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

Deconstructing Feedstocks to Sugars

This chapter describing the challenges of deconstructing cellulosic biomass to ethanol is critically linked to both the feedstock and fermentation areas. Lignocellulose recalcitrance to bioprocessing will remain the core problem and will be the limiting factor in creating an economy based on lignocellulosic ethanol production. Understanding biomass recalcitrance will help to drive crop design. Knowledge about feedstock breakdown mechanisms and products will drive fermentation and consolidation strategies, which ultimately will include consolidated bioprocessing (CBP). CBP incorporates the traits for deconstruction and fermentation of sugars to ethanol into a single microbe or culture. The previous chapter describes how tools of modern genomics-based systems biology can provide tremendous opportunities to engineer energy-plant genomes for new varieties. Those engineered plants will grow more efficiently while also producing optimal polysaccharide compositions for deconstruction to sugars and fermentation to ethanol and other products. Further advancements in plant engineering also can generate new energy crops and trees harboring polysaccharide storage structures (principally in the cell walls) that are *designed* for deconstruction. This achievement will be an important outcome of scientific research needed to optimize deconstruction of native cell walls in such crop residues as corn stover and wheat straw and such energy crops as switchgrass and poplar.

Once we understand more about the chemistry and ultrastructure of cell walls, improved thermochemical and biological means can be used to disassemble them. For example, this report will show that new and improved (existing) enzymes capable of depolymerizing cellulose in cell-wall microfibrils can significantly reduce the cost of deriving sugars from biomass—a critical factor in lowering the overall cost of cellulosic ethanol and making it cost-competitive with gasoline. Currently, the structure of cellulose itself is not adequately understood. For example, we do not know the actual faces of the cellulose crystal to which cellulases bind, and surface interactions among experimentally produced enzymes and substrates are not understood. In addition, such new findings about lignin as its synthesis, modification, and depolymerization are needed to develop conversion processes that are less energy intensive.

Sugars produced in optimal cell-wall deconstruction will be used by microbes specializing in converting (fermenting) these compounds to ethanol. In the near term, biomass conversion will produce some quantities of fermentation inhibitors that can include acetate, polyphenolics (aromatic compounds derived from lignin breakdown), and cellobiose. Even

References: p. 115

high concentrations of sugars from cell-wall deconstruction are inhibitory to many ethanologens. Thus the chapter, Sugar Fermentation to Ethanol, p. 119, describes research to develop processes at the cellular and genomic levels (e.g., evolutionary engineering to modify microbial strains that tolerate elevated levels of toxins or sugars). Also, genomic and bioinformatic tools can assist in the design of new metabolic pathways that permit efficient simultaneous fermentation of mixed sugars.

Looking ahead to longer-term improvements in biorefinery productivity, this report proposes that new organisms consolidating traits for both deconstruction and fermentation should be the subject of considerable research. For example, new ethanologen strains are needed that also are able to produce mixtures of highly competent cellulases and other hydrolytic enzymes. The application of advanced techniques in genetics, molecular biology, high-throughput screening, imaging, and mathematical modeling will accelerate the pace at which viable processing can be accomplished for energy security.

Determining Fundamental Physical and Chemical Factors in the Recalcitrance of Lignocellulosic Biomass to Processing

Lignocellulosic biomass is a complex structure with crystalline cellulose, hydrated hemicellulose, and lignin as major components. To date, the best enzyme cocktails proposed for saccharification of this material are synergistic mixtures of enzymes with defined activities, primarily those that degrade cellulose. The substrate's heterogeneous nature and enzyme cocktails' complexity suggest that traditional studies of bulk properties will not provide the detailed understanding required for knowledge-based advancements in this field.

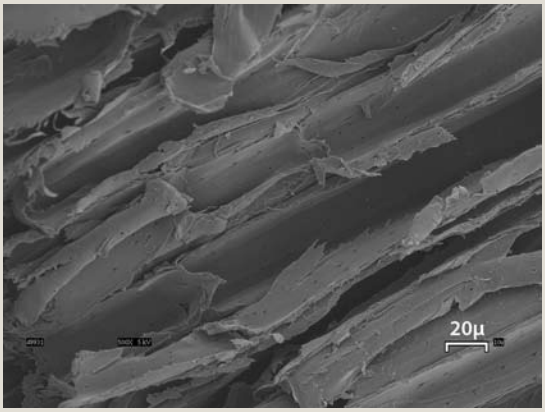
The question that must be answered is, How will structural and chemical details of enzyme substrate–binding sites affect enzyme adsorption and reaction rates? From cellulase kinetics alone, we never would be able to answer that question. We must be able to improve cellulases using an informational approach, considering that 20 years of "mixing and testing" individual cellulase proteins has yielded only modest progress toward an improved cellulase system. Enzyme kinetics always has been a measure of "ensemble average" results from experiments, and thus sugar-release values from biomass never can deliver information about the reactive site of the individual enzyme. Because of this dilemma, we do not know if cellulose recalcitrance is due to enzyme inadequacy, enzyme-substrate mismatch, or both. While in many cases we can perform bulk compositional analyses, molecular and structural correlations that are key to processability remain a challenge. The science and technology for these analyses must come from frontier capabilities in many disciplines in the physical, biological, and computational sciences [see sidebar, Image Analysis of Bioenergy Plant Cell Surfaces at the OBP Biomass Surface Characterization Lab (BSCL), p. 40].

Such enzymes as cellulases, hemicellulases, and other glycosyl hydrolases (GH) synthesized by fungi and bacteria work synergistically to degrade structural polysaccharides in biomass. These enzyme systems, however, are as complex as the plant cell-wall substrates they attack. For example, commercial cellulase preparations are mixtures of several types of GH, each with distinctly different substrate specificities (cellulose and xylan) and action patterns (exoenzymes acting from the chain ends, endoenzymes cleaving within the chain, and GH cleaving side-chain branches, for example, arabinose from arabinoxylans and xylose from xyloglucans). Optimization of these enzymes will require a more detailed understanding of their regulation and activity as a tightly controlled, highly organized system.

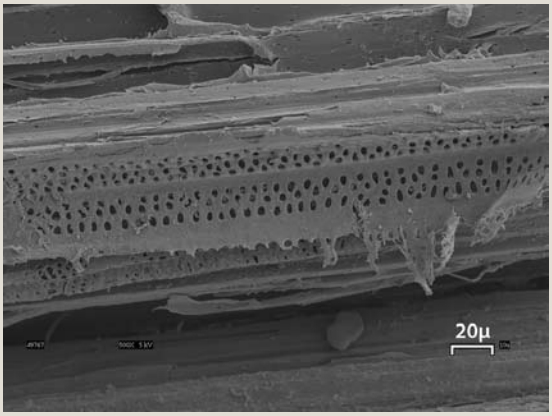
Tools must be developed first to analyze biomass at the levels and specificities needed to describe, understand, predict, and control its behaviors. These include molecular-level and nanoscale understanding of critical cell-wall physical and chemical properties, how they are formed in the plant, how they function, and how certain structural features inhibit or facilitate enzymatic interactions and subsequent saccharification. Due to the dynamic and interconnected nature of these conversion processes, we must be able to monitor them as they occur—from native structural biomass to free sugars. Critical interactions and reactions occur at many scales of time and length throughout multiple stages of intricately interlinked processing. All biomass structural and chemical properties must be analyzed in the native state, in various stages of dehydration and processing, and as the material is deconstructed into complex mixtures of reaction intermediates. These analyses must support research for biomass optimization for processing and ethanol yield, as well as investigations of robust growth and tolerance in the cultivar. Pretreatment optimization will both maximize fermentable sugar yield while minimizing inhibitors and other deleterious factors in subsequent steps.

This problem further involves structure and function at both the plant and microbial levels; their interactions; and the functions of proteins, enzymes, and living systems. They are studied under a very wide range of conditions, from pretreatment temperatures and chemistry to the enzymatic breakdown of complex cellulosic microfibrils and other polymers into sugars. To understand reaction pathways, we must be able to discern intermediate digestion and reaction products—physical structures as they degrade and chemical moieties as they are transformed and react. All these capabilities will enable the ultimate consolidation of all steps into one process.

An important and necessary departure from traditional approaches stems from our appreciation that biological systems are incredibly complex and interconnected. Cellular systems of plants and biomass-degrading microbes consist of highly coupled subsystems. In these living systems, each component (molecules and macromolecules) affects and is affected by other components. Resolution of research issues described in this section will advance biomass pretreatment and hydrolysis by developing a systems-level, predictive, and quantitative new understanding. Such understanding will facilitate rational design of biomass-conversion systems, inform the



(a) Enzyme Hydrolysis Only.



(b) Enzyme Hydrolysis Following Pretreatment.

Fig. 1. Scanning Electron Microscopy.
(a) A corn stover particle shows a smooth surface with a few micron-sized pores after enzyme hydrolysis converted 11% of cellulose to glucose in 3 h. (b) This corn stover particle has many more pores. It was pretreated in water at 190°C for 15 min and hydrolyzed by enzymes at 50°C for 3 h, resulting in 40% cellulose conversion to glucose.

The results illustrate that pretreatment changes lignocellulosic-structure susceptibility to attack by enzymes. Higher resolution in future imaging techniques will facilitate a deeper understanding of underlying molecular mechanisms. [Source: Images and conditions from unpublished work of M. Zeng, N. Mosier, C. Huang, D. Sherman, and M. Ladisch, 2006.]

design of a new generation of energy crops, and accelerate adoption of innovative biorefinery technologies.

Research Goals

Biological, mathematical, imaging, and other analytical tools must be developed and applied to identify and quantify the relative importance of potentially limiting (or facilitating) physical and chemical factors in bioconversion. All these methods will be used to characterize the effects of dehydration, heating, acidity, and cosolvents, among other biomass treatments. Detailed studies using these tools will help us better understand mechanisms of action and apply various pretreatments and enzyme systems to enhance lignocellulose bioconversion to ethanol. Such topics are discussed below.

Measurement of Biomass Properties

The organization of polymer components in biomass structures needs to be noninvasively imaged in 3D to elucidate the material's organization. Various experimental approaches including X-ray and neutron diffraction of crystals and drawn fibers, enhanced atomic force microscopy (AFM), and biochemical probe–enabled and chemically specific imaging will be used to characterize the different forms of lignocellulose in both "native" and treated (stored and processed) biomass cell walls.

Analytical methods [e.g., nuclear magnetic resonance (NMR), Fourier transform infrared (FTIR), and Raman] should provide information on chemical moieties, chemical bonds, and conformation of wall polymers. Surfaces that are enzyme-binding sites should be characterized by such tools as AFM, scanning electron microscopy, transmission electron microscopy, and electron spectroscopy for chemical analysis (see Fig. 1. Scanning Electron Microscopy, this page).

This information can be correlated using statistical-analysis packages based on principal components to reliably relate changes in lignin carbohydrate complexes (LCC) and lignin profiles, degrees of cellulose dehydration, and pretreatment chemistries to the abilities of hydrolytic enzymes to convert the resultant material biomass.

Models for Direct Enzymatic Interactions, Action

Mechanistic models of enzyme-system substrate relationships, based on new knowledge about plant cell-wall architecture and enzyme structure and function, need to be developed. The action of individual and combinations of enzymes on native and model substrates should be examined using single-molecule spectroscopy and imaging.

Mechanistic characterization of relevant enzymes, with attention to combined chemical and biological conversion processes of cell-wall materials, then will support models for interaction. Finally, new computational models for microfibrils and cell walls using improved codes and petascale leadership-class computers will provide the needed fidelity to support reaction-pathway calculations.

Technical Milestones

Within 5 years

- Using integrated analytical and mathematical methods, quantitatively analyze effects of selected biomass properties and pretreatment chemistries on enzymatic hydrolysis.
- Identify principles, genes, and controlling factors that influence biomass makeup, assembly, and processability, in conjunction with plant design.
- Perfect and extend existing spectroscopic, surface, and imaging characterization methods for biomass.
- Develop methods for monitoring the progress of physical modification during dehydration under various conditions of storage or pretreatment at elevated temperatures. Develop improved protocols to be tested in field trials.
- Define the primary structure of lignins and LCCs and the shape, dimensions, and heterogeneity of their domains.
- Develop models connecting lignocellulosic properties with deconstruction and hydrolysis processes.
- Definitively characterize the detailed organizational structures of principal types of plant cellulose and their relative energies and interrelationships.
  - Understand the energetics of different cellulose forms and ways in which these structures give rise to characteristics of the next organization level in the mesoscopic range of sizes, and
  - Determine how microfibrils interact with other principal structural components (lignins and noncellulosic polysaccharides).

Within 10 years

- Develop an understanding of the physicochemical basis of polysaccharide interactions with water. This information can be used to rationally design genomic variations for cell-wall polysaccharide composites with more desirable properties.
- Apply new biological and chemical tools to alter LCC and other relevant properties of selected biomass species.
- Subject these altered biomass species and controls to various pretreatment chemistries and temperatures.

- Hydrolyze the resulting pretreated biomass with various specific enzymes and correlate spectroscopic, surface, and image data with hydrolysis of structural carbohydrates.
- Quantitatively relate key biomass and pretreatment properties with hydrolysis yields for different biomass species.
- Improve enzymes based on the knowledge of substrate-imposed limitations and an understanding of optimized cell walls.
- Develop molecular models for enzyme-substrate structure-function relationships.

Within 15 years

- Develop mathematical models of cell walls and the transport of water, pretreatment chemicals, and enzymes through the walls. Such models should help in designing cell walls that can be digested more efficiently by developed enzyme systems. Designs optimized by modeling can then help guide genomic engineering for plant and enzyme optimization.
- Establish the foundation for integrated biomass processing by predicting and then verifying overall hydrolysis yields for native and modified biomass species using different pretreatment chemistries and severities (time, temperature, acidity) and a range of hydrolytic enzymes.
- Perform rational engineering of enzymes to specific requirements based on new understandings about enzymes and their mechanisms of action and substrate cellulose-hemicellulose-lignin complexes.

The Role of GTL and OBP Facilities and Capabilities

The program outlined most likely will need all facility resources at one point or another.

Protein Production

Protein production capabilities will be used for enzymes and tags (biomarkers).

Molecular Machine and Cellular System Analysis

These capabilities can image and diagnose interactions between enzymes and biomass substrates. In addition, BSCL, recently established at NREL, certainly will prove essential. Imaging facilities would be used to correlate digestibility to biomass ultrastructure. The sidebar, Image Analysis of Bioenergy Plant Cell Surfaces at the OBP Biomass Surface Characterization Lab, p. 40, illustrates the cell wall's highly coupled nature and how each component affects and is affected by other components, thereby highlighting the need to study the intact system.

DOE Joint Genome Institute

DOE JGI will map and identify genes concerned with deconstruction barriers or facilitating factors.

Crosscutting Tools, Technologies, and Science

- Better characterization of cell-wall polysaccharide association profiles in hydrated and dehydrated states using spectroscopic and ultramicroscopic methods.
- Larger-scale molecular mechanics modeling to explore the association of cell-wall polymers with each other and with water at various length scales.
- Robust and rapid methods for quantifying cell-wall polymers and interpolymer linkages.
- Genomic and chemical tools to identify genes involved in rate-limiting LCC linkages.
- New biomass variants with enhanced accessibility for structural carbohydrate–digesting enzymes.
- New methods for spectroscopic characterization of cell-wall polysaccharides in the living hydrated state. Most past work has been based on investigation of isolated dehydrated samples and very little on cell-wall constituents in their native state.
- Coupled structure and processing models of cell walls and components.
- Data-analysis software to identify and quantitatively describe the relative importance of different factors governing response to pretreatment and enzymatic hydrolysis of biomass.

Developing Better Enzymatic Systems for Biological Pretreatment: Ligninases and Hemicellulases

The most efficient conditions for simultaneous saccharification and fermentation of lignocellulosic biomass will be those in which transformations now performed during and following thermochemical pretreatment are seamlessly integrated into the overall process. Effective enzymatic breakdown of LCC, to expose cellulose to enzyme action, represents an important step toward this objective.

Lignocellulose, an extremely complex and widely varying nanoscale composite, is well designed to resist attack. Ligninases and hemicellulases, for which few examples are known, are inadequately understood. Technically, understanding and optimizing these enzymes will enable the ultimate goal of consolidating pretreatment and saccharification. Consequently, research is needed to identify, characterize, improve, and economically produce the most effective enzyme systems for biomass preconditioning. These enzymes would be applied before or after traditional pretreatment to minimize and, eventually, replace thermochemical processes, thus lessening the effects of overall pretreatment severity at the macromolecular level and simplifying processing. To support improvements, research must be focused on identifying more enzymes of this class and characterizing their principles of action. A crucial component for ligninase identification will be genomic, functional genomic, and proteomic comparisons of white rot fungi known or presumed to express such activities (see sidebar, White Rot Fungus, p. 93).

The ultimate goal of this research is to produce a recombinant ligninase-hemicellulase microbial system with enhanced catalytic activity and stability, industrialized for biorefinery operations.

Research Goals

The chemical and structural nature of native lignin is poorly understood. Developing robust ligninases requires a foundation of knowledge about the range of lignin chemistries and lignin associations that inhibit or facilitate freeing cellulose and hemicellulose for saccharification (see sidebar, Lignification: Random vs Template Directed, p. 94). Similarly, despite the importance of hemicelluloses and hemicellulases for saccharification, both are poorly understood. Due to the complex structure and compositional diversity of lignin-hemicellulose and -cellulose substrates that must be contemplated for feedstock biomass, a variety of enzymatic activities will be needed to catalyze deconstruction for conversion to monosaccharides (see Fig. 2. Complex Mixture of Enzymes for Degrading Hemicelluloses, p. 96).

Hemicelluloses and Hemicellulases

After cellulose, hemicelluloses (xylan polymers) are the next most abundant polysaccharides in native biomass feedstocks (see Table 1. Cellulose, Hemicellulose, and Lignin Content in Various Sources of Biomass, p. 93). Although considerable research has focused on converting biomass cellulose to fermentable glucose, less has been done on bioconverting other plant cell-wall components. A better understanding of the complex structure and composition of this polysaccharide group, which will differ in type and abundance among different feedstocks, will help to identify and optimize the mechanistic basis of required enzymatic activities. Structural information and mechanistic models must be developed to pinpoint "bottlenecks" in hemicellulose bioconversion.

Lignin and Ligninases

Despite its critical importance, the enzymatic basis of lignin depolymerization in vivo has remained elusive, if not controversial. Research during the past 20 years on putative "ligninases," which can both polymerize and depolymerize lignin preparations, has not yielded reliable insights into the mechanisms of lignin cleavage. Thus the possibility of another entirely different class of lignin-degrading enzymes cannot be disregarded, and potential candidates have been identified in white rot fungi. Genes encoding a range of ligninases must be identified, and genomic sequences of white rot fungi that degrade lignin but do not express known peroxidases will be compared with the sequenced *P. chrysosporium* genome. A suitable host organism will be needed to produce heterologous ligninases in sufficient quantities; possible hosts include *E. coli*, *S. cerevisiae*, *P. pastoris*, and *A. oryzae*. Heterologous ligninase expression and manipulation have had limited success.

This research would include the following goals:

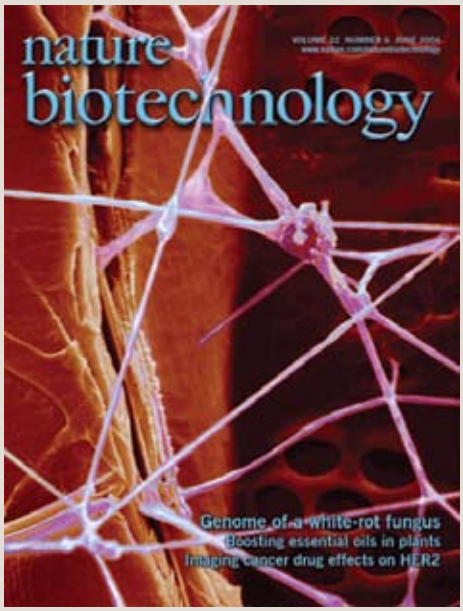
- Detailed characterization of the range of lignin and hemicellulose structures encountered in available biomass. Special attention must be paid

White Rot Fungus: Genome of Known Lignin Degrader Sequenced

Lignin degradation is the key to making polysaccharide components of cell walls available for breakdown. White rot fungi are the primary degraders of lignin, which is among the most abundant of natural materials on earth and plays a pivotal role in global carbon cycling. These organisms also degrade the cellulose and hemicellulose components of plant cell walls.

To aid in understanding these processes, the DOE Joint Genome Institute sequenced the genome of the white rot fungus *Phanerochaete chrysosporium*. This fungus degrades brown lignin, the protective matrix surrounding cellulose microfibrils of plant cell walls, leaving behind crystalline white cellulose. [D. Martinez et al., "Genome Sequence of the Lignocellulose Degrading Fungus *Phanerochaete chrysosporium* strain RP78," *Nat. Biotechnol.* **22**, 695–700 (2004).]

Analysis of the white rot fungus genome revealed genes encoding oxidases, peroxidases, and other enzymes that contribute to depolymerization of lignin, cellulose, and hemicellulose. Extensive genetic diversity was observed in gene families encoding these enzymes, possibly reflecting that multiple specificities are needed for effective degradation of cell-wall polymers from different plant species. Elucidating the regulation of genes, proteins, and metabolites from this organism and others will enhance understanding of the individual and collective mechanisms of degradative enzymes as well as their interactions with other organisms in their ecosystems. Such advances are necessary for generating the framework to engineer large-scale processes for biomass utilization.



Cover used with permission from *Nature Biotechnology*, www.nature.com/nbt/.

to the nature of covalent bonds linking LCC. Model systems will facilitate these analyses.

- Systems biology–style survey and molecular and functional characterization of the diversity and activity of relevant hemicellulases (glycoside hydrolases and esterases), including discovery of activities from such novel sources as soils, the rhizosphere, termite hindgut, and decaying biomass.
- Similar discovery and molecular and functional characterization of new lignin-degrading enzymes and their activities.
- Reliable expression systems and hosts for these enzyme classes.
- Mechanistic characterization of enzymes with attention to properties relevant to combined chemical and biological conversion processes.
- Examination of individual and combination enzyme activities on native and model substrates.
- Development of mechanistic models of enzyme system substrate relationships, based on new knowledge about plant cell-wall architecture.

Table 1. Cellulose, Hemicellulose, and Lignin Content in Various Sources of Biomass

Feedstock	Cellulose	Hemicellulose	Lignin
Corn stover	36.4	22.6	16.6
Wheat straw	38.2	24.7	23.4
Rice straw	34.2	24.5	23.4
Switchgrass	31.0	24.4	17.6
Poplar	49.9	20.4	18.1

Source: A. Wiselogel, S. Tyson, and D. Johnson, "Biomass Feedstock Resources and Composition," pp. 105–18 in *Handbook on Bioethanol: Production and Utilization* (Applied Energy Technology Series), ed. C. E. Wyman, Taylor and Francis (1996).

- Improvement of enzyme catalytic efficiency through protein engineering and of biomass through plant design.
- Coordination with feedstocks and biological-conversion research to inform plant design and to support improved fermentation and eventual incorporation of pretreatment enzymatic processes into consolidated processing schemes.

Lignin and hemicellulose investigations will follow similar research paths as described below.

Technical Milestones

Within 5 years

- Initiate metagenomic surveys in environments known to break down biomass naturally, and identify and molecularly characterize ligninases and hemicellulases, including accessory enzymes.
- Heterologously express enzymes in suitable protein production hosts.

Lignification: Random vs Template Directed

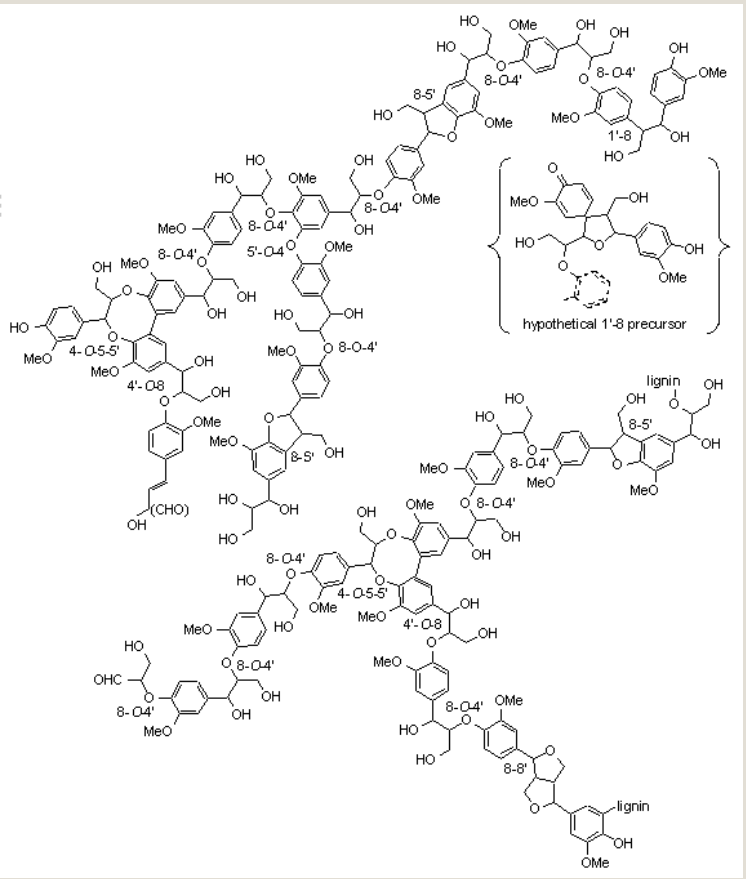


Fig. A. Contemporary View of Lignin Substructures. Theory proposed by G. Brunow and coworkers in 1998 (reproduced with permission).

In terms of energy content, lignins are thought to be the most abundant of all biopolymers. They are composed of *p*-hydroxyphenylpropanoid units interconnected through 8-O-4, 8-5, 8-8, 8-1, 5-5, and 4-O-5 linkages. Corresponding substructures in the polymer include alkyl aryl ethers, phenylcoumarans, resinols, tetrahydrofuran-*spiro*-cyclohexadienones, biphenyls, dibenzodioxocins, and diaryl ethers (see Fig. A, at left). The primary precursors themselves—the three monolignols *p*-coumaryl, coniferyl, and sinapyl alcohols—differ only according to their aromatic methoxy substitution patterns. These monolignols are oxidized enzymatically through single-electron transfer to generate the respective phenoxy radicals. The actual coupling of a monolignol radical with the growing end of a lignin chain, however, may not fall under direct enzymatic control.

Accordingly, many investigators have assumed that lignin primary structures must be "random" or combinatorial as far as sequences of interunit linkages are concerned. More recently, this theory has been reinforced by reports that certain kinds

- Determine catalytic properties of enzymes.
- Use overproduced and commercial enzymes to digest native hemicelluloses, and identify intermediate products in hemicellulose decomposition.
- Perform molecular evolution of enzymes, and select for improved kinetic efficiency and compatibility with other enzymes and with desired industrial applications. This is expected to result in the identification of factors limiting enzymatic conversion and the development of enzyme cocktails to overcome them efficiently.

Within 10 years

- Extend metagenomic surveys to discover novel ligninases and hemicellulases.
- Concurrently improve enzymes in the cell walls of energy crops in parallel with knowledge about substrate-imposed limitations.
- Develop molecular models for enzyme-substrate structure-function relationships.

of non-native monolignols can be incorporated into macromolecular lignin structures. Lignins and lignin derivatives exhibit two fundamental characteristics that traditionally have been viewed as evidence in favor of randomness in their configurations: They are both noncrystalline and optically inactive.[1]

Nevertheless, a number of observations are thought by some to point in the opposite direction. The individual molecular components in (nonpolyionic) lignin preparations tend to associate very strongly with one another in a well-defined way. These processes are thought to be governed by vital structural *motifs* derived from corresponding features disposed nonrandomly along the native biopolymer chain. Moreover, dimeric pinoresinol moieties are linked predominantly to the macromolecular lignin chain through at least one of their aromatic C-5 positions.

We do not know whether such features can be explained through combinatorial mechanisms under simple chemical control or if higher-level control mechanisms are required. One hypothesis proposes a way to replicate specific sequences of interunit linkages through a direct template polymerization mechanism. According to this model, an antiparallel double-stranded lignin template, maintained in a dynamic state at the leading edge of each lignifying domain, determines the configuration of the daughter chain being assembled on the proximal strand's exposed face. Furthermore, replication fidelity could be controlled by strong nonbonded orbital interactions between matching pairs of aromatic rings in the parent and the growing daughter chains. The overall process seems to be consistent with the lack of both crystallinity and optical activity in macromolecular lignin domains.

Finally, required sequence information may be encoded in polypeptide chains that embody arrays of adjacent lignol-binding sites analogous to those found in dirigent positioning proteins.[2]

Cited References

1. J. Ralph et al. 2004. "Lignins: Natural Polymers from Oxidative Coupling of 4-Hydroxyphenylpropanoids," *Phytochemistry Rev.* **3**, 29–60.

2. S. Sarkanen. 1998. "Template Polymerization in Lignin Biosynthesis," pp. 194–208 in *Lignin and Lignan Biosynthesis* **697**, ed. N. G. Lewis and S. Sarkanen, American Chemical Society, Washington, D.C.

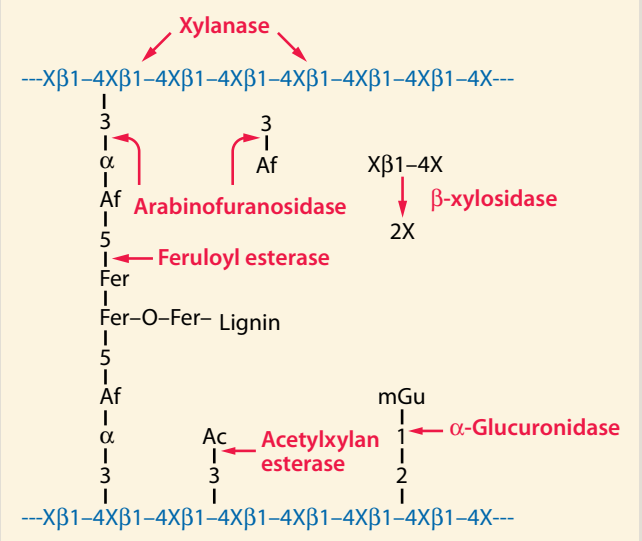
- Improve catalytic efficiency and stability of enzymes that degrade cell walls using protein engineering.

Constraints imposed by the fairly limited set of industrial host organisms on expression of native or heterologous GH proteins must be understood (for example, the role and modification of various glycoforms). Greater understanding in these areas would allow more efficient improvements in enzyme performance in industrialized hosts and resultant production economics. Understanding biomass recalcitrance will enable enzyme preparations to be engineered for superior performance using proteomics and the systems biology approach. These challenges will be more pronounced when dealing with "foreign" proteins (such as ligninases) and multienzyme complexes (cellulosomes) in heterologous hosts.

Within 15 years

- Perform rational engineering of enzymes to specific requirements based on new understanding of enzymes and substrate cellulose-hemicellulose-lignin complexes. The ability to design specific enzymes to attack specific substrates is expected by this time.

Fig. 2. Complex Mixture of Enzymes for Degrading Hemicelluloses. The example depicted is cross-linked glucurono arabinoxylan. The complex composition and structure of hemicellulose require multiple enzymes to break down the polymer into sugar monomers—primarily xylose, but other pentose and hexose sugars also are present in hemicelluloses. A variety of debranching enzymes (red) act on diverse side chains hanging off the xylan backbone (blue). These debranching enzymes include arabinofuranosidase, feruloyl esterase, acetylxylan esterase, and alpha-glucuronidase (see table below). As the side chains are released, the xylan backbone is exposed and made more accessible to cleavage by xylanase. Beta-xylosidase cleaves xylobiose into two xylose monomers; this enzyme also can release xylose from the end of the xylan backbone or a xylo-oligosaccharide.



[Source: Molecular structure adapted from L. B. Selinger, C. W. Forsberg, and K. J. Cheng, "The Rumen: A Unique Source of Enzymes for Enhancing Livestock Production," *Anaerobe* 2(5), 263–84 (1996).]

The table shows that some of these enzymes are multifunctional, with catalytic domains belonging to different enzyme families. Their great diversity and that of other enzymes involved in hemicellulose degradation present a remarkably complicated enzymatic system whose more-thorough analysis may yield greater understanding of hemicellulosic degradation.

Glycoside Hydrolase (GH) and Carbohydrate Esterase (CE) Enzyme Families for Degrading Hemicelluloses

Enzyme	Enzyme Families
Endoxylanase	GH5, 8, 10, 11, 43
Beta-xylosidase	GH3, 39, 43, 52, 54
Alpha-L-arabinofuranosidase	GH3, 43, 51, 54, 62
Alpha-glucuronidase	GH4, 67
Alpha-galactosidase	GH4, 36
Acetylxylan esterase	CE1, 2, 3, 4, 5, 6, 7
Feruloyl esterase	CE1

- Greatly increase understanding of fundamental protein-secretion limitations in hyperproducing strains. Similar challenges must be met to enable consolidated bioprocessing microbes.
- Similarly, the catalytic activity and thermal stability of ligninases must be enhanced through directed evolution strategies, which require identification of suitable model substrates for screening mutant enzymes in variant libraries.

Crosscutting Tools, Technologies, and Science

Computational science, advanced imaging, and high-throughput genomics tools must be incorporated to attain the objectives described here. Computational science is important in coordinating process engineering and in exploiting the products of protein engineering. Structural imaging of enzymes, substrates, and enzyme-substrate complexes will be needed to understand enzyme function at the molecular level; a library of model substrates for these kinetic and mechanistic studies is required.

A better understanding of the structure and composition of specific xylans and lignins would greatly improve the ability to design optimal enzyme mixes. Methods to monitor the process of polysaccharide hydrolysis (e.g., NMR and mass spectroscopy) would be beneficial. Transgenic plants and trees with reduced lignin contents or modified lignin configurations could exhibit differences in lignin-depolymerization rate; understanding enzymatic mechanisms would be crucial in manipulating conditions to the best effect.

Advanced imaging, spectroscopic, and enzymatic methods of characterizing cell walls before and after pretreatment also require new resources for producing and disseminating chemical standards. These new standards will be representative of the range of plant cell-wall chemistry and will permit improved cross-comparison research, not only on polysaccharide disassembly but also on subsequent sugar metabolism.

Understanding the Molecular Machinery Underpinning Cellulose Saccharification: Cellulases and Cellulosomes

With new biology tools, we have the opportunity to understand enzyme structure-function relationships that govern the critical processes of bioethanol production. Given the importance of alternative renewable energy sources, the prospect of engineering improved cellulases is an exciting and universally appealing concept. Cellulosomes, large molecular machines that integrate numerous enzymes and functions to break down biomass, are key to reducing the enzyme loading required for processing. Understanding cellulosomes and learning how to manipulate and modify them for greater efficiency will be important in consolidated processing. Research first will allow understanding and then improvement in the performance of both free and complexed (cellulosomal) cellulases on biofeedstocks. The rate-limiting step in hydrolysis is not catalytic cleavage but disruption of a single substrate

chain from its native matrix, thereby rendering it accessible to the catalytically active cellulase site. Thus, we must be able to analyze and understand the processes and interactions that facilitate this disruption of insoluble cellulose. To approach this problem in a systematic and rational fashion, detailed understanding of the structure and energetics of both the microcrystalline and noncrystalline portions of cellulose fibrils is first necessary.

- How do soluble enzymes act on an insoluble crystalline substrate? Hydrolysis of crystalline cellulose is the rate-limiting step in biomass conversion to ethanol because aqueous enzyme solutions have difficulty acting on this insoluble, highly ordered structure. Cellulose molecules in their crystalline form are packed so tightly that enzymes and even small molecules such as water are unable to permeate the structure.
- How do different biomass-degrading enzymes work together as a synergistic system? Cellulases and hemicellulases are secreted from cells as free enzymes or as extracellular cellulosomes. The collective activity of enzyme systems is believed to be much more efficient than the individual activity of any isolated enzyme; therefore, to truly understand how enzymes function, they must be studied as systems rather than individually or a few at a time. In addition, systems eventually must be analyzed under laboratory conditions more representative of real-world environments. For example, laboratories often use purified cellulose as the substrate for enzyme analysis rather than more heterogeneous, natural lignocellulosic materials; this can provide erroneous conclusions about natural enzyme activity. New analytical methods that are spatially and chemically sensitive will allow realistic understanding of the mechanisms of biomass degradation.

Research Goals

Discovering and Improving Free Cellulases

New generations of engineered cellulases will provide enhanced performance (activity) requiring lower protein loadings, so process cost will be reduced. Cellulase cost currently is estimated at 10 to 25 cents per gallon of ethanol produced. The new target will be similar to starch hydrolysis, or 1 to 2 cents per gallon of ethanol. For example, loadings of about 25-mg fungal cellulase proteins are required to hydrolyze about one gram of cellulose in pretreated biomass. Improved cellulases would convert an equivalent amount of cellulose to sugars using around one-tenth the enzyme loading. More efficient cellulase biomass digestion also may permit reduced-severity thermochemical pretreatment.

Acquiring new scientific insights into cellulose structure and the function of cellulase cocktails is a critical objective (see sidebar, New Cellulase Enzymes Dramatically Reduce Costs of Plant Biomass Breakdown, p. 100). The initial research phase involves an exhaustive search for examples of important families of free cellulases (those secreted into the extracellular mixture) using high-througput genomic and enzyme-production and -characterization methods. Cellulose and cell-wall interaction with

degrading enzymes must be characterized. Biochemical analysis of family members will reveal much about the natural diversity of solutions to cellulase action. Combined mutational analysis and computational modeling then will be used to define structure-function relationships of these enzymes and newly engineered mutants. With this kinetic and thermodynamic understanding as a guide, enzyme-engineering approaches will be used to test novel hypotheses for improving cellulase performance on the cell walls of plants considered suitable biofeedstocks.

Understanding and Utilizing Cellulosomes

The cellulosome, a unique type of molecular machine, can efficiently solubilize native and pretreated lignocellulose substrates. The cellulosome serves as a more efficient way of enzyme loading and presentation to the substrate. In this case, native cellulosomes attached to the microbe contain the full repertoire of plant cell-wall polysaccharide-degrading enzymes, and a single carbohydrate-binding module (CBM) targets the entire complement of enzymes to the substrate surface (see sidebar, The Cellulosome, p. 102).

- Artificial designer cellulosomes that exhibit a precise architecture and composition will help reveal various principles of cellulosome construction and action. In this context,
  - Structure and biochemical-biophysical aspects of assorted cellulosomal modules (i.e., cohesins, dockerins, CBMs, and catalytic domains) from different subunits and species will be examined.
  - Recombinant primary and adaptor chimeric scaffoldins will be constructed, designed to bear divergent cohesins for subsequent incorporation of native or hybrid (dockerin-tagged) cellulosomal enzymes.
  - Through this approach, improved high-performance (noncellulosomal) cellulases can be incorporated into a cellulosome to take advantage of enhanced synergistic properties inherent in the cellulosome complex.

Genes encoding for the complement of designer cellulosomal components also can be cloned into a suitable host cell system for heterologous production, assembly, and secretion of active designer cellulosomes of desired composition and architecture.

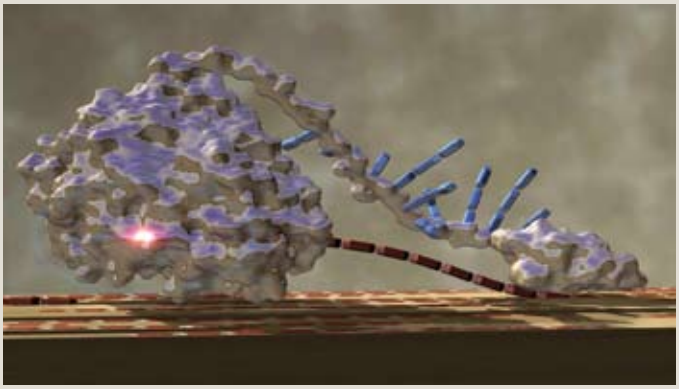
Several key scientific questions and issues are especially important for reaching the proposed goals.

- We do not understand enough about substrate microcrystalline cellulose's structure as it exists in the plant cell wall. Uncertainties about the crystal lattice structure in the microfibril and the crystal faces of cellulose targeted by enzyme diversity in the biosphere are examples of our lack of insight. Understanding the substrates' molecular architecture would help in using rational design to produce improved hydrolytic molecular machines. With such knowledge it might be possible, for example, to select specific mutations of cellulose-binding domains that tend to disrupt fibril packing or to use molecular dynamic simu-

New Cellulase Enzymes Dramatically Reduce Costs of Plant Biomass Breakdown: R&D 100 Award

Further Advances Needed to Improve Efficiency and Economics

Cellulase enzymes are used to break down the cellulose of plant cell walls into simple sugars that can be transformed (fermented) by microbes to fuels, primarily ethanol, as well as to chemicals, plastics, fibers, detergents, pharmaceuticals, and many other products.



Cellulase image from M. Himmel et al., "Cellulase Animation," run time 11 min., National Renewable Energy Laboratory (2000).

Like starch and sugar, cellulose is a carbohydrate (compound of carbon, hydrogen, and oxygen) made up of simple sugars (glucose) linked together in long chains called polysaccharides. These polymers form the structural portion of plant cell walls, and unraveling them is the key to economical ethanol fermentation. Technical barriers to large-scale use of cellulose technology include the low specific enzyme activity and high enzyme-production costs, as well as a general lack of understanding about enzyme biochemistry and mechanistic fundamentals.

In 2004, the DOE National Renewable Energy Laboratory (NREL), working with two of the largest industrial enzyme producers (Genencor International and Novozymes Biotech), achieved a dramatic reduction in cellulase enzyme costs. Cellulases belong to a group of enzymes known as glycosyl hydrolases, which break (hydrolyze) bonds linking a carbohydrate to another molecule. The new technology involves a cocktail of three types of cellulases—endoglucanases, exoglucanases, and beta-glucosidases. These enzymes work together to attack cellulose chains, pulling them away from the crystalline structure and breaking off cellobiose molecules (two linked glucose residues), splitting them into individual glucose molecules, and making them available for further processing. This breakthrough work resulted in 20- to 30-fold cost reduction and earned NREL and collaborators an R&D 100 Award (press release: www.nrel.gov/awards/2004hrvtd.html).

Further cost reductions are required, however, to support an economical and robust cellulose biorefinery industry. For example, costs of amylase enzymes for converting corn grain starch to ethanol are about 1 to 2 cents per gallon of ethanol produced, but the most optimistic cost estimates for cellulase preparations now are about tenfold higher than that. Routes to improving enzyme efficiencies include the development of enzymes with more heat tolerance and higher specific activities, better matching of enzymes and plant cell-wall polymers, and development of high-solid enzymatic hydrolysis to lower capital costs. A comprehensive understanding of the structure and function of these enzymatic protein machines, how their production and activity are controlled, and changes they promote on plant cell-wall surfaces will be critical for success.

(Also see sidebar, Image Analysis of Bioenergy Plant Cell Surfaces at the OBP Biomass Surface Characterization Lab, p. 40.)

lations to find environmental conditions (temperature, pressure, and molecular agents) that affect such disruption before the cellulase attacks the substrate.

- The kinetic and thermodynamic mechanism of processive cellulases or exoglucanases (GH family 7) is not known. These enzymes, protein machines, conduct most cellulose hydrolysis, with endoglucanases playing a lesser role. To understand this mode of catalysis at the molecular level, a mathematical model of the functioning of GH family 7 cel-

lulase must be developed and tested. Finally, the range of permissible modifications of these specific protein folds is not known, nor are constraints imposed by the fairly limited set of industrial host organisms on expression of wild-type members of GH families or newly engineered mutants.

- The major objective of the designer-cellulosome concept is to reconstruct improved cellulosomes by linking essential and most efficient enzymes to desired substrates. Numerous scientific issues and opportunities are still "at large."
  - Our capacity to control the final designer-cellulosome's composition and architectural arrangement will enable us to pose new hypotheses regarding enhanced cellulosome performance in degradation of plant cell-wall material.
  - The best set of cellulosomal cellulases most appropriate (synergistic) for use on lignocellulosic substrates is still unknown.
  - Although hybrid enzymes derived from noncellulosomal species can be incorporated into designer cellulosomes, whether their observed synergistic activity is comparable to that of native cellulosomal enzymes is still unknown.
  - Many other scientific questions have to be addressed, including the optimal stoichiometry and arrangement of components and whether the combined action of cellulosomal and noncellulosomal systems results in an additional improvement in performance. This research offers a unique opportunity to explore the expanding universe of molecular machines that can be constructed by the designer-cellulosome approach.

Technical Milestones

Within 5 years

- Near-term accomplishments in cellulase biochemistry should focus on improving the enzymes, primarily from fungi and actinomycetes, used in industrial preparations marketed for the biorefinery. This accomplishment would require searching for new enzymes from new sources (including metagenomic databases) and using directed evolution to probe mutational space for possible improvements over wild-type examples. The outcome would be better cocktails based on wild-type enzymes and enzymes improved by noninformational methods that could be used immediately by industry.
- Construction of designer cellulosomes will concentrate on several fronts, including incorporation of currently available enzymes; development of novel cellulosomal components—scaffoldins, cohesins, dockerins, CBMs, and linker segments; improvement of hybrid enzymes and assessment of their enhanced synergistic action within designer cellulosomes; and analysis of the combined action of designer cellulosomes with highly active noncellulosomal enzymes, including ligninases and hemicellulases.

Within 10 years

- The mid-term scope permits a more systematic approach to cellulase and cellulosome biochemistry. Such understanding will require combining classical biochemistry with computational science.
- A particularly important goal is identifying thermodynamic limitations of improvement, which would impact protein engineering.
- Initial studies to convert suitable host-cell microorganisms into cellulosome-producing strains will be pursued. The desired result will be improved biomass degradation and increased understanding of the structure-function relationship of cellulosome components.
- To achieve heterologous production, assembly, and secretion of active designer cellulosomes, suitable host cells that can accommodate the genes for such large proteins must be identified and inherent cloning and expression barriers overcome.

The Cellulosome: The "Swiss Army Knife" of Molecular Machines

The cellulosome is an extracellular supramolecular machine synthesized by some anaerobic microorganisms capable of degrading crystalline cellulose and other plant cell-wall polysaccharides. Each cellulosome contains many different complementary types of carbohydrate-active enzymes, including cellulases, hemicellulases, and carbohydrate esterases that are held together by a scaffoldin protein to form a single multienzyme complex (see Fig. A. Schematic of a Cellulosome, below). The cellulosome enhances cell-wall degradation by bringing several different enzymes into close proximity so they can work together to exploit enzyme-accessible regions of cellulose. The various product intermediates from one enzymatic subunit can be transferred readily to other enzymatic subunits for further hydrolysis (breakdown) of the cellulose. The cellulosome also promotes cell adhesion to the insoluble cellulose substrate, thus providing individual microbial cells with a direct competitive advantage in using soluble sugars produced in the hydrolysis. Cellulosomes need not be associated with cells for activity, and they function under both aerobic and anaerobic conditions.

Each enzymatic subunit contains a definitive catalytic module and a dockerin domain that binds tightly to a scaffoldin cohesin. Thus, cohesin-dockerin interaction governs assembly of the complex, while cellulosome interaction with

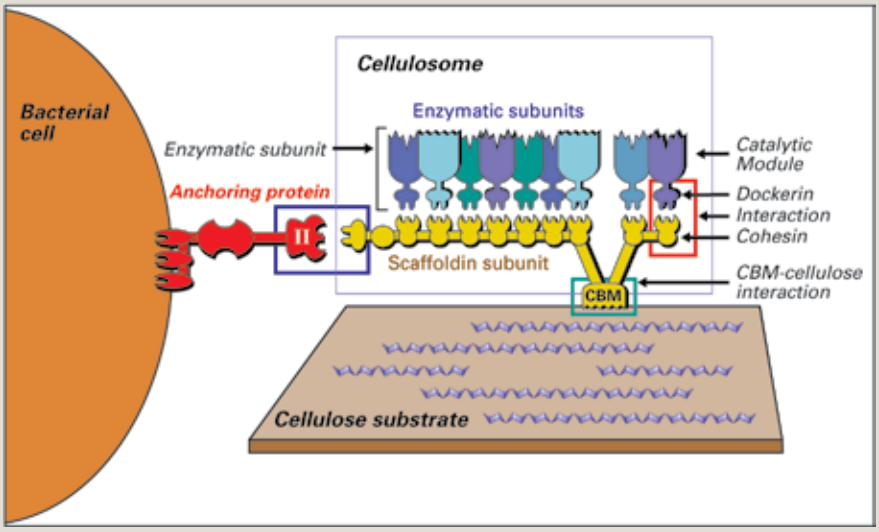


Fig. A. Schematic of a Cellulosome. The scaffoldin subunit, to which the cellulosomal enzymatic subunits are bound, is made up of different functional modules that dictate its architecture and various activities. These include a single CBM and nine very similar repeating domains called cohesins, which bind the enzymatic subunits.

Within 15 years

- HTP approaches and crystallographic and microscopic techniques will be used to elucidate the three-dimensional structure of cellulosomal components, the intact native and designer cellulosomes, and their action with pure (cellulose) and native (plant cell-wall) substrates.
- Understanding mechanistic principles of the entire accessible range of glycoside hydrolases and their function in the biosphere will be evaluated. The diversity of cellulase and cellulosome functional schemes will be modeled and optimized for specific biomass substrates (feedstocks). This improved understanding of cellulase action will provide new saccharification paradigms for the biorefinery.

Crosscutting Tools, Technologies, and Science

Computational science, advanced imaging, and HTP genomic tools must be incorporated to attain the objectives described here.

cellulose is mediated by scaffoldin-borne cellulose-binding molecules (CBM). Some scaffoldins also bear a divergent type of dockerin that interacts with a matching cohesin on an anchoring protein, thereby mediating cellulosome attachment to the cell surface.

The LEGO-like arrangement of cellulosomal modules offers an excellent opportunity to engineer new multienzyme complexes for desired purposes. The various cohesins (c), dockerins (d), and catalytic modules (A, B, C) are functionally independent and can be tethered together in any combination via recombinant genetics. The resulting complex can be applied either in the test tube (bioreactors) or in a cellular setting (fermentors).

Native scaffoldin-borne cohesins in cellulosome-producing bacteria generally recognize all enzyme-bearing dockerins nonspecifically. Thus, reconstruction of a cellulosome from its component parts would result in random incorporation of different enzymes into cellulosome complexes, yielding a heterogeneous population of artificial cellulosomes (Fig. B1. Random Incorporation, below).

To control the incorporation of desired enzymes into a precise position, distinctive cohesin-dockerin pairs must be used (Fig. B2. Controlled Incorporation, below.). Dockerin domains that bind a specific cohesin must be fused to the different catalytic modules, and a chimeric scaffoldin is engineered to contain complementary cohesins and a single CBM for substrate targeting. Subsequent self-assembly of the mature "designer" cellulosome complex can then occur selectively in vitro, resulting in a homogeneous population of cellulosome complexes.

Cellulosome References: p. 117

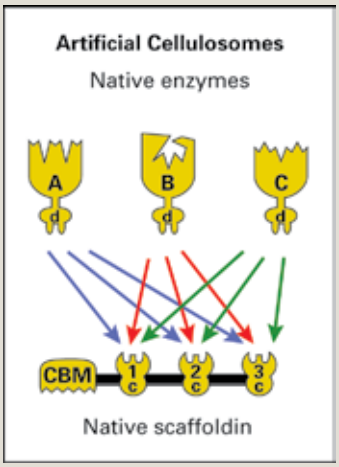


Fig. B1. Random Incorporation.

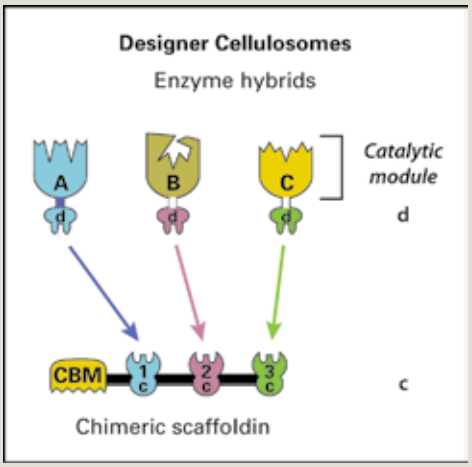


Fig. B2. Controlled Incorporation.

- Computational science is needed for understanding cellulase function at a holistic level, specifically, new programs and codes to handle models of ultralarge biological molecules and systems of more than 1 million atoms.
- New imaging methods are critical for elucidating cellulase action on the plant cell wall.
- Finally, tools to acquire, archive, and interpret new cellulase structures (both native and modified) are needed to better understand their natural structural and functional diversity.

We must develop new science and technical methodologies to explore protein-machine function at the molecular and atomic levels to attain these objectives. Solution physics and thermodynamics ruling protein-domain biological function at the nanometer scale are poorly understood. In this respect, the new field of nanoscience may provide important insights. For example, describing the mechanism of a processive cellulase at such level of detail will require more knowledge about the cellulose surface, water-protein-glycan dynamics, and biomolecular mechanics than is possible today.

The Role of GTL and OBP Facilities and Capabilities

Protein Production

Protein production capabilities will be critical for supplying samples of wild-type and recombinant proteins, including (1) new enzymes from new sources, (2) modified enzymes from directed evolution, (3) modified enzymes from site-directed mutation, and (4) large quantities of enzymes suitable for pilot testing. This resource will be critical in identifying the most effective new hemicellulases and ligninases without first having to solve high-level expression challenges for all candidates.

High-throughput (HTP) biochemical and biophysical assays will be developed and carried out, and functional annotation will enrich the first-pass sequence-based annotations. This problem is more difficult and complex for GH, in which a single enzyme can have several comparable activities on different polysaccharide substrates. Thus, a microchip HTP assay facility is needed.

Molecular Machine

These capabilities can provide analytical and computational science to study cellulase protein-machine function and advanced techniques for isolating and reconstituting complex interactions within native or improved hemicellulase and ligninase systems. Synergistic interactors including chaperonins, excretion paths, and cofactors will be isolated and identified in novel cellulases, hemicellulases, and ligninases.

Proteomics

Proteomic capabilities can be used to document how secretomes of selected white rot fungi and other organisms such as microorganisms in the termite hindgut respond to changing culture conditions and optimize biodegradation rates of lignocellulosic components. Correlations between

ligninase activity and concentrations of enzymes responsible for cosubstrate production will be of particular interest and will assist researchers in designing more effective experimental plans for improving the energetics and carbon-allocation efficiency of cell-wall degradation. This also will be important in monitoring and controlling the effects of substrate inducers and responses to heterologous expression, especially of the most "foreign" components. This resource can support studies aimed at improving production of ligninase and hemicellulase proteins from both near-term hosts such as *T. reesei* and longer-term systems including bacteria and yeast.

Cellular System

Cellular system capabilities will visualize cellulases, ligninases, cellulosomes, and other molecular species directly on substrates in real time under varying conditions and from varying species. Models will be developed to optimize hydrolysis under different conditions.

DOE Joint Genome Institute

DOE JGI could embark on sequencing new white rot fungal genomes, including those that do not encode lignin and manganese-dependent peroxidases. This work may lead to the discovery of more-active and less-labile ligninase variants. After that, more biological diversity from such bacterial populations as decaying biomass, termite hindgut, ruminants, and the rhizosphere could be examined.

Metagenomics

Metagenomic approaches will identify mechanisms and agents of effective hemicellulose and lignin conversion in communal or unculturable populations. They also can play a key role in sequencing new genomes thought to harbor cellulolytic components acquired by lateral gene transfer and in studying natural adaptations that allow (initially) heterologous expression.

Harvesting the Biochemical Potential of Microorganisms Through Metagenomics

The conceptual foundation for metagenomics is the realization that more than 99% of microbes in most habitats have not yet been cultured. By providing deep insight into biological capabilities and function without the need for culturing, metagenomics effectively expands our scientific capabilities and understanding beyond the small percentage of bacteria that currently can be cultivated in the laboratory. It further permits a more holistic and mechanistic analysis of microbial communities. Natural systems long ago learned how to attack lignocellulose and use the resultant sugars and other chemicals for their own purposes. The metagenomes of complex natural communities provide a fertile resource for data mining to search for new examples of relevant enzymes. The ability to construct and evaluate an engineered metagenome (from existing genome sequences) now enables creation of systems with superior capabilities.

This research has two objectives: (1) employ genetic and biochemical studies to characterize novel lignocellulose-degrading systems and (2) use metagenomics to augment these efforts by characterizing and recovering the genetic potential (including that from uncultured microbes) resident in microbial communities capable of rapid and extensive biomass degradation. Outcomes of these functional and comparative studies will include a repertoire of new enzymes and proteins available for engineering approaches (e.g., designer cellulosomes or free cellulase systems). In that context, knowledge and bioresources arising from this research are well integrated within the overall goal of improving plant cell-wall deconstruction and conversion of lignocellulosic products to bioethanol.

Research Goals

To elucidate the repertoire of gene products necessary to effect more rapid and extensive hydrolysis (and solubilization) of lignocellulosic materials, systematic advances in our understanding of microbe-mediated plant cell-wall deconstruction must be achieved. First, many sequenced microbial genomes encode more glycoside hydrolases for plant cell-wall deconstruction than have been identified through historical biochemical and genetic studies. Whether or not these multigene families are an evolutionary adaptation to provide more "total" activity diversity, to compensate for subtleties in substrate conformation and composition, or a combination of both is still unknown. In addition, degradative synergism in microbial communities often resides in the concerted actions of enzymes from multiple organisms, but mechanistic details and understanding of this concept are undeveloped.

Second, aerobic and anaerobic bacteria with superior degradation capabilities recently have been found to lack canonical processive cellulases required in all other well-described systems for substrate solubilization. The genetics and biochemistry underpinning their lignocellulose degradation, which could represent a new mechanism, must be explored.

Technical Milestones

GTL will accelerate the development of optimal cellulase systems by providing resources for screening thousands of natural and modified enzyme variants, enabling the HTP production and functional analysis of these enzymes, elucidating regulatory controls and essential molecular interactions, and generating models for analyzing the structure and activity of natural and engineered enzyme systems.

Within 5 years

- Undertake microbiome projects for functional and comparative purposes. The JGI Community Sequencing Program, supported by DOE, is now characterizing the microbiome of the termite hindgut.
- Initiate mechanistic characterization (at DNA, protein, and organismal levels) of no fewer than three natural microbial communities displaying rapid and extensive lignocellulose degradation. Each should be different

in terms of prevailing physicochemical conditions (e.g., temperature, pH, and salt concentration) or dominant feedstock. For example:

  - Soils representative of dominant residues in the DOE-USDA Billion-Ton study. Corn stover, rice, switchgrass, and forest residues would be appropriate.
  - Decaying biomass (see the sidebar, Examples of Metagenomic Analysis: Understanding the Dynamics of Microbial Colonization of Decaying Biomass, p. 108).

- Determine the mechanisms of lignocellulose solubilization by such known organisms as *F. succinogenes* and *C. hutchinsonii* to help guide metagenomic analyses, using a combination of biochemical and genomic approaches (see sidebar, Sequencing a Soil-Cellulose Degrader, this page).
- Develop heterologous expression systems for candidate enzymes and regulatory proteins, including necessary cofactors and post-translational modifications.
- Begin HTP production, characterization, and intercomparison of enzymatic systems discovered in metagenomic analyses. Produce appropriate molecular tags to support experimentation in natural and controlled systems—including imaging, protein isolation, and detection of interactions.

Within 10 years

- Complete comprehensive analyses of enzymatic systems to isolate different families and high performers.
- Rationally modify, express, and characterize native enzymes to understand design principles and optimize properties.
- Build computational models of biomass-decay microbial communities, and test against experimental data.

Within 15 years

- Deploy innovative processes based on these discoveries and bioresources to reduce enzyme costs and loading by tenfold. Provide new options that reduce or eliminate the need for nonbiological feedstock pretreatments.

The Role of GTL and OBP Facilities and Capabilities

In brief, the facilities will support and enable the attainment of these objectives. For instance, JGI already has contributed a foundational set of genome sequences for specialist cellulolytic microbes; this data set is augmented by genomes

Sequencing a Soil-Cellulose Degrader

The DOE Joint Genome Institute is sequencing *Cytophaga hutchinsonii*, an aerobic Gram-negative bacterium commonly found in soil that rapidly digests crystalline cellulose. Molecular analysis of cellulose degradation by *C. hutchinsonii* is now feasible, since techniques for genetically manipulating the organism recently have been developed. This microbe exhibits the ability to move rapidly over surfaces by a process known as gliding motility, which is thought to be important in allowing *C. hutchinsonii* to colonize its insoluble growth substrate. The mechanism of gliding motility is not known, but flagella are not involved. Analysis of the *C. hutchinsonii* genome sequence will facilitate studies of cellulose degradation and also will reveal more about bacterial gliding motility, which has remained an unsolved biological mystery for over 100 years (http://genome.jgi-psf.org/finished_microbes/cythu/cythu.home.html).

Example of Metagenomic Analysis: Understanding the Dynamics of Microbial Colonization of Decaying Biomass

As primary decomposers, microbial communities have evolved both as competitors and collaborators in biomass deconstruction. The ultimate aim of this research is to achieve bioprocesses for all steps in converting biomass to ethanol. A critical need is to replace plant biomass thermochemical pretreatment, which is now necessary to convert recalcitrant structural lignoellulose to a form in which cellulose is more accessible and amenable to hydrolytic enzyme action. By analyzing natural communities that colonize decaying biomass, we can ascertain natural mechanisms that can be used to supplant thermochemical treatments. The goal is to better understand the complex microbial communities responsible for lignocellulosic biomass deconstruction, harvest key biochemical decay mechanisms, and develop predictive modeling and control of these complex natural processes.

Metagenomics will be used to determine microbial-community composition and genetic diversity. Comparative genomic tools will initially determine community functionality and identify genes of unknown function. Coupled spatial and temporal measurements will reveal principles of community formation and dynamics, processes such as signaling, and other microbial interactions. Functional annotation of unknown proteins through protein production and characterization and other experimentation will augment initial gene-function assignment.

HTP methods from facility and other capabilities will be used for measuring enzyme activities:

- Determine cellular and biochemical functions of genes discovered in uncultured community members.
- Characterize the temporal composition and functional capability of microbial communities—Who is there, and what metabolic processes are being carried out?
- Characterize expression patterns of cellulolytic enzymes and related processes and pathways using transcriptomics, proteomics, and metabolomics.

Computational tools will predict the metabolic, physiologic, and behavioral characteristics of microbial communities from community DNA sequence data, and supporting measurements will be developed. The tools will allow design and engineering of microbial systems that ultimately can perform all steps in biomass processing to ethanol.

sequenced at The Institute for Genomic Research (called TIGR). Other resources can readily support hypothesis-driven research at the RNA, protein, organismal, and community levels (see sidebar, Example of Metagenomic Analysis, this page). This research will validate and produce the desired combination of proteins and other biologics necessary to make lignocellulose conversion comparable to starch as a feedstock in economic and process criteria.

Characterizing Cell Walls Using High-Throughput Methods

Understanding the chemical nature and architecture of cell walls at the nano-, micro-, macrochemical, and physical scales, as well as their behavior in pretreatment and fermentation, is essential in taking a systems-level approach to modifying the plant genome or increasing the biofuel system's efficiency. Knowledge and insight gained from HTP cell-wall characterization will drive several areas of biomass-to-biofuels research. These areas include selecting and modifying plants, matching the microbial conversion process to the substrates present, and minimizing inhibitors during biological conversion to valuable products. Many new instruments and analytical methods will be needed to characterize a range of biomass feedstocks (e.g., grasses, crop residues, short-rotation woods, and early energy crops) needed to achieve the Billion-Ton vision. Ultimately, many different types of feedstocks will be at different stages of harvest and hence will need to be analyzed rapidly. The robustness of these methods and their versatility in addressing the composition of many different materials requires significant further research. Data generated by these methods are used to value feedstocks, measure conversion efficiency, identify regulatory issues, and ultimately establish product pricing and investment risk.

Research must address the disconnect between the throughput of genomic and proteomic analysis and that of biomass chemical and structural characterization. DNA sequencing and metabolic profiling are routinely done at the microliter scale, where 1000 to 2000 samples can be processed and

analyzed in one day. Most biomass analysis requires hundreds of milligrams per sample and very few have been automated, so throughput is around 20 samples per week per person. Recent advancements in biomass analysis have demonstrated high throughput using multivariate analysis (MVA) coupled with analytical pyrolysis, FTIR, or near infrared spectroscopy to allow characterization of 200 to 500 samples per day per person. These spectroscopic methods, however, require calibration and validation using slow wet chemical procedures. MVA often is specific to feedstock and process; for example, a method developed for hardwood analysis won't accurately characterize a softwood or grass sample. To support the broad goals of this roadmap, many new analytical methods will need to be developed.

Research will focus on two areas: Improving throughput of traditional calibration methods to around 500 samples per week and using thse data to calibrate "next-generation" capabilities of analyzing 1000 or more samples per day. Strategies for improving traditional methods will require the adaptation of analytical instruments for biomass applications and development of new instruments specifically for this work. For example, multiplexing robotic sample preparation and multichannel capillary electrophoresis in carbohydrate analysis potentially could offer a tenfold increase in sample throughput, but such instruments and methods are not validated for biomass samples. New methods and instruments will generate data on molecular- and genetic-level relationships needed to (1) improve feedstocks by increasing biomass production and making feedstocks more amenable to the next generation of biomass-conversion technologies, and (2) develop data sets that enable the design of microbial cell conversion strategies that will produce desired products in high yield and purity. A special category of analytical methods with high sensitivity also will be needed to interpret high-resolution images of feedstocks and process substrates.

The complexity of the biomass matrix presents significant analytical challenges not faced in conventional genomics work. Chemical characterization methods will have to assess the relative weight percent of about 14 constituents in each feedstock. Structural constituents of potential interest are protein, lignin, ash, glucan (cellulose), xylan, mannan, arabinan, galactan, glucuronic acid, ferulic acid, acetyl groups, starch, and galacturonic acids (pectin). Some nonstructural materials, which will be specific to each feedstock, also may be of interest in saccharification and fermentation studies. These constituents include inorganic salts, sucrose, tars, waxes, gums, lignans, and others. Analytical methods also will be needed to track the constituents and their reaction products through pretreatment, saccharification, and fermentation. Structural characterization methods must quantify functional groups of interest (e.g., carboxylic acids, ketones, aldehydes, esters, and methoxyl and free-phenolic hydroxyl groups). With the requirement for so many measurements, a portfolio of complementary methods will be needed to close mass balance of feedstock and process components across multiple processing steps.

This research program addresses a critical component in the feedstock-sugars interface by providing analytical instruments and methods with the

precision, accuracy, and throughput required to optimize biomass selection and development, biomass pretreatment, and conversion processes leading to economical ethanol production in biorefineries. Biomass-relevant analytical tools of this type do not now exist for these applications.

Tracking individual biomass components from feedstock to products requires high-accuracy data obtained in an integrated and consistent fashion. A research program focused on analytical methods that coordinate data through all stages will improve process integration through consistent data on feedstocks, substrates, process inhibitors, and products. Integrated data will furnish a systems-level understanding of process streams and enable the desired correlations between reactivity and genome expression.

To enable future development of high-energy biomass feedstocks, method portfolios need to be standardized; validated in a consensus environment; and published in a forum available to plant breeders, field scientists, process engineers, enzyme scientists, and fermentation scientists. Technologies will provide data for identifying relationships among plant cell-wall components and interactions among biomass-derived substrates, microbes, and enzymes. Although progress has been made in chemical procedures for biomass analyses, existing portfolios of methods are labor intensive, time consuming, expensive, and generally not amenable to the HTP needs of breeders, agronomists, feedstock processors, and systems approaches to fermentation. To bypass this limitation, HTP techniques need to measure the amounts and structure of cell-wall components in processes. Correlations among process data and genomic and proteomic data will enable identification of genes most important for improved biomass conversion.

For example, more than 225,000 independent T-DNA insertion lines of *Arabidopsis* have been created that represent almost the entire genome space. HTP analysis must be able to screen these large sample sets, as well as the much larger number of samples produced for other species with more complicated genomes, to determine the role each gene plays in different cell-wall chemistry phenotypes. HTP methodologies also will enable screening of plant, enzyme, and microbe consortia to allow for deeper understanding of different plant and microbe genome interactions. High-sensitivity methods will integrate with imaging tools to enable spatial compositional and functional group determinations in individual plant cells. These studies require the analysis of hundreds if not thousands of samples and would not be possible without the cost and time saving provided by rapid biomass-analysis methods.

Technical Milestones

Within 5 years

For near-term deployment, portfolios of precise and accurate analytical methods for a wide range of structural and nonstructural constituents and biomass function groups will be developed and validated in a consensus mode and published for general use. Specific techniques and methods will be tailored to provide comprehensive characterization of biomass model

substrates (feedstocks). Working closely with commercial suppliers, investigators will create instruments and techniques specific to the needs of biomass-conversion genomics. Data obtained through customized analytical procedures will be used to calibrate and validate rapid MVA methodologies for HTP screening of feedstocks and biomass-derived research samples. This screening will help to select samples of interest from large populations representing species whose genome sequence has been determined. Selection will be based on identifiable differences in phenotypes of major plant cell-wall biopolymers (lignin, cellulose, hemicellulose, and others) important for conversion to fuels and valuable chemicals. Data generated will be used to support the breeding and development of new plant lines. In conjunction with HTP methods, information-technology strategies will capture, analyze, manage, and disseminate resulting data, enabling creation of a national resource for the biofuel and plant biology research community. To the extent possible, samples and data used in calibrating new methods will be archived safely for future applications. Archived samples may be required for benchmarking, new method development, and crossplatform comparisons.

Within 10 years

HTP methodologies will be integrated with other studies to provide a systems-level understanding of how to alter and improve plant cell-wall composition and structure for efficient and economic biofuel production. High-resolution, molecular-specific images enabled by the availability of molecular tags will allow spatial determinations of chemical and structural features across individual cell types and structures.

Within 15 years

Deployment of HTP and high-sensitivity analytical methods and appropriate data-reduction (informatic) techniques will be completed to integrate genomic and proteomic data with cell-wall chemistry information. A systems-level understanding of interactions between plant cell-wall structure and the microbes and enzymes is needed to convert biomass effectively to fuels and chemicals.

The Role of GTL and OBP Facilities and Capabilities

Many underlying principles and HTP techniques (e.g., robotics, informatics, and pattern recognition) for HTP chemical analysis of plants have been or are being developed by the genomic and metabolomic communities. Collaborations with researchers using the resource will require the adaptation and validation of these techniques for biomass-conversion applications. We expect technologies developed for plants (e.g., high-sensitivity spatial detection of biopolymers) to have applications for studying other organisms.

Information obtained from these analytical methodologies will facilitate the determination of plant-gene function in synthesis of plant cell walls and biopolymers, especially genes related to carbohydrate metabolism, polymerization, and modification. Analytical methods tailored to biomass applications will be used to identify protein function related to cell-wall construction. As

additional genomes are sequenced, these tools can be used to characterize gene function in other plant lines of interest for biofuel production.

Crosscutting Tools, Technologies, and Science

Many analytical methods that support a systems approach to biomass conversion will require biomass-relevant standards and tags not readily available from commercial sources. Obtaining the needed array of small-molecule standards will require prep-scale isolation of molecules of interest and advanced techniques of carbohydrate and natural product organic synthesis. As we learn more about the plant genome, important intermediate molecules and biomarkers will be revealed. Techniques for isolating and preparing relevant biomass standards will be transferred to the private sector and made available for biomass research and commercial biorefineries.

While much existing robotic instrumentation and data-reduction technologies are available for constructing HTP methodologies, several areas must be addressed to achieve project goals. New techniques for homogenizing and reducing the particle size of biomass samples without degradation will be needed to prepare biomass samples for automated and robotic systems. New modules, equipment, and accessories for existing analytical techniques such as columns for liquid chromatography will have to be adapted specifically for biomass samples. Rapid biomass-analysis methods with large dynamic ranges will detect multiple compounds with high precision, acceptable accuracy, and short acquisition times. Because these methods are often specific to feedstocks and processes, many new ones may be required. Integrated data from different analytical methods into a single comprehensive MVA calibration set will need testing and validation.

Data-reduction and storage methods with predictive platforms for production and processing pipelines are necessary for retrieving and integrating data from breeding or field studies. New "hands-off" bioinformatic methods will combine, analyze, and correlate data (including metabolomic and genomic information) from multiple systems. Multivariate analysis tools will be applied where pertinent to information extraction from complex systems.

Appropriate biomass analysis is a necessary first step in applying systems biology techniques to bioenergy production. In many cases, analytical methods for biomass substrates that demonstrate the required precision, accuracy, and speed are not available. Analysis using existing methods is too slow and expensive for large-scale screening outlined in many basic technology research programs. Before yields of a desired product can be maximized or undesirable side reactions minimized, analytical methods must accurately monitor changes in the chemical constituents. These methods also must be validated in the complex matrix of biomass-conversion streams. In many cases, biomass methods direct the course of research by revealing the presence of important products or the nature of yield-reducing side reactions. Biomass-conversion methods will lead us one step closer to those for accurately monitoring complex degradation reactions in natural systems.

Breakthrough, High-Payoff Opportunity: Simplifying the Bioconversion Process by Understanding Cell-Wall Deconstruction Enzymes Expressed in Plants

Scientific Challenges and Opportunities

Complexity is one of the major challenges to economical biomass conversion for large-scale production of liquid transportation biofuels. The process currently requires many steps, including pretreatment, detoxification, solid and liquid separation, cellulase production, cellulose hydrolysis, biomass sugar fermentation, and product recovery supported by such ancillary systems as wastewater treatment and utilities. This complexity results in added operating and capital costs as well as issues of robustness, and simpler procedures would be highly desirable. Although engineering potentially could improve the process incrementally, systems biology approaches have the potential to revolutionize it. Various strategies envisioned to consolidate processing steps will genetically incorporate capabilities within organisms. One potential strategy would include incorporating genes that encode lignocellulose-degradative enzymes into feedstock plants.

- Enzymes could target wall polysaccharides (backbones, side chains, and small substituents that can have big effects on activity of other enzymes), simple phenolic cross-links between polysaccharides such as those that tie together xylans in cereal walls.
- Promoters would be required that are activated from senescence-related or harvest insult–triggered genes or by a signal provided exogenously at a time deemed appropriate.
- Genes encoding enzymes or other disassembly proteins would be required, perhaps those of microbial origin whose discovery would be a major "engineering" goal of the larger project and whose expression would be driven by the promoters above.

Research Goals

While targeting specific objectives using the existing state of technology as a benchmark is important, research must be open to new ideas and concepts. Specific areas that need to be addressed in fundamental and applied areas include the following.

Fundamental Science

- What new learning would enable expression of structural polysaccharide-degrading enzymes in plant tissues? We need new promoters, genes, transformation tools, and knowledge about the sequestration and action of such enzymes when expressed in living plants.
- What is the relative effectiveness of cellulase-enzyme systems acting in cellulose-enzyme-microbe complexes as compared to cellulose-enzyme complexes, and what is the mechanistic basis for such enzyme-microbe synergy?

Applied Science

- Can hemicellulose and cellulose hydrolysis be accomplished without the acidic or alkaline high-temperature and -pressure pretreatment step? This is particularly important for advantageous expression of hemicellulases and cellulase enzymes in plants.
- Can biocatalysts be developed whose performance is not impacted significantly by inhibitory compounds formed as by-products of pretreatment, hydrolysis, or fermentation? If so, they have the potential to eliminate the costly and waste-intensive detoxification step.
- Can such upstream processes as storage (ensilage) or biopretreatment be developed that make cellulose microfibrils more amenable to enzymatic cellulose hydrolysis without the need for relatively harsh thermochemical pretreatment?

References

Aden, A., et al. 2002. *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover,* NREL Report No. TP-510-32438.

Agblevor, F., H. L. Chum, and D. K. Johnson. 1993. "Compositional Analysis of NIST Biomass Standards from the IEA Whole Feedstock Round Robin," pp. 395–421 in *Energy from Biomass and Wastes XVI: Proceedings of the 7th Institute of Gas Technology Clean Fuels and Energy from Biomass and Wastes Conference*, ed. D. L. Klass, Institute of Gas Technology, Chicago.

Casler, M. D., and H.-J. Jung. 1999. "Selection and Evaluation of Smooth Bromegrass Clones with Divergent Lignin or Etherified Ferulic Acid Concentration," *Crop Sci.* **39**, 1866–73.

Chum, H. L., et al. 1994. "Status of the IEA Voluntary Standards Activity–Round Robins on Whole Wood and Lignins," pp. 1701–16 in *Advances in Thermochemical Biomass Conversion: Papers from the International Conference on Advances in Thermochemical Biomass Conversion Vol. 2,* ed. A. V. Bridgwater, Blackie Academic & Professional, Glasgow, U. K.

de la Rosa, L. B., et al. 1994. "Integrated Production of Ethanol Fuel and Protein from Coastal Bermudagrass," *Appl. Biochem. Biotechnol.* **45–46**, 483–97.

Durot, N., F. Gaudard, and B. Kurek. 2003. "The Unmasking of Lignin Structures in Wheat Straw by Alkali," *Phytochem.* **63,** 617–23.

Gollapalli, L. E., B. E. Dale, and D. M. Rivers. 2002. "Predicting Digestibility of Ammonia Fiber Explosion (AFEX)- Treated Rice Straw," *Appl. Biochem. Biotechnol.* **98–100**, 23–35.

Greene, N., et al. 2004. *Growing Energy: How Biofuels Can Help End America's Oil Dependence*, Natural Resources Defense Council, New York (www.nrdc.org/air/energy/biofuels/biofuels.pdf).

Hames, B. R., et al. 2003. "Rapid Biomass Analysis: New Tools for Compositional Analysis of Corn Stover Feedstocks and Process Intermediates from Ethanol Production," *Proceedings of the Twenty-Fourth Symposium on Biotechnology for Fuels and Chemicals, 28 April–1 May, 2002, Gatlinburg, Tennessee,* ed. B. H. Davis et al., *Appl. Biochem. Biotechnol.* **105-108**, 5–16.

Himmel, M. E., M. F. Ruth, and C. E. Wyman. 1999. "Cellulase for Commodity Products from Cellulosic Biomass," *Curr. Opin. Biotechnol.* **10**(4), 358–64.

Kelley, S. S., et al. 2004. "Rapid Analysis of the Chemical Composition of Agricultural Fibers Using Near Infrared Spectroscopy and Pyrolysis Molecular Beam Mass Spectrometry," *Biomass Bioenerg.* **27**(1), 77–88.

Kim, Y., et al. 2005. "Plug Flow Reactor for Continuous Hydrolysis of Glucans and Xylans from Pretreated Corn Fiber," *Energy Fuels* **19**(5), 2189–2200.

Lam, T.-B. T., et al. 1993. "The Relationship Between in Vitro Enzymatic Digestibility of Cell Walls of Wheat Internodes and Compositional Changes During Maturation," *Acta. Bot. Neerl.* **42**, 175–85.

Lapierre, C., D. Jouin, and B. Monties. 1989. "On the Molecular Origins of the Alkali Solubility of *Graminae* Lignins," *Phytochem.* **28,** 1401–3.

Laureano-Perez, L., et al. 2005. "Understanding Factors that Limit Enzymatic Hydrolysis of Biomass," *Appl. Biochem. Biotechnol.* **121**, 1081–1100.

Lawoko, M., G. Henriksson, and G. Gellerstedt. 2005. "Structural Differences Between the Lignin-Carbohydrate Complexes Present in Wood and in Chemical Pulps," *Biomacromolecules* **6,** 3467–73.

Lynd, L. R. 1996. "Overview and Evaluation of Fuel Ethanol from Cellulosic BIomass: Technology, Economics, the Environment, and Policy," *Annu. Rev. Energy Environ.* **21**, 403–65.

Lynd, L. R., R. T. Elander, and C. E. Wyman. 1996. "Likely Features and Costs of Mature Biomass Ethanol Technology," *Appl. Biochem. Biotechnol.* **57–58**, 741–61.

Moniruzzaman, M., et al. 1997. "Fermentation of Corn Fibre Sugars by an Engineered Xylose-Utilizing *Saccharomyces* Yeast Strain," *World J. Microbiol. Biotechnol.* **13**, 341–46.

Mosier, N., C. M. Ladisch, and M. R. Ladisch. 2002. "Characterization of Acid Catalytic Domains for Cellulose Hydrolysis and Glucose Degradation," *Biotechnol. Bioeng.* **79**(6), 610–18.

Mosier, N., et al. 2005. "Features of Promising Technologies for Pretreatment of Lignocellulosic Biomass," *Bioresour. Technol.* **96**(6), 673–86.

Sarikaya, A., and M. R. Ladisch. 1997. "An Unstructured Mathematical Model for Growth of *Pleurotus ostreatus* on Lignocellulosic Material in Solid-State Fermentation Systems," *Appl. Biochem. Biotechnol.* **62**, 71–85.

Scobbie, L., et al. 1993. "The Newly Extended Maize Internode: A Model for the Study of Secondary Cell Wall Formation and Consequences for Digestibility," *J. Sci. Fd. Agric.* **61,** 217–25.

Sedlak, M., and N. W. Y. Ho. 2004. "Production of Ethanol from Cellulosic Biomass Hydrolysates Using Genetically Engineered *Saccharomyces* Yeast Capable of Cofermenting Glucose and Xylose," *Appl. Biochem. Biotechnol.* **113–16**, 403–5.

Singh, R., et al. 2005. "Lignin-Carbohydrate Complexes from Sugarcane Bagasse: Preparation, Purification, and Characterization," *Carbohydr. Polym.* **62,** 57–66.

"Standard Biomass Analytical Procedures: NREL Laboratory Analytical Procedures" (www.eere.energy.gov/biomass/analytical_procedures.html).

Wilson, J. R., and D. R. Mertens. 1995. "Cell Wall Accessibility and Cell Structure Limitations to Microbial Digestion of Forage," *Crop Sci.* **35**, 251–59.

Wilson, J. R., and Hatfield, R. D. 1997. "Structural and Chemical Changes of Cell Wall Types During Stem Development: Consequences for Fibre Degradation by Rumen Microflora," *Aust. J. Agr. Res.* **48**(2), 165–80.

Cellulosome References

Cellulosome Reviews

Bayer, E. A., et al. 2004. "The Cellulosomes: Multi-Enzyme Machines for Degradation of Plant Cell Wall Polysaccharides,"*Annu. Rev. Microbiol.* **58,** 521–54.

Bayer, E. A., et al. 1998. "Cellulose, Cellulases, and Cellulosomes," *Curr. Opin. Struct. Biol.* **8**, 548–57.

Demain, A. L., M. Newcomb, and J. H. Wu. 2005. "Cellulase, *Clostridia*, and Ethanol," *Microbiol. Mol. Biol. Rev*. **69,** 124–54.

Doi, R. H., and A. Kosugi. 2004. "Cellulosomes: Plant-Cell-Wall-Degrading Enzyme Complexes." *Nat. Rev. Microbiol.* **2,** 541–51.

Shoham, Y., R. Lamed, and E. A. Bayer. 1999. "The Cellulosome Concept as an Efficient Microbial Strategy for the Degradation of Insoluble Polysaccharides," *Trends Microbiol.* **7**, 275–81.

Designer Cellulosomes

Bayer, E. A., E. Morag, and R. Lamed. 1994. "The Cellulosome—A Treasure Trove for Biotechnology," *Trends Biotechnol*. **12**, 378–86.

Fierobe, H.-P., et al. 2005. "Action of Designer Cellulosomes on Homogeneous Versus Complex Substrates: Controlled Incorporation of Three Distinct Enzymes into a Defined Tri-Functional Scaffoldin," *J. Biol. Chem.* **280,** 16325–34.

Fierobe, H.-P., et al. 2002. "Degradation of Cellulose Substrates by Cellulosome Chimeras: Substrate Targeting Versus Proximity of Enzyme Components," *J. Biol. Chem.* **277**, 49621–30.

Fierobe, H.-P., et al. 2002. "Designer Nanosomes: Selective Engineering of Dockerin-Containing Enzymes into Chimeric Scaffoldins to Form Defined Nanoreactors," pp. 113–23 in *Carbohydrate Bioengineering: Interdisciplinary Approaches,* ed. T. T. Teeri et al., The Royal Society of Chemistry, Cambridge.

Fierobe, H.-P., et al. 2001. "Design and Production of Active Cellulosome Chimeras: Selective Incorporation of Dockerin-Containing Enzymes into Defined Functional Complexes," *J. Biol. Chem.* **276**, 21257–61.