Biological Sciences Faculty Publications   Biological Sciences

2018

Structure/function relationships in the rosette cellulose synthesis complex illuminated by an evolutionary perspective

Candace H. Haigler

Alison Roberts
University of Rhode Island, aroberts@uri.edu

Author Addresses

Candace H. Haigler
        Dept. of Crop and Soil Sciences and Dept. of Plant and Microbial Biology, North Carolina State University, Campus Box 7620, Raleigh, NC 27695, USA

Alison W. Roberts
        Dept. of Biological Sciences, University of Rhode Island, 120 Flagg Road, Kingston, RI 02881, USA

Corresponding Author Contact Information

candace_haigler@ncsu.edu
919-515-5645

Abstract

Cellulose microfibrils are a key component of plant cell walls, which in turn compose most of our renewable biomaterials. Consequently, there is considerable interest in understanding how cellulose microfibrils are made in living cells by the plant cellulose synthesis complex (CSC). This remarkable multi-subunit complex contains cellulose synthase (CESA) proteins, and it is often called a rosette due to its six-lobed shape. Each CSC moves within the plasma membrane as it spins a strong cellulose microfibril in its wake. To accomplish this biological manufacturing process, the CESAs harvest an activated sugar substrate from the cytoplasm for use in the polymerization of glucan chains. An elongating glucan is simultaneously translocated across the plasma membrane by each CESA, where the group of chains emanating from one CSC co-crystallizes into a cellulose microfibril that becomes part of the assembling cell wall. Here we review major advances in understanding CESA and CSC structure/function relationships since 2013, when ground-breaking insights about the structure of cellulose synthases in bacteria and plants were published. We additionally discuss: (a) the relationship of CSC substructure to the size of the fundamental cellulose fibril; (b) an evolutionary perspective on the driving force behind the existence of hetero-oligomeric CSCs that currently appear to dominate in land plants; and (c) how cellulose properties may be regulated by CESA and CSC activity. We also pose major questions that still remain in this rapidly changing and exciting research field.

Graphical Abstract
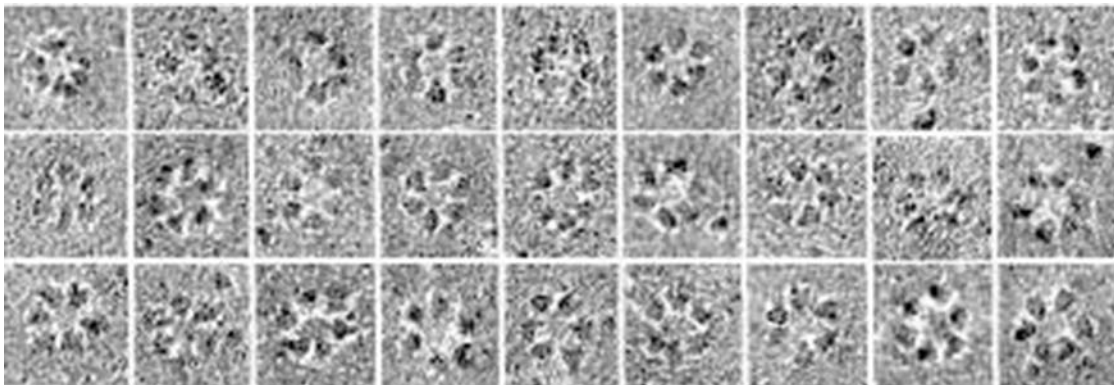
Keywords

cellulose synthase, constructive neutral evolution, cellulose microfibril, protein structure, rosette cellulose synthesis complex

Abbreviations

BcsA	Cellulose synthase in bacteria
BcsB	Non-catalytic partner protein for BscA
CSC	Cellulose synthesis complex
CESA	Cellulose synthase in plants
CNE	Constructive neutral evolution
CSR	Class-specific region of CESA
DP	Degree of polymerization
FF-TEM	Freeze fracture transmission electron microscopy
P-CR	Plant-conserved region of CESA
PCW	Primary cell wall
PCW CSC	Primary cell wall cellulose synthesis complex
SCW	Secondary cell wall
SCW CSC	Secondary cell wall cellulose synthesis complex
SSNMR	Solid-state nuclear magnetic resonance
TMH	Transmembrane helix
UDP-glucose	Uridine diphosphate-alpha-D-glucose

Introduction

This review updates our current understanding of structure/function relationships in the rosette cellulose synthesis complex (CSC). This multimeric plasma membrane-bound protein complex synthesizes cellulose in all the land plants characterized so far, as well as in their closest algal relatives. The CSC is a nanoscale fibril-spinning machine, containing many cellulose synthase (CESA) proteins. This multi-protein complex carries out its functions in a narrow zone of the cell that includes the cortical cytoplasm and cytoskeletal elements, the plasma membrane, and the exoplasmic space beneath the mature cell wall (Liu et al. 2015). The CESAs acquire substrate from the cytoplasm, synthesize ß-1,4-glucan, and export the polymers. The multiple glucan chains then co-crystallize into cellulose microfibrils near the plasma membrane surface as the CSC moves forward along a linear path. The formation of a strong cellulose microfibril via CSC activity is one of the most remarkable structural manufacturing processes in nature. Cellulose synthesis has equal importance to photosynthesis in supporting the plant lifestyle and the subsequent roles of plants in essential ecosystem cycles and human industry. As examples, plants are the basis of the food chain in every terrestrial biome, and cellulosic biomaterials have provided humans with fuel, renewable building materials, textiles, paper, and countless other products since civilization began.

We emphasize new results since 2013, which was a watershed year in the history of cellulose synthesis research. The pivotal factor in this revolutionary change was solving the structure of a bacterial cellulose synthase, BcsA from *Rhodobacter sphaeroides*, at the atomic level within a BcsA-B complex (Morgan et al. 2013). The prokaryotic cellulose synthases have been diverging along an independent evolutionary path as compared to the plant CESA for at least 1.6 billion years, as estimated from the time that a cellulose-synthesizing cyanobacterium is predicted to have been engulfed in route to becoming the plant chloroplast (Yoon et al. 2004). Nonetheless, the catalytic core structure, where the ß-1,4-glucan chain is formed, is structurally similar in prokaryotes and plants. Although no plant CESA structure has been solved so far, their similarity with BcsA was confirmed by structural comparison

of the modeled catalytic core of a plant CESA with the equivalent region in BcsA. It is now possible to infer likely structure/function relationships in plant CESAs that are shared with BcsA and to generate structurally informed hypotheses about unique aspects of plant CESAs for further experimental testing.

Several reviews on cellulose synthesis were published in the early part of the last decade (2008 - 2012), and they can be consulted for primary literature citations on aspects of cellulose synthesis that were well-proven by 2012 (Carpita 2011; Endler and Persson 2011; Guerriero et al. 2010; Mutwil et al. 2008). Here, we emphasize new discoveries from 2013 onward and also refer readers to recent reviews written since then (Cosgrove 2014; Haigler et al. 2016; Jarvis 2013; Jarvis 2018; Kumar and Turner 2015; Li et al. 2014; McFarlane et al. 2014; Meents et al. 2018; Schneider et al. 2016; Slabaugh et al. 2014; Turner and Kumar 2018).

We take an operational approach to structure/function relationships in the rosette CSC and summarize the current state of knowledge about: the structural basis of CESA polymerization and translocation of ß-1,4-glucan; the relationship of rosette CSC structure to the fundamental cellulose fibril; origins and roles of diverse CESAs and CSCs; and regulation of cellulose properties by CESA and CSC activity. We conclude with a summary of major unanswered questions about structure/function relationships in CESAs and rosette CSCs.

Structural basis of synthesis and translocation of ß-1,4-glucan by CESA

While the work to solve the structure of the prokaryotic BcsA-B complex by X-ray crystallography was occurring (3.25 Å resolution; Morgan et al. 2013), others were working to produce a stable *de novo* model of 506 amino acids of GhCESA1 from *Gossypium hirsutum* cotton, representing most of the large cytosolic (catalytic) domain (Sethaphong et al. 2013). [Note that CESA names are preceded by the genus and species abbreviation and include a number within the CESA family of that species. Most commonly, numbers in other species are consistent with their closest homolog in the model plant *Arabidopsis thaliana*. However, GhCESA1 was the first plant CESA identified (Arioli et al. 1998; Pear et al. 1996), and it is an ortholog of AtCESA8.] A rough initial homology model was assembled from predicted GhCESA1 protein fragments, which were modeled using other solved structures as templates. Then the assembled model was refined *in silico*, manually and by use a suite of tools for *de novo* (or *ab initio*) protein structure modeling, until quality parameters consistent with other protein structures solved by empirical methods at 2Å resolution were reached. The groups working on the BcsA-B structure and the GhCESA1 model met at the 2012 Plant Cell Wall Gordon Conference, where the catalytic regions of the two structures were immediately compared and found to be structurally similar with 3.9 Å root-mean-square deviation. This structural similarity established the usefulness of the partial GhCESA1 model for broad insights and hypothesis generation about plant cellulose synthesis. For example, many of the mutations in CESAs that caused cellulose-deficiency were in or near the catalytic site of the GhCESA1 model and could be interpreted with a structural context for the first time (Table S3 in Sethaphong et al. 2013).

The single catalytic site in BcsA and the GhCESA1 model was composed of one glycosyltransferase-A fold containing seven beta sheets and seven alpha helices. These secondary structure elements were arranged into one binding site for the substrate [uridine diphosphate-alpha-D-glucose (UDP-glucose)] and another binding site for the forming (or acceptor) glucan chain (McNamara et al. 2015). This structural similarity across Kingdoms occurred despite high overall sequence divergence between the plant and prokaryotic cellulose synthases. As cellulose polymerization occurs, UDP-glucose is cleaved to release energy along with the transfer of glucose to the growing chain. The structural co-alignment between the catalytic domain of BcsA and modeled GhCESA1 showed that key motifs were in similar locations (Table S3 in Sethaphong et al. 2013). These included the D, D, D, QxxRW motifs (or more specifically DDG, DCD, TED, QVLRW in GhCESA1) that are characteristic of processive Glycosyltransferase Family 2 enzymes, including cellulose synthases. Of these, the first two aspartates

coordinate the UDP of UDP-glucose in the catalytic site (via a divalent cation for DCD), the aspartate in TED is the catalytic base, and the arginine of QVLRW also coordinates UDP while its tryptophan forms van der Waals interactions with the most recently added glucose in the acceptor chain. These functional predictions based on the crystal structure alone were further validated through hybrid quantum mechanical/molecular mechanical computational methods, which were used to generate and optimize models of the BcsA active site during three stages of the polymerization reaction: enzyme-substrate, transition state, and enzyme-product (Yang et al. 2015). The rate of the modeled polymerization reaction was between about 8 - 27 glucoses per second, as compared to 90 residues per second observed *in vitro* for BcsA (Omadjela et al. 2013).

The BcsA crystal structure was part of a dimer that also included the non-catalytic, periplasmic, BcsB protein, which interacted with BcsA via one transmembrane helix (TMH) and another periplasmic helix. Protein truncation experiments showed that only these two helices of BcsB were required for the *in vitro* activity of BcsA (Omadjela et al. 2013). No protein with a similar role to BcsB has been identified in plants, and none may exist. The absence of a BcsB equivalent in plants would be consistent with two purified CESAs, one from the moss *Physcomitrella patens* (PpCESA5) and one from hybrid aspen (PttCESA8), being active *in vitro* after reconstitution into proteolipsomes (Cho et al. 2017; Purushotham et al. 2016). A BcsB partner may be uniquely required in prokaryotes due to their multi-layered outer wall that must be traversed by the elongating glucan chains (either singly or in groups) before they are extruded outside the cell, where they may or may not form crystalline microfibrils in different bacterial species. In contrast, the glucan chains synthesized by plants form microfibrils near the surface of the plasma membrane (Haigler et al. 2014), shortly after each chain passes through the TMH translocation channel that is predicted to exist in CESA (see below). Purified PttCESA8 and PpCESA5 were active *in vitro* without the addition of another priming molecule (Cho et al. 2017; Purushotham et al. 2016), arguing against the necessity of a macromolecular primer for *in vivo* plant CESA activity as long-discussed. Possibly, a glucose monomer might bind and provide the initial acceptor for cellulose polymerization, as has been speculated for BcsA (McNamara et al. 2015).

Remarkably, 18 residues of the forming glucan chain also crystallized with BcsA, ten of which were in the translocation channel. This fortuitous outcome led to profound insights into the cellulose translocation as well as its polymerization. A glucan chain translocation channel was formed from cytosolic residues immediately above the catalytic site in conjunction with a pore about 8Å wide formed by the TMHs that traversed the plasma membrane. A series of amino acids with both polar and non-polar side groups lines the TMH pore of BcsA, consistent with glucoses that are polymerized with ß-1,4 linkage presenting different interaction interfaces as the chain translocates up the channel after each successive glucose addition. In addition, a continuous series of hydrogen bonds interacts with the glucoses in the channel (McNamara et al. 2015). Computational approaches have been used to demonstrate the feasibility of an 180° inversion in spatial orientation of each successive glucose immediately after it is added to the elongating chain (Knott et al. 2016; Yang et al. 2015). The models suggest that steric constraints imposed by aromatic residues at the entrance to the translocation channel cause the rotation of the terminal glucose of the acceptor chain. The rotation of each newly-added glucose residue generates an 'in plane' conformation of the glucan chain that is favorable for traversing the membrane via the channel formed by TMHs. This explains how cellulose has a repeating unit of cellobiose despite the addition of glucose units one-by-one in the same orientation within one active site of each cellulose synthase. Both modeling studies concluded that the translocation process would not constitute a rate-limiting step for cellulose synthesis by BcsA (Knott et al. 2016; Yang et al. 2015). Limited conservation exists with the TMH domain of bacterial and plant cellulose synthases, making it reasonable to predict that glucan chain translocation in CESA occurs in a similar way. At the same time, the TMH region between Kingdoms has much less structural conservation than in the

catalytic domain (Slabaugh et al. 2014), presenting the need for further investigation of the structural basis for glucan chain translocation in CESA.

Relationship of rosette CSC structure to the fundamental cellulose fibril

The rosette CSC, which was originally seen by freeze fracture transmission electron microscopy (FF-TEM) about 40 years ago (Giddings et al. 1980; Mueller and Brown 1980) was originally called a rosette 'terminal complex' due to its apparent association with the ends (termini) of the impressions of cellulose microfibrils in the plasma membrane. The 'rosette' descriptor refers to the six-lobed CSC structure, as viewed in FF-TEM replicas where the TMHs of multiple CESAs (Kimura et al. 1999) cross the plasma membrane (Fig. 1). Given that cellulose microfibrils are an essential strength component and scaffold for
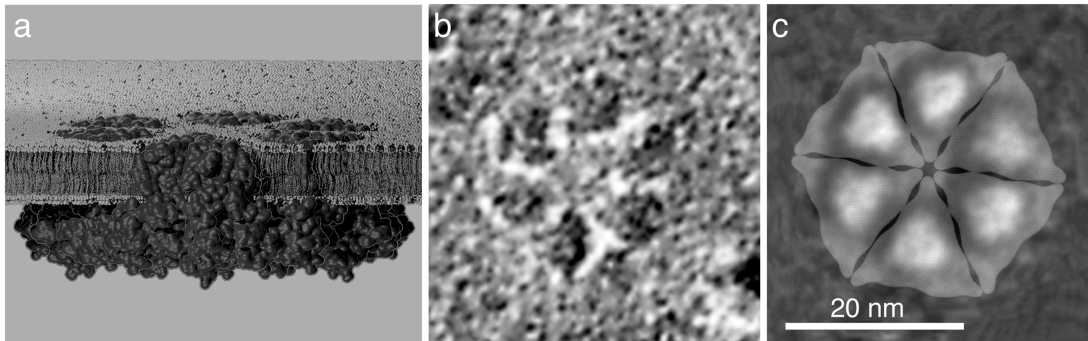
Fig. 1 Three views of rosette CSCs. (a) Cartoon of a rosette CSC embedded in the plasma membrane, based on a CESA model. A digital cut through the CSC and its surrounding membrane reveals the TMH region of one of the six trimeric lobes. The tops of the TMH regions barely protrude above the exterior plasma membrane surface (see also Haigler et al., 2014 for *in situ* images), and the modeled cytosolic regions are closely packed near the interior surface of the plasma membrane. (b) A FF-TEM image of a rosette CSC TMH region with six-lobes, viewed top-down within the plasma membrane of a moss protonemal cell. In this example, the fracture process removed the exterior leaflet of the plasma membrane and the TMH regions were highlighted with shadowing metal to produce the replica that was viewed by transmission electron microscope. The cytosolic regions were below the membrane when it was shadowed and are not visible. (c) A top-down, data-driven, schematic representation of a rosette CSC containing 18 CESAs. The TMH regions of the six lobes are represented by an image average of many FF-TEM images as shown in (b). The cytosolic regions of each lobe are represented by the semi-transparent triangles, which were placed by hand and reflect the cross-sectional shape of a model derived from small angle X-ray scattering analysis of a trimer formed in solution after heterologous expression of a purified AtCESA1 cytosolic domain (Vandavasi et al. 2016). The scale bar in (c) applies to (b) and (c). Under the Creative Commons License (http://creativecommons.org/licenses/by/4.0/), this figure includes modifications of images originally published by the authors and their coworkers in *Scientific Reports* (Nixon et al. 2016).

other polymers within cell wall structure, the size of the fibril made by one rosette CSC has been debated for many years. The size of the fundamental fibril depends on the number of simultaneously active CESAs within one rosette CSC. Although it is possible that not all CESAs in the CSC are active (Li et al. 2014), there is no evidence for this occurring in nature so far.

The long-standing belief that each rosette CSC synthesizes a 36-chain cellulose fibril is now considered unlikely. This idea arose about thirty five years ago from theoretical resonance between two

hypotheses: (1) the approximately 4 nm wide microfibrils in the cell walls of the alga *Spirogyra*, which has rosette CSCs, were composed of 36 glucan chains (Herth 1983); and (2) a 3.5 nm square 'elementary fibril' of cellulose contained 40 chains, extrapolating from early cellulose structural information and TEM images (Mühlethaler 1967). An initial critique of the 36-chain hypothesis was based on calculations using the typical area occupied by one TMH and the assumption of eight TMHs within one CESA, supporting a maximum of 24 CESAs within the rosette CSC (Bowling and Brown 2008). New evidence summarized below suggests that one rosette CSC synthesizes an 18-chain fibril. Given that the term 'elementary fibril' has long been associated with a cellulose fibril containing 36-40 chains, we prefer the term 'fundamental fibril' to describe the fibrillar product of a single rosette CSC.

Number of CESAs in the CSC and implications for cell wall structure

As summarized and illustrated recently (Turner and Kumar 2018; Jarvis 2018) recent spectroscopic analysis is consistent with diverse cell walls most commonly containing microfibrils with 18 to 28 chains (Fernandes et al. 2011; Newman et al. 2013; Thomas et al. 2014; Turner and Kumar 2018; Wang and Hong 2016; Wang et al. 2015). A study on primary cell wall (PCW) cellulose in mung bean hypocotyls by solid-state nuclear magnetic resonance (SSNMR) spectroscopy and wide angle X-ray scattering, together with computational simulations of diffractograms from different fibril sizes and aggregates, supported the existence of 18-chain fibrils and the possible coalescence of two of them to form a 36-chain unit. However, 24-chain fibrils could also fit the data (Newman et al. 2013). Many spectral interpretations, assumptions, and calculations must be made as part of this type of research, highlighting the need for other types of complementary data about the size of the fundamental fibril synthesized by one rosette CSC.

Recent studies focusing on CSC structure and composition support an 18 CESA model for the fundamental cellulose fibril. The demonstrations of 1:1:1 ratios for CESA isomers involved in PCW and SCW formation (Gonneau et al. 2014; Hill et al. 2014) were logically correlated with three CESA isomers within each of the six lobes of the rosette CSC, although an 18-, 24-, or 36-mer were equally feasible based on stoichiometry alone. Other kinds of structural analyses supported the 18-mer model. When the large cytosolic domain of a seed plant PCW CESA (AtCESA1) was heterologously expressed, it formed a trimer in solution as supported by modeling of small angle X-ray scattering data (Vandavasi et al. 2016). The same structure had a triangular cross-section, as viewed by negative staining in TEM, and six of these shapes could be accommodated beneath the transmembrane regions of an image-averaged rosette CSC from *P. patens* protonema (Nixon et al. 2016). Similarly, six trimeric lobes that had been assembled from a partial GhCESA1 model (including the TMH and catalytic domains, but excluding the N-terminal domain) fit best with the image-averaged rosette CSC. *In silico* analysis of free energies of various oligomers of this GhCESA1 model also favored a rosette CSC with 18 CESAs, three per each of six lobes (Nixon et al. 2016).

The feasibility of an 18-chain fundamental cellulose fibril was shown through molecular dynamics simulations of ß-1,4-glucan chains with ten cellobiose repeating units (Oehme et al. 2015b). This model was based on the cellulose Iß crystal structure and included water. The approximately hexagonally-shaped fibril model had six layers (with 2, 3, 4, 4, 3, or 2 chains each). The 18-chain fibril was preferred over a 36-chain model in terms of consistency with the data from spectroscopic analysis of cell walls summarized earlier (Oehme et al. 2015b). When density functional theory calculations were used to analyze the feasibility of 18-chain fibrils with different shapes, the 234432 model was considered slightly less likely than a lower energy, 5-layered, 34443 model (Kubicki et al. 2018). A microfibril model with six layers of three chains each was considered to be unlikely. In the 34443 model, there was one 'core chain' that was mainly two residues below the surface, although, on one of four sides of the microfibril cross-section, only one chain shielded the 'core chain' from a cleft in the fibril

surface. Interestingly, one chain protrudes outward on the opposite side of the fibril, allowing us to speculate that it could fit into the cleft of an adjacent fibril in an interaction that could facilitate fibril bundling. Two thirds of the chains are on the surface in both the 234432 and 34443 18-chain fibril models, leading to poor lateral chain order and a high association with water *in silico* (Oehme et al. 2015b). These features could confer potential to interact with other cellulose fibrils or cell wall matrix components *in vivo*, which could in turn confer flexibility to the assembly of cell walls with diverse biomechanical properties to serve many roles in plant structure and physiology (Jarvis 2018; Nixon et al. 2016).

Formation of fundamental fibrils and macrofibrils

FF-TEM images of the plasma membrane surface during secondary cell wall (SCW) deposition suggest that cellulose microfibrils form just beyond the extrusion sites in the rosette CSC (Haigler et al. 2014). Hemispherical domes at the ends of some fibrils may represent a pool of glucan immediately above the rosette CSC. This was supported by molecular dynamics simulations of six atomistic glucan chains, in which the chains formed a pool and then interacted in pairs before forming a unified six-chain fibril through hydrogen bonding and van der Waals interactions. Potentially, a pool of glucan from which fibrils are continuously drawn as the CSC moves forward would promote the continuity of fibril formation even if the 18 CESAs in the rosette CSC synthesize glucan chains at slightly different rates (Haigler et al. 2014).

The assembly of CESA proteins into multimeric complexes is essential for the crystallization of microfibrils with multiple chains aligned in parallel without chain folding. The ß-1,4 linkage produces a stiff and insoluble molecule at cellohexaose and above, favoring cellulose self-aggregation including through chain folding (Diotallevi and Mulder 2007; Fernandes et al. 2011; Taylor 1957; Umemura et al. 2004). Natural cellulose instead forms the cellulose I allomorph from extended, non-folded, glucan chains as the multimeric CSCs move forward in the plasma membrane. Modeling showed that interactions with the membrane surface favored the coalescence of glucan chains aligned in parallel (Haigler et al. 2014). This resonates with the observation that CESAs may be modified in ways that generate and/or foster their association with microdomains in the plasma membrane (Turner and Kumar 2018). Protein motif analysis of AtCESA1, 4, 6, 7, and 8 and biochemical experiments on AtCESA7 show that CESAs can be acylated, or have hydrophobic, long-chain, saturated fatty acids reversibly attached through covalent bonds. As many as 100 acyl groups could be covalently bound to CESAs and interact with the membrane around each CSC. More exploration is needed of the detailed mechanistic implications of this phenomenon, which could impact many levels of CSC assembly and function *in vivo* as occurs for membrane-bound ion channels (Li and Qi 2017).

Larger cellulose fibrils can exist within plant cell walls even if the fundamental fibril synthesized by one rosette CSC has only 18 chains. These small fundamental fibrils may 'bundle', or associate along at least part of their length even though they do not merge into one crystalline core (Oehme et al. 2015a; Zhang et al. 2014). Atomic force microscopy images of hydrated cellulose microfibrils on the innermost layer of intact, minimally perturbed, onion epidermal cell walls are consistent with this possibility (Zhang et al. 2014), and the convergence regions may have special roles in cell wall polymer interactions and mechanics (Park and Cosgrove 2015). The atomic force microscopy images revealed 3.5 nm wide apparently single fibrils to 35 nm wide fibrillar aggregates, with about 60% single fibrils, 20% twinned fibrils, and less common higher order associations (Zhang et al. 2016). Correspondingly, irregular groups of two, three, or more rosette CSCs have been seen by FF-TEM in cells of land plants during SCW synthesis (Haigler et al. 2014; Herth 1985; Schneider and Herth 1986), and cellulose microfibrils may also interact post-synthesis in the cell wall space. The use of Sum Frequency Generation spectroscopy to examine intact SCWs showed a stronger 2944 cm^{-1} peak (carbon-hydrogen stretch

region) relative to the weaker 3320 cm^{-1} peak (hydroxyl stretch region), which was suggested to arise from closely packed, oppositely aligned, cellulose fibrils. The opposite alignment is consistent with the observation that rosette CSCs move close to each other in the plasma membrane, but in opposite directions, during both PCW and SCW cellulose synthesis (Paredez et al. 2006; Watanabe et al. 2015).

The formation of larger cellulose macrofibrils from closely associated fundamental fibrils may be assisted by cell wall matrix components and/or lignin through adhesion or dehydration (Donaldson 2007). There is also evidence to support fundamental fibrils merging before crystallization into larger cellulose crystallites, e.g. cotton fiber SCWs contain cellulose fibrils/crystallites with 4-5 nm overall dimensions that have been modeled as containing 46-52 chains (Fang and Catchmark 2014; Lee et al. 2015; Martinez-Sanz et al. 2017). *Populus* tension wood has similar 4.5 nm cellulose fibrils, whereas normal wood and opposite wood have smaller microfibrils (3.7 - 3.9 nm lateral dimensions, respectively) (Foston et al. 2011). A higher density of rosette CSCs could foster formation of larger cellulose fibrils through production of closely-spaced fundamental fibrils that could interact before crystallization finalizes. Low density rosette CSCs have been seen in cotton fibers during PCW and SCW synthesis (Herth 1985), but cotton is not well-suited for FF-TEM due to challenges in mounting undamaged long fibers prior to freezing. However, FF-TEM revealed rosette CSC densities of 93 - 135 per μm^2 in tracheary elements engaged in SCW deposition (Schneider and Herth 1986), about ten times higher than in cells synthesizing PCWs (reviewed by Emons 1991; Herth 1985). Fluorescently labeled CESAs also occur at higher density in the plasma membrane (where they are presumed to exist within rosette CSCs) of differentiating tracheary elements monitored by live cell imaging (Li et al. 2016; Watanabe et al. 2015). An increased bias toward parallel movement of rosette CSCs during SCW synthesis, as compared to frequent bidirectional movement during PCW synthesis, would also promote the possibility of co-crystallization of the glucan products of more than one CSC and/or fibril bundling by surface interactions (Li et al. 2016; Watanabe et al. 2015). Currently we do not know what determines the directionality of CSC movement in the plasma membrane once they arrive at the general location of cellulose synthesis, which is established in interaction with cytoskeletal elements (Cosgrove 2014; Jarvis 2013; Jarvis 2018; Kumar and Turner 2015; Li et al. 2014; McFarlane et al. 2014; Meents et al. 2018; Schneider et al. 2016; Slabaugh et al. 2014; Turner and Kumar 2018).

Origins and roles of diverse CESAs and CSCs

Rosette CSCs occur in all land plants examined (reviewed in Emons 1991), and also in their closest relatives within the charophycean green algae. Phylogenetic relationships within this algal group have been difficult to resolve (Wickett et al. 2014), and CSC structure has not been determined for members of several charophycean lineages (Klebsormidiales, Chlorokybales, and Mesostigmatales). Furthermore, the unusual structures reported for *Coleochaete scutata* CSCs (Okuda and Brown 1992) may be misidentified based on strong resemblance to plasmodesmata (Willison 1976). However, available data are consistent with the rosette CSC first appearing in a charophycean green alga at least 630 million years ago (Morris et al. 2018), then being retained throughout the descent to land plants. Additionally, rosette CSCs exist in unicellular and filamentous desmids (e.g. *Micrasterias* and *Spirogyra*, respectively) and the giant-celled *Charales*. All of these organisms are expected to produce 18-chain fundamental fibrils from one rosette CSC, whereas thick microfibrils are produced by large CSCs in chlorophycean green algae (Tsekos 1999). The enhanced interaction potential of smaller fibrils, leading to greater cell wall diversity, may have contributed to the success of charophycean lineages in diverse aquatic and some terrestrial environments, as well as the ability of plants to colonize terrestrial environments (Harholt et al. 2016; Jarvis 2018; Nixon et al. 2016).

Diverse types of rosette CSCs and CESAs

Substantial evidence accumulated through genetic, biochemical, and imaging experiments supports the existence of two major types of CSCs in *Arabidopsis*, one responsible for the biosynthesis of PCW cellulose and another responsible for the biosynthesis of SCW cellulose (McFarlane et al. 2014; Meents et al. 2018). Results of these experiments are consistent with both types of CSCs being obligate hetero-oligomeric protein complexes. Both the PCW and SCW CSCs contain multiple copies of three non-interchangeable CESA isoforms, all of which are required for CSC assembly and delivery to the plasma membrane (Desprez et al. 2007; Persson et al. 2007; Taylor et al. 2003). In *Arabidopsis*, AtCESA1, 3 and a member of the AtCESA6-like class are required in PCW CSCs; and AtCESA4, 7 and 8 are required in SCW CSCs (reviewed by McFarlane et al. 2014). These six isoforms define six CESA sequence classes that correspond to six phylogenetic clades that have been strongly conserved in seed plants (Carroll and Specht 2011; Kumar et al. 2016).

The two major types of rosette CSCs participate in two distinct phases of plant morphogenesis. PCW *CESAs* are expressed in expanding organs (Hamann et al. 2004), and their mutations are either lethal or cause primary defects in cell expansion (reviewed by McFarlane et al. 2014). In contrast, SCW *CESAs* are expressed in differentiating xylem (reviewed in Taylor et al. 2004) and are tightly co-regulated with other genes involved in SCW deposition (Brown et al. 2005; Ruprecht et al. 2011). SCW *CESA* mutations produce the "irregular xylem" phenotype in which the water conducting vessels have collapsed inward (Turner and Somerville 1997). This typical division of CSC activity evidently appeared and was stabilized early in the evolution of the seed plants, but the two major types of CSCs are also used in other ways. Promoter-reporter experiments supported the activity of PCW CESAs in the deposition of the thick cell walls of *Arabidopsis* trichomes (Betancur et al. 2011). Biochemical and gene suppression experiments show that both PCW and SCW CESAs are involved in the synthesis of thick cell walls in poplar wood (Song et al. 2010; Xi et al. 2017). Analysis of mutant phenotypes and live cell imaging showed that the cellulose in *Arabidopsis* seed mucilage is synthesized by a PCW type CSC in which CESA1 is replaced by its closest paralog CESA10 (Griffiths et al. 2015). These occurrences also highlight the diversity and complexity of plant cell walls beyond the traditional PCW and SCW categories, with highly specialized cell walls being assembled through use of a 'cell wall toolbox' (Betancur et al. 2011).

Origin of CESA and CSC diversity

Like seed plants, the moss *P. patens* use different CESAs for PCW and SCW biosynthesis, and SCW biosynthesis in leaf midribs requires two distinct CESAs (Norris et al. 2017). Phylogenetic analysis supports independent evolution of this functional divergence in mosses and seed plants, which shared a common ancestor with a single CESA (Carroll and Specht 2011; Roberts and Bushoven 2007; Yin et al. 2009). This common ancestor also likely had rosette CSCs, which must have been homo-oligomeric, with multiple copies of only one CESA (Roberts and Bushoven 2007; Roberts et al. 2012). Gene duplication and sequence divergence within the euphyllophyte lineage (ferns and seed plants) produce the six CESA classes (or clades) now observed in seed plants. The conifers and angiosperms that have been examined have at least one member of each class (Carroll and Specht 2011; Jokipii-Lukkari et al. 2017; Kumar et al. 2016), and several of the classes also include fern CESAs (Yin et al. 2014).

Phylogenetic analyses show that the PCW and SCW CESAs of seed plants were separated before three clades diverged independently within both groups to generate heteromeric PCW and SCW CSCs (Roberts et al. 2012). Evolution of the specialized SCW CSCs in mosses represents a third independent origin of functionally specialized CESAs (Norris et al. 2017). This convergent evolution is consistent with strong selective advantage conferred through the uncoupling of transcriptional regulation of the

different CESA isoforms (Norris et al. 2017). For example, diversification of cellular function is promoted by only specific cells accumulating the high densities of CSCs that support the aggregation of fundamental fibrils and SCW thickening after cell expansion ends (Li et al. 2016; Schneider and Herth 1986). An explanation was needed, however, for the multiple cases of independent evolution of hetero-oligomeric CSCs.

The most compelling explanation is found in a hypothesized evolutionary ratchet that drives protein systems toward complexity (Doolittle 2012). This process, a form of constructive neutral evolution (CNE), requires duplicate genes (or paralogs) to exist in the same genome, which is an outcome of whole genome duplications that have frequently occurred during land plant evolution (Lang et al. 2018; Ren et al. 2018). Functional differentiation of duplicate genes often occurs by positive selection when a mutation affecting the expression pattern (sub-functionalization) or biochemical function (neo-functionalization) of one of the encoded proteins confers an adaptive advantage (Conant and Wolfe 2008). To the contrary, the CNE hypothesis explains the evolution of hetero-oligomeric multi-subunit protein complexes without positive selection (Doolittle 2012; Finnigan et al. 2012). Figure 2 illustrates a hypothetical scenario for the evolution of obligate hetero-oligomeric CSCs by the evolutionary ratchet of CNE, combining the effects of gene duplications, divergence of the paralogs by accumulation of neutral mutations, and occasional mutations with differential effects on self- and non-self-interactions between CESAs.

The diagrams and accompanying text in the figure provide a stepwise illustration of how an initially homo-oligomeric CSC could have evolved through the CNE process into a hetero-oligomeric CSC with each of three CESA isomers in a fixed position. Step one illustrates an ancestral homo-oligomeric CSC lobe containing only one CESA, and step two illustrates the co-existence of two CESA paralogs (encoding isomers A and B) after a genome duplication. Initially A and B would have been identical, and they remain interchangeable even as they accumulate neutral mutations that lead to them becoming distinct. Although some of the mutations occur at interfaces where the CESA isomers interact, they do not initially alter self- or non-self-interactions between the isomers. In the third step, a mutation blocks the self-interaction of isomer A. This is also a neutral mutation because it does not prevent formation of lobes containing only B or A+B. However, the mutation results in neofunctionalization of A such that it can now occupy only one position within a CSC lobe. Although deleterious mutations that block both self- and non-self-interactions are expected to be more common, these would be eliminated by selection because the lobe could not form. In step four, an additional genome duplication together with the continued accumulation of neutral mutations generates two pairs of isomers A'/A'' and B'/B'', with the members of each pair initially able to function interchangeably within the CSC. Eventually, neutral mutations block the interaction of B' with itself (step five) and B'' with itself (step six). At this point, B' and B'' are neofunctionalized, but A' and A'' cannot neofunctionalize in this way because this pair occupies only one position within the CSC due to the inability of their common ancestor (A) to self-interact. These isomers, as well as those arising from gene duplications subsequent to the evolution of the obligate hetero-oligomeric CSC, would be susceptible to sub-functionalization (divergence of expression patterns) as seen for the 6-like CESAs in *Arabidopsis* PCW CSCs. More commonly, duplicated genes are eliminated by selection. Loss of either A' or A'' would result in an obligate hetero-oligomeric CSC in which each position is occupied by a single CESA isomer, as in *Arabidopsis* SCW CSCs.

As a result of the hypothetical progression of the CNE process, the CESA isoforms *might* differ only in the position that they can occupy within the complex. Similar scenarios involving interaction-altering mutations in different interfaces could explain independent evolution of obligate hetero-oligomeric PCW and SCW CSCs in seed plants, which both have three unique positions. To the contrary, the SCW CSCs of mosses have only two unique positions (Norris et al. 2017) as represented in step four of Fig. 2. Known gene duplications in the moss lineage are much more recent than those that generated the six seed plant CESA classes (Lang et al. 2018). Thus, the moss CESA family may be at an intermediate

stage in the evolution of hetero-oligomeric CSCs. As the progression toward a hetero-oligomeric CSC occurs via changes in allowed interaction interfaces, the newly distinct CESA isomers could also evolve other differences, as will be discussed later.
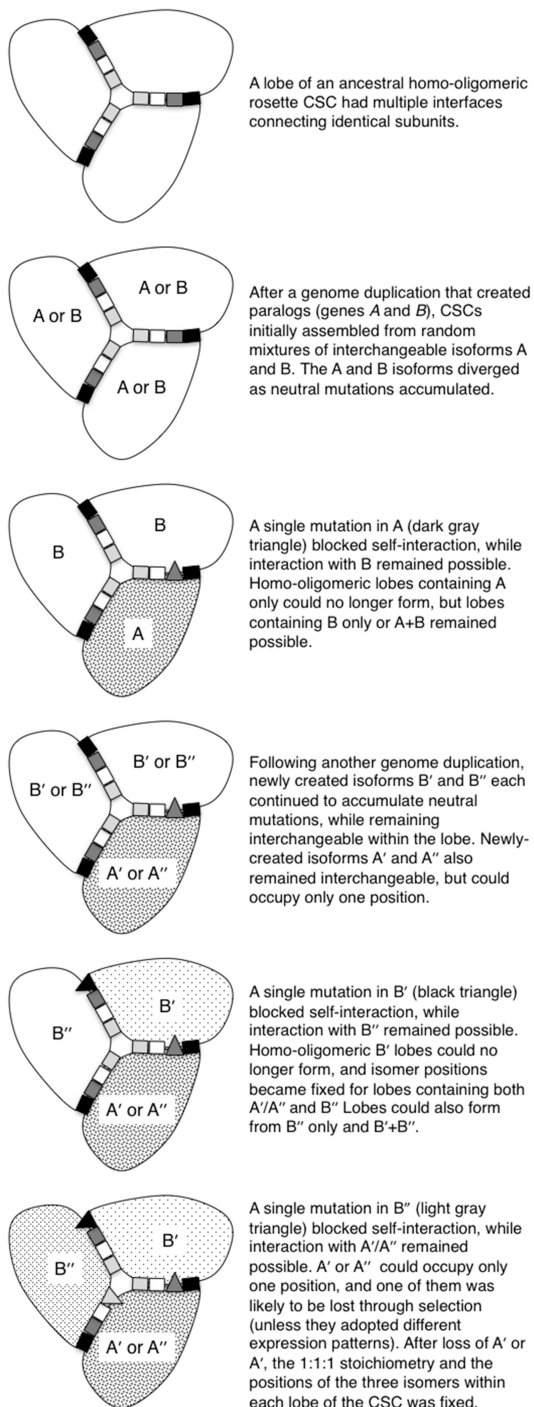
Fig. 2 A hypothetical scenario for the evolution of obligate hetero-oligomeric CSCs from an ancestral homo-oligomeric CSC according to the constructive neutral evolution hypothesis (Doolittle 2012; Finnigan et al. 2012). The text in the figure describes a stepwise scenario for the evolution of complexity in the CESA family and in CSC structure, as illustrated in each accompanying diagram.

A lobe of an ancestral homo-oligomeric rosette CSC had multiple interfaces connecting identical subunits.

After a genome duplication that created paralogs (genes *A* and *B*), CSCs initially assembled from random mixtures of interchangeable isoforms A and B. The A and B isoforms diverged as neutral mutations accumulated.

A single mutation in A (dark gray triangle) blocked self-interaction, while interaction with B remained possible. Homo-oligomeric lobes containing A only could no longer form, but lobes containing B only or A+B remained possible.

Following another genome duplication, newly created isoforms B′ and B″ each continued to accumulate neutral mutations, while remaining interchangeable within the lobe. Newly-created isoforms A′ and A″ also remained interchangeable, but could occupy only one position.

A single mutation in B′ (black triangle) blocked self-interaction, while interaction with B″ remained possible. Homo-oligomeric B′ lobes could no longer form, and isomer positions became fixed for lobes containing both A′/A″ and B″. Lobes could also form from B″ only and B′+B″.

A single mutation in B″ (light gray triangle) blocked self-interaction, while interaction with A′/A″ remained possible. A′ or A″ could occupy only one position, and one of them was likely to be lost through selection (unless they adopted different expression patterns). After loss of A′ or A′, the 1:1:1 stoichiometry and the positions of the three isomers within each lobe of the CSC was fixed.

Implications of CESA diversity for CSC sub-structure and function

CSCs are often illustrated as having identical hetero-oligomeric lobes. However, an arrangement with two each of three different homo-trimeric lobes remains possible for CSCs containing 18 CESAs (Olek et al. 2014; Turner and Kumar 2018; Vandavasi et al. 2016). Although a role for CNE in the evolution of CSCs (Fig. 2) has not been tested experimentally, this hypothesis is helpful in thinking about these alternatives. CNE is expected to have operated on CSC lobes containing a random mixture of identical isoforms following gene duplication, leading to CSCs with identical hetero-trimeric lobes. Evolution of CSCs with three different homo-trimeric lobes would require positive selection for lobes containing identical isoforms. This cannot be ruled out, but it is less parsimonious that the scenario illustrated in Fig. 2. Although originally interpreted in the context of the 36-mer CSC model, results from attempts to purify CSCs by tandem affinity chromatography from *Arabidopsis* lines expressing both His-CESA7 and STREP-CESA7 (Atanassov et al. 2009) are consistent with an 18-mer CSC with hetero-trimeric lobes. The ~440 kDa (4-mer) and ~700 kDa (6-mer) complexes that were observed may represent two associated lobes (6-mer) and single lobes associated with an additional CESA7 subunit (4-mer) as the smallest units with two CESA7 subunits that could be derived from a CSC with hetero-trimeric lobes.

Cellulose synthesis researchers have long assumed that regions that vary in sequence between isomers will be where interfaces occur within and between lobes of different types of CSCs (Carroll and Specht 2011; Hill et al. 2018a; Kumar et al. 2016; Vergara and Carpita 2001). Most sequence divergence across all CESAs occurs in four regions: 1) the N-terminus, which is truncated in some isoforms; 2) the hypervariable region between the Zn-binding domain and the first TMH; 3) the class-specific region (CSR) within the catalytic domain; and 4) the short C-terminus following the last TMH (Carroll and Specht 2011). Several "class-specific" sequence regions, which have greater amino acid diversity between vs. within CESA classes, have been a particular focus for potential roles in CESA-CESA interactions (Carroll and Specht 2011; Hill et al. 2018a; Kumar et al. 2016; Vergara and Carpita 2001). However, the CESAs within CSCs are apparently tightly packed with many contact regions (Nixon et al. 2016). Whereas evolution of hetero-oligomeric CSCs involved critical mutations that altered the binding properties of some of these contact regions, others retained their original binding properties. (Fig. 2). For example, in the moss CESAs, the CSRs are interchangeable between functional CESA classes despite being class-specific at the sequence level (Scavuzzo-Duggan et al. 2018).

The CSR region is also intrinsically disordered, with two important implications: 1) it is inherently suitable as an interaction domain, although we don't currently know whether this occurs between or within lobes; and 2) due to relaxed selection, it is prone to become class-specific at the sequence level by accumulating mutations that do not alter its functional role in the CSC (Scavuzzo-Duggan et al. 2018). Indeed, the critical mutations that alter interactions may reside within regions that are otherwise highly conserved. We can also predict that class-specific interfaces differ in CSCs that evolved the hetero-oligomeric state independently. This prediction has been proven to be true. Domain swap experiments in *Arabidopsis* (Hill et al. 2018a; Kumar et al. 2016; Wang et al. 2006) and *P. patens* (Scavuzzo-Duggan et al. 2018) have shown that the specific regions conferring functional class-specificity differ among the CESA isomers composing moss CSCs, seed plant PCW CSCs, or seed plant SCW CSCs.

To summarize and look to the future, it appears likely that the CESA-CESA interfaces within the CSCs of extant plants existed in the ancestral rosette and that they have become modified in different CSC lineages such that they contribute to the specific interactions between isoforms. By combining the results from domain swap experiments with computational modeling of interactions between CESA isomers, it should be possible to predict small sequence motifs or single residues that have high probabilities of participating in CESA-CESA interactions and test their function using targeted mutagenesis and complementation assays. We also need to understand why three CESA isoforms are required for assembly and function of seed plant CSCs *in vivo* whereas cellulose synthesis can be

reconstituted from a single hybrid aspen CESA *in vitro* (Purushotham et al. 2016) and heterologously-expressed CESAs can homotrimerize (Vandavasi et al. 2016) or homodimerize (Olek et al. 2014). Although more complex explanations may eventually be discovered, *in vitro* conditions may not constrain oligomerization in the same way as occurs *in vivo*. For example, less stable complexes may persist *in vitro*, while they would be removed by cellular protein quality control mechanisms *in vivo*. Interactions in the TMH domain may also more strongly restrict assembly as compared to only cytosolic interfaces in the partial CESAs.

In addition, CESAs may function in non-canonical ways beyond forming typical hetero-trimeric PCW and SCW CSCs, particularly in plants with expanded *CESA* gene families (poplar has 17 CESAs, Kumar et al. 2009) and phenotypes that depend on very high amounts of cellulose, e.g. wood. Larger CESA gene families arose through successive whole genome duplications, including ones that were specific to particular lineages (Ren et al. 2018). Afterwards, the interaction interfaces in CESA paralogs may have continued to evolve so that CSCs with non-canonical composition may now function in specialized plant tissues. This is an area of important future research, particularly coupled with parallel analysis of cellulose microfibril properties, as exemplified in recent work on wood formation in *Populus tremula* (aspen trees) (Zhang et al. 2018). Proteomic analysis showed that, during deposition of the cellulose-rich gelatinous layer in developing tension wood, PttCESA8b accounts for 78% of the total SCW CESAs. This observation is consistent with a role for both hetero-oligomeric SCW CSCs and homo-oligomeric PtCESA8b CSCs in biosynthesis of the gelatinous layer, which apparently contains wider-diameter cellulose microfibrils as compared to normal aspen wood (Zhang et al. 2018). Interestingly, cellulose fibril formation occurs *in vitro* from heterologously expressed PttCESA8, a single CESA isomer from hybrid aspen, after reconstitution into proteoliposomes (Purushotham et al. 2016). Although other possibilities have not been completely ruled out, available evidence suggests that PpCESA5 functions as a homo-oligomer (Li 2017), and this isomer also supports *in vitro* cellulose fibril formation when reconstituted into proteoliposomes (Cho et al. 2017). Therefore, cellulose fibril formation *in vitro* seems to occur most readily from isomers that may function as homo-oligomers *in vivo*. This presents the challenge of generating similar outcomes *in vitro* from hetero-oligomeric CESA assemblies in order provide an additional tool for analysis of the biochemical characteristics of canonical hetero-trimeric PCW and SCW CSCs.

Regulation of cellulose properties by CESA and CSC activity

Our view of how CESAs function to synthesize ß–1,4 glucan and how the CSC generates a microfibril has been strongly influenced by the concept that rosette CSCs are hetero-oligomeric, as demonstrated by analysis of the effect of CESA mutations on typical primary and secondary walls of *Arabidopsis*. However, the likely existence of functional homo-oligomeric rosette CSCs in the common ancestor of seed plants and mosses indicates that the different subunits of hetero-oligomeric rosette CSCs do not have distinct essential roles in glucan chain elongation and that the distinct interfaces that now exist between different subunits are not essential for formation of the rosette morphology (Roberts and Bushoven 2007). Nonetheless, CESA isomers have evolved distinct roles, including non-interchangeable positions in typical heteromeric PCW and SCW CSCs (see above). Beyond this, the expanding family of CESA isomers has evolved regulatory differences. For example, when catalytically inactivated versions of each *Arabidopsis* SCW CESA were expressed in their cognate null mutant backgrounds, differences in the extent of rescue were observed as would be consistent with unequal contributions of the different isomers to overall CSC activity (Kumar et al. 2018). Our current insights into such changes are summarized below, including discussion of particular cellulose properties that are predicted to be modulated through CESA and CSC activity.

Evidence for differences between CESA isomers and CSCs

As a general principle, protein activities in living organisms are frequently regulated by a variety of post-translational modifications. One of the most common is the addition and removal of phosphate groups, which changes protein conformation and may either activate, inactivate, or modulate CESA activity (Chen et al. 2010). As reviewed elsewhere (Jones et al. 2016; Kumar and Turner 2015; Li et al. 2014; McFarlane et al. 2014; Speicher et al. 2018), there is evidence from analysis of cellular proteomes that AtCESA1, 3, 4, 5, and 7 isomers are phosphorylated at one or more sites within the N-terminus and/or catalytic domains. These phosphorylation sites are often conserved in the orthologs from other plant species (Carroll and Specht 2011), which is consistent with conserved regulatory potential of both PCW and SCW CSCs. Whether or not phosphorylation-dependent regulation actually occurs in particular cell types or developmental stages depends on the presence and activity of the enabling protein kinases and phosphatases, which are not yet identified for CESAs. Functional analysis of CESAs with changed phosphorylation potential by complementation of their cognate mutant lines provided early evidence that CSC velocity and movement direction are under regulatory control in the cell (Chen et al. 2010). Phosphorylation status can also affect the stability and degradation of CESA (Hill et al. 2018b; Taylor 2007).

The velocity and trajectory of CSC movement in the plasma membrane have been analyzed in an expanding set of cells in recent years, building on many analyses done in etiolated hypocotyls of *Arabidopsis* wild-type and mutant lines. For example, the PCW CSCs of the grass, *Brachypodium distachyon*, moved with a similar velocity distribution and mean rate (164-184 nm/min) as compared to *Arabidopsis* hypocotyl PCW CSCs when analyzed in the same lab (Liu et al. 2017). Therefore, there may be conserved aspects of PCW CESA and CSC function in expanding plant cells, which is consistent in theory with the phosphorylation discussion above. The dynamic microtubule array has complex impacts on CESA movement rates that are still being explored mechanistically (Liu et al. 2017; Woodley et al. 2018). Other potential regulatory factors include structural differences between CESA isomers, post-translational modifications, interactions with other proteins, specialized plasma membrane domains, temperature, and/or cell expansion rate and direction.

Caution is appropriate when comparing CESA velocities in different cell types due to such variables and potential pleiotropic effects of mutations in experimental genotypes. Many of these variables were eliminated in observations of AtCESA1 (PCW CSCs) and AtCESA7 (SCW CSCs) moving at different rates in the same membrane of a differentiating tracheary element (Watanabe et al. 2018). The difference in average velocity was substantial, with AtCESA7 moving about 70% faster than AtCESA1 during two of three developmental stages assayed. The reduced velocity of AtCESA7 toward the end of SCW synthesis highlights the cellular regulation of this important aspect of CSC behavior. These interesting observations were made during the 'transition stage' of cell wall development that bridges PCW and SCW deposition in some cells such as the tracheary elements (Meents et al. 2018). Further supporting isomer-specific cellular regulation, the tagged AtCESA1 and AtCESA7 proteins were internalized independently even though they were delivered to the plasma membrane in the same vesicles. AtCESA1 was selectively removed at the end of the transition stage, while AtCESA7 continued to function in the plasma membrane for SCW cellulose synthesis. These differences in velocity and cellular trafficking observed concurrently between and within CESA isomers in one cell type (Watanabe et al. 2018) are likely to be determined by differences in protein structure or post-translational modifications, but these remain to be fully elucidated.

Other recent evidence raises the possibility of functional variation between CSCs. Optimized FF-TEM sample preparation methods showed that the CSCs in moss protonema synthesizing primary walls had a range of diameters, with lobes often, but not always, appearing triangular (Nixon et al. 2016). The images analyzed are of the TMH region of the membrane-embedded CSC, as commonly revealed by the

FF-TEM method. Measurements of 324 CSCs showed an external diameter range of 17.6 to 25.6 nm (45% variation), with a mean of 21.4 ± 1.3 nm. These values derive from measurements based on hexagonal geometry, i.e. point-to-point on the exterior of the lobes, which is more precise than drawing an encompassing circle given the slight ellipsoidal shape of some rosette CSCs. The average lobe area within the CSCs was 39.9 ± 6.5 nm^2, based on measurements of the individual lobes of 50 randomly selected CSCs, with a range of 30.8-49.6 nm^2 (61% variation). Some of the lobes (30 - 70%) appeared triangular, with others appearing more like squares. These observations are expected to be close to reality, given the small grain size and light coating of platinum/carbon used to generate the replicas (see Fig. S3 in Nixon et al. 2016).

As we speculated previously, these large differences in rosette CSC morphology, even within one cell type, could reflect functional variability, e.g. activated versus non-activated CSCs (Nixon et al. 2016). For example, do other membrane-bound proteins tightly, but perhaps reversibly, associate with CSCs to cause changes in diameter and lobe area and shape? One strong candidate for such a role would be the endo-ß-1,4-glucanase, KORRIGAN1, which has one TMH and co-localizes near or with PCW CESAs in *Arabidopsis* and impacts CESA velocity and cellulose accumulation (Lei et al. 2014; Vain et al. 2014). Variable CSC diameter in the TMH region and variable spacing between the TMH lobes (Nixon et al. 2016) could also arise from regulated changes in the conformation of flexible CESA regions on the periphery of the cytosolic catalytic core, e.g. the CSR, the plant-conserved region (P-CR), and/or the N-terminal regions (Scavuzzo-Duggan et al. 2018). Shape variations may also arise from stresses inherent in cellulose polymerization and microfibril production (Diotallevi and Mulder 2007; Turner and Kumar 2018). In the future, it will be worthwhile to compare CSC morphology in different organisms and developmental stages and after different experimental treatments. Consequently, we may gain a better understanding of how rosette CSC variations are regulated, with potential consequences for cellulose microfibril structure (Turner and Kumar 2018).

Degree of polymerization of cellulose

The degree of polymerization (DP) of cellulose in wild type plants varies. For example, solvated cellulose from cotton fiber PCWs versus SCWs had typical DP values of at least 617 or 6,170, respectively (Timpa and Triplett 1993). The apparent DP was 7500 in SCW-rich *Arabidopsis* stems subjected to ball milling and relatively gentle cell wall extraction and solubilization (Schneider et al. 2017). The DP varied between 2000 - 2500 in highly purified cellulose from three types of poplar wood (normal wood, tension wood, and opposite wood) (Foston et al. 2011), but these may be underestimates due to the extensive sample processing.

We know little about how cellulose DP is controlled, although the properties of CESAs as well as regulatory factors could have an impact. The length of one cellulose chain would logically be affected by the extent of enzyme processivity, or how long continuous polymerization proceeds before the acceptor chain is released from the active site causing chain termination. Control at this level would determine DP if the enzyme operated in isolation, but in the living plant other factors such as chain cleavage by membrane-anchored endo-1,4-ß-glucanases that are required for normal cellulose synthesis (Mølhøj et al. 2001; Yu et al. 2013), CSC retrieval from the membrane under abiotic stress (Bashline et al. 2014), or diminished carbon status of the plant (Ivakov et al. 2017) could potentially cause abrupt chain termination. There may also be CESA structural features regulating substrate access to the catalytic site, and implicitly the extent of glucan chain continuity. Part of the P-CR region from rice OsCESA8 (77 amino acids of 125 total) was crystallized, revealing two interacting alpha helices connected by an ordered loop (Rushton et al. 2017). This solved structure was consistent with computational models of the entire P-CR regions from the six major Arabidopsis CESA isoforms (Sethaphong et al. 2016). These two helices of the P-CR were modeled as existing above the catalytic site, where they could potentially be involved in

regulating substrate access (Rushton et al. 2017). A still undescribed regulatory signal might cause the two helices of the P-CR to block substrate access to the catalytic site, thereby inducing chain termination and providing a theoretical means of DP control.

The existence of the translocation channel immediately above the active site stabilizes the elongating chain and promotes cellulose synthase processivity (Knott et al. 2016), but it remains to be determined whether inherent structural and/or regulatory differences between isomers help to control DP. If such isomer-specific differences exist, we may find common characteristics that: (a) limit DP in the three isoforms involved in PCW cellulose synthesis (e.g. AtCESA1, 3, and 6, or 6-like isomers); and (b) enhance DP in the three isoforms involved in SCW cellulose synthesis (e.g. AtCESA4, 7, and 8). This prediction assumes that three CESA isomers within each major type of CSC collectively synthesize one continuous cellulose microfibril. The observation that AtCESA1 (a PCW CESA) is retrieved selectively from the same membrane while AtCESA7 (a SCW CESA) continues to function (Watanabe et al. 2018) demonstrates isomer-specific regulation of persistence time in the plasma membrane, which could result in different cellulose chain lengths synthesized by the two major types of CSCs.

Cellulose microfibril structure and interaction potential

Recently, the known allomorphs of native cellulose I have been expanded beyond I alpha and I beta that were first identified over three decades ago (Atalla and Vanderhart 1984). The combination of analysis of never-dried or rehydrated ^{13}C-labeled PCWs by two-dimensional magic angle spinning SSNMR spectroscopy and density functional theory calculations led to the proposal of five interior and two surface cellulose conformations (named a-g) (Wang et al. 2016). From the fibril surface inwards, these occur: in solvent-exposed regions, 'f and g'; in small regions with less hydration where cellulose interacts with hemicellulose, 'd'; between the surface and the core, 'a and b'; and within the dehydrated core, 'c'. The structural variation relates to different conformations of the CH_2OH group outside the five-carbon ring of glucose. These conformations were found in varying percentages in primary cell walls of diverse types (Wang et al. 2016), and conformations 'f, b and c' (on the surface, between the surface and the core, and within the dehydrated core, respectively) were found by similar analyses of dried *Arabidopsis* secondary walls (Dupree et al. 2015). Potentially, conformational changes in the CSC or changes in CSC velocity could influence the final conformations of the chains in the microfibril, which in turn might affect the potential for cellulose fibrils to interact with others or cell wall matrix polymers. The possible existence of such mechanistic links is supported by the subtle variation in the structure of cellulose microfibrils in cotton fibers (with predominantly SCWs) produced by three *Gossypium* species (Martinez-Sanz et al. 2017) and in three types of wood in one *Populus* species (Foston et al. 2011).

Cellulose crystallinity

The crystallinity of cellulose is likely to be also affected by CESA and CSC structure and behavior. This conclusion arises from analysis of CESA mutants, with hazards for interpretation arising from potential pleiotropic and/or compensatory effects arising from the primary mutations (Fujita et al. 2013). An *Arabidopsis* line (*ixr1-2*) with a T942I mutation in *AtCESA3* was resistant to the cellulose synthesis inhibitor, isoxaben, and the stems had about 69% of the control crystalline cellulose content and a slight reduction in cellulose crystallite size (2.2 nm versus 2.34 nm in the control). Analysis by ^{13}C magic-angle-spinning SSNMR spectroscopy showed that the cellulose was less crystalline, and it was enzymatically converted into glucose more efficiently. These changes were correlated with about 8% faster average CESA movement in hypocotyls (Harris et al. 2009; Harris et al. 2012; Scheible et al. 2001). The broad applicability of the results was supported by the expression of the *ixr-2* variant of *AtCESA3* in tobacco, which reduced the cellulose content, hindered upright growth, and increased saccharification efficiency

of the woody stems (Sahoo et al. 2013). Another mutation, A903V in *AtCESA1* (*aegeus*, or *ags1-2*) conferred resistance to the cellulose-synthesis inhibitor, quinoxyphen, and the mutant plants had cellulose-related phenotypes similar to the *ixr1-2* mutants, in association with about 16% faster average CSC movement. While there is uncertainty about where the *ixr1-2* mutation lies in CESA structure due to ambiguity about the number of TMHs in CESA (Slabaugh et al. 2014), the *ags1-2* mutation is within a TMH where it may directly impact glucan chain translocation (Morgan et al. 2013). There is increasing evidence that the TMH region, as collectively considered across multiple TMHs, is a 'hot spot' for regulation of cellulose biosynthesis and crystallization (Shim et al. 2018).

In *Arabidopsis*, the relationship between CSC velocity and cellulose crystallinity is not always the same (see Fujita et al. 2013 for a summary). However, substantially different growth temperatures between experiments could impact CSCs in unexpected ways, such as stability in the membrane (Hill et al. 2018b). For example, the analysis of the *ixr1-2* and *aegeus* lines described above was done at 21°C, a typical growth temperature for *Arabidopsis thaliana*. When the *any1* mutant line (D604N in *AtCESA1*) was grown at 29°C, the hypocotyls showed a 21% *reduction* in CSC velocity along with reduced cellulose crystallinity in the stems (17.3% versus 21% in the controls) (Fujita et al. 2013). Therefore, the results for the *any1* line show an opposite correlation of CSC velocity and cellulose crystallinity as compared to the *ixr1-2* and *aegeus* lines. Notwithstanding potential effects of growth temperature differences, this opposite correlation could arise from the different location of the *any1* mutation, which is in a loop connecting core beta strand 5 of the catalytic region with an alpha helix likely to lie near the membrane. This alpha helix contains another quinoxyphen-resistant mutation (*lycos*, G620E in *AtCESA1*), which showed decreased cellulose crystallinity when grown at 21°C. In this case, the CSC velocity was not analyzed (Sethaphong et al. 2013; Slabaugh et al. 2014). Similarly, another isoxaben-resistant mutant (*ixr1-6*, S377F in *AtCESA3*) had reduced cellulose crystallinity, and the mutation site is within core beta strand 1 (Sethaphong et al. 2013). Therefore, the cytosolic region of CESA may also help to regulate cellulose crystallinity, potentially through impacts on the glucan chain as it exits the catalytic site or through extended impacts of amino acid substitutions on CESA structure (Sethaphong et al. 2013). Overall, the results invoke the potential to modify cellulose crystallinity in useful ways once structure/function relationships in CESAs are fully understood. Such useful alterations will likely be subtle, given that the crystalline nature of cellulose confers strength to growing plants and biomaterials derived from them.

Unanswered questions about structure/function relationships in CESA and the rosette CSC

The advances in understanding plant cellulose synthesis via the rosette CSC have accelerated in the structural research era that began in 2013 with the publication of two groundbreaking research efforts (Morgan et al. 2013; Sethaphong et al. 2013). These studies built upon the earlier identification of cellulose synthases in bacteria (Saxena et al. 1990; Wong et al. 1990) and plants (Arioli et al. 1998; Pear et al. 1996) and set the stage for a complete understanding of the historically challenging field of cellulose biosynthesis research (Delmer 1999).

These new structural insights about cellulose synthases and other results in the last five years have also generated many important and exciting questions about glucan chain synthesis and microfibril formation via the rosette CSC, as exemplified by the list below. Trans-disciplinary research bridging many areas of expertise and technical approaches will be key to answering these questions. Most of the questions are directly related to the content of this review, but a few point to exciting related areas that can be explored further in other reviews that we cited earlier. Echoing the sentiments of Dr. Deborah Delmer, a pioneer in this field who stimulated our interests and collaborated with both of us in cellulose synthesis research, we are "having fun along the way, and we welcome new travelers to share in the adventure" (Delmer 1987).

Important questions about rosette CSC structure/function relationships include:

- What is the atomic structure of CESA and its homo-oligomeric and hetero-oligomeric assemblies?
- Where in the cell and how is the assembly of CSCs with rosette morphology controlled?
- Given that rosette CSCs occur in Golgi vesicles, how is cellulose synthesis typically prevented in the endomembrane system and activated at the plasma membrane?
- What was the structure of the ancestral CSC that directly preceded the rosette and how did its cellulose synthases differ from CESAs?
- Did rosette CSC complexity arise through constructive neutral evolution and does this explain multiple independent origins of hetero-oligomeric rosette CSCs?
- Do homo-oligomeric CSCs exist in seed plants, and, if yes, do they have an adaptive advantage in certain cells and/or any different outcomes for microfibril formation?
- Do homo-oligomeric CESA assemblies, or even rosette CSCs, form in *in vitro* synthesis systems and facilitate microfibril formation?
- What are the regulatory functions associated with the distinct structural features of different CESA isoforms?
- How is acylation of CESAs regulated in the cell, and how does it change CESA and CSC behavior?
- How is phosphorylation of CESAs regulated in the cell, and what more can we learn about its impacts on CESA conformation and CSC behavior?
- Are there other post-translational modifications of CESAs or other proteins associated with CSCs that have critical outcomes for cellulose synthesis?
- Do lobes of the CSC actually have variable sizes and shapes, and, if yes, what is the cause and outcome of this variation?
- Are different CSC diameters associated with different activity states? Could these differences be related to the activation of cellulose synthesis in the plasma membrane, given that rosette CSCs are not typically active in the endomembrane system?
- Do changes in conformation of flexible and disordered CESA regions on the periphery of the catalytic domain regulate CSC diameter?
- How do other plasma-membrane-bound or -associated proteins modulate CSC activity and/or cellulose crystallization? What are their precise roles in cellulose synthesis?
- How might differences between PCW versus SCW CESAs and CSCs help to regulate the degree of polymerization and cellulose microfibril formation in these two major cell wall types?
- Does variation in CESA and/or CSC structure affect the positions of the glucan chain extrusion sites in ways that can modulate chain coalescence and cellulose crystallinity?
- How might variation in the rate of polymerization and CSC velocity affect the crystallization of cellulose?
- Do the twins, triplets, or higher order aggregates of CSCs observed by FF-TEM have actual consequences for macrofibril formation? If yes, is increased density of CSCs sufficient to induce these associations, or are there other controls?
- Does the degree of bidirectional movement of CSCs have functional consequences for cell wall properties, and how is the directionality of movement regulated?
- Are cell wall properties changed when PCW and SCW CSCs operate together in the same membrane?
- Do 18-chain fundamental cellulose fibrils play a unique role in cell wall assembly and mechanical characteristics as compared to larger fibrils?

Acknowledgments

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Arioli T et al. (1998) Molecular analysis of cellulose biosynthesis in *Arabidopsis* Science 279:717-720

Atalla RH, Vanderhart DL (1984) Native cellulose: A composite of two distinct crystalline forms Science 223:283-285

Atanassov II, Pittman JK, Turner SR (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

Bashline L, Lei L, Li S, Gu Y (2014) Cell wall, cytoskeleton, and cell expansion in higher plants Mol Plant 7:586-600 doi:10.1093/mp/ssu018

Betancur L, Singh B, Rapp RA, Wendel JF, Marks MD, Roberts AW, Haigler CH (2011) Phylogenetically distinct cellulose synthase genes support secondary wall thickening in arabidopsis shoot trichomes and cotton fiber J Integr Plant Biol 52:205-220

Bowling AJ, Brown RM, Jr. (2008) The cytoplasmic domain of the cellulose-synthesizing complex in vascular plants Protoplasma 233:115-127

Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics Plant Cell 17:2281-2295

Carpita NC (2011) How plants make cellulose and other (1->4)-b-D-glycans Plant Physiol 155:171-184

Carroll A, Specht CD (2011) Understanding plant cellulose synthases through a comprehensive investigation of the cellulose synthase family sequences Front Plant Sci 2:5 doi:10.3389/fpls.2011.00005

Chen S, Ehrhardt DW, Somerville CR (2010) Mutations of cellulose synthase (CESA1) phosphorylation sites modulate anisotropic cell expansion and bidirectional mobility of cellulose synthase Proc Natl Acad Sci U S A 107:17188-17193

Cho SH et al. (2017) Synthesis and Self-Assembly of Cellulose Microfibrils from Reconstituted Cellulose Synthase Plant Physiol 175:146-156 doi:10.1104/pp.17.00619

Conant GC, Wolfe KH (2008) Turning a hobby into a job: How duplicated genes find new functions Nature Rev Genet 9:938-950

Cosgrove DJ (2014) Re-constructing our models of cellulose and primary cell wall assembly Curr Opin Plant Biol 22:122-131 doi:10.1016/j.pbi.2014.11.001

Delmer DP (1987) Cellulose biosynthesis Ann Rev Plant Physiol 38:259-290

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

Desprez T et al. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana* Proc Natl Acad Sci U S A 104:15572-15577

Diotallevi F, Mulder B (2007) The cellulose synthase complex: a polymerization driven supramolecular motor Biophys J 92:2666-2673

Donaldson L (2007) Cellulose microfibril aggregates and their size variation with cell wall type Wood Sci Technol 41:443-460

Doolittle WF (2012) Evolutionary biology: A ratchet for protein complexity Nature 481:270-271

Dupree R, Simmons TJ, Mortimer JC, Patel D, Iuga D, Brown SP, Dupree P (2015) Probing the molecular architecture of Arabidopsis thaliana secondary cell walls using two- and three-dimensional (13)C solid state nuclear magnetic resonance spectroscopy Biochemistry 54:2335-2345 doi:10.1021/bi501552k

Emons AMC (1991) Role of particle rosettes and terminal globules in cellulose synthesis. In: Haigler CH, Weimer PJ (eds) Biosynthesis and Biodegradation of Cellulose. Marcel Dekker, New York, pp 71-98

Endler A, Persson S (2011) Cellulose synthases and synthesis in Arabidopsis Mol Plant 4:199-211 doi:10.1093/mp/ssq079

Fang L, Catchmark JM (2014) Structure characterization of native cellulose during dehydration and rehydration Cellulose 21:3951-3963

Fernandes AN et al. (2011) Nanostructure of cellulose microfibrils in spruce wood Proc Natl Acad Sci U S A 108:E1195-1203 doi:10.1073/pnas.1108942108

Finnigan GC, Hanson-Smith V, Stevens TH, Thornton JW (2012) Evolution of increased complexity in a molecular machine Nature 481:360-364

Foston M et al. (2011) Chemical, ultrastructural and supramolecular analysis of tension wood in Populus tremula x alba as a model substrate for reduced recalcitrance Energy Environ Sci 4:4962-4971

Fujita M et al. (2013) The anisotropy1 D604N mutation in the Arabidopsis cellulose synthase1 catalytic domain reduces cell wall crystallinity and the velocity of cellulose synthase complexes Plant Physiol 162:74-85 doi:10.1104/pp.112.211565

Giddings TH, Jr., Brower DL, Staehelin LA (1980) Visualization of particle complexes in the plasma membrane of *Micrasterias denticulata* associated with the formation of cellulose fibrils in primary and secondary cell walls J Cell Biol 84:327-339

Gonneau M, Desprez T, Guillot A, Vernhettes S, Hofte H (2014) Catalytic subunit stoichiometry within the cellulose synthase complex Plant Physiol 166:1709-1712 doi:10.1104/pp.114.250159

Griffiths JS et al. (2015) Unidirectional movement of cellulose synthase complexes in Arabidopsis seed coat epidermal cells deposit cellulose involved in mucilage extrusion, adherence, and ray formation Plant Physiol 168:502-520 doi:10.1104/pp.15.00478

Guerriero G, Fugelstad J, Bulone V (2010) What do we really know about cellulose biosynthesis in higher plants? J Integr Plant Biol 52:161-175

Haigler CH, Davis J, Slabaugh E, Kubicki JD (2016) Biosynthesis and assembly of cellulose. In: Rose JK (ed) Molecular Cell Biology of the Growth and Differentiation of Plant Cells. CRC Press, Boca Raton, FL, pp 120-138. doi:10.1201/b20316

Haigler CH, Grimson MJ, Gervais J, Le Moigne N, Hofte H, Monasse B, Navard P (2014) Molecular modeling and imaging of initial stages of cellulose fibril assembly: evidence for a disordered intermediate stage PLoS One 9:e93981 doi:10.1371/journal.pone.0093981

Hamann T, Osborne E, Youngs HL, Misson J, Nussaume L, Somerville C (2004) Global expression analysis of *CESA* and *CSL* genes in Arabidopsis Cellulose 11:279-286

Harholt J, Moestrup Ø, Ulvskov P (2016) Why plants were terrestrial from the beginning Trends Plant Sci 21:96-101

Harris D, Stork J, Debolt S (2009) Genetic modification in cellulose-synthase reduces crystallinity and improves biochemical conversion to fermentable sugar Global Change Biol 1:51-61

Harris DM et al. (2012) Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1A903V and CESA3T942I of cellulose synthase Proc Natl Acad Sci U S A 109:4098-4103

Herth W (1983) Arrays of plasma-membrane "rosettes" involved in cellulose microfibril formation of *Spirogyra* Planta 159:347-356

Herth W (1985) Plant cell wall formation. In: Robards AW (ed) Botanical Microscopy. Oxford University Press, Oxford, UK, pp 285-310

Hill JL, Jr., Hammudi MB, Tien M (2014) The Arabidopsis cellulose synthase complex: a proposed hexamer of CESA trimers in an equimolar stoichiometry Plant Cell 26:4834-4842 doi:10.1105/tpc.114.131193

Hill JL, Jr., Hill AN, Roberts AW, Haigler CH, Tien M (2018a) Domain swaps of Arabidopsis secondary wall cellulose synthases to elucidate their class specificity Plant Direct 2:e00061

Hill JL, Jr., Josephs C, Barnes WJ, Anderson CT, Tien M (2018b) Longevity in vivo of primary cell wall cellulose synthases Plant Mol Biol 96:279-289 doi:10.1007/s11103-017-0695-4

Ivakov A et al. (2017) Cellulose synthesis and cell expansion are regulated by different mechanisms in growing Arabidopsis hypocotyls Plant Cell 29:1305-1315 doi:10.1105/tpc.16.00782

Jarvis MC (2013) Cellulose biosynthesis: counting the chains Plant Physiol 163:1485-1486 doi:10.1104/pp.113.231092

Jarvis MC (2018) Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture Phil Trans R Soc A 376 doi:10.1098/rsta.2017.0045

Jokipii-Lukkari S, Sundell D, Nilsson O, Hvidsten TR, Street NR, Tuominen H (2017) NorWood: a gene expression resource for evo-devo studies of conifer wood development New Phytol 216:482-494 doi:10.1111/nph.14458

Jones DM, Murray CM, Ketelaar KJ, Thomas JJ, Villalobos JA, Wallace IS (2016) The emerging role of protein phosphorylation as a critical regulatory mechanism controlling cellulose biosynthesis Front Plant Sci 7:684 doi:10.3389/fpls.2016.00684

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

Knott BC, Crowley MF, Himmel ME, Zimmer J, Beckham GT (2016) Simulations of cellulose translocation in the bacterial cellulose synthase suggest a regulatory mechanism for the dimeric structure of cellulose Chem Sci 7:3108-3116 doi:10.1039/C5SC04558D

Kubicki JD, Yang H, Sawada D, O'Neill H, Oehme D, Cosgrove D (2018) The shape of native plant cellulose microfibrils Sci Rep 8:13983 doi:10.1038/s41598-018-32211-w

Kumar M, Atanassov I, Turner S (2016) Functional analysis of cellulose synthase CESA protein class-specificity Plant Physiol 173:970-983

Kumar M, Mishra L, Carr P, Pilling M, Gardner P, Mansfield SD, Turner S (2018) Exploiting CELLULOSE SYNTHASE (CESA) class specificity to probe cellulose microfibril biosynthesis Plant Physiol 177:151-167 doi:10.1104/pp.18.00263

Kumar M et al. (2009) An update on the nomenclature for the cellulose synthase genes in *Populus* Trends Plant Sci 14:248-254

Kumar M, Turner S (2015) Plant cellulose synthesis: CESA proteins crossing kingdoms Phytochemistry 112:91-99 doi:10.1016/j.phytochem.2014.07.009

Lang D et al. (2018) The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution Plant J 93:515-533 doi:10.1111/tpj.13801

Lee CM, Kafle K, Belias DW, Park YB, Glick RE, Haigler CH, Kim SH (2015) Comprehensive analysis of cellulose content, crystallinity, and lateral packing in *Gossypium hirsutum* and *Gossypium barbadense* cotton fibers using sum frequency generation, infrared and Raman spectroscopy, and X-ray diffraction Cellulose 22:971-989

Lei L et al. (2014) The *jiaoyao1* mutant Is an allele of *korrigan1* that abolishes endoglucanase activity and affects the organization of both cellulose microfibrils and microtubules in Arabidopsis Plant Cell 26:2601-2616 doi:10.1105/tpc.114.126193

Li S, Bashline L, Lei L, Gu Y (2014) Cellulose synthesis and its regulation The Arabidopsis book / American Society of Plant Biologists 12:e0169 doi:10.1199/tab.0169

Li S et al. (2016) Cellulose synthase complexes act in a concerted fashion to synthesize highly aggregated cellulose in secondary cell walls of plants Proc Natl Acad Sci U S A 113:11348-11353 doi:10.1073/pnas.1613273113

Li X (2017) Characterization of cellulose synthesis complexes and *Physcomitrella patens*. University of Rhode Island

Li Y, Qi B (2017) Progress toward Understanding Protein S-acylation: Prospective in Plants Front Plant Sci 8:346 doi:10.3389/fpls.2017.00346

Liu D et al. (2017) Imaging cellulose synthase motility during primary cell wall synthesis in the grass Brachypodium distachyon Sci Rep 7:15111 doi:10.1038/s41598-017-14988-4

Liu Z, Persson S, Sanchez-Rodriguez C (2015) At the border: the plasma membrane-cell wall continuum J Exp Bot 66:1553-1563 doi:10.1093/jxb/erv019

Martinez-Sanz M, Pettolino F, Flanagan B, Gidley MJ, Gilbert EP (2017) Structure of cellulose microfibrils in mature cotton fibres Carbohydr Polym 175:450-463 doi:10.1016/j.carbpol.2017.07.090

McFarlane HE, Doring A, Persson S (2014) The cell biology of cellulose synthesis Annu Rev Plant Biol 65:69-94 doi:10.1146/annurev-arplant-050213-040240

McNamara JT, Morgan JL, Zimmer J (2015) A molecular description of cellulose biosynthesis Annu Rev Biochem 84:895-921 doi:10.1146/annurev-biochem-060614-033930

Meents MJ, Watanabe Y, Samuels AL (2018) The cell biology of secondary cell wall biosynthesis Ann Bot 121:1107-1125 doi:10.1093/aob/mcy005

Mølhøj M, Ulvskov P, Dal Degan F (2001) Characterization of a functional soluble form of a *Brassica napus* membrane-anchored endo-1,4-ß-glucanase heterologously expressed in *Pichia pastoris* Plant Physiol 127:674-684

Morgan JLW, Strumillo J, Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation Nature 493:181-U192 doi:Doi 10.1038/Nature11744

Morris JL et al. (2018) The timescale of early land plant evolution Proc Natl Acad Sci U S A 115:E2274-E2283 doi:10.1073/pnas.1719588115

Mueller SC, Brown RM, Jr. (1980) Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants J Cell Biol 84:315-326

Mühlethaler K (1967) Ultrastructure and formation of plant cell walls Annu Rev Plant Physiol 18:1-24

Mutwil M, Debolt S, Persson S (2008) Cellulose synthesis: a complex complex Curr Opin Plant Biol 11:252-257

Newman RH, Hill SJ, Harris PJ (2013) Wide-angle x-ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls Plant Physiol 163:1558-1567 doi:10.1104/pp.113.228262

Nixon BT et al. (2016) Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthesis complex Sci Rep 6:28696

Norris JH et al. (2017) Functional specialization of cellulose synthase isoforms in a moss shows parallels with seed plants Plant Physiol 175:210-222

Oehme D, Doblin MS, Wagner J, Bacic A, Downton MT, Gidley MJ (2015a) Gaining insight into cell wall cellulose macrofibril organisation by simulating microfibril adsorption Cellulose 22:3501-3520

Oehme DP, Downton MT, Doblin MS, Wagner J, Gidley MJ, Bacic A (2015b) Unique aspects of the structure and dynamics of elementary Ibeta cellulose microfibrils revealed by computational simulations Plant Physiol 168:3-17 doi:10.1104/pp.114.254664

Okuda K, Brown RM, Jr. (1992) A new putative cellulose-synthesizing complex of *Coleochaete scutata* Protoplasma 168:51-63

Olek AT et al. (2014) The structure of the catalytic domain of a plant cellulose synthase and its assembly into dimers Plant Cell 26:2996-3009 doi:10.1105/tpc.114.126862

Omadjela O, Narahari A, Strumillo J, Melida H, Mazur O, Bulone V, Zimmer J (2013) BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis Proc Natl Acad Sci U S A 110:17856-17861 doi:10.1073/pnas.1314063110

Paredez AR, Somerville CR, Ehrhardt DW (2006) Visualization of cellulose synthase demonstrates functional association with microtubules Science 312:1491-1495

Park YB, Cosgrove DJ (2015) Xyloglucan and its interactions with other components of the growing cell wall Plant Cell Physiol 56:180-194 doi:10.1093/pcp/pcu204

Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase Proc Natl Acad Sci U S A 93:12637-12642

Persson S et al. (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in Arabidopsis Proc Natl Acad Sci U S A 104:15566-15571

Purushotham P, Cho SH, Díaz-Moreno SM, Kumar M, Nixon BT, Bulone V, Zimmer J (2016) A single heterologously expressed plant cellulose synthase isoform is sufficient for cellulose microfibril formation in vitro Proc Natl Acad Sci U S A 113:11360–11365

Ren R et al. (2018) Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms Mol Plant 11:414-428 doi:10.1016/j.molp.2018.01.002

Roberts AW, Bushoven JT (2007) The cellulose synthase (*CESA*) gene superfamily of the moss *Physcomitrella patens* Plant Mol Biol 63:207-219

Roberts AW, Roberts EM, Haigler CH (2012) Moss cell walls: structure and biosynthesis Front Plant Sci 3:166

Ruprecht C, Mutwil M, Saxe F, Eder M, Nikoloski Z, Persson S (2011) Large-scale co-expression approach to dissect secondary cell wall formation across plant species Front Plant Sci 2:23 doi:10.3389/fpls.2011.00023

Rushton PS, Olek AT, Makowski L, Badger J, Steussy CN, Carpita NC, Stauffacher CV (2017) Rice *Cellulose SynthaseA8* plant-conserved region is a coiled-coil at the catalytic core entrance Plant Physiol 173:482-494 doi:10.1104/pp.16.00739

Sahoo DK, Stork J, DeBolt S, Maiti IB (2013) Manipulating cellulose biosynthesis by expression of mutant Arabidopsis proM24::CESA3(ixr1-2) gene in transgenic tobacco Plant Biotechnol J 11:362-372 doi:10.1111/pbi.12024

Saxena IM, Lin FC, Brown RM, Jr. (1990) Cloning and sequencing of the cellulose synthase catalytic subunit gene of *Acetobacter xylinum* Plant Mol Biol 15:673-683

Scavuzzo-Duggan TR et al. (2018) Cellulose synthase 'class specific regions' are intrinsically disordered and functionally undifferentiated J Integr Plant Biol 60:481-497 doi:DOI: 10.1111/jipb.12637

Scheible W-R, Eshed R, Richmond T, Delmer D, Somerville C (2001) Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in *Arabidopsis Ixr1* mutants Proc Natl Acad Sci USA 98:10079-10084

Schneider B, Herth W (1986) Distribution of plasma membrane rosettes and kinetics of cellulose formation in xylem development of higher plants Protoplasma 131:142-152

Schneider R, Hanak T, Persson S, Voigt CA (2016) Cellulose and callose synthesis and organization in focus, what's new? Curr Opin Plant Biol 34:9-16 doi:10.1016/j.pbi.2016.07.007

Schneider R et al. (2017) Two complementary mechanisms underpin cell wall patterning during xylem vessel development Plant Cell 29:2433-2449 doi:10.1105/tpc.17.00309

Sethaphong L, Davis JK, Slabaugh E, Singh A, Haigler CH, Yingling YG (2016) Prediction of the structures of the plant-specific regions of vascular plant cellulose synthases and correlated functional analysis Cellulose 23:145-161

Sethaphong L, Haigler CH, Kubicki JD, Zimmer J, Bonetta D, DeBolt S, Yingling YG (2013) Tertiary model of a plant cellulose synthase Proc Natl Acad Sci U S A 110:7512-7517 doi:10.1073/pnas.1301027110

Shim I, Law R, Kileeg Z, Stronghill P, Northey JGB, Strap JL, Bonetta D (2018) Alleles causing resistance to isoxaben and flupoxam highlight the significance of transmembrane domains for CESA protein function Front Plant Sci 9:1152

Slabaugh E, Davis JK, Haigler CH, Yingling YG, Zimmer J (2014) Cellulose synthases: new insights from crystallography and modeling Trends Plant Sci 19:99-106 doi:10.1016/j.tplants.2013.09.009

Song D, Shen J, Li L (2010) Characterization of cellulose synthase complexes in Populus xylem differentiation New Phytol 187:777-790

Speicher TL, Li PZ, Wallace IS (2018) Phosphoregulation of the plant cellulose synthase complex and cellulose synthase-like proteins Plants 7 doi:10.3390/plants7030052

Taylor JB (1957) The water solubilities and heats of solution of short chain cellulosic oligosaccharides Transactions of the Faraday Society 53:1198-1203

Taylor NG (2007) Identification of cellulose synthase AtCesA7 (IRX3) in vivo phosphorylation sites--a potential role in regulating protein degradation Plant Mol Biol 64:161-171

Taylor NG, Gardiner JC, Whiteman R, Turner SR (2004) Cellulose synthesis in the Arabidopsis secondary cell wall Cellulose 11:329-338

Taylor NG, Howells RM, Huttly AK, Vickers K, Turner SR (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis Proc Natl Acad Sci U S A 100:1450-1455

Thomas LH, Forsyth VT, Martel A, Grillo I, Altaner CM, Jarvis MC (2014) Structure and spacing of cellulose microfibrils in woody cell walls of dicots Cellulose 21:3887-3895

Timpa JD, Triplett BA (1993) Analysis of cell-wall polymers during cotton fiber development Planta 189:101-108

Tsekos I (1999) The sites of cellulose synthesis in algae: Diversity and evolution of cellulose-synthesizing enzyme complexes J Phycol 35:635-655

Turner S, Kumar M (2018) Cellulose synthase complex organization and cellulose microfibril structure Phil Trans R Soc A 376 doi:10.1098/rsta.2017.0048

Turner SR, Somerville CR (1997) Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall Plant Cell 9:689-701

Umemura M, Yuguchi Y, Hirotsu T (2004) Interaction between cellooligosaccharides in aqueous solution from molecular dynamics

simulation: comparison of cellotetraose, cellopentaose, and cellohexaose J Phys Chem A 108:7063-7070

Vain T et al. (2014) The cellulase KORRIGAN is part of the cellulose synthase complex Plant Physiol 165:1521-1532 doi:10.1104/pp.114.241216

Vandavasi VG et al. (2016) A structural study of CESA1 catalytic domain of Arabidopsis cellulose synthesis complex: evidence for CESA trimers Plant Physiol 170:123-135 doi:10.1104/pp.15.01356

Vergara CE, Carpita NC (2001) b-D-glycan synthases and the *CesA* gene family: lessons to be learned from the mixed-linkage (1→3),(1→4)b-D-glucan synthase Plant Mol Biol 47:145-160

Wang J, Howles PA, Cork AH, Birch RJ, Williamson RE (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

Wang T, Hong M (2016) Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls J Exp Bot 67:503-514 doi:10.1093/jxb/erv416

Wang T, Park YB, Cosgrove DJ, Hong M (2015) Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis Primary Cell Walls: Evidence from Solid-State Nuclear Magnetic Resonance Plant Physiol 168:871-884 doi:10.1104/pp.15.00665

Wang T, Yang H, Kubicki JD, Hong M (2016) Cellulose structural polymorphism in plant primary cell walls investigated by high-field 2D solid-state NMR spectroscopy and density functional theory calculations Biomacromolecules 17:2210-2222 doi:10.1021/acs.biomac.6b00441

Watanabe Y et al. (2015) Visualization of cellulose synthases in Arabidopsis secondary cell walls Science 350:198-203 doi:10.1126/science.aac7446

Watanabe Y et al. (2018) Cellulose synthase complexes display distinct dynamic behaviors during xylem transdifferentiation Proc Natl Acad Sci U S A 115:E6366-E6374 doi:10.1073/pnas.1802113115

Wickett NJ et al. (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants Proc Natl Acad Sci U S A 111:E4859-4868 doi:10.1073/pnas.1323926111

Willison JHM (1976) Plasmodesmata: a freeze-fracture view Can J Bot 54:2842-2847

Wong HC et al. (1990) Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA 87:8130-8134

Woodley M, Mulvihill A, Fujita M, Wasteneys GO (2018) Exploring microtubule-dependent cellulose-synthase-complex movement with high precision Particle Tracking Plants 7 doi:10.3390/plants7030053

Xi W, Song D, Sun J, Shen J, Li L (2017) Formation of wood secondary cell wall may involve two type cellulose synthase complexes in Populus Plant Mol Biol 93:419-429 doi:10.1007/s11103-016-0570-8

Yang H, Zimmer J, Yingling YG, Kubicki JD (2015) How cellulose elongates--A QM/MM study of the molecular mechanism of cellulose polymerization in bacterial CESA J Phys Chem B 119:6525-6535 doi:10.1021/acs.jpcb.5b01433

Yin Y, Huang J, Xu Y (2009) The cellulose synthase superfamily in fully sequenced plants and algae BMC Plant Biol 9:99

Yin Y, Johns MA, Cao H, Rupani M (2014) A survey of plant and algal genomes and transcriptomes reveals new insights into the evolution and function of the cellulose synthase superfamily BMC Genomics 15:260 doi:10.1186/1471-2164-15-260

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes Mol Biol Evol 21:809-818

Yu L, Sun J, Li L (2013) PtrCel9A6, an endo-1,4-β-glucanase, is required for cell wall formation during xylem differentiation in populus Mol Plant 6:1904-1917 doi:10.1093/mp/sst104

Zhang T, Mahgsoudy-Louyeh S, Tittmann B, Cosgrove DJ (2014) Visualization of the nanoscale pattern of recently-deposited cellulose microfibrils and matrix materials in never-dried primary walls of the onion epidermis Cellulose 21:853-862

Zhang T, Zheng Y, Cosgrove DJ (2016) Spatial organization of cellulose microfibrils and matrix polysaccharides in primary plant cell walls as imaged by multichannel atomic force microscopy Plant J 85:179-192 doi:10.1111/tpj.13102

Zhang X et al. (2018) Cellulose synthase stoichiometry in aspen differs from Arabidopsis and Norway spruce Plant Physiol 177:1096-1107 doi:10.1104/pp.18.00394