Plant Cell Physiol. 43(12): 1407–1420 (2002)
JSPP © 2002

Mini Review

Cellulose Biosynthesis in Plants: from Genes to Rosettes

Monika S. Doblin[1, 2], Isaac Kurek, Deborah Jacob-Wilk and Deborah P. Delmer
University of California, Davis, CA 95616, U.S.A.

Modern techniques of gene cloning have identified the *CesA* genes as encoding the probable catalytic subunits of the plant CelS, the cellulose synthase enzyme complex visualized in the plasma membrane as rosettes. At least 10 CesA isoforms exist in *Arabidopsis* and have been shown by mutant analyses to play distinct role/s in the cellulose synthesis process. Functional specialization within this family includes differences in gene expression, regulation and, possibly, catalytic function. Current data points towards some CesA isoforms potentially being responsible for initiation or elongation of the recently identified sterol β-glucoside primer within different cell types, e.g. those undergoing either primary or secondary wall cellulose synthesis. Different CesA isoforms may also play distinct roles within the rosette, and there is some circumstantial evidence that *CesA* genes may encode the catalytic subunit of the mixed linkage glucan synthase or callose synthase. Various other proteins such as the Korrigan endocellulase, sucrose synthase, cytoskeletal components, Rac13, redox proteins and a lipid transfer protein have been implicated to be involved in synthesizing cellulose but, apart from *CesAs*, only Korrigan has been definitively linked with cellulose synthesis. These proteins should prove valuable in identifying additional CelS components.

Keywords: *Arabidopsis thaliana* — Cellulose — CesA — Cotton (*Gossypium hirsutum*) — Plant polysaccharide biosynthesis.

Abbreviations: CalS, callose synthase enzyme complex; CD, cellodextrin; c-di-GMP, cyclic diguanylic acid; Csl, cellulose synthase-like; CelS, cellulose synthase enzyme complex; CesA (formerly CelA), cellulose synthase catalytic subunit; CGAbp, cellulose synthesis inhibitor CGA 325'615 binding protein; CR-P, plant conserved region; dpa, days post anthesis; DCB, 2,6-dichlorobenzonitrile; Glc, glucose; GT, glycosyltransferase; HVR, hypervariable region; Kor, Korrigan endocellulase; LTP, lipid transfer protein; Mt, metallothionein; MT, microtubule; SCD, sterol cellodextrin; SG, sitosterol-β-glucoside; SuSy, sucrose synthase; TC, terminal complex; TMH, transmembrane helix, UDP-Glc, uridine diphospho-glucose.

Introduction

Understanding the biosynthesis of wall polysaccharide components has attracted considerable interest in light of the fundamental importance of these molecules not just to plant function, but also to man. Unfortunately, very little is known of the mechanism(s) and regulation of the biosynthetic steps that control polysaccharide biosynthesis, deposition and assembly, or the interaction of these components to provide cells with a functional wall. Furthermore, manipulation of polysaccharide quantity and quality has been hampered by the lack of cloned genes for plant glycosyltransferases (GTs). As recently as 1995, not one single enzyme involved in plant cell wall biosynthesis had been purified to homogeneity, nor had a single gene coding for such an enzyme been identified and cloned. Fortunately, since that time, a number of GTs have been cloned using traditional biochemical or, more modern, in silico, molecular and genetic techniques. By far the most significant of these has been the cloning of the *CesA* genes of cotton and *Arabidopsis*, presumed to encode catalytic subunits of cellulose synthase (CelS), the enzyme complex responsible for the synthesis of cellulose. [We refer to CelS as the entire synthase complex, and will use the accepted term CesA when referring just to the catalytic subunits within that complex.] Identification of these genes has led to remarkable progress in the field, and this review will focus on this recent work. Due to space limitations, we will not cover very interesting recent work on cellulose synthesis in bacteria (Ausmees et al. 1999, Ausmees et al. 2001, Nakai et al. 1999, Römling et al. 2000, Zogaj et al. 2001). The reader is also referred to a number of reviews that give the background on cellulose synthesis (Brown 1996, Delmer 1999, Brown and Saxena 2000, Saxena and Brown 2000, Richmond and Somerville 2000, Dhugga 2001). However, for the general reader, a few words about our current understanding of the nature of the CelS complex may be helpful and are provided below.

Structures responsible for cellulose synthesis have been identified by electron microscopy in freeze-fractured plasma membranes of many organisms (Brown 1996, Kimura et al. 1999a). Linearly arranged terminal complexes (TCs) in single or multiple rows are observed in bacteria, *D. discoideum* and some algae, or hexagonal structures with six-fold symmetry, termed rosettes, are observed in mosses, ferns, algae and vascular plants (Brown 1996, Delmer 1999, Tsekos 1999). Although TCs and rosettes reside in the plasma membrane, Haigler and

[1] Corresponding author: E-mail, msdoblin@unimelb.edu.au; Fax, +61 3 9347 1071.
[2] Current address: University of Melbourne, School of Botany, Royal Parade, Parkville 3010, Victoria, Australia.

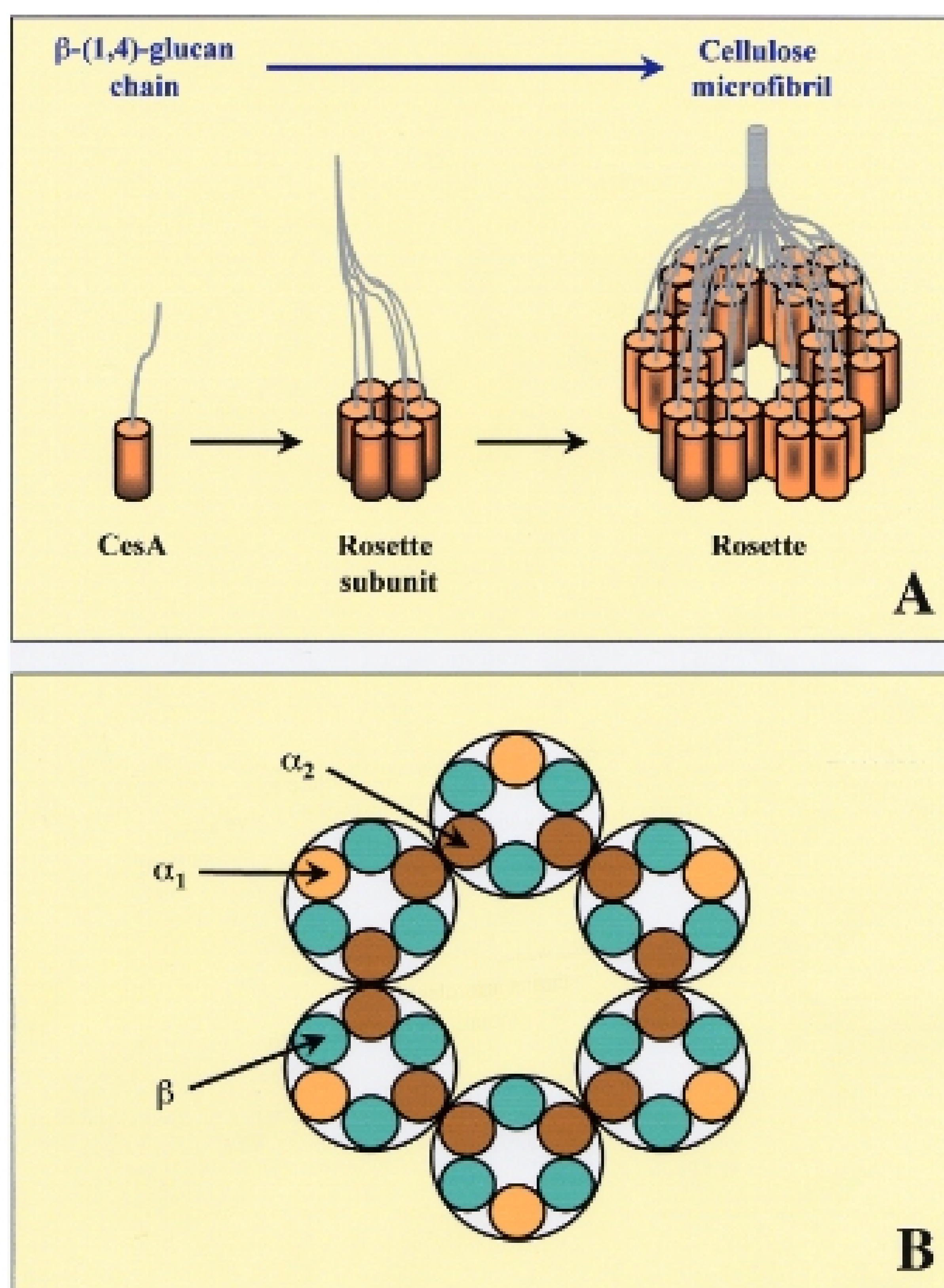

Fig. 1 A model for the structure of the rosette. (A) Six subunits, possibly containing six CesA polypeptides, interact to form a rosette, a single CesA enzyme complex. Each CesA polypeptide is shown to be involved in the synthesis of one β-(1,4)-glucan chain. The CesA protein has eight predicted TMHs which could potentially form a pore in the plasma membrane through which the nascent chain is extruded into the wall. Once the 36 chains emerge from the rosette, they coalesce to form an elementary cellulose microfibril. (B) In this modified rosette structure model of Scheible et al. (2001), at least two types of CesA polypeptides, α and β, are required for spontaneous rosette assembly. Two different types of α isoform can be distinguished, α_1 which interacts with two β isoforms only, and α_2 interacting with another α_2 isoform and two β isoforms.

Brown (1986) revealed that during synthesis, rosettes are assembled in the Golgi and then transported to the plasma membrane. Biochemical studies indicate that the higher plant CelS complex is a large (>500 kDa), integral membrane, multi-subunit enzyme utilising uridine diphospho-glucose (UDP-Glc) as substrate (Delmer 1999). Assumed to be included within each complex are a specific number of obligatory CesA catalytic subunits that utilize UDP-Glc as substrate for glucan chain elongation, as well as other components that may be involved in providing the substrate, in initiating or terminating chain elongation or that may be involved in regulating the activity of the complex. Based on our current understanding of synthesis, UDP-Glc is thought to bind to an active site on the cytoplasmic face of the plasma membrane with the polysaccharide being extruded through the membrane, presumably through a pore-type structure, into the wall (Delmer 1999, Brown and Saxena 2000). Each rosette (~25 nm in diameter comprising six subunits) is believed to contain a number of synthetic units, possibly six per subunit, each of which polymerizes a glucan chain that associates with adjacent chains of the same rosette to form elementary microfibrils (Fig. 1A).

Identification of CesA Genes

The road towards identifying components of the plant CelS complex has been a long and arduous one primarily because of the inability to assay this enzyme in vitro despite high levels of CelS activity in vivo. When incubated with UDP-Glc, isolated cell membranes from numerous plant species produce limited amounts of (1,4)-β-glucan, with the predominant product instead being callose, a linear (1,3)-β-linked polymer of D-Glc with occasional (1,6)-β-linked branches, that is normally found in only small amounts in specialized cell-types and wall structures (Delmer 1987, Delmer 1999). Rosette disassembly occurs concomitantly with the loss of CelS activity upon cell rupture. Studies on the stability of the rosettes within the plasma membrane of *Funaria hygrometrica* have shown that rosette number is reduced after 4–5 min in the presence of the vesicle transport inhibitors monensin and cytochalasin (Rudolph et al. 1989), indicating that rosettes have a short half-life and are subject to rapid degradation. Thus, early attempts to purify the enzyme were, and continue to be, severely hampered.

Due to these difficulties, researchers looked towards simpler systems, notably the bacterium *Acetobacter xylinum*, to gain insight into the mechanism of cellulose synthesis. A cellulose-synthesis operon [*bcsA–D*, later renamed *AxCesA1-D1* by Delmer (1999)] was identified by genetic complementation of *A. xylinum* insertion mutants lacking CelS activity and by purification and partial sequencing of the AxCesA1 catalytic subunit (reviewed by Ross et al. 1991). These advances in the bacterial cellulose synthesis field did not, however, lead to the immediate identification of an homologous plant gene. Rather, these plant genes were identified using a purely molecular approach that targeted a plant tissue highly enriched in cellulose, the cotton fiber, wherein two cDNA clones (*CelA1* and *CelA2*, renamed *GhCesA1* and *GhCesA2*) were identified (Pear et al. 1996). The full-length ORF of *GhCesA1* encodes a 974 amino-acid polypeptide of ~110 kDa and, like the bacterial CesA proteins, is predicted to be a membrane-bound protein with eight transmembrane helices (TMHs), two at the N-terminus and six at the C-terminus, that border a central, cytoplasmic domain. However, GhCesA1 and GhCesA2 differ from the bacterial CesAs in that they contain two large "plant-specific" insertions within the central domain: one a conserved "plant conserved region" (CR-P) and the other, a "hypervariable region" (HVR) (Pear et al. 1996, Delmer 1999). Furthermore, the plant CesA proteins have an extended N-terminal region and a shorter C-terminal region in comparison with the bacterial CesAs.

The identification of the cotton *CesA* genes as being homologous to the bacterial *CesA* genes and, thus, encoding potential catalytic subunits of CelS, was based on three lines of evidence. First, the encoded protein sequences contained three regions within the central domain that had a low (50–60%) level of similarity to the bacterial CesA proteins. Overall, DNA identity is less than 30%, which accounts for the lack of detection of plant homologues in earlier heterologous screens. Furthermore, included in these regions were small and more highly conserved regions that contained the "D,D,D,QXXRW" motif (Saxena et al. 1995). Enzymes having this signature have since been classified as members of the glycosyltransferase family 2 (GT family 2) (Campbell et al. 1997, Henrissat et al. 2001; http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html). The critical nature of the amino acids comprising the D,D,D,QXXRW motif for substrate binding and catalysis has been shown experimentally by numerous site-directed mutagenesis studies and by resolution of the crystal structure of a member of GT family 2 (Nagahashi et al. 1995, Saxena et al. 1995, Saxena et al. 2001, Saxena and Brown 1997, Saxena and Brown 2000, Charnock and Davies 1999, Charnock et al. 2001). The D,D,D,QXXRW motif has now become a robust sequence characteristic of the polysaccharide synthase members of this GT family, examples of which include all known chitin synthases, hyaluronan synthases and cellulose synthases. Secondly, Pear et al. (1996) demonstrated that the DNA segment encoding the central region of GhCesA1 bound the substrate UDP-Glc; binding did not occur under the same conditions when a fusion protein with the region containing the first conserved D residue was deleted. Thirdly, expression of *GhCesA1* and *GhCesA2* correlated with the timing of cellulose biosynthesis. Northern blot analysis showed that their expression was lowest during the stage of primary wall deposition in the fiber, but rose to much higher levels in the transition stage from primary to secondary wall synthesis at ~17 days post-anthesis (dpa) and reaching maximal levels at 24 dpa, when the rate of cellulose synthesis is highest in vivo. Taken together, these data point towards the cotton *CesA* genes encoding functional homologues of the bacterial *CesA* genes, and thus the presumed CelS catalytic subunits.

Experimental evidence for this proposed role has now been provided by the isolation of a number of cellulose-deficient mutants in *Arabidopsis thaliana*, the genetic lesions for which have been shown to be in *CesA* genes (*AtCesA1*: *rsw1*, Arioli et al. 1998a; *AtCesA3*: *ixr1*, Scheible et al. 2001; *AtCesA4*: *irx5*, Taylor and Turner 2001; *AtCesA6*: *procuste*, Fagard et al. 2000; *ixr2*, Scheible et al. 2001, Desprez et al. 2002; *AtCesA7*: *irx3*, Taylor et al. 1999; *AtCesA8*: *irx1*, Taylor et al. 2000). In addition, the disassembly of rosettes within the *rsw1* mutant (Arioli et al. 1998a), the significantly lower number of rosettes (20% compared with wild type) in a cellulose-deficient *brittle culm* mutant line of barley (Kimura et al. 1999b), and the specific labelling of these structures by polyclonal antisera directed towards a recombinant cotton CesA polypeptide (Kimura et al. 1999a), confirms that the rosettes observed in freeze-fracture experiments are the sites of cellulose synthesis in plants.

Discovery of the CesA/Csl superfamily

Sequence database searches with the cotton *CesA* genes revealed that a relatively large number of *Arabidopsis* and