CHAPTER 3

THE CELLULOSE SYNTHASE SUPERFAMILY

HEATHER L. YOUNGS, THORSTEN HAMANN, ERIN OSBORNE, AND CHRIS SOMERVILLE*

Carnegie Institution, Department of Plant Biology, Stanford, CA 94305

Abstract

The completion of the *Arabidopsis thaliana* genome revealed ten cellulose synthase or *AtCESA* genes. Mutations in seven of the ten *AtCESA* genes have been studied. Studies indicate a requirement for three genes, *AtCESA1*, *AtCESA3*, and *AtCESA6*, in primary wall formation; whereas *AtCESA4*, *AtCESA7*, and *AtCESA8* may be involved in secondary cell wall formation. Genes with significant similarity to cellulose synthase-like (*CSL*) genes have been classified into eight distinct families. Thirty such genes have been identified in *Arabidopsis*. Members of the superfamily differ in their size, topology, and predicted physical properties.

Keywords

Arabidopsis, cellulose synthase (CESA), cellulose synthase-like (CSL), gene expression, predicted proteins.

Abbreviations

Arabidopsis thaliana (*At*), cellulose synthase (CES), constitutive expression of VSP1 (*cev*), cellulose synthase-like (CSL), ectopic lignin (*eli*), Fourier transform infrared (FTIR), *Gossypium hirsutum* (*Gh*), glycosyl transferase family II (GT-2), β-glucuronidase (GUS), kojak, a root hairless mutant (*kjk*), irregular xylem (*irx*), *Medicago truncatula* (*Mt*), *Nicotiana alata* (*Na*), isoelectric point (pI), procuste (*rc*), *Populus tremuloides* (*Pt*), resistance to *Agrobacterium tumefaciens* transformation (*rat*), radially swollen (*rsw*), the *Arabidopsis* information resource (TAIR), transmembrane domain (*T*MD), zinc binding domain (ZnBD).

1 INTRODUCTION

Cellulose is a simple polymer of unbranched β-1,4-linked glucan chains, which coalesce to form microfibrils. Extensive hydrogen bonding, between the glucan chains of the microfibrils and amongst the microfibrils themselves, yields a range of

* Author for correspondence: e-mail: crs@andrew2.stanford.edu

R.M. Brown, Jr. and I.M. Saxena (eds.), Cellulose: Molecular and Structural Biology, 35–48.
© 2007 *Springer.*

cellulose confomers that can form loose noncrystalline networks or robust crystalline structures. These provide a structural framework to the wall, which is crosslinked by hemicellulosic polymers and infiltrated with a dense pectic gel (Bacic et al. 1988; Carpita and McCann 2000). Noncellulosic polymers have relatively simple backbone structures that may be adorned with a varying complexity of carbohydrate branches (Table 3-1). It is likely that the polymers with β-linked homopolysaccharide backbones, such as that of xyloglucan, are synthesized by processive enzymes, whereas, the sugars comprising the branches are added to the backbone by nonprocessive enzymes, either following completion of the backbone chain or in concert with its biosynthesis (Henrissat et al. 2001; Perrin 2001). Polymers with α-1,4-linked backbone sugars, or heteropolysaccharide and mixed linkage backbones, such as the pectins (Table 3-3), are most likely synthesized by a different class of enzymes (Scheller et al. 1999).

Table 3-1. Basic structural composition of various cell wall polymers[a]

Polymer[a]	Backbone	Sidechains
Cellulose and Hemicelluloses		
Cellulose	β-1,4-glucan	None
Xylan	β-1,4-xylan	α-1,2-arabinose α-1,2-(4-O-methyl)-glucuronic acid
Xyloglucan	β-1,4-xylan	α-1,6-xylose α-1,2-fucosyl-β1,2-galactoysl-α1,6-xylose α-1,2-arabinosyl-α1,6-xylose
Mannan	β-1,4-mannan	α-1,6-galactose
Glucuronomannan	β-1,4-mannosyl-β-1,2-glucuronan	β-1,6-galactose α-1,3-arabinose
Glucomannan	β-1,4-glucosyl-(β-1,4-mannose)$_2$	α-1,6-galactose
Mixed-linked Glucan	β-1,3-glucosyl-(β-1,4-glucose)$_2$	None
Arabinogalactan II	β-1,3-galactan and β-1,6-galactan	β-1,3-arabinose
Pectins		
Galactan	β-1,4-galactan	None
Arabinogalactan I	β-1,4-galactan	(α-1,5-arabinosyl)$_2$-α-1,3-arabinose
Arabinan	α-1,5-arabinan	α-1,2-arabinose α-1,3-arabinose
Homogalacturonan	α-1,4-galacturonan	None
Xylogalacturonan	α-1,4-galacturonan	α-1,3-xylose
Rhamnogalacturonan I (RGI)	α-1,2-rhamnosyl-α-1,4-galacturonan	α-1,5-arabinan β-1,4-galactan arabinogalactan I
Rhamnogalacturonan II (RGII)	α-1,4-galacturonan	Various

[a]Brett and Waldron (1990); Carpita and McCann (2000).

2 IDENTIFICATION OF CELLULOSE SYNTHASE

Although cell-free synthesis of cellulose was claimed as early as 1964 (Karr 1976 and references therein), it was not possible to isolate the enzymes responsible for cellulose biosynthesis in plants by conventional biochemical techniques. A breakthrough in the identification of the enzymes came with the successful cloning of the cellulose synthesis operon of the bacterium *Acetobacter xylinum* (Saxena and Brown, Jr. 1990; Wong et al. 1990). Amino acid sequence analysis showed that cellulose synthase is a member of the glycosyl transferase family II (GT-2), which includes inverting processive nucleotide diphosphosugar glycosyl tranferases (Campbell et al. 1997; Saxena et al. 1995). Several conserved hydrophilic domains, including the proposed catalytic QXXRW motif (Table 3-2), were identified in the bacterial system. These conserved sequences were termed "U domains" to indicate "ubiquitous" presence in CESA proteins. Plant homologs of the bacterial cellulose synthase catalytic proteins were subsequently identified in an expressed sequence tag library from cotton (Pear et al. 1996).

The enzymes have several putative transmembrane domains (TMD). This is consistent with previous microscopic and biochemical data indicating that cellulose synthase is an integral membrane protein and that cellulose biosynthesis occurs at the plasma membrane (Mueller and Brown, Jr. 1980; Ross et al. 1991; Brown, Jr. et al. 1996; Delmer 1999). Visible by electron microscopy, the enzymes form large linear terminal complexes in the plasma membrane of bacteria and many algae whereas they form hexagonal rosette structures in higher plants and some algae (Mueller and Brown, Jr. 1980; Ross et al. 1991; Kimura et al. 1999). Delmer (1999) has speculated that the transmembrane domains may create a

Table 3-2. Protein model and conserved motifs for the rosette-forming eukaryotic cellulose synthase (*CESA*)

Protein Model[a]

ZnBD U1 U2 U3 U4

Motif[b]	Amino Acid Sequence[c]
ZnBD	CQICGDDVGLAETGDVFVACNECAFPVCRPCYEYERKDGTQCCPQC
U1	DYPVDKVACYVSDDGSA
U2	TNGAYLLNVDCDHYFNNS
U3	SVTEDILTGFKMHARGWISIY
U4	RLNQVLRWALGSIEIL

[a]Model of the *Arabidopsis thaliana* CESA1 predicted protein. Black boxes represent putative transmembrane domains.

[b]Conserved 'U' motifs originally identified in bacterial cellulose synthases were used to identify the higher plant enzymes, which also contain a conserved zinc-binding domain (ZnBD) specific to the eukaryotic enzymes (Saxena et al. 1995).

[c]Sequences are for the *Arabidopsis* CESA1 protein. Proposed critical residues are underlined.

pore through which the glucan chain is extruded into the extracellular space. Each hexagonal plant cell rosette structure is thought to comprise six complexes of five or six enzymes, and synthesize microfibrils containing 30–36 glucan chains. In addition to the U domains, the plant enzymes contain a conserved N-terminal Zn-binding domain indicating a possible mechanism for association of the catalytic subunits (Table 3-1) (Kurek et al. 2001).

3 TOWARD A FUNCTIONAL ANALYSIS OF CELLULOSE SYNTHASE

Homology-based genomic identification of the *CESA* genes opened the door for meaningful genetic and biochemical studies and has been conclusively supported by both. *CESA* genes have been identified in numerous plant species. The completion of the *Arabidopsis thaliana* genome revealed ten cellulose synthase or *AtCESA* genes (Richmond 2000). Mutations in seven of the ten *AtCESA* genes have been studied (Table 3-3). The *rsw1-1* mutant, which was originally isolated on the basis of a temperature-sensitive root-swelling phenotype (Baskin et al. 1992), was found to carry an A549V mutation in the *AtCESA1* gene (Arioli et al. 1998). At the nonpermissive temperature, mutant plants produce less cellulose and more soluble β-1,4-glucan than wild-type plants. The mutation was proposed to interfere with assembly to the rosette synthase complex and aggregation of the β-1,4-glucan into microfibrils at the nonpermissive temperature (Arioli et al. 1988). Several additional alleles of *AtCESA1*, which have markedly reduced cellulose, have been reported (Williamson et al. 2001; Beeckman et al.

Table 3-3. The cellulose synthase (CESA)proteins of *Arabidopsis*

Protein	Gene Locus	Alleles[a]	Protein Length	Predicted pI[c]	Predicted TMD[b]	*Arabidopsis* ESTs[d]
CESA1	At4g32410	*rsw1*	1081	6.7	8	90
CESA2	At4g39350		1084	7.5	8	9
CESA3	At5g05170	*ixr1, eli1*, *cev1*	1065	7.6	8	49
CESA4	At5g44030	*irx5*	1049	8.0	8	10
CESA5	At5g09870		1069	7.3	8	10
CESA6	At5g64740	*prc1*, *ixr2*	1084	7.4	8	36
CESA7	At5g17420	*irx3*	1026	6.7	8	14
CESA8	At4g18780	*irx1*	985	7.1	8	12
CESA9	At2g21770		1088	6.9	8	1
CESA10	At2g25540		1065	6.5	8	8

[a]*rsw* = radially swollen (Arioli et al. 1998); *ixr* = isoxaben resistant (Scheible et al. 2001); *eli* = ectopic lignin (Cano-Delgado et al. 2000); *cev* = constitutive expression of VSP1 (Ellis et al. 2002); *irx* = irregular xylem (Turner and Somerville 1997); *prc* = procuste (Fagard et al. 2000).

[b]Based on intron/exon and transmembrane modeling (Richmond and Somerville 2000) using HmmTop v2.0 (Tusnády and Simon 2001).

[c]Isoelectric point predicted by ProtParam (http://us.expasy.org/tools/protparam.html).

[d]Expressed sequence tags reported by TAIR (http://www.*Arabidopsis*.org/).

2002; Gillmor et al. 2002). The embryos of nonconditional *AtCESA1* mutants are radially swollen in appearance, indicating decreased elongation even at early stages. Although the pattern of cell division appears relatively normal, incompletely formed cell walls are observed frequently (Beeckman et al. 2002). The epidermis of the mutants is markedly affected with an apparent complete loss of guard cells and pavement cell crenulation (Beeckman et al. 2002).

The radially swollen phenotype also occurs when wild-type plants are grown in the presence of the cellulose biosynthesis inhibitor, isoxaben. Mutations in *AtCESA3* and *AtCESA6* confer resistance to isoxaben (Scheible et al. 2001; Desprez et al. 2002). This is consistent with evidence that multiple AtCESA enzymes participate in the rosette structure (Taylor et al. 2003). Antisense studies also indicate a requirement for all three genes, *AtCESA1*, *AtCESA3* and *AtCESA6,* in primary wall formation (Burn et al. 2002). This finding is further supported by strong expression of *AtCESA1*, *AtCESA3* and *AtCESA6* in young expanding leaves (Hamann et al. 2004) and evidence from GUS:promoter fusion studies which indicates the genes are expressed in the same cells simultaneously (Scheible et al. 2001).

The different structures (e.g., degree of polymerization and crystallization) of cellulose in primary and secondary cell walls prompted the hypothesis that a separate complex of enzymes was specifically devoted to secondary wall biosynthesis (Karr 1976). This idea was supported by the isolation of the irregular xylem (*irx*) mutants with defects in the *AtCESA4*, *AtCESA7* and *AtCESA8* genes (Turner and Somerville 1997; Taylor et al. 1999, 2000, 2003). Stems of these mutants contained 30–50% less cellulose than wild-type plants (Turner and Somerville 1997; Taylor et al. 2003). Recent studies reveal that these three genes are coexpressed temporally and spatially in *Arabidopsis* stems and the proteins can be copurified (Hamann et al. 2004; Taylor et al. 2003). *AtCESA7* and *AtCESA9* gene expression apparently increase with leaf age (Hamann et al. 2004), providing additional evidence for the involvement of these genes in secondary wall formation. Together these data strongly support the involvement of three separate, coregulated, cellulose synthase proteins in secondary cellulose deposition.

Very little is known about the regulatory mechanisms underlying cell wall biogenesis. Preliminary evidence suggests that the *CESA* genes are regulated by circadian rhythm, hormones such as ethylene and cytokinin, salt stress and other factors (Hamann et al. 2004). There is also a proposed link between organization of the cortical microtubule cytoskeleton and cellulose deposition (Ledbetter and Porter 1963). Evidence from a variety of studies indicates that cortical microtubules are, in some way, involved in organizing cellulose deposition and microfibril orientation (Emons et al. 1992; Fowler and Quatrano 1997). Microscopic analysis of *rsw1* plants supports this proposed connection and suggests the relationship is bidirectional; decreased rates of cellulose synthesis apparently cause destabilization of cortical microtubule organization (Sugimoto et al. 2001). Interaction between cellulose biosynthesis and biotic stress-responsive pathways is indicated by analysis of a leaky *AtCESA3* mutant allele, *cev1* (Ellis et al. 2002). The *cev1* allele apparently

causes constitutive activation of both the jasmonate and ethylene signal pathways important in plant cell defense (Ellis and Turner 2001). *cev1* plants also apparently contain increased levels of pectin (Ellis and Turner 2001). Other instances of apparently compensatory increases in pectin have been documented in cellulose deficient mutant plants (Gillmor et al. 2002) and in cell cultures adapted to growth on an inhibitor of cellulose synthesis (Shedletzky et al. 1992). Another leaky *AtCESA3* allele, *eli1*, was isolated based on its production of ectopic lignin, presumably in response to cellulose deficiency (Cano-Delgado et al. 2000). These examples suggest the existence of complex regulatory processes that sense the functional properties of the cell wall and regulate complementary pathways to achieve cell walls with appropriate aggregate functionality.

There are now over 200 cellulose synthase sequences from at least 50 organisms in the public sequence databases. With the exception of the *CESA* genes of *Acetobacter* and those specifically expressed during fiber development in cotton (*GhCESA1*, *GhCESA2*) and during xylem development in poplar (*PtCESA2*, *PtCESA3*), few *CESA* genes from other organisms have been studied in detail (Holland et al. 2000). Not surprisingly, homologs of *CESA* genes are evident in the genomes of cyanobacteria and algae (Nobles et al. 2001; Roberts et al. 2002). The early divergence of *CESA* N-terminal sequences, the putative Zn-binding domain in particular, in the green algae appears to correlate with rosette versus linear terminal cellulose synthase complex formation (Roberts et al. 2002). The presence of several *CESA* sequences in the rosette forming green alga *Mesotaenium caldariorum* has interesting implications regarding temporal and/or spatial specificity of individual CESA proteins and may provide important clues to the composition of the early-evolving rosette complex.

4 IDENTIFICATION OF THE CELLULOSE SYNTHASE-LIKE GENES

In addition to the 10 *AtCESA* genes, 30 genes with significant similarity to cellulose synthase were identified in *Arabidopsis* (Table 3-4) (Richmond 2000). These cellulose synthase-like (*CSL*) genes have been classified into eight distinct families according to sequence divergence and intron/exon structures (Richmond 2000; Hazen et al. 2002). Together the *CESA* and *CSL* genes form the cellulose synthase superfamily. Interestingly, a *CSL* gene has also been identified in the cellulose-producing cyanobacterium, *Nostoc punctiforme* (Nobles et al. 2001), indicating an ancient lineage for these gene families.

The CSL proteins contain the GT-2 family signature as well as the conserved U domains containing catalytic aspartic acid residues and QXXRW motif (Table 3-2). Members of the superfamily differ in their size, topology, and predicted physical properties. A major difference between the proteins of the CSL and CESA families is the lack of the zinc-binding domain in most CSL family members (Richmond and Somerville 2000). This may indicate that CSL proteins do not participate in forming complexes to the same degree as the CESA proteins and supports a possible function of these enzymes in making single polymer chains rather than mul-

Table 3.4. The cellulose synthase-like (CSL) proteins of *Arabidopsis*

Name	Gene Locus	Predicted Protein Model[a,b]	Protein Length[a]	Predicted pI[c]	TMD[b]	ESTs[d]
CSLA1	At4g16590		554	9.2	7	9
CSLA2	At5g22740		534	9.4	6 to 7	15
CSLA3	At1g23480		556	8.5	6 to 7	7
CSLA7[e]	At2g35650		484	9	5 to 7	4
CSLA9 (*rat*4)[f]	At5g03760		533	9.2	6 to 7	9
CSLA10	At1g24070		585	8.8	7	2
CSLA11	At5g16190		504	9.3	6 to 7	3
CSLA14	At3g56000		535	6.5	5 to 6	2
CSLA15	At4g13410		500	8.8	4 to 6	0
CSLB1	At2g32610		757	7.3	8	5
CSLB2	At2g32620		757	7.2	8	0
CSLB3	At2g32530		755	7.3	8	0
CSLB4	At2g32540		755	7.4	8	0
CSLB5	At4g15290		757	7.2	8	2
CSLB6	At4g15320		759	8.4	8	0
CSLC4	At3g28180		673	8.6	7	20
CSLC5	At4g31590		692	8.7	9	14
CSLC6	At3g07330		682	9.0	9	24
CSLC8	At2g24630		690	8.3	9	4
CSLC12	At4g07960		694	9.2	7 to 9	3
CSLD1	At2g33100		1036	7.9	8	1
CSLD2	At5g16910		1145	7.6	8	12
CSLD3 (*kjk*)[g]	At3g03050		1145	7.8	8	17
CSLD4	At4g3810		1111	6.6	8	2
CSLD5	At1g02730		1181	7.8	6 to 8	8
CSLD6	At1g32180		1181	7.8	8	0
CSLE	At1g55850		729	6.2	8	7
CSLG1	At4g24010		760	8.3	8	3
CSLG2	At4g24000		722	6.5	6 to 8	4
CSLG3	At4g23990		732	7.3	8	3

[a]Protein sequence based on intron/exon modeling performed by Todd Richmond (http://cellwall.stanford.edu/php/structure.php). Black boxes = putative transmembrane domains; White boxes = conserved 'U' domains; Grey boxes = hydrophobic regions manually.
[b]Transmembrane domains predicted with HmmTop v2.0 (Tusnády and Simon 2001).
[c]Isoelectric point predicted by ProtParam (http://us.expasy.org/tools/protparam.html).
[d]Expressed sequence tags reported by TAIR on July 1, 2003 (http://www.*Arabidopsis*.org/).
[e]An embryo lethal mutation (Goubet et al. 2003).
[f]*rat* refers to a mutant displaying resistance to *Agrobacterium tumefaciens*. (*Zhu* et al. 2003).
[g]*kjk* refers to kojak, a root-hairless mutant (Favery et al. 2001).

tichain fibrils. Biochemical evidence indicates that these polymers are mostly likely synthesized in the Golgi apparatus and exported into the extracellular space (Karr 1976; Carpita and McCann 2000). Thus, the localization of the CSL proteins to the Golgi has been proposed (Richmond and Somerville 2000).

Arabidopsis and rice appear to share only four of the gene families: *CSLA*, *CSLC*, *CSLD*, and *CSLE*. Rice appears to lack the *CSLG* and *CSLB* families and possess two additional families: *CSLF* and *CSLH*. Monocots and dicots do possess different cell wall architectures (Carpita and McCann 2000). Whether this classification of the *CSLs* truly represents a division between the monocots and dicots requires further study since biochemical functions have not yet been ascribed to the CSL proteins. Interestingly, the *CSL*s form two separate clades when compared with the *CESA* genes from plants and other organisms. The *CSLD*, *CSLG*, *CSLE* and *CSLB* families cluster with the plant *CESA* genes, whereas the *CSLA* family clusters with nonplant *CESA* genes (Richmond 2000). Although the *CSLC* family was not included in this analysis, its similarity to the *CSLC* family in *Arabidopsis* suggests that its members will also cluster with the nonplant *CESA* genes. This divergence at the gene level is further supported by analysis of the predicted protein structures. Analysis of the protein sequences (Table 3-4) supports the family assignments based on gene sequences and intron/exon structures.

Of all the *CSL* families, the *CSLD* family is most homologous to *CESA*, both at the gene and protein level (Richmond and Somerville 2000). At 1000 to 1200 amino acids, the CSLD proteins in *Arabidopsis* and rice are similar in size or larger than the CESA proteins and considerably larger than the other *CSL* gene products. The predicted isoelectric point (pI ~ 7) and relative positions of the eight transmembrane domains are similar to those of the CESAs (Tables 3-1 and 3-4). Members of the *CSLD* family in both *Arabidopsis* and rice contain very few introns. These factors all suggest the possibility that *CSLD* family members represent genetic ancestors of the *CESA* family and may also produce β-1,4-linked glucan (Richmond and Somerville 2000). Expression of the *CSLD* family members in *Arabidopsis* is quite varied. *AtCSLD2* is also expressed in older, expanded leaves, whereas *AtCSLD5* is expressed in flowers and young, expanding leaves (Hamann et al. 2004). *AtCSLD2* and *AtCSLD3* are strongly expressed in roots and negatively regulated by salt stress (Hamann et al. 2004). Additionally, *AtCSLD3* is negatively regulated by light and is apparently the only *CSL* negatively regulated by cytokinin (Hamann et al. 2004).

Expression of a tobacco *CSLD* (*NaCSLD1*) has been observed in growing pollen tubes (Doblin et al. 2001). *The enzyme* was proposed to function as a tip-growth specific cellulose synthase; however, root hairs, another tip-growing system, were not analyzed and no biochemical evidence for such a functional assignment was reported. *NaCSLD1*is an apparent ortholog of *AtCSLD4* (Doblin et al. 2001). Unfortunately, because of the incomplete information available for the tobacco genome, it is not currently possible to assess this assignment. The only mutant allele of a *CSLD* family member so far reported is *kojak*

(*kjk*), an allele of *AtCSLD3* exhibiting a defect in root hair formation (Favery et al. 2001). Northern and DNA chip analyses indicate that expression of the *Arabidopsis CSLD3* gene is not restricted to tip-growing cells (Favery et al. 2001; Hamann et al. 2004).

Of all the *CESA*/*CSL* superfamily members, the *CSLA* and *CSLC* genes are the most divergent from the *CESA* genes (Richmond and Somerville 2000). The predicted protein sequences of CSLA and CSLC family members in *Arabidopsis* exhibit some interesting features. Whereas, the CESA proteins and most other members of the other CSL families possess eight putative transmembrane domains, two in the N-terminus and six clustered in the C-terminus, most of the AtCSLA and AtCSLC proteins exhibit only four to five C-terminal transmembrane domains, respectively (Table 3-4). In addition, many of the AtCSLC predicted protein sequences contain hydrophobic regions around 50 amino acids C-terminal of the second putative transmembrane domain, which may represent two additional transmembrane domains. An additional hydrophobic region, located between the conserved U2 and U3 domains of the catalytic loop is apparent in the protein sequences of CSLA2 and CSLA9 (Table 3-4). The very interesting topologies of members of these two CSL families could have important functional consequences and merit further examination. For example, if the hydrophobic regions represent transmembrane domains which participate in forming a pore through which product is extruded, is the pore structure altered in these two families compared to the CESA and other CSL proteins and how does it affect catalysis, substrate specificity, product export, and regulation by binding partners? If these additional hydrophobic regions are not transmembrane domains, do they participate in protein–protein interactions thus specifying binding partners or are they simply involved in maintaining structural stability of the catalytic loops?

The CSLA and CSLC proteins exhibit basic pI values ranging from 8.3 to 9.2 for the CSLCs and 6.5 to 9.4 for the CSLAs (Table 3-4). There is a stretch rich in basic amino acids between the third and fourth C-terminal transmembrane domains. The other CSL family members contain a short acidic loop and putative transmembrane domain in this region. If the topology of the enzyme is such that the catalytic loop is in the cytosol (Delmer 1999), this basic loop is predicted to be extracellular. Its proximity to the proposed pore formed by the transmembrane domains is particularly intriguing. One possible role for this loop is in the formation of salt bridges with other protein partners, such as nonprocessive glycosyl transferases that may be involved in adding sugar branches. Alternatively, this loop may interact with the emerging carbohydrate chain, perhaps to facilitate chain extension. Hemicelluloses are mostly insoluble at neutral pH. A locally alkaline pH could conceivably facilitate production of these polymers.

Two mutations in the *Arabidopsis CSLA* family have been reported. A mutation of *CSLA9* (*rat4*) was isolated based on its ability to confer resistance to transformation by *Agrobacterium tumefaciens* (Zhu et al. 2003). A mutation

in the *AtCSLA7* gene results in an embryo lethal phenotype, severely affecting the pattern of cell division in the early globular stage and disrupting cellularization of the endosperm (Goubet et al. 2003). Pollen tube growth is also impaired in the mutant. *AtCSLA7* is expressed strongly in flowers, in accordance with a role in embryogenesis (Hamann et al. 2004). These results suggest that the *AtCSLA7* has a nonredundant, widespread function in *Arabidopsis* and may be particularly critical to establishing new wall placement and/or cell wall extension. Biochemical analysis of the walls of these mutants has not yet been reported.

Expression of the *AtCSLA* and *AtCSLC* genes also may indicate related functionality of the enzymes in these families (Hamann et al. 2004). Both *AtCSLA9* and *AtCSLC4* are expressed throughout the plant but show especially strong expression in stems (Hamann et al. 2004), consistent with a role for these enzymes in secondary wall formation. This, in turn, might suggest a role in hemicellulose production.

The *CSLG* family represents the only proposed dicot-specific family. There are three *CSLG* genes in the *Arabidopsis* genome arranged in tandem on chromosome four. The family may be larger in other dicots. For example, *Medicago truncultata* exhibits expressed sequence tags for six *CSLG* family members (Richmond and Somerville 2001). In *Arabidopsis* expression of the *CSLG* family members is relatively low, with *CSLG2* and *CSLG3* expressed in flowers and *CSLG1* and *CSLG3* expressed in leaves. Predicted protein sequences of the *AtCSLG* family members exhibit the closest similarity with those of the *CSLE* family member in *Arabidopsis* and rice. Like the *CESA* and *CSLD* proteins, members of the *CSLG* family from *Arabidopsis* and *Medicago* have eight putative transmembrane domains (Table 3-4). The *Arabidopsis* proteins in TAIR are annotated to contain a putative actinin-type actin binding motif (PROSITE PS00019 signature) in the C-terminal region of the protein between the fourth and fifth transmembrane domains. Although provocative, this assignment is dubious for two reasons. First, analysis of the *CSLG* family members of *Medicago* indicates some loss of this consensus sequence. Second, there is a second signature motif in the actinin-type proteins that appears to be essential for actin binding (PROSITE PS00020) which is absent in the *CSLG* predicted proteins.

In *Arabidopsis*, the *CSLB* family represents a tightly clustered group of six genes. The family is apparently absent from rice although the proposed cereal specific *CSLH* family appears related (Hazen et al. 2002). Predicted proteins of the CSLB family show structures very similar to the CSLD proteins, with eight putative transmembrane domains and neutral predicted pI values. There are few expressed sequence tags for this family in the *Arabidopsis* database, perhaps indicating a specialized function. The family exhibits very low levels of expression compared to the other *CSL* families (Hamann et al. 2004). *AtCSLB4* appears to be preferentially expressed in seedlings, whereas *AtCSLB5* is apparently preferentially expressed roots (Hamann et al. 2004). Several other family members, *AtCSLB1*, *AtCSLB2* and *AtCSLB6*, are negatively regulated by ethylene,

possibly indicating a role in cell expansion (Hamann et al. 2004). This is also supported by expression of *AtCSLB1*, *AtCSLB2* and *AtCSLB5* which appear to be preferentially expressed in young, expanding leaves compared to older leaves (Hamann et al. 2004).

There is only one *CSLE* gene in *Arabidopsis* (Richmond and Somerville 2000). The rice genome apparently encodes two *CSLE* genes. The OsCSLE2 predicted protein contains an altered *Q*XX*RW* domain with the sequence *Q*ILVLYK*RW* (Hazen et al. 2002). It will be interesting to see whether this protein is catalytically active. The sequences of the *CSLE* gene and encoded protein are sufficiently different from the other *CESA*/*CSL* superfamily members that the presence of only one copy of the gene in *Arabidopsis* is rather interesting. Expression of the *CSLE* gene is widespread with highest expression levels in seedlings, roots and older leaves (Hamann et al. 2004). The AtCSLE protein has the lowest predicted pI of the CSLs at 6.2. The overall topology is similar to that of the CSLB proteins except for a small hydrophobic region just N-terminal of the U1 domain.

Two apparent "cereal-specific" CSL families, *CSLF* and *CSLH* have been proposed. The rice *CSLF* family is highly related to both the *CESA*s and *CSLD*s (Hazen et al. 2002). Cereals produce a unique mixed-linkage glucan, which contains an alternating β-1,4-glucosyl-β-1,3-glucan backbone. It is therefore tempting to assign the CSLF proteins to production of this polymer, although this new family has not yet been the subject of biochemical analyses. The *CSLH* family is related to the *CSLB* family. Whether the *CSLF* and *CSLH* families are truly specific to monocots, cereals or the rice genome, or whether they are actually members of the *CSLD* and *CSLB* families will become evident as more full-length sequences in these families become available.

A role for the CESA enzymes in cellulose biosynthesis is well established. However, the biochemical function of the related CSL proteins is less certain. The phenotypes of the available mutations in *CSL* genes are consistent with the hypothesis that the *CSL* genes have roles in cell wall synthesis. Although mutations in many of the *CSL* genes show significant changes in the FTIR spectra of cell walls (Raab, Youngs, Milne and Somerville, unpublished), it has not yet been possible to identify reproducible differences in the amounts of cell wall polysaccharides. We believe that this reflects limitations in the analytical methods currently available for analysis of cell wall polysaccharide composition. In addition, we consider it possible that some changes in cell wall composition resulting from mutations in *CSL* genes may result in compensatory changes in other polysaccharides that tend to obscure the direct effects of the mutations.

Acknowledgments

This work was supported, in part, by grants from the US National Science Foundation and the US Department of Energy (DE-FG02-03ER20133). We thank Jennifer Milne and Michelle Facette for useful discussion of this manuscript.

REFERENCES

Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Herth W., Camilleri C., Hofte H., Plazinski J., Birch R., et al. 1998. Molecular analysis of cellulose biosynthesis in *Arabidopsis*. Science 279:717–720.

Bacic A., Harris P.J., and Stone B.A. 1988. Structure and function of plant cell walls. The Biochemistry of Plants 14:299–371.

Baskin T.I., Betzner A.S., Hoggart R., Cork A., and Williamson R.E. 1992. Root morphology mutants in *Arabidopsis thaliana*. Aust J Plant Physiol 19:427–437.

Beeckman T., Przemeck G.K.H., Stamatiou G., Lau R., Terryn N., De Rycke R., Inze D., and Berleth T. 2002. Genetic complexity of cellulose synthase A gene function in *Arabidopsis* embryogenesis. Plant Physiol 199:1883–1903.

Brett C. and Waldron K. 1990. *Physiology and Biochemistry of Plant Cell Walls*. Chapman & Hall, London.

Burn J.E., Hocart C.H., Birch R., Cork A., and Williamson R.E. 2002. Functional analysis of the cellulose synthase genes, CESA1, CESA2, and CESA3 in *Arabidopsis*. Plant Physiol 129:797–807.

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

Cano-Delgado A.I., Metzlaff K., and Bevan M.V. 2000. The *eli1* mutation reveals a link between cell expansion and secondary cell wall formation in *Arabidopsis*. *Development* 127:3395–3405.

Carpita N. and McCann M. 2000. The cell wall. In Buchanan, B.B. Gruissem, W. Jones, R. (eds.) *Biochemistry and Molecular Biology of Plants*. American Society of Plant Physiologists, Maryland, pp. 52–108.

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

Desprez T., Vernhettes S., Fagard M., Refregier G., Desnos T., Aletti E., Py N., Pelletier S., and Hofte H. 2002. Resistance against herbicide isoxaben and cellulose deficiency caused by distinct mutations in same cellulose synthase isoform CESA6. Plant Physiol 128:482–490.

Doblin Ms., DeMelis L., Newbigin E., Bacic A., and Read S.M. 2001. Pollen tubes of *Nicotiana alata* express two genes from different β-glucan synthase families. Plant Physiol 125:2040–2052.

Ellis C. and Turner J.G. 2001. The *Arabidopsis* mutant *cev1* has constitutively active jasmonate and ethylene signal pathways. Plant Cell 13:1025–1033.

Ellis C., Karafyllidis I., Wasternack C., and Turner J.G. 2002. The *Arabidopsis* mutant *cev1* links cell wall signaling to jasmonate and ethylene responses. Plant Cell 14:1557–1566.

Emons A.M.C., Derksen J., and Sassen M.M.A. 1992. Do microtubules orient plant cell wall microfibrils? Physiol Plantarum 84:486–493.

Fagard M., Desnos T., Desprez T., Goubet F., Refregier G., Mouille G., McCann M., Rayon C., Vernhettes S., and Hofte H. 2000. *PROCUSTE1* encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of *Arabidopsis*. Plant Cell 12:2409–2424.

Favery B., Ryan E., Foreman J., Linstead P., Boudonck K., Steer M., Shaw P., and Dolan L. 2001. *KOJAK* encodes a cellulose synthase-like protein required for root hair cell morphogenesis in *Arabidopsis*. Genes Dev 15:79–89.

Fowler J.E. and Quantrano R.S. 1997. Plant cell morphogenesis: plasma membrane interactions with the cytoskeleton and cell wall. Annu Rev Cell Dev Biol 13:697–743.

Gillmor C.S., Poindexter P., Loriaeu J., Palcic M.M., and Somerville C. 2002. α-Glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. J Cell Biol 156:100–1013.

Goubet F., Misrahi A., Park S.K., Zhang Z., Twell D., and Dupree P. 2003. AtCSLA7, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in *Arabidopsis* Plant Physiol 131:547–557.

Hamann T., Osborne E., Youngs H.L., Misson J., Nussaume L., and Somerville C. 2004. Global expression analysis of *CesA* and *CSL* genes in *Arabidopsis*. Cellulose 11:279–286.

Hazen S.P., Scott-Craig J.S., and Walton J.D. 2002. Cellulose synthase-like genes of rice. *Plant Physiol* 128:336–340.

Henrissat B., Coutinho P.M., and Davies G.J. 2001. A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. Plant Mol Biol 47:55–72.

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

Karr A.L. 1976. Cell wall biogenesis. In: *Plant Biochemistry*, 3rd edn. Academic Press, New York, p. 405–426.

Kimura S., Laosinchai W., Itoh T., Cui X., Linder R., and Brown, Jr. R.M., 1999. Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. Plant Cell 11:2075–2085.

Kurek I., Kawagoe 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 99:11109–1114.

Mueller S.C. and Brown, Jr. R.M., 1980. Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. J Cell Biol 84:315–326.

Nobles D.R., Romanovicz D.K., and Brown, Jr. R.M., 2001. Cellulose in cyanobacteria. Origin of vascular plant cellulose synthase? Plant Physiol 127:529–542.

Pear J., 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.

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

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

Richmond T. 2000. Higher plant cellulose synthases. Genome Biol 1:reviews3001.1–3001.6.

Richmond T.A. and Somerville C.R. 2001. Integrative approaches to determining CSL function. Plant Physiol 47:131–143.

Roberts A.W., Roberts E., and Delmer D.P. 2002. Cellulose synthase (*CESA*) genes in the green alga *Mesotaenium caldarium*. Eukaryotic Cell 1:847–855.

Ross P., Mayer R., and Benziman M. 1991. Cellulose biosynthesis and function in bacteria. Microbial Rev 55:35–58.

Saxena I.M., Lin F.C., and Brown, Jr. R.M., 1990. Cloning and sequencing of the cellulose synthase catalytic subunit gene of *Acetobacter xylinum*. Plant Mol Biol 15:673–683.

Saxena I.M., Brown, Jr. R.M., Fevre M., Geremia R.A., and Henrissat B. 1995. Multidomain architecture of β-glycosyl transferases: implications for mechanism of action. J Bacteriol 177:1419–1424.

Shedletzky E., Shmuel M., Trainin T., Kalman S., and Delmer D. 1992. Cell-wall structure in cells adapted to growth on the cellulose-synthesis inhibitor 2,6-dichlorobenzonitrile - a comparison between 2 dicotyledonous plants and a gramineous monocot. Plant Physiol 100:120–130.

Scheible W.R., Eshed R., Richmond T., Delmer D., and Somerville C. 2001. Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in the *Arabidopsis ixr1* mutants. Proc Natl Acad Sci USA 98:10079–10084.

Scheller H.V., Doong R.L., Rodley B.L., and Mohnen D. 1999. Pectin biosynthesis: a solubilized α-1,4-galctouronosyl transferase from tobacco catalyzes the transfer of galacturonic acid from UDP-galacturonic acid onto the reducing end of homogalacturonan. Planta 207:512–517.

Sugimoto K., Williamson R.E., and Wasteneys G.O. 2001. Wall architecture in the cellulose-deficient *rsw1* mutant of *Arabidopsis thaliana*: microfibrils but not microtubules lose their transverse alignment before microfibrils become unrecognizable in the mitotic and elongation zones of roots. Protoplasma 215:172–183.

Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., and Turner S.R. 1999. The irregular xylem 3 locus of *Arabidopsis* encodes a cellulose synthase gene required for secondary cell wall synthesis. Plant Cell 11:769–780.

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., 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.R. 1997. Collapsed xylem phenotype of *Arabidopsis* identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell 9:689–701.

Tusnády G.E. and Simon I. 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics 17:849–850.

Williamson R.E., Birch J.E., Baskin T.I., Arioli T., Betzner A.S., and Cork A. 2001. Morphology of *rsw1*, a cellulose-deficient mutant of *Arabidopsis thaliana*. Protoplasma 215:116–127.

Wong H., Fear A., Calhoon R., Eichinger G., Mayer R., Amikam D., Benziman M., Gelfand D., Meade J., Emerick A., Bruner R., Benbassat A., and Tal R. 1990. Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA 87:8130–8134.

Zhu Y., Nam J., Humara J.M., Mysore K.S., Lee L.-Y., Cao H., Valentine L., et al. 2003. Identification of *Arabidopsis rat* mutants. Plant Physiol 132:494–50