allele of *KORRIGAN* that affects the secondary cell wall of *Arabidopsis thaliana.* Plant J 37: 730-740

Taylor N.G. (2008) Cellulose biosynthesis and deposition in higher plants. New Physiologist 178: 239-252

Taylor N.G., Howells R.M., Huttly A.K., Vickers K., and Turner S.R. (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci USA 100: 1450-1455

Turner S.R., and Somerville C. (1997) Collapsed xylem phenotype of *Arabidopsis* indentifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell 1: 689-701

Vanhome R., Morreel K., Ralph J., and Boerjan W. (2008) Lignin engineering. Curr Opin Plant Biol 11: 278-285

Vergara C.E., and Carpita N.C. (2001) β-D-glycan synthases and the CesA gene family: lessons to be learned from the mixed-linkage (1→3),(1→4)β-D-glucan synthase. Plant Mol Biol 47: 145–160

Vincken, J.-P., York, W.S., Beldman, G., and Voragen, A.G.J. (1997) Two general branching patterns of xyloglucan, XXXG and XXGG. Plant Physiol 114: 9-13

Zhong R., Kays S.J., Schroeder B.P., and Ye Z.H. (2002) Mutation of a Chitinase-like gene causes ectopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. Plant Cell 14: 165-179

Zabotina O.A., Ven van de W.T.G., Freshour G., Drakakaki G., Cavalier D., Mouille G., Hahn M.G., Keegstra K., and Raikhel N.V. (2008) *Arabidopsis* XXT5 gene encodes a putative α-1,6-xylosyltransferase that is involved in xyloglucan biosynthesis. Plant J 56: 101-115

Implementation and optimization of a system to characterize protein complexes involved in plant cell wall biosynthesis

Jaap Timmers[1,2,3], Jean-Paul Vincken[2,4], Richard GF Visser[2], and Luisa M. Trindade[2]

[1]*CRC-Wageningen, Wageningen University and Research Centre, 6703 HD Wageningen, The Netherlands,*

[2]*Wageningen UR Plant Breeding, Wageningen University and Research Centre, 6708 PD Wageningen, The Netherlands*

[3]*Graduate School Experimental Plant Sciences, Wageningen University, The Netherlands*

[4]*Present address: Laboratory of Food Chemistry, Wageningen University, P.O. Box 8129, 6700 EV Wageningen, The Netherlands*

Abstract

The analysis of protein interactions can provide insights into pathways and developmental programs, as well as give an indication about the function of the proteins present in the same complex. Several powerful tools are available to screen protein interactions, however each method has its benefits and drawbacks, which are discussed in this chapter. The membrane based yeast two hybrid (MbYTH) was chosen to characterize the protein complexes involved in cell wall biosynthesis because of it is high throughput method able to identify interactions between the membrane proteins. Several practical hurdles, which are described in this chapter, had to be overcome to optimize this system for efficient screening and validation of the found interactors.

Methods for the analysis of protein-protein interactions

Proteins are essential constituents of the cell, either as structural components, or as enzymes for catalysis of highly diverse chemical reactions. Physical associations between different proteins, protein-protein interactions (PPI), are an important aspect in all biological pathways. Interactions between proteins can serve diverse functions such as protection of proteins from their environment, facilitation of substrate channeling, or building supramolecular structures such as the cytoskeleton. The binding between two proteins is influenced by shape and chemistry of the binding surface, which is in turn determined by the amino acid composition and tertiary structure of the proteins. The identification of interactions between different proteins can provide further insights into protein functions and signaling pathways *in vivo*.
During the last decade, a wide variety of methods has been developed to detect, analyze, and quantify protein interactions, including yeast two-hybrid assays, fluorescence-based technologies, and the biochemical purification of tagged proteins followed by identification of associated proteins via mass spectrometry. Each method has its restrictions and advantages as described below.

Affinity purification approach

One strategy to identify protein-protein interactions is to isolate protein complexes from a living cell, followed by characterizing their constituents. Originally, the purification of proteins complexes relied on specific antibodies for a given protein of interest for which the interacting partners had to be determined. By fusion of a tag, or a protein domain, to the protein of interest, problems associated with antibody-based purifications, like specific conditions for binding specificity and post translational modifications, might be avoided. This tag can then be used for specific isolation of the protein of interest, together with its associated partners, after which the partners can be identified (Fig. 1A). As these purification steps are performed in a gentile manner, it is thought that the native complexes are recovered, which may subsequently be tested for their activities or used in structural analysis (Puig *et a*l., 2001). Sometimes the harsh conditions needed to isolate the proteins from the cell might cause proteins to dissociate from the complex. As the interaction is not tested *in vivo*, deviated condition might cause association and dissociation of proteins, resulting in false results. Affinity purification can identify multimeric protein assemblies, but does not indicate the precise physical associations within them. As several interacting partners can be found in one

screening, the method can be considered high-throughput, although it requires the availability of a specific antibody or the introduction of a tagged protein in a cell or organism.

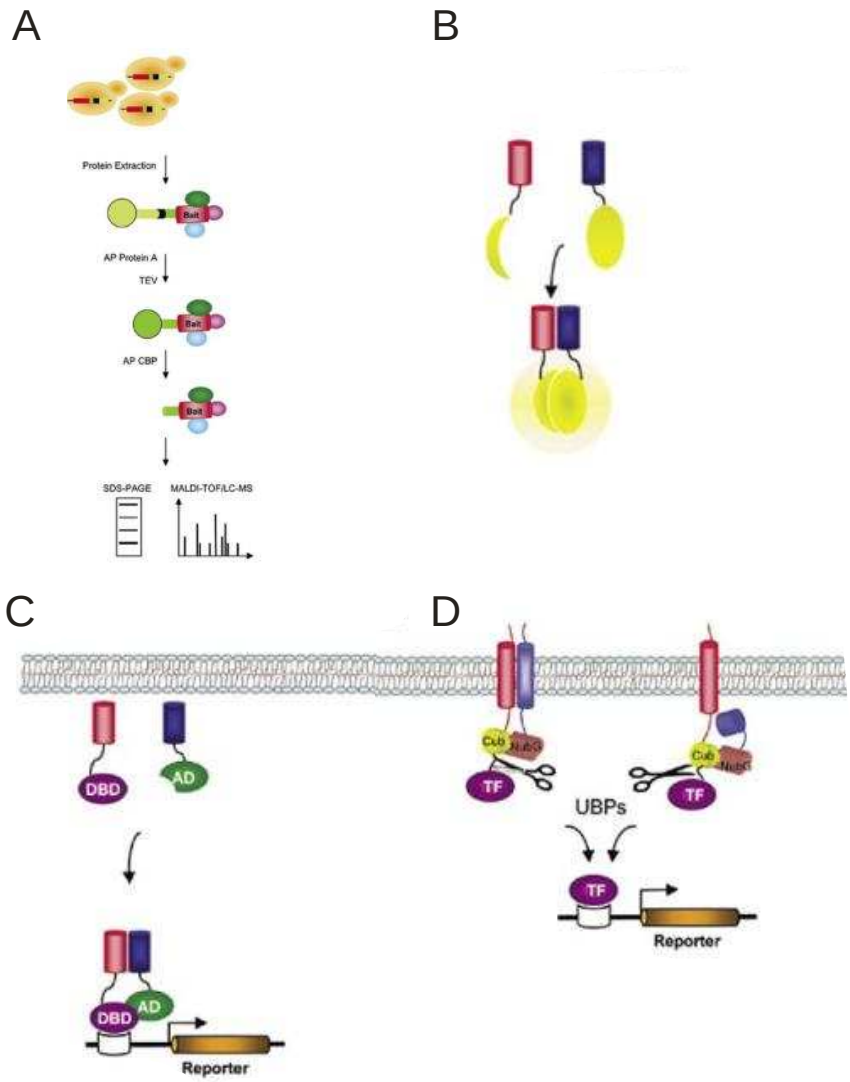


Figure 1. Systems for detection of protein interactions. (A) Tandem affinity procedure; The TAP sequence is fused to the coding sequence. Protein extracts from the organism expressing the TAP-tagged protein are subjected to two sequential affinity purification steps. Purified protein complexes are separated by denaturing gel electrophoresis and identified by mass spectrometry. (B) Bimolecular fluorescence complementation. Bait-prey interaction reconstitutes the YFP. Reconstitution can be detected by measuring fluorescence upon excitation at a suitable wavelength. (C) Classical yeast two-hybrid. A bait protein is fused to the DNA-binding domain of a transcriptional activator, whereas a prey protein is fused to the activation domain. The physical interaction between bait and prey reconstitutes the transcription factor, resulting in the expression of the reporter gene, thereby enabling the yeast cell to grow on selective medium. (D) Membrane based yeast two hybrid. The bait, membrane protein, is fused to the C-terminal half of ubiquitin, followed by a transcription factor, whereas a prey protein is fused to the N-terminal half of ubiquitin. Bait-prey interaction reconstitutes native ubiquitin, which is then cleaved by endogenous ubiquitin specific protease. The transcription factor enters the nucleus and activates reporter gene expression. Figure adapted and modified from Suter *et al*., 2008.

Fluorescence microscopy for analysis of protein-protein interactions

The use of fluorescent proteins (FPs) enables a variety of approaches to probe protein-protein interactions in living cells. A commonly used method for the co-localization of two labeled proteins is the bimolecular fluorescence complementation assay (BiFC). In this method the two proteins tested for interaction are each fused to a non-fluorescence half of a FP. The interaction between the two proteins of interest induces the complementation of the fluorescent protein by bringing the two dissected halves in each others vicinity, and restores fluorescence (Fig. 1B). This complementation is based on the inability of the two separate protein domains to reconstitute in the absence of the covalent linkage between the two subunits (Hu and Kerppola, 2003; Hu *et al.,* 2002). The reconstitution from YFP fragments in the BiFC system has a higher dynamic range with practically no fluorescence in the absence of an interaction to high levels of fluorescence after fusion to interacting proteins (Magliery *et al,*. 2005). Therefore, this method is, in principle, very sensitive. The possibility to study intact cells, *in vivo,* also avoids the risk of changes in protein interactions as a result of cell lysis and mixing of the contents of different cellular compartments. A disadvantage of the BiFC system is that protein folding of the fluorescent protein interferes with the dynamics of association-dissociation processes, as upon interaction the fragments are not able to release and stay fluorescent. Furthermore, this approach is not suitable for high throughput screening as it needs two fusion proteins expressed in a cell or organism.

Classical yeast two hybrid system

Another prominent technique for discovering protein interactions is the yeast two-hybrid system (YTH). As with the BiFC system the yeast two hybrid system is based on the concept that two separately expressed domains of a split protein cannot complement each other without an additional trigger for interaction. In the classical YTH system, this protein is a transcription factor that is split genetically into a DNA-binding and a transcription activation domain. The two domains are not able to reconstitute a functional transcription factor independently, they require the help of an interaction. The interaction of two fused proteins reconstitutes the activity of the split transcription factor, thus allowing the use of a simple growth selection in yeast to identify interactions (Fig. 1C).

This system is very cost effective, convenient to use, and easily adaptable for high throughput screening procedures making the YTH the most widely used method to assess both individual protein-protein interactions, in a one to one screening, and entire interactomes, which display the affinity of a bait against a cDNA prey library. Since the interactions between all bait and

prey proteins are confined in the nucleus of a lower eukaryote, PPI analysis by YTH suffers from a lack of contextual specificity. This also means that integral membrane and membrane-associated proteins are unsuitable for analysis by YTH due to their hydrophobic nature and non-nuclear localization. As nearly one-third of all genes encode proteins containing at least one TMD (Suter *et al*., 2008), this is an important limitation of the system. Another major drawback is the number of false positive interactions that occur upon self-activation of reporter genes by individual bait and prey proteins. A wide range of methods have been developed to improve the YTH system and made it more suitable for different purposes. The YTH is a high throughput method for the analysis of protein protein interactions and it allows the identification of new interactors by the screening of a cDNA library.

Membrane-based yeast two-hybrid system

The membrane-based yeast two-hybrid (MbYTH) system was developed to overcome the limitations of the classical YTH as it allows the identification of PPIs of membrane bound bait proteins with either membrane or cytosolic preys. The MbYTH system is based on the reassembly of the N- and C-terminal halves (Nub and Cub, respectively) of the ubiquitin protein when the fused proteins are able to interact. The reassembled quasi-native ubiquitin is recognized by ubiquitin-specific proteases, which then cleaves the C-terminally attached transcription factor (TF) from Cub and thereby provide an immediate readout of the Nub-Cub re-association. NubG (Ile-13-Gly), a mutant of Nub, has very low intrinsic affinity for Cub and therefore can interact with Cub only if both ubiquitin fragments are linked to proteins that interact. The TF A-LexA-VP16 (PLV) is attached to the C-terminus of Cub. PLV is liberated upon cleavage from Cub and activates the *lacZ* and *HIS3* reporter genes, thereby providing a useful selection for interaction (Fig. 1D). In conclusion this system is able to detect PPI between membrane bound proteins by expression in yeast. Similar to the classical yeast two hybrid system this is a high throughput system. One disadvantage of this system is that the interaction occurs in yeast where the environment might be different from the one *in planta*. Another limitation is that the system detects one to one interactions, whenever an additional protein is needed, as intermediate, the interaction will not be detected.

MbYTH assay as preferred approach to identify candidate genes involved in cell wall biosynthesis

The methods described here differ in their sensitivity and specificity; thus selection of a suitable method is crucial for a given investigation. The characterization of the organization of

the cellulose synthesizing rosette complex will allow a better understanding of cellulose biosynthesis. For this a method is needed which is able to determine the one to one interaction between the known constituents, whereas the identification of new components of the rosette complex requires a high throughput approach. In previous studies the BiFC assay has been used for the characterization of primary and secondary CESA complexes (Desprez *et al*., 2007; Timmers *et al*., 2009). Although this method generated valuable information about the composition of these complexes it is not feasible to identify new players involved in cellulose biosynthesis, as it entails the testing of each of the several thousand protein combinations *in planta.* The BiFC assay was consequently not chosen as primary method as the goal of this research was to identify new candidate genes involved in cell wall biosynthesis.

The affinity purification approach is potentially able to identify all proteins located in the complex together with the protein of interest. Using the affinity tag method, it has been possible to confirm the presence of three different CESA proteins in the primary and secondary rosette complex (Desprez *et al*., 2007; Atanassov *et al*., 2009); however it has not been possible to identify other components of the complex using this method. The yeast two hybrid systems are able to detect interactions in a high throughput way and therefore able to screen for new candidate genes. Although the classical yeast two hybrid was used to study the ability of CESA proteins to interact, it was only possible to test the interaction between partial proteins, as no membrane bound proteins can be tested in this system (Kurek *et al*., 2002). The applicably of this system in cell wall research is limited, as many proteins involved cell wall biosynthesis are membrane proteins, which cannot be screened in this assay. Therefore the MbYTH system was chosen to identify candidate genes involved in cell wall biosynthesis. The system enables to study the interaction between two proteins without the interference of other endogenous proteins. Two hybrid screens measure direct binary interactions, whereas immunoprecipitation-based methods and fluorescence methods measure the presence of a bait protein in a complex. The MbYTH approach is able to determine interactions between two known proteins as well as identifying unknown interactors, both soluble and insoluble.

Set up, implementation and practical considerations of the MbYTH method

The MbYTH system has proven to be a powerful tool to determine the interaction between membrane proteins (Peng *et al*., 2006). Although, the protocol of the MbYTH system is very straightforward, this method, as any other, has its pitfalls. Therefore the MbYTH system was optimized for identifying candidate genes involved in cell wall biosynthesis, first for one to

one protein interaction approaches, and later for the identification of new interactors by the screninf of a cDNA library.

Plasmid construction

The first critical step for MbYTH was the cloning of the membrane proteins with the right open reading frame (ORF) in the appropriate plasmid. A limited number of plasmids were available at the starting point of this research. The plasmids differed in promoter and fusion site (N- or C-terminal end) of the ubiquitin fragment. The protocol supplied by Dualsystems advices to use the strongest promoter to start the experiments, however, when thee is an over-expression of the fusion proteins, the use of an expression vectors with weaker promoters is recomended. An important feature to consider before cloning the open reading frames into the vectors is the topology of the proteins. In order to produce a read-out in the nucleus, the Cub-TF and NubG fusions both must be present at the cytosolic site of the membrane, the system does not allow for analysis of proteins in which both N- and C-termini are located inside an organelle or outside the cell. This can be achieved by choosing the appropriate plasmid so the ubiquitin part is either N- or C-terminally fused based upon the topology of the proteins.

In many cases the SfiI restriction enzyme can be used to clone the insert in the plasmid. SfiI has the recognition sequence GGCCNNN'NGGCC, with a variable core of five nucleotides (the cleavage site is indicated by '). There are two SfiI sites in the pADSL-Nx vectors: GGCCATTA'CGGCC and GGCCGCCT'CGGCC. The two SfiI sites have different overhangs which allow unidirectional cloning of cDNA in the vectors. Their long recognition sequence has to be preceded by eight nucleotides for an efficient digestion by the enzyme, therefore the primer for adding this restriction-site in front of a cDNA is very large which hampers the specificity of the PCR reaction. As the plasmids for the bait (pTFM) and prey (pADSL-Nx) contain different multiple cloning sites, it is not possible to ligate the same insert in these plasmids due to different restriction sites and possible frame shifts. The limited number of restriction sites in the multiple cloning sites of the plasmids might lead to a one restriction enzyme strategy or the use of blunt end restriction ligation strategy which both decrease the efficiency of the construction of a functional expression vector. These limitations can be solved by introduction of an identical multiple cloning site in both plasmids, which has been done in the newer versions of the system.

Membrane proteins can be toxic when expressed in *E. coli*, and the first step for the generation of Nub/Cub fusions is the cloning in *E. coli*. This toxicity dramatically decreases the efficiency of generating full-length functional proteins, as plasmids with a non functional

reading frame, due to errors and mutations will appear more frequently. Therefore, sequencing of the constructs is essential.

Determining the optimal screening conditions in a one to one interaction assay

When the bait and prey plasmids are constructed they can be transformed to the yeast strain. As the bait and prey plasmids both have a different selection marker in yeast, leucine and tryptophan respectively, growth on SD medium lacking the appropriate amino acids selects for the presence of the plasmid. Upon interaction of the bait and the prey the TF is released, which in turn activates the *HIS3* reporter gene allowing growth on medium lacking histidine. However, background cleavage of the bait protein can release trace amounts of TF, which causes selection leakage as it activates the *HIS3* reporter gene. Therefore 3-amniontriazole (3-AT) is added to the selection medium, which is a competitive inhibitor of the imidazole glycerolphophate dehydratase involved in histidine biosynthesis (Klopotowski and Waiter, 1965), to inhibit the basic leakage but not the strong activation of the histidine reporter gene. As the cleavage of the protein can be different for each protein expressed in yeast, the concentration is unique for each bait protein. To determine the concentration of 3-AT specific for the baits, a concentration range of 3-AT was added to selective medium. SD-medium lacking leucine, tryptophan (for selection on presence of respectively the bait and prey plasmids), and histidine (for selection on interaction). At the appropriate concentration of 3-AT interaction can initiate yeast growth, whereas in absence of an interaction there is no background growth, this can be tested with a positive and negative control.

The negative control can be a random protein not related to the bait, in these experiments the yeast protein ALG5 fused to NubG was used. Co-expression of the bait proteins with prey protein ALG5-NubG should not result in an interaction, and therefore not in activation of the system, as it is not involved in the pathways of interested. As a positive control the ALG5 protein was fused to the wild-type ubiquitin domain. In contrast to the I13G mutant (NubG), the wild-type N-terminal ubiquitin domain (NubI) can readily interact with the C-terminal ubiquitin domain. Thus the co-expression of the bait, containing the CUB, with a prey fused to the NubI will lead to an interaction and therefore may be used to test for bait expression and accessibility without the need for the fused proteins to interact. Growth of the yeast (expression of the positive control) on selective medium therefore not only indicates that the bait protein is expressed, but also that the Cub-TF is available for interaction. The use of two different auxotrophic markers increases the reliability of the system dramatically as two different pathways have to be circumvented to auto-activate the system. Therefore the yeast

strain DSY-5 (phenotype; MATalpha leu2 trp1 ura3-52 his3::GAL1-GAL4 pep4 prb1-1122), which was used in initial screens, was replaced by the strain NMY51 (phenotype; MATalpha his3delta200 trp1-901 leu2-3,112 ade2 LYS2::(lexAop)4-HIS3 ura3::(lexAop)8-lacZ (lexAop)8-ADE2 GAL4). This yeast strain allows autotrophic selection on both histidine and adenine, therefore increasing the reliability of the results. In addition, a colorimetric marker was used to confirm the interactions.

One to one interactions were tested by transforming yeast, containing the bait plasmid, with a prey plasmid and plate them on selective medium. To get an objective result the number of colonies grown was counted after five days at 30°C. Although there was a clear difference between a positive result and a negative one, usually a large variation was found between repeat experiments, due to transformation efficiency and colony density influenced the colony number. To obtain a repeatable readout of the system without disruptions by transformation efficiency and colony density a different method was applied in which the yeast containing the bait and prey plasmid was first grown on non selective plates. One hundred of the colonies then were picked, diluted and plated on selective plates to screen for growth. Results showed that this method gave reliable and repeatable numbers in a one to one interaction (Chapter 3).

Determining the optimal screening conditions in a library screen

In order to identify new candidate genes involved in cell wall biosynthesis the MbYTH system can also be used in a higher throughput method. This can be achieved by screening a cDNA library for interactors with the bait. The cDNA library consists of DNA copies of all mRNAs transcribed at a certain stage and in a specific organ of an organism inside the prey plasmid. When transformed into yeast, each yeast colony will express one protein, encoded by the cDNA, fused to the NubG. Transformation of the library into yeast containing the bait results in expression of the bait and a prey, which can be tested for interaction. In contrast to the one to one approach, this yeast is directly plated on selective medium (-leu, -trp, and –his) to eliminate colonies expressing a protein combination which does not interact. Using the parameters defined above, the bait can be used to screen a cDNA library of for instance *Arabidopsis thaliana*. The *A. thaliana* library (Dualsystems P02210) was constructed of six-day-old seedlings from a mixture of dark grown (etiolated seedlings) and seedlings exposed to blue and far red light. The library consists out of $1.7*10^7$ independent clones with an average insert size of 1.7 kb (ranges from 1.2-2.5kb). The Arabidopsis cDNA library was transformed into yeast containing a CESA as bait and grown for five days at 30° C on selective plates (-leu, -trp-, -his containing the appropriate amount of 3-AT). These colonies grown were

transferred to plates lacking adenine as well as histidine, leucine and tryptophan (-A-H-L-T with the appropriate amount of 3-AT), and therefore have a double auxotrophic selection for interaction. Activity of the *LacZ* gene (blue staining) is a third screening method for interaction. Only when the colony passes all three screens, the interaction between the bait and the prey was considered positive.

Identification of the interactor

Although the bait is known the prey can be any gene of the cDNA library. To identify which gene is expressed as prey the plasmid has to be sequenced. As the prey plasmid has a low copy number in yeast, it is difficult to obtain enough plasmid DNA for sequencing. To obtain the pADSL-Nx plasmids encoding the resulting positive library clones, total yeast DNA was isolated and transformed into *E. coli* strain XL-blue. *E.coli* colonies containing a pADSL-Nx library plasmid were obtained by ampicilin selection. pADSL-Nx DNA was purified from *E.coli*, as this DNA isolation step is much more efficient, and the plasmids were sequenced. The strategy to transfer the prey plasmid to *E.coli* and then sequence was very laborious. Therefore a different approach was optimized in which the insert in the prey plasmid is multiplied by PCR. The yeast is first washed with Milli Q water, and then boiled in 1M NaOH, to fracture the yeast cells and release the plasmid DNA. The NaOH is then neutralized by the buffer Tris at pH7. This mixture, containing the prey plasmid can be used as template for a PCR reaction and using plasmid specific primers the insert region can be amplified. The resulting PCR fragment can be sequenced using nested primers located at the 3'prime end of the forward primer to prevent artifacts. The obtained sequences were blasted against the NCBI database and the prey protein identified.

Recognition of false interactors

In the screening of a cDNA library in yeast, several combinations of two proteins are able to grow even though the bait and prey do not interact *in planta*. There are several reasons for these false positive results. One source of false positives in YTH screens is auto activators, which emerge during the course of the screen by spontaneous mutations allowing the yeast to grow on selective medium (Rual *et al.*, 2005). The use of two auxotrophic markers however greatly reduces the number of these false positives. In addition, false positives may arise from proteolysis of the fusion and release of the TF by unknown processes (e.g. the quality control mechanisms in the ER). This type of false positives can be eliminated by testing for reporter activity with a non related protein which should result in a negative signal. As it is known that

there are proteins which are able to auto activate the system, a list of known false positive proteins contains among others ATPases and Ubiquitin (Table I Dualsystems). Proteins known to cause auto activation of the system can be disregarded as candidate proteins, unless there are indications that they might play a role in the relevant pathway.

Table I. List of known auto activators of the MbYTH system. As indicated by the Dualsystems website.

Interactor	**Frequency**	**Comments**
ATPases	Frequent	Mostly vacuolar ATPases, may be connected with sorting of particular bait proteins to the vacuoles of yeast.
Proteolipid proteins	Frequent	
Ubiquitin	Frequent	Frequently isolated from X-NubG libraries, less frequently from NubG-X libraries. Confirmed false positive, interacts with the Cub portion of the bait via the wild type ubiquitin. Isolated sequences often encode partially truncated ubiquitin, thereby creating a wild type Nub (N-terminal part of ubiquitin) fused to the NubG portion.
ADE2	Frequent	Frequently isolated when screening on SD-trp-leu-ade. Encodes the mammalian homologue of the yeast ade2 protein, capable of complementing the ade2 deficiency in the NMY32 and NMY51 reporter strains. Can be avoided by screening on SD-trp-leu-his-ade instead of SD-trp-leu-ade medium.
Translocon components	Rare	May interact with the bait upon translocation through the membrane due to spatial proximity. This is not a true false positive, as it reflects a biologically relevant interaction.
Signal peptidases	Rare	May reflect an interaction of the signal peptidase with type I baits upon cleavage of the signal sequence peptide.
Cytochrome C	Rare	
SelenoproteinK	Rare	

Other protein combinations do not induce an activation of the system although the bait and prey interact *in vivo.* These false negatives may arise from low abundance of the bait and prey due to low expression or low stability of the fusion protein in yeast or to the lack of accessibility of the protease cleavage site. Additional false negative might arise due to toxicity of the proteins, the expression of two proteins might influence the growth rate of the yeast. The size and the number of transmembrane domains of both bait and prey increase the toxicity for the yeast and will decrease growth rate preventing them to be picked up in a library-screen. Due to the expression in yeast proteins can also be improperly folded due to the

conditions in yeast and post translational modifications which can hamper interaction. Due to the absence of yeast growth false negatives are difficult to detect.

Final remarks

Several methods are able to successfully identify interaction between proteins. All methods mentioned in this chapter have been used to study proteins involved in cell wall biosynthesis. The MbYTH system was chosen to identify candidate proteins involved in this process, as most proteins involved in cell wall biosynthesis are membrane bound and because there is a need for a high throughput system. Although this assay is a straight forward powerful tool, working with the membrane based yeast two hybrid system entailed several (practical) problems to identify interaction between proteins involved in cell wall biosynthesis. Several simple steps introduced by the author have improved the efficiency of this system.
Modification of the multiple cloning site might improve the efficiency of producing both the bait and prey plasmid. Also the ability to sequence the prey without additional transformation steps increases the speed of this system. The MbYTH will be valuable method to get more insight in plant cell wall biosynthesis, whereas the other described methods for protein interaction studies might confirm and compliment the results found.

References

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. J Biol Chem 284: 3833-3841

Delmer, D.P. (1999) CELLULOSE BIOSYNTHESIS: Exciting Times for A Difficult Field of Study. Annu Rev Plant Physiol Plant Mol Biol 50: 245-276

Hu D.C., Chinenov Y., and Kerppola T.K. (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell 9: 789-798

Hu D.C., and Kerppola T.K. (2003) Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. Nat Biotechnol 21: 539-545

Kurek I., Kaqagoe Y., Jacob-Wilk D., Doblin M., and Delmer D. (2002) Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. Proc Natl Acad Sci USA 99: 11109-11114

Magliery T.J., Wilson C.G.M., Pan W.L. Mishler D., Ghosh I., Hamilton A.D., and Regan L. (2005) Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. J Am Chem Soc 127: 146-157

Peng L., Ma J., Chi W., Guo J., Zhu S., Lu Q., Lu C., and Zhang L. (2006) LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in *Arabidopsis thaliana*. Plant Cell 18: 955-919

Puig O., Caspary F., Rigaut G., Rutz B., Bouveret E., Bargado-Nilsson E., Wilm M., and Séraphin B. (2001) The tandem affinity purification (Tap) method: A general procedure of protein complex purification. Methods 24: 218-229

Suter B., Kittanakom S., and Stagljar I. (2008) Interactive proteomics: what lies ahead? Biotechniques 44: 681-691

Interactions between membrane bound cellulose synthases involved in the synthesis of the secondary cell wall

Jaap Timmers[1,2,3], Samantha Vernhettes[4], Thierry Desprez[4], Jean-Paul Vincken[2,5], Richard G.F. Visser[2] and Luisa M. Trindade[2].

[1]*CRC-Wageningen, Wageningen University and Research Centre, 6703 HD Wageningen, The Netherlands,*

[2]*Wageningen UR Plant Breeding, Wageningen University and Research Centre, 6708 PD Wageningen, The Netherlands*

[3]*Graduate School Experimental Plant Sciences, Wageningen University, The Netherlands*

[4]*Laboratoire de biologie Cellulaire, Institute Jean-Pierre Bourgin, INRA, Rte de Saint Cyr, Versailles, 78026 cedex, France*

[5]*Present address: Laboratory of Food Chemistry, Wageningen University, 6700 EV Wageningen, The Netherlands*

Abstract

It has not been reported how the secondary CESA proteins are organized in the rosette structure. A membrane-based yeast two hybrid approach was used to analyze the interactions between the CESA proteins involved in secondary cell wall synthesis of Arabidopsis and confirmed *in planta* by Bimolecular Fluorescence Complementation assay. Results indicated that although all CESA proteins can interact with each other, only CESA4 is able to form homodimers. A model is proposed for the secondary rosette structure. The RING-motif proved not to be essential for the interaction between the CESA proteins.

Published in FEBS-Letters (2009) 583: 978-982

Introduction

Cellulose synthases (CESAs) are components of membrane-localized complexes (rosettes), and catalyze cellulose fibers elongation. Three CESA family members (#4, #7, and #8) have shown to be required for the formation of a rosette protein complex involved in secondary cell wall cellulose biosynthesis in Arabidopsis thaliana (Taylor *et al*., 2003), hereafter referred to as the secondary CESA proteins. The secondary CESAs are not functionally redundant and gene expression suggest that CESA4, CESA7 and CESA8 are the only CESAs involved in cellulose synthesis in the secondary cell wall (Brown *et al*., 2005). Immunoprecipitation experiments showed that these CESA proteins co-precipitate (Taylor *et al*., 2003; Atanassov *et al*., 2009). Although, this is a step towards the clarification of the CESA protein complex, the specific composition and structure of the rosette complex remain elusive.

All CESAs contain eight transmembrane domains (TMDs) and two putative N-terminal zinc-fingers. These zinc-fingers are thought to mediate protein–protein interactions between the CESAs (Saurin *et al*., 1996). However, the disrupted-rosette phenotype of the rsw1 mutant (V549A) (Arioli *et al*., 1998) and domain swapping experiments (Wang *et al*., 2006) suggest that also other regions of the CESA protein play a role in rosette assembly. The most accepted model of the rosette has been proposed by Scheible and co-workers (Scheible *et al*., 2001) in which the rosette structure has six symmetrically arranged subunits that in turn consist of six CESA proteins. However, there is no experimental evidence as to how the different CESA proteins are arranged within the complex or the subunits. To form such a regular structure, the interactions between the CESA proteins are expected to be highly specific. To get more insight into the different interactions, a method to detect one-to-one protein interactions of membrane-bound proteins is essential. The split-ubiquitin membrane-based yeast two-hybrid system (MbYTH) allows the screening for interaction between the different membrane-bound CESA isoforms in yeast (Fetchko and Stagljar, 2004). The bimolecular fluorescence complementation (BiFC) assay (Hu *et al*., 2002) was implemented to confirm the interactions in living plant cells. This technique provided evidence that the primary CESA proteins can interact in vivo, and therefore are present in the same complex (Desprez *et al*., 2007). In this report the possible interactions between the secondary CESA proteins is discussed, and a model for the rosette organization is proposed. Finally the role of the RING-finger motif in protein interaction is discussed.

Materials and methods

Membrane-based yeast two-hybrid (MbYTH) screen

Yeast strain NMY51 (Dualsystems Biotech AG) was transformed according to the protocol (DUALmembrane Kit 1). Interactions were quantified by 100 colonies spotted on SD medium (lacking Leucine, Tryptophan, Histidine and Adenine) containing the appropriate concentration of 3-ammonium-triazole (130mM, 10mM, and 75mM, for baits CESA4, CESA7, and CESA8, respectively) and grown at 30°C for five days, the number of spots grown was scored. Detection β-galactosidase activity was performed with the filter-lift assay (Breeden and Nasmyth, 1985). Experiments have been done twice to confirm results.

Constructs for the MbYTH system

The full-length cDNAs were obtained from the Riken Bioresource Center (Seki *et al*., 1998; Seki *et al*., 2002) ATCESA4 (RAFL15-30-K05), ATCESA7 (RAFL09-35-F05), and ATCESA8 (RAFL09-65-M12). Restriction sites were generated by PCR with primers as indicated in Supplementary data (Table SI). The resulting PCR-products were digested and ligated in the pTFB1 vector (Bait) and the pADSL-Nx vector (Prey) (Dualsystems Biotech AG). Bait and prey expression is regulated by the TEF1 and ADH1 promoter, respectively. The sequences of the inserts were confirmed by Sanger sequence analysis. Experiments have been done four times to confirm results.

Site directed mutagenesis

The QuikChange Multi site-Directed MutagenesisKit from Stratagene (200514) was used to introduce point mutations into the RING-motif of CESA7 using primers CESA7C37, CESA7C56, CESA7C64, and CesA7C79 (Table SI) to introduce mutations C37A, C56A, C64A, C79A, respectively.

Bimolecular fluorescence complementation screen

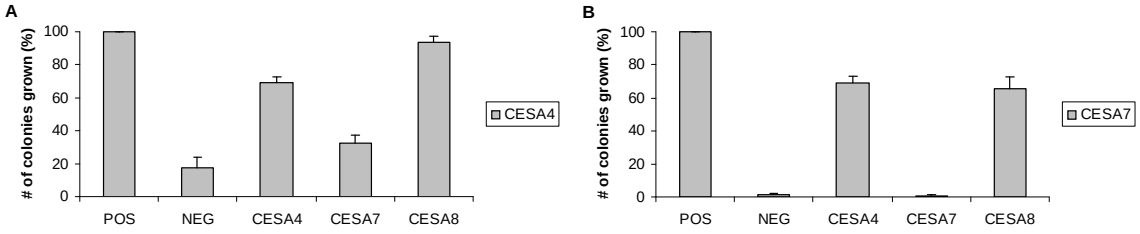
Genes were cloned in the pBIFP-2 and pBIFP-3 plasmids and regulated by the constitutive 35S promoter (Hu *et al*., 2002). The sequence of the primers used are in Table SI. Leaves of 3-week-old tobacco (*Nicotiana benthamiana*) plants were transformed by infiltration (Desprez *et al*., 2007). YFP (yellow fluorescent protein) fluorescence was detected 3 days after infiltration by using the 514-nm laser line of a SP2 AOBS confocal laser scanning microscope

(Leica, Solms, Germany) equipped with an argon laser. To check the YFP reconstitution, spectral analysis was performed with the 496-nm laser line.

Results

Interactions between the secondary CESA proteins

The regular structure of the rosette suggests that the assembly of this complex is highly regulated. In order to understand these complexes, the first step is the identification of specific interaction between the building-blocks of the complex, the different CESAs. The membrane-based yeast two-hybrid (MbYTH) method was used to identify the interactions between membrane-bound CESAs as it avoids the need to co-purify membrane proteins present in the same complex. In this system the protein of interest (bait) is fused to Cub-transcription factor (TF) and expressed in yeast together with another protein (prey) fused to NubG (Fetchko and Stagljar, 2004). Upon interaction between the bait and prey, the Cub-TF and NubG reconstitute and the TF is released by a protease so it can activate reporter gene expression. As the interaction is detected by a protease, the location of interaction is therefore not restricted to the nucleus but might also occur at the plasma membrane (Fetchko and Stagljar, 2004).The selection with two different auxotrophic markers increased the reliability of the system dramatically in that the prey had to circumvent two different pathways to auto-activate the system, as well as a colorimetric marker. The screening was optimized for each bait by addition of appropriate amounts of inhibitor (3-AT) to the selected medium so that growth of the yeast expressing a bait protein and the positive or negative control were significantly different, to rule out auto activation and to make it possible to screen for interactions between different proteins. All possible combinations of fusion proteins were grown on selective medium to determine their interactions. Fig. 1A shows the results of the interactions when CESA4 was used as bait, indicating strong interaction with itself and CESA8, and a weaker yet still significant interaction with CESA7.



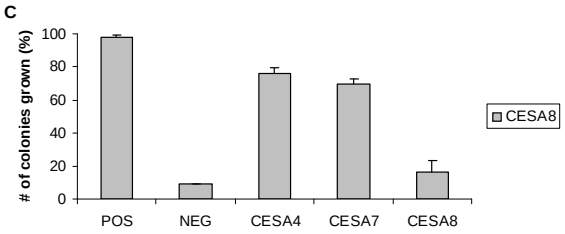


Figure 1. Interactions between the secondary CESA visualized by yeast growth. Yeast expressing CESA 4 (A), CESA7 (B) or CESA8 (C) as bait with the ALG5 protein fused to NubI and NubG, pos. and neg. control respectively, and different CESA proteins NubG fused proteins, as indicated. The percentage of colonies that show visible growth after 5 days at 30°C on selective medium is shown. Standard deviation is visualized by the error bar.

When CESA7 was used as the bait, strong interactions were detected with CESA4 and CESA8, however, CESA7 did not homodimerize (Fig. 1B). Similar results were obtained with CESA8 as a bait; CESA8 interacted with the other CESAs, but was unable to homodimerize (Fig. 1C).

Identification of CESA interactions in planta

The interactions were also tested in planta using BiFC assays. In this system a YFP fragment, either YFP/N or YPF/C, was linked to the N-terminal part of the secondary CESA proteins and transiently expressed in N. benthamiana. To determine whether heterodimers could be formed, two different CESA proteins were co-expressed YFP/N-CESA4/YFP/C-CESA7, YFP/N-CESA7/YFP/C-CESA4 (Fig. 2F), YFP/N-CESA4/YFP/C-CESA8 (Fig. 2G), YFP/N-CESA8/YFP/C-CESA4, YFP/N-CESA7/YFP/C-CESA8 (Fig. 2H), or YFP/N-CESA8/YFP/C-CESA7 (all interactions are shown in Fig. S2) To prevent false positives all fusion proteins were tested for interaction with the negative control, the aquaporin PIP2-1 protein, and all combinations showed no interaction. Although not all combinations were able to restore the YFP fluorescence, results indicated that all isoforms can interact with each other. Some combinations gave a weak signal, indicating that this dimerization is less efficient, particularly the combination of CESA7 and CESA8 even lacked fluorescence above the threshold in some of the repeat experiments. Also homodimerization of the CESA proteins was tested. A strong signal was found for the combination YFP/N-CESA4/YFP/C-CESA4 (Fig. 2B) whereas YFP/N-CESA7/YFP/C-CESA7 (Fig. 2C) was unable to restore YFP fluorescence, and YFP/N-CESA8/YFP/C-CESA8 (Fig. 2D) only gave a very weak fluorescence signal. Whenever CESA8 was expressed the signal was weaker and punctuate structures appear.

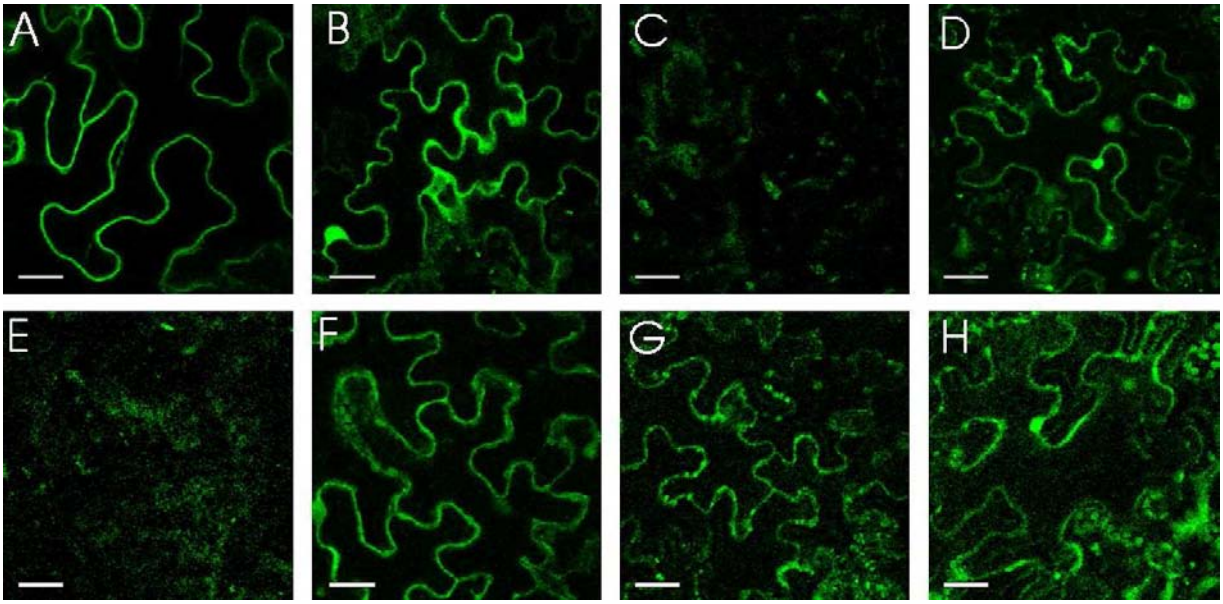


Figure 2. BiFC in *N. benthamiana* shows in vivo dimerization between the secondary CESAs. Positive controls PIP/PIP (A), negative control CESA7/PIP (E), homodimerization of CESA4 (B), CESA7 (C) and CESA8 (D) and the different heterodimerizations YFP/N-CESA7/YPF/C-CESA4 (F), YFP/N-CESA8/YFP/C-CESA4 (G), and YFP/N-CESA7/YFP/C-CESA8 (H). Scale bar = 100µm.

Role of RING-finger in CESA interactions

The N-terminal region of each CESA protein contains a double zinc-finger motif (CX2CX12FXACX2CX2PXCX2CXEX5GX3CX2C) highly homologous to the RING-finger domain. RING-fingers have been implicated in mediating protein-protein interactions, in a redox regulated bridging between cysteine residues (Saurin *et al*., 1996). Protein–protein interaction studies with only these zinc-finger domains, showed that they were able to interact with themselves and with the zinc-fingers of other family members (Kurek *et al*., 2002). In order to get more insight in the mechanism of interaction between the CESA proteins, the RING-finger motif of CESA7 was mutagenized and cysteines (C) at different positions were substituted by alanines (A) using site-directed mutagenesis. Also combinations of the substitutions were made (Fig. 3).

```
                          *  *                 *  *    *  *           *  *
CESA7Wt                   CEICGDQIGLTVEGDLFVACNECGFPACRPCYEYERREGTQNCPQC
CESA7C37A                 AEICGDQIGLTVEGDLFVACNECGFPACRPCYEYERREGTQNCPQC
CESA7C56A                 CEICGDQIGLTVEGDLFVAANECGFPACRPCYEYERREGTQNCPQC
CESA7C64A                 CEICGDQIGLTVEGDLFVACNECGFPAARPCYEYERREGTQNCPQC
CESA7C79A                 CEICGDQIGLTVEGDLFVACNECGFPACRPCYEYERREGTQNAPQC
CESA7C37A + C56A          AEICGDQIGLTVEGDLFVAANECGFPACRPCYEYERREGTQNCPQC
CESA7C37A + C64A          AEICGDQIGLTVEGDLFVACNECGFPAARPCYEYERREGTQNCPQC
CESA7C37A + C79A          AEICGDQIGLTVEGDLFVACNECGFPACRPCYEYERREGTQNAPQC
CESA7C56A + C64A          CEICGDQIGLTVEGDLFVAANECGFPAARPCYEYERREGTQNCPQC
CESA7C56A + C79A          CEICGDQIGLTVEGDLFVAANECGFPACRPCYEYERREGTQNAPQC
CESA7C64A + C79A          CEICGDQIGLTVEGDLFVACNECGFPAARPCYEYERREGTQNAPQC
```

Figure 3. Sequence alignment of the CESA7 RING-motif indicating the substitutions. The alignment of the RING-motif of CESA7 (AA38 until AA79). * indicate the essential cysteines, the substitutions (C→A) are highlighted (grey) in the different mutated proteins.

The interactions between the different mutated CESA7 proteins and the other secondary CESAs were tested with the MbYTH. The interaction between CESA8 and CESA7 decreased only slightly when one cysteine was mutated (C37A, C56A, C64A, or C79A). Double substitution affected the interaction more in only one combination (C37A+C79A), however the interaction was not abolished (Fig. 4). The interaction between CESA4 and CESA7 also decreased upon introduction of mutations (Fig. SI).

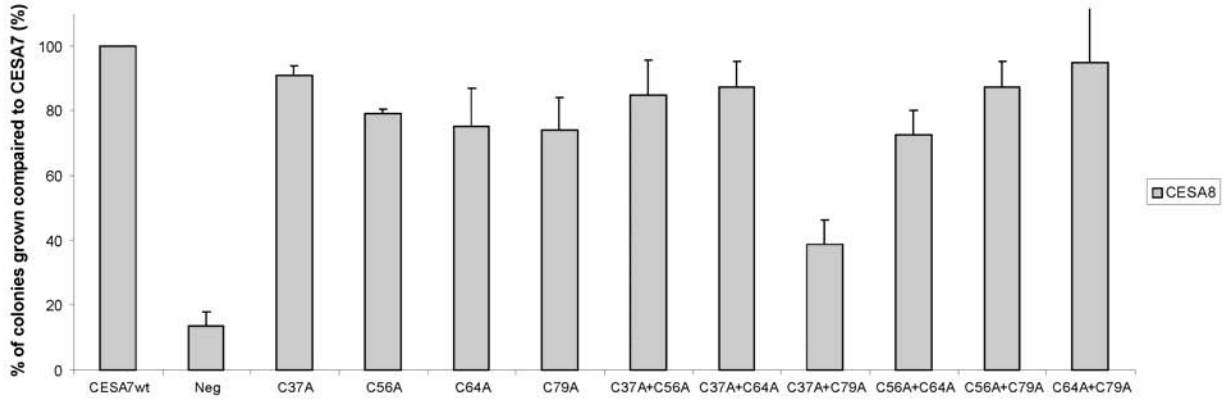


Figure 4. Effects of the substitution mutations on the interactions between CESA7 and CESA8. Yeast expressing CESA8 as bait with the wild type CESA7 (CESA7wt), NubG-ALG5 (Neg), and the different mutated CESA7 proteins as prey, that show visible growth after 5 days at 30°C on selective medium as a percentage of the interaction with wild type CESA7 protein.

Discussion

Different models for the structure of the CESA complex are possible, but the rosette structure model, proposed by Scheible and co-workers (Scheible *et al*., 2001) and modified by Doblin and co-workers (Doblin *et al*., 2002), is the most widely accepted. In this model three types of CESAs (α1, α2, and β) are assembled hexagonally in different proportions one α1, two α2, and three β isoforms (Fig. 5A). Three types of protein–protein interactions were proposed: α2–β and α1–β to form each subunit, and α2–α2 between subunits to form rosettes, therefore type-α1 can only bind type-β, type-α2 can bind types α2 and β, whereas type-β can bind to types α1 and α2. Although more complex variants of this model are also possible, less complex models seem unlikely, based on simple geometric considerations (Scheible *et al*., 2001). The interaction in yeast using the MbYTH method were confirmed with in planta

studies by the BiFC assays, except for the combination YFP/N-CESA7 and YFP/C-CESA8, which might be due to specific interference of the interaction-site of CESA7, and the ability of the CESA8 protein to form homodimers, there a very weak but detectable signal was observed in planta. The reasons for this discrepancy are not clear but it might be that this weak signal is not caused by a direct interaction but by a bridging endogenous protein of *N. benthamiana* that brings CESA8's in the vicinity of each other resulting in the assembly of the YFP. Furthermore punctuated structures were found in all CESA8 interactions indicate towards a different role for CESA8 compared to the others. In a recent publication pull down experiments indicate that in absence of one CESA the remaining isoforms form mono- and di-mers (Atanassov *et al*., 2009), which confirms the results described in this article. In a dual tagging assay a 240kDa band was found which might be assign to CESA7 homodimer. As the authors mention, it is difficult to find solid prove that this band is indeed a CESA7 homodimer as the three CESA proteins possess very similar molecular masses (Atanassov *et al*., 2009). In the MbYTH system expression of CESA7 alone is not sufficient for the formation of a dimer, a result confirmed with the BiFC system by heterologous expression in *N. benthamiana*.

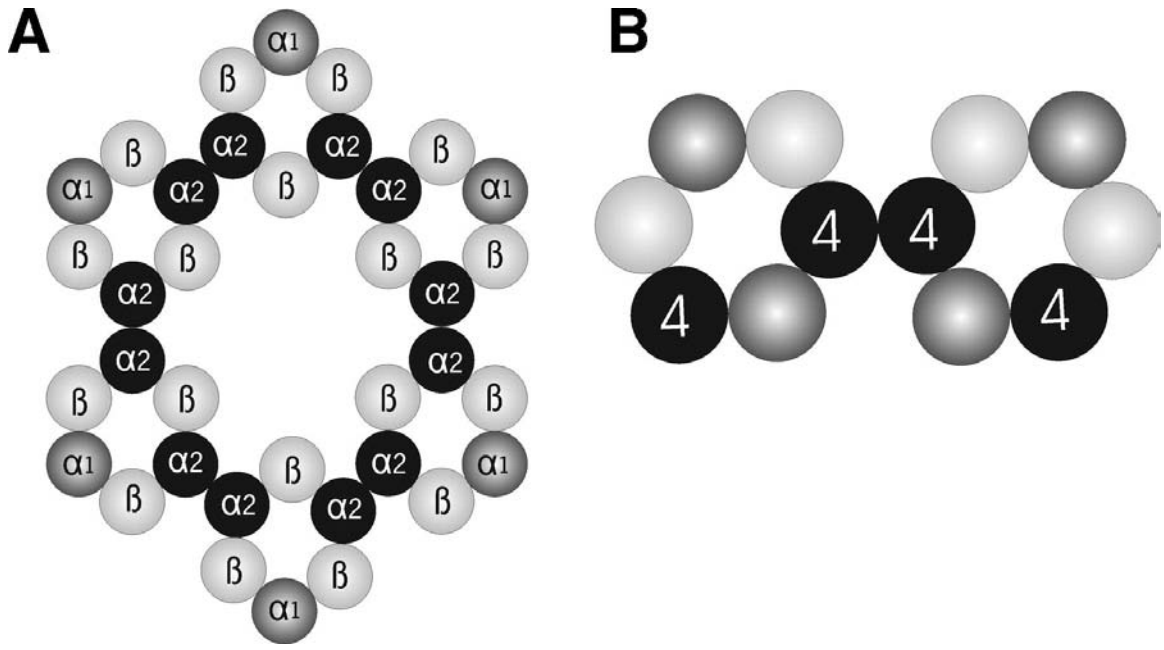


Figure 5. Proposed models for the structure of the rosette. (A) Six subunits, containing six CESA polypeptides, interact to form a rosette as suggested by Doblin and co-workers. (B) The modified model based upon the results described in this work between the different isoforms, CESA4 (4) interacts with all isoforms, the homodimerization links the subunits together, the two other position are filled by CESA7 and CESA8.

If our results are projected onto the rosette model of Scheible and co-workers, CESA4 self-interacts and therefore it has to occupy position α2. However CESA7(α1) and CESA8(β) both interact with the other two CESA of the rosette, but not with themselves, indicating that next to the three proposed types of protein–protein interactions (α2–β, α1–β, α2–α2) there is also interaction between α1 and α2. We adapted the model with these findings by replacing one β position by a α1 position (Fig. 5B). This results in a complex in which the subunits contain the same number of proteins of each of the isoforms and the homomeric interaction is responsible for the interaction between two subunits. Although the subunits themselves are not symmetric the overall complex is (Fig. 5B). The stoichiometry of the model is that each individual CESA protein is present in the same number, which is supported by the co-expression and regression score between the different secondary CESA genes (Persson *et al*., 2005). However to date this has not been confirmed at protein level. The specificity of the interaction suggests a non-random incorporation of CESA proteins into the rosettes, and might hint towards a specific function of each of the CESA proteins in the rosette.

The RING-finger motif was the best candidate to facilitate the interactions between the CESAs, as it has been shown that the RING-finger domains themselves can interact with each other when expressed separately (Kurek *et al*., 2002). It was found that a mutation in the RING-motif abolishes the interaction between two RING-fingers. When this motif is indeed essential for interaction between two CESA proteins, mutations in the RING-finger of CESA7 should abolish the interactions found between CESA7 and CESA4 and CESA8. Although some of the single mutations in the zinc-finger resulted in a decrease in interaction, the interaction itself was still observed and certainly not completely abolished. Combinations of cysteine substitutions did not result in a further decrease of the interaction (Fig. 4), suggesting that other domains than the RING-finger are involved in the interaction. Domain swapping experiments of Wang et al. (Wang *et al*., 2006), showed that the catalytic and/or C-terminal domains were the most important for entering the specific site in the complex, which is consistent with our results.

Our results do not exclude involvement of the RING-motif in protein interaction. The localization of the RING-motif at the cytoplasmic face of the plasma membrane suggests a role in recruiting proteins other than CESA. It has also been speculated that the RING-motifs under reduced conditions promote their own degradation (Voinnet *et al*., 2003). The introduced mutations within this motif reduce the zinc coordination to zinc-fingers, and may therefore lead to the degradation of the subsequent CESA protein. This might be a reason for decreased interaction between the mutated CESA7 proteins and the other CESA isoforms.

In conclusion, our screens have revealed that the CESA proteins involved in secondary cell wall synthesis specifically interact with each other. The interactions found result in an adapted version of the model for the rosette composition in which the homodimerization of CESA4 links the subunits to form the complete rosette. Interaction studies also indicated that the RING-motif is not essential for the interaction between different CESA proteins. More research will be required to understand the architecture of the rosette and the domains involved in the specific interaction between the different CESA within the complex.

Acknowledgments

The authors thank Dr M. Seki (RIKEN Genomic Science Center, Japan) for generously providing the cDNA of the CESAs. This work was supported by National Agency for Research Project “IMACEL” ANR-06-BLAN-0262.

References

Arioli T., Peng L., Betzner S., Burn J., Wittke W., Herth W., Camilleri C., Höfte H., Plazinski J., Birch R., Cork A., Glover J., Redmond J. and Williamson R.E. (1998) Molecular analysis of cellulose biosynthesis in Arabidopsis. Science 279: 717–720

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. J Biol Chem 284: 3833–3841

Breeden L., and Nasmyth K. (1985) Regulation of the yeast HO gene. Cold Spring Harb Symp Quant Biol 50: 643–650

Brown R.M., Zeef L.A.H., Ellis J., Goodacre R., and Turner S.R. (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17: 2281–2295

Desprez T., Juraniec M., Crowell E.F., Joury H., Pochylova Z., Parcy P., Höfte H., Gonneau M., and Vernhettes S. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci USA 104: 15572–15577

Doblin M.S., Kurek I., Jacob-Wilk D., and Delmer D.P. (2002) Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol 43:1407–1420

Fetchko M. and Stagljar I. (2004) Application of the split-ubiquitin membrane yeast two-hybrid system to investigate membrane protein interactions. Methods 32: 349–362

Hu C.D., Chinenov Y., and Kerppola T.K. (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell 9: 789–798

Kurek I., Kawago Y., Jacob-Wilk D., Doblin M., and Delmer D. (2002) Dimerization of cotton fibre cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. Proc Natl Acad Sci USA 99: 11109–11114

Persson S., Wei H., Milne J., Page G.P., and Somerville C.R. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci USA 102: 8633–8638

Saurin A.J., Borden K.L.B., Boddy M.N., and Freemont P.S. (1996) Does this have a familiar RING? TIBS 21: 208–214

Scheible W.R., Eshed R., Richmond T., Delmer D., and Somerville C. (2001) Modifications of cellulose synthase confer resistance to isoxaben and thiazoledinone herbicides in Arabidopsis lxr1 mutants. Proc Natl Acad Sci USA 98: 10079–10084

Seki M., Carninci P., Nishiyama Y., Hayashizaki Y., and Shinozaki K. (1998) High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated CAP trapper. Plant J 15: 707–720

Seki M., Narusaka M., Kamiya A., Oono Y., Muramatsu M., Hayashizaki Y., Kawai J., Carnici P., Itoh M., Ishii Y., Arakawa T., Shibata K., Shinagawa A., and Shinozaki K. (2002) Functional annotation of a full-length Arabidopsis cDNA collection. Science 296: 141–145

Taylor N.G., Howells R.M., Huttly A.K., Vickers K., and Turner S.R. (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci USA 100: 1450–1455

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. Plant J 33: 949–956

Wang J., Howles P.A., Cork A.H., Birch R.J., and Williamson R.E. (2006) Chimeric proteins suggest that the catalytic and/or C-terminal domains give CesA1 and CesA3 access to their specific sites in the cellulose synthase of primary walls. Plant Physiol 142: 685–695

KORRIGAN interacts specifically with the components of the cellulose production machinery

Jaap Timmers[1,2,3], Thierry Desprez[4], Nasim Mansoori[2,3], Jean-Paul Vincken[2,5], Richard GF Visser[2], Samantha Vernhettes[4], and Luisa M. Trindade[2]

[1]*CRC-Wageningen, Wageningen University and Research Centre, 6703 HD Wageningen, The Netherlands,*

[2]*Wageningen UR Plant Breeding, Wageningen University and Research Centre, 6708 PD Wageningen, The Netherlands*

[3]*Graduate School Experimental Plant Sciences, Wageningen University, The Netherlands*

[4]*Laboratoire de biologie Cellulaire, Institute Jean-Pierre Bourgin, INRA, Rte de Saint Cyr, Versailles, 78026 cedex, France*

[5]*Present address: Laboratory of Food Chemistry, Wageningen University, 6700 EV Wageningen, The Netherlands*

Abstract

Cellulose is synthesized by the so-called rosette complex, which comprises at least three different cellulose synthases (CESAs). Although several other proteins are known to be essential for cellulose biosynthesis, it has not been possible yet to link them physically to the rosette complex. One of these proteins is the cellulase KORRIGAN, and it has been hypothesized that this enzyme is an integral part of the rosettes, as mutations in this gene result in an altered cellulose content in both the primary and secondary cell wall. Using different methods, both *in vitro* and *in planta*, it was shown that KORRIGAN specifically interacted with CESA proteins. The interacting regions of korrigan were probed by analyzing the interaction of truncated forms of korrigan with CESA proteins. It was shown that the transmembrane domain of the korrigan is responsible for the interaction with the CESA proteins.

Introduction

Cellulose is a major component of the plant cell wall. In the primary cell wall, where it is considered a vital component of the load bearing network and an important determinant of the orientation of cell expansion. After a period of expansion some cell types lay down a thick cell wall layer in the inner side of the primary wall, the secondary cell wall. The cellulose microfibrils are synthesized by a multi-protein-complex at the plasma membrane, called rosette complexes, which consist of six globules. Each globule contains multiple cellulose synthase proteins (CESAs). Genetic analysis of cellulose-deficient mutants in plants resulted in the hypothesis that the rosette complex contains at least three different non-redundant cellulose synthases. CESA1, CESA3 and CESA6 were found to be involved in the cellulose biosynthesis in the primary cell wall, whereas CESA4, CESA7 and CESA8 are essential for the cellulose synthesis of the secondary cell wall. Hereafter, these proteins are referred to as the primary and secondary CESAs, respectively. Mutation in any of their genes causes defects in the assembly of cellulose microfibrils (Taylor *et al*., 2003).

On the basis of Arabidopsis mutant analysis, some proteins were predicted to be associated with the complex, although their interaction with the complex has never been demonstrated. One of these proteins is KORRIGAN (KOR1), a membrane-bound endo-1,4-β-D-glucanase. Enzymatic analysis of a recombinant soluble form of a KOR1 homologue from *Brassica napus* showed substrate specificity for low substituted carboxymethyl cellulose and amorphous cellulose, but no hydrolysis of crystalline cellulose or cellotriose to cellopentaose (Mølhøj *et al*., 2001). Mutations in the *KOR1* gene, like *rsw2*, have a clear effect on the primary wall, as they exhibit abnormal plant morphology, defects in cell wall formation, reduced cellulose content, increased pectin synthesis, and aberrant cell division features, also found in the *CESA1* mutant *rsw1* (Nicol *et al*., 1998; Peng *et al*., 2000; Zuo *et al*., 2000; Sato *et al*., 2001; Lane *et al*., 2001). Two independent mutations in the *KOR1* genes showed similar phenotypes as the *irx* mutants, i.e. collapsed xylem cell walls due to reduced cellulose synthesis in the secondary cell wall (Szyjanowicz *et al*., 2004). Based on the phenotype of the mutants, KOR1 appears to be involved in cellulose biosynthesis, but thus far it has not been possible to assign a specific role to this protein in this process. It has been suggested that KOR1 is involved in recycling sterol glucoside primers (Robert *et al*., 2004). KOR1 might also have a kind of proof-reading activity involved in hydrolyzing disordered amorphous cellulose to relieve stress generated during assembly of glucan chains in cellulose microfibrils (Mølhøj *et al*., 2002). Alternatively, KOR1 may determine the length of individual cellulose

chains during or subsequent to microfibril assembly or termination of the cellulose synthesis, releasing the cellulose microfibril from the synthase complex. Read and Bacic (2002) suggested that KOR1 might be an integral part of the rosettes. However, until now there is no experimental evidence for this, neither with co-precipitation experiments, nor with localization studies (Szyjanowicz *et al*., 2004). The membrane-based yeast two hybrid system (MbYTH; Fetchko and Stagljar, 2004) which has already been proven to be a powerful tool to determine interactions between glycosyltransferase involved in cell wall biosynthesis (Timmers *et al*., 2009), was used to determine the interaction between KORRIGAN and the different CESA proteins *in vitro*. The bimolecular fluorescence complementation method (BiFC) was used to confirm these results *in planta*. Our results show that the interaction between KOR1 and the CESA proteins is specific, and takes place in the membrane.

Material and Methods

Constructs for the MbYTH system

The contructs for the MbYTH system concerning the secondary CESAs were described previously (Timmers et al., 2009). The full-length cDNAs were obtained from the Riken Bioresource Center (Seki *et al*., 1998; Seki *et al*., 2002) *AtCESA1* (RAFL09-89-G08), *AtCESA3* (RAFL05-19-M03), *AtCESA6* (RAFL05-02-P19) and *KOR1* (RAFL05-02-G06). Restriction sites were generated by PCR with specific primers (primers for *AtCESA1*)
5’-AAGACTGCAGAATGGAGGCCAGTGCCGGC/
5’-AACAGGCGCCCTAAAAGACACCTCCTTTGCC
and 5’AAGAGGCCATTACGGCCATGGAGGCCAGTGCCGGC/
5’-AAGAGGCCGAGGCGGCCAAGTAAAAGACACCTCCTTTGCCAT (for bait and prey plasmid respectively); primers for *AtCESA3*
5’AGAACCATGGAATGGAATCCGAGGAGAAACC/
5’-AAGAACTAGTTCAACAGTTGATTCCACTTCC and
5’-AGAACGGCCATTACGGCCATGGAATCCGAAGGAGAAACC/
5’-GAGGCCGAGGCGGCCGTCAACAGTTGATTCCACATTCCAGAAT (for bait and prey plasmid respectively); primers for *AtCESA6* 5’-AGAACCATGGAATGAACACCGGTGGTCGG/
5’AAGAACTAGTTCACAAGCAGTCTAAACCA and
5’-AGAACGGCCATTACGGCCATGAACACCGGTGGTCGGTTAATCGC/
5’-GAGGCCGAGGCGGCCGTCACAAGCAGTCTAAACCACAGATCTCGAGAAT (for bait and prey plasmid respectively); primers for *KOR1*

5’-AAGACGTCATGTACGGAAGAGATCCATGGGG/
5’TTTACTAGTCAAGGTTTCCATGGTGCTGGTGG and
5’-AACAGGCCATTACGGCCATGTACGGAAGAGATCCATGGGG/
5’-AAGAGGCCGAGGCGGCCATCAAGGTTTCCATGGTGCTGGTGG (for bait and prey plasmid respectively); primers for *KOR1C*
5’-AAAGACGTCAAGATCTTCGTCTGGACTGTTGGT/
5’TTTACTAGTCAAGGTTTCCATGGTGCTGGTGG and
5’-AACAGGCCATTACGGCCAAGATCTTCGTCTGGACTGTTGGT/
5’-AAGAGGCCGAGGCGGCCATCAAGGTTTCCATGGTGCTGGTGG (for bait and prey plasmid respectively); primers for *KOR1N*
5’-AAGACGTCATGTACGGAAGAGATCCATGGGG/
5’TTTACTAGTTTAACGATCAAGGTAATGAA and
5’-AACAGGCCATTACGGCCATGTACGGAAGAGATCCATGGGG/
5’-AAGAGGCCGAGGCGGCCTTTAACGATCAAGGTAATGAA (for bait and prey plasmid respectively); primers for *KOR1TMD*
5’-AAAGACGTCAAGATCTTCGTCTGGACTGTTGGT/
5’-TTTACTAGTTTAACGATCAAGGTAATGAA and
5’-AACAGGCCATTACGGCCAAGATCTTCGTCTGGACTGTTGGT/
5’-AAGAGGCCGAGGCGGCCTTTAACGATCAAGGTAATGAA (for bait and prey plasmid respectively); primers for *KOR1SOL* 5’-AAGACGTCATGTACGGAAGAGATCCATGGGG/
5’-GCGGCTAACGATAATACAACC/5’-ACTGTGCCGCGTCATCATCC/
5’-TTTACTAGTCAAGGTTTCCATGGTGCTGGTGG and
5’-AACAGGCCATTACGGCCATGTACGGAAGAGATCCATGGGG/
5’-GCGGCTAACGATAATACAACC/5’-ACTGTGCCGCGTCATCATCC/
5’-AAGAGGCCGAGGCGGCCATCAAGGTTTCCATGGTGCTGGTGG (for bait- and prey-plasmid respectively). The resulting PCR-products were digested and ligated in the pTFB1 vector or the pADSL-Nx vector (bait and prey plasmid, respectively) (Dualsystems Biotech AG). The sequences of the inserts were confirmed by Sanger sequence analysis.

Membrane based yeast two hybrid screen

The bait and prey constructs were co-transformed into the yeast strain NMY51 (Dualsystems Biotech AG) according to the provided transformation procedure (DUAL membrane Kit 1). The yeast, containing both plasmids, was plated onto synthetic medium lacking leucine and tryptophan (SD med.-L-T), and grown at 30ºC for three days. To quantify the interactions between different preys 100 colonies of each combination were spotted onto selection

medium containing the appropriate amount of 3-ammonium-triazole (3-AT) and grown at 30ºC for three days. The number of spots grown was then counted. Detection of β-galactosidase activity was performed with the filter-lift assay (Breeden and Nasmyth, 1985).

Bimolecular Fluorescence Complementation screen

Leaves of 3-week-old plants (*Nicotiana benthamiana*) were transformed by infiltration as described (Voinnet et al., 2003). YFP fluorescence was detected 3 days after infiltration by using the 514-nm laser line of a SP2 AOBS confocal laser scanning microscope (Leica, Solms, Germany) equipped with an argon laser. To check the YFP reconstitution, spectral analysis was performed with the 496-nm laser line. The fluorescence with all constructs was detected at the same photo-multiplier tube (PMT) settings (760), except for the negative interactions for which the PMT was increased up to 880.

Results

Interaction between KORRIGAN and the primary cellulose synthases

In order to test the hypothesis that KORRIGAN is part of the primary cellulose synthase complex, the interaction between KOR1 and three members of the primary cell wall cellulose-synthesizing rosette (CESA1, CESA3 and CESA6) were tested using the MbYTH system. Interactions were tested with KOR1 fused to C-terminal part of the ubiquitin (Cub) and the transcription factor (bait), whereas the CESA1, 3 and 6 proteins were fused to the N-terminal part of the ubiquitin (Nub) (preys). Upon interaction between the bait and the prey the transcription factor (TF) is released into the nucleus where it activates reporter genes allowing the yeast to grow on selective medium. The results indicated that KOR1 is able to interact with all three of the CESA proteins as most yeast colonies were able to grow on selective medium (Fig. 1).

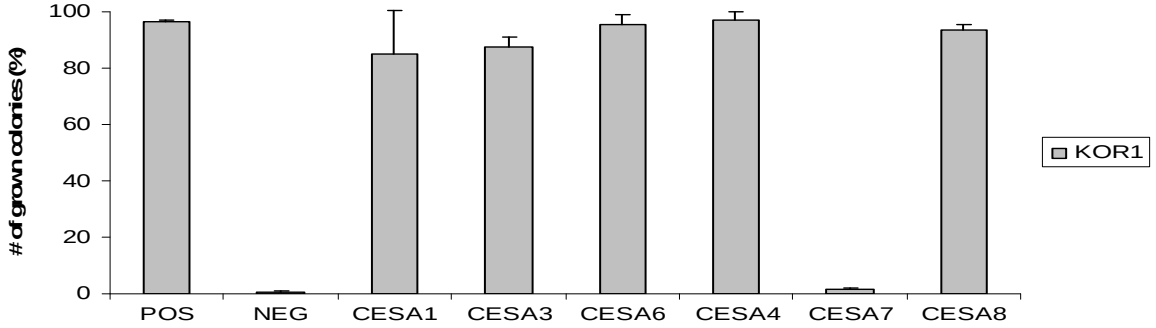


Figure 1. Interactions between KOR and the different CESA proteins using the Membrane-based Yeast Two Hybrid. The bars represent the percentage of yeast colonies grown for 3 days on selective medium at 30 ºC. KOR1 was expressed in yeast as bait and the different CESA proteins as prey (as indicated in the legend).

The lack of growth in the negative control indicated that the interaction with KOR1 was specific as an unrelated protein expressed as prey is not able to activate the system. Reverse experiments, in which the different CESA proteins were the bait and the KOR1 the prey, confirmed these findings in that all combinations were able to induce the reporter genes, allowing the yeast to grow on selective medium (Fig. 2). The results were confirmed *in planta* with the bimolecular fluorescence complementation method (BiFC). Using this system, the interaction between the primary CESAs and KOR1 were tested in *Nicotiana benthamiana*. Two YFP fragments, either YFP/N or YPF/C, each linked to the N-terminus of the proteins, were transiently expressed in the plant. Upon interaction between the two proteins, the fragments restore fluorescence, which can be detected. Expression of CESA1 together with KOR1 (both fused to YPF fragment) resulted in a strong fluorescent signal at the plasma membrane (Fig. 3A). The location of the interaction, at the plasma membrane, is an indication that the proteins were folded properly and not disrupted by the fluorescent tag as they were transported to the same sub-cellular localization as they are predicted to be located *in vivo*. These results indicated that the KOR1 protein is a physical component of the primary rosette complex, as it can bind to the major constituents (the CESA proteins) of the complex. Preliminary results have shown an interaction between KOR1 and CESA6 and the interaction between CESA3 and KOR1 still have to be determined *in planta* with BiFC.

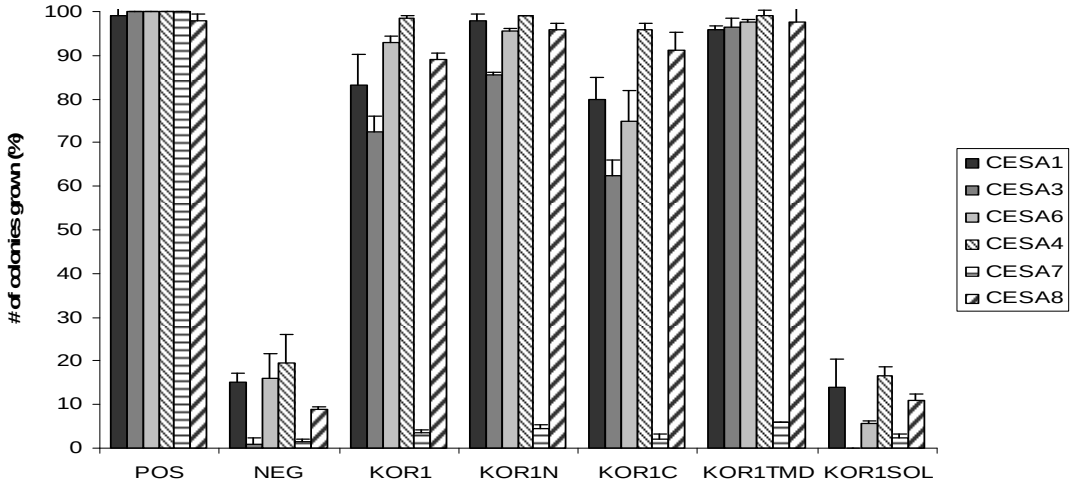


Figure 2. Interactions between the different KOR1 domains and the different CESA proteins using the Membrane-based Yeast Two Hybrid. The bars represent the percentage of yeast colonies grown for 3 days on selective medium at 30 ºC. The different CESA proteins were expressed in yeast as bait (as indicated in the legend) and the different KOR1 protein domains as prey.

KORRIGAN also interacts with the secondary CESA proteins

Reduction in cellulose content of the secondary cell wall in the KOR1 *irx* mutants links KOR1 also to cellulose synthesis in the secondary cell wall. To test interaction between KORRIGAN and the secondary CESA proteins (CESA4, CESA7 and CESA8) the KOR1 was expressed as bait (fused to the Cub and TF) in combination with the secondary CESA as prey (fused to the Nub). The combination KOR1 with CESA4 or CESA8 activated the reporter genes, and therefore able to grow on selective medium whereas no growth was detected when KOR1 and CESA7 were tested (Fig. 1). The interaction between KORRIGAN and CESA4 or KORRIGAN and CESA8 were comparable, however no significant interaction was detected for CESA7. The results were confirmed by the interaction between the CESA proteins (fused to the Cub and the transcription factor) and the KOR1 protein fused to the Nub which resulted in the same interactions (Fig. 2).

The physical interaction of KOR1 with the secondary CESA protein found with the MbYTH system was tested also *in planta.* Both KOR1 and the different secondary CESA proteins were fused to fragments of the YFP and expressed in *Nicotiana* leaves in different combinations. Restored fluorescence indicated interaction between the two fusion proteins. The combination CESA7 and KOR1 only showed a background fluorescent signal (Fig. 3E), whereas both CESA4 and CESA8 show a distinct signal when expressed in combination with KOR1 (Fig. 3D and 3F, respectively). Judging from the fluorescent signal, the interacting proteins were located at the plasma membrane, the location of the cellulose synthezing complex *in vivo*. The interactions found *in planta* by the BiFC system confirmed the results found with the yeast two hybrid system.

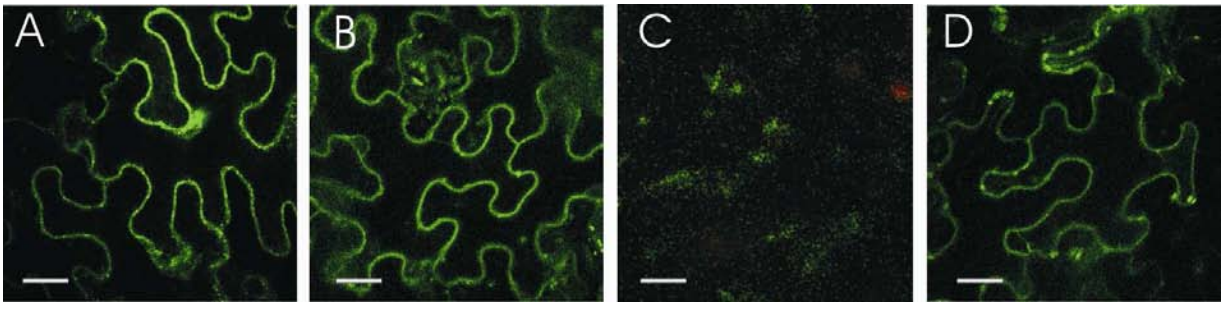


Figure 3. BiFC in *N. benthamiana* shows *in vivo* dimerization between KORRIGAN and various CESA proteins. The KOR1 expressed together with different CESA proteins; CESA1 (A), CESA4 (B), CESA7 (C), and CESA8 (D), respectively. The scale bar is 100µm.

The transmembrane domain is essential for interaction

The KOR1 protein is a membrane-anchored protein containing a short N-terminus located in the cytosol, a transmembrane domain (TMD), and an extra cellular catalytic (Fig. 4). To get more insight in the interaction between KOR1 and the CESA proteins several truncated KOR1 proteins were engineered. The N-terminal part of the protein together with the TMD was used to test whether the cytosolic portion of the protein (KORN; Amino Acids 1 to 94, Fig. 4B) is responsible for the interaction. The TMD with the C-terminal portion of the protein (KORC; AA 70 to 621, Fig. 4C) was tested for interactions between the catalytic domain and the CESA proteins. The TMD was also tested separately, by expressing only the TMD (KORTMD; AA 70 to 94, Fig. 4D). These truncated proteins were fused to the Nub and used as prey to test for interaction with the different CESA proteins. All the truncated proteins of KOR1 interacted with all CESA proteins, except for the CESA7 (Fig. 2). To confirm these results *in planta*, truncated forms of KOR1 were tested for interaction using the BiFC assay. The results indicated that the fluorescent signal was comparable with the full length KOR1 protein (Fig. 6A-C). As all tested KOR1 domains were able to interact with the same proteins as the full length KOR1, and they all contained the TMD, it was deduced that the TMD might be involved in these interactions. To test this hypothesis and to determine whether this domain is essential for the interaction another recombinant protein was made, of which the TMD was removed resulting in a soluble protein (KORSOL AA 1 to 69 + 95 to 621, Fig. 4E). Because there are other glycosyl hydrolase family 9 members, which lack the TMD it was assumed that KORSOL was folded correctly. No interaction between the KORSOL and any of the CESA proteins was found (Fig. 2). The reverse experiments confirmed these results and showed no interaction between the KORSOL (fused to the Cub and the transcription factor) and the CESA proteins (fused to the Nub).

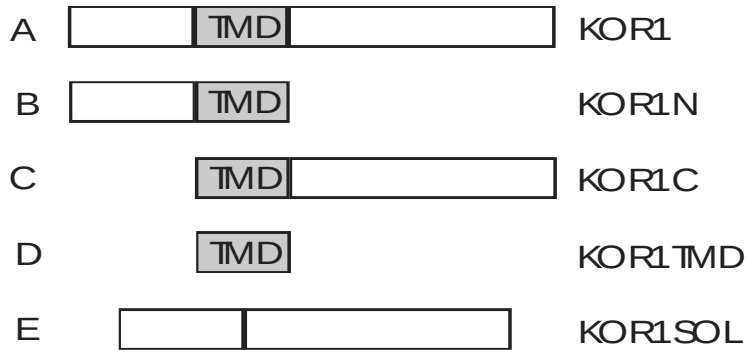


Figure 4. Representation of different truncated KORRIGAN proteins used in this study. A) KOR1 is the complete protein (AA 1 to 621), B) KOR1N: the N-terminal part with the TMD (AA 1 to 94), C) KOR1C: TMD plus the C-terminal part of KOR1 (AA 70 to 621), D) KOR1TMD only the TMD of KOR1 (AA 70 to 94), E) KOR1SOL is the KOR1 protein without the TMD; in which the N-terminus is fused directly to the C-terminal part (AA 1 to 69 + 95 to 621).

KORRIGAN is able to form a homodimer in planta

It is known that type II membrane proteins, like KOR1, often form dimers to perform their function. Using the MbYTH assay, KOR1 was expressed both as bait and as prey to test the ability of KOR1 to form homodimers. It appeared that KOR1 could form a homodimer, as the yeast was able to grow. To test whether a protein domain can be found responsible for this interaction the truncated KOR1 proteins were tested for interaction with each other. Growth was found in all combinations of truncated proteins except for those which lack a TMD. The combination of the KOR1C and the KOR1N, showed a weak yet significant interaction, whereas the reverse combination lacked detectable interaction (Fig. 5).

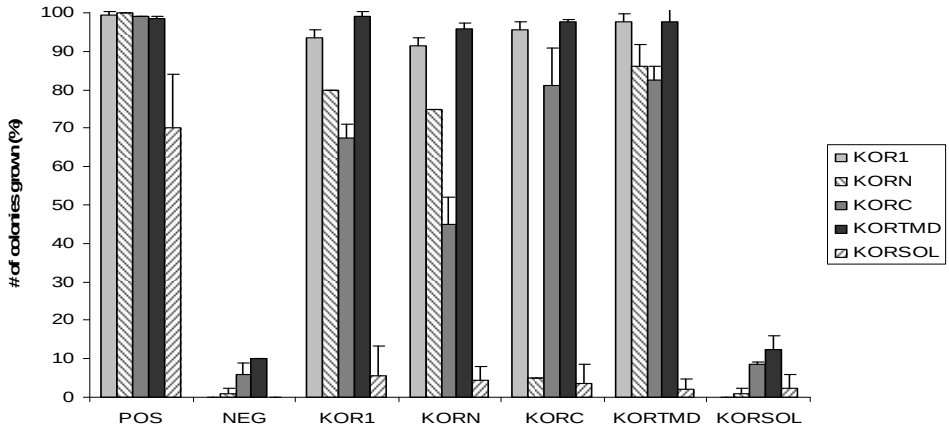


Figure 5. Interactions between the different KOR1 domains using the Membrane-based Yeast Two Hybrid. The bars represent the percentage of yeast colonies grown for 3 days on selective medium at 30 ºC. KOR1 domains represented in the legend were expressed in yeast as bait and the different KOR1 domains indicated at the x-axis as prey.

To confirm these findings the interactions were tested *in planta*. Our results showed that fluorescence was restored when two different fusion proteins (YFP/C-KOR1 and YFP/N-KOR1) were expressed in *Nicotiana*, indicating the formation of a homodimer of KOR1 proteins (Fig. 6A). Different parts of the protein [KOR1N or KOR1C (Fig. 4B and C)] were also tested for interaction with the full-length or portions of KOR1 protein and fluorescence was found in all combinations tested. The partial proteins KOR1N and KOR1C could interact with the full-length KOR1 protein (Fig. 6E and F, respectively) as well as with each other (Fig 6G) and form homodimers (Fig 6H and I, respectively), although the interaction between the KOR1N and KOR1C (Fig. 6G) was significantly weaker. As all (truncated) proteins in this test contained the TMD, these results indicate that this domain is also essential in the dimerization of KOR1.

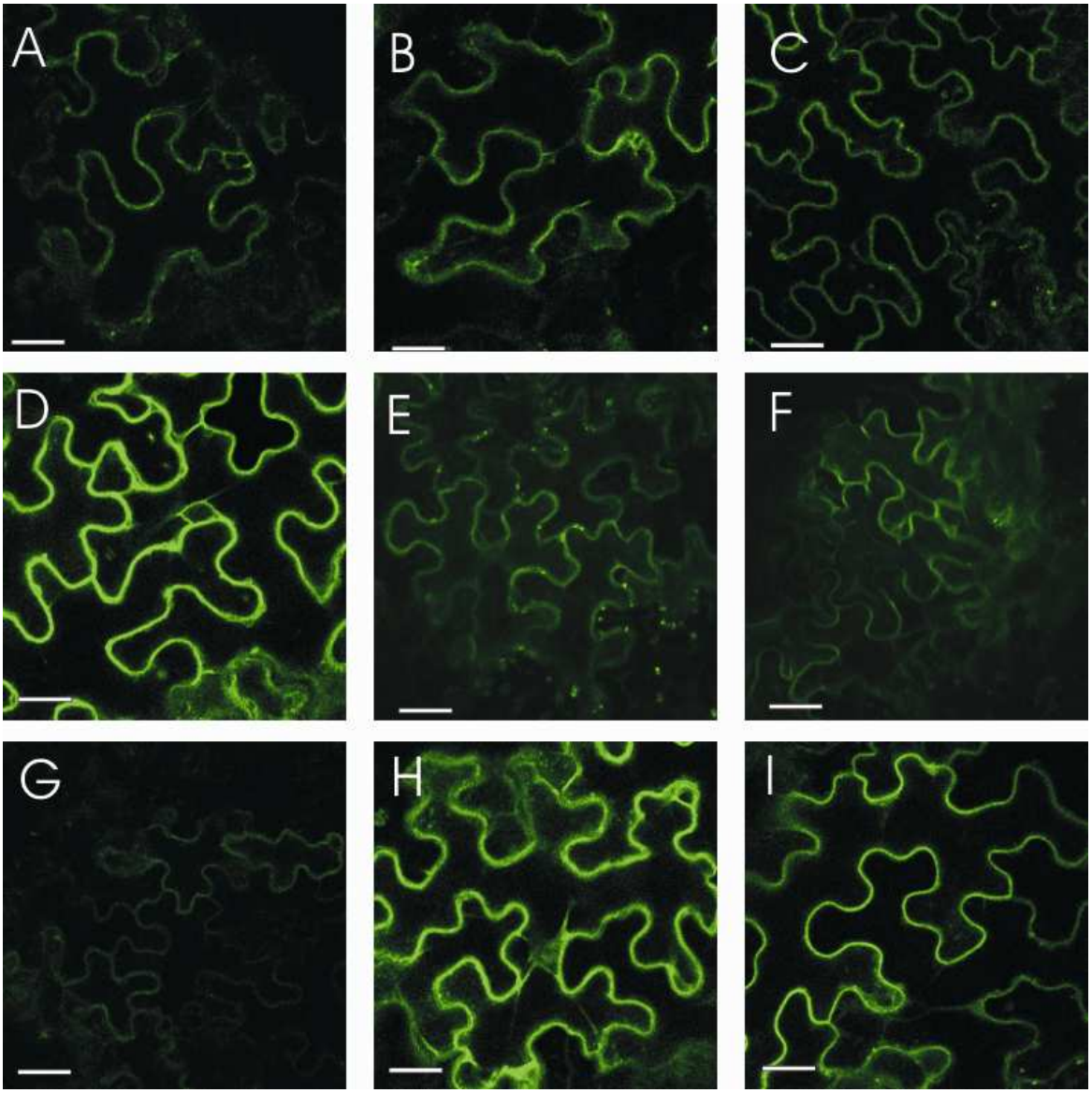


Figure 6. BiFC in *N. benthamiana* shows *in vivo* dimerization between KORRIGAN domains and the CESA1 protein. Dimerization between the different truncated forms of KORRIGAN KOR1/KOR1 (D), KOR1/KOR1N (B), KOR1/KOR1C (C), KOR1N/KOR1C (G), KOR1N/KOR1N (H), KOR1C/KOR1C (I) and the partial KOR1 proteins with the CESA1 protein KOR1/CESA1 (A), KOR1N/CESA1 (E), and KOR1C/CESA1 (F). The scale bar is 100μm.

Discussion

Our interaction studies presented in this study showed a physical interaction between the KOR1 and CESA proteins, indicating that the KOR1 protein is part of the rosette complex both *in vitro* in one-to-one interactions and *in planta*. Although KORRIGAN has shown to play an important role in cellulose biosynthesis, based on gene expression and mutant phenotype analysis, direct interaction between KOR1 and CESA proteins has never been proven due to experimental limitations (Szyjanowicz *et al*., 2004).

The result that KOR1 interacted with CESA1, 3 and 6 indicated that KOR1 is involved in the cellulose biosynthesis of the primary cell wall, whereas the interaction with CESA4 and CESA8 indicates that KOR1 is also part of the rosette structure of the secondary cell wall. The different reaction between the primary and the secondary cell wall proteins is difficult to explain as the specific functions of the different CESA proteins are not known. The binding of KORRIGAN with CESA1, 3, and 6 might be related to the lower degree of polymerization of the cellulose [8000 glucose residues in a chain (Brown, 2004)] in the primary cell wall compared to the secondary cell wall [14.000-15.000 glucose residues in a chain (Brett, 2000)], as korrigan has been implicated in severing the cellulose form the rosette structure (Taylor *et al*., 2008). The binding of KOR1 to the different CESA proteins is specific, as KORRIGAN does not bind to all of them. Not only does this imply that the methods used are sensitive enough to specifically determine interactions between these highly homologous proteins, it also indicates that the KOR1 protein has a specific position within the rosette.

The mutant phenotype of *kor1* results in a severe reduction in crystalline cellulose, both in the primary and secondary cell wall, underlining the importance of glucan trimming during cellulose biosynthesis.

A more detailed view on the interaction between KOR1 and the CESA proteins revealed that all portions which contain the TMD were able to interact and this led to the conclusion that the TMD was essential for the interaction. Therefore it is assumed that the interaction takes place within the membrane, between the TMD of KORRIGAN and the TMDs of the CESA proteins. This domain is also important in the homodimerisation of the KOR1 protein as all partial proteins containing this domain are able to interact, however other domains of the KOR1 protein might also play a role in the dimerization as the combination KOR1C and KOR1N only showed a weak interaction in both assays. The function of this dimerization is thus far unknown, however one could speculate that the dimer enables a more efficient hydrolysis of the glucan chains or binding to the cellulose, or the interlinking of KORRIGAN might result in a more stable rosette complex.

In conclusion, we have determined that the KOR1 protein interacts with both primary and secondary CESA proteins both *in vitro* and *in planta* and proved the hypothesis that KOR1 makes up a part of the rosette structure. The physical interaction also indicates that KOR1 is directly involved in cellulose biosynthesis, and probably does so in the form of a homodimer. Furthermore, our study showed that the TMD of KOR1 is essential in the interaction with the different CESA proteins.

References

Benziman M., Haigler C.H., Brown Jr.R.M., White A.R., and Cooper K.M. (1980) Cellulose biogenesis: Polymerization and crystallization are coupled processes in Acetobacter xylinum. Proc Natl Acad Sci USA 77: 6678-6682

Brett C.T. (2000) Cellulose microfibrils in plants: biosynthesis, deposition and integration into the cell wall. Int Rev Cytol 199: 161–199

Brown R.M. Jr. (2004) Cellulose structure and biosynthesis: what is in store for the 21st century? J Polym Sci Part A: Polym Chem 42: 487–495

Brummell D.A., Catala C., Lashbrook C.C., and Bennet A.B. (1997) A membrane-anchored E-type endo-1,4-β-glucanase is localized on Golgi and plasma membranes of higher plants. Proc Natl Acad Sci USA 94: 4794-4799

Desprez T., Juraniec M., Crowell E.F., Joury H., Pochylova Z., Parcy P., Höfte H., Gonneau M., and Vernhettes S. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci USA 104: 15572-15577

Fetchko M., and Stagljar I. (2004) Application of the split-ubiquitin membrane yeast two-hybrid system to investigate membrane protein interactions. Methods 32: 349-362

Haigler C.H., and Brown Jr.R.M. (1986) Transport of rosettes from the Golgi apparatus to the plasma membrane in isolated mesophyll cells in *Zinnia elegans* during differentiation to tracheary elements in suspension culture. Protoplasma 134: 111-120

Jzyjnowicz P.M.J., McKinnon I., Taylor N.G., Gardiner J., Jarvis M.C., and Turner S.R. (2004) The *irregular xylem 2* mutant is an allele of *KORRIGAN* that affects the secondary cell wall of *Arabidopsis thaliana*. Plant J 37: 730-740

Lane D.R., Wiedemeier A., Peng L., Höfte H., Vernhettes S., Desprez T., Hocart C.H., Birch R.J., Baskin T.I., Burn J.E., Arioli T., Betzner A.S., and Williamson R.E. (2001) Temperature-sensitive alleles of *rsw2* link the KORRIGAN endo-1,4-glucanase to cellulose synthesis and cytokinesis in *Arabidopsis*. Plant Physiol 126: 278-288

Mølhøj M., Ulvskov P., and Degan F.D. (2001) Characterization of a functional soluble form of a *Brassica napus* membrane-anchored Endo-1,4-β-glucanase heterologously expressed in *Pichia pastoris.* Plant physiol 127: 674-684

Mølhøj M., Pagant S., and Höfte H. (2002) Towards understanding the role of membrane-bound endo-beta-1,4-glucanases in cellulose biosynthesis. Plant Cell Physiol 43: 1399-1406

Nicol F., His I., Jauneau A., Vernhettes S., Canut H., and Höfte H. (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. EMBO J 17: 5563-5576

Peng L., Hocart C.H., Redmond J.W., and Williamson R.E. (2000) Fractionation of carbohydrates in Arabidopsis root cell walls shows that three radial swelling loci are specifically involved in cellulose production. Planta 211: 406-414

Read S.M., and Bacic T. (2002) Prime time for cellulose. Science 295: 59-60

Robert S., Mouille G., and Höfte H. (2004) The mechanism and regulation of cellulose synthesis in primary walls: lessons from cellulose-deficient *Arabidopsis* mutants. Cellulose 11: 351-364

Sato S., Kato T., Kakegawa K., Ishii T., Liu Y.G., Awano T., Takabe K., Nishiyama Y., Kuga S., Sato S., Nakamura Y., Tagata S., and Shibata D. (2001) Role of the putative membrane-bound endo-1,4-beta-glucanase KORRIGAN in cell elongation and cellulose synthesis in *Arabidopsis thaliana*. Plant Cell Physiol 41:251-263

Schneider D., Finger C., Prodöhl A., and Vokmer T. (2007) From interactions of single transmembrane helices to folding of α-helical membrane proteins: Analyzing transmembrane helix-helix interactions in Bacteria. Curr Prot Pept Sci 8: 45-61

Seki M., and Shinozaki K. 2009. Functional genomics using RIKEN Arabidopsis thaliana full-length cDNAs. J Plant Res 122(4): 355-366

Szyjanowicz P.M.J., McKinnon I., Taylor N.G., Gardiner J., Jarvis M.C., and Turner S.R. (2004) The *irregular xylem2* mutant is an allele of KORRIGAN that affects the secondary cell wall of *Arabidopsis thaliana*. Plant J 37:730-730

Taylor N.G., Howells R.M., Huttly A.K., Vickers K., and Turner S.R. (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci USA 100: 1450-1455

Taylor N.G. (2008) Cellulose biosynthesis and deposition in higher plants. New Physiol 178: 239-252

Timmers J., Vernhettes S., Desprez T., Vincken J.-P., Visser R.G.F., and Trindade L.M. (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett 583: 978-982

Zuo J., Niu Q.W., Nishizawa N., Wu Y., Kost B., and Chua N.H. (2000) KORRIGAN, an *Arabidopsis* endo-1,4-beta-glucanase, localized o the cell plate by polarized targeting and is essential for cytokinesis. Plant Cell 12:1137-1152

Identifying new components of the cellulose synthase rosette complex involved in cellulose biosynthesis in the secondary plant cell wall

Jaap Timmers[1,2,3], Jean-Paul Vincken[2,4], Richard GF Visser[2], and Luisa M. Trindade[2]

[1]*CRC-Wageningen, Wageningen University and Research Centre, 6703 HD Wageningen, The Netherlands,*

[2]*Wageningen UR Plant Breeding, Wageningen University and Research Centre, 6708 PD Wageningen, The Netherlands*

[3]*Graduate School Experimental Plant Sciences, Wageningen University, The Netherlands*

[4]*Present address: Laboratory of Food Chemistry, Wageningen University, P.O. Box8129, 6700 EV Wageningen, The Netherlands*

Abstract

Cellulose is one of the major components of the plant cell wall. This polysaccharide is synthesized at the plasma membrane by a rosette complex. Although the main constituents of this protein complex, the cellulose synthases (CESA), are well characterized, it has been difficult to identify other components of the rosette complex. In this study, a protein interaction approach was used to identify other members of the rosette complex. A library screening was performed with the membrane-based yeast two-hybrid system using each of the three different CESA proteins as bait (CESA 4, 7 and 8). This resulted in a list of approximately one hundred candidates, of which twelve showed interaction with at least two out of three CESA proteins. The list consists of proteins known to be involved in cell wall metabolism, such as the endo-chitinase-like gene CTL2, whereas other proteins (transporters, transcription factors) had no previous link to cellulose biosynthesis and might be promising targets for future research.

Introduction

Plant cell walls are composed of different classes of polysaccharides including cellulose, hemi-cellulose, and pectins and the percentage of each of the polysaccharides depends on the plant species. Cellulose is a major component of both primary and secondary cell walls. Chemically, cellulose is simple polysachariде: it is linear polymer of β-(1,4) glucose. The substrate for cellulose synthesis is UDP-glucose, the glucose moiety of which is incorporated into linear glucan chains by cellulose synthases (CESAs). The CESA proteins are organized in rosette complexes localized in the plasma membrane of plant cells, and believed to be formed by six subunits which in turn consist of six CESA proteins. Biochemical evidence and transcript analyses revealed that distinct CESA proteins are necessary for the correct assembly of the rosette structure. The different CESA proteins have a similar expression pattern, and it has been suggested that at least three different CESA proteins are required for a functional CESA complex. Three proteins, CESA4, CESA7, and CESA8, are required for cellulose synthesis during secondary cell wall formation in vascular tissues. The *irx1* (*Atcesa7*), *irx3* (*Atcesa8*) and *irx5* (*Atcesa4*) mutants show changes in cellulose content in the secondary cell wall, indicating that all three CESAs are essential for proper cellulose synthesis (Taylor *et al.*, 2003). Previous studies have revealed that these proteins specifically interact to form the rosette structure (Timmers *et al.*, 2009).

Reverse genetics have uncovered additional components affecting cellulose biosynthesis. Proteins like the β-(1,4)-glucanase KORRIGAN and endo-chitinase-like protein CTL-1 have been linked to the cellulose machinery because of a reduced cellulose content in the mutant background (Schindelman *et al.*, 2001; Zhong *et al.*, 2002). Also the *kobito*, *knopf* and *botero1* mutants show a compromised cellulose content. Although these mutants have an effect on the cellulose synthesis, it is unknown whether the corresponding proteins are an integral part of the cellulose synthesizing machinery.

Interaction with constitutive components of the rosette structure, the CESA proteins, might be an indirect indication that a particular protein might be involved in cellulose biosynthesis. The membrane-based yeast two-hybrid system (MbYTH) was used to determine the interaction between the secondary wall CESA proteins (Timmers *et al*., 2009), and proved a valuable tool to analyze interactions between CESAs. In this study, we extend this approach by using CESAs as baits to identify other constituents of the rosette complex.

Materials and Methods

Constructs for the MbYTH system

The full-length *A. thaliana* cDNAs were obtained from the Riken Bioresource Centre [12, 13] A full-length *AtCES4* (NM_123770) was obtained for the RBC (RAFL15-30-K05)A SstII restriction site at the 5' end and a SpeI restriction site at the 3'end were generated by PCR with primers 5'-AAACCGCGGATGGAACCAAACACC and
5'-AAACTAGTTAACAGTCGACGCCACA. A full-length *AtCESA7* (NM_121748) was obtained from the RBC (RAFL09-35-F05). A SstII restriction site at the 5' end and a Eco47III restriction site at the 3'end were generated by PCR with primers
5'-AAGACCGCGGATGGAAGCTAGCGCCGGTCTTGT and
5'- AGCGCTTCAGCAGTTGATGCCACACTTG. A full-length *AtCESA8* (NM_117994) was obtained from the RBC (RAFL09-65-M12). A PstI restriction site at the 5' end and a NcoI restriction site at the 3'end were generated by PCR with primers
5'- AAGACTGCAGAATGATGGAGTCTAGGTCTCCC and
5'-AGAACCATGGCATTAGCAATCGATCAAAAGACAGTTC. The proof-reading polymerase Pfu (Fermentas) was used in all PCR set-ups. The resulting PCR-products were digested and ligated in the pTFB1 vector (Bait) (Dualsystems Biotech AG). The bait expression is regulated by the TEF1 promoter. The sequences of the inserts were obtained by Sanger sequence analysis. The library, a Nubg Arabidopsis cDNA-library (Dualsystems P02210), was constructed for six-day-old seedlings. a mixture of dark grown (etiolated seedlings) and seedlings exposed to blue and far red light. The library consists out of $1.7*10^7$ independent clones with an average insert size of 1.7 kb (ranges from 1.2-2.5kb).

Membrane based Yeast two hybrid (MbYTH) screen

The yeast strain NMY51 (Dualsystems Biotech AG) was transformed according to the protocol (DUAL membrane Kit 1). Yeast containing the bait plasmids were transformed with the *Arabidopsis thaliana* library plasmids resulting in approximately $1.8*10^6$ transformants and grown on SD medium (lacking leucine, tryptophan, and histidine) containing the appropriate concentration of 3-ammonium-triazole (130mM, 10mM, and 75mM, for baits CESA4, CESA7, and CESA8, respectively). After five days of growth at 30ºC, the colonies grown were scored. In order to identify interacting proteins, the yeast, containing both bait and prey, was plated on SD medium (lacking leucine, tryptophan, histidine, and adenine) containing the appropriate concentration of 3-ammonium-triazole (130mM, 10mM, and

75mM, for baits CESA4, CESA7, and CESA8, respectively). After five days of growth at 30ºC, the colonies grown were scored. Detection of β-galactosidase activity was performed with the filter-lift assay (Breeden and Nasmyth, 1985). The interactors were identified by PCR, Sanger sequence analysis and consecutively a blast search on the NCBI website (Altschul *et al*., 1997). The list of known auto-activators can be found on the Dual membrane website (www.Dualsystems.com).

Results

Identifying candidate proteins interacting with the cellulose synthezing complex

Several proteins are suggested to be physically linked to the rosette structure. However, so far the only protein, other than CESAs, for which this has been established an interaction is the cellulase KORRIGAN (Chapter 4). In order to find other members of the secondary CESA complex, each of the secondary CESA proteins was used as bait to screen a plasmid library, which expressed the cDNA of Arabidopsis as prey. Only colonies able to grow in the absence of the four auxotrophic markers, and able to activate the *LacZ* gene, were considered as potential interactors with the CESA complex. The preys expressed in these colonies were identified using colony-PCR and subsequent sequence analysis. The corresponding genes were identified by highest similarity to genes of the *A. thaliana* genome available at the NCBI database (Altschul *et al*., 1997). This resulted in a list of 50, 50, and 109 proteins, with several proteins found more than once, potentially interacting with CESA4, CESA7, and CESA8, respectively (Table SII). The list of interactors contained proteins with diverse functions, which were divided in ten groups based on enzymatic activity or metabolic pathway involved (Figure 1A, Table SII), or in seven groups based on sub cellular localization (Figure 1B, Table SII). Several of these groups might contain proteins involved in cellulose biosynthesis as their location and function is associated with cell wall biosynthesis, like the group of proteins located at the plasma membrane.

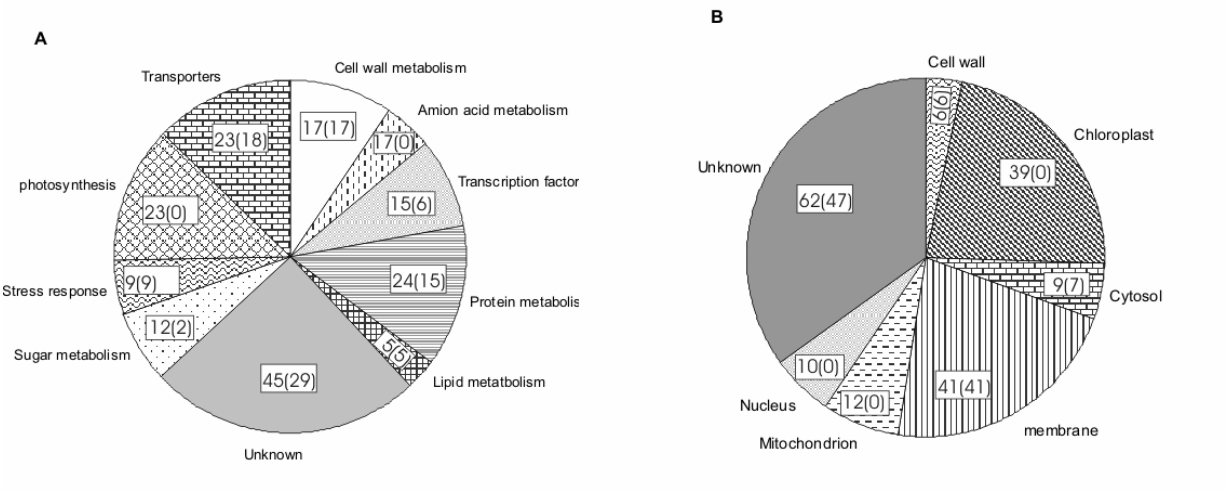


Figure 1. Grouping of the interactors with CESA4, CESA7 and CESA8 based on protein function and localization in the cell. The sum of interactors involved in a similar pathway (A) or sub-cellular location (B) is indicated; the number between brackets is the sum of interactors with the three CESA proteins, CESA4, CESA7 and CESA8, after elimination of biologically non-relevant proteins.

Biological relevance of the protein interactions

In spite of the power of MbYTH to characterize protein-protein interactions, as in all heterologous analyses systems, the biological relevance of the identified interactors needs to be further evaluated. One of the drawbacks of this system was the false positives generated by auto-activation, without interaction between the bait and the prey. These false positives have been identified (see M&M) and consequently discarded as candidates (so-called Filter I to reduce the number of potentially relevant proteins, Table SII). Twenty proteins could be related to auto-activation of the system.

The current view is that the rosettes are assembled in the Golgi apparatus, where the complex exists in an inactive state. The complexes are subsequently transported via cytoplasmic vesicles from the Golgi to the plasma membrane, where they are activated for cellulose synthesis (Haigler and Brown, 1986). Proteins, which do not reside in Golgi, cytosol or plasma membrane, are therefore considered as unlikely candidates of the secondary CESA complex (Filter II, Table SII). Using both criteria the number of candidate genes decreased from one hundred eighty-one to one hundred and one (Table I), as the original list of proteins contained several chloroplast and nuclear proteins (Table SII).

Table I. List of relevant interactors with the three CESA proteins. The list of interactors with the different CESA proteins after elimination of biological non-relevant proteins.

Group	Locus	Name	Cellular location	Function
Cell wall metabolism	At2g20750	AtExpB1	cell wall	expansin
	At2g42840	PDF1	cell wall	protodermal factor 1
	At1g56700			pyrolidone-carboxylate peptidase family
	At5g40390	SIP1	endomembrane	seed imbibition 1-like (GH 36)
	At4g14130	AtXTH15	endomembrane	xyloglucan endotransglycosylase hydrolase (GH16)
	At2g06850	AtXTH4	cell wall	xyloglucan endotransglycosylase hydrolase (GH16)
	At2g05790		endomembrane	endo-glucanase (GH16)
	At2g32990	AtGH9B8	endomembrane	endo-1,4-glucanase (GH17)
	At1g75680	AtGH9B7	endomembrane	endo-1,4-glucanase (GH9)
	At3g16920	CTL2	endomembrane	endo-chitinase like (GH19)
	At3g27540		membrane	(GT17)
	At5g47780	GAUT4	membrane	galacturonosyltransferase (GT8)
	At3g61130	GAUT1	membrane	(alpha-1,4)-galacturonosyltransferase activity (GT8)
	At5g60920	COBRA	membrane	glycosylphosphatidylinositol anchored protein
	At2g01610		membrane	invertase pectinmethylesterase inhibitor
	At1g28580		endomembrane	carboxylic ester hydrolase activity
	At4g13660			pinoresinol reductase
Transporters	At4g17340	Delta-Tip2	membrane	delta tonoplast integral protein
	At2g37180	RD28	membrane	plasma membrane intrinsic protein
	At3g61430	PIP1A	membrane	plasma membrane intrinsic protein
	At1g01620	PIPC1	membrane	plasma membrane intrinsic protein
	At3g53420	PIP2A	membrane	plasma membrane intrinsic protein
	At3g26520	TIP2	membrane	tonoplast intrinsic protein 2
	At2g39010	PIP2	membrane	water channel
	At4g35100	PIP3	membrane	water channel
	At3g16240	Delta-Tip	membrane	water channel
	At2g45960	PIPB1	membrane	water channel
	At1g76850	SEC5A	plasma membrane	exorcist complex

	At5g58060	YKT61	trans Golgi	membrane fusion
	At1g06400	AtRab11E		regulation of vescular trafficking
	At1g76520		endomembrane	auxin efflux carrier family protein
	At2g34250	SEC61	membrane	protein translocase activity
	At1g22530	SEC14	membrane	transporter activity
	At1g14660	AtNHX8	membrane	Na+/H+ exchanger
	At3g45600	TET3	endomembrane	tetraspanin3
Transcription factor	At1g06040	AtSTO		salt tolerance zinc finger
	At3g11400	EIF3G1		translation initiation factor 3G1
	At2g21320		endomembrane	zinc finger B-box type
	At1g69570			DOF-type zinc finger containing protein
	At3g22840	ELIP1		early light inducible protein
	At1g15380			lactosyl glutathione lyase/ glyoxalase
Protein metabolism	At1g14320	60S	cytosol	60S acidic ribosomal protein
	At5g01020			protein kinase
	At3g63490			ribosomal protein L1 family
	At1g58684	40Srib S2		ribosomal protein
	At3g47360	AtHSD3	endomembrane	short-chain dehydrogenase/ reductase (SDR) family protein
	At4g13180	AIS3		short-chain dehyrogenase/ reductase SDR family
	At5g18140	DNAJ		heat shock N-terminal domain containing
	At1g08570			thioredoxin family
	At1g80440			kelch repeat-containing F-box family
	At5g04530		endomembrane	beta-ketoacyl-CoA synthase family
	At5g20050			protein kinase
	at3g53870	RPS3B	cytosol	ribosomal protein
	At3g52190	PHF1	endoplasmic reticulum	phosphate transporter traffic facilitator 1
	At3g09200	RPPOB	cytosolic ribosome	60S acidic ribosomal protein
	At2g42810	PAPP5	cytoplasm	protein phosphatase 5
Lipid metabolism	At1g49660	AtCXE5		carboxyl esterase 5
	At4g19860	LACT		cholesterol acyltransferase fam.
	At2g03980			GDSL-motif lipase/hydrolase family

	At3g21720			cytrate lyase
	At5g13640	ATPDAT	membrane	phospholipid diacylglycerol acyltransferase
Stress response	At1g78380	AtGSTU19	cytoplasm	glutahione transferase
	At5g27380	GSH2	cytosol	glutathione synthase
	At2g37130	PER21	endomembrane	peroxidase activity
	At3g16420	PBP1	cytosol	PYK10-binding protein1
	At3g32980	PER32		peroxidase activity
	At2g32150			dehalogenase-like hydrolase family protein
	At1g76680	OPR1		12-oxophytodienoate reductase 1
	At1g62380	ACO2		1-aminocyclopropane-1-carboxylate oxidase
	At5g57970			methyl adenosine glycosylase fam.
Sugar metabolism	At2g21170	TIM		triosephosphate isomerase
	At1g71170			6-phosphogluconate dehydrogenase NAD-binding domain containing
Unknown	At1g60010			unknown
	At4g38280			unknown
	At4g25670	CPuORF12		unknown
	At1g62780			unknown
	At5g16110			unknown
	At5g42765			unknown
	At3g15450			unknown
	At1g19990			unknown
	At4g27450			unknown
	At1g18490	DUF1637		unknown
	At2g20670			unknown
	At1g19400			unknown
	AT1g11440			unknown
	At1g18740	DUF793		unknown
	At3g59300			unknown
	At3g49140			unknown
	At3g05280			unknown Yip! Family
	At3g60800	DHHC type		unknown Zinc finger protein
	At4g38160	PDE191		Pigment defective 191
	At4g32460		endomembrane	unknown
	At5g55940	EMB2731	endoplasmic reticulum	Embryo defective 2731
	At4g08950	PHI1	cell wall	Phosphate induced protein (PHI1)
	At5g14430		cell wall	dehydatation responsive
	At4g38770	PRP4	cell wall	proline rich protein
	At4g20260	DREPP	membrane	polypeptide family protein
	At5g59500		Membrane	unknown

	At1g09070	SRC2	endoplasmic reticulum	soybean gene regulated by cold-2
	At2g38750	ANNAT4	membrane	calcium ion binding
	At2g32150			dehalogenase-like hydrolase family protein

Indications for involvement in cellulose synthesis

Several of remaining interactors (Table I) have previously found indications that implicate their involvement in cell wall biosynthesis. Some of them can be directly linked to cellulose biosynthesis based on their enzymatic function. Two members of the glycosyl hydrolase family 9, *AtGH8B7* and *AtGH8B8*, were thought to be involved in cellulose metabolism based on their endo-glucanase activity. Another member of this family, KOR1, was found to interact directly with CESA4 and CESA8 (Chapter 4). As KOR1 is important for cellulose synthesis, the action of these other family members enzymes might also be required for proper cellulose synthesis. At2g05790 encodes a family 17 glycosyl hydrolase. These enzymes are able to hydrolyze callose (β (1,3) glucan). As it has been suggested that the CESA proteins can be involved in callose synthesis (Doblin *et al*., 2002), it might be speculated that the family 17 hydrolase has a similar function in callose biosynthesis as KORRIGAN in cellulose biosynthesis.

The endo-chitinase-like protein (*CTL1*) homolog, *CTL2*, was found interacting with CESA7 and has previously been linked to cellulose biosynthesis in the secondary cell wall, based on its high level of co-expression with CESA4, 7, and 8 (Persson *et al.,* 2005). Disruption of this gene caused a Fourier transform infrared phenotype, indicative of alterations in the cell wall composition. Both our findings and those of Persson *et al*., (2005) are indications that CTL2 is involved in cellulose biosynthesis in the secondary cell wall.

COBRA, found to interact with CESA4, has been previously associated with primary cell wall cellulose biosynthesis, as it had an expression pattern resembling that of the primary cell wall CESAs (Persson *et. al.,* 2005). The interaction between COBRA and CESA4 is an indication that this protein might also be involved in secondary cell wall cellulose biosynthesis. Also KOR1 has been associated with both primary and secondary cell wall cellulose biosynthesis (Nicol *et al.,* 1998; Szyjanowicz *et al.,* 2004).

A group of transporters (Table I) might be directly linked to cellulose biosynthesis. The mechanism by which the growing glucan chain is translocated over the plasma membrane, has not been experimentally demonstrated although several models are proposed. One of the models is that the eight transmembrane helices of the CESA protein itself form a pore in the membrane through which the glucan chains are extruded into the cell wall (Delmer, 1999). In

other models, porin-like proteins are proposed to translocate the chains (Bessueille and Bulone; 2008). The water channels found interacting with the different CESA proteins might represent such porin-like proteins. The fact that they were found with all CESA proteins and in almost all cases more than once (Table SII) supports this possible role.

The proteins involved in vesicle transport, like SEC5 and AtRAB11E, might transport the cellulose biosynthesis complex from Golgi to the plasma membrane, the actual site of cellulose biosynthesis. It has previously been found that the CESA proteins are essential and non-redundant and interaction studies have revealed that the CESA proteins specifically interact forming the rosette structure (Timmers *et al*., 2009). This indicates that each CESA protein might fulfill a unique function in cellulose biosynthesis. Interestingly, not all CESA isoforms had a similar number of interactors (Fig. 2). Twenty-one proteins were found more than once, thirteen proteins were found with two different CESA isoforms as bait (Fig. 2). Two proteins, a water channel (Delta-TIP) and a defense response protein (PBP1) showed interaction with all three CESA proteins. Eight proteins are able to interact with both CESA4 and CESA8, four with CESA7 and CESA8, whereas no proteins were found interacting only with both CESA4 and CESA7 (Fig.2). The relatively low number of common interactors between CESA4 and CESA7 might suggest that these proteins have a different function during the formation of cellulose.

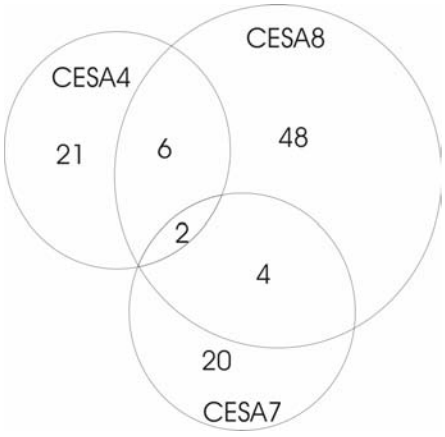


Figure 2. Number of interactors found with the different CESA proteins as bait
The number of proteins, after elimination of biologically non-relevant proteins, found interacting with the different CESA specifically, and simultaneously.

Discussion

Although it has been known for some time that cellulose is produced in a multi protein complex only the main constituents of the rosette complex have been identified. Using a protein-protein interaction approach new candidates have been identified. The interaction with

the secondary CESA proteins is a strong indication for the involvement, of the interactor, in cellulose synthesis. The screen of the cDNA library resulted in the identification of a large number proteins able to interact with the secondary CESA proteins in yeast. The size limitations of the cDNA library, which is restricted to 2.5kb explains the absence of the CESA proteins, as they are 3kb in size. The absence of KOR1 in turn suggests that the screen is not saturated or that KOR1 is in present in the library. Because the protein interactions were tested *in vitro* in an heterologous expression system, additional experiments are needed to establish the relevance of the interactions *in planta* as is illustrated by the PIP2 water channel did not show an interaction in the BiFC assay (Timmers *et al*., 2009). *In vivo* localization of protein can be used to identify false positives and confirm the potential biological relevance of a candidate. Additional knowledge of an interactor can confirm its involvement in cellulose biosynthesis. Proteins of the GH9 family are therefore likely to be involved in cellulose biosynthesis as they are known cellulases, however thus fare only one member of this family, the KOR1 protein, has been identify based on its mutant phenotype. This might indicate that the other family members found are not essential in cellulose synthesis or might be redundant. Proteins like COBRA and CTL2, although previous implicated in cellulose biosynthesis, do not have a known function and therefore it is difficult to speculate about their specific role in the rosette complex without additional research. Also other proteins need addition data to asses their role in cellulose biosynthesis. This study revealed a list of proteins (physically) linked to cellulose biosynthesis, and these proteins might be a promising starting point for future research in this topic.

References

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

Bessueille L., and Bulone V. (2008) A survey of cellulose biosynthesis in higher plants. Plant biotech 25: 315-322

Breeden L., and Nasmyth K. (1985) Regulation of the yeast HO gene. Cold Spring Harb Symp Quant Biol 50: 643-650

Brown D.M., Zeef L.A.H., Ellis J., Goodacre R., and Turner S.R. (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17: 2281-2296

Carpita N., Tierney M., and Campbell M. (2001) Molecular biology of the plant cell wall: Searching for the genes that define structure architecture and dynamics. Plant Mol Biol 47: 1-5

Delmer D.P. (1999) Cellulose biosynthesis: exciting times for a difficult field of study. Annu Rev Plant Physiol Plant Mol Biol 50: 245-276

Doblin M.S., Kurek I., Jacob-Wilk D., and Delmer D.P. (2002) Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell 43: 1407-1420

Gardiner J.C., Taylor N.G., and Turner S.R. (2003). Control of cellulose synthase complex localization in developing xylem. Plant Cell 15: 1740-1748

Ha M.A., MacKinnon I.M., Sturcova A., Apperley D.C., McCann M.C., Turner S.R. and Jarvis M.C. (2002) Structure of cellulose-deficient secondary cell walls from the irx3 mutant of Arabidopsis thaliana. Phytochemistry 61: 7-14

Holland N., Holland D., Helentjaris T., Dhugga K.S., Xoconostle-Cazares B., and Delmer D.P. (2000) A comparative analysis of the plant cellulose synthase (CesA) gene family. Plant physiol 123: 1313-1323

Jones, L., Ennos, A.R., and Turner, S.R. (2001). Cloning and characterization of irregular xylem4 (irx4): a severely lignin deficient mutant of Arabidopsis. Plant J 26: 205-216

Klopotowski T., and Wiater A. (1965) Synergism of aminotriazole and phosphate on the inhibition of yeast imidazole glycerol phosphate dehydratase. Arch Biochem Biophys 112: 562-566

Nicol F., His I., Jauneau A., Verhettes S., Canut H., and Höfte H. (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in Arabidopsis. EMBO J 17: 5563-5576

Pagant S., Bichet A., Sugimoto K., Lerouxel O., Desprez T., McCann M., Lerouge P., Vernhettes S., and Höfte H. (2002) *Kobito1* encodes a novel plasma membrane protein necessary for normal synthesis of cellulose during cell expansion in *Arabidopsis*. Plant Cell 14: 2001-2013.

Persia D., Caimpiero C., Del Casino C., Faleri C., Willemse M.T.M., and Cresti M. (2008) Sucrose Synthase is associated with the cell wall of tobacco pollen tubes. Plant Phys 147: 1603-1618

Persson S., Wei H., Milne J., Page G.P., and Somerville C.R. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci USA 102:8633-8638

Rual J.F., Venkatesan K., Hao T., Hirozane-Kishikawa T., *et al.* (2005) Towards a proteome-scale map of the human protein-protein interaction network. Nature 437: 1173-1178

Schindelman G., Morikami A., Jung J., Baskin T.I., Carpita N.C., Derbyshire P., McCann M.C., and Benfey P.N. (2001) COBRA encodes a putative GPI-anchored protein, which is pearly localized and necessary for oriented cell expansion in Arabidopsis. Genes Dev 15: 1115-1127

Schmid M., Davison T.S., Henz S.R., Pape U.J., Demar M., Vingron M., Schölkopf B., Weigel D., and Lohmann J.U. (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501-506

Szyjanowicz P.M.J., McKinnon I., Taylor N.G., Gardiner J., Jarvis M.C., Turner S.R. (2004) The irregular *xylem 2* mutant is an allele of *KORRIGAN* that affects the secondary cell wall of *Arabidopsis thaliana.* Plant Journal 37: 730-740

Taylor N.G., Laurie S., and Turner S.R. (2000) Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis. Plant Cell 12: 2529-2539

Taylor N.G., Howells R.M., Hutty A.K., Vickers K., and Turner S.R. (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. Proc Natl Acad Sci USA 100: 1450-1455

Turner S.R., and Somerville C.R. (1997). Collapsed xylem phenotype of Arabidopsis identifies mutant deficient in cellulose deposition in the secondary cell wall. Plant Cell 9: 689-701

Timmers J., Vernhettes S., Desprez T., Vincken J.-P., Visser R.G.F., and Trindade L.M. (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett 583: 978-982

Zhong R., Kays S.J., Schroeder B.P., and Ye Z.H. (2002) Mutation of a chitinase-like gene causes ectopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. Plant cell 14: 165-179

Determining the interactions between proteins known to be involved in xyloglucan biosynthesis, and identifying new candidate proteins of the xyloglucan biosynthetic machinery

Jaap Timmers[1,2,3], Jean-Paul Vincken[2,4], Elly Janssen[2], Richard G.F. Visser[2], and Luisa M. Trindade[2]

[1]*CRC-Wageningen, Wageningen University and Research Centre, 6703 HD Wageningen, The Netherlands,*

[2]*Wageningen UR Plant Breeding, Wageningen University and Research Centre, 6708 PD Wageningen, The Netherlands*

[3]*Graduate School Experimental Plant Sciences, Wageningen University, The Netherlands*

[4]*Present address: Laboratory of Food Chemistry, Wageningen University P.O. Box 8129, 6700 EV Wageningen, The Netherlands*

Abstract

Besides cellulose, xyloglucan is a major component of the plant cell wall. The synthesis of xyloglucan occurs in the Golgi, and the resulting polysaccharide is transported to, and inserted into, the plant cell wall. Although the structure of xyloglucan is well established, the enzymes involved in its synthesis are poorly understood. Several glycosyltransferases involved in xyloglucan synthesis, like the xylosyltransferases XXT1, XXT2 and the glucosyltransferase (CSLC4) thought to be responsible for the polymerization of the glucan of the backbone, have been identified. Based on the regular, multiple glycosyl residue structure of xyloglucan, our hypothesis was that, like cellulose, xyloglucan is synthesized by a complex of proteins. In order to characterize this complex, a membrane based yeast two hybrid system was used to characterize the interactions between the XXT1, XXT2, and CSLC4. A model for the protein complex involved in the synthesis of xyloglucan is proposed. Furthermore, several hitherto unknown proteins, like two xyloglucan endo-transglucosylase-hydrolases, were found to interact with XXT1, and are suggested to play a role in xyloglucan biosynthesis.

INTRODUCTION

The plant cell wall is a complex structure that is composed of cellulose, hemicellulose, pectin, lignin, protein, and various inorganic compounds (Carpita and MacCann, 2000). Xyloglucan (XyG) is one of the most abundant hemicelluloses and is believed to play an important role in cell wall structure, where it functions as a cross-linker to form a cellulose-xyloglucan network that constitutes the major load-bearing structure during cell expansion (Veytsman and Cosgrove, 1998). In Arabidopsis XyG is composed of a β-(1,4)-glucan backbone that is substituted with α-(1,6)-xylosyl residues in a regular pattern (Vincken *et al.*, 1997), called XXXG in a standardized nomenclature (Fry *et al*., 1993), where the letters G and X refer to an unbranched β-D-glucose and an α-D-xylose-(1,6)-β-D-glucose segment, respectively. The xylosyl residues can be further substituted at the second and/or third xylose residue from the non-reducing end, by the addition of (1,2)-β-D-galactosyl (segment L) residues and the subsequent addition of α-(1,2)-L-fucose (segment F) at the galactosyl unit (Fig. 1).

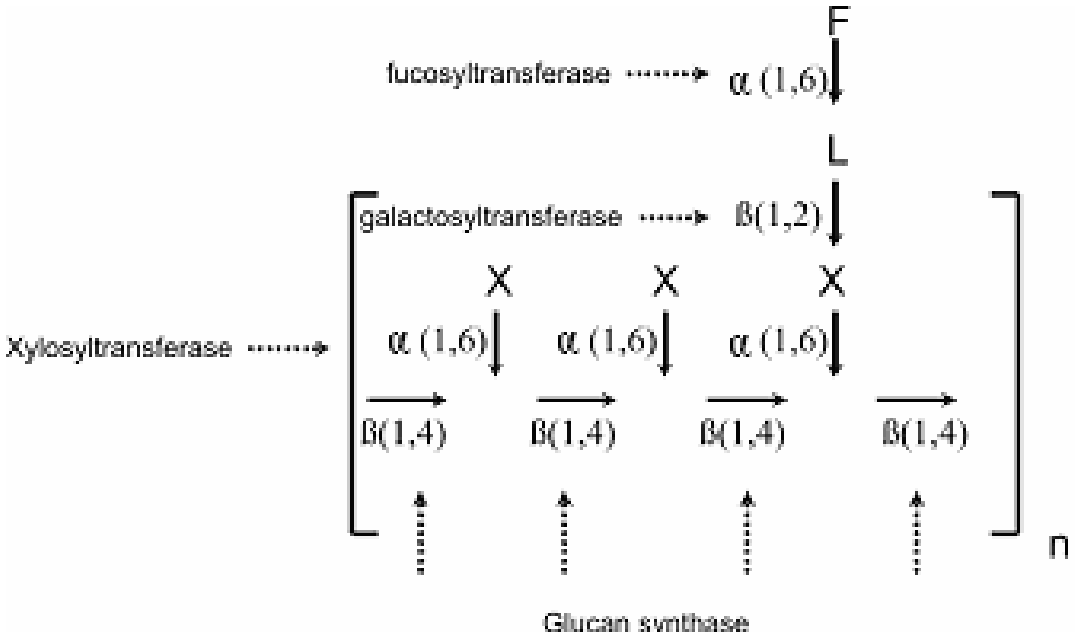


Figure 1. Repetitive structure of xyloglucan. The repetitive structure of XyG and the glycosyltransferases involved in its synthesis.

Considerable progress has been made in the identification of glycosyltransferases involved in the biosynthesis of XyG. The *Atfut1* gene encodes a XyG fucosyltransferase from Arabidopsis (Perrin *et al*., 1999). Plants with a mutated *Atfut1* gene lack fucosylated XyG in

all major organs; however, the plants do not show any changes in their growth or physiology (Vanzin *et al*., 2002). The Arabidopsis *mur3* defect results in a failure of attachment of the galactosyl residue on the third xylosyl unit within the XXXG core structure, and also these plants were phenotypically normal (Madson et al., 2003). Identification of a pea xylosyltransferase led to the discovery of an Arabidopsis gene family consisting of seven putative XyG xylosyltransferases (Faik *et al*., 2002). It seems that XXT1, XXT2, and XXT5 family members are involved in xyloglucan synthesis and are able to add one xylose unit to unsubstituted glucans; under high donor substrate concentrations they can also add the second and even third xylose residue to the adjacent glucosyl residues (Cavalier and Keegstra, 2006). The *xxt1* and *xxt2* double mutant produced slightly smaller plants and lacked detectable XyG, whereas the single mutants did not show a growth reduction, although the glucan backbone was less substituted (Zabotina *et al*., 2008 and Cavalier *et al.*, 2008). It has been suggested that XXT1 and XXT2 are partially redundant, because of the modest reductions of XyG content in the single mutants, which is enhanced in the double mutant (Briggs *et al*., 2006). As the *xxt5* mutant shows similar reduction of backbones substitution with xylose residues, it has been suggested that all three proteins (XXT1, XXT2 and XXT5) are partly redundant (Zabotina *et al.*, 2008). To identify proteins responsible for the synthesis of the glucan-backbone transcriptional profiling of developing Nasturtium (*Tropaeolum majus*) seeds, which contain high amounts of XyG, was used. A gene with high similarity to the Arabidopsis *cellulose synthase-like* (*CSL*) gene family was identified (Cocuron *et al.*, 2007). The subfamily CSLC in Arabidopsis contains five family members. The Nasturtium protein and its homologue from Arabidopsis (AtCSLC4) both show β-glucan synthase activity when expressed in *Pichia pastoris*. Although the *atcslc4* mutant did not show an altered XyG level, Cocuron and co-workers (2007) suggested that the AtCSLC4 protein from Arabidopsis is involved in XyG biosynthesis.

In order to investigate whether XyG is produced in a protein complex, we tested the interactions between the xylosyltransferases and the glucosyltransferase with each other and the ability to form homodimers. Also a library screen was performed to identify proteins, which are involved in xyloglucan synthesis, using the same method.

MATERIALS AND METHODS

Construction of plasmids for the split-ubiquitin system

To construct the bait plasmid, the vector pTMF1 encoding the Cub fragment was used. The reporter in this plasmid consists of the LexA-DNA binding domain and the VP16-activation (TF) (Dualsystems Biotech AG).

The full-length *XXT1* gene (NM_116137) was obtained from the Riken Bioresource Center (Seki *et al.,* 1998 and Seki *et al.,* 2002) (RAFL09-34-I12). The PstI restriction site at the 5'end and the SpeI restriction site at the 3'end were generated by PCR with primers

5'-ACCGCGGATGATAGAGAAGTGTATAGGAG and

5'-TACTAGTCACGTCGTCGTCGTACTAA (restriction sites are underlined). A full-length *XXT2* (NM_116484) was obtained from the RBC (RAFL09-33-I10). The SstII restriction site at the 5'end and the SalI restriction site at the 3'end were generated by PCR with primers

5'-AAACCGCGGATGATTGAGAGGTGTTTAGG and

5'-TTGTCGACTCACGTCGTCGTCGTACTAA. A full-length *AtCSLC4* (NM_113737) was obtained from the RBC (RAFL08-13I06). The NcoI restriction site at the 5'end and the SpeI restriction site at the 3'end were generated by PCR with primers

5'-AAACCATGGAATGGCTCCAAATTCAGTAGCAGTGAC and

5'- AAACTAGTTTCTAGCTGATCTGTTCTCCGATCAAATCC. The proof-reading polymerase Pfu (Fermentas) was used in all PCR set-ups. The resulting PCR-products were digested and ligated in the pTFB1 vector (Dualsystems Biotech AG).

The vector pADSL-Nx, which encodes the Nub fragment, was used to construct the prey plasmid (Dualsystems Biotech AG). For making the prey constructs, an EcoRI linker and a SalI linker were generated by PCR, at the 5'and 3' end, respectively. For the *XXT1* gene the primers 5'-AAAGAATTCAATGATAGAGAAGTGTATAGGAG and

5'- TTTGTCGACTCACGTCGTCGTCGTACTAA were used, for *XXT2* primers

5'-AAAGAATTCAATGATTGAGAGGTGTTTAGG and

5'-TTTGTCGACTCAAACTTGATTGGTTTGTAC, and for *AtCSLC4* primers

5'- TTTGAATTCAATGGCTCCAAATTCAGTAGCAGTGAC and

5'-TTGTCGACTTTCTAGCTGATCTGTTCTCCGATCAAATCC. The resulting PCR-products were digested and ligated in the pADSL-Nx vector (Dualsystems Biotech AG). The sequences of the inserts were confirmed by Sanger sequence analysis.

Yeast two-hybrid assay

The bait and prey constructs were co-transformed into the yeast strain NMY51 (Dualsystems Biotech AG) according to transformation procedure (DUAL membrane Kit 1). The yeast, containing both plasmids, was plated onto synthetic medium lacking leucine, tryptophan (SD med.-L-T), and grown at 30ºC for three days. To quantify the interactions between different preys, 100 colonies of each combination were spotted on selection medium (synthetic medium lacking leucine, tryptophan, histidine, and adenine (SD med.-L-T-H-A)) containing the appropriate 20mM of 3-Ammonium-triazole and grown at 30ºC for three days. The number of spots grown was then counted, to confirm the results the experiment is replicated. Detection of β-galactosidase activity was performed with the filter-lift assay (Breeden and Nasmyth, 1985).

Yeast two-hybrid library screening

The library used for the screening for new interactors was a Nub Arabidopsis cDNA-library (Dualsystems P02210) constructed for six-day-old seedlings of a mixture of dark grown (etiolated seedlings) and seedlings exposed to blue and far red light. The library consisted of $1.7*10^7$ independent clones with an average insert size of 1.7 kb (ranges from 1.2-2.5kb). The yeast strain NMY51 (Dualsystems Biotech AG) containing the bait plasmid was transformed with the cDNA library and plated onto synthetic medium lacking leucine, tryptophan, and histidine (SD med.-L-T-H), and grown at 30ºC for five days. Colonies were plated onto synthetic medium lacking leucine, tryptophan, histidine, and adenine (SD med.-L-T-H-A), and grown at 30ºC for three days. Colonies were tested for β-galactosidase activity using the filter-lift assay (Breeden and Nasmyth, 1985).

RESULTS

Xyloglucan is produced in a protein complex

The XXT1, XXT2, and CSLC4 were tested for interaction using a membrane based yeast two hybrid (MbYTH) system. Interactions were first tested with XXT1 fused to the C-terminal part of the ubiquitin (Cub) and the transcription factor (bait), whereas the XXT2 and CSLC4 proteins were fused to the N-terminal part of the ubiquitin (Nub) (preys). Upon interaction between the bait and the prey the ubiquitin is restored, and subsequently the transcription

factor (TF) is released into the nucleus where it activates reporter genes allowing the yeast to grow on selective medium. Yeast growth indicates that XXT1 is able to interact with both proteins (Fig. 2A). When XXT2 was used as bait, similar results were found in that it also interacted with XXT1 and AtCSLC4 (Fig. 2B). The interactions were confirmed using AtCSLC4 as bait and the XXT1 and XXT2 as preys; both combinations were able to grow on selective medium. The capacity of these enzymes to form homodimers was also tested, and both xylosyltransferases and the AtCSLC4 protein were able to do so (Fig. 2). All combinations activated the colorimetric marker, which confirmed the interactions, whereas the negative control indicated that the interactions were specific.

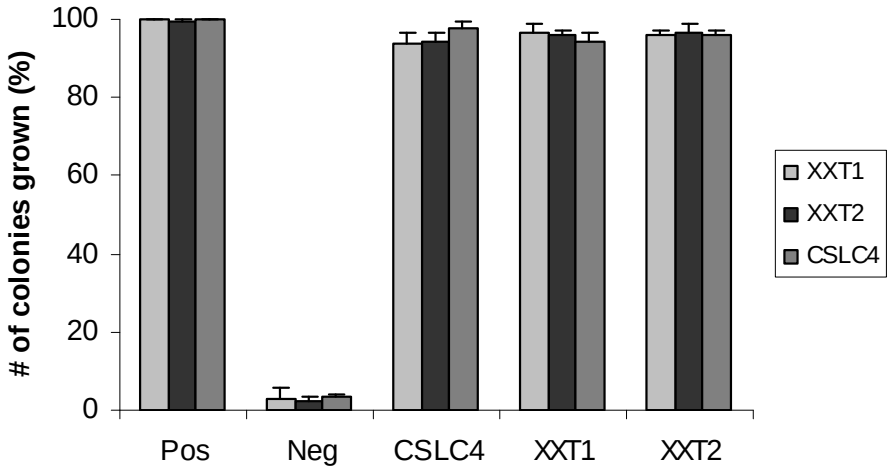


Figure 2. Interactions between the glycosyltransferase involved in XyG synthesis visualized by yeast growth. Yeast expressing XXT1, XXT2 or CSLC4 as bait with different NubG fusion proteins. The percentage of colonies that show visible growth after 5 days at 30ºC on selective medium is shown. Standard deviation is visualized by the error bar.

Identifying new candidates which are part of the xyloglucan synthezing complex

To identify, other proteins comprised in the XyG synthesizing machinery, the XXT1 protein was used as bait to screen a cDNA library. Yeast containing the XXT1 as bait was transformed with an Arabidopsis cDNA library and plated on selection plates lacking histidine (20mM 3-AT). A total of fifty-eight independent colonies were able to grow. Additional selection on the autotrophic marker adenine and colorimetric marker LacZ identified two colonies to be false positives. The remaining fifty-six colonies were considered true interactors. These fifty-eight colonies corresponded to forty-one different proteins interacting with XXT1 in yeast (Table I).

Biological relevance of the protein interactions

As the protein interactions were tested in yeast, these results might not always reflect biologically relevant interactions, and false positives need to be filtered out. Some proteins are known to activate the reporter genes without an interaction between the bait and the prey. Seven of these known auto-activators were found in the list and disregarded as interactors (Table SIII, filter 1). A second filter was based on localization. Expression of proteins in yeast might cause them to be present at a different location in the cell, which makes them available for interaction with proteins they would normally not encounter *in planta.* As XyG synthesis takes place in the Golgi, only Golgi-localized proteins should be considered as interactors. Proteins which are located in different cell organelles *in planta*, like the cell nucleus or the chloroplast, are unlikely candidates to be involved in XyG biosynthesis. Therefore, twenty-one proteins were discarded as candidate genes base on localization (Table SIII, filter 2). The remaining list of interactors (Table I) was studied in more detail as they might represent proteins able to interact *in planta* with XXT1.

Table I. List of relevant interactors with the XXT1 protein. The list of interactors with the different CESA proteins after elimination of biological non-relevant proteins.

Locus	Protein name	Protein name		Localization
At1g61250	SCAMP3	secretory carrier 3		post Golgi / cytosol
At4g15780	ATVAMP724	vesicle-associated membrane protein 724		post Golgi / plasma-membrane
At5g09440		phosphate-responsive protein		
At1g20630	CAT1	catalase 1	H202 to water en oxygen	
At1g78630	EMB1473	embryo defective 1473	structural constituent of ribosome	
At3g26650	GAPA	Glyceraldehyde 3 phosphate dehydrogenase A subunit		
At4g13940	HOG1	homology-dependent gene silencing 1	adenosyl homocysteinase	
At2g38080	LAC4	laccase 4	copper ion binding/ oxidoreductase	
At4g37540	LBD39	LOB domain-containing protein 39		
At1g80070	SUS2	abnormal suspensor2	transcriptionfactor	

At1g12780	UGE1	UDP-D-glucose/UDP-D-galactose-4-Epimerase 1		
At4g28850	AtXTR18	xyloglucan endo-transglycosylase hydrolase	xyloglucosyl transferase	
At4g14130	AtXTH15	xyloglucan endo-transglycosylase hydrolase	xyloglucosyl transferase	

Indications for involvement in xyloglucan synthesis

The remaining interactors (Table I) were screened for indications which strengthen the relevance of the interaction with XXT1. Some proteins detected can be linked directly to xyloglucan synthesis based on their catalytic function. The UGE1 produces UDP-galactose which might be used as a donor substrate for the galactosyl side chain of xyloglucan. The two xyloglucan endo-transglucosylase-hydrolases (XTH), found interacting, can cleave donor xyloglucan chains and rejoin the newly formed reducing end to the non-reducing terminus of an available acceptor xyloglucan chain or oligosaccharide (EC 2.4.1.207).

Co-expression of genes might also be an indication for the function of the interaction with the XXT1. The two interactors found in the library screen, LACCASE4 (LAC4) and ABNORMAL SUSPENSOR2 (SUS2), are both involved in secondary cell wall synthesis. The *lac4* gene showed a similar gene expression pattern as the *Atcesa7*, which is part of the complex required for cellulose biosynthesis in the secondary cell wall (Brown *et al*., 2005). Furthermore, the phenotype of the corresponding mutant, a mild irregular xylem, is indicative for a secondary cell wall defect. Both are indications that LAC4 is involved in cell wall biosynthesis. Also *sus2* has been confirmed as differentially expressed during cell wall development, and is therefore thought be involved in the construction of the secondary cell wall (Yang *et al.*, 2008). Although xyloglucan is one of the major components of the primary cell wall it is also found as a secondary cell wall storage polysaccharide (Buckeridge *et al*., 2000; Reid *et al*., 1985). Based on the expression patterns of the *lac4* and *sus2* it is likely that these genes are involved in this phase of the cell wall production.

DISCUSSION

Although there are indications that some of the polysaccharides of the cell wall are synthesized in a complex, this has only been confirmed for the cellulose synthesis. The conserved structure of XyG within plants implies that the synthesis of XyG needs to be precisely orchestrated which might indicate XyG is synthesized in a protein complex. Results of the one to one interaction study show that the proteins, which attach xylosyl side chains to the growing glucan backbone, physically interact with the protein which is thought to produce this glucan chain. This corroborates the idea that AtCSLC4 synthesizes the backbone of XyG, as well as the hypothesis that XyG is produced in a protein complex. Our results lead us to propose the topology for the XyG synthesizing complex as indicated in Fig. 3.

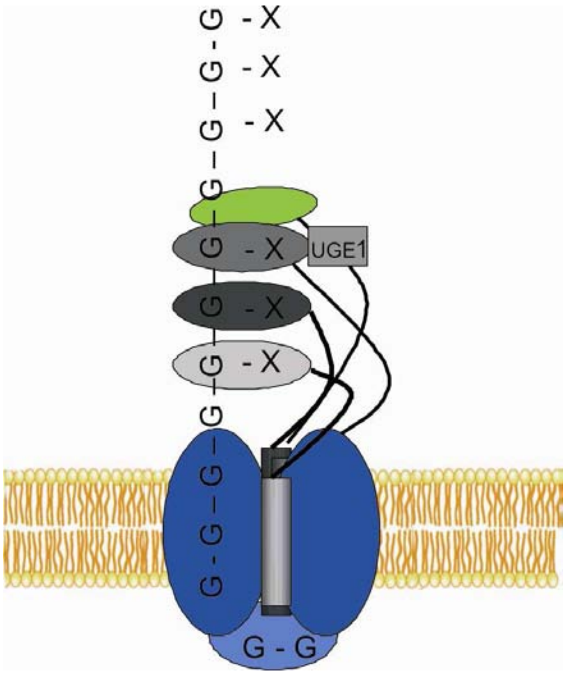


Figure 3 Model of the xyloglucan synthezing complex. A model based on the interactions between the known enzymes involved in XyG synthesis and found to interacting. The integral membrane protein, CSLC4 (Blue), forms a homodimer and the two active sites form a disaccharide which is attached to the growing glucan chain, the XXT proteins (Grey) add xylosyl side chains to the backbone (the presents of a third xylosyltransferase in the complex still has to be established.). The XTH protein (green) cleaves the xyloglucan and attaches it to an acceptor molecule.

The ability of AtCSLC4 to homodimerize might be an indication that this enzyme is functional as a dimer during XyG biosynthesis. The CESA are thought to be a processive enzymes, adding two glucosyl residues simultaneously to the growing chain (Richmond and Somerville, 2000). Two active sites are needed for this, in which subsequent glycosyl units

are flipped 180 degrees, such as in chitin (Yeager and Finney, 2004). This might also apply to the synthesis of the β-(1,4) glucan backbone of XyG. Another possibility might be that the homodimer of CSLC4 produces two β-glucan chains, as is thought each of the cellulose synthase of the rosette complex produces one glucan chain of the microfibril (Somerville *et al*., 2004) and the dimerizations stabilize the protein complex.

As at least two of the known xylosyltransferases interact with the AtCSLC4 protein, it is speculated that the XyG biosynthesis complex consists of at least two (and possibly three) different xylosyltransferase each adding one specific xylosyl side-chain, which is consistent with previous research (Zabotina *et al*., 2008 and Cavalier *et al.*, 2008).

The XTH proteins, AtXTH15 and AtXTH26, found interacting with XXT1 are members of the xyloglucan endo-transglycosylase hydrolase family 16. Members of this family were found to internally cleave xyloglucan and ligate the newly generated reducing ends onto the O-4 of the non-reducing terminal glucosyl residue of an acceptor, which can be a xyloglucan or an oligosaccharide of xyloglucan (Uozu *et al*., 2000). The ability of the XTH proteins to remodel XyG molecules might have several functions during synthesis, such as disconnecting nascent chains from the actual biosynthetic machinery, and thus terminating chain elongation. A similar function is ascribed to korrigan, a protein also found to be part of the complex synthesizing cellulose (Mølhøj *et al*., 2002; Chapter 3). *Atxth15* shows a similar expression pattern as *xxt1*, which lends further support to the hypothesis that XTH is part of the complex.

XXT1 was also found interacting with UGE1, an enzyme that interconverts UDP-glucose and UDP-galactose. A family of five UGE isoforms is encoded in the Arabidopsis genome, some of them known to be involved in the galactosylation of xyloglucan (Rösti *et al*., 2007). It has been suggested that UGE1 is the dominant isoform in green plant parts (Dörmann and Benning, 1998). Knock-down of UGE1 however, neither induces a morphological phenotype nor alters cell wall polymers or any other galactose containing carbohydrate (Rösti *et al*., 2007). The UDP-galactose produced by UGE1 can be used as a donor substrate for the galactose side chain of XyG. The production of UDP-galactose might be an indication that the galactosyl side chain is also added within this complex. The on-site production of the donor substrate for the galactosyltransferase might ensure a more efficient incorporation of galactose in xyloglucan, a concept also suggested for cellulose biosynthesis. SUSY, the enzyme that produces the donor substrate, UDP-glucose, for cellulose synthesis suggested to

be part of the complex involved in the biosynthesis of this polysaccharide (Carlson and Chourey, 1996).

Two proteins found (SCAMP3 and VAMP724) are known to be involved in the control of protein and lipid trafficking by vesicle sorting at the trans-Golgi network and transportation to the plasma membrane (Sanderfoot *et al*., 2000). The interaction with the xylosyltransferase might target the vesicle containing XyG to the plasma membrane and therefore might be a temporal interaction. In the list of interactors there are some genes with a putative function which can not be directly linked to XyG biosynthesis to date. Future research has to reveal whether their interaction with XXT1 has a biological relevance and what their place in the XyG machinery is. Several new candidate genes were found using the library screen, however it cannot be ruled out that more glycosyltransferases are involved in the synthesis of XyG, as a few known interactors of XXT1 where not identified in this screen.

The characterization of the protein complex involved in xyloglucan synthesis is a step forward in the understanding of xyloglucan biosynthesis and the proteins involved. The further characterization of the interactors identified in this study might be a valuable starting point for future research.

References

Brown D.M., Zeef L.A.H., Ellis J., Goodacre R., and Turner S.R. (2005) Identification of novel genes in Arabidopsis involved secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17: 2281-2295

Brummell D.A., Camirand A., and Maclachlan G.A. (1990) Differential distribution of xyloglucan glycosyl transferases in pea Golgi dictyosomes and secretory vesicles. J Cell Sci 96:705-710

Buckeridge M.S., Santos H.P., and Tiné M.A. (2000) Mobilisation of storage cell wall polysaccharides in seeds. Plant Physiol Biochem 38:141-156

Campbell P., and Braam J. (1999) In vitro activities of four xyloglucan endotransglycosylases from Arabidopsis. Plant J 18: 371-382

Carlson S.J., and Chourey P. (1996) Evidence for plasma membrane-associated forms of sucrose synthase in maize. Mol Gen Genet 252: 303-310

Carpita , N.C., and McCann, M. (2000) The cell wall. In biochemistry and Molecular Biology of Plants, B.B. Buchanan, G. Wilhelm, and R.L. Jones, eds Rockville, MD: American Society of Plant Physiologists 52-108

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

Cavalier, D.M., Lerouxel, O., Neumetzier, L., Yamauchi, K., Reinecke, A., Freshour, G., Zabotina, O.A., Hahn, M.G., M.G., Burgert, I., Pauly, M., Raikehl, N.V., and Keegstra, K. (2008) Disruption of two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. Plant Cell 20: 1519-1537

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 glucansynthase. Proc Nat Acad Sci USA 104: 8550-8555

Dörmann P., and Benning C. (1998) The role of UDP-glucose epimerase in carbohydrate metabolism of Arabidopsis. Plant J 13: 641-652

Edwards M.E., Dickson C.A., Chengappa S., Sidebottom C., Gidley M.G., and Reid J.S. (1999) Molecular characterization of a membrane-bound Galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. Plant J 19: 691-697

Faik A., Price N.J. Raikhel N.V., and Keegstra K. (2002) An Arabidopsis gene encoding an α-xylosyltransferase involved in xyloglucan biosynthesis. Proc Natl Acad Sci USA 99: 7797-7802

Fetchko M., and Stagljar I. (2004) Application of the split-ubiquitin membrane yeast two-hybrid system to investigate membrane protein interactions. Methods 32: 349-362

Fry S.C., York W.S., Albersheim P., Darvill A., Hayashi T., Joseleau J-P., Kato Y., Lorences E.P., Maclachlan G.A., McNeil M., and others (1993) An unambiguous nomenclature for xyloglucan derived oligosaccharides. Physiol Plant 89: 1-3

Keegstra K., and Raikhel N., (2001) Plant glycosyltransferases. Curr Opin Plant Biol 4: 219-224

Kiiskinen L.-L., Palonen H., Linder M., Viikari L., and Kruus K. (2004) Laccase from *Melanocarpus albomyces* binds effectively to cellulose. FEBS Lett 576: 251-255

Li X., Cordero I., Caplan J., Molhoj M., and Reiter W.D. (2004) Molecular analysis of 10 conding regions from Arabidopsis that are homologous to the MUR3 xyloglucan galoctosyltransferase. Plant Physiol 134: 940-950

Madson M., Dunand, C., Li X., Verma R., Vanzin G.F., Caplan J., Shoue D.A., Carpita, and Reiter W.D. (2003) The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. Plant Cell 15: 1662-1970

McNeil M., Darvill A., Fry S.C., and Albersheim P. (1984) Structure and function of the primary cell walls of plants. Annu Rev Biochem 53: 625-663

Mølhøj M., Ulvskov P., and Degan F.D. (2001) Characterization of a functional soluble form of a *Brassica napus* membrane-anchored Endo-1,4-β-glucanase heterologously expressed in *Pichia pastoris.* Plant physiol 127: 674-684

Nguema-Ona E., Andème-Onzighi C., Aboughe-Angone S., Barbor M., Ishii T., Lerouge P., and Driouich A. (2006) The reb1-1 mutation of Arabidopsis. Effect on the structure and localization of galactose-containing cell wall polysaccharides. Plant physiol 140: 1406-1417

Obel N., Neumetzler L., and Pauly M. (2006) Hemicelluloses and cell expansion Plant Cell Monographs 57-88

O'Neill M.A., York W.S. (2003) in Rose J.K. (ed) The composition and structure of plant primary cell walls. Blackwell/CRC, Boca Ration, FL, pp 11-31

Reid J.S.G., Cell wall strorage carbohydrates in seeds – biochemistry of the seed 'gums' and 'hemicelluloses'. Adv Bot Res 11: 125-155

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 Physiol 134: 443-451

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

Persson S., Wei H., Milne J., Page G.P., and Somerville C.R. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarry data sets. Proc Nat Acad Sci USA 102: 8633-8638

Richmond T.A., and Sommerville C.R. (2000) The cellulose synthase superfamily. Plant Physiol 124:495-498

Rösti J., Barton C.J., Albrecht S., Dupree P., Pauly M., Findlay K., Roberts K., and Seifert G.J. (2007) UDP-glucose 4-Epimerase isoforms UGE2 and UGE4 cooperate in providing UDP-Galactose for cell wall biosynthesis and growth of *Arabidopsis thaliana*. Plant Cell 19: 1565-1579

Sanderfoot A.A., Assaad F.F., and Raikhel N.V. (2000) The Arabidopsis genome. An abundance of soluble N-Ethylmaleimide-sensitive factor adaptor protein receptors. Plant Physiol 124: 1558-1569

Sarria R., Wagner, T.A., O'Neill M.A., Faik A., Wilkerson C.G., Keegstra K., and Raikehl N.V. (2001) Characterization of a family of Arabidopsis genes related to xyloglucan fucosyltranferase1. Plant Physiol 127: 1595-1606

Schmid M., Davison T.S., Henz S.R., Pape U.J., Demar M., Vingron M., Schölkopf B., Weigel D., and Lohmann J.U. (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501-506

Seki M., Carninci P., Nishiyama Y., Hayashizaki Y. and Shinozaki K. (1998). High-efficiency cloning of Arabidopsis full-length cDNA by biotinylated CAP trapper. Plant J 15: 707-720

Seki M., Narusaka M., Kamiya A., Oono Y., Muramatsu M., Hayashizaki Y., Kawai J., Carnici P., Itoh M., Ishii Y., Arakawa T., Shibata K., Shinagawa A., and Shinozaki K. (2002). Functional annotation of a full-length Arabidopsis cDNA collection. Science 296: 141-145.

Thompson D.S. (2005) How do cell walls regulate plant growth? J Exp Bot 56: 2275-2285

Timmers J., Vernhettes S., Desprez T., Vincken J.-P., Visser R.G.F., and Trindade L.M. (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett 583: 978-982

Vanzin G.F., Madson M., Carpita N.C., Raikhel N.V., Keegstra K., and Reiter W.D. (2002) The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proc Nat Acad Sci USA 99: 3340-3345

Veytsman B.A., and Cosgrove D.J. (1998) A model of cell wall expansion based on thermodynamics of polymer networks. Biophys J 75: 2240-2250

Vincken, J.-P., York, W.S., Beldman, G., and Voragen, A.G.J. (1997) Two general branching patterns of xyloglucan, XXXG and XXGG. Plant Physiol 114: 9-13

Yang X., Tu L. Zhu L., Fu L., Min L., And Zhang X. (2008) Expression profile analysis of genes invovled in cell wall regeneration during protoplast culture in cotton by supression subtractive hybridization and macroarry. J Exp Bot 59: 3661-3674

Yeager A.R., and Finney N.S. (2004) The first direct evaluation of the two-active site mechanism of chitin synthase. JOC 69: 613-618

Zabotina O.A., Ven van de W.T.G., Freshour G., Drakakaki G., Cavalier D., Mouille G., Hahn M.G., Keegstra K., and Raikhel N.V. (2008) Arabidopsis XXT5 gene encodes a putative α-1,6-xylosyltransferase that is involved in xyloglucan biosynthesis. Plant J 56: 101-115

General discussion

The research described in this thesis was performed to generate more in-depth knowledge of cell wall biosynthesis, and notably to identify the different proteins interacting to form the cell wall biosynthetic complexes. Previous studies using common genetic approaches have proven to be successful in obtaining valuable knowledge in cell wall research. However several questions remain which due to the limitations of these approaches need to be addressed in a different way. In this thesis, a different approach to characterize the plant cell wall biosynthesis was chosen, i.e. one embarking on protein-protein interactions, which might provide a more detailed view on the protein complexes involved in plant cell wall biosynthesis. The membrane-based yeast two-hybrid system (MbYTH) was chosen to study the *in vitro* interactions of proteins involved in the synthesis of cellulose and xyloglucan, and the BiFC was used to confirm these interactions *in planta*. Although other protein interaction assays could have been used, the MbYTH system was chosen as it is a high throughput method, not only enabling verification of interactions between known interactors but also enabling discovery of new interactors. The MbYTH method was optimized to identify new components of and insights in the cell wall biosynthesizing machinery (Chapter 2).

Cellulose synthesizing complex

Specific interactions between CESAs in the rosette complex

To get more insight in the protein complex responsible for the synthesis of cellulose in the secondary cell wall, the interactions between the cellulose synthases (CESA4, 7 and 8) have been determined. The first step was the understanding of the organization of the CESA proteins in the rosette complex. Although, it had been suggested that homodimerization of these CESA proteins is the first step of the rosette formation (Kurek *et al*., 2002), it was found that only CESA4 was able to form homodimers. As no homo-dimerization of CESA7 and CESA8 was detected it was concluded that the method was able to identify the specific interactions between the highly homologous CESA. The specific interaction was also an indication that the CESA proteins have a specific position in the organization of the rosette and the model proposed Doblin and co-workers (Doblin *et al*., 2002) was modified accordingly (Timmers *et al*., 2009). In the new model the homo-dimerization of CESA4 leads

the subunits to form the rosette structure (Timmers *et. al.,* 2009; Chapter 3). Several assumptions were made based on previous research by others. For the secondary cell wall complex the stoichiometry of the complex has never been resolved, although gene expression indicates a 1:1:1 ratio between the different secondary *CESA* genes (Person *et al.,* 2005). Another indication for this proposed stoichiometry of the secondary cell wall CESA complex is the stoichiometry of the complex synthesizing the cellulose for the primary cell, which has been determined at 1:1:1 for the three primary CESA proteins (Elizabeth Crowell, INRA, France, personal communication). The specific position of each of the CESA proteins might also hint towards an individual function of the proteins in the synthesis of cellulose, however as each of the CESA isoforms is essential for assembly of the rosette, it is difficult to determine their specific functions. This specific position of one CESA in the complex might explain the failure for compensation by another CESA protein. Identifying orthologs with a similar interaction pattern might be able to complement the knock out of a specific CESA protein. This might also enable the incorporation of exogenous CESA in the rosette complex, thereby modifying the cellulose produced.

Identification of proteins interacting with the rosette complex

Several proteins linked to cellulose biosynthesis have been identified and among them is the cellulase KORRIGAN. The mutant phenotype for the knockout of this gene resembles the phenotype obtained with plants containing mutated CESA proteins (Nicol *et al*., 1998; Szyjanowicz *et al*., 2004), and it has been suggested that KOR1 digests the growing glucan chains during cellulose synthesis (Nicol *et al*., 1998). Therefore, it was thought that the KOR1 protein might be an additional component of the rosette complex. To confirm this hypothesis, the interactions between the CESA proteins and KOR1 were tested. It was found that KOR1 was able to bind to three different primary CESA proteins (#1, #3, and #6), as well as specifically to the CESA proteins of the secondary cell wall, as it bound to both CESA4 and CESA8; however no interaction was found between CESA7 and KOR1 (Chapter 4). The specific interactions with KOR1 suggest that the different CESA proteins might have a specific function, which is also implied by their specific position in the complex. The ability of the KOR1 protein to form homodimers indicates that this protein might function as a homodimer in the rosette complex.

Additional proteins involved in cellulose biosynthesis were found by using the MbYTH system to screen an *A. thaliana* cDNA library for interactors with the secondary CESA proteins (Chapter 5). This resulted in a list of proteins with very diverse catalytic functions. The list did not contain any of the proteins found interacting with the CESA proteins in previous studies, like KOR1 and the different CESA proteins themselves. The absence of the CESA a proteins among the interactors can be explained by the size limitations of the cDNA library, as it is restricted to 2.5kB, whereas the absence of KOR1 might indicate an incomplete screening of the library or absence of the *kor1* gene in the library. The list of interactors also contains proteins which Several proteins could be directly linked to cellulose biosynthesis and the rosette complex based on their protein function. Two examples are COBRA and CTL2, found to interact with CESA4 and CESA7, respectively. These proteins have been previously implicated in cell wall biosynthesis (Roudier *et al*., 2005). The direct interaction with the rosette links both COBRA and CTL2 to cellulose biosynthesis. Several cellulases have been found interacting with CESA8. An explanation for the interactions found between all CESA proteins and water channels might be that these proteins form a pore through which the growing glucan chain is transported trough the plasma membrane. The protein interaction approach did not only reveal proteins previously related with cellulose biosynthesis, it also identified interactions with gene function thus far not implicated in cellulose biosynthesis. Therefore this screen is a powerful tool to limit the number of candidate genes involved in a specific pathway or process like polysaccharide biosynthesis.

Identification of motives critical for protein-protein interactions

To understand the nature of the interaction between KOR1 and the CESAs, different domains of these proteins were studied in more detail. It has been suggested that the RING-finger motif of the CESA protein might be responsible for the interaction between these proteins (Kurek *et al.*, 2002). However, modification of the essential cysteines in the RING-finger motif did not result in a complete disruption of the interactions (Chapter 3). Therefore it was concluded that the RING-finger was not essential for the interaction between the different isoforms and might be involved in the binding of proteins other then the CESA. As KOR1 did not comprise a specific domain related to protein-protein interaction, based on sequence analysis, different parts of the protein were tested for interaction with the CESA proteins. It was found that the transmembrane domain (TMD) was responsible for the interaction

between KOR1 and CESA proteins, as all partial proteins containing this domain are able to interact. The lack of interaction in absence of TMD led to the conclusion that this domain is essential for the interaction with the CESA proteins (Chapter 4). As the TMD of KOR1 is essential in the interaction with the CESA proteins, it is expected that the CESA counterpart is also one (or more) of its eight TMDs, as it is known that transmembrane helices can interact with each other in a specific way (Schneider *et al.*, 2007). This insight in the incorporation of proteins in the complex allows a new approach in cellulose modification. The TMD of KOR1 can be used as a tool to attach proteins to the CESA proteins and therefore the rosette complex as it is sufficient for the interaction with the CESA proteins. Using this domain one might be able to anchor an enzyme into the protein complex during cellulose synthesis, which then can be involved in or interfere with the synthesis of this polymer. A possibility might be the incorporation of a more active cellulase to replace KOR1, in order to obtain shorter glucan chains incorporated in the microfibril.

Xyloglucan synthesizing complex

Interactions between proteins involved in xyloglucan biosynthesis

Xyloglucan is another major constituent of the plant cell wall, and considerable progress has been made to identify glycosyltransferases involved in xyloglucan biosynthesis. Although several indications suggest the polysaccharide to be produced in a protein complex, this has not been established experimentally. The interaction between several of the known glycosyltransferases involved in this process has been tested to confirm this. Our results indicate that two different xylosyltransferases, XXT1 and XXT2, can both form homo- and heterodimers as well as bind with the glucosyltransferase (AtCSLC4), thought to be responsible for the synthesis of the glucan backbone. Also the AtCSLC4 protein is able to form homodimers. These results do not only corroborate the involvement of AtCSLC4 in xyloglucan biosynthesis, it also indicated that xyloglucan in produced in a protein complex. The core of the protein complex is suggested to be composed of two AtCSLC4 homodimerized proteins, which produce the glucan backbone, and three different xylosyltransferases, which add the xylosyl groups to three of the four glucose residues (Zabotina *et al.*, 2008 and Cavalier *et al.*, 2008; Briggs *et al.*, 2006). As the amino acid sequence of AtCSLC4 is very homologues to that of the CESA proteins and the

xylosyltransferase are type II membrane proteins as is the KOR1 protein, one might suggest that these proteins have a similar way of interacting. Therefore, the interaction might also take place in the membrane between the transmembrane domains. This could be an interesting subject for future studies.

To identify other constituents of this complex a cDNA library was screened for interactors with MbYTH system using the XXT1 protein as bait. After the screen, based on biological relevance, this resulted in a list of thirteen proteins which can interact to the xyloglucan producing complex. UDP-galactose epimerase, found interacting with XXT1, can be directly linked to xyloglucan synthesis. The presence of this protein in the complex might be an indication that galactosyl side chains are also added in the complex, although these residues are less evenly distributed than xylosyl residues. This indicates that, although not found in this study, a galactosyltransferase might be part of the complex synthesizing xyloglucan. Although the xyloglucan endo-transglucosylase-hydrolases (XTH) are known to be involved in cell wall metabolism, in that they cleave and rejoin xyloglucans in the cell wall, they have thus far not been implicated in xyloglucan biosynthesis. The interaction with the XXT1 protein links these proteins to the location where the xyloglucan is synthesised, where they might release the growing xyloglucan chain from the complex, and attach it to another xyloglucan in the Golgi. Other proteins, like the vesicle proteins, might be only temporally part of this complex to regulate translocation of the complex, or its product xyloglucan, towards the cell wall. Other proteins do not have an obvious link to xyloglucan biosynthesis but need follow up experiments to confirm their interaction *in planta* and to indicate what their involvement is in xyloglucan biosynthesis.

The membrane based yeast two hybrid assay

The ability of the MbYTH system to test one to one protein interactions between membrane bound proteins revealed several new insights in the biosynthesis of both cellulose and xyloglucan. These one to one screens show repeatable results which have been confirmed by the reverse experiments and the BiFC method *in planta.* The large number and diversity of interactors found in the library screens reveal the power and the weakness of the system. As the screen is independent of the protein function very diverse proteins can be identified in this assay as interactors. The list of interactors is too large to be confirmed *in planta* in detailed follow-up experiments without removing biological irrelevant interactions. As the

Arabidopsis proteins are expressed heterologously in yeast we expect to come across interactions which will not take place *in planta.* The simultaneous expression in one cell of two proteins might enable them to interact whereas *in vivo* these proteins never occur at the same time in the same cell. Also differences in post translational modifications between Arabidopsis and yeast can influence protein interactions. Although follow-up experiments are required to confirm proteins interacting *in planta*, a number of false positives can be identified without additional experiments. Several proteins are known to cause autoactivation of the system and therefore cannot be tested with this method. Other proteins can be excluded based on their different subcellular localization or distinct gene expression profiles.
In the library screens performed with either the secondary CESA proteins or XXT1, several of the previously identified interactors have not been detected suggesting an unsaturated screen, which is also supported by the low frequency of each of the interactors. Several factors contributed to their absence in the 1.8 million colonies screened: i) the library contains only fragments between 1.2 and 2.5 kb, ii) only the clones comprising the full cDNA sequence in frame with the Cub will be able to generate a functional fusion protein.

The MbYTH has proven to be a powerful tool in the identification of interactions between protein involved in the synthesis of different polysaccharides in the plant cell wall. This strategy based on protein interaction allows the identification of new components of the biosynthetic machinery independently of their function. This method entails an approach to shorten the list of candidate genes involved in the biosynthesis of a specific polysaccharide. Additional experiments are required in order to identify the specific function within the polysaccharide biosynthesis.

References

Cavalier D.M., Lerouxel O., Neumetzier L., Yamauchi K., Reinecke A., Freshour G., Zabotina O.A., Hahn M.G., Burgert I., Pauly M., Raikehl N.V., and Keegstra K. (2008) Disruption of two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. Plant Cell 20: 1519-1537

Doblin M.S., Kurek I., Jacob-Wilk D., and Delmer D.P. (2002) Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol 43:1407–1420

Kurek I., Kaqagoe Y., Jacob-Wilk D., Doblin M., and Delmer D. (2002) Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains Proc Natl Acad Sci USA 99: 11109-11114

Nicol F., His I., Jauneau A., Vernhettes S., Canut H., Höfte H. (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. EMBO J 17: 5563-5576

Persson S., Wei H., Milne J., Page G.P., and Somerville C.R. (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci USA 102: 8633-8638

Roudier F., Fernandez A.G., Fujita M., Himmelspach R., Borner G.H.H., Schindelman G., Song S., Baskin T.I., Dupree P., and Wasteneys G.O. (2005) COBRA, an *Arabidopsis* extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. Plant Cell 17: 1749–1763

Schneider D., Finger C., Prodöhl A., Vokmer T. (2007) From interactions of single transmembrane helices to folding of α-helical membrane proteins: Analyzing transmembrane helix-helix interactions in Bacteria. Curr Prot Pept Sci 8: 45-61

Szyjanowicz P.M.J., McKinnon I., Taylor N.G., Gardiner J., Jarvis M.C., Turner S.R. (2004) The *irregular xylem2* mutant is an allele of KORRIGAN that affects the secondary cell wall of *Arabidopsis thaliana*. Plant J 37: 730-730

Timmers J., Vernhettes S., Desprez T., Vincken J.-P., Visser R.G.F., and Trindade L.M. (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett 583: 978-982

Zabotina O.A., Ven van de W.T.G., Freshour G., Drakakaki G., Cavalier D., Mouille G., Hahn M.G., Keegstra K., and Raikhel N.V. (2008) Arabidopsis XXT5 gene encodes a putative α-1,6-xylosyltransferase that is involved in xyloglucan biosynthesis. Plant J 56: 101-115

Supplementary data

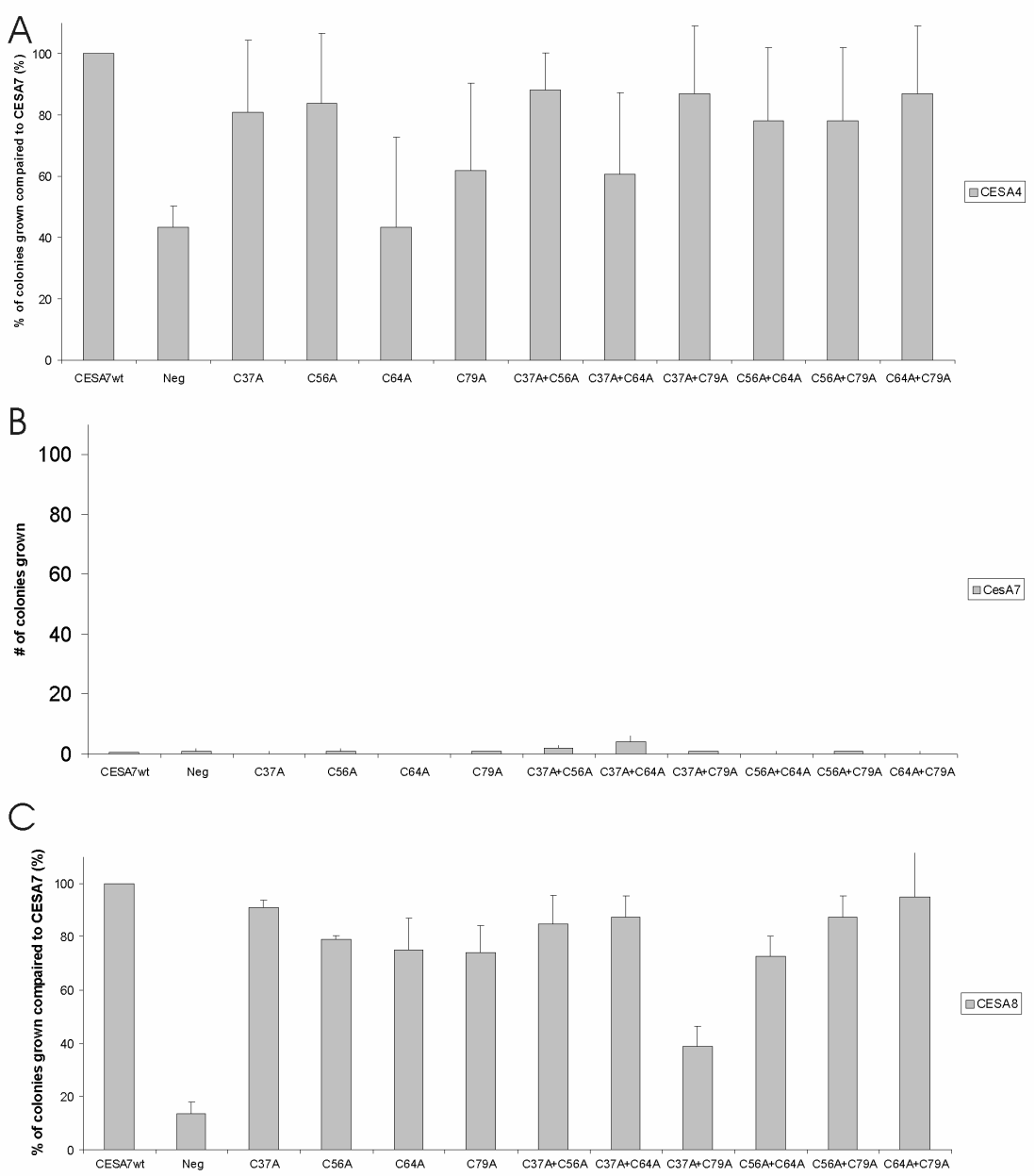


Figure S1. Effects of the substitution mutations on the interactions between the different CESAs. Yeast expressing CESA4 (A), CESA7 (B), and CESA8 (C) as bait with the wild type CESA7 (CESA7wt), NubG-ALG5 (Neg), and the different mutated CESA7 proteins as prey, that show visible growth after 5 days at 30°C on selective medium as a percentage of the interaction with wild type CESA7 protein.

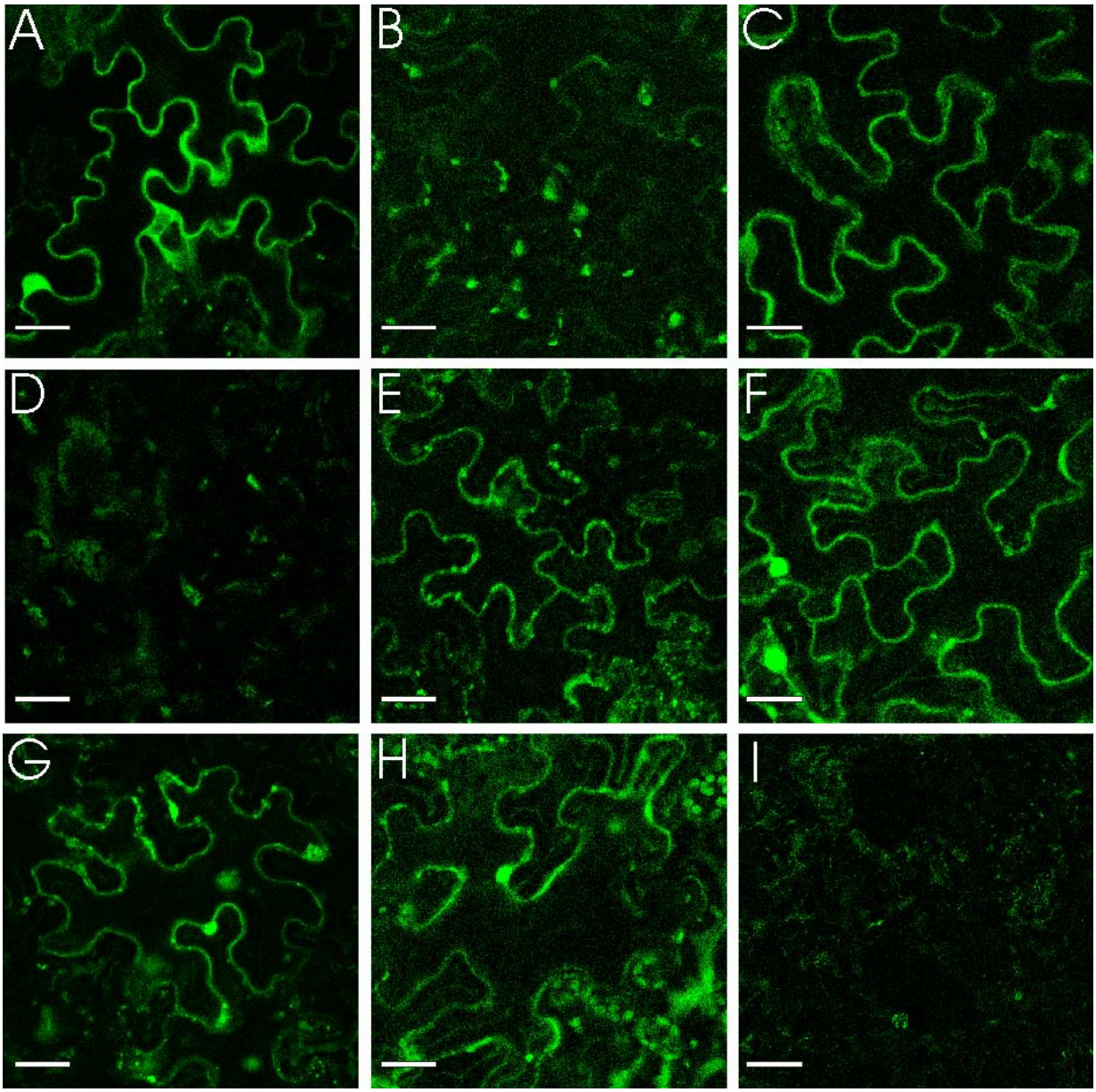


Figure S2. BiFC in *N. benthamiana* shows in vivo dimerization between the secondary CESAs. Homodimerization of CESA4 (A), CESA7 (D) and CESA8 (G) and heterodimerization YFP/N-CESA4/YFP/C-CESA7 (B), YFP/N-CESA7/YFP/C-CESA4 (C), YFP/N-CESA4/YFP/C-CESA8 (E), YFP/N-CESA8/YFP/C-CESA4 (F), YFP/N-CESA7/YFP/C-CESA8 (H), and YFP/N-CESA8/YFP/C-CESA7. Scale bar = 100µm

Table SI. Primers used for cloning and Site-Directed Mutagenesis. Restriction sites used for cloning are underlined.

Primer name	Sequence(5'-3')
FwCESA4SstII	AAACCGCGGATGGAACCAAACACC
RvCESA4SpeI	AAACTAGTTAACAGTCGACGCCACA
FwCESA7SstII	AAGACCGCGGATGGAAGCTAGCGCCGGTCTTGT
RvCESA7Eco47III	AAGAAGCGCTTCAGCAGTTGATGCCACACTTG
FwCESA8PstI	AAGACTGCAGAATGATGGAGTCTAGGTCTCCC
RvCESA8NcoI	AGAACCATGGCATTAGCAATCGATCAAAAGACAGTTC
FwCESA4XhoI	AAACTCGAGATGGAACCAAACACCATG
RvCESA4XhoI	AAACTCGAGTTAACAGTCGACGCCA
FwCESA7SfiI	AAGAGGCCATTACGGCCATGGAAGCTAGCGCCGGTCTTGT
RvCESA7SfiI	AAGAGGCCGAGGCGGCCATCAGCAGTTGATGCCACACTTG
FwCESA8SfiI	AAGAGGCCATTACGGCCATGATGGAGTCTAGGTCTCCC
RvCESA8SfiI	AAGAGGCCGAGGCGGCCATTAGCAATCGATCAAAAGACAGTTC
CESA7C37A	CTAGATGGACAATTCGCTGAGATCTGTGGAGATCAGATTGG
CESA7C56A	GACCTCTTCGTAGCTGCCAATGAGTGTGGTTTTCCGGCG-
CESA7C64A	GTGGTTTTCCGGCGGCTAGACCTTGCTATG
CESA7C79A	AAGGAACACAAAACGCTCCTCAGTGTAAGACTCG
FwCesA4gw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCC ATGGAACCAAACACCATGGCC
RvCesA4gw	GGGGACCACTTTGTACAAGAAAGCTGGG TTCAACTTAACAGTCGACGCCACATTGC
FwCesA7gw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGAAGCTAGCGCC GGTCT
RvCesA7gw	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAACTCAGCAGTTGATG CCACAC
FwCesA8gw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGATGGAGTCTAGG TCTCCC
RvCesA8gw	GGGGACCACTTTGTACAAGAAAGCTGGGTTAGCAATCGATCAAAAG

Table SII. List of all the proteins found to interact with the three CESA proteins. The list of proteins found interacting with the secondary CESA proteins in yeast. Bait indicates the CESA protein which was used as bait to identify the interactor (between bracktes it the number of times the interactor was found), Filter I (F. I) indicates the proteins known to be auto-activators of the system, whereas Filter II (F. II) indicates proteins that do not localize at the same intracellular localization as the CESA proteins.

Group	Locus	Name	Cellular location	Function		Bait	F. I	F. II
Cell wall metabolism	At2g20750	AtExpB1	cell wall	expansin		7		
	At2g42840	PDF1	cell wall	protodermal factor 1	embryogenesis	7(2) 8(1)		
	At1g56700			pyrolidone-carboxylate peptidase family	assimilate ammonium	8(2)		
	At5g40390	SIP1	endo membrane	seed imbibition 1-like / glycohydrolase family 36	GH family 36	4		
	At4g14130	AtXTH15	endo membrane	xyloglucan endotransglycosylase hydrolase	GH family 16	4		
	At2g06850	AtXTH4	cell wall	xyloglucan endotransglycosylase hydrolase	GH family 16	4(1) 8(2)		
	At2g05790		endo membrane	endo-glucanase	GH family 17	8		
	At2g32990	AtGH9B8	endo membrane	endo-1,4-glucanase	GH family 9	8		
	At1g75680	AtGH9B7	endo membrane	endo-1,4-glucanase	GH family 9	8		
	At3g16920	CTL2	endo membrane	endo-chitinase	GH family 19	7		
	At3g27540		membrane		GT family 17	8		
	At5g47780	GAUT4	membrane	galacturonosyltransferase	GT family 8	8		
	At3g61130	GAUT1	membrane	(alpha-1,4)-galacturonosyltransferase activity	GT family 8	8(2)		
	At5g60920	COBRA	membrane	glycosylphosphatidyl-inositol anchored protein	cellulose microfibril organization	4		
	At2g01610		membrane	invertase pectinmethylesterase inhibitor	pectin metabolism	8		
	At4g13660			pinoresinol reductase	lignan biosynthesis	4		

	At1g28580		endo membrane	carboxylic ester hydrolase activity	pectin metabolism	7		
Transporters	At4g17340	DELTA-Tip2	membrane	delta tonoplast integral protein	water channel	4(2) 8(2)		
	At2g37180	RD28	membrane	plasma membrane intrinsic protein	water channel	8		
	At3g61430	PIP1A	membrane	plasma membrane intrinsic protein	water channel	7(1) 8(1)		
	At1g01620	PIPC1	membrane	plasma membrane intrinsic protein	water channel	8(2)		
	At3g53420	PIP2A	membrane	plasma membrane intrinsic protein	water channel	4(1) 8(1)		
	At3g26520	TIP2	membrane	Tonoplast intrinsic protein 2	water channel	4(2) 8(6)		
	At2g39010	PIP2	membrane	water channel	water channel	7(2)		
	At4g35100	PIP3	membrane	water channel	water channel	7(7) 8(1)		
	At3g16240	DELTA-Tip	membrane	water channel	water channel	4(4) 7(2) 8(2)		
	At2g45960	PIPB1	membrane	water channel	water channel	4		
	At1g76850	SEC5A	plasmamembrane	exocyst complex component	vesicle transport	7		
	At5g58060	YKT61	trans Golgi	membrane fusion	vesicle transport	7		
	At1g06400	AtRab11E		regulation of vesicular trafficking		8		
	At1g76520		endo membrane	auxin efflux carrier family protein	membrane trafficking	4		
	At2g34250	SEC61	membrane	protein translocase activity	membrane trafficking	7		
	At1g22530	SEC14	membrane	transporter activity	membrane trafficking	8		
	At4g27540	PRA1		arginine N-methyltransferase	membrane trafficking	7	‡	

	At1g14660	AtNHX 8	membrane	Na+/H+ exchanger	membrane trafficking	8		
	At5g47030		mitochondrial		hydrogen ion transporting ATP synthase	8	‡	‡‡
	At2g10940		chloroplast	protease inhibitor	Lipid transfer protein family	4(1) 7(1)		‡‡
	At1g17530	ATTIM 23-1	mitochondrial membrane	protein translocase		4		‡‡
	At3g26570	PHT1	chloroplast	phosphate transporter 2		8		‡‡
	At3g45600	TET3	endo membrane	tetraspanin3		8		
Trans-cription factor	At1g06040	AtSTO		salt tolerance zincfinger	transcription factor	7(1) 8(1)		
	At5g44190	GLK2	nucleus	chloroplast development	Transcription factor	7(2)		‡‡
	At3g11400	EIF3G1		translation initiation factor 3G1	transcription factor	8		
	At2g21320		endo membrane	zinc finger B-box type	transcription factor	8		
	At1g69570			DOF-type zinc finger containing protein	transcription factor	8		
	At3g22840	ELIP1		early light inducable protein	chlorophyll binding	7		
	At1g15380			lactosyl glutathionel yase	detoxification	7		
	At1g51950	IAA18	nucleus	indoleacetic acid-induced protein	transcription factor	8		‡‡
	At2g77080	MAF1	nucleus	MADS affecting flowering	transcription factor	8		‡‡
	At3g01470	ATHA B-1	nucleus	homeobox-leucine zipper protein HAT5	transcription factor	8(2)		‡‡
	At1g77080	MAF1	nucleus	MADS affecting flowering	transcription factor	8		‡‡
	At4g01280		nucleus	MYB transcription factor		4		‡‡
	At3g32605		mitochondrion		transcription factor	8		‡‡
	At3g16770	AtEBP	nucleus		transcription factor	8		**
	At5g10960	CCR4N OT	nucleus	transcription complex		7		‡‡

Protein metabolism	At5g01020			protein kinase		4		
	At5g60390	AF-1-alpha	nucleus	elongation factor		7(2)		‡‡
	At4g20360	AtRABE1	chloroplast thylakoid membrane	Rab GTPasw homolog E1b	translation elongation	8		‡‡
	At1g07930	EF-1-alpha	mitochondrion	elongation factor		7		‡‡
	At1g14320	60S	cytosol	60S acidic ribosomal protein		8		
	At3g66654		chloroplast	protein folding		7		‡‡
	At4g33410		endo membrane	signal peptide peptidase family		8	‡	
	At4g02890	UBQ14		protein interaction		4	‡	
	At4g39093	RD19	endo membrane	responsive to dehydration peptidase	protein metabolism	8(2)		
	At3g63490			ribosomal protein L1 family	protein metabolism	7(1) 8(1)		
	At1g58684	40Sribs2		ribosomal protein	protein metabolism	8		
	At3g47360	AtHSD3	endo membrane	short-chain dehydrogenase/reductase (SDR) family protein	protein metabolism	8		
	At4g13180	AIS3		short-chain dehyrogenase/ reductase SDR family	protein metabolism	8		
	At5g18140	DNAJ		heat shock N-terminal domain containing	protein metabolism	8(2)		
	At1g08570			thioredoxin family	protein metabolism	8		
	At5g20050			protein kinase		7		
	At5g04530		endo membrane	beta-ketoacyl-CoA synthase family	protein metabolism	8		
	At5g64960	CDKC2	nucleus	cyclin-dependent kinaseC2	Kinase	7(2)		‡‡

	At1g80440			kelch repeat-containing F-box family	protein metabolism	8		
	at3g53870	RPS3B	cytosol	ribosomal protein		4		
	At3g52190	PHF1	endoplasmic reticulum	phosphate transporter traffic facilitator 1		8		
	At3g09200	RPPOB	cytosolic ribosome	60S acidic ribosomal protein		8		
	At2g42810	PAPP5	cytoplasm	protein phosphatase 5		8		
	At5g65200		chloroplast	armadillo/beta-catenin repeat family	ubiquitin ligase	8(2)	‡	
Lipid metatbolism	At1g49660	AtCXE5		carboxylesterase 5	lipid synthesis	4		
	At4g19860	LACT		chlolesterol acyltransferase fam.	lipid synthesis	4		
	At3g21720			citrate lyase	lipid synthesis	8		
	At2g03980			GDSL-motif lipase/hydrolase family	lipid synthesis	8		
	At5g13640	ATPDAT	membrane	Phospho-lipid diacyl-glycerolacyl transferase	lipid synthesis	8		
Stress response	At1g78380	AtGSTU19	cytoplasm	glutahione transferase	stress management	7		
	At5g27380	GSH2	cytosol	glutathione synthase	stress management	7		
	At2g37130	PER21	endo membrane	peroxidase activity	defence response	8(2)		
	At3g16420	PBP1	cytosol	PYK10-binding protein1	defence response	4(1) 7(2) 8(2)		
	At1g76680	OPR1		12-oxophytodienoate reductase 1	jasmonate forming	7		
	At2g32150			dehalogenase-like hydrolase family protein	hydrolase activity	7		
	At1g62380	ACO2		1-aminocyclopropane-1-carboxylate oxidase	Ethylene forming	8		

	At3g32980	PER32		peroxidase activity	oxidative stress response	8		
	At5g57970			methyladenosine glycosylase fam.	DNA methylation	8		
Sugar metabolism	At2g39730	RCA	chloroplast	rubisco activase		8		‡‡
	At2g44350	ATCS	mitochondrion	citrate synthase 4		8		‡‡
	At5g67590	FRO1	mitochondrion	FROSTBITE1	NADH dehydrogenase	7		‡‡
	At5g07440	GDH2	mitochondrion	Glutamate dehydrogenase 2		8		‡‡
	At2g21170	TIM		triosephosphate isomerase	glycolysis	8		
	At1g71170			6-phosphogluconate dehydrogenase NAD-binding domain containing	glycolysis	8		
	At5g09660	PMDH2	chloroplast	peroxisomal NAD malate dehydrogenase 2	glycolysis	7(1) 8(1)		‡‡
	At4g38970		chloroplast stroma	fructose-biphosphate aldolase	glycolysis	4		‡‡
	At5g06340		chloroplast	5,5-P1-P4-tetraphosphate hydrolase		8		‡‡
	At3g26650	GAPA	chloroplast	glyceraldehyde 3-phosphate dehydro-genase A subunit	glycolysis	4(1) 7(2) 8(1)		‡‡
	At2g36530	LOS2	mitochondrion	phosphoryryvate hydratase	glycolysis	7		‡‡
	At3g15020	NAD	mitochondrion	malate dehydrognase		7		‡‡
Unknown	At1g60010			unknown		4		
	At4g38280			unknown		4		
	At4g25670	CPuORF12		unknown		8		
	At1g62780			unknown		4		
	At5g16110			unknown		7		
	At3g15450			unknown		4(1) 8(1)		

At5g42765			unknown		8		
At1g19990			unknown		4		
At4g27450			unknown		4		
At1g18490	DUF1637 fam		unknown		8		
At2g20670			unknown		4(2)		
At1g19400			unknown		7		
AT1g11440			unknown		4		
At1g18740	DUF793 fam		unknown		7		
At3g59300			unknown		8		
At3g49140			unknown		8		
At3g05280			unknown Yip! Family		4		
At3g60800	DHHC type		unknown Zinc finger protein		7		
At4g38160	PDE191		pigment defective 191		4(1) 8(1)		
At4g32460		endo membrane	unknown		8		
At5g55940	EMB2731	endoplasmic reticulum	embryo defective 2731		4		
At4g08950	PHI1	cell wall	phosphate induced protein (PHI1)		8		
At5g14430		cell wall	dehydatation responsive		7(2)		
At4g38770	PRP4	cell wall	proline rich protein		4		
At4g20260	DREPP	membrane	polypeptide family protein		4		
At5g59500		membrane	unknown		7		
At1g50020		chloroplast thylakoid membrane	unknown		8		‡‡
At2g14910		chloroplast	unknown		8		‡‡
At3g21200		chloroplast	unknown		8		‡‡
at5g21920	YGGT	chloroplast membrane	unknown function		4(1) 7(1)		‡‡
at3g32930		chloroplast	unknown		4(1) 8(1)		**
At4g23890		chloroplast thylakiod membrane	unknown		8		**
At1g55480		chloroplast thylakoid membrane			8		‡‡
At5g02500	HSC70-1		heat shock cognate 70kD protein	ATP binding	8	‡	

	At2g34560	ERH3		Ectopic Root Hair 3	ATP binding / katanin putative	4	‡	
	At4g32260		chloroplast membrane	ATP synthase family		4(2)	‡	‡‡
	At1g09070	SRC2	endoplasmic reticulum	soybean gene regulated by cold-2		8		
	At2g38750	ANNAT4	membrane	calcium ion binding		8		
	At3g07670		chloroplast	SET domain-containing protein		8		‡‡
	At5g03290		mitochondrion			8		‡‡
	At1g16880		chloroplast membrane	uridylyltransferase-related		4		‡‡
	At4g03030		chloroplast	kelch repeat-containing F-box family		8		‡‡
	At2g40600		chloroplast	appr-1-p processing enzyme family		8		‡‡
	At2g32150			dehalogenase-like hydrolase family protein		8		
	At1g16880		chloroplast membrane	uridylyltransferase-related		4		‡‡
Photo-synthesis	At2g34420	LHB1B2	chloroplast thylakiod membrane		light harvesting	8(2)		‡‡
	At1g61520	LHCA3	chloroplast			4		‡‡
	At1g15820	LHCB6	chloroplast thylakoid membrane	light harvesting complex PSII	chlorophyll binding	8		‡‡
	At2g05100	LHCB2.1	chloroplast thylakoid membrane	photosystem II light harvesting complex gene 2.1	chlorophyll binding	8		‡‡
	At3g61470	LHCA2	chloroplast thylakoid membrane		chlorophyll binding	4(2) 7(1) 8(2)		‡‡
	At5g14740	CA2	chloroplast thylakoid membrane	beta carbonic anhydrase 2	cabonate dehydrase activity zinc ion binding	4(1) 7(1) 8(1)		‡‡

	At5g38660	APE1	chloroplast thylakiod membrane	acclimation of photosynthesis to environment		8		‡‡
	At2g34430	LHB1B1	chloroplast	light harvesting complex		7(2) 8(1)		‡‡
	At3g54890	LHCA1	chloroplast	light harvesting		4		‡‡
	At3g47470	LHC4	chloroplast	chlorophyll binding		4		‡‡
	At5g13630	GUN5	chloroplast thylakiod membrane	genome uncoupled 5	magnesium chelatase activity	8(2)		‡‡
	At1g44575	NPQ4	chloroplast thylakoid membrane	NON photochemical quencing		8		‡‡
	At5g54270	LHC-B3	chloroplast	light harvesting complex		7		‡‡
	At1g30510	AtRFNR2	chloroplast	root FNR2	oxidoreductase	8(2)		‡‡
	At3g10920	MSD1	mitochondrion	manganese superoxide dismutase	removal of superoxide radicals	8		‡‡
	At2g25080	AtGPX1	chloroplast thylakoid membrane	gluthathione peroxidase 1	response to oxidative stress	8		‡‡
	At1g11360	USP		can bind ATP	response to stress	8	‡	
	At1g20620	CAT3	chloroplast	CATALASE 3	Catalase H2O2 breakdown	7		‡‡
	At1g78900			ATP binding hydrogen ion transport		4(1) 8(1)	‡	
	At1g30540			ATPase		8(2)	‡	
	At3g42050			ATP binding		8	‡	
	At2g20020	RAN2	cytoplasm	Ras related GTP-binding nuclear protein 2	GTP binding	8	‡	
	At3g18820	RABG3f/		Rab GTPase homolog G3f		8	‡	
Amino acid metabolism	At2g36880	MAT3/MTO3		mehionine Adenosyl transferase		7(1) 8(2)	‡	
	At1g17290	ALAAT1	mitochondrion	alanine aminotransferase		4	‡	‡‡
	At1g23310	GGT1	chloroplast	alanine-2-oxoglutarate aminotransferase		4	‡	‡‡

	At1g07780	PAI1	chloroplast	phosphoribosylanthranilate isomerase 1		8		‡‡
	At5g49020		chloroplast	arginine N-methyltransferase fam.		7		‡‡
	At5g17920	AtCIMS	cytosol	methionic biosynthesis		7(2)	‡	
	At4g13940	HOG1		adenosyl homocysteinase activity	Amino acid biosynthesis	7(1) 8(1)	‡	
	At5g63570	GSA1		porphyrin biosynthetic pro		4	‡	

Table SIII. List of all proteins interacting with XXT1. Filter I indicates protein known to be auto-activators of the system, filter II indicates proteins that do not localize at the same intra-cellular location as the XXT1 protein. # indicates the number of times the protein was found.

Locus	Protein name	Protein name	Putative function	Localization	#	Filter 1	Filter 2
At1g61250	SCAMP3	secretory carrier 3		post Golgi / cytosol	2x		
At4g15780	ATVAMP724	vesicle-associated membrane protein 724		post Golgi / plasma-membrane	2x		
At5g09440		phosphate-responsive protein			1x		
At1g20630	CAT1	Catalase 1	H202 to water en oxygen		1x		
At1g78630	EMB1473	embryo defective 1473	structural constituent of ribosome		1x		
At3g26650	GAPA	glyceraldehyde 3 phosphate dehydrogenase A subunit			1x		
At4g13940	HOG1	homology-dependent gene silencing 1	adenosyl homocysteinase		1x		
At2g38080	LAC4	laccase 4	copper ion binding/ oxidoreductase		1x		
At4g37540	LBD39	LOB domain-containing protein 39			1x		
At1g80070	SUS2	abnormal suspensor2	transcription factor		1x		
At1g12780	UGE1	UDP-D-glucose / UDP-D-galactose-4-Epimerase 1			1x		
At4g28850	AtXTR18	xyloglucan endo-transglycosylase hydrolase	xyloglucosyl transferase		1x		
At4g14130	AtXTH15	xyloglucan endo-transglycosylase hydrolase	xyloglucosyl transferase		2x		
At1g09340	AtCSP41B		RNA-binding	chloroplast	1x		‡‡
At2g34430	LHB1B1/ CAB1	photosystem II light harvesting complex gene	chlorophyll binding	chloroplast	5x		‡‡
At1g19150	LHCA2	photosystem I light harvesting complex	chlorophyll binding	chloroplast	1x		‡‡
At1g61520	LHCA3	photosystem I light harvesting complex gene 3	chlorophyll binding	chloroplast	1x		‡‡

At3g47470	LHCA4/CAB4	photo system I light harvesting complex gene 4	chlorophyll binding	chloroplast	2x		‡‡
At1g15820	LHCB6	light harvesting complex PSII		chloroplast	1x		‡‡
At1g67090	RBCS1A	ribulose-biphosphate carboxylase		chloroplast	1x		‡‡
At1g54780	TLP18.3	thylakoid lumen 18.3kDA protein		chloroplast	1x		‡‡
At4g38970		fructose-biphosphate aldolase putative		chloroplast	6x		‡‡
At2g17340			pantothenate kinase related	chloroplast	1x		‡‡
At2g42810	PAPP5/ PP5	protein phosphatase 5	dephosphorylates active Pfr-phytochromes	cytoplasm/ER membrane	1x		‡‡
At1g07940	EF-1-alpha	elongation factor	protein synthesis	cytosol	2x		‡‡
At3g16420	PBP1	PYK10-binding protein 1		cytosol	2x		‡‡
At4g08950	EXO	phosphate-responsive protein		extra cellular	2x		‡‡
At2g13540	ABH1	ABA hypersensitive 1		nucleus	1x		‡‡
At5g43280	AtDCI1	Delta (3,5), Delta (2,4) dienoyl-coa isomerase 1	CoA isomeras / enoyl-CoA hydratase	peroxisome	1x		‡‡
At4g17340	Delta-tip2/TIP2	tonoplast intrinsic protein2	water channel	plasma-membrane	2x		‡‡
At5g60660	PIP2;4/ PIP2F		water channel	plasma-membrane	2x		‡‡
At1g71340		Glycerophosphoryl diester phosphodiesterase	Cell wall organisation	plasma-membrane	2x		‡‡

Summary

The plant cell wall provides a rigid support that allows the plant to stand upright and acts as a barrier against invading organisms. The cell wall does not simply constitute the physical confinement of the cell, it is a highly dynamic structure with great importance for growth and development, cell to cell communication, and transport processes. In addition, cell wall polymers make up most of the plant biomass and provide the raw material for many economically important products including food, feed, bio-materials, chemicals, textiles, and biofuel. This broad range of functions and applications make the biosynthesis of the plant cell wall a highly interesting target of scientific research.

Several approaches, like reverse genetics and expression profiling, have previously been used to identify the key components in the biosynthesis of the cell wall. This revealed many glycosyltransferase and especially the identification of the CESA superfamily was a big step forward in plant cell wall research. Due to the huge number of genes and diversity of the polysaccharides constituting the plant cell wall it has been difficult to link enzymes to the biosynthesis of a specific polysaccharide. Furthermore, not all mutants in cell wall genes show a visible phenotype at the organ or organism level due to the redundancy of polysaccharide and gene functions. Furthermore, while expression analysis might indicate the involvement of a gene in cell wall biosynthesis, it is difficult to associate it to biosynthesis of a specific polysaccharide. Therefore a different approach was chosen to obtain new insights in the plant cell wall biosynthesis.

As there are several indications that the cell wall polysaccharides are synthesized in a protein complex we used a protein-protein interaction strategy to gain insight in the cell wall biosynthesis of *Arabidopsis thaliana* and to identify additional genes involved in this process. Although there are several methods available to identify interactions between proteins involved in cell wall biosynthesis, the membrane based yeast two hybrid (MbYTH) system was chosen as this method is able to identify the interaction between membrane bound proteins in a high throughput manner. Before this method could be used for the characterization of protein complexes involved in cell wall biosynthesis several practical improvements were incorporated and a more efficient protocol has been proposed (Chapter 2). Using this protein-protein interaction assay several distinct goals have been reached in this thesis, i) the characterization of the rosette structure by resolving the organization of the

different cellulose synthase proteins in the complex, ii) the identification of unknown components of the cellulose synthezing machinery, iii) the confirmation of a xyloglucan synthesizing complex and the identification of several of its components.

At the start of the research, it was already known that the three different CESA proteins (#4, #7, and #8) are the main constituents of the rosette complex synthesizing cellulose in the secondary cell wall, however, the organization of these proteins within the complex was thus far unknown. Using the MbYTH system the interactions between the different secondary CESA proteins was tested and it was found that all proteins were able to interact with each other but only CES4 was able to form homodimers. Based on these results a model was proposed in which each of the CESA proteins has its position in the complex and the homo-dimerization found for CESA4 is responsible for interaction between the subunits (Chapter 3). Site directed mutagenesis has been used to study the involvement of RING-finger domain in the interaction between the CESA proteins, the results revealed that this domain is not essential for the interaction between CESA proteins.

Thus far no other proteins have been identified as components of the rosette structure, next to the CESA protein, although several proteins have been implicated in cellulose biosynthesis. One of these proteins is the cellulase KORRIGAN, as the mutant phenotype of KOR1 showed decreased amounts of cellulose in both the primary and secondary cell wall, compared to wild type plants. The interaction between the cellulase KORRIGAN (KOR1) and the CESA proteins was tested, and it was found that KORRIGAN could bind with three different primary CESA proteins whereas it could only bind to two of the secondary CESA proteins (Chapter 4). The interactions between the CESA proteins and truncated KOR1 proteins revealed that the transmembrane domain of KOR1 is essential for interaction between these proteins. Additionally, it was found that KOR1 could form homodimers.

In order to identify additional proteins involved in cellulose biosynthesis a cDNA library screen for interactors with the different secondary CESA proteins was performed. A large number of proteins were shown to be able to interact with the different CESA proteins in yeast (Chapter 5). Some of these proteins have been previously implicated in cell wall biosynthesis, others can be directly linked to cellulose biosynthesis based on their function, whereas others need additional research to reveal their function in this process.

In addition to cellulose, xyloglucan is a major constituent of the primary cell wall. There are several indications that xyloglucan is synthesized in a protein complex, the regular structure

of this polysaccharide. To confirm this idea the interaction between proteins known to be involved in xyloglucan synthesis is tested. It was found that two different xylosyltransferases (XXT1 and XXT2) were able to interact with each other and the glucosyltransferase (CSLC4), which synthesizes the glucan backbone of xyloglucan (Chapter 6). In addition, several new candidate genes involved in xyloglucan biosynthesis have been identified in a genome wide interaction study with the xylosyltransferase XXT1. Several proteins like xyloglucantransferase, which might be involved in removal of the growing xyloglucan from the synthesizing complex, and UDP-glucose epimerase is thought to supply the substrate for the addition of a galactose site chain have been physically linked to the xyloglucan synthesizing complex.

On the whole, this work has generated an effective tool in cell wall research and identified new players in the biosynthesis of both cellulose and xyloglucan

(Chapter 7).

Samenvatting

De celwand verschaft de cel een starre versterking die het de plant mogelijk maakt om te staan en fungeert als een barrière tegen binnendringende organismen. De celwand vormt niet alleen de fysieke grens van de cel, is het een dynamische structuur met groot belang voor de groei en ontwikkeling, cel aan cel communicatie, en transport processen. Bovendien vormen de celwand polymeren het grootste deel van de planten biomassa en verstrekken de grondstoffen voor vele economisch belangrijke producten met inbegrip van voedsel, voer, biologisch materialen, chemische producten, textiel, en biofuels. Dit brede spectrum van functies en toepassingen maken de biosynthese van de plantencelwand een hoogst interessant onderwerp van wetenschappelijk onderzoek. Verscheidene benaderingen, zoals reverse genetica en expressie onderzoek, zijn eerder gebruikt om de belangrijkste componenten in de biosynthese van de celwand te identificeren. Dit leverde vele glycosyltransferases op en vooral de identificatie van CESA superfamily was een grote stap voorwaarts in het onderzoek van de planten celwand. Vanwegen het enorme aantal genen en diversiteit van de polysacchariden die de planten celwand vormen is het moeilijk geweest om enzymen met de biosynthese aan een specifiek polysaccharide te verbinden, aangezien niet alle mutaties in celwandgenen een zichtbaar fenotype op orgaan of organismeniveau vertonen. Terwijl de expressie analyse op de betrokkenheid van een gen in celwandbiosynthese zou kunnen wijzen, is het moeilijk om deze aan biosynthese van een specifiek polysaccharide te verbinden. Daarom werd een andere benadering verkozen om nieuwe inzichten in de biosynthese van de planten celwand te verkrijgen. Aangezien er verscheidene aanwijzingen zijn dat de celwand polysacchariden in eiwit complexen worden gevormd is een eiwit-eiwitinteractie strategie gebruikt om meer inzicht in de celwandbiosynthese van *Arabidopsis thaliana* te verkrijgen en extra genen te identificeren die betrokken zijn bij dit proces. Hoewel er verscheidene methodes beschikbaar zijn om de interactie tussen proteïnen te identificeren, werd het membraan gebonden gist twee hybride systeem (MbYTH) gekozen aangezien deze methode de interactie tussen membraan gebonden eiwitten op een snelle manier kan identificeren. Verscheidene praktische verbeteringen en een efficiënter protocol zijn voorgesteld voor de karakterisatie van eiwit complexen betrokken bij celwand biosyntheses (Hoofdstuk 2). Met behulp van deze eiwit-eiwitinteractie analyse zijn verscheidene verschillende doelstellingen bereikt in dit proefschrift, i) de karakterisering van de organisatie van de rozetstructuur door de interacties van de verschillende proteïnen van cellulose synthase in complex aan te tonen, ii) de identificatie van onbekende componenten van de cellulose synthezing machines, iii) de bouw van het xyloglucan producerende complex en de identificatie van verscheidene van de componenten. Bij het begin van het onderzoek, wist men reeds dat de drie verschillende proteïnen CESA (#4, #7, en #8) de belangrijkste constituenten van de complexe dat de cellulose produceert in de secundaire celwand, echter, de organisatie van deze proteïnen binnen het complex is tot dusver onbekend was. Met behulp van het systeem MbYTH werden de interactie tussen de verschillende secundaire CESA eiwitten getest wat uitwees dat al deze eiwitten met elkaar konden interacteren maar alleen CES4 kon homodimers vormen. Gebaseerd op deze resultaten werd een model voorgesteld met daar in een positie voor elk van de CESAs en de homo-dimerizatie gevonden voor CESA4 is verantwoordelijk voor interactie tussen de verschillende subunits (Hoofdstuk 3). Mutagenese was gebruikt om de betrokkenheid van RING-vinger domein in de interactie tussen de verschillende CESAs te bestuderen, de resultaten toonde aan dat dit domein niet essentieel voor de interactie tussen deze CESA eiwitten. Tot zover zijn geen andere proteïnen

geïdentificeerd als componenten van de rozetcomplex, naast de CESAs, hoewel verscheidene andere eiwitten betrokken zijn bij cellulose biosynthese. Één van deze proteïnen is de cellulase KORRIGAN (KOR1), aangezien het mutante fenotype van KOR1 een vermindering van hoeveelheid cellulose, in zowel de primaire als secundaire celwand, laat zien in vergelijking met wild type planten. De interactie tussen de cellulase KOR1 en de CESA eiwitten is getest en de resultaten lieten zien dat KOR1 met drie verschillende primaire CESAs kon binden terwijl het slechts aan twee van de secundaire CESA eiwitten kon binden (Hoofdstuk 4). De interacties tussen de CESAs en partiële KOR1 eiwitten toonde aan dat het transmembraan-domein van KOR1 voor interactie tussen deze proteïnen essentieel is. Bovendien, is aangetoond dat KOR1 homodimers kon vormen.

Om extra proteïnen te identificeren die betrokken zijn bij cellulose biosynthese werd een cDNA bank gescreend voor interactors met de verschillende secundaire proteïnen CESA. Een groot aantal proteïnen werd gevonden die met de verschillende CESA eiwitten een interactie aan konden gaan in gist (Hoofdstuk 5). Sommige van deze proteïnen hadden aanvullende indicaties dat zij zijn betrokken bij de celwand biosynthese, anderen konden direct met cellulosebiosynthese worden verbonden op basis van hun functie, terwijl anderen extra onderzoek nodig hebben om hun functie in dit proces te duidelijk te maken.

Naast cellulose is xyloglucan een belangrijke component van de primaire celwand. Er zijn verscheidene aanwijzingen dat xyloglucan in een eiwit complex wordt geproduceerd, bijvoorbeeld de regelmatige structuur van dit polysaccharide. Om deze hypothese te bevestigen werd de interactie tussen de eiwitten de betrokken zijn de biosynthese van dit polysaccharide getest. Twee verschillende xylosyltransferases (XXT1 en XXT2) en een glucosyltransferase (CSLC4) konden met elkaar een interactie aangaan (Hoofdstuk 6). Bovendien zijn verscheidene nieuwe kandidaat genen betrokken bij xyloglucan biosynthese geïdentificeerd, in interactie studie met de xylosyltransferase XXT1. Verscheidene eiwitten zoals xyloglucantransferase, die zou kunnen worde geïmpliceerd met het verwijderen van de groeiende xyloglucan keten van het complex, en een UDP-Glucose epimerase, welke het substraat voor galactose zijketen zou kunnen te leveren, vertoonde interactie met XXT1. In het algemeen, heeft dit werk een efficiënt hulpmiddel in celwandonderzoek opgeleverd en verscheidene nieuwe componenten geïdentificeerd in de biosynthese van zowel cellulose en xyloglucan (Hoofdstuk 7).

Acknowledgments

As I am not a man of many words I will try to keep this short; I would like to thank everybody who played a role in this PhD thesis, yes this means you, THANK YOU.

However there are several people I want to mention by name. First and foremost I would like to thank my supervisors Jean-Paul Vincken and Luisa Trindade. Jean-Paul, as initiator of this research you allowed me to develop my own ideas and supported my decisions and with your endless scientific enthusiasm, you helped me to give it yet another go, in the numerous attempts in setting up the MbYTH system. Luisa you helped me to fill in the blanks and to get the most out of the results obtain, again and again. Both of you probably will not know how much your support has meant for my. It is little to say that I could never have realized this work and write this thesis without your support, ideas and suggestions.

Also I would like to thank Richard Visser who gave me the opportunity to do my PhD research at the laboratory of Plant breeding, you had the difficult task of convincing me of the significance of my results in the final stages of this research.

I thank Elly, Irma and Carin for their direct involvement in my research and the efforts made to overcome difficulties in the lab. But also all other the technicians in the laboratory to keep the lab smoothly running and the small talk to get my mind of things. Marian thanks for all the holiday suggestions, they are all on my to do list.

I also thank all the members of “Biobased Economy” group, although several of you already have left some time ago, like Wole, Monique, Farhad and Katja, you all left a remaining impression on me, each in your own way. Xingfeng, thanks for the times you pulled me away from the computer, I sometime really needed that. Nasim, I must say I really enjoyed answering your “one question”, it is hard to say who benefited more from this. The members of the laboratory of Plant cell biology thanks for our friday meetings, you have helped me to keep an open mind for different views on the cell wall biosynthesis.

Of course I also want to thank Samantha Vernhettes and Thierry Desprez for the fruitful collaboration which resulted in one but hopefully in the near future two published articles.

Er bestaat ook een leven buiten de wetenschap en ik wil ook iedereen buiten mijn werkomgeving bedanken voor het begrip en de belangstelling voor mijn project. De opmerking “houdt moed, hout moet” is het altijd goed blijven doen als ik weer eens aan het vertellen was over ‘mijn blauwe zaadjes’. Ook moet ik mijn ouders bedanken voor hun steun en vertrouwen, van jullie heb ik geleerd om te streven naar het hoogst haalbare.
Als laatste, maar zeker niet als minste, wil ik Noortje bedanken voor haar begrip en geduld. Jij bent degene die altijd is blijven geloven dat ik dit proefschrift kon afronden, dit proefschrift is dus eigenlijk ook een beetje van jou.

As I did not mentioned your name pleas read the first paragraph of this section as it was written especially for you.

Curriculum Vitae

Johannes Fransicus Petrus Timmers (Jaap) was born on July 12, 1978, in Uden, the Netherlands. In July 1998 he obtained his diploma for the pre-university education at the Comenius College. In the same year he started the study biology at the University of Nijmegen, with specialization towards molecular plant biology, which in 2003 resulted in the degree of MsC. After his graduation he started a PhD at the laboratory of Plant Breeding at Wageningen University, the results of this research project are described in this thesis.

List of Publications

Timmers J., Vernhettes S., Desprez T., Vincken J.-P., Visser R.G.F., and Trindade L.M. (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett 583: 978-982

Timmers J., Vernhettes S., Desprez T., Vincken J.-P., Visser R.G.F., and Trindade L.M. (2009) Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. Abstract book Workshop “Systems biology for plant design” July 2009 page 60

Obembe O.O., Jacobsen E., Timmers J., Gilbert H., Blake A W., Knox J P. Visser R.G.F., and Vincken J.P. (2007) Promiscous, non-catalytic, tandem carbohydrate-binding modules modulate the cell-wall structure an development of transgenic tobacco(*Nicotiana tabacum*) plants. J Plant Res 120: 605-617

1) Start-up phase	*date*
First presentation of your project (higly recommended)	
Idendification of genes involved in cell wall biosynthesis	Sep 2003
Writing or rewriting a project proposal	
Writing a review or book chapter	
MSc courses	
Laboratory use of isotopes	
Subtotal Start-up phase	*1.5 credits*

2) Scientific Exposure	*date*
EPS PhD student days	
EPS PhD Student Day, Wageningen University	Sep 19, 2006
EPS PhD Student Day, Wageningen University	Sep 13, 2007
EPS Post-doc day	Sep 18, 2008
EPS theme symposia	
EPS Theme 3 symposium 'Metabolism and Adaptation', University of Amsterdam	Nov 10, 2006
EPS Theme 1 symposium 'Developmental Biology of Plants', Wageningen University	Oct 11, 2007
EPS Theme 3 symposium 'Metabolism and Adaptation', Wageningen University	Nov 06, 2007
NWO Lunteren days and other National Platforms	
NWO-ALW lunteren 2004	Apr 05-06, 2004
NWO-ALW lunteren 2005	Apr 04-05, 2005
NWO-ALW lunteren 2006	Apr 03-04, 2006
NWO-ALW lunteren 2007	Apr 02-03, 2007
Seminars (series), workshops and symposia	
Mathematics and Plant biology (5 seminars attended)	Apr-Dec 2005
Workshop RNA interference	May 27, 2004
Dr. Lucia Colombo	Jan 14, 2004
Prof.dr. Tom Gerats	Dec 16, 2004
Prof.dr. Candace Haigler	May 19, 2006
EU COST E50 Workshop "Systems Biology for Plant Design"	Jul 08-11, 2009
Seminar plus	
International symposia and congresses	
Gordon Research Conference 'Plant Cell Walls'	Jul 31-Aug 04, 2006
Presentations	
Bi-annual presentations, CRC-Wageningen (8 times)	2003 -2007
Poster: EU COST E50 Workshop, 'Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall'.	Jul 08-11, 2009
IAB interview	Sep 19, 2006
Excursions	
Subtotal Scientific Exposure	*14.7 credits*

3. In-Depth Studies	*date*
EPS courses or other PhD courses	
Summer School 'Glycosciences'	Jun 28-Jul 01, 2004
PhD course 'Protein Engineering'	Mar 08-10, 2004
PhD course 'Gateway to Gateway Technologies' (hands on)	Nov 20-24, 2006
Journal club	
Literature discussion of the Plant Breeding group	2003-2007
Individual research training	

Subtotal In-Depth Studies *6.4 credits*

4) Personal development	*date*
Skill training courses	
Presentations Skills	May, 2005
Scientific writing	Apr-May 2006
Organisation of PhD students day, course or conference	
Membership of Board, Committee or PhD council	

Subtotal Personal Development *3.8 credits*

Total number of credit points 26.4 credits