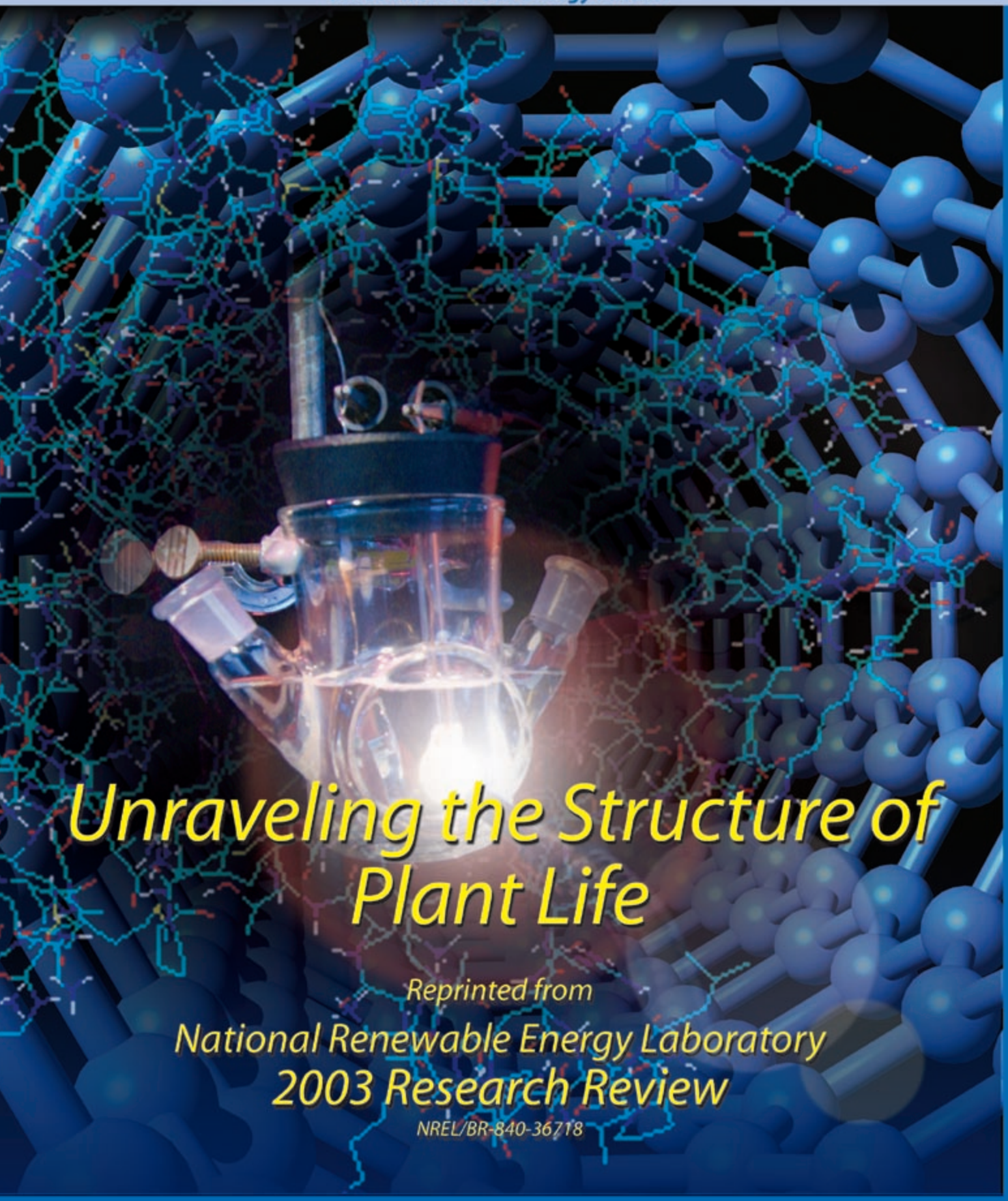
A national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy

NREL National Renewable Energy Laboratory

Innovation for Our Energy Future



Unraveling the Structure of Plant Life

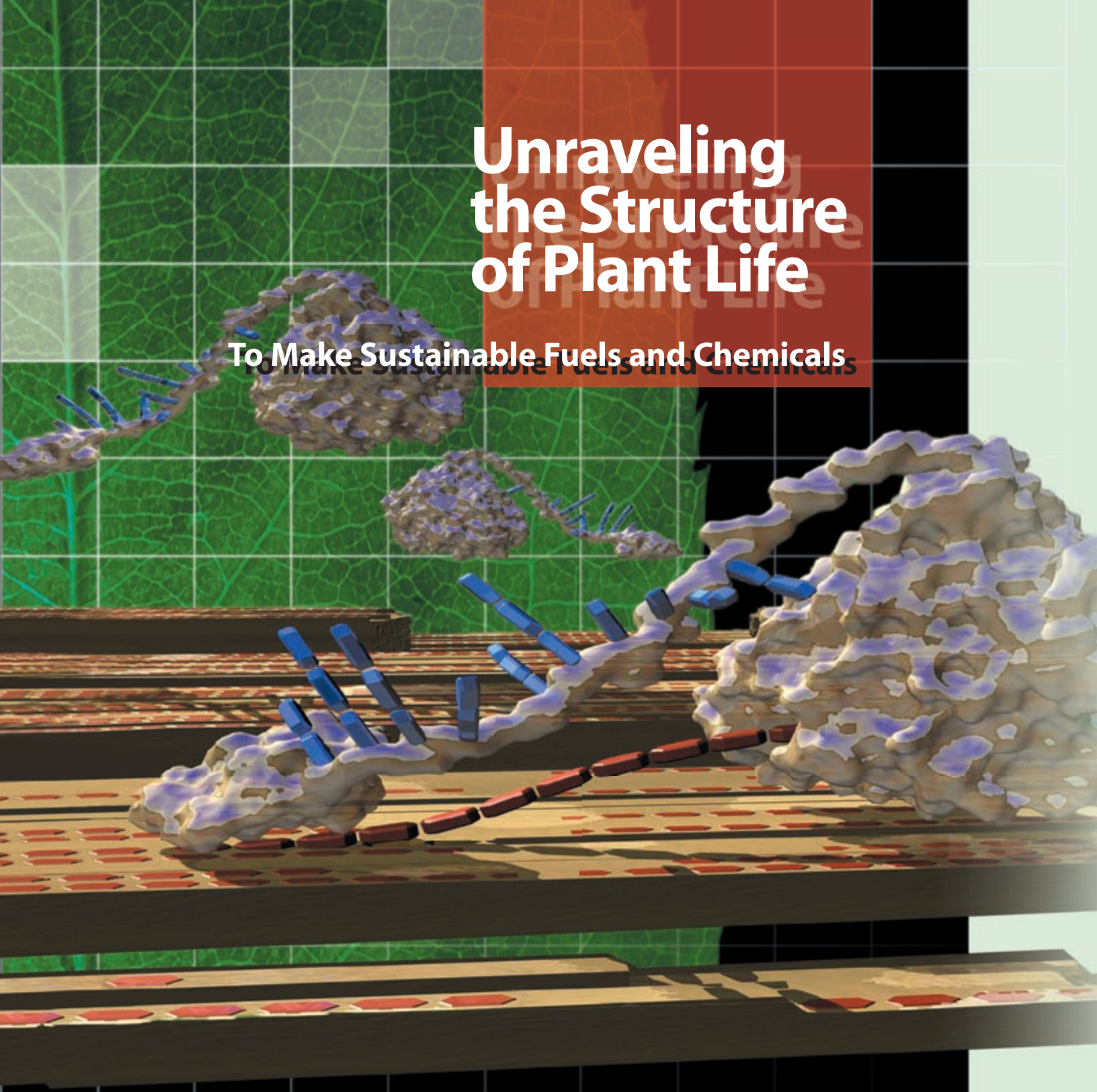
Reprinted from

National Renewable Energy Laboratory
2003 Research Review

NREL/BR-840-36718

Unraveling the Structure of Plant Life

To Make Sustainable Fuels and Chemicals



Imagine the look on the bartender's face if you asked for your happy-hour cocktail to be 19 parts endoglucanase, 90 parts exoglucanase, and 1 part betaglucosidase. If you wanted a wood chips hors d'oeuvre, or to get energy from the fiber in your wheat crackers or corn chips, that enzyme cocktail would be just the thing. Endoglucanase, exoglucanase, and betaglucosidase are cellulases, a family of enzymes that act in concert to hydrolyze (decompose in water) the cellulosic fiber in plant material to sugar. The sugar can then be used to make chemicals or ethanol fuel for our cars and trucks. According to NREL scientists, lowering the cost of the cellulases in that enzyme cocktail is the most promising avenue toward a competitive and domestic renewable alternative to gasoline. It is also a key element for developing a U.S. biorefinery industry that could make a wide range of chemicals and products from biomass — plants and organic wastes — as an alternative to chemicals derived from petroleum (see sidebar "What is a Biorefinery," page 12).

The United States already makes more than 2 billion gallons of ethanol per year from the starch in corn kernels and other grain. This is added as a 5% to 10% mixture to about one out of eight gallons of gasoline sold. But the primary use of corn is as animal feed, so it is relatively expensive, and there is a limit to how much ethanol can be made from it. NREL is developing technology to produce "cellulosic ethanol" from the fibrous material that makes up the bulk of plant matter. This dramatically increases the potential supply, adding inexpensive materials such as corn stalks and cobs, municipal wastes, sawdust and wood chips, and "energy crops" like grasses and fast-growing trees. Although the materials are relatively cheap, the conversion technology is not — yet. Currently, one of the biggest costs is for the cellulase enzymes used to convert cellulose to sugar. Although grain ethanol plants are sophisticated industrial operations, their underlying technology is relatively simple, similar to that used for moonshine stills. A common and inexpensive enzyme with a long history of industrial use for ethanol production — amylase — efficiently converts starch to sugar. Cellulases, however, are highly complex and do not have that history.

So NREL and DOE contracted with the world's two largest enzyme companies, Genencor International and Novozymes — to reduce the cost of producing cellulases. The goal is to bring the cost of the enzymes down to about $0.10/gallon of ethanol produced, which is key to making ethanol derived from cellulose economically competitive. Both companies are reporting excellent progress, but the goal is ambitious (see sidebar "Bringing in the Big Guns") and the natural resistance of cellulose to decomposition makes the task challenging.

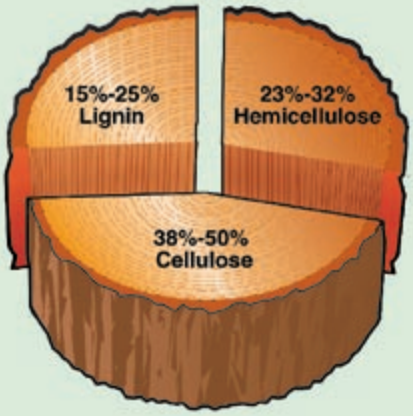
The Recalcitrance of Biomass

The name of the game is hydrolysis, breaking down complex carbohydrates (compounds of carbon, hydrogen, and oxygen) into their component sugars — analogous to how petroleum refineries break down complex hydrocarbons (compounds of carbon and hydrogen) into simpler chemicals, which are then built back up into desired fuels, plastics, and other chemicals.

The difference is that, unlike petroleum, plant carbohydrates are renewable. The vast bulk of plant material (most anything that is considered fibrous) consists of cellulose, hemicellulose, and lignin, with substantial starch and sugar found primarily in fruits and certain roots and tubers. Like starch and sugar, however, cellulose and hemicellulose are carbohydrates. But the sugars of which they are made are linked together in long chains called polysaccharides, which form crystalline structures — nature's plastics. Unraveling those polymeric structures is the key to economic biological conversion of cellulosic biomass to valuable fuels and chemicals. NREL scientist and cellulase expert Mike Himmel likes to refer to this complex and highly protective structure (the plant cell wall) as "nature's cunning plan" to keep plants standing and resistant to microbes. He refers to the resulting challenge for bioconversion as the "recalcitrance of biomass."

Cellulose consists of a loose crystalline structure of thousands of strands, with each strand containing hundreds of glucose sugar molecules ($C_6H_{12}O_6$). The cellulose is, in turn, wrapped in a sheath of hemicellulose and lignin, which further protects the cellulose. Hemicellulose is easier to hydrolyze than cellulose. A combination of heat, pressure, and acidic or basic conditions (each a way of causing hydrolysis, contributing an element of "severity") breaks down hemicellulose into its component sugars. (Hemicellulose sugars are different from the glucose of cellulose and harder to ferment — but that's another story). NREL researchers use dilute sulfuric acid for this thermochemical "pretreatment," which hydrolyzes hemicellulose to its component sugars, making them soluble in water along with some of the lignin. The "solubilized" hemicellulose liquid can then be separated, leaving the cellulose and remaining lignin as solids. This pretreatment step not only hydrolyzes the hemicellulose, but also effectively removes the hemicellulose/lignin sheath from the cellulose, leaving the cellulose accessible to further hydrolysis.

Currently, cellulose is hydrolyzed in a similar fashion to the hydrolysis of hemicellulose, but with far greater severity, which is typically



So-called "cellulosic biomass" (trees, crop residues, municipal solid wastes, grasses, etc.) is primarily made up of cellulose, hemicellulose, and lignin. The proportion of these constituents varies with the type of biomass.

Bringing in the Big Guns

To reduce the cost of cellulases, NREL and DOE contracted with the two largest enzyme companies in the world — Genencor International and Novozymes. Genencor sells enzymes and other biotechnology products for the health care, agri-processing, industrial, and consumer markets. Novozymes produces enzymes for a variety of markets and performs biotechnology research for the pharmaceutical, agricultural, and biochemical industries.

Both companies have already exceeded their initial objective, which was to reduce the cost of producing cellulases to one-tenth of what it was to start with. Having met their initial objective, they are performing follow-up work that could cut costs another several fold — to about $0.10/gallon of ethanol produced. This will make enzymatic hydrolysis more cost-competitive and overcome a critical barrier for making cellulosic ethanol an economic reality along with grain ethanol. And, with the promise of the ongoing research to engineer even cheaper and more efficient cellulases, combined with advances in other aspects of biomass conversion technology, cellulosic ethanol may even become competitive with gasoline.

achieved with strong acid or higher temperature. This requires more expensive processing equipment — one reason why NREL researchers determined that biological (enzymatic) hydrolysis could be more cost-effective in the long run. (Another reason is that acid hydrolysis is fairly well developed, with little room for further cost saving, whereas enzymatic hydrolysis has great potential for cost reductions.) And cellulose can certainly be broken down biologically. Although humans cannot digest cellulose, cattle, termites, beaver, and mushrooms can. Some bacteria, fungi, and insects produce cellulases themselves; other animals host cellulase-producing bacteria in their digestive tracts. But, in keeping with the complex, recalcitrant structure of cellulose, even after the hemicellulose/lignin sheath is removed, enzymatic conversion is not simple.

A Marvel of Nanomachinery

Most cellulases are systems of three types of enzymes, proteins that work together to catalyze biological conversion processes such as cellulose hydrolysis. First, an endoglucanase attacks one of the cellulose chains within the crystal structure, breaking it via hydrolysis, and creating new chain ends. During this hydrolysis, a molecule of water is consumed, and one of the chain ends becomes "reducing" and the other "non-reducing." Then — in what is a remarkable example of nanoscale machinery — an exoglucanase attaches to a loose end, pulls the cellulose chain out of the crystal structure, and then works its way down the chain, breaking off cellobiose (dimers of two glucose molecules) as it goes. Actually, there are two types of exoglucanase to match the two types of loose chain ends. A cellobiohydrolase I (CBH I) attaches to the "reducing" end, and a cellobiohydrolase II (CBH II) attaches to the "non-reducing" end. Finally, a betaglucosidase splits cellobiose into two separate glucose molecules, making them available for processing into chemicals or fuels.

The dominant cellulase systems considered thus far for industrial processing have come from fungi, in particular from *Trichoderma reesei*. Cellulase researchers, however, also explore enzymes produced by other fungi and bacteria for traits or capabilities that might improve the enzymatic hydrolysis process. NREL scientists, for example, have investigated *Acidothermus cellulolyticus*, a bacterium they found in hot springs in Yellowstone National Park. Even though bacterial exoglucanases are not usually as good as fungal ones, they have a tolerance for high temperatures that could be used to speed up bioprocessing. The problem then becomes one of taking a fungus that normally grows on rotting wood scattered through the forest, mass producing it in a factory setting, extracting enzymes for industrial processing, and making the enzymes more effective by incorporating features such as the high-temperature tolerance of NREL's hot springs bacterial enzymes.

The High-Tech World of Protein Engineering

The two goals NREL set for Genencor and Novozymes (see sidebar "Bringing in the Big Guns," page 11) are: (1) reduce the cost of producing the cellulases; and (2) make them more effective, so that less enzyme is needed. While the first avenue may include such mundane measures as optimizing growth conditions or processes, both lie predominantly in the high-tech world of protein engineering and production.

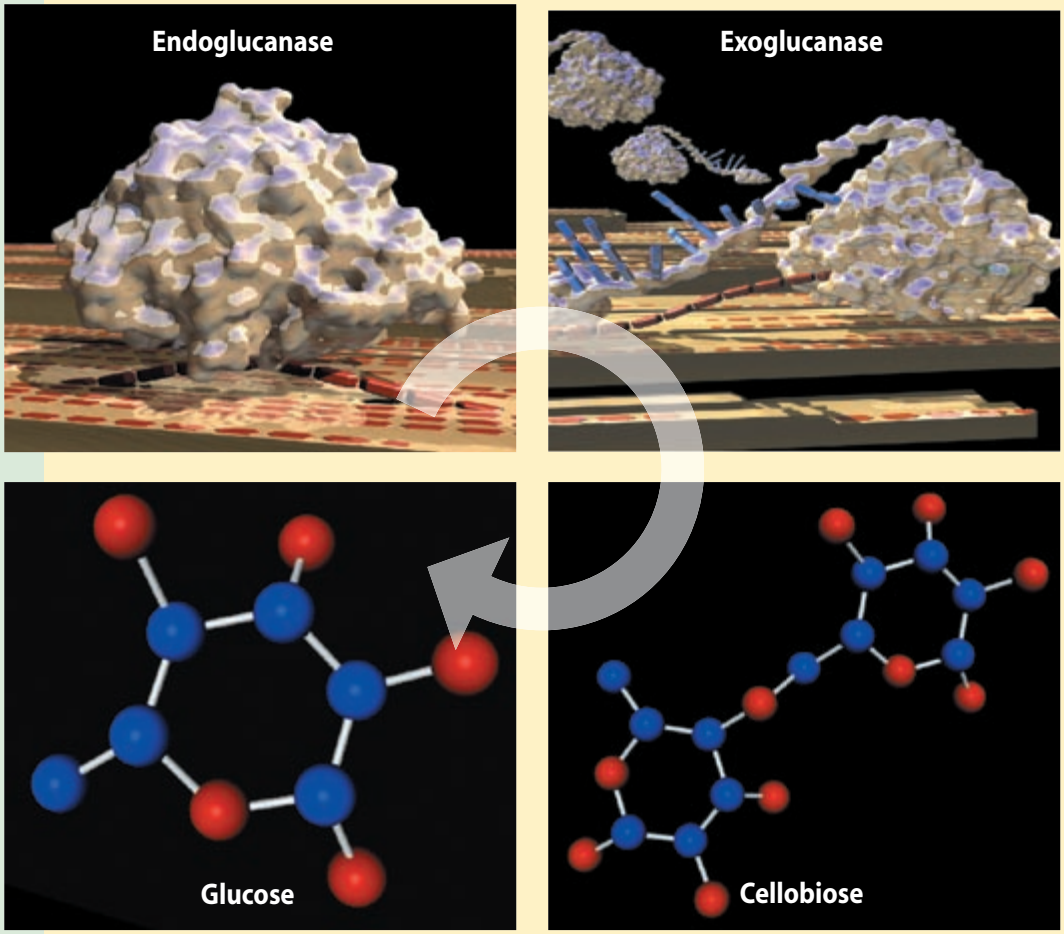
With genetic manipulation, you seek to turn genes on or off or to import genes that express a particular trait. In metabolic engineering, which NREL scientists practice to develop more effective fermentative organisms, researchers manipulate a series of different genes to give the organism the ability to digest a new food source or to produce more of a desired product such as ethanol. But NREL, Genencor, and Novozymes scientists are also going beyond these techniques. Instead, they are actually creating enzymes that never existed in nature and getting organisms to produce them.

What is a Biorefinery?

Once cellulase enzymes and bioprocessing technology are cost-effective enough to produce inexpensive sugar from cellulosic biomass, that sugar can become a "platform" chemical from which various fuels and chemicals can be made in "biorefineries." The biorefinery concept is analogous to today's petroleum refineries, which produce multiple fuels and products from petroleum. Industrial biorefineries have been identified as the most promising route to the creation of a new domestic biobased industry.

Wet-mill corn refineries now make ethanol, beverage sweeteners, and a variety of other food and animal feed products from cornstarch. Industrial giants Cargill-Dow (a joint venture between Dow Chemical and Cargill, Inc.) and DuPont are already making biomass plastics for clothing and other uses from lactic acid and 1,3 propanediol, both of which can be made from sugar or from cellulosic biomass. With lignin and hemicellulose sugars, as well as glucose from cellulose, the biorefinery will be even better positioned to make a variety of products.

The ability of an enzyme to catalyze a biochemical reaction is partly a function of the physical shape of the enzyme. In the illustration (right) of an exoglucanase — which Himmel refers to as "one of the most important proteins in the biosphere" — the "tadpole body" is the catalytic domain, and the "tail" is a binding site that "grabs onto" the cellulose and draws it into the body. The CBH I catalytic domain contains ten active subsites that physically adjoin to the cellulose and initiate the chemical reactions that break the chains apart into cellobiose. The question is, Can this nanomachine be improved?

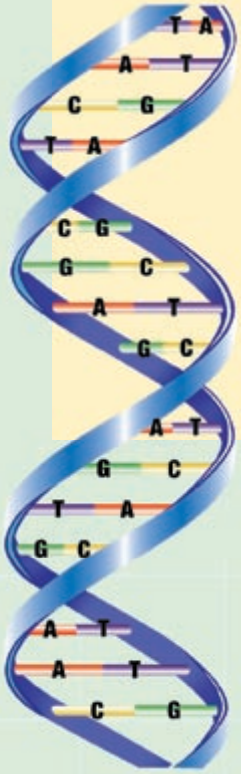


Cellulase enzymes break down cellulose to sugars in three steps. First, endoglucanase (top left) attacks a cellulose chain and severs it via hydrolysis; second, exoglucanase (top right) attaches to a cellulose chain end, works its way along the chain, and breaks off cellobiose molecules (bottom right); third, betaglucosidase splits the cellobiose into two glucose molecules.

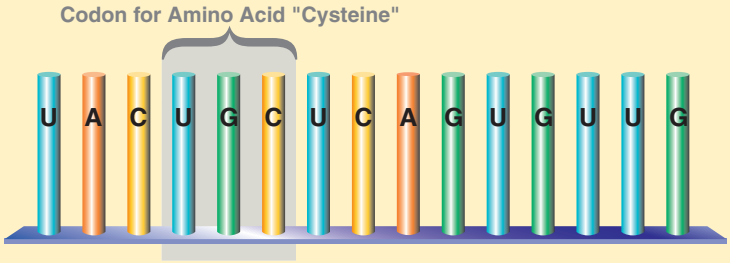
Assuming that improvement is possible, how do you get a fungi to produce a more efficient protein? To see how this may be done, it is necessary to understand how cells produce proteins. Ribonucleic acid (RNA) provides the code for producing proteins. But RNA itself is synthesized from a gene's DNA (deoxyribonucleic acid), a polymer comprised of two strands of sequences of four nucleotides connected and wound around each other in a double helix configuration (see illustration, next page). The four nucleotides are guanine (G), cytosine (C), adenine (A), and thymine (T). Nucleotides on each strand attach to their complementary nucleotides on the other strand: G connects to C and A to T. To synthesize RNA, a polymerase enzyme attaches to a DNA and travels along the double helix.

As is does so, it unzips the double helix into two strands, which zip back together as the polymerase passes along the DNA. One of the unzipped DNA strands serves as a template that the polymerase reads as it travels along, synthesizing a strand of RNA, whose sequence of nucleotides are the complements of the nucleotides read on the template. As the polymerase reads the sequence of nucleotides on the DNA template, it performs the trick of synthesizing a uracil (U) nucleotide for each adenine nucleotide it reads on the DNA (rather than producing the complementary thymine).

By producing many products, biorefineries take advantage of differences in biomass components and intermediates and maximize value derived from biomass. For example, a biorefinery might produce several low-volume, high-value chemical products and a low-value, high-volume liquid fuel, while generating electricity and process heat for its own use and perhaps enough to sell electricity. High-value products enhance profitability. High-volume fuel helps meet national energy needs. Power production reduces costs and avoids greenhouse-gas emissions.



Protein Engineering



Sequenced triplets of RNA nucleotides, or codons, provide the code for a cell's ribosomes to manufacture chains of amino acids to make proteins. RNA itself is synthesized via copying the sequence of nucleotides of one of the strands of a gene's DNA. Hence, to engineer a protein, a first step is to devise a way to induce variations during the copying of a DNA's nucleotides.

Thus, the synthesized RNA is comprised of a sequence of A's, C's, G's, and U's — such as UGC UCA GUG; the sequence of which is divided into functional triplet nucleotides known as codons. Each codon codes for one of 20 amino acids. The codon UGC, for example, codes for the amino acid cysteine, while UCA codes for serine and GUG for valine. The codons on the RNA strand are read by the cell's ribosome machinery, which are "factories" that use the information provided by the codons to produce the amino acids. The string of amino acids thus produced is the primary structure of a protein. This primary structure then twists and folds to become a functional three-dimensional protein.

By manipulating the nucleotide sequence in the gene that codes for producing a particular cellulase in the fungi, scientists can change the sequence of amino acids in the cellulase, and thus alter its effectiveness — producing a designer protein. One way to change the nucleotide sequence is to use a bacterial polymerase that has been genetically engineered to incorrectly copy the genetic sequence to be modified. Or, scientists can subject a polymerase to abnormal conditions, such as high salt concentrations, which will induce the polymerase to produce variations when copying a DNA nucleotide. Using polymerases, researchers apply several techniques to engineer proteins, such as directed evolution and site-directed mutagenesis.

Directed evolution. In directed evolution, researchers randomly alter the nucleotide sequence of the particular fungal gene that codes for the exoglucanase or other protein produced. They then grow and test cultures of each of the altered fungi and select those with more desirable traits, such as greater enzyme production or enzymes with stronger hydrolysis activity. They tackle this Herculean task of seemingly infinite possibilities by using robotic equipment that can automatically culture the possible genetic variants and screen them to identify the ones that exhibit the most desired characteristics. Researchers can then test the top candidates more thoroughly.

Site-directed mutagenesis. In contrast to the rapid-screening-of-random-changes approaches, with site-directed mutagenesis, researchers (1) build molecular models of the enzymes; (2) use computer simulations to examine the enzyme's structure and predict how it could be altered to more effectively interact with the cellulose strands; (3) identify the amino acid responsi-

NREL scientists use a robotics deck such as this to quickly culture genetic variants of fungal genes for producing enzymes and screen them to identify the variants with the most promising characteristics.

What if Biomass Was Not So Recalcitrant?

The crystalline structure of cellulose and resulting fibrous nature of biomass may be one of nature's key features for protecting plant life in the biosphere — imagine a 200-foot-tall tree or even an 8-foot cornstalk otherwise — but that is not to say there are not other possibilities. NREL researcher Steve Thomas, who has worked on improving cellulases for breaking down cellulose, talks about starting to look at the other side of the coin — altering the structure and composition of biomass. Thomas and colleagues are already examining many of the thousands of varieties of corn that seed companies have developed over the years to find ones that might be better for cellulose bioprocessing. Rather than just valuing high starch content in the kernels and tall and straight growth, cellulose bioprocessing could benefit from other traits. For example, corn plants with more cellulose in the stalk could benefit bioethanol

ble for the portion of the enzyme where change is desired; (4) identify the codon of nucleotides in the organism's genes that correspond to those amino acids in the enzyme, and; (5) set out to make the specific desired change in the DNA. At this point, they are back to inducing somewhat random changes. They know, however, that they have targeted the right site and more precisely what they are looking for in screening the variants.

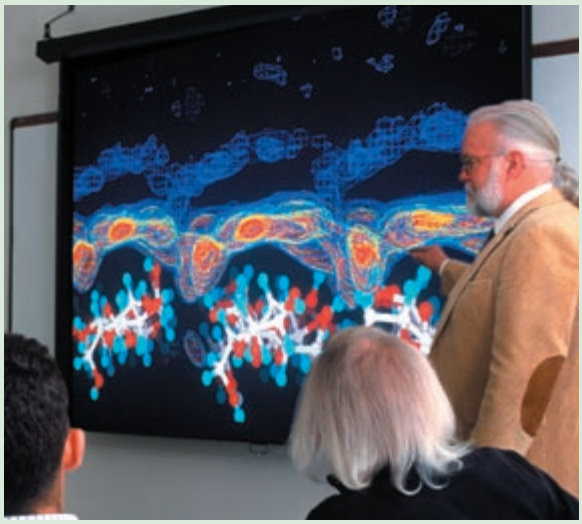
Modeling for Fundamental Understanding

Techniques such as site-directed mutagenesis require understanding the fundamentals of how cellulases function. One way to do this is through 3-D animation (see images of endoglucanase and exoglucanase, page 13), which can help researchers hypothesize how cellulases interact with cellulose. The hypotheses may then be tested to hone the picture. Another way is via modeling dynamic interactions, such as the surface layer interaction between water and cellulose, believed to be a key factor in the resistance of cellulose to hydrolysis and therefore in designing effective methods to overcome that recalcitrance. Modeling activities are invaluable in helping to understand the fundamentals of cellulase action. NREL researchers will continue to work in this vein, developing thermodynamic, mathematic, and mechanistic models of the molecular machinery of enzymatic hydrolysis.

Moving to Eureka! and a Biomass Economy

Just as understanding basic science behind cellulase activity is critical for specific applications such as site-directed mutagenesis, it is also where dramatic changes in approach could come from. Himmel expects the "Voila!s" and "Eureka!s" of enzymatic hydrolysis to come from fully understanding the way the exoglucanases work. Industry does not generally do such basic research, so DOE has asked NREL to lead in this area. NREL's work to gain understanding of this crucial research area is supported by collaborations with Cornell University, the Colorado School of Mines, the University of California at Davis, Rutgers University, the University of Arkansas, and research institutes in Japan, Israel, and Sweden.

On the other hand, NREL cannot begin to develop the real-world improvements associated with actual industrial operations that can come from enzyme producers such as Genencor and Novozymes NREL researchers are validating the achievements of the enzyme companies. And, of course, research understanding will do no good unless industry incorporates it into practice. So, overcoming the recalcitrance of biomass with economical enzyme production — which will lead to more cost-effective production of ethanol and other products from cellulosic biomass — must come from a combination of basic research at NREL and universities and from steady improvement by the enzyme industry.



One way in which researchers gain insight into how cellulases function is through the development and use of models, such as this one of the dynamic interaction at the interface between water and cellulose.

For More Information

NREL Biomass Research. "Alternative Fuels Users Facility (AFUF)." http://www.nrel.gov/biomass/workingwithus.html#afuf. Accessed January 27, 2004.

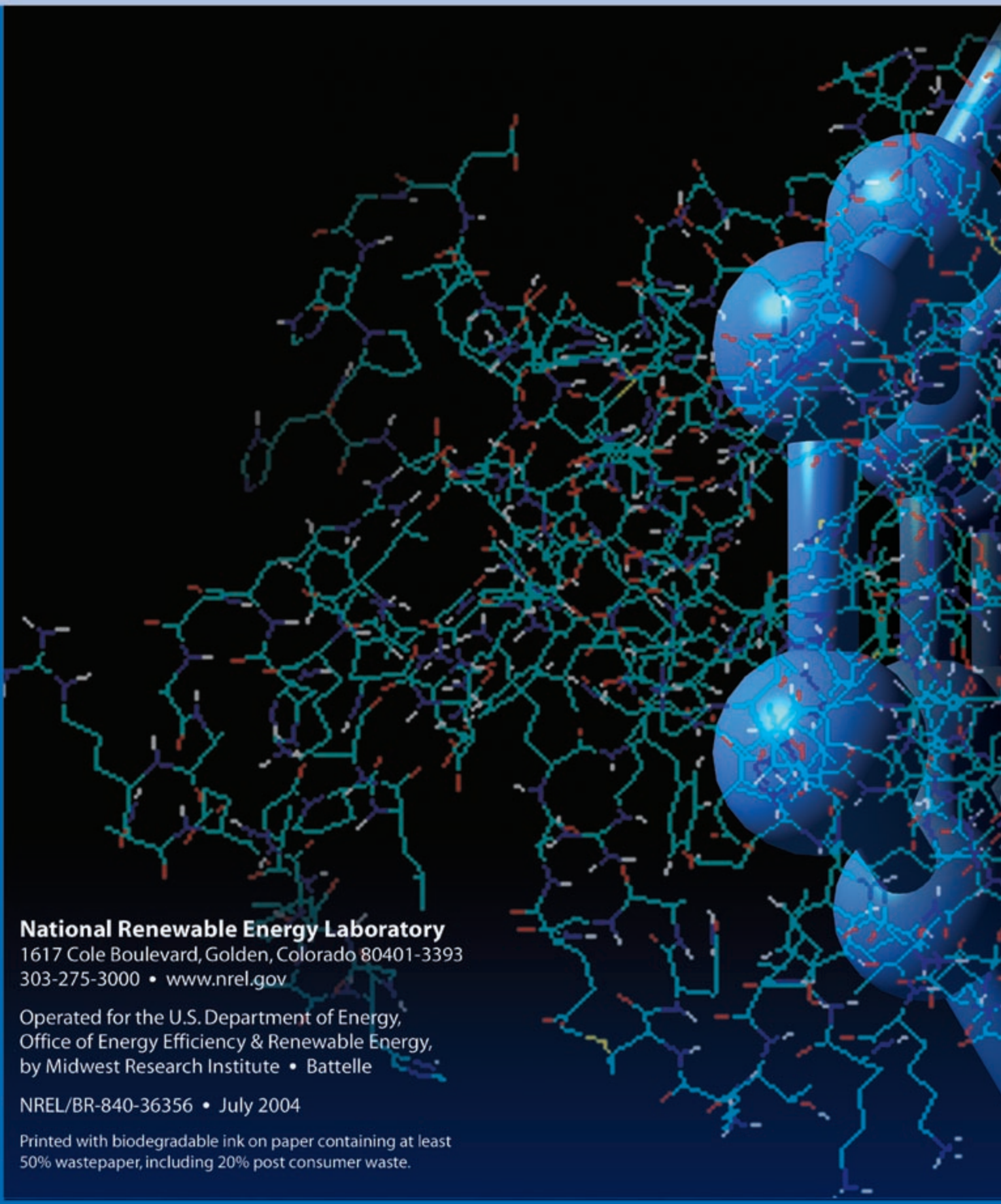
Himmel, M.E.; Baker, J.O.; Saddler, J.,eds. "Glycosyl Hydrolases for Biomass Conversion." ACS Symposium Series 769, Washington, DC: American Chemical Society, 2001: Distributed by Oxford University Press.

Sheehan, J.; Himmel, M.E. "Enzymes, Energy, and the Environment: Cellulase Development in the Emerging Bioethanol Industry." *Biotechnology Progress* (15:3), 1999; pp. 817–827.

Hames, B.R.; Thomas, S.R.; Sluiter, A.D.; Roth, C.J.; Templeton, D.W. "Rapid Biomass Analysis: New Tools for Compositional Analysis of Corn Stover Feedstocks and Process Intermediates from Ethanol Production." Davison, B.H., et al., eds. *Applied Biochemistry and Biotechnology. Proceedings of the 24th Symposium on Biotechnology for Fuels and Chemicals.* 28 April–1 May, 2002, Gatlinburg, TN. Vol 105–108, Spring 2003; pp. 5–16.

production, or ones with more lignin could benefit biomass power generation. Using NREL's rapid analysis technology for biomass and sophisticated economic models, researchers have already found that ethanol production could be more than $0.30/gallon less expensive using stover with different composition.

Going a step further, many plants produce enzymes that break down their own cell walls, with the enzymes serving functions such as ripening fruit or allowing leaves to break off in the fall. We could breed plants with high natural cellulase production, genetically engineer higher or earlier production of the cellulases, or import cellulase production genes from fungi. One possibility is plants that already start to break themselves down before they are harvested. Another possibility could be one part of a plant producing enzymes needed to break down other parts. Either would make the work of the biorefinery easier and less costly.



National Renewable Energy Laboratory
1617 Cole Boulevard, Golden, Colorado 80401-3393
303-275-3000 • www.nrel.gov

Operated for the U.S. Department of Energy,
Office of Energy Efficiency & Renewable Energy,
by Midwest Research Institute • Battelle

NREL/BR-840-36356 • July 2004

Printed with biodegradable ink on paper containing at least
50% wastepaper, including 20% post consumer waste.