Cellulose Structure and Biosynthesis: What is in Store for the 21st Century?

R. MALCOLM BROWN, JR.

Molecular Genetics and Microbiology, University of Texas at Austin, Austin, Texas 78712

ABSTRACT: This article briefly summarizes historical developments in fundamental research related to the structure and biosynthesis of cellulose. Major advances concerning the structure of cellulose include the discovery of a new suballomorph of cellulose I, the lattice imaging of glucan chains showing no fringe micelle structure, parallel chain orientation in cellulose I, and the discovery of nematic ordered cellulose. Major advances in biosynthesis include the discovery of the terminal synthesizing complex, the isolation and purification of cellulose synthase, the *in vitro* synthesis of cellulose I, and synthetic cellulose assembly. This article focuses on recent advances in molecular biology with cellulose, including the cloning and sequencing of cellulose synthase genes from bacteria, cyanobacteria, and vascular plants; proof of the terminal synthesizing complex as the site of the catalytic subunit of cellulose synthase; cellulose and callose synthase expression during growth and development; and phylogenetic aspects of cellulose synthase evolution. This article concludes with thoughts about future uses for the accumulating genetic information on cellulose biosynthesis for textiles and forest products and discusses possibilities of new global resources for cellulose production. © 2003 Wiley Periodicals, Inc. J Polym Sci Part A: Polym Chem 42: 487–495, 2004

Keywords: cellulose; structure; biosynthesis; molecular genetics; new cellulose crops; biomaterials; biopolymers

INTRODUCTION

I am pleased to dedicate this review to my colleague and good friend, Professor Otto Vogl, who has made seminal contributions to the field of polymer chemistry (Fig. 1). I am also honored to be invited to speak in Vienna on a subject very dear to my heart: cellulose (Fig. 2). This biopolymer is the most abundant macromolecule on earth, with more than 10^{11} tons estimated to be synthesized each year on our planet.[1] Given the vast quantity of this product and the importance of this material to humankind, it is indeed re-markable that more attention has not been devoted to the scientific exploration of cellulose, especially with respect to its structure and biosynthesis. Although this was one of the earliest molecules to be subjected to X-ray analysis,[2] and considerable efforts have been directed toward its structure during the past century, only now are we beginning to unlock the secrets of this fascinating biopolymer.

I will present a brief overview of cellulose, finishing with what is in store for us as we move into the 21st century. Indeed, in the era of synthetic polymers, cellulose continues to be a dominant force in many vital industries, including the pulp and paper, lumber, and textile industries. Will it maintain this dominance? What factors will come into play as we develop post-genomic-era engineering methods for new cellulose crops? I will

Correspondence to: R. M. Brown, Jr. (E-mail: rmbrown@mail.utexas.edu)

Figure 1. Otto Vogl visiting the author's home in Austin, Texas, in May 2001.

cover these and other relevant questions in this brief review. The references in this article will direct readers to the most recent literature in the field for further study. The reader is also encouraged to visit my website for additional references and information.[3]

DIVERSITY OF CELLULOSE: WHERE IS IT FOUND?

Cellulose is synthesized by a great diversity of living organisms. In fact, we usually associate our resources as trees and cotton plants, and rightfully so, because these are the major sources of the industrial production and processing of cellulose. However, cellulose is synthesized by bacteria and prokaryotes as well (e.g., *Acetobacter*, *Rhizobium*, and *Agrobacterium*). Even some pathogenic bacteria have been found to synthesize cellulose.[4] One of the most interesting recent discoveries is that the most ancient forms of life on earth, represented by the cyanobacteria, also synthesize cellulose[5] (Fig. 3). Of course, eukaryotic organisms produce cellulose, and certain fungi, amoe-

bae, cellular slime molds, and green algae have representatives that produce perfectly good cellulose.[6] In fact, one of nature's most perfect crystalline forms of cellulose comes from a green alga, *Valonia ventricosa*; more than 1200 glucan chains produce very highly ordered crystalline microfibrils with a degree of polymerization (dp) of more than 25,000.[1,7] Cellulose is also synthesized by freshwater and marine algae, including land plants such as mosses, ferns, angiosperms, and gymnosperms.[1] Cellulose is even made by some animals, the tunicates.[8] Why is there all this diversity? This is highly suggestive of an ancient evolutionary process that has culminated in the production of diverse forms of cellulose among a great variety of organisms on earth. Therefore, it is logical to conclude that on a functional basis, cellulose is essential for many cellular functions.

DIVERSITY OF CELLULOSE BIOSYNTHESIS: HOW LONG HAS IT BEEN AROUND?

Cellulose biosynthesis must be an ancient process. As we learn more about the phylogeny of living organisms and the origin of life on earth, we cannot come away unimpressed by the extreme age of life on our planet. It has been around for more than 3.5 billion of the 5–6 billion years that the earth has existed.[9] The earliest known life forms may have been photosynthetic or chemosynthetic prokaryotic organisms. The stromatolites found in Canada and Australia[9] are composed of algal mats, mostly cyanobacteria. Cyanobacteria have remained relatively unchanged for billions of years. Recently, David Nobles in my laboratory found that cellulose biosynthesis is widespread among the various cyanophycean genera,[5] and this implies that cellulose has been around a long time and has probably dramatically influenced the evolution of life. Did the first form of life produce cellulose? We do not have an answer to this provocative question, but as we learn more through gene sequencing and genome analysis, it may be possible to reconstruct life's tree back to an even earlier time than we know from fossil evidence, a time when life may have just begun. I think that cellulose may have had an important role in the successful formation of life on earth, particularly in sustaining life forms in earth's harsh primitive atmosphere. I have discussed this in a preliminary way recently, and I believe that cellulose may have offered protection from dangerous ultraviolet radiation when the
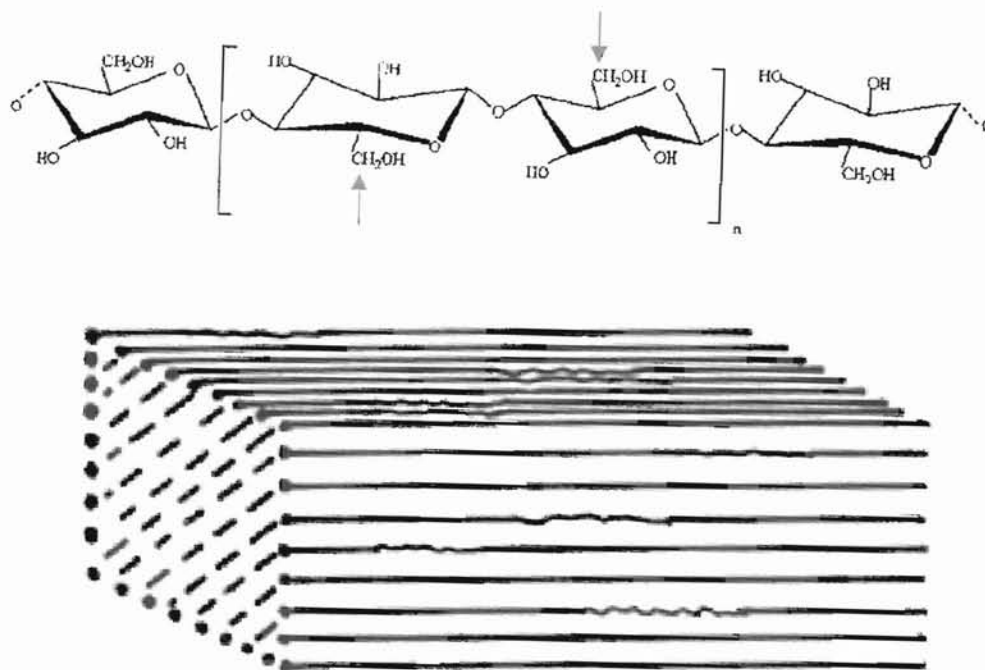
Figure 2. The top image is the structural formula of cellulose. The arrows point to the basic repeat unit, which is a cellobiose molecule. The molecule has a twofold screw–axis symmetry. The bottom image is a schematic diagram; the lines show the individual glucan chain polymers that constitute the crystalline cellulose I microfibril.

earth still had a reducing atmosphere devoid of oxygen.[5] Cellular evolution is a hot new field of science that will certainly benefit from an understanding of the biosynthetic pathways and their roles in survival. That is not the major topic of this review, but it has sufficient merit that I would urge the reader to bookmark this section and to follow up on subsequent publications as they appear.

STRUCTURE OF NATIVE CELLULOSE

Native cellulose is defined as that cellulose made by living organisms. Of course, it can be altered by strong alkali treatments to produce other crystalline forms. Cellulose is chemically composed only of glucose monomers (Fig. 2). The linkage is β-1,4. Starch is composed only of glucose, but the linkage is α-1,4. The number of glucose monomeric units required to produce an insoluble product is about 8. Above that, the glucan chains have a greater affinity for one another than they do for the aqueous solvent. A typical number of glucose units in native cellulose depends on the source, such as primary or secondary cell walls. Primary cell wall cellulose polymers have about 8000 glu-

cose units per chain (dp = 8000). Secondary wall cellulose has a higher dp, up to 15,000.

Native cellulose is found in two crystalline forms, cellulose I and cellulose II. By far, most
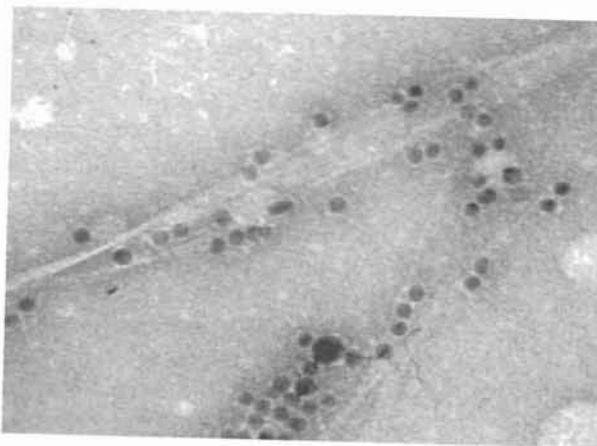
Figure 3. Cellulose microfibrils from a cyanophycean alga, *Nostoc muscorum*, negatively stained with uranyl acetate. The opaque electron spheres are 10-nm gold particles bound to the enzyme CBH-I, which specifically binds to the crystalline surface of cellulose I. Reprinted with permission from D. Nobles, D. Romanovicz, and R. M. Brown, Jr. 2001. Plant Physiology 127, 529–542. © 2001 American Society of Plant Biologists.
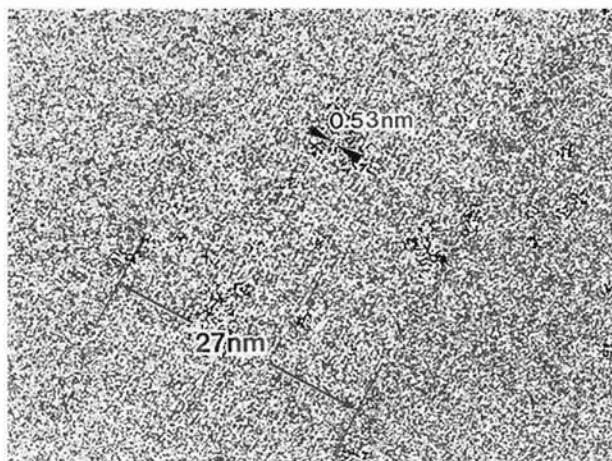
Figure 4. Lattice image of a microfibril from a green alga, *Boergesenia forbsei*. The signal-to-noise ratio is not high for the imaging of the single glucan chains, but they can be detected. The 0.53-nm glucan chain spacings are revealed in this image. Reprinted with permission from S. Kuga and R. M. Brown, Jr. 1987. J Electron Mic Tech 6, 349–356. © 1987 Oxford Press.

native cellulose exists as cellulose I. The glucan chains are oriented parallel in cellulose I,[10] but in cellulose II, the chains are antiparallel. The thermodynamically most stable allomorph is cellulose II, which has an additional hydrogen bond per glucose residue. Native cellulose II is rare and is found only in several algae as well as some bacteria.[11] Cellulose I allomorphs consist of distinct numbers of parallel glucan chains arranged to form the nanostructure known as a microfibril. The number and arrangement of the glucan chains are under the genetic control of the enzyme complex that synthesizes the cellulose. Microfibrils come in many shapes and sizes, ranging from thin membrane-like structures, as found in *Erythrocladia*,[12] to giant square structures with more than 1200 glucan chains, as found in *Valonia*,[1] or large rectangular microfibrils with many hundreds of glucan chains, as found in *Boergesenia* (Fig. 4). Furthermore, the reducing ends of glucan chains in a microfibril are found in the tips of microfibrils away from the site of synthesis on the cell.[10]

I might also add that a noncrystalline form of cellulose has been recently described. It is called *nematic ordered cellulose*.[13] This cellulose is highly ordered but noncrystalline and is produced from solution in lithium dimethylacetamide by the slow introduction of water vapor (and controlled hydrogen bonding). With high-resolution transmission electron microscopy (HRTEM), it is possible to image single glucan chains in this cellulose because they are surrounded by the uranyl acetate negative stain (Fig. 5).

The history of the structure of cellulose is long and contorted. One of the earlier favored models for native cellulose was a microfibril with a crystalline core surrounded by a fringe micelle.[14] We know from recent lattice image studies using HRTEM[15] and atomic force microscopy (AFM)[16] that the crystalline arrangement of the glucan chains is constant and subtends the entire microfibril structure. There are no fringe micelles. An image of the lattice structure of glucan chains in a single microfibril of *Boergesenia* is shown in Figure 4. The structure of cellulose II is not so straightforward. The glucan chains that we know are antiparallel, but are they long chains, or are they folded chains? One relatively recent study[17] showed that in a mutant of *Acetobacter*, cellulose II was arranged in the form of folded chains. The situation may be very different in a nonnative form of cellulose II known as rayon, which is produced from the reprecipitation of alkali-solubilized cellulose. Rayon is known to have a relatively low dp (ca. 300–500), and the nanostruc-
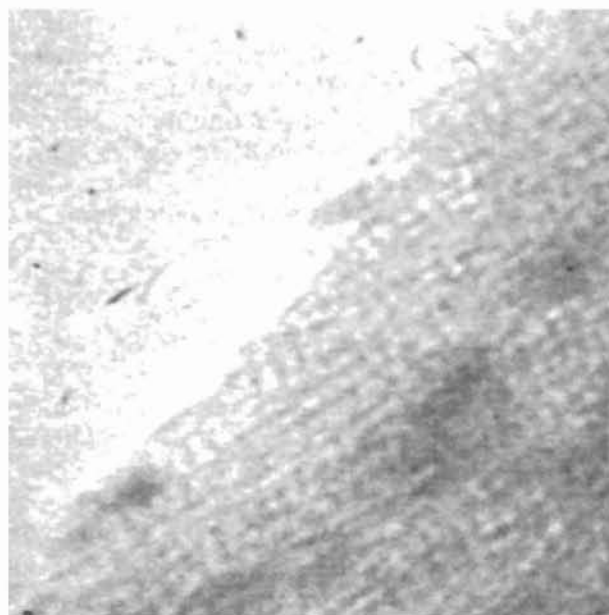
Figure 5. Nematic ordered cellulose imaged by HRTEM after negative staining with uranyl acetate. The individual glucan chains are ordered but are not tightly associated to form a crystalline phase. Reprinted with permission from T. Kondo, E. Togawa, and R. M. Brown, Jr. 2001. Biomacromolecules 2, 1324–1330. © 2001 American Chemical Society.
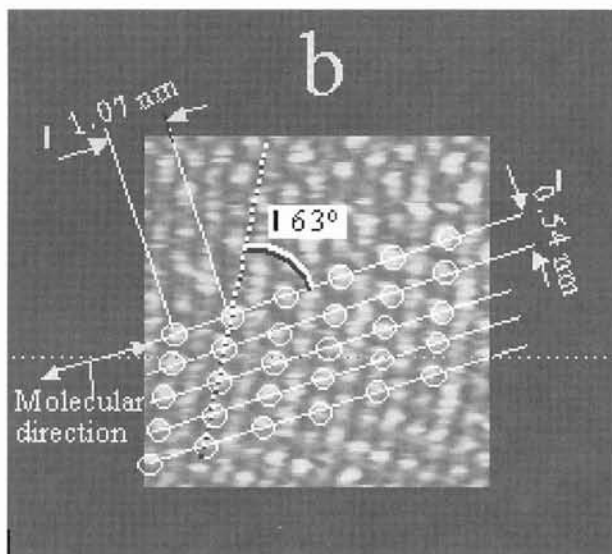
Figure 6. AFM height image of a bacterial cellulose subunit surface acquired by the contact mode in air, which may correspond to the intermolecular and intramolecular periodicity for the (100) crystal face of the triclinic cellulose Iα form of cellulose (photograph and preparation by R. M. Brown and T. Kondo).

tures are not imaged as long microfibrils but rather as short fibrils with tapering ends.[18]

Native crystalline cellulose I also has two different suballomorphs, cellulose Iα and cellulose Iβ.[19] Of these, cellulose Iβ is the more thermodynamically stable, and it is rarely synthesized in nature in a pure form, with the exception of tunicates. Cellulose Iα exists as a single-chain triclinic unit cell, whereas cellulose Iβ has a two-chain monoclinic unit cell. These structures have been imaged with HRTEM[16] and AFM (Fig. 6).

FUNCTION OF NATIVE CELLULOSE

Although it may appear obvious that the structure of cellulose is for the protection of a cell, I would like to briefly elaborate to provide new insight into this concept. Certainly, there can be no argument that a coating or covering of cellulose to produce a cell wall is of great importance in protecting the delicate protoplasm from the environment. Equally important are the site of deposition and the patterns of the arrangement of microfibrils in the cell wall. All plant cells enlarge throughout their cell cycles. This is called growth. Plant cell growth can be random, isodiametric, or directional, depending on the arrangement of the

cellulose. For example, in tip growth in apical cells, root hairs, or pollen tubes, the polymer orientation is random. If a cell is genetically programmed to enlarge by producing a tube, the arrangement of cellulose microfibrils is critical to the growth leading to the elongation of the tube. The microfibrils are first deposited in the axis transverse to the cell's long axis; however, as the tube grows or enlarges by the forces of the turgor pressure from within, the transverse microfibrils shift toward the longitudinal axis of the cell. More frequently, the cellulose microfibrils become deposited in a helical fashion (described later with respect to an ordinary child's toy). When roots, stems, and hypocotyls elongate and grow, this is the result of the sums of individual elongations of the individual cells. Cellulose deposition often controls cell elongation by occurring in a specific pattern on the cell surface. The control of this deposition pattern is linked to the cell's cytoskeletal system, but for now, consider as a model for this concept, if you will, a Slinky. This toy is often used to jump down stairs. When a Slinky is pulled, it is elongated, but its diameter remains constant. The metal spiral coil merely changes its transverse orientation to a more parallel orientation. The diameter of the Slinky remains unchanged. In the real world of the plant cell, the turgor pressure is the driving force for cell enlargement or growth. Growth is defined as an irreversible increase in the cell volume. If one could place a balloon inside a Slinky and then blow it up, the balloon would only be able to expand to form a cylinder. In the real world, directed growth is cell elongation, as demonstrated by the Slinky model. A cellulose microfibril by itself cannot be extended. The nanostructure is a rigid rod; however, microfibrils can slide past neighboring microfibrils if a noncellulose grease is present. In plant cell walls, other polymers such as xyloglucans and rhamnoglacturonans form this grease. Cell wall expansion and growth is a very complex subject, and I will not describe it further, except to point out the central role for cellulose as the skeleton for maintaining and controlling the growth processes.

BIOSYNTHESIS OF NATIVE CELLULOSE: NATURE'S NANOMACHINE PAR EXCELLENCE

At first thought, it seems rather simple that a homopolymer with a specific linkage would re-

quire only a simple enzyme to polymerize glucose into β-1,4-linked chains. The case is not so simple, for as we have learned, glucan chains are highly ordered into nanostructures called microfibrils. How does the cell's machinery not only accomplish the simultaneous polymerization of multiple glucan chains but also arrange these into metastable crystalline nanostructures known as microfibrils?

Before the 1950s, no one had any good ideas about the biosynthetic machinery for cellulose; however, as biological electron microscopy approaches became more widespread, attention turned toward the role of the plasma membrane in cellulose biosynthesis. After all, the cellulose microfibrils were observed on the outside of the cell, yet they were in close contact with the cell's contents; therefore, it seemed logical that the cellular biosynthetic machinery also would be involved in the extracellular materials comprising the cell wall. In 1958, Roelofsen[20] proposed that the cellulose microfibril originated from an enzyme complex and that it was assembled by tip growth; however, no direct proof for such a structure had been forthcoming. In 1976, my graduate student David Montezinos and I discovered an enzyme complex associated with the ends of growing microfibrils in a green alga, *Oocystis apiculata*.[21] The development of an electron microscopy preparative method known as freeze fracture made this discovery possible. Living cells are rapidly frozen in Freon and then are placed in a high vacuum. They are fractured with a microtome knife, and the fracture surface is etched by vacuum sublimation. A platinum–carbon film evaporates on the etched surface. A carbon backing film is applied, and the sample is brought to room temperature and atmospheric pressure. The replica is removed from the sample, cleaned, and examined in a transmission electron microscope. The unique principle of freeze fracture is that the hydrophobic interiors of the cell membranes often are fractured, revealing the transmembrane proteins. Thus, images of ordered structures shown to be directly associated with impressions of cellulose microfibrils offer tangible evidence that a multienzyme complex in the plasma membrane is responsible for cellulose biosynthesis. In the specific instance with *Oocystis*, we found that the ordered enzyme complexes consisted of three linear rows of particle subunits (Fig. 7, top).

Since 1976, freeze fracture has revealed a multitude of enzyme complexes associated with cellulose microfibrils from a variety of organisms rang-
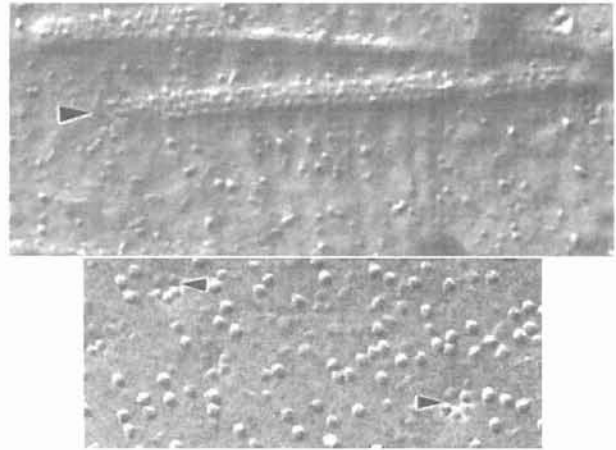

Figure 7. The top image is the E-fracture face of a plasma membrane of *Oocystis apiculata* showing a linear cellulose synthesizing TC. There are three rows of subunits that comprise this complex. The TC moves in the plane of the plasma membrane as the cellulose is generated and attaches to the inner cell wall layer. This alga was the first organism for which TCs were described. No fixation was used. Reprinted with permission from R. M. Brown, Jr. and D. Montezinos 1976. Proc Nat Acad Sci 73, 143–147. © 1976 National Academy of Science. The bottom image is the P-fracture face of a plasma membrane of a *Zea mays* root showing several rosette TCs. Unlike the linear TC, the subunits are hexagonally arranged. Rosette TCs are common to all vascular plants, including some algae (*Chara, Nitella, Spyrogyra*, and *Micrasterias*; courtesy of Susette Mueller)

ing from algae to tunicates to bacteria and plants.[22] These enzyme complexes are termed *terminal complexes* (TCs). In the case of *Oocystis*, TCs are called *linear* because they have three rows of particles. These represent not only the catalytic subunits involved in cellulose assembly but also ancillary proteins that help to position the glucan chains to achieve precise ordering in the metastable crystalline state. In vascular plants, which assemble cellulose in trees and cotton, for example, TCs are termed *rosette* because unlike *Oocystis*, they have six hexagonally arranged subunits in the plasma membrane (Figs. 7 and 8). The rosette TC in vascular plants was first described by my graduate student Susette Mueller in 1980.[23]

GENES INVOLVED IN CELLULOSE BIOSYNTHESIS

With tremendous progress in molecular genetics in the 1980s, the stage was set for gene sequenc-
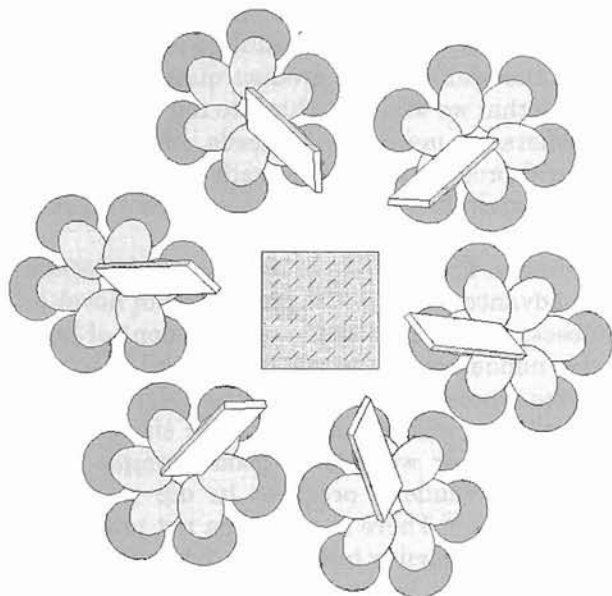
Figure 8. Hypothetical model of a rosette TC showing the possible origin of glucan chains. This TC produces a crystalline microfibril with approximately 36 glucan chains. Each subunit may synthesize a 6-glucan-chain sheet held together by van der Waals forces. The sheets then stack by hydrogen bonding to form the three-dimensional microfibril.[34–36] Reprinted with permission from R. M. Brown, Jr. and I. M. Saxena 2000. Plant Physiology and Biochemistry 38, 57–67. © 2000 Elsevier.

ing on a much larger scale, and in 1990, my colleagues Inder Saxena and Fong Chyr Lin and I first isolated and sequenced a gene for cellulose biosynthesis from *Acetobacter xylinum*.[24] During the past 13 years, genes for cellulose biosynthesis have been characterized from a great variety of organisms, including vascular plants.[25] I will not go into details here, except to note that the sequence of genes for cellulose biosynthesis has led to notable comparisons of highly conserved sequences among many organisms with the discovery by hydrophobic cluster analysis[26] that the aspartate residue is involved in the catalytic reaction and that two domains and three aspartates are required for cellulose biosynthesis. In addition, a QXXRW motif regulates the catalysis. Mutations of any part of this conserved sequence lead to abnormal plant growth and development and altered cellulose assembly.[27] Notably, the RSW-1 radial root swelling mutant of *Arabidopsis* greatly reduces crystalline cellulose and produces instead noncrystalline cellulose.[28]

An excellent example of a spin-off from the genetic data comes from the work of one of my former graduate students, Walairat Laosinchai, who cloned a cotton cellulose synthase gene, expressed it in *E. coli*, and then isolated the genetically engineered cotton cellulose synthase from *E. coli* and produced antibodies against it.[29] Using a difficult exploitation of freeze fracture known as fracture labeling, Kimura and Itoh,[30] along with members of my laboratory, were successful in immunolabeling the rosette TCs in a vascular plant. This study provided the best direct evidence to date that the rosette TC contains the catalytic subunit for cellulose synthase.

One of the most exciting aspects of cellulose synthesis research lies in the genetic revolution and what it can and is telling us about cellulose. As I mentioned previously, one of my graduate students, David Nobles identified cellulose synthesis in the cyanobacteria, the oldest living examples of ancient life on earth. The genetic sequences also tell us something else. They suggest but do not prove that cellulose synthase genes were introduced into a primitive eukaryotic cell from the cyanobacteria by lateral gene transfer. This is a most interesting point, because if this is really true, we then have an understanding of how land plants evolved, including photosynthetic eukaryotes! Obviously, more supporting evidence will be needed to confirm this concept, but the critical and key points already have been elucidated. Cellulose synthase genes can be injected into organisms and in this mode can dramatically alter successful survival. The evolutionary history of the origin of life on earth will, in my opinion, be elucidated by an understanding of cellulose biosynthesis. I realize that this is a far-reaching statement, but I believe that it is one that will be shown to be of immense importance in understanding not only cellular evolution but also the evolution of many eukaryotic life forms, including early animals. It is significant that the recent genome sequence of the tunicate *Ciona*[31] shows the possibility of the lateral gene transfer of cellulose synthase into this animal. This suggests that if this had not happened, the tunicate would never have evolved or developed and we would not have tunicates with cellulose! I expect great advances to be made in the next few years as we sequence more genomes and better understand the role of cellulose biosynthesis from these studies.

pers, research in progress, movies, links, and many other things related to cellulose. The revised Cellulose Electronic Network may be particularly useful. Please visit http://www.botany.utexas.edu.

4. Zogai, X.; Nimtz, N.; Rodhe, M.; Bokranzl, W.; Romling, U. Mol Microbiol 2001, 39, 1452.

5. Nobles, D.; Romanovicz, D.; Brown, R. M., Jr. Plant Physiol 2001, 127, 529.

6. Brown, R. M., Jr. In Experimental Phycology 1: Cell Walls and Surfaces, Reproduction, Photosynthesis; Wiessner, G.; Robinson, D. G.; Starr, R. C., Eds. Springer-Verlag: Berlin, 1990; p 20.

7. Shibazaki, H.; Kuga, S.; Onabe, F.; Brown, R. M., Jr. Polymer 1995, 36, 4971.

8. Kimura, S.; Itoh, T. Protoplasma 1995, 186, 24.

9. Schopf, J. W.; Walter, M. R. In The Biology of Cyanobacteria; Carr, N.; Whitton, B., Eds.; Blackwell: Boston, 1982.

10. Kuga, S.; Brown, R. M., Jr. Carbohydr Res 1988, 180, 345.

11. Canale-Parola, E. Bacteriol Rev 1970, 34, 82.

12. Okuda, K.; Tsekos, I.; Brown, R. M., Jr. Protoplasma 1994, 180, 49.

13. Kondo, T.; Togawa, E.; Brown, R. M., Jr. Biomacromolecules 2001, 2, 1324.

14. Frey-Wyssling, A. Science 1954, 119, 80.

15. Kuga, S.; Brown, R. M., Jr. J Electron Microsc Tech 1987, 6, 349.

16. Baker, A. A.; Helbert, W.; Sugiyama, J.; Miles, M. J. Biophys J 2000, 79, 1139.

17. Kuga, S.; Takagi, S.; Brown, R. M., Jr. Polymer 1993, 34, 3293.

18. Okuda, K.; Li, L.; Kudlicka, K.; Kuga, S.; Brown, R. M., Jr. Plant Physiol 1993, 101, 1131.

19. Atalla, R.; Vander Hart, D. L. Science 1984, 223, 283.

20. Roelofsen, P. A. Acta Bot Neerl 1958, 7, 77.

21. Brown, R. M., Jr.; Montezinos, D. Proc Natl Acad Sci USA 1976, 73, 143.

22. Tsekos, I. J Phycol 2000, 35, 635.

23. Mueller, S. C.; Brown, R. M., Jr. J Cell Biol 1980, 84, 315.

24. Saxena, I. M.; Lin, F. C.; Brown, R. M., Jr. Plant Mol Biol 1990, 15, 673.

25. Please visit http://cellwall.stanford.edu.

26. Saxena, I. M.; Brown, R. M., Jr.; Fevre, M.; Geremia, R.; Henrissat, B. J Bacteriol 1995, 177, 1419.

27. Saxena, I. M.; Kudlicka, K.; Okuda, K.; Brown, R. M., Jr. J Bacteriol 1994, 176, 5735.

28. Arioli, T.; Peng, L.; Betzner, A. S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; Höfte, H.; Plazinski, J.; Birch, R.; Cork, A.; Glover, J.; Redmond, J.; Williamson, R. E. Science 1998, 279, 717.

29. Laosinchai, W. Ph.D. Dissertation, University of Texas at Austin, 2002.

30. Kimura, S.; Laosinchai, W.; Itoh, T.; Cui, X.; Linder, R.; Brown, R. M., Jr. Plant Cell 1999, 11, 2075.

31. Dehal, P.; et al. Science 2002, 298, 2157.

32. Lee, K.-B.; Park, S.-J.; Mirkin, C. A.; Smith, J. C.; Mrksich, M. Science 2002, 295, 1702.

33. Unger, M. A.; Chou, H.-P.; Thorsen, T.; Scherer, A.; Quake, S. R. Science 2000, 288, 113.

34. For proof of the two-step crystallization model for cellulose biosynthesis and an energy analysis proving the two-step crystallization model, see refs. 35 and 36, respectively.

35. Cousins, S. K.; Brown, R. M., Jr. Polymer 1997, 38, 897.

36. Cousins, S. K.; Brown, R. M., Jr. Polymer 1995, 36, 3885.