Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda

A Research Roadmap Resulting from the Biomass to Biofuels Workshop Sponsored by the U.S. Department of Energy

December 7–9, 2005, Rockville, Maryland

DOE/SC-0095, Publication Date: June 2006
Office of Science, Office of Biological and Environmental Research, Genomics:GTL Program
Office of Energy Efficiency and Renewable Energy, Office of the Biomass Program

DOE Genomics:GTL
GTL Biofuels
Home Page This Document

Chapter PDFs

John Houghton
Office of Science
Office of Biological and Environmental Research
301.903.8288
John.Houghton@science.doe.gov

Sharlene Weatherwax
Office of Science
Office of Biological and Environmental Research
301.903.6165
Sharlene.Weatherwax@science.doe.gov

John Ferrell
Office of Energy Efficiency and Renewable Energy
Office of the Biomass Program
202.586.6745
John.Ferrell@hq.doe.gov

Lignocellulosic Biomass Characteristics

Makeup, Structure, and Processability

Lignocellulosic biomass has long been recognized as a potential low-cost source of mixed sugars for fermentation to fuel ethanol. Plant biomass has evolved effective mechanisms for resisting assault on its structural sugars from the microbial and animal kingdoms. This property underlies a natural recalcitrance, creating technical barriers to the cost-effective transformation of lignocellulosic biomass to fermentable sugars. Moderate yields and the resulting complex composition of sugars and inhibitory compounds lead to high processing costs. Several technologies have been developed over the past 80 years, often in wartime, that allow this conversion process to occur, yet the clear objective now is to make the process cost-competitive in today's markets.

Cell walls in lignocellulosic biomass can be converted to mixed-sugar solutions plus lignin-rich solid residues by sequential use of a range of thermochemical pretreatments and enzymatic saccharification. The low rate at which biomass is converted to sugars and the coproduction of fermentation inhibitors increase equipment size and result in high pretreatment and enzyme costs. New approaches for designing improved energy feedstocks, deconstructing plant cell walls, and transforming their polysaccharides to fermentable sugars are needed. A systematic understanding of enzyme interactions with plant cell architecture and hierarchy, as well as cellulose, hemicellulose, and lignin structure during chemical and enzymatic hydrolysis, will allow the prediction of plant-tissue response to hydrolytic attack and the creation of new systems.

Significant technology development will be needed for creation of large-scale bioenergy and biorefinery industries that can handle a billion tons made up of a variety of biomass each year. In the DOE-USDA Billion-Ton Study, corn stover and perennial crops such as switchgrass and hybrid poplar make up about half the potential 1.3 billion tons of biomass that could be available by the mid-21st Century (Perlack et al. 2005). Understanding the structure and function of these and other biomass resources will be critical to enhancing their processability.

The result of analysis and research described here will be to increase the efficiency with which the solid (substrate) interacts with large protein macromolecules (enzymes) at its surface while the surface itself is being eroded into soluble oligosaccharides [see sidebar, Image Analysis of Bioenergy Plant Cell Surfaces at the OBP Biomass Surface Characterization Lab (BSCL), p. 40]. This knowledge, combined with development of new proteins that catalyze these transformations as well as microbial systems

Image Analysis of Bioenergy Plant Cell Surfaces at the OBP Biomass Surface Characterization Lab (BSCL)

Many aspects of current biomass conversion technology are becoming better understood, and a nascent biomass processing industry is emerging for some niche markets. To reach the mid- and long-term goals stated in the DOE Office of the Biomass Program's *Multi Year Program Plan: 2007-2012*, however, enhanced fundamental understanding of feedstocks and all biorefinery processes is critical. For example, detailed knowledge about plant cell-wall ultrastructure and function to formulate improved enzyme mixtures and pretreatments will reduce the cost of producing sugars. (See images, right.)

In many cases, we know how to describe biomass compositionally. That is, we can conduct chemical or spectroscopic analyses and determine the percentages of individual sugars, protein, uronic acids, and lignin. When we study biomass conversion of corn stover, hardwoods, or rice straw, for example, we are in fact working primarily with the plant's structural parts, most of which are cell wall. Therefore, more knowledge is needed about the natural organization and structure of polymers and chemicals in plant tissue that affect chemical pretreatment, enzymatic digestibility, and the generation of compounds inhibiting fermentative microorganisms used to produce the final fuel or chemical. The study of plant cell walls at the submicron or macromolecular scale is challenging. Imaging and image analysis are at the cutting edge of botany, molecular biology, biochemistry, chemistry, and material and computer sciences. Descriptions of microscopies important for ultrastructure imaging are in the Imaging Technologies section of the Crosscutting Technologies chapter, p. 163, and sidebar, Some Imaging Technologies Relevant to Feedstock Characterization, p. 163.

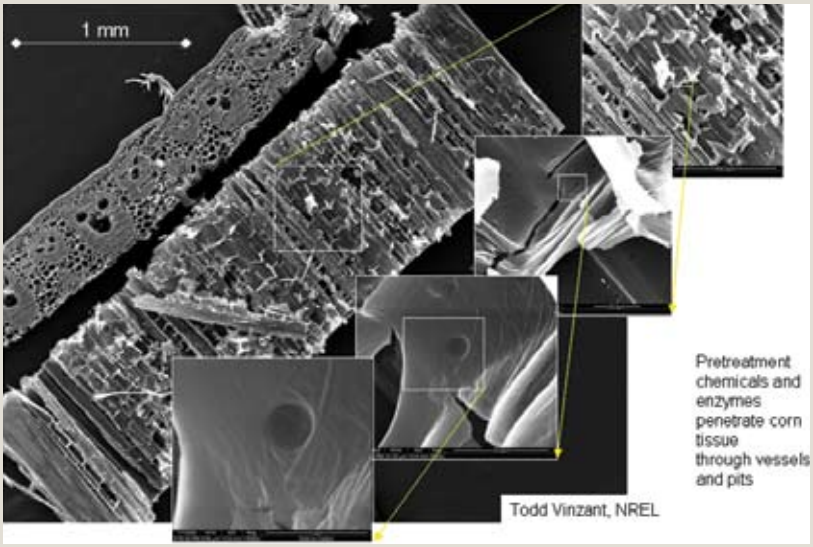


Fig. A. Collage of Scanning Electron Microscopy Images Showing a Rind and Adjacent Pith Section Cut from a Field-Dried Corn-Stem Cross Section. The rind shows a higher density of vascular elements made from thick-walled cells. The pith section (shown longitudinally) shows a greater number of thin-walled parenchyma cells. Overall, most cellulose needed for biomass conversion is located in the rind, although the pith represents most of the stem volume. Closeups of a cell-wall pit also are shown (~150,000×). These structures are thought to aid transfer of chemicals and enzymes used in processing within the biomass bulk.

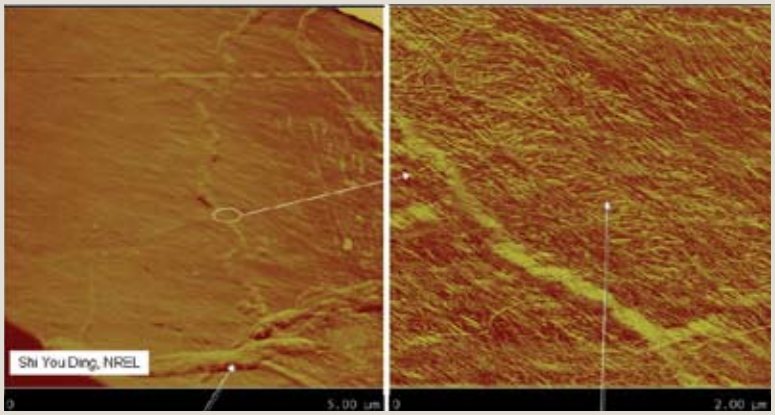


Fig. B. AFM: Corn Parenchyma Cell Wall.

for fermentation and consolidation, will enable the design of procedures and hardware that dramatically speed up the process, improve yield, and lower costs.

Structure and Assembly of Cell Walls

Plant cell walls are a complex and dynamic mixture of components that perform many functions (see Fig. 1. Simplified Cell Wall and Fig. 2. Conceptual Illustration of Cell-Wall Biogenesis, p. 42; and sidebar, Understanding Biomass, pp. 53 to 55). The cell walls are intricate assemblages of celluloses, hemicelluloses (i.e., xyloglucans, arabinoxylans, and glucomannans), pectins (i.e., homogalaturonans, rhamnogalacturonan I and II, and xylogalacturonans), lignins, and proteoglycans (e.g., arabinogalactan-proteins, extensins, and proline-rich proteins). Most mass in the plant cell wall is in the form of polysaccharides (cellulose and hemicelluloses). The next most abundant polymer is lignin, which is composed predominantly of phenylpropane building blocks. Lignins perform an important role in strengthening cell walls by cross-linking polysaccharides, thus providing support to structural elements in the overall plant body. This also helps the plant resist moisture and biological attack. These properties of lignin, however, interfere with enzymatic conversion of polysaccharide components. Additionally, since lignin is not converted readily to ethanol, we must find other uses in the process if we are to maximize energy yield from biomass.

Several thousand gene products are estimated to participate in synthesis, deposition, and function of cell walls, but very few associated genes have been identified and very little is known about their corresponding enzymes. Many questions remain, for example, regarding how polysaccharides and lignin are synthesized, how wall composition is regulated, and how composition relates to cell-wall biological functions. To answer these questions, we need to discover the functions of many hundreds of enzymes, where proteins are located within cells, whether or not they are in complexes, where and when corresponding genes are expressed, and what factors and genes control expression and activities of the proteins involved. Application of new or improved biological, physical, analytical, and mathematical tools will facilitate a detailed mechanistic understanding of cell walls. That knowledge will permit optimization of various processes involved in producing biomass and converting it to fuels.

Productivity and conversion-process efficiencies can be increased by altering fundamental aspects of plant growth, development, and response to biotic and abiotic stress. Altering cell-wall composition to increase the relative amount of cellulose and decrease lignin, for example, could have significant effects (see sidebar, Optimizing Lignin Composition for More Efficient Bioethanol Production, p. 43). Eventually, a systems cell-wall model incorporating biophysical aspects with structural properties and knowledge of proteins involved in synthesis will aid in rational development of highly productive feedstock species whose cell walls are optimized for conversion.

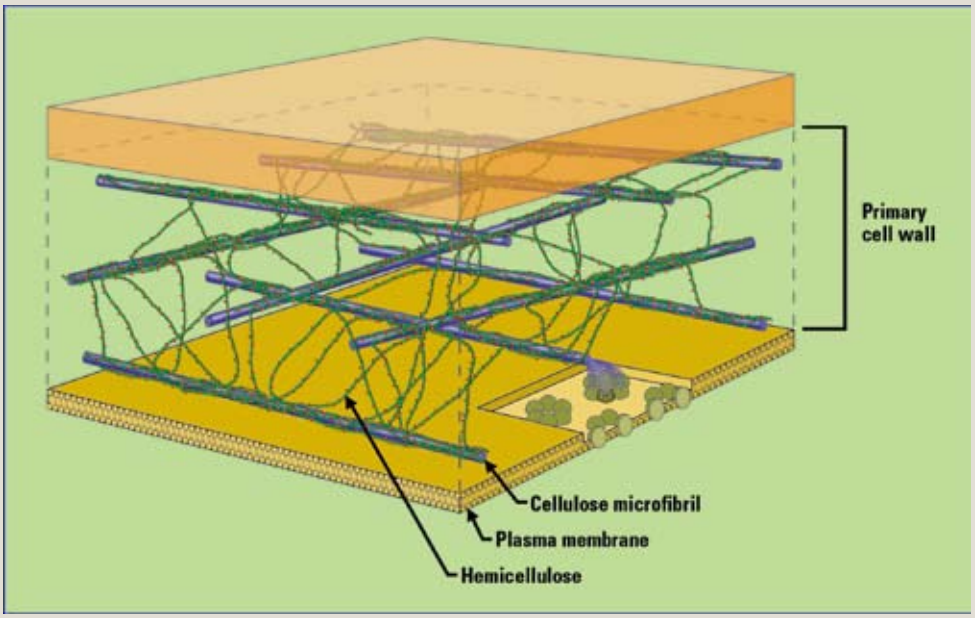
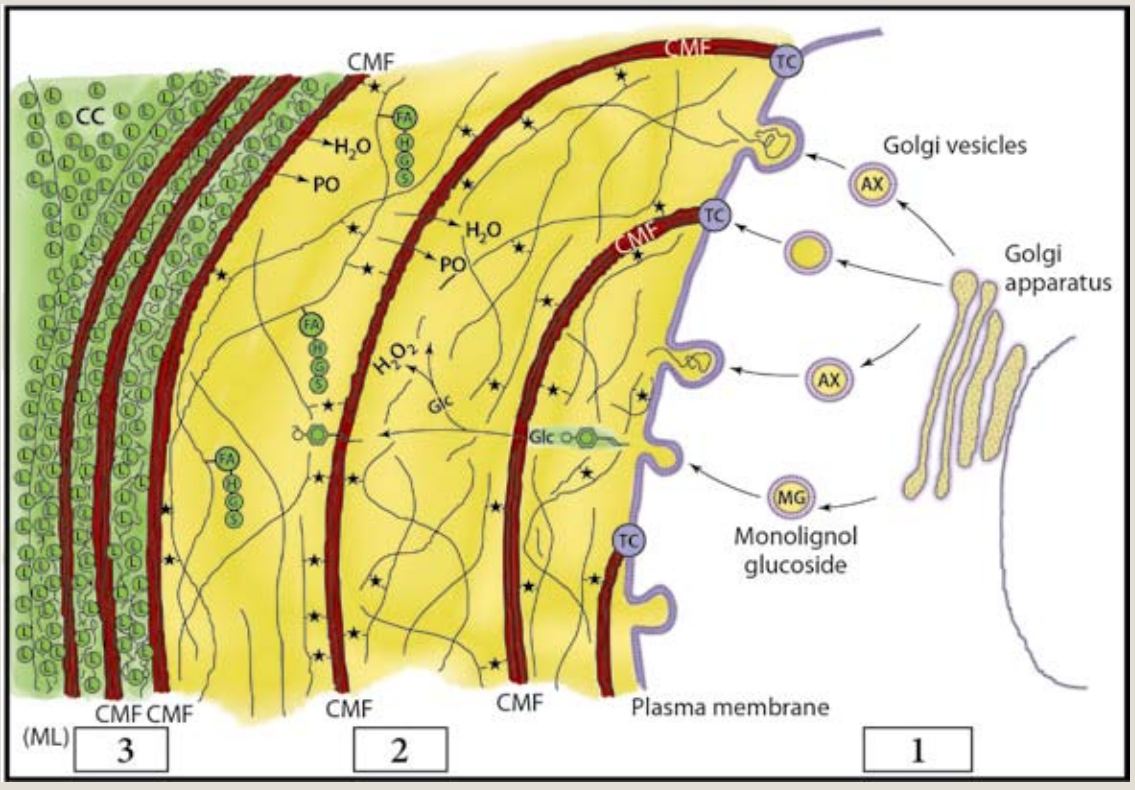


Fig. 1. Simplified Cell Wall. For more details, see sidebar, Understanding Biomass, p. 53. [Adapted with permission from C. Somerville et al., *Science* **306**, 2206–11 (2004); © 2004 AAAS.]

Fig. 2. Conceptual Illustration of Cell-Wall Biogenesis. The Golgi apparatus participates in hemicellulose and lignin biosynthesis. Cellulose microfibrils (CMF) are laid separately in the swollen gel of hemicelluloses. As lignin is deposited, the cell wall becomes hydrophobic. Water removal from the swollen gel, together with peroxidase (PO) and calcium, causes anisotropic shrinkage perpendicular to the CMFs. This shrinkage drives further oligolignol polymerization by orienting the lignin aromatic ring parallel to the cell-wall surface. [Adapted from N. Terashima et al., "Comprehensive Model of the Lignified Plant Cell Wall," pp. 247–70 in *Forage Cell Wall Structure and Digestibility*, ed. H. G. Jung et al., American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America (1993).]



Plants can have two types of cell walls, primary and secondary. Primary cell walls contain cellulose, which consists of hydrogen-bonded chains of thousands of β-1,4-linked glucose molecules, in addition to hemicelluloses and other materials woven into a nanoscale network with the cellulose. Cellulose in higher plants is organized into microfibrils, each measuring about 3 to 6 nm in diameter and containing up to 36 cellulose chains. Each

Optimizing Lignin Composition for More Efficient Bioethanol Production

Plant lignin (guaiacyl and syringyl) interferes with the release and hydrolysis of cell-wall polysaccharides. Metabolic engineering of the lignin biosynthetic pathway has been suggested as a method for modifying lignin content in feedstocks. Studies in *Arabidopsis* demonstrated that overexpression of the enzyme ferulate 5-hydroxylase (F5H) increases lignin syringyl monomer content and abolishes the tissue specificity of its deposition (Fig. A).

To determine whether or not this enzyme has a similar regulatory role in woody plants, F5H was overexpressed in poplar trees using a cinnamate 4-hydroxylase promoter to drive F5H expression. Transgenic trees displayed enhanced lignin syringyl monomer content, indicating that F5H overexpression is a viable metabolic engineering strategy for modifying lignin biosynthesis. These high-syringyl lignin poplars demonstrated a significant increase in chemical pulping efficiency. [R. Franke et al., "Modified Lignin in Tobacco and Poplar Plants Over-Expressing the *Arabidopsis* Gene Encoding Ferulate 5-Hydrosylase," *Plant J.* **22**(3), 223–34 (2000).] Similar metabolic engineering strategies hold promise for developing improved feedstocks for bioethanol production.

Many aspects of lignin biosynthesis remain matters of debate. Although most genes involved in the biosynthetic pathway have been cloned and functions assigned, mechanisms that regulate the pathway still are largely unknown, as is its relationship with other cell-wall biochemical pathways and plant development. Topics to be studied include regulation of lignin deposition and tissue specificity, identity of proteins involved in monolignol transport and polymerization, and ways in which lignin content and composition can be modified (see Fig. 4. Phenylpropanoid Pathway Leading to Lignin Biosynthesis in Plants, p. 49). Also needed is a detailed understanding of lignin-biodegradation mechanisms, including that accomplished by white rot fungi, which break down yellow lignin and leave behind crystalline white cellulose (see sidebar, White Rot Fungus, p. 93). Comprehensive explorations of lignin biosynthesis and degradation are required to maximize energy yield from biomass crops.

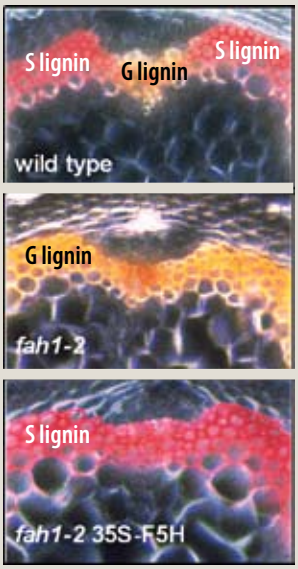


Fig. A. Lignin Composition Controlled by Genetic Manipulation and Monitored via Histochemical Staining for Lignin Monomer Composition in *Arabidopsis* Stem Cross Sections. In the lignified cells of wild-type *Arabidopsis* stems, the presence of syringyl (S) or guaiacyl (G) monomers can be visualized by histochemical staining of S lignin (red) and G lignin (yellow).

Histochemical staining allows a diagnosis of the effects of experiments to manipulate lignin composition. For example, eliminating one enzyme of the lignin biosynthetic pathway in an *Arabidopsis* mutant (*fah1-2*) leads to a pure G lignin, and overexpression of the same enzyme leads to a homogeneous deposition of S lignin (35S-F5H). The *fah1* gene encodes ferulate-5-hydroxylase, a cytochrome P450–dependent monooxygenase that catalyzes hydroxylation of coniferaldehyde and coniferyl alcohol in the pathway leading to syringyl lignin. [Figures published in C. Chapple et al., "Lignin Monomer Composition is Determined by the Expression of a Cytochrome P450–Dependent Monooxygenase in *Arabidopsis*," *Proc. Natl. Acad. Sci. USA* **95**, 6619–23 (1998); ©1998 National Academy of Sciences, U.S.A.]

cellulose chain is a linear collection of thousands of glucose residues. Pairs of glucose residues (cellobiose) make up the repeating unit of cellulose.

Like steel girders stabilizing a skyscraper's structure, the mechanical strength of the primary cell wall is due mainly to the microfibril scaffold. A microfibril's crystalline and paracrystalline (amorphous) cellulose core is surrounded by hemicellulose, a branched polymer composed of pentose (5-carbon) and, in some cases, hexose (6-carbon) sugars. In addition to cross-linking individual microfibrils, hemicelluloses in secondary cell walls also form covalent associations with lignin, a complex aromatic polymer whose structure and organization within the cell wall are not completely understood (see Fig. 3. Association of Lignin with Polysaccharides, this page). The crystallinity of cellulose and its association with hemicellulose and lignin are two key challenges that prevent the efficient breakdown of cellulose into glucose molecules that can be fermented to ethanol.

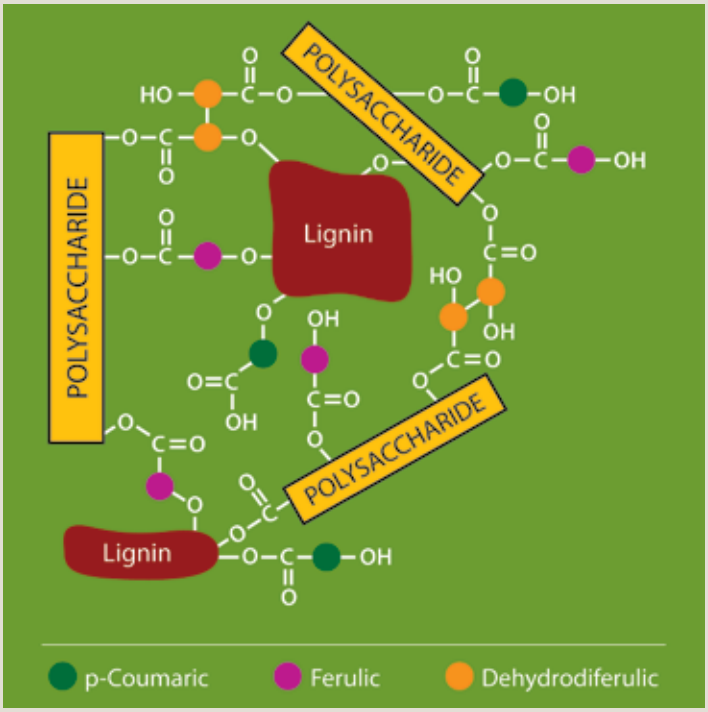


Fig. 3. Association of Lignin with Polysaccharides. The schematic diagram shows possible covalent cross-links between polysaccharides and lignin in cell walls. Lignin is bonded to the cellulose and hemicellulose polysaccharides and serves as a stiffening and hydrophobic agent, complicating biomass breakdown. [Source: Adapted from K. Iiyama, T. Lam, and B. Stone, "Covalent Cross-Links in the Cell Wall," *Plant Physiol.* **104**(2), 318 (1994). Reprinted with permission of American Society of Plant Biologists, ©ASPB 1994.]

Many enzymes involved in cell-wall synthesis or modification are thought to be located in protein complexes. Within the plasma membrane are rosettes composed of the enzyme cellulose synthase; these protein complexes move laterally along the membrane to synthesize cellulose molecular chains (36 per rosette), which crystallize into microfibrils. Movement of the rosette molecular machine is associated with cortical microtubules that underlie the membrane, but that linkage also is poorly understood. The interaction of cellulose synthase with the cytoskeleton has an impact on cellulose fibril orientation and perhaps length. Understanding the function of these complexes and their interactions with metabolic pathways that produce sugars will be important for eventually controlling cell-wall composition. A number of cellulose-synthase genes have been cloned for a variety of plants. (See sidebar, Understanding Biomass, beginning on p. 53, for an illustrated explanation.)

Factors in Recalcitrance of Lignocellulose Processing to Sugars

Organization and interactions among polymers of the cell wall—constructed for strength and resistance to biological, physical, and chemical attack—constitute a barrier to access by depolymerizing enzymes and must be partially deconstructed in the bioconversion pretreatment step before saccharification can occur. Although

various pretreatments have been developed, we still do not have a detailed understanding of fundamental physical and chemical features of lignocellulosic biomass that limit its breakdown to sugars and ultimate bioconversion efficiency. Improved methods must be developed to characterize biomass and its interaction with various chemical treatments, as well as with deconstruction and saccharification enzymes.

Natural factors believed to contribute to lignocellulosic feedstock's recalcitrance to chemicals or enzymes include plant and cell-wall architecture and molecular structure.

Plant Architecture

The organs of a plant (leaf, stem or trunk, and root) are composed of myriads of cells with different functions in the plant's economy. Each has its own particular type of cell wall whose composition is related directly to cell function [e.g., support (fibers), protection (epidermis), and transport (xylem and phloem)]. Leaf, stem, and root tissues invariably contain cells of more than one type. In tissues, individual cells are closely associated at their cell-wall interfaces to give a compact tissue structure. This structure must be disassembled by milling (comminution) to allow liquid access to cell walls.

- The waxy barrier comprising grass cuticle and tree bark impedes penetration of enzymes.
- Even milled plant stems and woody tissues limit liquid penetration by their nature.

Cell-Wall Architecture

The nanoscale composite nature of the plant cell wall restricts penetration of chemicals and enzymes to their substrates. The lignin-hemicellulose coating on the cell wall's cellulosic microfibrils affects the following:

- Conformation of cellulose and noncellulosic polysaccharides making up the microfibril limits accessibility of hydrolytic enzymes to their substrates.
- Lignin-carbohydrate complexes limit enzymatic hydrolysis of biomass polysaccharides.

Molecular Structure

Cellulose crystallinity severely restricts cellulase attacks. Cellulases must physically release individual cellulose chains from microfibril crystals for subsequent catalytic hydrolysis to sugars. Limiting factors:

- Inherent difficulty of enzymes in acting on poorly hydrated cellulose surfaces.
- Amount and composition (including heterogeneity) of lignin.
- Chemical heterogeneity and strength of covalent interactions between lignin and noncellulosic polysaccharides.
- Robustness of hydrogen bonding in cellulose microfibrils arising from extended hydrogen-bond periodicity.

In addition, all native celluloses undergo physical modifications that can inhibit saccharification as they are dehydrated in traditional methods of isolation or storage after harvest and in pretreatment processes. Characterizing the effects of conditions and storage environments would point to modifications in harvesting and storage for biomass resources. Understanding physicochemical characteristics of the cell-wall polysaccharide system would guide genomic modifications of bioenergy crops to facilitate processing and resist deleterious physical modification as much as possible. For example, structural elements of many lignocellulosic materials react to pretreatment in ways that reduce enzymatic digestibility.

- High mechanical pressure such as that from plug feeders collapses the natural vascular structure.
- Dilute-acid chemical pretreatments may permit cellulose to reanneal, leading to "hornification" of cellulose in microfibrils.
- Ambient or elevated temperatures may accelerate denaturation (e.g., the tendency of most (beta-1,4)-pentosans and hexosans to have inverse water-solubility relationships with temperature).
- Some pretreatments may permit lignin to become soluble and "plate out" on cellulose surfaces during the cool-down phase.

These "process-induced" causes of recalcitrance must be understood and overcome through process modification or biomass design.

Optimizing Hemicellulose Acetylation in Cell Walls

Hemicellulose Acetylation Degradation Products Are Toxic to Microbes

Acetyl side groups from hemicellulose biomass polymers are released during current pretreatment steps. These small acetyl molecules often are toxic and inhibit the microbial activity that converts sugars to ethanol. Hemicellulosic components such as xyloglucan and glucuronoarabinoxylan and pectic cell-wall components often are O-acetylated. For instance, O-acetyl groups may be present on the glucan backbone of xyloglucan or on galactose or arabinose residues of side chains. The degree of sugar-residue O-acetylation of pectins varies from 0 to 90% depending on the tissue, species, and method of preparation. The role of O-acetyl substituents in vivo is not known, but in vitro experiments suggest that one function may be their involvement in hindering enzymatic polysaccharide breakdown. O-acetyl substituents also affect polysaccharide solubility and pectin's gelation properties (Pauly and Scheller 2000).

Plant genes exhibit weak sequence similarity to putative bacterial acetyltransferase genes. Genetic tools in plants such as *Arabidopsis* will enable the identification of gene products catalyzing polysaccharide acetylation and the determination of acetylation's role in cell-wall structure and function. Such studies will provide insights into the possibility of developing biomass crop varieties with significantly reduced polysaccharide acetylation and thus improving the fermentation process.

After cellulose, hemicellulose is the next most abundant polysaccharide in native biomass feedstocks. Structural information on these polymeric substrates is necessary, and mechanistic models must be developed to identify "bottlenecks" in hemicellulose bioconversion (see sidebar, Optimizing Hemicellulose Acetylation in Cell Walls, this page).

A systematic approach to understanding these factors will promote more effective use of lignocellulosic biomass in bioconversion systems. Fortunately, we now have new biological, physical, analytical, and mathematical tools that can help in reliably identifying and quantifying the relative importance of various potentially limiting factors. We also have tools to identify and optimize facilitating factors, for example, through plant breeding.

The goal is to provide a rational basis for design of practical, effective, and economical pretreatments, including controlling the physical modification of native celluloses and related cell-wall polysaccharides during thermal and chemical treatments. Current thermochemical treatments ultimately will be replaced with more benign enzymatic treatments to the degree feasible. Necessary detailed analyses are discussed in the chapter, Deconstructing Feedstocks to Sugars, p. 85.

Optimization of Plant Cell Walls

Optimal efficiency of biofuel production depends on maximizing fuel yield from a unit of biomass and minimizing energy inputs. The plant cell wall, a complex assembly that plays a primarily, but not exclusively, structural role during plant growth and development, may be particularly amenable to the application of engineering principles in redesigning the cell wall to meet energy needs. To breed plants in which cell-wall composition is optimized for conversion efficiency, understanding how cell walls are made, how composition is regulated, and the roles of various polymers in supporting plant growth and development will be necessary. The long-term goal is to develop a systems-level understanding to facilitate rational improvement of plant cell-wall composition in dedicated energy crops. Such knowledge of plant cell walls is in a very primitive stage because of scientific and technical challenges that have impeded scientific progress. Future research on cell-wall synthesis and function requires interdisciplinary approaches ranging from genomics to synthetic carbohydrate chemistry and biophysics. Model organisms are important in facilitating advances in basic biology and in bringing the most sophisticated biological tools to the problem. Several new plant models closely related to species selected for energy crops are advocated. A powerful first step is to obtain comprehensive DNA sequences for these organisms.

Understanding Cell-Wall Structure and Function

Increasing the production of biofuels begins with increasing biomass productivity, either by making more cell walls or making cell walls with more carbon. In addition, changes in cell-wall composition could have major effects on the efficiency with which biomass can be converted to fuels; relative amounts of certain sugars could be increased or wall polymers could be made more amenable to enzymatic hydrolysis, thus improving the yield of sugars delivered to the biorefinery as raw feedstock.

Important questions remain about the structures of cell-wall polymers, how they are made, and their functions in plant growth and development. To optimize the amount, composition, and structure of walls for biofuel production, we must identify the genes involved in synthesis of cell-wall polymers, the design principles for cell walls, and factors that control the amounts and organization of various types of enzymes and resultant polymers. Preliminary evidence suggests that cell-wall biophysical properties important to plant growth and development may be achieved in many different ways with regard to chemical composition. Thus, cell-wall composition of energy crops

probably can be altered so they are better suited for fuel production. Desirable improvements include increasing the amount of such useful polysaccharides as cellulose or certain hemicelluloses and minimizing the content of such undesirable components as lignin or acetyl groups.

Evidence indicates that photosynthetic CO_2 fixation is regulated by plants in response to demand for fixed carbon, so understanding photosynthate flux into cell-wall polymers relative to other pathways of primary carbon metabolism and storage is important. Understanding mechanisms that regulate carbon flux and synthesis of various polysaccharides may make possible the development of plants that accumulate significantly more polysaccharide per cell. Expected significant increases in the ratio of carbon to nitrogen and mineral nutrients would have a beneficial effect on agricultural inputs (e.g., planting, fertilizing, cultivating, and harvesting), costs, and sustainability.

Progress in this area requires broad approaches to achieve a foundation of knowledge about cell-wall structure and function that will be the basis for a systems approach to predicting and controlling biomass composition. Before a systems approach can be implemented, a comprehensive understanding is needed about what reactions are performed by the many hundreds of enzymes involved in cell-wall synthesis and deposition, where and when relevant genes are expressed, and what genes control expression and activity of proteins involved in polysaccharide and lignin synthesis and modification. Indeed, one "grand challenge" in systems biology may be understanding how to engineer cell walls that meet the need of chemical biorefineries for optimized feedstocks yet still meet the plant's need for development, robustness, and maximal rates of growth.

GTL capabilities could provide extensive support for research on cell-wall synthesis, structure, and function. Sequencing support for model organisms (see below) and for identifying relevant genes in energy crops is an immediate goal. The development of populations of transgenic experimental plants with epitope-tagged proteins would greatly facilitate the determination of subcellular protein localization and the application of proteomic techniques to identify protein complexes. DNA chips, in conjunction with advanced genetic technologies, can be used for a systems-level understanding of transcriptional control of cell-wall synthesis and modification pathways. Epitope tagging also may be used to facilitate mRNA purification from single cells, facilitating insights into processes specific to cell types. Ultimately, GTL capabilities in systems analysis will permit an integrated systems model that can be used to support directed modification of cell walls for specific applications.

Efforts to understand and modify cell walls need to be coordinated with bioconversion and plant cell-wall deconstruction initiatives to optimize feedstock composition based on pretreatment and conversion methods and effects. These objectives also need coordination to develop analytical and visualization methods, computational facilities, and organic-chemistry methods for production of enzyme substrates and standards used in phenotyping and gene characterization.

Control of Lignin Synthesis and Structure

Although lignin is not converted readily to ethanol, lignin biomass may be amenable to chemical or thermal processing to achieve such liquid fuels as low-grade diesel or fuel oil. One aspect of optimizing biomass composition for ethanol production is minimizing lignin content. Alternatively, developing plants with modified lignin that can be removed easily during biomass processing may be possible.

Lignin is a complex aromatic polymer associated with polysaccharides in secondary cell walls (see Fig. 3, p. 44, and Fig. 4. Phenylpropanoid Pathway Leading to Lignin Biosynthesis in Plants, this page). Lignin constitutes a significant barrier in biomass conversion to fuels by inhibiting enzyme access to polysaccharides and by releasing toxins during degradation that inhibit organism growth during fermentation of cell-wall hydrolysates to ethanol. Genetic studies have indicated that lignin reductions may cause deleterious changes in plant growth and development. However, lignin possibly may be reduced with or without harmful effects on plant growth if compensating changes could be made in the amount of cell-wall polysaccharides. Some early experiments are under way. The degree to which

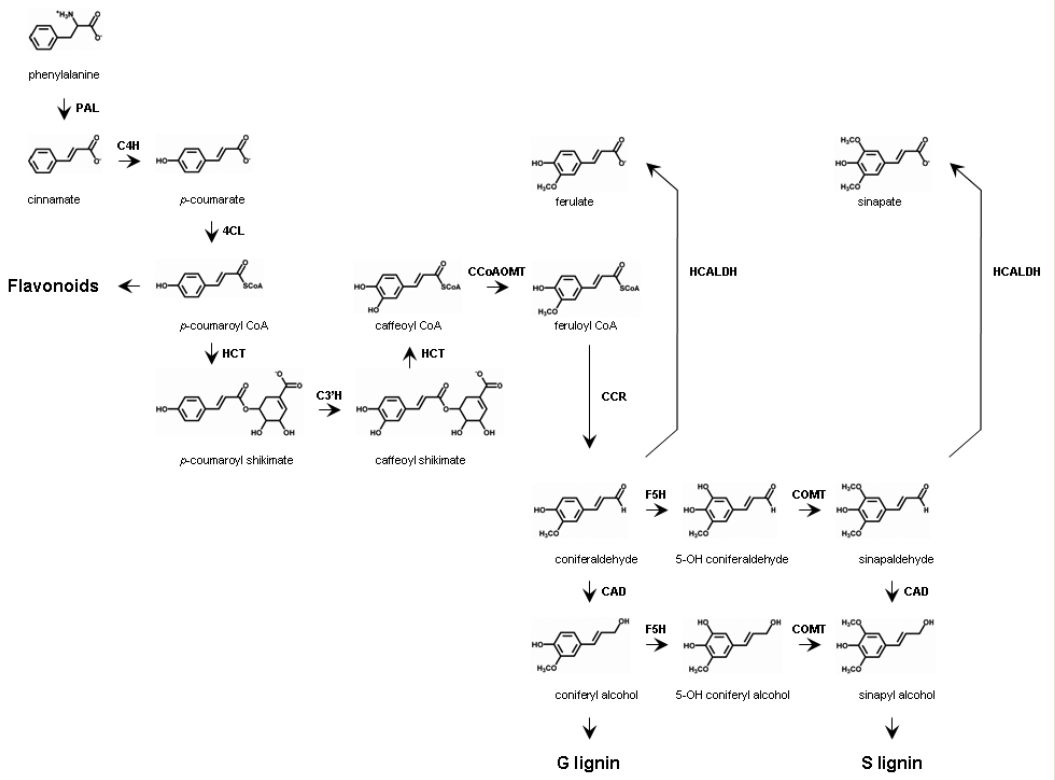


Fig. 4. Phenylpropanoid Pathway Leading to Lignin Biosynthesis in Plants. Horizontal reactions are ring modifications; vertical reactions are side-chain modifications. [Figure source: C. Fraser and C. Chapple, Purdue University]

4CL, 4-(hydroxy)cinnamoyl CoA ligase; C3'H, p-coumaroyl shikimate/quinate 3'-hydroxylase; C4H, cinnamate 4-hydroxylase; CAD, cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl CoA O-methyltransferase; CCR, cinnamoyl CoA reductase; COMT, caffeic acid/5-hydroxyferulic acid O-methyltransferase; F5H, ferulate 5-hydroxylase; HCALDH, hydroxycinnamaldehyde dehydrogenase; HCT, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase; PAL, phenylalanine ammonia-lyase.

cellulose amount can be increased with or without simultaneous changes in hemicellulose content and composition must be ascertained.

Exploring lignin biosynthesis and its regulation in a comprehensive fashion may make possible the formulation of methods for limiting and altering lignification to maximize biomass-to-energy conversion. For instance, the gene for ferulate-5-hydroxylase has been used to increase the syringyl monomer content of poplar lignin (see sidebar, Optimizing Lignin Composition for More Efficient Bioethanol Production, p. 43). The resulting trees had normal growth and development, but the pulping time was reduced by more than 60% (Huntley et al. 2003). Similarly, opportunities exist to modify lignification cell specificity so its impact on energy conversion can be minimized. A goal is to create a lignification toolbox to manipulate polymer depositions genetically and analyze the impact of those manipulations with advanced analytical organic chemistry. Such detailed knowledge could create novel opportunities for fundamentally changing how biomass is synthesized and subsequently processed for biofuels. For instance, novel monomers might be incorporated to generate lignins with unique, useful chemistries—readily cleavable linkages that could facilitate lignin depolymerization under more benign conditions (i.e., with enzymes).

Improved Methods, Tools, and Technologies

New analytical methods, tools, and technologies will accelerate the understanding of cell-wall synthesis, makeup, structure, and function and will speed breeding or rational modification of energy crop varieties.

At the basic research level, new and improved methods are needed to analyze wall composition and nanoscale structure. Ideally, these methods could be applied to analysis of a small number of cells. Molecules in cell walls range from 2 to 5 angstroms (0.2 to 0.5 nm) in diameter (i.e., a polysaccharide chain) and to many microns in length. Primary cell walls are from 50 to 100 nm in thickness and, in some cells, are thought to be chemically differentiated from one side to another. New imaging modalities that take advantage of various chemically specific imaging tags will support the long-term vision of in situ images of living plant cell walls. Images will reveal key molecular processes occurring in real time during the full life cycle of cell-wall formation, maturation, transformation, dehydration, and processing into simple feedstocks. The understanding obtained through research using such imaging is expected to result in quantitative, predictive modeling as a guide to formulating advanced feedstocks and their subsequent processing. A systematic approach is required to identify plant biomarkers and specific antibodies or other molecular tags useful in feedstock improvement.

Poorly understood now, the fine structure of intact walls must be studied to determine how the parts fit together to comprise the whole wall's physical properties. Some aspects of the general problem may be resolved simply by encouraging the application of such existing methods as very high resolution electron and scanning probe microscopy (see sidebar, Image Analysis of Plant Cell Surfaces at the OBP Biomass Surface Characterization Lab, p. 40).

Similarly, greater use of nuclear magnetic resonance (NMR) and magnetic resonance imaging may allow the development of 2D and 3D maps of cell-wall composition from important experimental and production species such as *Arabidopsis* and poplar. Use of NMR may be expanded by isotopic labeling and further development of solvents capable of dissolving cell-wall components. Complete annotation of 2D maps could facilitate greatly the analysis of genetic variation in cell-wall composition and the assignment of function to genes implicated in wall biosynthesis and modification.

Other approaches meriting investment include expanded collections of antibodies and aptamers to cell-wall components, the use of enzyme-based polysaccharide fingerprinting, pyrolysis gas chromatography–mass spectrometry (GC-MS), and related methods. A challenge is associated with characterizing enzymes that synthesize cell-wall polysaccharides: Many enzymes add sugars to preexisting polysaccharides (i.e., "acceptors" or "primers") that are not readily available as standards and reagents. Focused investments in carbohydrate chemistry will be required to construct substrates—including labeled substrates—for measuring the activity of many wall biosynthetic enzymes. Expertise in carbohydrate synthetic chemistry also would be a needed complement to proteomic and metabolomic capabilities envisioned in GTL capability suites. Expanded capabilities in synthetic carbohydrate chemistry could open up new high-throughput methods for characterizing carbohydrate-active enzymes based on high-density and high-diversity "glycochips." In this method, the activity or binding of a target protein could be evaluated simultaneously with hundreds or thousands of potential substrates and very small amounts of reagent.

High-throughput methods of cell-wall analysis are needed for plant breeding and improvement, allowing timely analysis with the most sophisticated analytical techniques. Ultimately, infield characterization is required to support breeding, molecular marker mapping, and studies involving such environmental variables as fertilizers and various biotic and abiotic stresses. Methods must be accurate and relatively inexpensive for the large numbers of samples typically handled during a breeding program. Additionally, they should be applicable to a wide variety of materials, from corn stover to wood. In principle, a high-throughput sample analysis may be enabled by detailed analysis of the relationship between cell-wall composition and features of Fourier transform infrared spectroscopy spectra or pyrolysis GC-MS chromatograms, combined with computational methods.

Technical Milestones

Within 5 years

- Develop rapid, accessible tools and methods for consistent biomass compositional analysis in bulk and fractions (see section, Characterizing Cell Walls Using High-Throughput Methods, p. 108).

- Identify genes for enzymes that catalyze synthesis of major polysaccharide backbones.
- Identify a substantial fraction of enzymes that catalyze synthesis of polysaccharide sidechains and determine sidechain biological function in model plant species.
- Identify enzymes that acetylate polysaccharides, and establish biological function for such modifications.
- Identify genetic regulatory factors that control lignin synthesis and deposition.

Within 10 years

- Clarify regulation of polysaccharide biosynthesis, including key steps that regulate carbon flow from photosynthesis into cell-wall polymers.
- Define mechanisms that control cellulose amount and fibril length and angle.
- Modify celluloses with altered numbers of glycan chains in secondary walls, and produce and test them in model species.
- Make available for testing biomass crop plants with decreased lignin and increased amounts of cellulose or other polysaccharides.
- Develop new tools and methods to help us understand cell-wall structure, including highly parallel computational simulations and high-sensitivity 2D NMR and MS instrumentation for analysis of lignin in small tissue samples.
- Identify all genes that catalyze synthesis of polysaccharide sidechains.

Within 15 years

- Determine regulatory genes that control amounts of major polysaccharides, including cellulose.
  - Develop methods for manipulating polysaccharide composition of any particular cell type within a specific tissue.
- Make available plants with improved wall composition. These plants will have increased yields of fermentable sugars, requiring less costly preprocessing; cell-wall degradation will result in insignificant levels of inhibitory compounds (in the fermentation process).
- Develop a detailed model of lignin monomer transport, polymerization initiation, and the interactions of lignin polymers with polysaccharide components of the plant cell.

Understanding Biomass: Plant Cell Walls

A First Step to Optimizing Feedstocks for Fuel Production

Optimizing plant biomass for more efficient processing requires a better understanding of plant cell-wall structure and function (see next two pages). Plant cell walls contain long chains of sugars (polysaccharides) that can be converted to transportation fuels such as ethanol. The saccharification process involves using enzymes to break down (hydrolyze) the polysaccharides into their component sugars for fermentation by microbes to ethanol (see sidebar, From Biomass to Cellulosic Ethanol, p. 26). Significant challenges for efficient conversion are presented by both the large number of enzymes required to hydrolyze diverse sugar linkages and the physical inaccessibility of these compounds to enzymes because other cell-wall components are present.

Plant cell walls contain four different polymer types—cellulose microfibrils, hemicelluloses, pectins, and lignins. Microfibrils perform an important role in strengthening cell walls, thus providing support to the overall plant body. Some properties of lignin, however, interfere with enzymatic conversion of polysaccharide components. Additionally, since lignin is not readily converted to ethanol, we must find other ways it can be used if we are to maximize the yield of energy from biomass.

Several thousand genes are estimated to participate in cell-wall synthesis, deposition, and function, but very few genes have been identified and very little is known about their corresponding enzymes. Many questions remain, for example, regarding how polysaccharides and lignin are synthesized, how wall composition is regulated, and how composition relates to the biological functions of cell walls. To answer these questions, we need to discover the functions of many hundreds of enzymes, where proteins are located within cells, whether or not they are in complexes, where and when the corresponding genes are expressed, and which genes control the expression and activities of proteins involved. Application of new or improved biological, physical, analytical, and mathematical tools will facilitate a detailed mechanistic understanding of cell walls. That knowledge will permit optimization of various processes involved in producing biomass and converting it to fuels.

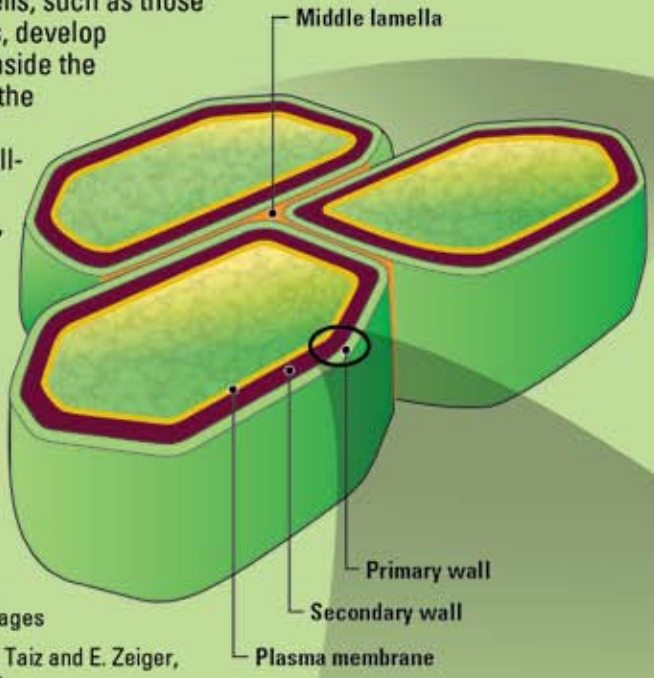
Major opportunities exist to increase productivity and conversion-process efficiencies by altering fundamental aspects of plant growth, development, and response to biotic and abiotic stress. Altering cell-wall composition to increase the relative amount of cellulose and to decrease lignin, for example, could have significant effects (see sidebar, Optimizing Lignin Composition for More Efficient Bioethanol Production, p. 43). Eventual development of a comprehensive physiological cell-wall model incorporating biophysical aspects with structural properties and knowledge of proteins involved will aid in rational development of highly productive feedstock species whose cell walls are optimized for conversion.

Understanding Biomass: Plant Cell Walls

1 Switchgrass
A potential bioenergy crop

2 Overview of plant cell walls

Plants can have two types of cell walls, primary and secondary. Primary cell walls contain cellulose consisting of hydrogen-bonded chains of thousands of glucose molecules,* in addition to hemicellulose and other materials all woven into a network. Certain types of cells, such as those in vascular tissues, develop secondary walls inside the primary wall after the cell has stopped growing. These cell-wall structures also contain lignin, which provides rigidity and resistance to compression. The area formed by two adjacent plant cells, the middle lamella, typically is enriched with pectin.



* Containing β-1,4-linkages

Figure adapted from L. Taiz and E. Zeiger, *Plant Physiology* (1991).

3 Simplified model of a primary cell wall

Cellulose in higher plants is organized into microfibrils, each measuring about 3 to 6 nm in diameter and containing up to 36 glucan chains having thousands of glucose residues. Like steel girders stabilizing a skyscraper's structure, the primary cell-wall's mechanical strength is due mainly to the microfibril scaffold. A microfibril's crystalline and paracrystalline (amorphous) cellulose core is surrounded by hemicellulose, a branched polymer composed of pentose (5-carbon) and hexose (6-carbon) sugars. In addition to cross-linking individual microfibrils, hemicellulose in secondary cell walls (not shown) forms covalent associations with lignin, a rigid aromatic polymer whose structure and organization within the cell wall are poorly understood. The crystallinity of cellulose and its association with hemicellulose and lignin are two key challenges preventing efficient cellulose breakdown into glucose molecules convertible to ethanol.

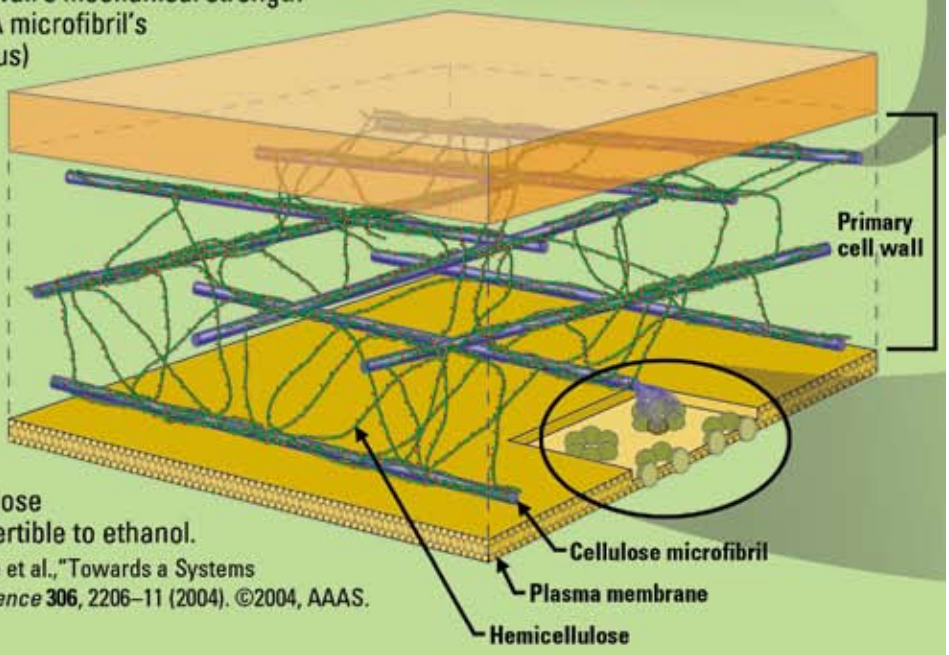
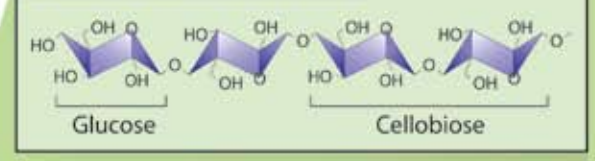


Figure adapted with permission from C. Somerville et al., "Towards a Systems Approach to Understanding Plant Cell Walls," *Science* **306**, 2206–11 (2004). ©2004, AAAS.

Questions Remain

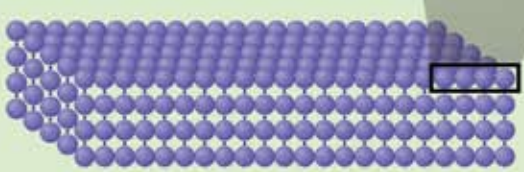
- *How is cellulose synthesis regulated?*
- *How is hemicellulose synthesized and regulated?*
- *How can we alter cell-wall structure (e.g., increase cellulose and hemicellulose, decrease lignin) for easier breakdown into component sugars?*

7 Fragment of a cellulose molecule



Alternating glucose residues are in an inverted orientation so the cellobiose (a disaccharide) is the repeating structural unit.

6 Crystalline cellulose



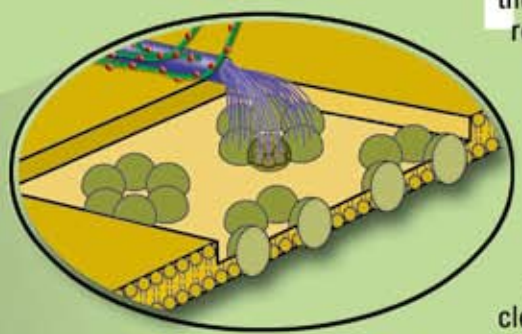
The glucan chains contain thousands of glucose residues.

5 Microfibril structure

Cellulose microfibrils are composed of linear chains of glucose molecules* that hydrogen bond to form the microfibrils.

4 Cellulose synthesis

Many enzymes involved in cell-wall synthesis or modification are thought to be located in complexes. Within the plasma membrane are rosettes composed of the enzyme cellulose synthase; these protein complexes move through the membrane during the synthesis of glucan chains (36 per rosette) that aggregate to form cellulose microfibrils. Cellulose synthase interacts with the cytoskeleton in a poorly characterized way impacting cellulose fibril orientation and perhaps length. Understanding the function of these complexes and their interactions with sugar-producing metabolic pathways will be important for eventually controlling cell-wall composition. A number of cellulose synthase genes have been cloned for a variety of plants.



Cellulose synthase complexes

Cited References

Huntley, S. K., et al. 2003. "Significant Increases in Pulping Efficiency in C4H–F5H– Transformed Poplars: Improved Chemical Savings and Reduced Environmental Toxins," *J. Agric. Food Chem.* **51**(21), 6178–83.

Pauly, M., and H. V. Scheller. 2000. "O-Acetylation of Plant Cell Wall Polysaccharides: Identification and Partial Characterization of a Rhamnogalacturonan O-Acetyl-Transferase from Potato Syspension-Cultured Cells," *Planta* **210**(4), 659-67.