Molecular Breeding of Woody Plants
N. Morohoshi and A. Komamine, editors.

BIOSYNTHESIS OF CELLULOSE

Inder M. Saxena* & R. M. Brown Jr.

Section of Molecular Genetics and Microbiology, School of Biological Sciences, University of Texas at Austin, Austin, TX 78712, U.S.A.

ABSTRACT

Cellulose is synthesized by a large number of living organisms ranging from the bacterium *Acetobacter xylinum* to forest trees. *A. xylinum* produces abundant amounts of cellulose and this bacterium has been used as a model system for studies on cellulose biosynthesis and structure of the cellulose product. Cellulose is synthesized by the enzyme cellulose synthase, a membrane protein that catalyzes the direct polymerization of glucose from the substrate UDP-glucose into a cellulose product. Genes for cellulose synthases have been identified from many bacteria, *Dictyostelium discoideum*, and higher plants. Analysis of the predicted protein sequences has allowed identification of conserved residues in cellulose synthases from different organisms. The conserved residues are found in the globular region of the cellulose synthases. Using site-directed mutagenesis experiments we have shown that the conserved amino acid residues are required for cellulose synthase activity in *A. xylinum*. Although cellulose synthase activity can be monitored *in vitro* using membrane fractions from *A. xylinum*, it is not easy to monitor this activity when membrane fractions from plants are used. We have initiated experiments to analyze cellulose synthases from plants in *A. xylinum* in an effort to characterize the different cellulose synthases, for example the ones involved in cellulose biosynthesis during primary cell wall formation and those that are active during secondary wall synthesis. A general model describing the possible sequence of events in the cellulose synthase catalytic site will be presented to provide sufficient details not only into the biosynthesis of cellulose but also other polysaccharides.

KEYWORDS

Cellulose, biosynthesis, cellulose synthase, *Acetobacter xylinum*

INTRODUCTION

Cellulose is a major polysaccharide produced mainly by plants. A number of other organisms also synthesize cellulose and the capacity to synthesize this polysaccharide may be present in a larger variety of organisms than presently known. For example, sequences similar to cellulose synthase and other proteins identified in the cellulose-synthesizing operon of *A. xylinum* have been identified in *E. coli* and other bacteria even though no cellulose production has been reported in these bacteria [1]. Although plants produce cellulose as a major product, genes for cellulose synthase were identified only in the last few years [2,3]. More recently, the identification of a large gene

family of cellulose synthases and related proteins in plants [4] has allowed a number of interesting observations to be made in terms of the expression of different genes in different tissues and in the requirement of different genes for the synthesis of cellulose in the primary and secondary cell wall [5-7]. Even before multiple cellulose synthase genes were identified in plants, two cellulose synthases genes were identified in *A. xylinum* [8]. Whether the enzymes coded by the different cellulose synthase genes differ in their catalytic activity, their regulation, and their association with similar or different catalytic subunits and accessory protein remains to be understood.

CELLULOSE SYNTHASES

Cellulose is synthesized by the enzyme cellulose synthase, and in all cases this enzyme is predicted to be a membrane protein that utilizes UDP-glucose as the sugar donor in a direct transfer reaction [9]. The glucan chain is elongated from the non-reducing end [10] processively and although suggestions for the requirement of a primer by cellulose synthase have been made [11], no primer has yet been identified. The cellulose synthases from plants show similarity to the *A. xylinum* cellulose synthase in a globular region that contains the putative catalytic region and the conserved amino acid residues. Cellulose synthase activity from *A. xylinum* can be assayed *in vitro* and the enzyme has been partially purified. On the other hand, cellulose synthase has not been sufficiently purified from plant membranes and the analysis of the cellulose product is complicated because of the synthesis of other polysaccharides, especially callose. In terms of regulation of the enzyme activity, the *A. xylinum* enzyme is specifically activated by c-di-GMP and does not seem to require any additional factors. In biochemical studies with membrane preparations from plants it has not been possible to identify any specific activator of cellulose synthase; however, cellobiose is required for increased uptake of glucose from UDP-glucose into cellulose (characterized as the ANR-insoluble product) in *in vitro* reactions. Since the mechanism by which cellulose synthase performs glycosyl transfer is predicted to be essentially the same in bacteria and plants, we are interested in understanding this mechanism and the various modes of regulation involved in the activity of cellulose synthase from different sources.

The process of cellulose biosynthesis can be viewed as a number of simultaneous or sequential events that requires an active site in the cellulose synthase and may involve accessory factors and proteins. We have investigated the role of different regions of the *A. xylinum* cellulose synthase in a number of experiments to determine the amino acid residues essential for enzyme activity (by site-directed mutagenesis experiments) [12,13] and regions required for enzyme activity (by transposon insertion mutagenesis and deletion mutagenesis) [14].

The different regions and residues of cellulose synthase that may be essential for the enzyme activity include:

(a) The catalytic region, which includes the substrate (UDP-glucose) binding residues and the catalytic residue (that functions as a base). We investigated the role of the conserved residues in the D, D, D, QXXRW motif, identified in β-glycosyltransferases, by site-directed mutagenesis. Replacement of the conserved aspartic acid residues and the conserved residues in the QXXRW motif led to a loss of cellulose synthase activity in *A. xylinum*.

(b) Region(s) for binding of the growing glucan chain (and containing the glucose residue in the growing chain that functions as an acceptor). This region may be essential for the processivity of cellulose synthase and other processive glycosyltransferases.

(c) Transmembrane regions for membrane insertion (so far no cellulose synthase activity has been observed in the globular region that has been expressed as a cytosolic protein).

(d) Activator-binding site – in *A. xylinum*, the c-di-GMP- binding site may be present in either the same polypeptide chain that carries the catalytic region or in a separate polypeptide chain. A deletion of amino acid residues 707 – 1108 of the AcsAB protein, where this region corresponds to the c-di-GMP-binding region, shows no activity and no reaction with the antibody against the 93-kDa polypeptide (Saxena and Brown, unpublished observations).

(e) Protein-protein interaction sites for binding one subunit with homologous or/and heterologous subunit(s). In *A. xylinum*, cellulose synthase activity is still observed even when one of the two cellulose synthase genes is mutagenized by insertion [8]. Moreover insertion in the *acsAII* gene does not result in a phenotypic change (the cells are still able to make a ribbon of cellulose) suggesting that the AcsAB and the AcsAII proteins are not required together for forming the cellulose synthesizing complex in *A. xylinum*. When insertions take place in the *acs* operon genes, the *acsAII* gene and operon is not able to assemble a cellulose synthesizing complex. However, in plants multiple cellulose synthase catalytic subunits may interact to give rise to homomeric or heteromeric structures (see Figure 1). The interaction of these subunits has been suggested based on mutant and protein interaction analysis in *A. thaliana* [6,7].

(f) Residues that bind to Mg^{2+} in the catalytic region.

(g) Residues that function in a ratchet-like mechanism for movement of the growing polymer chain from the active site.

CELLULOSE SYNTHESIZING COMPLEXES

Is the smallest unit in a cell responsible for producing a cellulose I microfibril a single cellulose synthase enzyme or an aggregate of cellulose synthase catalytic subunits? The production of non-crystalline cellulose and cellulose II probably can take place from single catalytic subunits of cellulose synthase as observed in *in vitro* reactions and in mutants where the organization of the subunits has been disturbed. However, assembly of cellulose I microfibrils probably require other proteins that allow assemblage of the catalytic subunits for efficient synthesis and export of the cellulose product. In *A. xylinum*, these proteins have been identified by mutant analysis; however, they have not been characterized biochemically. In higher plants, a larger number of proteins may be required for cellulose I biosynthesis, including more than one kind of cellulose synthase catalytic subunit (Figure 1). Apart from the catalytic subunits, other proteins may be required for the processing of the proteins before they are exported to the plasma membrane. In all cases, cellulose I microfibrils have been found to be associated with organized structures observed on the cell membrane. These structures are believed to be the cellulose synthesizing complexes. They are referred to as terminal complexes (TCs) and are visualized as rosettes or linear structures by freeze-fracture electron microscopy.

The nature of the components in these structures has so far only been inferred based on the attachment of cellulose microfibril with these structures. Recently, techniques that allow labeling of freeze-fracture replicas by antibodies have led to the localization of some of the components in these complexes. In plants, the cellulose synthase catalytic subunit has been localized to the rosette structure [15] and an activator (c-di-GMP) binding protein has been localized to the linear complexes observed in *A. xylinum*. So far, no other protein has been found to be associated with the cellulose synthases in the rosette complex of higher plants; however suggestions have been made for the role of a number of proteins that may associate with the catalytic subunits in the rosette complex. It will certainly be interesting to determine the nature of interaction between the different catalytic subunits in the organization of the rosette. Do the different subunits interact directly with each other or do they require the assistance of other proteins? Certainly the variable regions in cellulose synthases from plants may provide sites for interaction with other proteins. Since the amino acid sequence in the variable regions vary amongst the different cellulose synthases, these regions may provide sites for specific interactions with other proteins. Modeling of the globular region of cellulose synthase from cotton suggests that the variable regions observed in higher plant cellulose synthases are present as loops on the surface where they are accessible for interaction with other proteins [13].

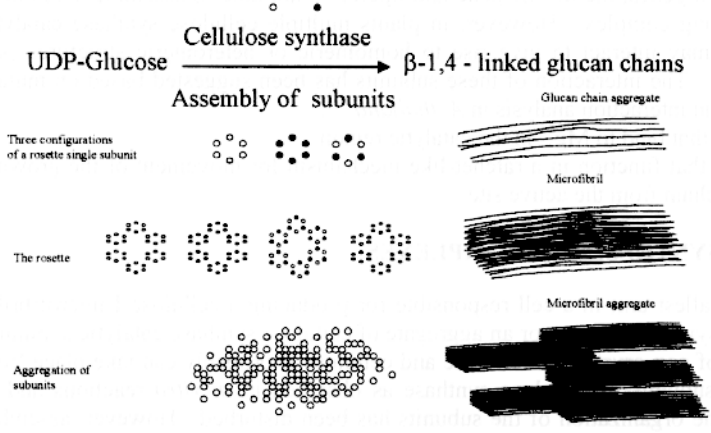

FIGURE 1. Similar or different cellulose synthase catalytic subunits may associate to form rosettes that give rise to a cellulose microfibril. In some cases, during the formation of secondary cell wall, subunits may aggregate without forming rosettes and these aggregates can form microfibril bundles.

FUNCTIONAL ANALYSIS OF CELLULOSE SYNTHASES AND CELLULOSE SYNTHASE-LIKE PROTEINS FROM HIGHER PLANTS

The globular region of cellulose synthases and cellulose synthase-like proteins contain the conserved amino acid residues involved in catalysis and this region is homologous to the globular region of the cellulose synthase from *A. xylinum*. In order to understand the function of the globular region of the plant proteins and develop a system where specific residues or regions can be altered for understanding their function, we have developed a system for substituting the globular region of the bacterial cellulose synthase with the homologous region from the plant cellulose synthases and cellulose synthase-like proteins. The goal is to produce chimeric proteins that have the transmembrane and regulatory (c-di-GMP-binding) regions of the *A. xylinum* cellulose synthase and the catalytic region from plant proteins. Expression of these chimeric proteins in *A. xylinum* would allow for their systematic analysis. We have chosen to express the globular region from two different cellulose synthases and a cellulose synthase-like protein using this system at present (Figure 2). Expression of the chimeric protein using the globular region of the cellulose synthase-like protein has been observed in *A. xylinum*; however, this chimeric protein was not found to produce any cellulose in the *in vitro* assay validating the suggestion that these proteins are probably involved in the synthesis of non-cellulosic polysaccharides.

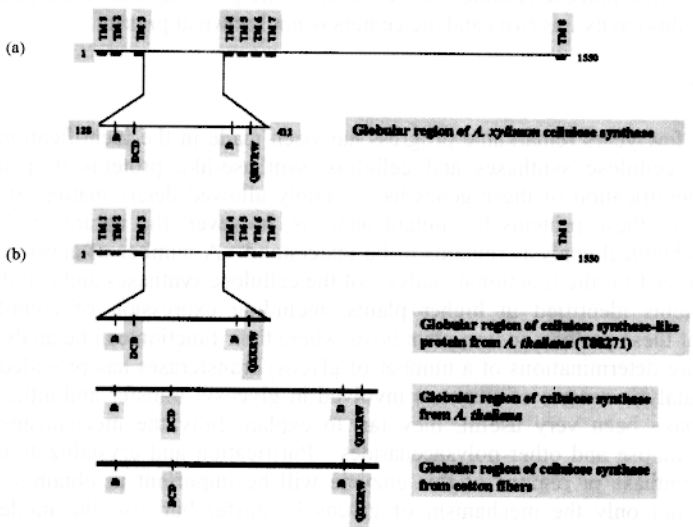

FIGURE 2 Substitution of the (a) globular region of cellulose synthase of *A. xylinum* with (b) globular region of cellulose synthases and cellulose synthase-like proteins from plants.

STRUCTURE AND FUNCTION OF CELLULOSE SYNTHASES AND OTHER GLYCOSYLTRANSFERASES

Interpretations of recently determined structures of glycosyltransferases (of unknown function or those that attach a single sugar residue to an acceptor molecule) suggest a single catalytic center in these enzymes [16]. This proposal is also extended to enzymes for the synthesis of cellulose, hyaluronan, and other β-linked polysaccharides. Although the single active site model may be useful in understanding non-processive addition of sugar residues, it does not explain the 180° rotation of glucose residues in the glucan chains of cellulose or the addition of two different sugars with two different linkages in hyaluronan. A Class II hyaluronan synthase from *Pasteurella multocida* has been shown to have duplication of domain A and the third conserved aspartic acid providing it with two catalytic centers in a single polypeptide chain [17]. Duplicated domains have not been identified by sequence analysis in other processive β-glycosyltransferases and this has led to the thinking that these enzymes have a single catalytic center. However, the two catalytic centers do not have to be generated from duplicated domains. The addition of two sugar residues in a sequential or simultaneous fashion probably requires two catalytic centers that may be present in the same polypeptide chain or two polypeptide chains. So far, no crystal structure has been obtained for any processive β-glycosyltransferase. In our model of the globular region of cellulose synthase, an extended catalytic cavity has been observed. This cavity can accommodate two UDP-glucose residues as well as hold the growing end of the glucan chain [13]. Whether this cavity has two catalytic centers is not known at present.

CONCLUSIONS

In the last few years remarkable progress has been made in the identification of genes coding for cellulose synthases and cellulose synthase-like proteins in plants. Progress in the identification of these genes has certainly allowed determination of the function of some of these proteins by mutant analysis; however, the nature of these proteins from biochemical analysis remains to be understood. A number of approaches will have to be utilized for the functional analysis of the cellulose synthases and cellulose synthase-like proteins identified in higher plants, including expression of complete proteins or parts of these proteins in non-plant hosts where their function can be analyzed. The recent structure determinations of a number of glycosyltransferases has provided an insight into the catalytic centers of enzymes involved in glycosyl transfer and although these structures have been very useful, they fail to explain fully the mechanisms of biosynthesis of cellulose and other polysaccharides. Purification and crystallization of native cellulose synthase or regions of this enzyme will be important in obtaining an understanding of not only the mechanism of glycosyl transfer but also the mode of processivity and the manner in which alternate residues are inverted in the glucan chains of cellulose. Certainly all these studies will aid in understanding the conditions that regulate biosynthesis of cellulose and allow for production of cellulose with desirable properties.

ACKNOWLEDGEMENTS

The authors acknowledge grant support from the U. S. Department of Energy (DE-FG03-94-ER20145).

REFERENCES

1. H. J. Sofia, V. Burland, D. L. Daniels, G. Plunkett III & F. R. Blattner, Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes, *Nucleic Acids Res.*, **1994**, 22, 2576-2586.
2. J. R. Pear, Y. Kawagoe, W. E. Schreckengost, D. P. Delmer & D. M. Stalker, Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase, *Proc. Natl. Acad. Sci. USA*, **1996**, 93, 12637-12642.
3. T. Arioli, L. Peng, A. S. Betzner, J. Burn, W. Wittke, W. Herth, C. Camilleri, H. Höfte, J. Plazinski, R. Birch, A. Cork, J. Clover. J. Redmond & R. E. Williamson, Molecular analysis of cellulose biosynthesis in *Arabidopsis*, *Science*, **1998**, 279, 717-720.
4. T. A. Richmond & C. R. Somerville, The cellulose synthase superfamily, *Plant Physiology*, **2000**, 124, 495-1324.
5. N. Holland, D. Holland, T. Helentjaris, K. Dhugga, B. Xoconostle-Cazares & D. P. Delmer, A comparative analysis of the plant cellulose synthase (CesA) gene family, *Plant Physiology*, **2000**, 123, 1313-498.
6. M. Fagard, T. Desnos, T. Desprez, F. Goubet, G. Refregier, G. Mouille, M. McCann, C. Rayon, S. Vernhettes & H. Höfte, *PROCUSTE1* encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of Arabidopsis, *Plant Cell*, **2000**, 12, 2409-2423.
7. N. G. Taylor, S. Laurie & S. R. Turner, Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in Arabidopsis, *Plant Cell*, **2000**, 12, 2529-2539.
8. I. M. Saxena & R. M. Brown, Jr., Identification of a second cellulose synthase gene (*acsAII*) in *Acetobacter xylinum*, *J. Bacteriol.*, **1995**, 177, 5276-5283.
9. I. M. Saxena & R. M. Brown, Jr., Cellulose synthases and related enzymes, *Curr. Op. Pl. Biol.*, **2000**, 3, 523-531.
10. M. Koyama, W. Helbert, T. Imai, J. Sugiyama & B. Henrissat, Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose, *Proc. Natl. Acad. Sci. USA*, **1997**, 94, 9091-9095.
11. W. Lukowitz, T. C. Nickle, D. W. Meinke, R. L. Last, P. L. Conklin & C. R. Somerville, *Arabidopsis cyt1* mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked glycosylation for cellulose biosynthesis, *Proc. Natl. Acad. Sci. USA*, **2001**, 98, 2262-2267.
12. I. M. Saxena & R. M. Brown, Jr., Identification of a cellulose synthase(s) in higher plants: Sequence analysis of processive β-glycosyltransferases with the common motif 'D,D,D35Q(R,Q)XRW,' *Cellulose*, **1997**, 4, 33-49.

13. I. M. Saxena, R. M. Brown, Jr. & T. Dandekar, Structure-function characterization of cellulose synthase: Relationship to other glycosyltransferases, *Phytochemistry*, **2001** (in press).
14. I. M. Saxena, K. Kudlicka, K. Okuda & R. M. Brown, Jr., Characterization of genes in the cellulose-synthesizing operon (*acs* operon) of *Acetobacter. xylinum*: Implications for cellulose crystallization, *J. Bacteriol.*, **1994**, 176, 5735-5752.
15. S. Kimura, W. Laosinchai, T. Itoh, X. Cui, C. R. Linder & R. M. Brown, Jr., Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*, *Plant Cell*, **1999**, 11, 2075-2085.
16. S. J. Charnock, B. Henrissat & G. J. Davies, Three-dimensional structures of UDP-sugar glycosyltransferases illuminate the biosynthesis of plant polysaccharides, *Plant Physiology*, **2001**, 125, 527-531.
17. W. Jing & P. L. DeAngelis, Dissection of the two transferase activities of the *Pasteurella multocida* hyaluronan synthase: two active sites exist in one polypeptide, *Glycobiology*, **2000**, 10, 883-889.