Glycosyl Hydrolases for Biomass Conversion

About the Cover

Computer-generated rendering of *Trichoderma reesei* cellobiohydrolase I enzyme depicts its action on cellulose. Animation created by Pixel Kitchen in Boulder, Colorado.

Glycosyl Hydrolases for Biomass Conversion

Michael E. Himmel, EDITOR
National Renewable Energy Laboratory

John O. Baker, EDITOR
National Renewable Energy Laboratory

John N. Saddler, EDITOR
The University of British Columbia

American Chemical Society, Washington, DC

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48–1984.

Foreword

The ACS Symposium Series was first published in 1974 to provide a mechanism for publishing symposia quickly in book form. The purpose of the series is to publish timely, comprehensive books developed from ACS sponsored symposia based on current scientific research. Occasionally, books are developed from symposia sponsored by other organizations when the topic is of keen interest to the chemistry audience.

Before agreeing to publish a book, the proposed table of contents is reviewed for appropriate and comprehensive coverage and for interest to the audience. Some papers may be excluded in order to better focus the book; others may be added to provide comprehensiveness. When appropriate, overview or introductory chapters are added. Drafts of chapters are peer-reviewed prior to final acceptance or rejection, and manuscripts are prepared in camera-ready format.

As a rule, only original research papers and original review papers are included in the volumes. Verbatim reproductions of previously published papers are not accepted.

ACS BOOKS DEPARTMENT

Preface

The depolymerization of biomass polysaccharides by glycosyl hydrolases has for some time been recognized as *the* critically important step in the conversion of biomass components into fermentable feedstocks for production of fuels and chemicals. For this reason, such enzymes have become the focus of studies ranging from the most fundamental, "purely scientific" investigations of the molecular and catalytic properties of individual enzyme molecules themselves, to strictly applied experiments at the pilot-plant or near-pilot-plant scale. The particular focus of this volume is on studies that have been carried out at what might be termed "large bench-scale" in order to address the ultimate application of the enzymes in industrial processes. Five of the chapters deal directly with the use of mixtures of glycosyl hydrolases in the saccharification of cellulose- and mannan-containing biomass. Topics include adsorption to macromolecular substrates of specific enzymes in a mixture, the effects on enzyme efficiency of producing enzyme mixtures in the presence of the target biomass, factors influencing enzymatic hydrolysis of lignocellulosics, and a study of enzymes acting on soluble and insoluble mannans. Two additional chapters describe explorations of solid-substrate fermentations and expression in transgenic plants as means of producing glycosyl hydrolases at an industrial scale.

In addition to the applications-oriented studies described above, the volume also includes accounts from the more fundamental and exploratory regions of the glycosyl hydrolase research spectrum. Direct structure-function studies address the roles of four conserved aspartate residues in *Thermomonospora fusca* endoglucanase E2, and the roles of four apparently distinct functional domains in a single molecule of a bifunctional (endoglucanase/exoglucanase) cellulase from *Teredinobacter turnerae*. Other chapters describe newly discovered enzymes (thermostable cellulases and xylanases from *Thermoascus aurantiacus* and alkalo-philic dextranases from *Streptomyces anulatus*) that may extend the limits of use-able reaction conditions. In another chapter, molecular mechanics calculations have been applied to the effect on substrate binding of a point mutation in E1 from *Acidothermus cellulolyticus* and to the generalized interaction between a glucose molecule and an alanine side-chain. Both the applied and fundamental studies are put in perspective relative to the entire bioconversion process by the opening chapter's comprehensive strategic analysis of the U.S. Department of Energy's National Ethanol Program.

This book includes several chapters describing "rational design" approaches to enhancing enzyme performance, and it is apparent that the application of random mutation strategies to glycosyl hydrolases will also be the subject of future studies. Site-directed-mutagenesis (SDM) is an informational approach to protein engineering and relies on high-resolution crystallographic structures of target proteins and some stratagem for specific amino acid changes. Resurgence in SDM

technology has followed the recent advent of computational methods for identifying these site-specific changes for a variety of protein engineering objectives. Non-informational mutagenesis techniques (referred to generically as "directed evolution"), in conjunction with high-throughput screening, allows testing of statistically meaningful variations in protein conformation. Directed evolution technology has undergone significant refinement from initial error-prone PCR methodology and now includes Gene Shuffling, site-saturation mutagenesis, staggered extension process (StEP) technology, and others. In our opinion, the primary challenge in the application of directed evolution technology to cellulase improvement lies almost entirely in (1) adaptation of robotic screening methods to accurately select transformed host cells (clones) producing enzymes displaying enhanced performance on microcrystalline cellulose, and (2) the development of hosts able to express active and industrially relevant recombinant enzymes from shuffled or otherwise modified eukaryotic genes. One wonders where promising new directed evolution technologies can and will take us. Efforts have been made to predict the probability of successful engineering by estimating the proportion of total protein sequence space already visited, and perhaps rejected, in nature. Certainly, many protein forms not selected, or even encountered, in natural evolutionary processes are of interest to humankind. However, where do actual structural boundaries exist—those imposed by laws of thermodynamics, for example—and can we learn to recognize situations lying outside of these natural laws a priori? In the case of cellulase engineering, perhaps the low specific activities of known cellulases reflect more the inherent difficulty of the protein/ substrate boundary, than problems of pure protein design. Improving cellulase efficiency, then, may only come with directed *co*-evolution of (or modification to) *both* enzyme and substrate (which may then not be cellulose as we know it). This is an experiment nature has probably had no incentive to conduct, and may thus represent a completely unexplored block of protein sequence space!

MICHAEL E. HIMMEL
National Renewable Energy Laboratory
1617 Cole Boulevard
Golden, CO 80401

JOHN O. BAKER
Biotechnology Center for Fuels and Chemicals
National Renewable Energy Laboratory
Golden, CO 80401

JOHN N. SADDLER
Department of Wood Science
The University of British Columbia
4th Floor, Forest Sciences Centre
4042–2424 Mail Mall
Vancouver, British Columbia V6T I24, Canada

x

Chapter 1

The Road to Bioethanol: A Strategic Perspective of the U.S. Department of Energy's National Ethanol Program

John Sheehan

**Biotechnology Center for Fuels and Chemicals,
National Renewable Energy Laboratory, Golden, CO 80401**

As the Bioethanol Program at the Department of Energy (DOE) nears the end of two decades of research, it is time to take a hard look at where we have been and where we are going. This paper summarizes the status of bioethanol technology today and what we see as the future directions for research and development. All of this is placed in the perspective of strategic national issues that represent the drivers for our program—the environment, the economy, energy security and sustainability. The key technology pathways include the use of new tools for protein engineering and directed evolution of enzymes and organisms, as well as new approaches to physical/chemical pretreatment of biomass.

Ethanol is used today as an alternative fuel, a fuel extender, an oxygenate and an octane enhancer. From just over 10 million gallons of production in 1979, the U.S. fuel ethanol industry has grown to more than 1.8 billion gallons of annual production capacity (*1*). Almost all of this capacity is based on technology that converts the starch contained in corn to sugars, which are then fermented to ethanol.

From its first days, this industry has been looking for ways to expand the available resource base to include many other forms of biomass. The U.S. Department of Energy has, throughout this period, invested in research and development on technology that will allow the fuel ethanol industry to achieve its goal of expanded production using a diversified supply of biomass feedstocks.

We refer to ethanol made from these as-yet untapped biomass resources as "bioethanol." This paper provides a strategic perspective on this new bioethanol technology.

© 2001 American Chemical Society

Strategic Issues

There are several major strategic issues that motivate and influence DOE's research program for bioethanol. These include:
- national security,
- the environment, and
- the marketplace.

Though each of these issues has shifted in importance over the years, all three remain consistent drivers for our plans. Let me touch on each of these issues briefly.

National Security.

Oil Supply. A recent *Science* article summarized the strategic situation with regard to oil supply this way:

> "Nature took half a billion years to create the world's oil, but observers agree that humankind will consume it all in a 2-century binge of profligate energy use."(*2*)

Our dependence has been growing at an alarming rate since the early 1980s, ironically a time when public concern about petroleum has been very low. DOE's Energy Information Administration paints a dismal picture of our growing dependence on foreign oil (*3*). Consider these basic points:
1. Petroleum demand is increasing, especially due to new demand from Asian markets
2. New oil will come primarily from the Persian Gulf
3. As long as prices for petroleum remain low, we can expect our imports to exceed 60% ten years from now
4. U.S. domestic supplies will likewise remain low as long as prices for petroleum remain low

Not everyone shares this view of the future, or sees it as a reason for concern. The American Petroleum Institute does not see foreign imports as a matter of national security (*4*). Others have argued that the prediction of increasing Mideast oil dependence worldwide is wrong (*5*). Nevertheless. the International Energy Agency (IEA) recently announced that it sees annual petroleum supplies reaching a peak some time between 2010 and 2020. The IEA is one more voice in a growing chorus of concern about the imminent danger of shrinking oil supplies (*2*). While some disagree with this pessimistic prediction, concern about our foreign oil addiction is widely held by a broad range of political and commercial perspectives (*6*).

While there may be uncertainty and even contention over when and if there is a national security issue, there is one more piece to the puzzle that influences our perspective on this issue. Put quite simply, 98% of the energy consumed in the U.S. transportation sector comes from petroleum (mostly in the form of gasoline and diesel fuel). The implication of this indisputable observation is that even minor hiccups in the supply of oil could have crippling effects on our nation. This lends special

significance to the Bioethanol Program as a means of diversifying the fuel base in our transportation sector.

Energy Diversity. An important corollary to the notion of increasing energy security is the concept of energy diversity. Today, in the U.S., natural gas, propane, and biodiesel are establishing a place in the transportation fuel market. Bioethanol is yet another option in the fuel mix that we seek to provide. J.S. Jennings, the Chairman of Royal Dutch Shell, a company recognized as a leading strategic thinker in the energy industry, has stated that "...the only prudent energy policy is one of diversity and flexibility" (7).

Economic Security. Our view of national security today must include questions about the health and robustness of our economy. Energy today plays an essential role in our economy. Petroleum imports represent 20% of our growing trade deficit. This cannot help but have an impact on our economy. A diverse portfolio of fuels, including bioethanol, would bring money and jobs back into the U.S. economy built on this new renewable energy technology. The associated development of energy crops will likewise provide a needed boost to our agricultural sector, a mainstay of the U.S. economy.

The Environment.

Air Pollution. A life cycle study conducted by DOE in 1993 evaluated the overall impact of bioethanol on several key regulated pollutants targeted by the Clean Air Act Amendments of 1990 (1990 CAAA) (8). This study found that, compared with reformulated gasoline (RFG), a 95% ethanol/5% gasoline blend (E95) reduced sulfur oxide emissions by 60 to 80%. Volatile organic emissions from E95-fueled vehicles are 13 to 15% lower. Net (life cycle) emissions of NOx and carbon monoxide are essentially the same.

These results are encouraging, but of greater importance is the impact that bioethanol has directly on tailpipe emissions (as opposed to net pollutant levels across the life cycle of the fuel). Low blends of ethanol have some peculiar emission problems that go away at higher blend rates (mostly due to Reid vapor pressure increases that occur between 10% and 20% volume blends). A survey of the available emissions data for high blends of ethanol reveals that, while there is a fair amount of data, it is often not consistently obtained. Still, the survey found the following broad trends for ethanol used in high blend levels with gasoline: (9)

- CO levels may decrease as much as 20%, probably because of the oxygen content of ethanol
- Similar decreases in NOx can be anticipated as well.
- High blends of ethanol cut end-use emissions of volatile organic carbon (VOCs) by 30%.
- Aldehyde emissions from ethanol combustion in spark-ignited engines are, however, substantially higher for ethanol.

The first round of comprehensive emissions tests for flexible fueled ethanol vehicles used in federal fleets was completed in 1996. These tests included a

comparison of 21 ethanol-fueled Chevrolet Luminas with an equal number of standard gasoline model Luminas (*10*). The results of the extensive study of exhaust emissions confirm the trends seen across the literature (see Figure 1).
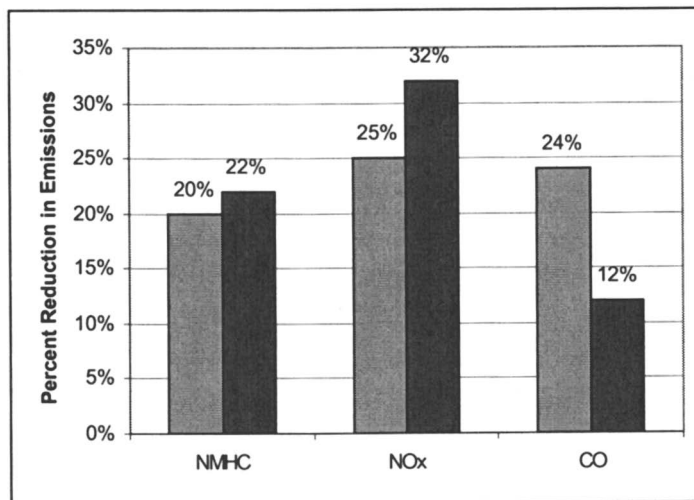
Figure 1: *Emission Reductions for E85-fueled Federal Fleet Vehicles. The two sets of data represent analytical results from two independent laboratories. (NMHC = non-methane hydrocarbons)*

Sustainable Development. Public concern about the quality of our environment has grown steadily over the past decade (*11*). Vice President Al Gore posits an environmental crisis that has been brought on by an exploding world population, a technology revolution that has led to over-exploitation of our natural resources and an apparent disregard for the future. He cites the 1992 "Earth Summit" in Rio de Janeiro as a major turning point in our thinking about the environment.

World-renowned naturalist Edward O. Wilson echoes these sentiments in his call for technology development that moves us away from fossil fuels and reduces the energy intensity of our economy. Wilson describes very eloquently his notion of an ethic of sustainability:

> "The common aim must be to expand resources and improve quality of life for as many people as heedless population growth forces upon Earth, and do it with minimal prosthetic dependence. That, in essence, is the ethic of sustainable development." (*12*)

Bioethanol technology represents just one approach to moving our economy to a more sustainable basis. We, like many others touting technological solutions, should heed his remonstration of over-dependence on what he calls "environmental prostheses" that will extend the capacity of our planet, but will not eliminate the risk

of environmental catastrophe. Environmentalists and technologists must work together to provide balance and reason in our approach.

The biggest impediment to sustainable development is our economic system, which places no value on the environment or on the future. "The hard truth," writes Al Gore," is that our economic system is partially blind" (*13*). The blindness of the marketplace to environmental issues makes deployment of bioethanol technology problematic, but not impossible. It forces a discipline on our development efforts in which we seek out opportunities for bioethanol that meet multiple needs. Still, it is clear that something must change in our economic calculus if renewable and sustainable technologies are to take hold before a crisis forces the issue.

Climate Change. Climate change is a particular example of the kind of risks that are involved in ignoring the "ethic" of sustainable development. Political and public concern about climate change varies with the time of day and day of the week. A year with El Niño certainly promotes the cause. One reason for the seemingly arbitrary nature of our views on climate change is that it involves a discussion of relative risks, rather than explicit cause-and-effect problems. The reason for this is simple: understanding the climatic implications of global warming is *not* simple. Some have even suggested that we can never understand the complex interaction of variables involved in understanding our climate (*14*). The salvos continue to go back and forth among the scientific experts as to the degree of warming that has occurred and its impact (*15, 16*). For example, many critics of climate change claim that satellite data on global temperature contradict claims of increased temperature over the past decade. Researchers have recently demonstrated that decreasing temperature trends seen in satellite data are actually due to errors caused by not accounting for changing altitude of the satellite. When corrected for this change, the satellite data is consistent with other surface temperature measurements showing an increase in average temperature (*17*).

What the policymakers and the public need to do is to make some rational choices about risk. The research reported in 1957 that confirmed CO_2 accumulation in the atmosphere couched the question of climate change in exactly these terms (*18*), and there is still no better way to look at the problem. Given the catastrophic nature of the implications related to climate change, how much risk is too much? The potential risk associated with climate change has gotten the attention of the insurance industry, a group all too familiar with the damage and expense that could be involved (*19*). E.O. Wilson's take on the kind of risk associated with our environment is along similar lines:

> "In ecology, as in medicine, a false positive diagnosis is an inconvenience, but a false negative diagnosis can be catastrophic. That is why ecologists and doctors don't like to gamble at all, and if they must, it is always on the side of caution. It is a mistake to dismiss a worried ecologist or a worried doctor as an alarmist." (*12*)

In other words, can we afford a false negative diagnosis regarding climate change? Technologies like bioethanol are insurance. Prudence dictates that we take some forward movement in encouraging the use of such sustainable technologies.

The current political setting for discussing climate change frames the question as an all or nothing proposition. Either climate change is a real problem or it is not. If it is real, then we should treat it as a "crisis"; otherwise, we are wasting our time. The Kyoto agreement signed by representatives of countries from around the world is doomed to fail if we continue to view the issue in this ill-conceived framework. A group of prominent energy and environmental leaders recently met at the highly respected Aspen Institute to address the issue of climate change. In a letter to the White House, they urged the Clinton administration not to send the Kyoto agreement to Congress, where it will too readily be dismissed (20). Instead, they suggest that the U.S. take a leadership role in establishing a *long-term* strategy for dealing with climate change. "Climate change," they wrote, "is a long term problem, and the focus should be on achieving sustainable levels of greenhouse gas concentrations at the least cost, not only on near-term emission reductions." This approach recognizes climate change as a question of risk rather than a black and white problem that must be dealt with using Draconian measures. In the end, renewable energy options like bioethanol benefit from this type of longer-term strategy. Reasonable and sustained support is what is needed if bioethanol is to play a part in our energy future.

The Market. The bottom line for bioethanol is what, if any, market opportunities exist for this fuel. It can be used as a fuel additive or extender in blends of around 10%, or it can be used as a fuel substitute. In today's U.S. fuel market, ethanol can be used in flexible fuel vehicles that can operate using blends of 85% ethanol (and 15% gasoline).

Alternative Fuels Market. For a long time, the greatest impediment to ethanol's use as an alternative fuel was the lack of ethanol-compatible vehicles in the U.S. This has changed dramatically. Today, both Ford and Chrysler offer standard models designed to run on either 85% ethanol (E85) or gasoline. They are offering this fuel flexibility at no additional cost to the consumer (21, 22). While the availability of vehicles is no longer an issue, there is still a paucity of fuel stations and fuel distribution infrastructure for E85. Today, 45 publicly available E85 stations are available in the U.S. Thirty more limited access stations are available (23). The lack of basic infrastructure and the higher price of ethanol versus gasoline are major constraints on this market.

Fuel Additive Market. Use of ethanol as an additive in gasoline has become a major market. Starting from literally nothing a little over 20 years ago, ethanol as a fuel additive has become a billion gallon per year market. It has value as an oxygenate in "CO nonattainment" markets, and as a fuel extender and octane booster. The value of ethanol in the oxygenate and octane booster market is around 70 to 80 cents per gallon.

Ethanol Selling Price and Tax Incentives. Passage of 1998's overhaul of the highway bill brought with it an extension of the ethanol tax incentive program. This program adds about 50 cents per gallon to the value of ethanol sold in the fuel market. When added on top of the market value for ethanol as an oxygenate and an octane booster, this tax incentive allows ethanol to sell on the market for around $1.20 to $1.40 per gallon. The ethanol tax incentive will remain in place through 2007. Without continued authorization from Congress, this incentive will go away. A major strategy of the Bioethanol Program is to take advantage of this tax incentive by developing near term technology that can compete in the current ethanol market. In the meantime, our research is geared toward achieving cost reductions that will eliminate the need for further extensions of the tax incentive.

The Technology Today

Our working definition of biomass is plant matter produced via photosynthetic uptake of carbon from atmospheric CO_2. It is important to understand this definition. The photosynthetic uptake of carbon imparts many of the benefits of biomass-derived fuels, such as sustainability and greenhouse gas reductions (24). The Bioethanol Program is, more specifically, concerned with the conversion of carbon present as sugars in biomass to fuel ethanol.

At the risk of oversimplifying the Bioethanol story, we prefer to view ethanol technology in terms of only four basic steps (see Figure 2). Production of biomass results in the fixing of atmospheric carbon dioxide into organic carbon. Conversion of this biomass to a useable fermentation feedstock (typically some form of sugar) can be achieved using a variety of different process technologies. These processes for sugar production constitute the critical differences among all of the ethanol technology options. Using biocatalysts (microorganisms including yeast and bacteria) to ferment the sugars released from biomass to produce ethanol in a relatively dilute aqueous solution is probably the oldest form of biotechnology developed by humankind. This dilute solution can be processed to yield ethanol that meets fuel-grade specifications. Finally, the economics of biomass utilization demands that any unfermented residual material left over after ethanol production must be used, as well.
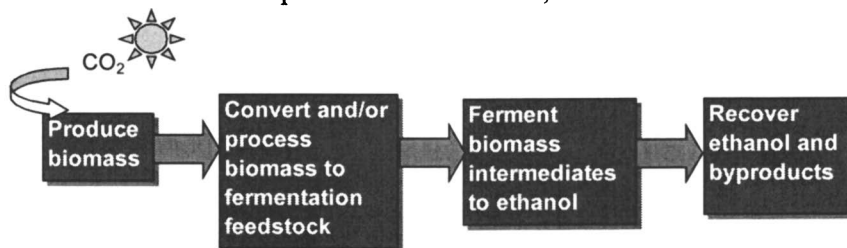

Figure 2. *General scheme for converting biomass to ethanol*

The Nature Of Sugars In Biomass. The degree of complexity and feasibility of biomass conversion technology depends on the nature of the feedstock from which we

start. The least complicated approach to fuel ethanol production is to use biomass that contains monomeric sugars, which can be fermented directly to ethanol. Sugarcane and sugar beets are examples of biomass that contain substantial amounts of monomeric sugars. Up until the 1930s, industrial grade ethanol was produced in the United States via fermentation of molasses derived from such sugar crops (*25*). The high cost of sugar from these crops has made these sources prohibitively expensive in the United States (*26, 27*).

Sugars are more commonly found in the form of biopolymers that must be chemically processed to yield simple sugars. In the United States, today's fuel ethanol is derived almost entirely from the starch (a biopolymer of glucose) contained in corn. Starch consists of glucose molecules strung together by α-glycosidic linkages. These linkages occur in chains of α-1,4 linkages with branches formed as a result of α-1,6 linkages (see Figure 3).

The terms α and β are used to describe different stereoisomers of glucose. A not-so-obvious consequence of the α linkages in starch is that this polymer is highly amorphous, making it more readily attacked by human and animal enzyme systems. The ability to commercially produce sugars from starch is the result of one of the earliest examples of modern industrial enzyme technology—the production and use of α-amylases, glucoamylases and glucose isomerase in starch processing (*28*). Researchers have long hoped to emulate the success of this industry in the conversion of cellulosic biomass to sugar (*29*).
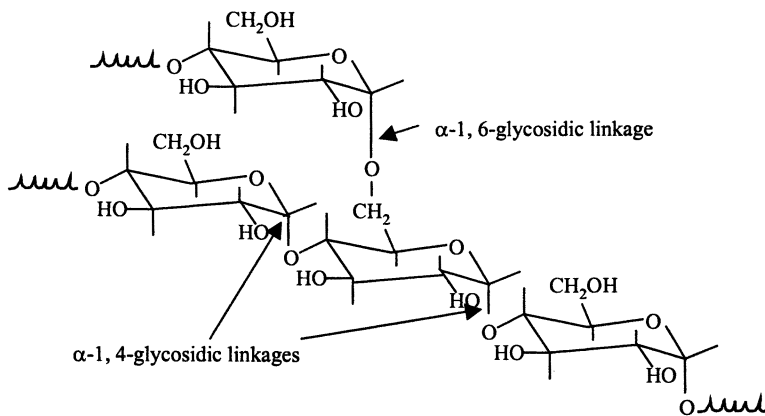

Figure 3. The polymeric structure of glucose in starch tends to be amorphous

Cellulose, the most common form of carbon in biomass, is also a biopolymer of glucose. In this case, the glucose moieties are strung together by β-glycosidic linkages. The β-linkages in cellulose form linear chains that are highly stable and much more resistant to chemical attack because of the high degree of hydrogen bonding that can occur between chains of cellulose (see Figure 4). Hydrogen bonding between cellulose chains makes the polymers more rigid, inhibiting the flexing of the molecules that must occur in the hydrolytic breaking of the glycosidic linkages.
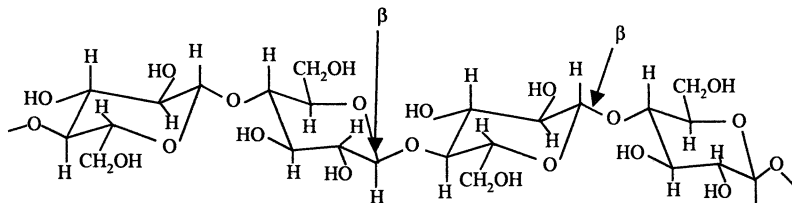
Figure 4. Linear chains of glucose linked by β-glycosidic bonds comprise cellulose

Yet a third form of sugar polymers found in biomass is hemicellulose. Hemicellulose consists of short, highly branched, chains of sugars. It contains five carbon sugars (usually D-xylose and L-arabinose) and six carbon sugars (D-galactose, D-glucose and D-mannose) and uronic acid. The sugars are highly substituted with acetic acid. Its branched nature renders hemicellulose amorphous and relatively easy to hydrolyze to its constituent sugars. When hydrolyzed, the hemicellulose from hardwoods releases products high in xylose (a five-carbon sugar). The hemicellulose contained in softwoods, by contrast, yields more six carbon sugars (*30*).

The four forms of sugar in biomass represent a range of accessibility that is reflected in the history of ethanol production. Simple sugars are the oldest and easiest to use feedstock for fermentation to ethanol. Next comes starch, now the preferred choice of feedstock for fuel ethanol. Starch-containing grain crops, like sugar crops, have higher value for food and feed applications. Because many animals (including humans) can digest starch, but not cellulose, starch will likely continue to serve a unique and important role in agriculture (*31*). The remaining two forms—cellulose and hemicellulose—are the most prevalent forms of carbon in nature, and yet they are also the most difficult to utilize. Cellulose's crystalline structure renders it highly insoluble and resistant to attack, while hemicellulose contains some sugars that have not, until recently, been readily fermentable to alcohol.

Three Technology Platforms. As indicated earlier, the technology pathways pursued in the Bioethanol Program differ primarily in the approach used to produce sugars from biomass (step 2 in Figure 2). Regarding sugar recovery, releasing the sugars from the biopolymers in plant matter involves hydrolysis of the linkages between the sugar moieties. Hydrolysis is a simple chemical reaction in which a water molecule is added across the glycosidic linkages in order to break the bonds. The discovery of sugar production by acid hydrolysis of cellulose dates back to 1819 (*32, 33*). By 1898, a German researcher had already attempted to use this chemistry in a commercial process for producing sugars from wood. This early process included fermentation of the sugars to ethanol (*34*). In the one hundred years since then, researchers have continued to pursue different approaches to achieving high yields of fermentable sugars from the acid hydrolysis of biomass. It is easy to lose this historical perspective on acid hydrolysis technologies.

The Bioethanol Program supports development of three technologies based on different approaches to producing sugars:

- Low Temperature, Concentrated Acid Hydrolysis
- High Temperature, Dilute Acid Hydrolysis
- Enzymatic Hydrolysis.

The two acid hydrolysis technology platforms have the longest history of development, while the use of enzymes to produce sugars from biomass is, in the scheme of things, a relatively recent concept.

Concentrated Acid Hydrolysis Process. The concentrated acid process for producing sugars and ethanol from lignocellulosic biomass has a long history. The ability to dissolve and hydrolyze native cellulose in cotton using concentrated sulfuric acid followed by dilution with water was reported in the literature as early as 1883 (*35*). The concentrated acid disrupts the hydrogen bonding between cellulose chains, converting it to a completely amorphous state. Once the cellulose has been decrystallized, it is extremely susceptible to hydrolysis at this point. Thus, dilution with water at modest temperatures provides complete and rapid hydrolysis to glucose, with little degradation.

It seems as though most of the research on concentrated acid processes has been done using agricultural residues, particularly corncobs. In 1918, researchers at USDA proposed a process scheme for production of sugars and other products from corncobs based on a two stage process. These researchers introduced the idea of using dilute acid pretreatment of the biomass to remove hemicellulose before decrystallization and hydrolysis of the cellulose fraction (*36*). The ability to isolate hemicellulosic sugars from cellulosic sugars was an important improvement to the process, because the five carbon sugars were not fermentable.

In 1937, the Germans built and operated commercial concentrated acid hydrolysis plants based on the use and recovery of hydrochloric acid. Several such facilities were successfully operated. During World War II, researchers at USDA's Northern Regional Research Laboratory in Peoria, Illinois further refined the concentrated sulfuric acid process for corncobs (*37*). They conducted process development studies on a continuous process that produced a 15-20% xylose sugar stream and a 10-12% glucose sugar stream, with the lignin residue remaining as a byproduct. The glucose was readily fermented to ethanol at 85-90% of theoretical yield. The Japanese developed a concentrated sulfuric acid process that was commercialized in 1948. The remarkable feature of their process was the use of membranes to separate the sugar and acid in the product stream. The membrane separation, a technology that was way ahead of its time, achieved 80% recovery of acid (*38*). Research and development based on the concentrated sulfuric acid process studied by USDA (and which came to be known as the "Peoria Process") picked up again in the United States in the 1980s, particularly at Purdue University (*39*) and at TVA (*40*). Among the improvements added by these researchers were: 1) recycling of dilute acid from the hydrolysis step for pretreatment, and 2) improved recycling of sulfuric acid. Minimizing the use of sulfuric acid and recycling the acid cost effectively are critical factors in the economic feasibility of the process.

Commercial success in the past was tied to times of national crisis, when economic competitiveness of ethanol production could be ignored. Conventional

wisdom in the literature suggests that this process cannot be economical because of the high volumes of acid required (*41*).

Today, despite that "wisdom", two companies in the U.S. are working with DOE and NREL to commercialize this technology by taking advantage of niche opportunities involving the use of biomass as a means of mitigating waste disposal or other environmental problems. Arkenol, a company which holds a series of patents on the use of concentrated acid to produce ethanol, is currently working with DOE to establish a commercial facility that will convert rice straw to ethanol. Arkenol plans to take advantage of opportunities for obtaining rice straw in the face of new regulations that would restrict the current practice of open field burning of rice straw. The economics of this opportunity are driven by the availability of a cheap feedstock that poses a disposal problem. Arkenol's technology further improves the economics of raw straw conversion by allowing for the recovery and purification of silica present in the straw. The facility would be located in Sacramento County (*42*).

Masada, a company which holds several patents related to MSW (municipal solid waste)-to-ethanol conversion, is working with DOE to contruct a MSW-to-ethanol plant, which will be located in Orange County, NY. The plant will process the lignocellulosic fraction of municipal solid waste into ethanol using technology based on TVA's concentrated sulfuric acid process. Concentrated acid hydrolysis produces high yields of sugar with little decomposition. The robustness of this process makes it well suited to complex and highly variable feedstocks like municipal solid waste. Masada's New York project takes advantage of relatively high tipping fees available in the area for collection and disposal of municipal solid waste. Masada is finalizing engineering and project financing, and expects to break ground on the plant in the year 2000.

Dilute Sulfuric Acid Process. Dilute acid hydrolysis of biomass is, by far, the oldest technology for converting biomass to ethanol. As indicated earlier, the first attempt at commercializing a process for ethanol from wood was done in Germany in 1898. It involved the use of dilute acid to hydrolyze the cellulose to glucose, and was able to produce 7.6 liters of ethanol per 100 kg of wood waste (18 gal per ton). The Germans soon developed an industrial process optimized for yields of around 50 gallons per ton of biomass. This process soon found its way to the United States, culminating in two commercial plants operating in the southeast during World War I. These plants used what was called "the American Process"—a one stage dilute sulfuric acid hydrolysis. Though the yields were half that of the original German process (25 gallons of ethanol per ton versus 50), the productivity of the American process was much higher. A drop in lumber production forced the plants to close shortly after the end of World War I (*43*). In the meantime, a small, but steady amount of research on dilute acid hydrolysis continued at the USDA's Forest Products Laboratory.

In 1932, the Germans developed an improved "percolation" process using dilute sulfuric acid, known as the "Scholler Process." These reactors were simple systems in which a dilute solution of sulfuric acid was pumped through a bed of wood chips. Several years into World War II, the U.S. found itself facing shortages of ethanol and

sugar crops. The U.S. War Production Board reinvigorated research on wood-to-ethanol as an "insurance" measure against future worsening shortages, and even funded construction of a plant in Springfield, Oregon. The board directed the Forest Products lab to look at improvements in the Scholler Process (44). Their work resulted in the "Madison Wood Sugar" process, which showed substantial improvements in productivity and yield over its German predecessor (45). Problems with start up of the Oregon plant prompted additional process development work on the Madison process at TVA's Wilson Dam facility. Their pilot plant studies further refined the process by increasing yield and simplifying mechanical aspects of the process (46). The dilute acid hydrolysis percolation reactor, culminating in the design developed in 1952, is still one of the simplest and most effective means of producing sugars from biomass. It is a benchmark against which we often compare our new ideas. In fact, such systems are still operating in Russia.

In the late 1970s, a renewed interest in this technology took hold in the U.S. because of the petroleum shortages experienced in that decade. Modeling and experimental studies on dilute hydrolysis systems were carried out during the first half of the 1980s. DOE and USDA sponsored much of this work.

After a century of research and development, dilute acid hydrolysis has evolved into a process in which hydrolysis occurs in two stages to accommodate the differences between hemicellulose and cellulose (47). The first stage can be operated under milder conditions, which maximize yield from the more readily hydrolyzed hemicellulose. The second stage is optimized for hydrolysis of the more resistant cellulose fraction. The liquid hydrolyzates are recovered from each stage and fermented to alcohol. Residual cellulose and lignin left over in the solids from hydrolysis reactors serve as boiler fuel for electricity and steam production.

While a variety of reactor designs have been evaluated, the percolation reactors originally developed at the turn of the century are still the most reliable. Though more limited in yield than the percolation reactor, continuous cocurrent pulping reactors have been proven at industrial scale (48). NREL recently reported results for a dilute acid hydrolysis of softwoods in which the conditions of the reactors were as follows:

- Stage 1: 0.7% sulfuric acid, 190°C, and a 3 minute residence time
- Stage 2: 0.4% sulfuric acid, 215°C, and a 3 minute residence time

These bench scale tests confirmed the potential to achieve yields of 89% for mannose, 82% for galactose and 50% for glucose. Fermentation with *Saccharomyces cerevisiae* achieved ethanol conversion of 90% of the theoretical yield (49).

BC International (BCI) and the DOE's Office of Fuels Development have formed a cost-shared partnership to develop a biomass-to-ethanol plant based on dilute acid technology. The facility will initially produce 20 million gallons per year of ethanol. BCI will utilize an existing ethanol plant located in Jennings, Louisiana. Dilute acid hydrolysis will be used to recover sugar from bagasse, the waste left over after sugar cane processing. A proprietary, genetically engineered organism will ferment the sugars from bagasse to ethanol (50, 51).

Enzymatic Hydrolysis Process. Enzymes are the relative newcomers with respect to biomass-to-ethanol processing. While the chemistry of sugar production

from wood has almost two centuries of research and development history and a hundred years of process development, enzymes for biomass hydrolysis can barely speak of fifty years of serious effort. The search for biological causes of cellulose hydrolysis did not begin in earnest until World War II. The U.S. Army mounted a basic research program to understand the causes of deterioration of military clothing and equipment in the jungles of the South Pacific—a problem that was wrecking havoc with cargo shipments during the war. This campaign resulted in the formation of the U.S. Army Natick Laboratories (*52*). Out of this effort to screen thousands of samples collected from the jungle came the identification of what has become one of the most important organisms in the development of cellulase enzymes—*Trichoderma viride* (eventually renamed *Trichoderma reesei*). *T. reesei* is the ancestor of many of the most potent cellulase enzyme-producing fungi in commercial use today.

Ironically, the research on cellulases was prompted by a need to prevent their hydrolytic attack on cellulose. Today, we turn to these enzymes in hope of increasing their hydrolytic power. This turning point in the focus of cellulase research did not occur until the early 1960s, when sugars from cellulose were recognized as a possible food source (*53*), echoing similar notions expressed by researchers in earlier days on acid hydrolysis research (*54*). In the mid-1960s, the discovery that extracellular enzyme preparations could be made from the likes of *T. reesei* (*55*) accelerated scientific and commercial interest in cellulases. In 1973, the army was beginning to look at cellulases as a means of converting solid waste into food and energy products (*56*). In a keynote address at a major symposium on cellulases, the Honorable Norman R. Augustine, then Under Secretary of the Army, spoke with vision about the potential impact that these enzymes could have on our society (*52*):

> "As the army's development of "ENIAC" proved to be the stimulus for the worldwide computer industry, I look forward to this emerging technology whose birth stems from a lonely fungus found in New Guinea many years ago, to have an equivalent worldwide impact on our way of life."

By 1979, genetic enhancement of *T. reesei* had already produced mutant strains with up to 20 times the productivity of the original organisms isolated from New Guinea (*57,58*). For roughly 20 years, cellulases made from submerged culture fungal fermentations have been commercially available. In another ironic twist, the most lucrative market for cellulases today is in the textile industry, where they have found valuable niches such as in the production of "stone-washed" jeans.

In many ways, however, our understanding of cellulases is in its infancy compared to other enzymes. There are some good reasons for this. Cellulase-cellulose systems involve soluble enzymes working on insoluble substrates. The jump in complexity from homogeneous enzyme-substrate systems is tremendous. It became clear fairly quickly that the enzyme known as "cellulase" was really a complex system of enzymes that work together synergistically to attack native cellulose. In 1950, this complex was crudely described as a system in which an enzyme known as "C_1" acts to decrystallize the cellulose, followed by a consortium of hydrolytic enzymes, known as

"C_x" which breaks down the cellulose to sugar (*59*). This early concept of cellulase activity has been modified, added to and argued about for the past forty years (*60, 61*).

Though many researchers still talk in terms of the original model of a nonhydrolytic C_1 enzyme and a set of C_x hydrolytic enzymes, our current picture of how these enzymes work together is much more complex. Three major classes of cellulase enzymes are recognized today:

- Endoglucanases, which act randomly on soluble and insoluble glucose chains
- Exoglucanases, which include glucanhydrolases that preferentially liberate glucose monomers from the end of the cellulose chain and cellobiohydrolases that preferentially liberate cellobiose (glucose dimers) from the end of the cellulose chain
- β-glucosidases, which liberate D-glucose from cellobiose dimers and soluble cellodextrins.

For a long time, researchers have recognized that these three classes of enzymes work together synergistically in a complex interplay that results in efficient decrystallization and hydrolysis of native cellulose. In reaching out to "non-scientific" audiences, promoters of cellulase research often oversimplify the basic description of how these enzymes work together to efficiently attack cellulose (*62*). The danger in such oversimplifications is that they may mislead many as to the unknowns and the difficulties we still face in developing a new generation of cost effective enzymes. While our understanding of cellulase's modes of action has improved, we have much more to learn before we can efficiently develop enzyme cocktails with increased activity.

The first application of enzymes for hydrolysis of wood in an ethanol process was obvious—simply replace the acid hydrolysis step with an enzyme hydrolysis step. This configuration, now often referred to as "separate hydrolysis and fermentation" (SHF) is shown in Figure 5 (*63*). Pretreatment of the biomass is required to make the cellulose more accessible to the enzymes. Many pretreatment options have been considered, including both thermal and chemical steps.

The most important process improvement made for the enzymatic hydrolysis of biomass was the introduction of simultaneous saccharification and fermentation (SSF), as patented by Gulf Oil Company and the University of Arkansas (*64, 65*). This new process scheme reduced the number of reactors involved by eliminating the separate hydrolysis reactor and, more importantly, avoiding the problem of product inhibition associated with enzymes. In the presence of glucose, β-glucosidase stops hydrolyzing cellobiose. The build up of cellobiose in turn shuts down cellulose degradation. In the SSF process scheme, cellulase enzyme and fermenting microbes are combined. As sugars are produced by the enzymes, the fermentative organisms convert them to ethanol. The SSF process has, more recently, been improved to include the cofermentation of multiple sugar substrates. This new variant of SSF, known as SSCF for Simultaneous Saccharification and CoFermentation, is shown schematically in Figure 6.
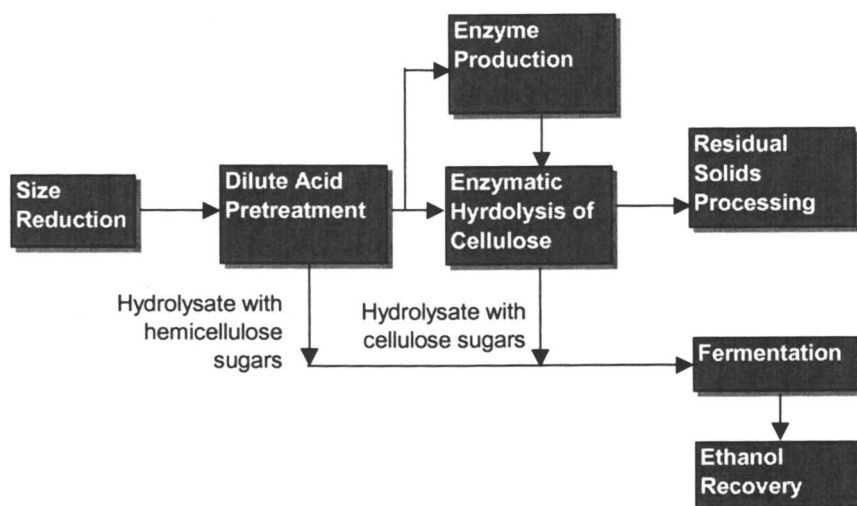
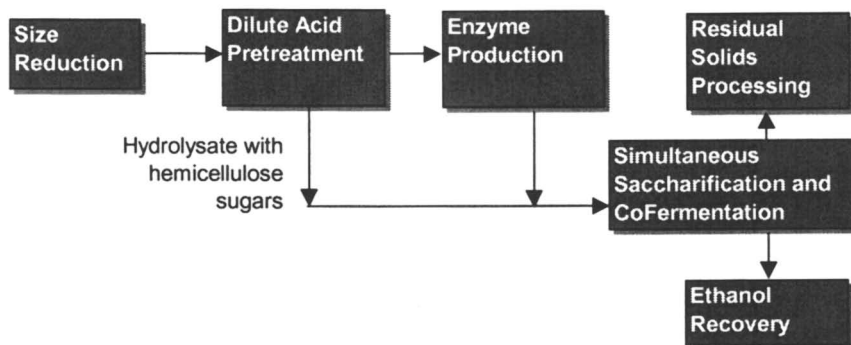*Figure 5. The enzyme process configured as Separate Hydrolysis and Fermentation (SHF*

Figure 6. The enzyme process configured for Simultaneous Saccharification and CoFermentation (SSCF)

As suggested earlier, cellulase enzymes are already commercially available for a variety of applications. Most of these applications do not involve extensive hydrolysis of cellulose. For example, the textile industry applications for cellulases require less than 1% hydrolysis. Ethanol production, by contrast, requires nearly complete hydrolysis. In addition, most of the commercial applications for cellulase enzymes represent higher value markets than the fuel market. For these reasons, there is quite a large leap from today's cellulase enzyme industry to the fuel ethanol industry. Our partners in commercialization of near-term ethanol technology are choosing to begin with acid hydrolysis technologies because of the high cost of cellulase enzymes.

Two companies have plans to deploy enzyme technology for ethanol production. Petro-Canada, the second largest petroleum refining and marketing company in Canada, signed an agreement with Iogen Corporation in November of 1997 to co-fund research and development on biomass-to-ethanol technology over a period of 12 to 18 months. Petro-Canada, Iogen and the Canadian government will then fund construction of a plant to demonstrate the process, which is based on Iogen's proprietary cellulase enzyme technology (*66*). .

BC International (BCI), mentioned in the previous section, will begin operation of their Jennings, Louisiana plant using dilute acid hydrolysis technology. The choice of dilute acid technology is strategic, in that it allows for the eventual addition of enzyme hydrolysis when cellulase production becomes cost effective. BCI is currently evaluating options for utilizing enzymes (*67*). BCI plans to utilize cellulase enzymes in a project partially funded by the Department of Energy that will lead to a commercial rice straw to ethanol facility in Gridley, CA by 2003.

Technology Pathways—The Promise Of Biotechnology

From a "big picture" technological perspective, there is every reason to believe that the progress made over the past few decades in genetic engineering technology

could be dwarfed by future advances. Biotechnology is an explosive field. New tools and breakthroughs are occurring at an exponential pace. Knowledge in the biological sciences is doubling every five years. In the field of genetics, the amount of information is doubling annually (*68*).

In 1997, *Business Week* declared the 21st century to be "The Biotech Century." They cite Nobel Prize winning chemist Robert F. Curl, who states that the 20th century "was the century of physics and chemistry. But it is clear that the next century will be the century of biology" (*69*). Jeremy Rifkin, a frequent critic of biotechnology, still acknowledges the profound impact that genetic engineering will have (*70*):

> "The marriage of computers and genetic science, in just the last ten years, is one of the seminal events of our age and is likely to change our world more radically than any other technological revolution in history."

It is in this broader context of biotechnology's bright future that we build a roadmap for bioethanol technology. We see the path for technology development as one that uses computer technology, biochemistry and molecular biology as the essential tools for fundamental improvement.

Cellulase Enzyme Development. Dr. Ghose, one of the pioneers in cellulase research, spoke these words almost thirty years ago:

> "Microorganisms have no difficulty digesting cellulose. They accomplish it rapidly and effectively. Why is it then that we cannot utilize their systems to develop a practical conversion of cellulose to sugar? The answer is rather simple; we can--if we pour into this problem the effort it rightly deserves." (*71*)

Despite his optimism, we have yet to crack the secrets of microbial cellulose hydrolysis. We still share his optimism. Learning how to use cellulase enzymes to efficiently digest cellulose to sugar requires a consistent effort that simply hasn't been applied up to now. Furthermore, we have access to exciting new biotechnological tools unimagined by Dr. Ghose in 1969. These new tools will make it possible to produce new enzymes specifically designed for use in industrial production processes.

Because of the importance of cellulase enzymes in the process, DOE and NREL sponsored a series of colloquies with experts and stakeholders in industry and academia to determine what types of improvements in enzyme production and performance offer the greatest potential for success in the short term (*72*). There was a clear consensus in these discussions that the prospects for enzyme improvement through protein engineering are very good. We identified the following targets for protein engineering:

- *Increased Thermal Stability.* Simply by increasing the temperature at which these enzymes can operate, we can dramatically improve the rate of cellulose hydrolysis. The genetic pool available in our labs and in

others around the world include thermo-tolerant, cellulase-producing organisms that represent a good starting point for engineering new enzymes.

- *Improved Cellulose Binding Domain.* Cellulase enzymes contain a catalytic domain and a binding domain. Improvements in the latter will lead to more efficient interaction between the soluble cellulase enzymes and the insoluble surface of the biomass.
- *Improved Active Site.* In addition to modifying the binding domain, we plan to modify amino acid sequences at the active site. Even minor modifications of the enzyme can lead to dramatic improvements in catalytic activity of the enzyme.
- *Reduced Non Specific Binding.* Enzyme that adsorbs on lignin is no longer available for hydrolysis. Genetic modifications of the enzyme will be geared toward adjusting its surface charge to minimize such unwanted binding.

We have identified two approaches for achieving these goals, both representing the state-of-the-art in biotechnology research. The first is a rational design approach known as site-directed mutagenesis. It uses sophisticated 3-D modeling tools to identify specific amino acids in the protein sequence that can affect the enzyme properties listed above (*73, 74, 75*). The second is a more recent strategy known among biotechnologists as "directed-evolution" (*76*). It combines advanced genetic engineering techniques with highly automated laboratory robotics to randomly evolve new enzymes with the features required. The enzyme performance goals that are indicated in the future cases are based on the projected progress for these research strategies. By 2005, improvements in thermostability of the enzymes should yield a three-fold improvement in specific activity. By 2010, enhancements in the cellulose binding domain, the active site and protein surface charge will lead to an increase in enzyme performance of ten fold or more.

In parallel with the protein engineering work, our program plan calls for research aimed at improving the productivity of the enzyme expression systems. Two targets for research are being pursued:

- Improved microbial organisms genetically engineered for high productivity of enzymes
- Genetically engineered crops harvested as feedstock, which contain high levels of cellulase enzymes

Higher efficiency microorganisms for use in submerged culture fermentors should be available by 2005.

New Organisms For Fermentation. Research over the past 10 years on ethanol producing microorganisms has yielded microorganisms capable of converting hexose and pentose sugars to ethanol (*77, 78, 79*). These ethanol-producing microorganisms ferment xylose and glucose mixtures to ethanol with high efficiency. This represents a major advance in technology, as previous conversion of pentose sugars by natural yeasts was not industrially attractive. Furthermore, these new ethanologens have eliminated the need for separate pentose and hexose fermentation trains.

Substantial improvement in biomass conversion can be achieved by making the following additional improvements in ethanol producing microorganisms:

- Ethanol producing microorganisms capable of producing 5% ethanol at temperatures greater than or equal to 50°C, and
- Ethanol producing microorganisms capable of converting cellulose to ethanol.

We have recently shown that a doubling of the rate of biomass hydrolysis for every 20°C increase in temperature of saccharification can be expected if *T. reesei*-like cellulases are used. The development of ethanologens capable of fermentation at temperatures greater than 50°C can potentially reduce the cost of cellulase enzyme by one-half. This is because the current ethanologens can only meet desired performance at temperatures of 30-33°C.

The most advanced processing option is one in which all biologically mediated steps (e.g., enzyme production, enzymatic cellulose hydrolysis, and biomass sugar fermentation) occur in a single bioreactor (80). This process, also known as direct microbial conversion (DMC) or Consolidated Bioprocessing (CBP), can be carried out to various extents by a number of microorganisms, including fungi, such as *Fusarium oxysporum* and bacteria, such as *Clostridia* sp. However, known DMC strains often exhibit relatively low ethanol yield and have not yet been shown effective in handling high concentrations of biomass.

Our program plan calls for introducing a high temperature ethanologen by 2005. This new organism should be able to operate at 50°C, while maintaining the best characteristics of the current ethanologens.

Ethanol Cost Savings In The Future. The improvements in enzyme and ethanologen performance will impact the process in 2005 and 2010. Genetically engineered feedstocks with higher carbohydrate content might happen in 2015—though the timing for this last item needs to be determined more precisely. Figure 7 shows the decline in bioethanol pricing based on these research targets. The upper and lower bounds on the error bars reflect the results of sensitivity studies to assess the effect of feedstock price. The lower bound is a price projection for $15 per dry U.S. ton ($17.50 per MT) feedstock and the upper bound is a price projection for $44 per dry U.S. ton ($40 per MT) feedstock.

Conversion technology improvements could provide a 35 cents per gallon cost reduction over the next ten years. Combining these improvements with genetically engineered feedstocks brings the savings to 40 cents per gallon.

Conclusions

New technology for the conversion of biomass to ethanol is on the verge of commercial success. Over the course of the next few years, we should see new acid hydrolysis-based bioethanol plants come on line, which use niche feedstocks that address an environmental issue, such as solid waste disposal. As improvements in enzyme technology become available, we expect to see bioethanol production coming
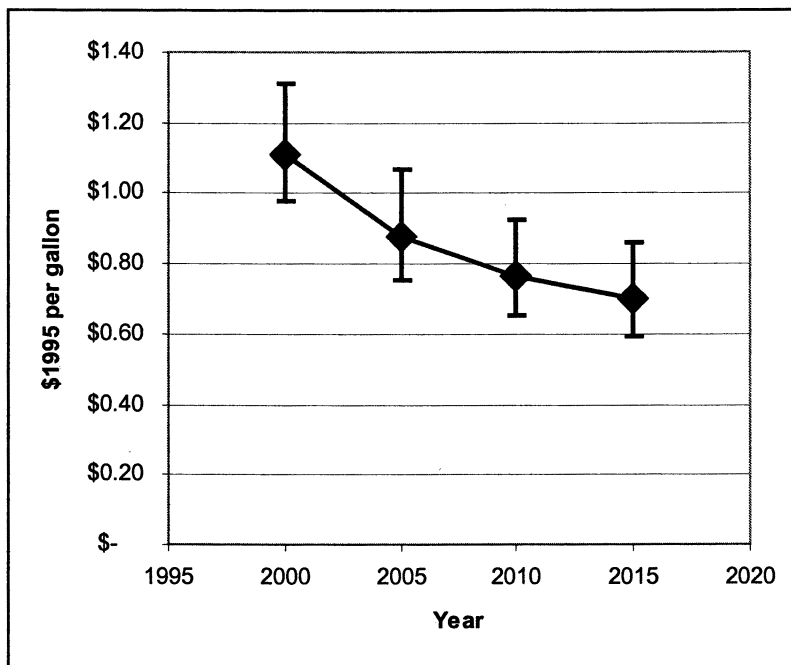
Figure 7. Price Trajectory for Enzyme-Based Process Technology

on line that provides ethanol at prices that can compete with other fuel additives and blending components without any subsidy. This technology should be available just as the existing incentives for fuel ethanol are scheduled to end. As concern about climate change, sustainability and other environmental issues increase, the opportunities for bioethanol will continue to grow. The next ten years should prove an interesting time for bioethanol, a time when bioethanol takes on much greater importance in the fuel market.

References

1. *Ethanol Industry Outlook: 1999 and Beyond*; Renewable Fuels Association: Washington, DC, 1999.
2. Kerr, R.A. *Science*, **1998**, 281, 1128-1131.
3. *Annual Energy Outlook 1996 with Projections to 2015*; U.S. Department of Energy, Energy Information Administration: Washington, D.C. 1996; DOE/EIA-0383(96).
4. *Reinventing Energy: Making the Right Choices*. The American Petroleum Institute, Washington, DC. 1996.

5. Stevens, P. *Energy Policy*, **1997**, 25/2, 135-142.
6. Romm, J.J.; Curits, C.B. *The Atlantic Monthly*, **1996**, April, 57-74.
7. Jennings, J.S. "Future Sustainable Energy Supply". An address to the 16th World Energy Council Congress. Tokyo, Japan, October 1995.
8. Tyson, K.S.; Riley, C.J.; Humphreys, K.K. *Fuel Cycle Evaluations of Biomass-Ethanol and Reformulated Gasoline, Volume I*; National Renewable Energy Laboratory: Golden, CO, 1993; Report No. NREL/TP-463-4950.
9. Bailey, B. "Chapter 2: Performance f Ethanol as a Transportation Fuel" In *Handbook on Bioethanol: Production and Utilization*; Wyman, C., Ed; Taylor and Francis: Washington, D.C., 1996.
10. Kelly, K. J.; Bailey, B.K.; Coburn, T.C.; Clark, W.; Lissiuk, P. "Federal Test Procedure Emissions test Results from Ethanol Variable-Fuel Vehicle Chevrolet Luminas." Presented at *Society for Automotive Engineers International Spring Fuels and Lubricants Meeting*, Dearborn, MI, May 6-8, 1996.
11. Farhar, B. *Trends in Public Perceptions and Preferences on Energy and Environmental Policy*; National Renewable Energy Laboratory: Golden, CO, 1993; NREL/TP-461-485.
12. Wilson, E.O. *Consilience: The Unity of Knowledge*; Alfred A. Knopf: New York, 1998, pp 266-298.
13. Gore, A. "Eco-nomics: Truth or Consequences." In *Earth in the Balance: Ecology and the Human Spirit*; Plume (Penguin Books): New York, 1993, pp 182-196.
14. Oreskes, N.; Shrader-Frechette, Belitz, K. *Science*, **1994**, 263, 641-646.
15. Michaels, P.J. *National Geographic Research & Exploration*, **1993**, 9/2, 222-233.
16. Schneider, S.H. *Science*, **1989**, 243, 771-781.
17. Gaffen, D.J. *Nature*, **1998**, 394, 615-616.
18. Revelle, R.; Suess, H.E. *Tellus*, **1957**, 9/4, 18-27.
19. Linden, E. "Burned by Warming," *Time*, March 14, 1994, p 79.
20. *Top Energy, Environment Leaders Urge President and Congress to Depoliticize Climate Debate, Take Long View, Start Now*. Press Release; The Aspen Institute: Washington, DC, September 1, 1998.
21. *Chrysler Introduces New Flexible-Fuel Technology*. Press Release; Chrysler Corporation: Auburn Hills, MI, June 10, 1997.
22. Bradshaw, K. *Ford to Hike Output of Vehicles Using Ethanol*; Press Release, June 4, 1997, America Online.
23. Renewable Fuels Association website. http://www.ethanolrfa.org.
24. Hall, D.O. *Solar Energy*, **1979**, 3, 307-328.
25. Rhodes, A.; Fletcher, D. *Principles of Industrial Microbiology*; Pergamon Press: New York, 1966.
26. Brazil's fuel ethanol program, on the other hand, relies heavily on the use of sugar from sugarcane. See, Goldemberg, J.; Monaco, L.; Macedo, I., "The Brazilian Fuel-Alcohol Program." In *Renewable Energy: Sources for Fuels and Electricity*; Johansson, T.; Kelly, H.; Reddy, A.; Wil;liams, R., Eds; Island Press: Washington, D.C. 1993.

27. Wyman, C., "The DOE/SERI Ethanol from Biomass Program." In *Ethanol Annual Report FY 1990*; Texeira, R.; Goodman, B., Eds; Solar Energy Research Institute (now the National Renewable Energy Laboratory): Golden, CO, 1991. Controls in the U.S. currently keep sugar prices at around $360/ton ($1990), making it far too expensive a feedstock for fermentation to ethanol.

28. Grohmann, K.; Himmel, M. "Chapter 1: Enzymes for Fuels and Chemical Feedstocks." In *ACS Symposium Series #460: Enzymes in Biomass Conversion*; Leatham, G.F.; Himmel, M.E., Eds; American Chemical Society: Washiington, DC, 1991, p 2-11.

29. Underkofler, L.A. "Chapter 21: Development of a Commercial Enzyme Process: Glucoamylase." In *Advances in Chemistry Series No. 95: Celluases and Their Applications*; Hajny, G.J.; Reese, E.T., Eds; American Chemical Society: Washington, D.C., 1969, pp 343-358.

30. Fan, L.T.; Gharpuray, M.M.; Lee, Y-H. "Chapter 2: Nature of Cellulosic Material." In *Cellulose Hydrolysis*; Springer-Verlag: New York, 1987, pp 5-20.

31. Grohman, K.; Himmel, M. "Chapter 1: Enzymes for Fuels and Chemical Feedstocks." In *ACS Symposium Series No. 460: Enzymes in Biomass Conversion*; Leatham, G.; Himmel, M., Eds; American Chemical Society: Washington, DC, 1991.

32. Braconnot, H. *Gilbert's Annalen der Physik,* **1819**, 63, 348.

33. Braconnot, H. *Ann chim. Phys.*, **1819**, 12, 172.

34. Simonsen, E. *Zeitschrift fur angewaudte Chemie*; **1898**, pp 962-966, 1007-1012.

35. Harris, E.E. "Wood Saccharification." In *Advances in Carbohydrate Chemistry,*; Academic Press: New York, 1949; Vol 4; pp 153-188.

36. LaForge, F.B.; Hudson, C.S. *The Journal of Industrial and Engineering Chemistry*, **1918**, 10/11, 925-927.

37. Dunning, J.W.; Lathrop, E.C. *Industrial and Engineering Chemistry*, **1945**, 37/1, 24-29.

38. Wenzl, H.F.J. "Chapter IV: The Acid Hydrolysis of Wood." In *The Chemical Technology of Wood*; Academic Press; New York, 1970, pp 157-252.

39. Tsao, G.T.; Ladisch, M.R.; Voloch, M.; Bienkowski, P. *Process Biochemistry*, **1982**, September/October, 34-38.

40. Broder, J.D.; Barrier, J.W.; Lightsey, G.R. "Conversion of Cotton Trash and Other Residues to Liquid Fuel." In *Liquid Fuels from Renewable Resources: Proceedings of an Alternative Energy Conference*; Cundiff, J.S., Ed; American Society of Agricultural Engineers: St. Joseph, MI, 1992, pp 189-200.

41. Wright, J.D.; d'Agincourt, C.G. *Biotechnology and Bioengineering Symposium*, **1984**, 14, 105-123.

42. "Two New Cellulosic Ethanol Plants in Late Preconstruction Stages" *New Fuels & Vehicle Report,* March 14, 1997.

43. Sherrard, E.C.; Kressman, F.W. *Industrial and Engineering Chemistry*, **1945**, 37/1, 5-8.

44. Faith, W.L. *Industrial and Engineering Chemistry*, **1945**, 37/1, 9-11.

45. Harris, E.E.; Beglinger, E. *Industrial and Engineering Chemistry,* **1946**, 38/ 9, 890-895.

46. Gilbert, N.; Hobbs, I.A.; Levine, J.D. *Industrial and Engineering Chemistry*, **1952**, 44/ 7, 1712-1720.

47. Harris, J.F.; Baker, A.J.; Conner, A.H.; Jeffries, T.W.; Minor, J.L.; Patterson, R.C.; Scott, R.W.; Springer, E.L.; Zorba, J. *Two-Stage Dilute Sulfuric Acid Hydrolysis of Wood: An Investigation of Fundamentals*; U.S. Forest Products Laboratory: Madison, Wisconsin, 1985; General Technical Report FPL-45.

48. Torget, R. *Milestone Completion Report: Process Economic Evaluation of the Total Hydrolysis Option for Producing Monomeric Sugars Using Hardwood Sawdust for the NREL Bioconversion Process for Ethanol Production*; Internal Report; National Renewable Energy Laboratory: Golden, Colorado, 1996.

49. Nguyen, Q. *Milestone Completion Report: Evaluation of a Two-Stage Dilute Sulfuric Acid Hydrolysis Process*; Internal Report; National Renewable Energy Laboratory: Golden, Colorado, 1998.

50. "Bagasse-to-Ethanol Plant Proposed." *Ethanol Report*, January 8, 1998.

51. Wald, M.L., "A New Bacterium Helps Turn Agricultural Waste Into Enrgy to Fuel Cars." *The New York Times*, October 25, 1998.

52. Augustine, N.R. *Biotechnology and Bioengineering Symposium*, **1976**, 6, 1-8.

53. Reese, R.T. *Biotechnology and Bioengineering Symposium*, **1976**, 6, 9-20.

54. Peterson, W.H.; Snell,J.F.; Frazier, W.C. *Industrial and Engineering Chemistry,* **1945**, 37/1, 30-35.

55. Mandels, M.; Reese, E.T. *Developments in Industrial Microbiology*, **1964**, 5, 5-20.

56. Brandt, D.; Hontz, L.; Mandels, M. *AIChE Symposium Series*, **1973**, 69, 127.

57. Mandels, M.; Weber, J.; Parizek, R., *Applied Microbiology*, **1971**, 21, 152.

58. Montenecourt, B.S.; Eveleigh, D.E. "Selective Screening Methods for the Isolation of High Yielding Cellulase Mutants of *Trichoderma reesei*." In *Advances in Chemistry Series: Hydrolysis of cellulose: Mechanism of Enzymatic and Acid Catalysis*; No. 181; American Chemical Society: Washington, DC, 1979; pp 289-301.

59. Reese, E.T.; Siu, R.G.H.; Levinson, H.S. *Journal of Bacteriology*, **1950**, 59, 485-497.

60. Lee, Y.-H.; Fan, L.T. "Properties and Mode of Action of Cellulase." *Advances in Biochemical Engineering*; Springer-Verlag,:New York, 1980 ; Vol. 17; pp 101-129.

61. Kuhad, R.C.; Singh, A.; Ericksson,K.-E. "Microorganisms and Enzymes Involved in the Degradation of Plant Fiber Cell Walls." In *Advances in Biochemical Engineering : Biotechnology in the Pulp and Paper Industry*; Eriksson, K.-E., Ed; Springer-Verlag: New York, 1997; pp 45-125.

62. See, for example, Wyman, C., "Overview of the Simultaneous Saccharification and Fermentation Process for Ethanol Production from Cellulosic Biomass." In *Ethanol Annual Report FY 1990*; Texeira, R.; Goodman, B., Eds; Solar Energy Research Institute (now the National Renewable Energy Laboratory): Golden, CO, 1991.

63. Wilke, C.R.; Yang, R.D.; von Stockar, U. *Biotechnology and Bioengineering*, **1976**, 6, 155-175.

64. Gauss, et al, U.S. Patent No. 3,990,944, November 9, 1976.
65. Huff, et al, U.S. Patent 3,990,945, November 9, 1976.
66. "Petro-Canada Announces alternative Fuel venture." Canada NewsWire Press Release. http://ww2.newswire.ca/releases/November 1997/24/c5556.html.
67. *Cellulase Enzyme Production on Novel Pretreatment Substrates.* U.S. Department of Agriculture Small Business Innovative Research Grant Abstract FY 1998. See http://www.reeusda.gov/crgram/sbir/98phase1.htm.
68. Rifkin, J. *The Biotech Century: Harnessing the Gene and Remaking the World;* Jeremy P. Tarcher/Putnam, Inc.: New York, 1998; pp 1-36.
69. Carey, J.; Freundlisch, N.; Flynn, J.; Gross, N. "The Biotech Century," *Business Week*, March 10, 1997, pp 78-90.
70. Rifkin, J. *The Biotech Century: Harnessing the Gene and Remaking the World*; Jeremy P. Tarcher/Putnam, Inc.: New York, 1998; pp x-xvi.
71. Ghose, T.K.; Kostick, J.A., "Chapter 24: Enzymatic Saccharification of Cellulose in Semi- and Continuously Agitated Systems." In *Advances in Chemistry Series: Cellulases and Their Applications*, Vol. 95, 1969; pp 415-446.
72. Hettenhaus, J.; Glassner, D. *Milestone Completion Report: Enzyme Hydrolysis of Cellulose: Short-Term Commercialization Prospects for Conversion of Lignocellulosics to Ethanol*; National Renewable Energy Laboratory: Golden, CO, 1997.
73. Himmel, M.E.; Karplus, P.A.; Sakon, J.; Adney, W.S.; Baker, J.O.; Thomas, S.R. *Appl. Biochem. Biotechnol.* , **1997**, 63/65, 315-325.
74. Warren, R.A.J. "Structure and Function in β-1,4-Glycanases." In *Carbohydrases from* T. reesei *and Other Microorganisms*; Claeyssens, M.; Nerinckx, W.; Piens, K., Eds.; The Royal Society of Chemistry: Cambridge, UK, 1998; pp 115-123.
75. Thomas, S.R.; Adney, W.S.; Baker, J.O.; Chou, Y.-C.; Himmel, M.E. U.S. Patent No. 5,712,142. Method for Increasing Thermostability in Cellulase Enzymes. July 1, 1997.
76. Arnold, F.H.; Moore, J.C. *Adv. Biochem. Eng./Biotechnol.,* **1997**, 58, 1-14.
77. Zhang, M., C. Eddy, K. Deanda, M. Finkelstein, S. Picataggio, *Science*, **1995**, 267, 240-243.
78. Ingram, Lonnie O.; Conway, Tyrrell; Alterthum, Flavio., *Ethanol Production by* Escherichia coli *strains co-expressing Zymomonas PDC and ADH Genes.* US Patent 5,000,000. Issued March 19, 1991.
79. Ingram, L.O.; Conway, T.; Clark, D.P.; Sewell, G.W.; Preston, J.F. *Applied and Environmental Microbiology*, **1987**, 53/10, 2420-2425.
80. Lynd, L; Elander, R.; Wyman, E. *Appl. Biochem. Biotech.,* **1996**, 57-58, 741-761.

Chapter 2

Role of Four Conserved Active-Site Aspartic Acid Residues in *Thermobifida fusca* Endoglucanase Cel6A

David B. Wilson, David Wolfgang, Sheng Zhang, and Diana Irwin

Department of Molecular Biology and Genetics, Cornell University, 458 Biotechnology Building, Ithaca, NY 14853

Each of the four conserved Asp residues in the *Thermobifida fusca* endoglucanase Cel6A active site were mutated to Ala, Asn and Glu. The mutant genes were expressed in *Streptomyces lividans*, purified to homogeneity, and their CD spectra, dissociation constants for several modified oligosaccharides and activities on several substrates were determined. The results showed that Asp117 functions as an essential catalytic acid, Asp156 helps shift the pKa of Asp117 towards 10, Asp79 is important but not essential for catalysis, while Asp265 appears to function mainly in substrate binding and clearly is not an essential catalytic base.

Thermobifida fusca Cel6A is a thermostable family 6 endocellulase with a very broad pH optimum and high activity on crystalline cellulose (*1*). Its molecular weight is 42K and it contains an N-terminal 30K catalytic domain joined by a 27-residue linker peptide to a family II cellulose binding domain (*2*). The 3-dimensional structure of the *T. fusca* Cel6A catalytic domain has been determined by X-ray crystallography at 1.18Å resolution and it was refined to 1.0Å resolution (*3, 4*). The structure is a globular, modified α-β barrel with a deep active-site cleft that contains at least four subsites that bind adjacent glucose residues in a cellulose chain (Figures 1 and 2). Cleavage of the cellulose chain occurs with inversion between the two central subsites.

There are four Asp residues that are in, or near, the active-site cleft and that are conserved in all family 6 enzymes, including *T. fusca* Cel6A, *C. fimi* Cel6A and *T. reesei* Cel6A (Table I). Three of them (Asp117, Asp156 and Asp265 in *T. fusca* Cel6A) are also structurally conserved between *T. fusca* Cel6A and *T. reesei* Cel6A, a family 6 exocellulase whose structure was the first cellulase structure determined (*5*) (Figure 3). The other conserved Asp (Asp79) is in a loop that covers part of the active-site cleft in *T. reesei* Cel6A, but the loop points away from the cleft in *T. fusca* Cel6A, so that Asp79 is 11Å away from the position of the corresponding residue in *T. reesei* Cel6A. The loop has a higher B value than the average Cel6Acd residue,
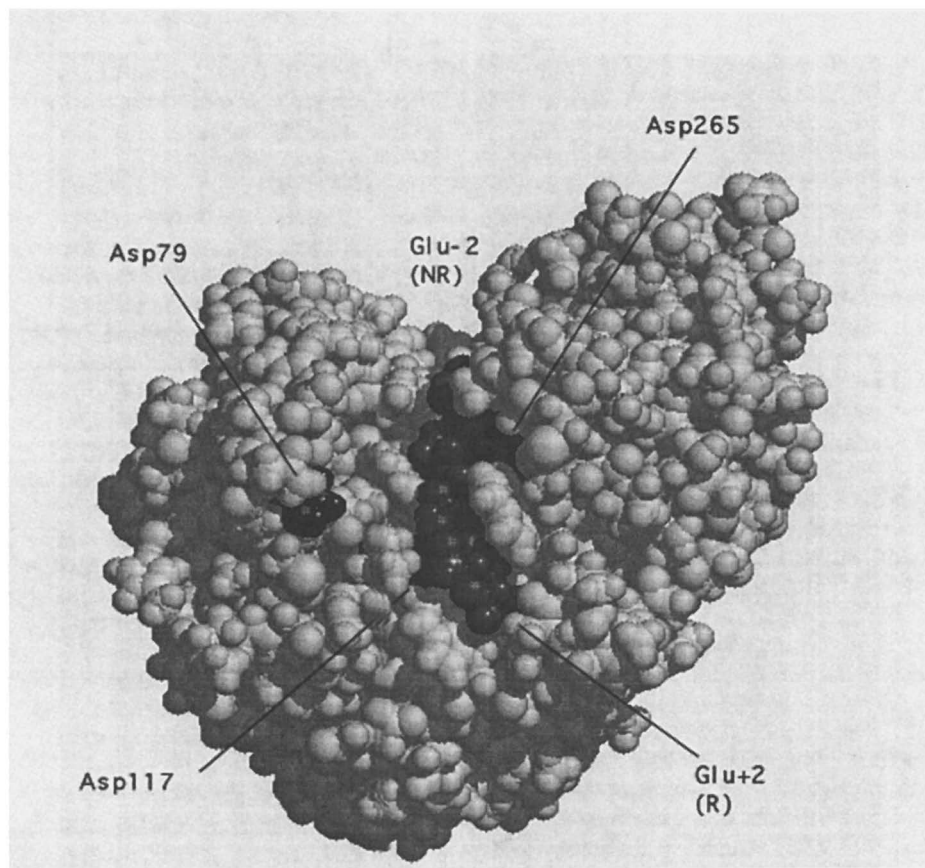
Figure 1.   Space filling model of Thermobifida fusca Cel6Acd with cellotetraose modeled into the active site (12).
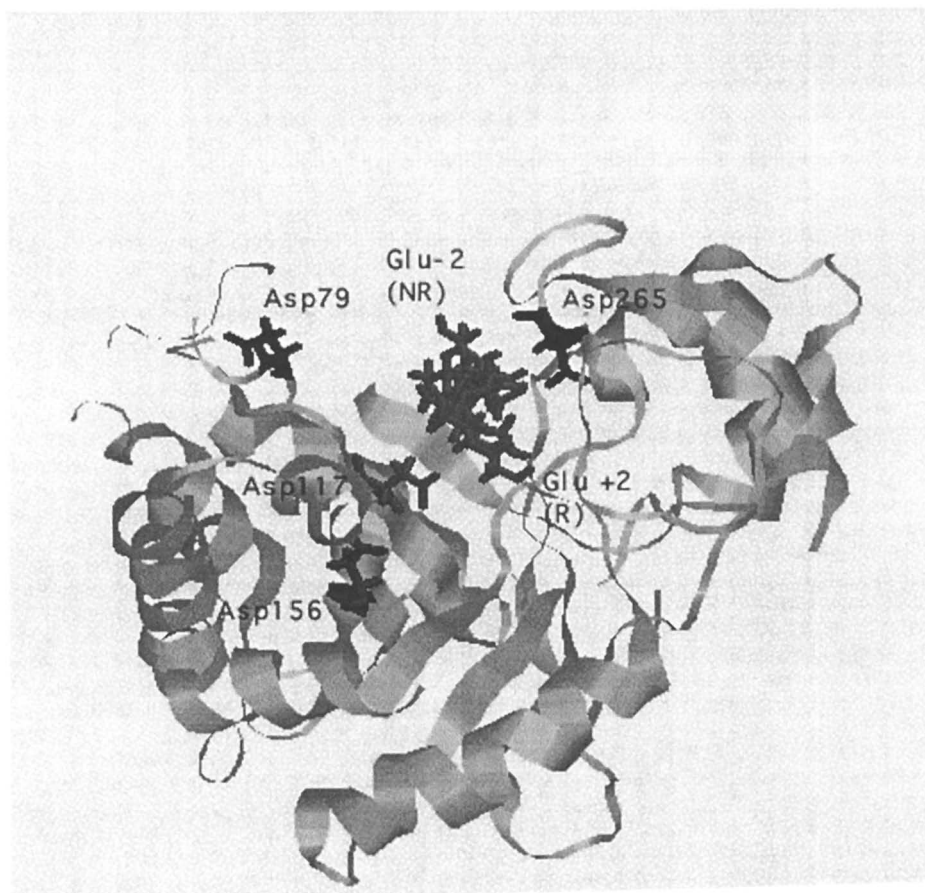
Figure 2. Ribbon figure model of Cel6Acd with the four conserved Asp molecules highlighted as stick figures. Shown in the same orientation as Figure 1, but tilted slightly to illustrate the relationship between Asp117 and Asp156.

showing that the loop residues are less rigid than the average residue despite the fact that the loop makes several crystal contacts with an adjacent molecule (*3*).

Table I. Corresponding Conserved Aspartic Acid Residues in Family 6 Cellulases

Thermobifida fusca Cel6A Swiss Prot P26222 PDB 1tml	*Trichoderma reesei* Cel6A Swiss Prot P07987 PDB 1CB2	*Cellulomonoas fimi* Cel6A Swiss Prot P07984
Asp79	Asp175	Asp216
Asp117	Asp221	Asp252
Asp156	Asp263	Asp287
Asp265	Asp401	Asp392

Gly86 and Gly87 are near one end of this loop and probably serve as a hinge that allows the loop to move from the open position seen in the *T. fusca* Cel6A structure to a closed position similar to that in *T. reesei* Cel6A. Each of these Gly residues was mutated to Ala and the Gly86Ala enzyme retained only 16% of wild-type activity on carboxymethylcellulose (CMC) and amorphous cellulose (SC), while the structurally more constrained Gly87Ala enzyme had only 4.5% wild-type activity on these substrates (*6*). Since the circular dichroism spectra of both mutant enzymes were very similar to that of wild-type and both mutant enzymes bound oligosaccharides with wild-type affinity, both mutant proteins appear to have folded properly. This result provides very strong evidence that loop mobility is essential for Cel6A activity, as it is very difficult to imagine any other reason that a Gly to Ala change in a non-core residue outside the active site would inhibit activity.

In our Cel6A structure paper (*3*), we proposed that Asp117 would function as the catalytic acid and Asp265 as the catalytic base in the mechanism for an inverting hydrolase proposed by Dr. Koshland (*7*). However, in the structure of Cel6Acd crystallized with cellobiose, Asp265 is hydrogen bonded to the glucose residue in subsite -1 and catalytic residues do not normally participate in substrate binding.

Site-directed mutagenesis of two of the conserved Asps in family 6 have been carried out on *T. reesei* Cel6A (*8, 9*) and on all four in *Cellulomonas fimi* Cel6A (*10*). For *C. fimi* Cel6A, each Asp was changed to both Ala and Glu and the mutant proteins were expressed in *E. coli*. The Asp392Glu mutant enzyme was not expressed, but all the others were. From the activities of the mutant enzymes, the authors proposed that Asp252 was the catalytic acid and Asp392 was the catalytic base. They proposed that Asp287 might shift the pKa of Asp252 and they did not propose a role for Asp216. For *T. reesei* Cel6A, the mutants were expressed in a *T. reesei* strain that lacked endogenous Cel6A, and Asp175 and Asp221 were mutated to Ala (*8, 9*). The Asp221Ala enzyme lacked activity on all tested oligosaccharide substrates, while the Asp175Ala enzyme retained about 20% of wild-type activity on MUG cellotetraose, but only 1% on the other oligosaccharides tested.

In our experiments, we mutated each of the four conserved Asp residues to Ala, Asn and Glu. We chose to make the Asn mutants because Asn has a very similar
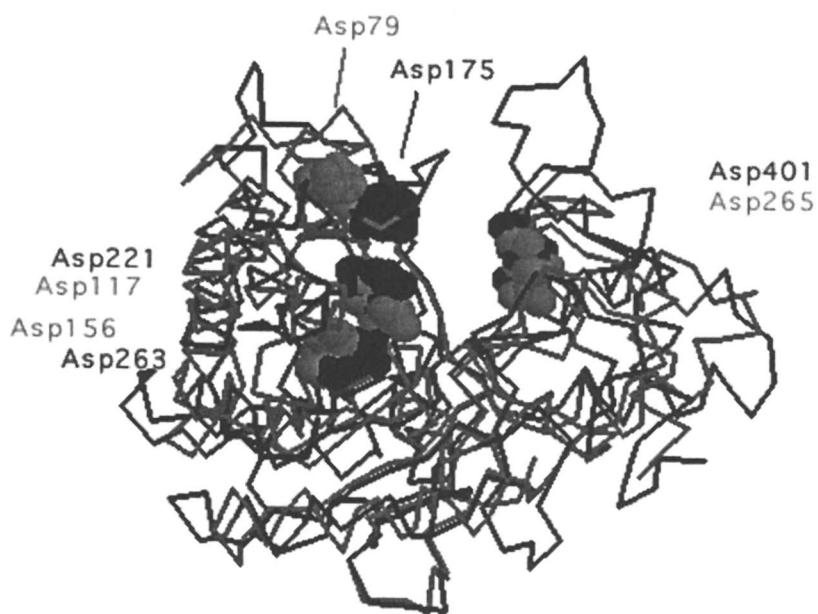
Figure 3. Overlay of the backbone structures of Cel6Acd (gray) and T. reesei Cel6Acd (black). The labeled Asp molecules are shown in space filling mode.

size and shape to Asp so that it should cause the smallest change in the structure of any possible amino acid substitution. The mutants were expressed in *Streptomyces lividans* and were purified free from a contaminating endoglucanase. The CD spectrum of each mutant protein was nearly identical to that of wild-type Cel6A, showing that each protein had folded correctly (data not shown). Furthermore, the dissociation constants for several oligosaccharide derivatives were determined for each mutant enzyme by fluorescence quenching (*11*) and were nearly identical to those for wild-type except for the Asp265 mutant enzymes, which have much weaker binding (Table II). Except for the Asp265 mutant enzymes, the binding data provides even stronger evidence than the CD spectra that the mutant enzymes have folded correctly.

Table II. Dissociation Binding Constants

Enzyme	MU(Glc)$_2$ nM	MU(Glc)$_3$ nM	Cellotriose μM	MU(XylGlc) nM
Cel6A	400	53	130	3.8
D79A	300	35	77	ND
D79E	290	200	*	ND
D79N	740	47	90	ND
D117A	270	57	120	ND
D117E	210	17	250	ND
D117N	1,500	*	*	ND
D156A	530	35	300	ND
D156E	1,000	100	190	ND
D156N	390	80	100	ND
D265A	ND	>50,000	ND	2,800
D265E	ND	ND	ND	30,000
D265N	>50,000	4,500	140	840
D265V	ND	ND	ND	50,000

*Curve does not fit the data. ND-not determined.
SOURCE: Reproduced from Wolfgang and Wilson, *Biochemistry* **1999**, *38*. 9746-9751.

The activities of the mutant enzymes on CMC, SC, filter paper (FP) and 2,4-dinitrophenyl cellobioside (DNPCB) are shown in Table III. Clearly Asp117 functions as the catalytic acid since all three mutant enzymes have very low activity on CMC and SC, but near wild-type activity on DNPCB which has a good leaving group. These results show that the mutants are able to carry out every function normally except protonation of the leaving group. The Asp156 Ala and Asn mutant enzymes also have near wild-type activity on DNPCB, but somewhat reduced activity on CMC and SC, showing that Asp156 only affects the catalytic acid. Figure 4 shows the pH activity profile of the wild-type and Asp156 mutant enzymes on SC. There is a dramatic narrowing and a shift of the pH optimum to more acid pH values for the Ala and Asn mutant enzymes, which provides strong evidence that the role of Asp156, which is only 3.9Å from Asp117, is to shift its pKa towards pH 10, creating the broad pH optimum of Cel6A. It is interesting that an Asp156 equivalent residue is present in *T. reesei* Cel6A whose pH profile resembles that of the Asn

mutant enzyme, but the distance between Asp263 and Asp221 in *T. reesei* Cel6A is 6.3Å, which is longer than the 3.9Å seen in *T. fusca* Cel6A. The fact that the Asp156Glu mutant enzyme has normal activity on all substrates and a nearly identical pH activity profile to that of the wild-type enzyme shows that a negative charge is required at this position, but its exact location is not that important. In the other three positions, the Glu mutant enzymes have essentially the same low activity as the Ala mutant enzymes, showing that in those locations the position of the carboxyl side chain is critical.

Table III. Activities of Wild Type and Mutant Enzymes

Enzyme	CMC %	SC %	FP %	2,4-DNPCB* kcat (min^{-1})
Cel6A[a]	100	100	100	1.1
D79A	1.3	0.7	42	
D79E	0.4	0.6	38	
D79N	3.9	2.0	55	0.14
D117A	0.03	0.02	9	1.2
D117E	0.06	0.06	17	0.3
D117N	0.02	0.02	5	0.3
D156A	2.2	1.0	67	0.5
D156E	130	120	84	1.3
D156N	10	8.6	74	0.5
D265A	2.5	2.0	26	
D265E	2.0	1.6	35	
D265N	11	7.0	56	0.05**
D265V	0.1	0.2	18	
D79N/D265N	0.23	0.14	27	

*All 2,4-DNPCB activities except for wild type and D117A were calculated at a single substrate concentration (1,000 µM).
**Data only reflects the specific activity rather than kcat under the given conditions since it is very likely that D265N is not saturated at 1,000 µM 2,4-DNPCB.
SOURCE: Reproduced from Wolfgang and Wilson, *Biochemistry* **1999**, *38*. 9746-9751.
[a]The specific activities of Cel6A on CMC, SC and FP were 355, 631 and 1.8 µmoles/min/mg respectively.

The properties of the Asp79 mutant enzymes are interesting in that all mutations significantly reduce activity on both CMC and SC, showing that this residue has some role in catalysis and providing further evidence for the loop movement discussed above. However, the fact that the Asn mutant enzyme still retained 2-3% of wild-type activity shows that this residue is not playing an essential role, such as that of the catalytic acid or catalytic base.

The properties of the Asp265 mutant enzymes clearly show that this residue is not the catalytic base, since the Asn mutant enzyme has 7-11% of wild-type activity on CMC and SC. In fact, the structural data and the binding result suggest that the major function of this residue is in substrate binding. The activity seen for the
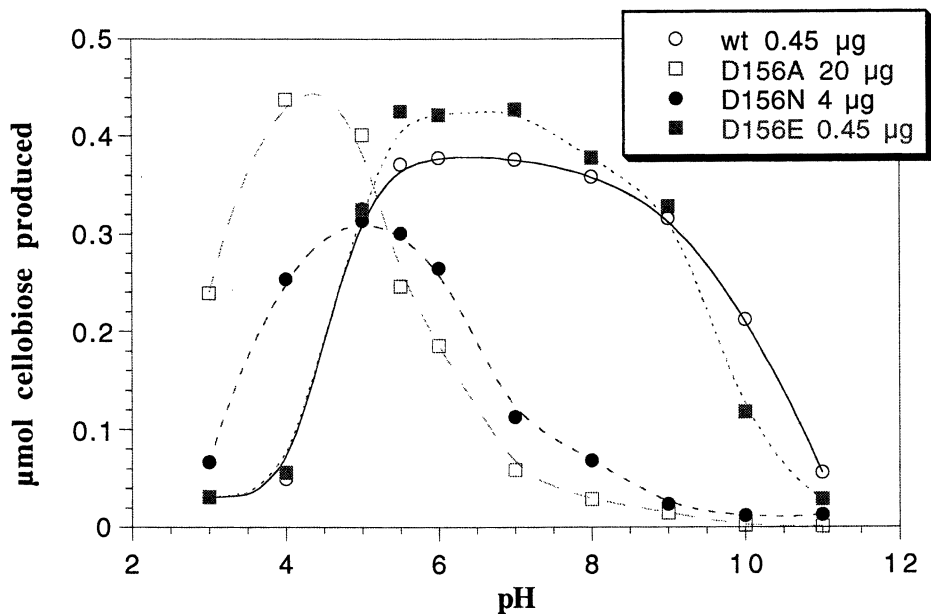
Figure 4. Activity of Asp156 mutant enzymes on SC as a function of pH.
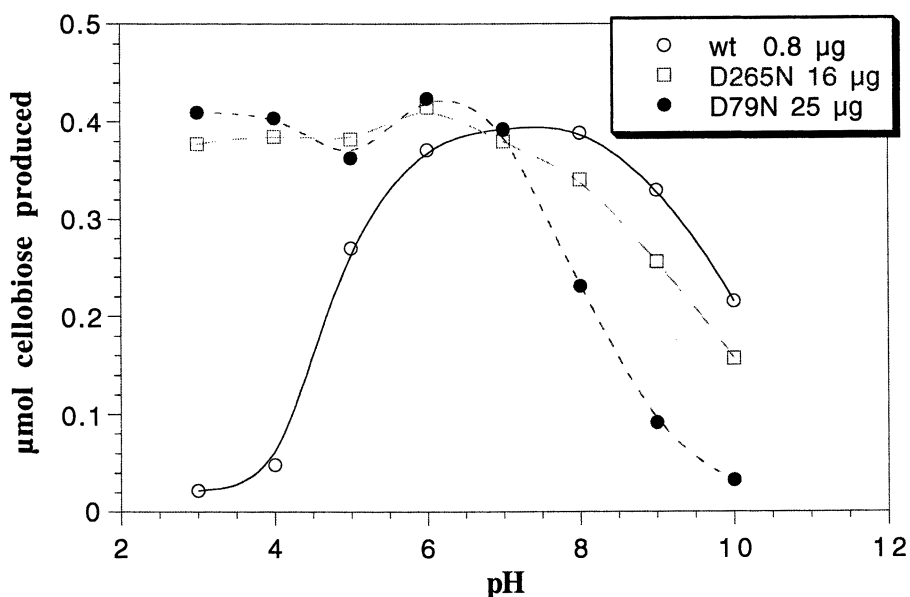
Figure 5. Activity of Asp265Asn and Asp79Asn mutant enzymes on SC as a function of pH.

Asp79Asn/Asp265Asn double mutant enzyme rules out the possibility that Asp79 and Asp265 share the catalytic base function, as it retains ~0.2% of wild-type activity on SC and CMC. This value is what is expected if the two residues have independent functions (8% x 3% = .24), whereas if they shared an essential function, the activity would be close to the .02% seen for the Asp117 mutant enzymes. The pH activity profiles of the 79Asn and the 265Asn mutant enzymes are shown in Figure 5 and are clearly different from that of the wild-type enzyme, so that the activity of these samples is clearly due to the mutant enzyme and not to a wild-type contaminant.

Table IV. Mutants With Higher Activity on CMC

| Plasmid | Enzyme | Relative Specific Activity on Substrates (%)[a] | | | |
		CMC	SC	FP	HEC
Whole enzyme					
pSZ61	Wild type	100	100	100	100
pSZ36	$R^{237}A$[b]	176	99	82	163
pBB1	$E^{263}D$	236	86	80	200
pSZ68	$R^{237}A\text{-}E^{263}D$	290	63	65	200
pBB3	$E^{263}Q$	162	72	68	246
pSZ69	$R^{237}A\text{-}E^{263}Q$	363	53	58	161
pSZ157	$H^{159}S$	168	46	65	ND
pSZ109	$K^{259}H$	373	66	30	327
Catalytic domain					
pSZ61	Cel6Acd (wt)	100	100	100	100
pSZ36	$R^{237}Acd$[b]	144	88	75	ND
pBB1	$E^{263}Dcd$	162	82	93	ND
pSZ68	$R^{237}A\text{-}E^{263}Dcd$	173	50	71	ND
pBB3	$E^{263}Qcd$	189	53	76	ND
pSZ69	$R^{237}A\text{-}E^{263}Qcd$	203	42	65	ND
pS109	$K^{259}Hcd$	262	30	15.6	230

[a]The specific activity (SA) of Cel6A on CMC, SC, FP and hydroxyethyl cellulose was 355, 631, 1.84 and 90 μmole cellobiose/min/μmole enzyme, respectively. The SA of Cel6Acd on CMC, SC, FP and HEC was 330, 560, 0.93 and 81 μmole cellobiose/min/μmole enzyme, respectively.
[b]The activity data of R237A and R237Acd on CMC, SC and FP were from Zhang and Wilson (*J. Biotech.* **1997**, *57*, 101-113).

An interesting result is that every mutant enzyme has a much higher relative activity on FP, than it has on CMC and SC. This shows that the rate-limiting step for filter paper hydrolysis is different than the rate-limiting steps for CMC and SC hydrolysis. Since the rate of FP hydrolysis is about 0.5% of the rate on CMC or SC, this is not unexpected. Further evidence for a different rate-limiting step is shown in Table IV, which shows the activities of the seven mutants in four different non-catalytic residues that have higher activity than wild-type on CMC. None of these mutants has higher activity on either FP or SC. These results also support the

proposal that SC hydrolysis has a different rate-limiting step than CMC hydrolysis, as we first suggested in reference 6. This is a surprising result since the rates of CMC and SC hydrolysis are very similar. CMC is soluble and thus should bind readily in the active site; however, the high proportion of modified residues may cause CMC to bind in a distorted manner that reduces the rate of bond cleavage, which then would be the rate-limiting step for CMC hydrolysis. In contrast, SC is insoluble so that entry of a cellulose molecule into the active site would be rate limiting and once the molecule is bound, it would be cleaved rapidly as it can bind optimally into the active site.

Our results on the Asp79, Asp117 and Asp156 mutant enzymes are in complete agreement with the results for *C. fimi* Cel6A and *T. reesei* Cel6A (*8-10*). However, our results for the Asp265 mutant enzymes are very different from the results reported for *C. fimi* Cel6A. The Asp265, Ala and Glu mutant enzymes had 1% activity and the Asn mutant enzyme had 10% activity, while the *C. fimi* Cel6A Ala mutant enzyme had only 0.001% activity and the Glu mutant enzyme could not be expressed. There are two possible explanations for these differences. One is that the two enzymes utilize different mechanisms and the other is that the very low activity of the *C. fimi* Cel6A mutants is due to their instability. It is very unlikely that *C. fimi* Cel6A and *T. fusca* Cel6A have different catalytic mechanisms since they are both members of family 6 and show 42% identity in their catalytic domains. It seems more likely that the thermophilic *T. fusca* Cel6A mutants are more stable than the mesophillic *C. fimi* mutants and thus we are able to measure the true activities of the mutants. Furthermore, we found higher activity with all of the Asn mutant enzymes, which were not made for *C. fimi* Cel6A. The fact that the *C. fimi* Cel6A Asp392Glu mutant enzyme could not be expressed and the *T. fusca* Cel6A Asp265Glu enzyme was expressed is also consistent with *T. fusca* Cel6A being more stable than *C. fimi* Cel6A.

In conclusion, our results show no evidence for an essential catalytic base in *T. fusca* Cel6A. The Asp residue that has the largest impact on activity besides the catalytic acid, Asp117, is Asp79. However, its position in the structure is uncertain, because of the potential for movement of the loop where it is located. It seems likely that the factors responsible for the catalytic enhancement caused by *T. fusca* Cel6A are: Asp117 activity as a catalytic acid, Asp79 (but its function is unclear), and distortion of the sugar in subsite -1.

References

1. Wilson, D. B. *Methods Enzymol.* **1988**, *160*, 314-323.
2. Ghangas, G. S.; Wilson, D. B. *Applied & Environ. Microbiol.* **1988**, *54*, 2521-2526.
3. Spezio, M.; Wilson, D. B.; Karplus, P. A. *Biochemistry* **1993**, *32*, 9906-9916.
4. Sakon, J.; Irwin, D.; Wilson, D. B.; Karplus, P. A. *Nat. Struct. Biol.* **1997**, *4*, 810-817.
5. Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. *Science* **1990**, *249*, 380-386.
6. Zhang, S.; Wilson, D. B. *J. Biotechnol.* **1997**, *57*, 101-113.
7. Koshland, D. E. *Biol. Rev.* **1953**, *28*, 416-436.

8. Ruohonen, L.; Koivula, A.; Reinikainen, T.; Valkeajarvi, A.; Teleman, A.; Claeyssens, M.; Szardenings, M.; Jones, T. A.; Teeri, T. T. In *Second TRICEL symposium on Trichoderma reesei cellulases and other hydrolases*; Suominen, P.; Reinikainen, T., Eds.; Foundation for Biotechnical and Industrial Fermentation Research, Espoo, Finland, 1993, pp. 87-96.

9. Koivula, A.; Reinikainen, T.; Rouhonen, L.; Valkeajarvi, A.; Claeyssens, M.; Teleman, O.; Kleywegt, G. J.; Szardenings, M.; Rouvinen, J.; Jones, T. A.; Teeri, T. T. *Protein Eng.* **1996**, *9*, 691-699.

10. Damude, H.; Withers, S.; Kilburn, D.; Miller, R.; Warren, R. A. *Biochemistry* **1995**, *34*, 2220-2224.

11. Barr, B. K.; Wolfgang, D. E.; Piens, K.; Claeyssens, M.; Wilson, D. B. *Biochemistry* **1998**, *37*, 9220-9229.

12. Taylor, J. S.; Teo, B.; Wilson, D. B.; Brady, J. W. *Protein Engineering* **1995**, *8*, 1145-1152.

Chapter 3

Gene Structure of a Bifunctional Cellulase Gene (celA) Isolated from *Teredinobacter turnerae*

S. N. Freer[1], R. V. Greene[2], and R. J. Bothast[1]

[1]Fermentation Biochemistry Research Unit and [2]Biopolymer Research Unit, National Center for Agricultural Utilization Research, Agricultural Research Service, U.S. Department of Agriculture, 1815 North University Street, Peoria, IL 61604

Bacterial cultures (*Teredinobacter turnerae*) isolated from the gland of Deshayes of the marine shipworm (*Psiloteredo healdi*) produce extracellular cellulase activity when cultured with cellulose. Screening a pUC 19 genomic DNA library for clearing of Ostrazin Brilliant Red-hydroxyethyl Cellulose allowed for isolation of the *T. turnerae* celA gene. Sequence analysis revealed that the celA gene encodes a putative protein of 1010 amino acids (106,068 kDa). The protein appears to be multidomained, consisting of an N-terminal endocellulase, two cellulose-binding domains, and a C-terminal exocellulase. Three linker regions, rich in serine, separate the four distinct activity domains. The endocellulase domain is most closely related to celE of *P. fluorescens*, while the exocellulase domain is most closely related to the cbhA gene of *Cellulomonas fimi*.

Marine shipworms (Teredo), boring their way into the hulls of ships, have a long and interesting relationship with mankind. The earliest fossil records of Teredinidae appear in the Jurassic (*1*), 213 to 144 million years ago. The writings of Ovid (*2*), whose birth was 43 BC, contain references to shipworms. Many of the early records are reports of extreme infestation by these creatures, which is akin to a severe marine termite outbreak. These marine borers may have changed the course of civilization. In February of 1588, the Spanish Armada sailed with 196 ships to attack England. The fleet was buffeted by North Sea winds and returned to port for emergency repairs. Months later, the fleet again sailed toward England. Again, storms and Queen Elizabeth's navy sank ship after ship. By August 1588, King Philip's II fleet was in shambles. Four hundred years later, naval historians speculate that the vessels in the Spanish Armada were unable to withstand the heavy North Sea weather because they were riddled with shipworms, thus costing Spain the war.

Shipworm damage has been a plague to maritime enterprises throughout recorded history. Warships, merchant vessels, and small coastal trading craft had short lives

39

before the 18th century. Two years without care and bottom timbers exposed to the sea were mined with tunnels. Even today, shipworms remain a challenge to those responsible for harbor maintenance, especially in warmer waters. An increasing number of shipworms have been noted in some American harbors, such as the very southern Hudson River. This probably reflects an improvement in water quality since bivalves are very sensitive to pollution.

The amount and severity of damage attributed to shipworms are reflected by the extent to which mankind has gone to prevent or remedy the problem. One of the early prevention methods used by ships of Columbus' era, which continued through the 1850s, was to attach wood sheathing to the hulls with hair and pine tar. Once infected, the attached wood was replaced. In 1758, the first copper sheathing was attached to British warships. Although expensive, this was very effective at preventing teredo boring into the wooden hulls. By the 1790s, most of the British naval ships were coppered. However, only a few merchant ships could afford this kind of protection.

Peoples from around the globe have devised many "traditional methods" to protect their wooden craft from shipworm infestation and some are still in use today (3). The standard preventive treatment for American colonists was pitch and pine tar. The hull was protected as long as no bare wood was exposed to the seawater. Portuguese and Japanese ship owners beached their vessels at high tide, then surrounded them with straw and ignited it. Their hope was to boil or steam the shipworms in their tunnels, without destroying the boats. Traditional boat builders in Georgia beached their vessels when a freeze was expected. A freeze was considered successful if slime oozed from the pinpoint holes in the hull. The Chinese and peoples of the Caribbean islands tried to prevent shipworm infestations by adding cayenne pepper to their bottom paints. One of the oldest ways to protect a craft from shipworm damage is to sail it up a freshwater creek or river and remain there for 3 to 4 weeks. Shipworms have a low tolerance for fresh water, but, in poor environments, they can seal themselves in their burrows for several weeks. Today, the most effective means to avoid shipworm damage is by use of chemicals in antifouling paints. However, these must be utilized judiciously, as they can pose environmental hazards.

In the 1700s, the study of shipworms was a matter of survival to the Dutch, whose dikes were constructed of wood. Thus, it is not surprising that the most extensive and accurate early scientific contributions were by Dutch scientists. In 1733, Sellius (4) first reported that shipworms were bivalves, not worms. These two-shelled mollusks are related to clams and oysters. Fourteen genera have been described, with *Teredo* and *Bankia* being the most common (1). These mollusks are distributed worldwide. No one knows the limits to their vertical distribution, but it is believed that they digested the Titanic's elaborate woodwork, which sank to the ocean floor to a depth of 13,000 feet in 1912 (3). Also, examination of preserved wood dredged from the ocean floor at various sites around the world has revealed the presence of shipworms at depths of up to 7488 meters (1). Shipworms consist of three main body structures: valves, a worm-like body, and pallets. At the end nearest the outside of the wood, each shipworm has a pair of calcified pallets that act as a plug in times of distress. Otherwise, two siphons extend into the outside water. One sucks in water for breathing, and possibly feeding, and the other expels wastes. The

long, soft body can extend from inches to up to 2 meters, depending upon the species and the environment. The most anterior portion of the adult shipworm is covered by two ribbed shells that the animal uses to rasp wood efficiently and form its burrows.

There is still controversy as to whether shipworms tunnel only to protect their soft bodies, digesting little of the liberated wood fiber and utilizing plankton as their major food source, or, whether they utilize the wood as a major carbon/energy source. The answer might be dependent upon the species of shipworm. The length of the gills and the size of the gill lamellae vary greatly between species. It is likely that species that have relative short gills with truncated lamellae depend largely on wood for food, while those with very long gills and long lamellae may feed upon plankton. Also, the anatomy of the digestive tract varies somewhat between species. Species with elongated stomachs and enlarged accessory pouches or caeca, especially the posterior wood-storing caecum, are probably better able to digest wood (*1*). Dore and Miller (*5*) analyzed samples of wood particles ejected by *Teredo navalis* and found they differed in chemical composition from the wood in which the animal was boring. They concluded that "on this basis, it appears that during the passage through the animals' digestive tract the wood has lost about 80% of its cellulose and from 15% to 56% of its hemicellulose."

For those who believe that shipworms could grow on a diet of wood alone, the animals' source of fixed nitrogen was a perplexing question. In 1848, Deshayes (*6*) described a gland consisting of dark or brown, irregular tissue that is found in shipworm gills. Sigerfoos (*7*), in 1909, named the tissue after Deshayes. The gland of Deshayes is not a gland in the traditional sense; rather, it is more aptly described as specialized tissue associated with shipworm gills. In 1983, Waterbury et al. (*8*) isolated and cultured a bacterium that exists as a pure culture in the gland of Deshayes. This bacterium is capable of both nitrogen fixation and cellulose digestion. It was postulated that the bacterium enabled the host shipworm to grow on wood because it provided a ready supply of protein as well as the primary source of cellulose-digesting enzymes. Subsequently, Waterbury's group (*9*) isolated the same bacterial species from shipworms belonging to four diverse genera. The bacterium was most closely related to *Pseudomonas aeruginosa* and *Oceanospirillum linum*.

Two independent laboratories (*8, 10*) have partially characterized the shipworm bacterium and reported similar results. Throughout exponential phase, this gram-negative bacterium is a rigid-rod. However, in stationary phase, the rods can elongate and form spirals. It is mobile, utilizing a single flagellum for movement. The bacterium utilizes a limited number of carbon sources for growth (cellulose, cellodextrins, sucrose, succinate, and possibly glucose) and requires elevated concentrations of sodium, chloride, magnesium, and calcium. As previously mentioned, the shipworm bacterium fixes nitrogen. This occurs under microaerophilic conditions. In the presence of a combined nitrogen source, vigorous growth occurs under aerobic conditions. The doubling time for nitrogen-fixing growth is 1 to 2 days, while under aerobic conditions in combined nitrogen supplemented medium, the doubling time is reduced to 8 to 15 hours. Optimal growth occurs at about 35°C and pH 8. When growing on cellulosic substrates, the bacterium produces several organic acids, of which succinic acid is the most prevalent (*10*).

It has been speculated that the shipworm bacterium provides the host animal with the cellulase enzymes necessary to digest the wood particles resulting from its boring activity (*8, 11*). However, the cellulase complex of the shipworm bacterium (*Teredinobacter turnerae*) has been only partially characterized. Cell-free medium from *T. turnerae* cultures extensively and visibly disrupts filter paper after brief incubations. However, little or no free sugars are released. When carboxymethylcellulose is used as the substrate, the reducing sugar content of the CMC increases greatly. These observations suggest that the majority of the secreted activity is of an endoglucolytic nature (*12*). Table 1 is a compilation of some of the properties of the cellulolytic activity found in crude cell-free culture medium (*13*). Although the production levels of the cellulases are somewhat low (200 to 500 CMCase units/liter), the overall specific activity (approximately 40 CMCase units/mg protein) of the cellulase complex is high, resulting in an active culture beer. For a rough comparison, *Trichoderma reesei* strains produce significantly more cellulase activity (from 3,600 to 60,000 CMCase units/liter) while the specific activity of the complex varies from 0.45 to 5.8 CMCase units/mg protein, depending upon the strain, growth medium, growth conditions, etc., (*14, 15, 16*). Culture beers of *T. turnerae* exhibit reasonable thermal stability. In the absence of substrate, little or no activity is lost when heated to 50°C (*12*). In the presence of substrate, residual activity has been observed after heating to 100°C (*17*). The crude enzymes function over a broad pH range and are insensitive to the presence of high concentrations of salt. The activity is relatively stable in solution and can be stored for extended periods of time as a lyophilized powder. Finally, the cellulase activity is resistant to protease treatment (*12*).

We deemed that the unique ecological aspects of the host/shipworm bacterium relationship, as well as the interesting biochemical properties of the bacterial cellulase complex and its potential commercial utility merited further study. Our approach was to clone the cellulase genes and express them in *E. coli* because of the lack of extracellular protein produced by *T. turnerae*. In this paper, we describe the partial characterization of a bifunctional cellulase gene isolated from *T. turnerae*.

Materials and Methods

Cell Culture, Enzyme Assay, and Chemicals. *Teredinobacter turnerae* (Woods Hole Oceanographic Institute strain number T8301), isolated from *Psiloteredo healdi* shipworm, was a gift of Dr. John Waterbury. Cultures were grown aerobically as previously described (*10*) with 0.05% NH₄Cl as a nitrogen source, 0.01% yeast extract as a vitamin source, and either 0.5% phosphoric acid-swollen cellulose (*18*) or sucrose as a carbon source. Unless mentioned otherwise, all chemical reagents were purchased from Sigma Chemical Company, St. Louis, MO.

Cloning of a *T. turnerae* Cellulase Gene. *T. turnerae* chromosomal DNA was partially digested with EcoR1 and ligated with Eco R1/CIP (calf intestine phosphatase) treated pUC19. The ligated mixture was used to transform *E. coli* DH5 competent cells. Transformed cells were selected on medium containing per liter: tryptone (10 g), NaCl (2.5 g), KCl (2.5 g), CaCl₂ (0.147 g), ampicillin (100 μg/ml),

Table 1. Properties of cellulolytic activity secreted by _T. turnerae_[a].

Cellulase Activity[b,c]	
units/liter	200 to 550
mg protein/liter	5 to 13
number of enzymes	3 to 7
Thermal Stability[b]	
maximum activity	60°C
inactivation	75°C
pH Stability[b]	
>50% maximum activity	4.5 to 8.5
Salt Tolerance[b]	
>50% maximum activity	0 M to 4 M NaCl
Half-life[d]	
liquid, 25°C	35 days
liquid, 4°C	63 days
Protease Sensitivity[b]	
trypsin	resistant
chymotrypsin	resistant
pronase	resistant

[a]Data compiled from cultures grown aerobically in artificial sea water medium containing 1% microcrystalline cellulose and 0.1% ammonium chloride. Cultures were harvested in early stationary phase.

[b]Data from Griffin et al. (_12_).

[c]Units are micromoles of glucose reducing sugar equivalents formed from CMC per minute at 25°C

[d]In the presence of cellulose, residual activity has been observed after a 30 minute exposure to 100°C (_17_).

and agar (15 g). The agar was overlaid with 5 ml of medium containing 0.01 g/ml of Ostrazin Brilliant Red-hydroxyethyl Cellulose (OBRC). Positive clones were selected based upon the formation of clear halos that surrounded the colonies.

Molecular Biology Procedures. Standard methods were used for DNA isolation, subcloning, and restriction enzyme analysis (*19*). Southern hybridization analyses were performed with the Genius System (Boehringer Mannheim, Indianapolis, IN) according to the manufacturers recommendations. Sequence analyses for both strands were performed using the Taq DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Reaction mixtures were purified with Quick Spin columns (Boehringer Mannheim) and analyzed with an Applied Biosystems 377 DNA sequencer. Oligonucleotide primers used for sequencing were prepared on an Applied Biosystems 392 DNA/RNA Synthesizer. Computer analysis of nucleotide data was performed using the Lasergene computer program from DNASTAR, Inc. (Madison, WI).

Results

Screening and Identification of Positive Clones. The partial *Teredinobacter turnerae* plasmid gene bank was screened for endo-1,4-β-glucanase activity using the chromogenic substrate OBRC. Identification of positive clones was based upon the ability to degrade ORBC, thereby forming a colorless halo on a red background. From approximately 6,000 total clones, two positive clones were identified. Restriction enzyme analysis revealed that both clones carried identical 6.5 kb DNA inserts (Fig. 1). Subclones of the parent insert were prepared and tested for their ability to clear ORBC. Results indicated that the endocellulase gene was located between the second Eco R1 site and the Sph 1 site. Additionally, the orientation of the insert did not significantly effect the plate assay, suggesting that *E. coli* recognized the *T. turnerae* promoter.

Sequencing of the celA Gene. Primers M13 (-20) and M13 reverse (Stratagene, La Jolla, CA) were used to sequence the 5' and 3' ends of the inserts shown in Fig. 1. Synthetic oligonucleotides specific to selected portions of the insert were then used as primers to continue sequencing until a complete open reading frame (ORF) was unambiguously established. A total of 5.133 kb of DNA was sequenced. The largest ORF identified was 3.033 kb and initiated 1.663 kb downstream from the first Eco R1 site. It terminated 307 bp upstream from the start of the Sph 1 site.

The predicted product of the celA gene is shown in Fig. 2. The protein is composed of 1010 amino acids with a molecular mass of 106,068 Daltons. The protein appears to be composed of four major activity domains that are separated by three serine-rich regions. The first 337 amino acid residues, which are designated Cat 1, exhibit significant sequence similarity with endo-1,4-glucanases belonging to glycoside hydrolase family 5 (Clan GH-A), as defined by Henrissat et al. (*20, 21*) (Fig. 3). Closest sequence similarities are with the catalytic domains of *Pseudomonas fluorescens* CelE (*22*) and *Erwinia chrysanthemi* CelZ (*23*). Pairwise (Bestfit) comparisons, in which signal sequences were omitted, revealed 61.3% and 60.4%
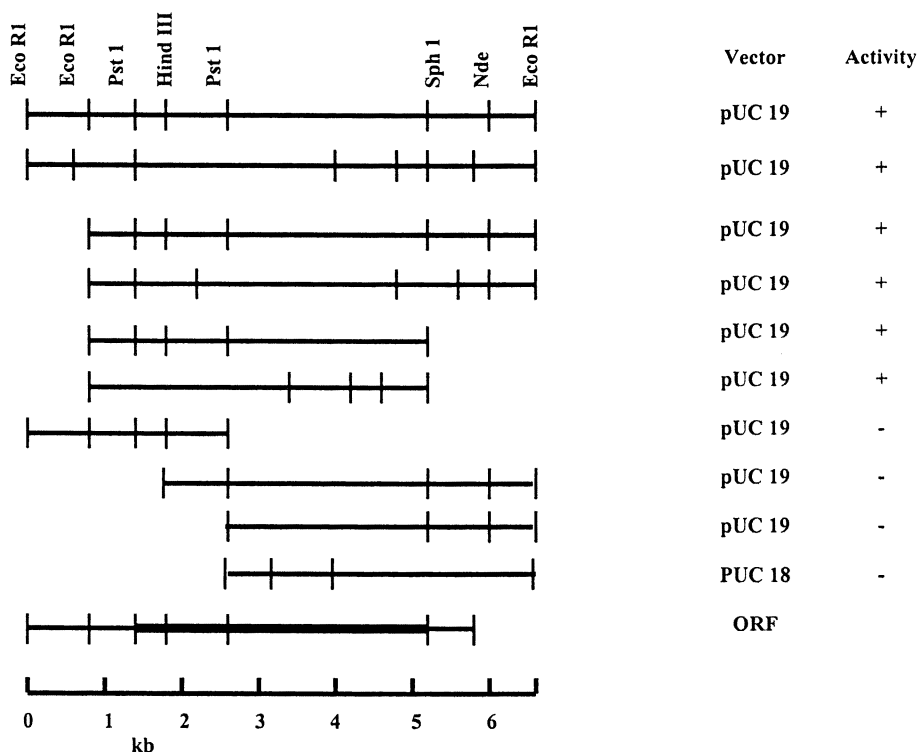
Figure 1. *Restriction endonuclease mapping and identification of the region encoding cellulase activity in a DNA fragment from T. turnerae. Fragments that were successfully subcloned in plasmid vector pUC19 or pUC18 are shown beneath the restriction map of the parent clone. In most cases, the subcloned fragments were prepared in both orientations. The (+) in the activity column indicates that the clone exhibited a positive reaction (clearing zone) when grown on OBRC. The largest ORF identified is depicted by the thicker line and encompasses 3.033 kb.*

```
MGTSLMIKSTLTGMITAVAAAVFTTSAAFADVPPLTVSGNQVLSGGEAKSFAGNSFFWSNTGWGQERFYN
        10        20        30        40        50        60        70

AETVRWLKDDWNATIVRAAMGVDFDGSYIPEHEDADPEGNVARVRALVDAAIAEDMYVIIDFHTHHAEDY
        80        90       100       110       120       130       140

QAESIEFFEEMATLYGGYDNVIYEIYNEPLQISWDNVIKPYAESVIGAIRAIDPDNLIIVGTPTWSQDVD
       150       160       170       180       190       200       210

AAARNPITSYSNIAYTLHFYAGTHGSWLRDKARNAMNSGIALFVTEWGTVNADGDGAPAVNETQQWMDFL
       220       230       240       250       260       270       280

KQNNISHLNWSVSDKLEGASIVQPGTPISGWNASDLTASGTLVKNIVSNWGTTIGNGSSSSSSSSSSSSS
       290       300       310       320       330       340       350

SSSSSSSSSSSSSSSSSSSSSGSTGGGNCAGVNVYPNWTARDWSGGAYNHANAGDQMVYQNSLYRANWYTNS
       360       370       380       390       400       410       420

VPGSDASWTSLGACGGNGSTTSSSSSSSSSSSSSSSSSSSSSSSTGGGSSSSSSSSSSSSSSSSSSSSSSSTGGGQC
       430       440       450       460       470       480       490

TEVCNWYGQGTYPLCNNTSGWGWENNQSCIGRQTCESQNGGAGGVVSNCTGSSTSSSSSSSSSSSSSSSS
       500       510       520       530       540       550       560

SSSSSSSSGTGSSTSSSSSSSSSSSSSSSTGSSGMPGPRVDNPFAAAQKWYVNPMWSASAANEPGGSVIANE
       570       580       590       600       610       620       630

PSFVWMDRIGAIEGPADGMGLRDHLNEALAQGADLFMFVVYDLPNRDCAALASNGELRISEDGFNIYKSD
       640       650       660       670       680       690       700

YIAPIVEIISDPAYAGIKIAAVIEVDSLPNLVTNLSEPDCQEANGPGGYRDGIRHAITELGKIPNVYSYV
       710       720       730       740       750       760       770

DIAHSGWLGWSDNFAQGVNLIYEVVANLGSGINPIAGFVSNSANYTPVEEPFLPDANLQVGGQPVRSSDF
       780       790       800       810       820       830       840

YEWNSYLAEKPFVTDWRSAMISKGMPSSIGMLIDTARNGWGGPERPTAQSTSNNLNTFVNESRIDRREHR
       850       860       870       880       890       900       910

GNWCNQPGGVGYRPTAAPSPGIDAYVWVKPQGESDGVSDPNFEIDPNDPNKQHDPMCDPFASNSSNSAYG
       920       930       940       950       960       970       980

TGAMPNAPHAGRWFPEAFQLLLENAYPPIN-
       990      1000      1010
```

Figure 2. *Amino acid sequence of CelA from T. turnerae.*

```
                 10          20          30          40
Tt Cat 1  M G T S L M I K S T L T G M I T A V A A A V F T T S A A F A D V P P L T V S G N   40
Ps cel E  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - D V A P L S V Q G N   10
Ech prez  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - S V N G N    5

                 50          60          70          80
Tt Cat 1  Q V L S G G E A K S F A G N S F F W S N T G W G Q E R F Y N A E T V R W L K D D   80
Ps cel E  K I L A N G Q P A S F S G M S L F W S N T E W G G E K Y N A Q V V S W L K S D   50
Ech prez  K I Y A G E K A K S F A G N S L F W S N N G W G G E K F Y T A D T V A S L K K D   45

                 90         100         110         120
Tt Cat 1  W N A T I V R A A M G V D F D G S Y I P E H E D A D P E G N V A R V R A L V D A   120
Ps cel E  W N A K L V R A A M G V E D E G G Y L - - - - - T D P A - N K D R V T Q V V D A   84
Ech prez  W K S S I V R A A M G V Q E S G G Y L - - - - - Q D P A G N K A K V E R V V D A   80

                130         140         150         160
Tt Cat 1  A I A E D M Y V I I D F H T H H A E D Y Q A E S I E F F E E M A T L Y G G Y D N   160
Ps cel E  A I A N D M Y V I I D W H S H N A H Q Y Q S Q A I A F F Q E M A R K Y G A N N H   124
Ech prez  A I A N D M Y A I I G W H S H S A E N N R S E A I R F F Q E M A R K Y G N K P N   120

                170         180         190         200
Tt Cat 1  V I Y E I Y N E P L Q I S W D N V I K P Y A E S V I G A I R A I D P D N L I I V   200
Ps cel E  V I Y E I Y N E P L Q V S W S N T I K P Y A Q A V I A A I R A I D P D N L I I V   164
Ech prez  V I Y E I Y N E P L Q V S W S N T I K P Y A E A V I S A I R A I D P D N L I I V   160

                210         220         230         240
Tt Cat 1  G T P T W S Q D V D A A A R N P I T S Y S N I A Y T L H F Y A G T H G S W L R D   240
Ps cel E  G T P T W S Q D V D V A A N D P I T G Y Q N I A Y T L H F Y A G T H G Q Y L R D   204
Ech prez  G T P S W S Q N V D E A S R D P I N A - K N I A Y T L H F Y A G T H G E S L R N   199

                250         260         270         280
Tt Cat 1  K A R N A M N S G I A L F V T E W G T V N A D G D G A P A V N E T Q Q W M D F L   280
Ps cel E  K A Q T A L N R G I A L F V T E W G S V N A N G D G A V A N S E T N A W V S F M   244
Ech prez  K A R Q A L N N G I A L F V T E W G T V N A D G N G G V N Q T E T D A W V T F M   239

                290         300         310
Tt Cat 1  K Q N N I S H L N W S V S D K L E G A S I V Q P G T P I S G - W N A S D L T A S   319
Ps cel E  K T N H I S N A N W A L N D K V E G A S A L V P G A S A N G G W V N S Q L T A S   284
Ech prez  R D N N I S N A N W A L N D K N E G A S T Y Y P D S K - - - - - - - N L T E S   271

           320         330
Tt Cat 1  G T L V K N I V S N W G T T I G N G   337
Ps cel E  G A L A K S I I S G W P S Y N T   300
Ech prez  G K K V K S I I Q S W P - Y K A G S   288
```

Figure 3. Comparison of the T. turnerae Cat 1 region of CelA and the catalytic regions of P. fluorescens CelE (22) and E. chrysanthemi CelZ (23).

identical residues with the CelE and CelZ catalytic domains, respectively. This region also contains the family 5 signature sequence (*24*), VIYEIYNEPL, between residues 161 to 170 of the whole protein (Fig. 2).

The carboxy-terminal end (424 residues) of the CelA polypeptide, which is designated Cat 2 (Fig. 4), exhibits sequence similarity with exo-1,4-glucanases belonging to glycoside hydrolase family 6 (*20, 21*). Closest sequence similarities are with the catalytic domains of *Thermomonospora fusca* cellobiohydrolase E3 (*25*) *and Cellulomonas fimi* cellobiohydrolase A (*26*). Pairwise comparisons revealed 46% and 49% identical residues between the CelA Cat 2 domain and the catalytic domains of the E3 and A exoglucanases, respectively. The family 6 signature sequence (*24*), VVYDLPNRDCAALASNG, is evident between residues 669 and 685 of the whole protein (Fig. 2).

Two regions (residues 376 to 439 and 484 to 542, Fig. 2) that exhibit sequence similarity with cellulose-binding domains exist between the Cat 1 and Cat 2 domains. The first, which is designated CBD 1 (Fig. 5A), contains 58% identical residues with the family V cellulose-binding domain of *E. chrysanthemi* CelZ (*23*). The second region, designated CBD 2 (Fig. 5B), contains 36% to 56% residues identical to those of several family X cellulose-binding domains of various *P. fluorescens* xylanses and cellulases, as well as *Cellvibrio mixtus* XynA (*27*).

Immediately following the Cat 1 domain, between the CBD 1 and CBD 2 domains, and preceding the Cat 2 domain are highly serine-rich sequences (Fig. 2). These sequences range in length from 31 to 49 amino acids, with 86% and 100% of the residues being serine. Linker regions of polysaccharide-degrading enzymes are typically rich in either proline or serine and threonine (*28*). The three *T. turnerae* linker regions are most similar to one found in *P. fluorescens* CMCase (aa 608 to 666), which contains about 75% serine (*29*).

Southern Analysis of the cel A Gene. Although the occurrence of a bifunctional gene is not unique among polysaccharide-degrading genes, the majority of genes studied to date encode a single catalytic domain. Nevertheless, to ensure that the bifunctional nature of the celA gene was not an artifact of cloning, Southern hybridization analysis was performed. Genomic DNA was digested with six restriction endonucleases and hybridized to DNA probes to the N-terminal (residues 64 to 166) and C-terminal (residues 811 to 932) regions of the predicted protein. When restriction endonucleases that do not contain sites within the open reading frame (Bam H1, Eco R1, Sph 1, and Sma 1) are employed, one would expect both probes to display a single band of identical molecular weight. If the cloned celA gene is the result of a random ligation event, then the two probes should display single bands of different molecular weight. If the genome contains multiple copies of the celA gene, the probes should display multiple bands of identical molecular weights. A fourth possibility is that the *T. turnerae* genome contains a single copy of the celA gene, but one of its catalytic domains is present in multiple copies. Then, one of the probes should show a single band and the other probe should show multiple bands, one of which is the same size as the first probe. The results (Fig. 6) reveal that the N-terminal probe hybridizes to only one genomic fragment. However, the C-terminal probe hybridizes to two fragments, one of which is the same size as the fragment that the N-terminal probe hybridized with. The exception to this was the Bam H1 digest,

```
              10        20        30        40        50        60        70
Tt Cat 2  TGSSGMPGPRVDNPFAAAQKWYVNPMWSASAANEPGG-----------SVIANEPSFVWMDRIGAIEG--  57
C fimi    ------APVHVDNFYAGAVQ-YVNFTWAASVNAAAGRQSADPALAAKMRTVAGQFTAVWMDRISAILGN-   62
T fusca   N-----PGEKVDNPFEGA-KLYVNFVWSAKAAAEPGG-----------SAVANESTAVWLDRIGAIEGND   53

              80        90       100       110       120       130       140
Tt Cat 2  --PADGMGLRDHLNEALAQ--GAD---LFMFVVYDLPNRDCAALASNGELRISEDGFNIYKSDYIAPIVE  120
C fimi    ---ADGNGLKFHIDNAVAQQKAAGVPLVFNLMIYDLFGRDCFALASNGEIPATDAGLARYKSFYIDPIAD  129
T fusca   SPTTGSMGLRDHLEEAVRQ--SGGDPLTIQVMILNLFGRDCAALASNGEI--GPFELDRYKSFYIDPIAD  119

             150       160       170       180       190       200       210
Tt Cat 2  IISD-PAYAGIKIAAVIEVDSLPNLVTNLSEPD----CQEANGPGGYRDGIRHAITELGKIPNVYSYVDI  185
C fimi    LL-DNFEYESLRIAATIEPDSLPNLLTNISEFA----CQ--QAAPYYHQGVKYALDKLLHAIPNVYNYIDI  192
T fusca   IMWDFADYENLRIVAIIEIDSLPNLVTNVGGNGGTELCAYMKQNGGYVNGVGYALRKLGFIPNVYNYIDA  189

             220       230       240       250       260       270       280
Tt Cat 2  AHSGWLGWSDNFAQGVNLIYEVVANLGSGINPIAGFVSNSANYTPVEEPFLPDANLQVGGQPVRSSDFYE  255
C fimi    GHSGWLGWDSNAGPSATLFAEVAKSTTAGFASIDGFVSDVANTTHLEEPLLSDSSLTINNTFIRSSKFYE  262
T fusca   AHHGWIGWDSNFGPSVDIFYEAANASGSTVDYVHGFISNTANYSATVEFYL-DVNGTVNGQLIRQSKWVD  258

             290       300       310       320       330       340       350
Tt Cat 2  WNSYLAEKPFVTDWRSAMISKGMPSSIGMLIDTARNGWGGPERPTAQSTSNNLNTFVNESRIDRREHRGN  325
C fimi    WNFDFDEIDYTAHMHRLLVAAGFPSSIGMLVDISRNGWGGFNRPTSITASTDVNAYVDANRVDRRVHRGA  332
T fusca   WNQYVDELSFVQDLRQALIAKGFRSDIGMLIDISRNGWGGFNRPTGPSSSTDLNTYVDESRIDRRIHPGN  328

             360       370       380       390       400       410       420
Tt Cat 2  WCNQPG-GVGYRPTAAPS----PGIDAYVWVKPQGESDGVSDPNFEIDPNDPNKQHDPMCDF-FASNSSN  389
C fimi    WCNPLGAGIGRFFEATPSGYAASHIDAFVWLKFPGESDGAST---DI-PNLQGKRFDRMCDFTFVS-PKL  397
T fusca   WCNQAGAGLGERPTVNFA----PGVDAYVWVKFPGESDGASE---EI-PNDEGKGFDRMCDFTYQGNARN  390

             430       440       450
Tt Cat 2  SAYGTGAMPNAPHAGRWFPEAFQLLLENAYPPIN.                                      424
C fimi    NNQLTGATPNAFLAGQWFEEQFVTLVKNAYFVIG                                       431
T fusca   GNNPSGALPNAFISGHWFSAQFRELLIANAYPF--L                                     423
```

Figure 4. Comparison of the *T. turnerae* Cat 2 region of CelA and the catalytic domains of *T. fusca* E3 (25) and *C. fimi* cellobiohydrolase A (26).

A

```
                10        20        30        40        50        60
Tt CBD 1   NCAGVNVYPNWTARDWSGGAYNHANAGDQMVYQNSLYRANWYTNSVPGSDASWTSLGACGGNGS   67
Ech prez   DCANANVYPNWVSKDWAGGQPTHNEAGQSIVYKGNLYTANWYTASVPGSDSSWTQVGSC----N   60
```

B

```
                10        20        30        40        50        60
Tt CBD2    STGGGQCTEVCNWYGQGTYPLCNNT-SGWGWENNQSCIGRQTCESQNGGAGGVVSNCTGS   59
Pfl XynE   AGGNTGGNCQCNWWG-TFYPLCQTQTSGWGWENSRSCISTSTCNSQGTGGGVVC----N   55
Pfl XynA   TPGSSSGNQQCNWYG-TLYPLCVITTNGWGWEDQRSCIARSTCAAQPAPFGIVGS----G   55
Pfl EglB   SSSVLTGAQACNWYG-TLTPLCNNTSNGWGYEDGRSCVARTTCSAQPAPYGIVST----S   55
Pfl EglA   SSSSVSGGLRCNWYG-TLYPLCVTIQSGWGWENSQSCISASTCSAQPAPYGIVGA----A   55
Pfl CelE   VSSAVSG-QQCNWYG-TLYPLC-SITNGWGWENNASCIARATCSGQPAPWGIVGG----S   53
Pfl CelC   SSSVVSGGGQCNWYG-TLYPLCVSITSGWGYENNRSCISPSTCSAQPAPYGIVGG----S   55
Cmi XynA   SSAVTGNACQCNWWG-TRYPLCTNTASGWGWENNTSCITTSTCNSQGAGGGVVC----N   55
```

Figure 5. *Amino acid comparisons of related cellulose-binding domains from various organisms with those (CBD 1, Panel A; CBD 2, Panel B) from T. turnerae CelA.*

Figure 6. *Southern hybridization. Genomic T. turnerae DNA was digested with various restriction endonucleases and separated on a 0.7% agarose gel. The left half of the gel was probed with a fragment prepared from the N-terminal portion of the CelA protein. The right half of the gel was probed with a fragment prepared from the C-terminus of the CelA protein.*

which did not go to completion. This indicated that the celA gene is not an artifact of cloning and that the Cat 2 portion of the gene is duplicated in the genome of *T. turnerae*. Northern hybridization analysis (data not shown) supported this conclusion.

Discussion

The data described above indicate that *T. turnerae* contains a gene that codes for CelA, a multifunctional cellulase composed of at least seven different domains. There appears to be only a single copy of the whole gene in the *T. turnerae* genome. However, the portion of the gene that codes for the exocellulase domain (Cat 2) appears to be duplicated in the genome (Fig. 6).

The *T. turnerae* celA gene codes for a predicted protein of 1010 amino acid residues with a calculated molecular mass of 106,068 Daltons. The N-terminus of the protein (Cat 1) exhibits a high degree of homology to the family 5 endocellulases of *P. fluorescens* CelE (22) and *E. chrysanthemi* CelZ (23). The C-terminus of CelA (Cat 2) is highly homologous to the family 6 exocellulases of *T. fusca* E3 (25) and *C. fimi* cellobiohydrolase A (26). Both of the *T. turnerae* CelA domains contain the signature sequences of their respective families. Interspersed between these catalytic domains are two different cellulose-binding domains (CBD 1 and CBD 2) and three linker regions that are rich in serine.

Bifunctional enzymes are not the norm. Of more than 1,000 described glycoside hydrolases, about 20 genes are reported to code for bifunctional enzymes. The dual activities that bifunctional enzymes display are often, but not always, related. Several different organisms have been reported to produce single proteins which contain the following activities: xylanase and arabinosidase (30, 31, 32, 33); xylanase and endocellulase (34); maltase and glucoamylase (35); sucrase and isomaltase (36, 37, 38); amylopullulanase (α-amylase and pullulanlase) (39); xylanase and lichenase (40); chitanase and either lysozyme (41) or lichenase (42); and mannanase and endoglucanase (43).

To date, six genes have been previously characterized which code for bifunctional enzymes that display both endo- and exocellulase activities. Jung et al. (44) cloned and isolated the gene product of the *T. fusca* E4 gene. Even though the protein appeared to contain only a single catalytic domain, it demonstrated both exo- and endocellulase activities (44, 45). *Bacillus* sp. D04 secretes a bifunctional cellulase that has a molecular mass of only 35 kDa. Yet, from inhibitor studies of the purified enzyme, the exo- and endocellulase activities appeared to exist in separate sites (46). The extreme thermophile, *Caldocellum saccharolyticum*, contains two genes, celA and celB, that code for bifunctional cellulases (47, 48, 49). Both genes contain endo- and exocellulase domains and either one or two cellulose-binding domains that are separated by proline-threonine-rich regions. Zverlov et al. (50) isolated a bifunctional cellulase (CelA) from the thermophilic bacterium, *Anaerocellum thermophilum*. Sequence analysis revealed a large (190 kDa), multidomain protein that consists of 2 distinct catalytic domains homologous to glycoside hydrolase families 9 and 48, and 3 domains homologous to the family III cellulose-binding domain, linked by Pro-Thr-Ser-rich regions. In gross architecture,

the *T. turnerae* CelA protein most closely resembles the *A. thermophilum* CelA and *C. saccharolyticum* CelA and CelB proteins.

Literature Cited

1. Turner, R. D. Ph.D. thesis, Harvard Univ., Cambridge, MA, 1966.
2. Bartsch, P. U.S. National Museum: Washington, DC, 1922, Bulletin No. 122.
3. Spence, L. *Wooden Boat* **1993**, *111*, 66-68.
4. Sellius, G. *Historia naturalis teredinis seu xylophagi marini tubulo-conchoidis speciatim Belgiei, Tragecti as Rhenum*; **1733**, 366 pp.
5. Dore, W. H.; Miller, R. C. *Univ. California Publ. Zool.* **1923**, *22*, 383-400.
6. Deshayes, G. P. *Explor. Sci. d'Alger. Zool.* **1948**, *1*, 35-76.
7. Sigerfoos, C. P. *Bull. Bur. Fish.* **1908**, *37*, 191-231.
8. Waterbury, J. B.; Calloway, C. B.; Turner, R. D. *Science* **1983**, *221*, 1401-1403.
9. Distel, D. L.; Delong, E. F.; Waterbury, J. B. *Appl. Environ. Microbiol.* **1991**, *57*, 2376-2382
10. Greene, R. V.; Griffin, H. L.; Freer, S. N. *Appl. Environ. Microbiol.* **1986**, *52*, 982-986.
11. Gallager, S. M.; Turner, R. D.; Berg, C. J. *J. Exp. Mar. Biol. Ecol.* **1981**, *52*, 63-77.
12. Griffin, H. L.; Freer, S. N.; Greene, R. V. *Biochem. Biophys. Res. Commun.* **1987**, *144*, 143-151.
13. Greene, R. V. *SIM News* **1994**, *44*, 51-59.
14. Sheir-Neiss, G.; Montenecourt, B. S. *Appl. Microbiol. Biotechnol.* **1984**, *20*, 46-53.
15. Gadgil, N. J.; Daginawala, H. F.; Chakrabarti, T.; Khanna, P. *Enzyme Microbiol. Technol.* **1995**, *17*, 942-946.
16. Montenecourt, B. S.; Eveleigh, D. E. *Appl. Environ. Microbiol.* **1977**, *34*, 777-782.
17. Iman, S. H.; Greene, R. V.; Hockridge, M. E. *Biotechnol. Lett.* **1993**, *7,* 579-584.
18. Wood, T. M. *Methods Enzymol.* **1988**, *160*, 19-25.
19. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning: a laboratory manual, 2ⁿᵈ ed.*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
20. Henrissat, B.; Bairoch, A. *Biochem. J.* **1996**, *316*, 695-696.
21. Henrissat, B.; Davies, G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 637-644.
22. Hall, J.; Black, G. W.; Ferreira, L. M.; Millward-Sadler, S. J.; Ali, B. R.; Hazlewood, G. P.; Gilbert, H. J. *Biochem. J.* **1995**, *309*, 749-756.
23. Guiseppi, A.; Cami, B.; Aymeric, J. L.; Ball, G.; Creuzet, N. *Mol. Microbiol.* **1988**, *2*, 159-164.
24. *SWISS-PROT Protein Sequence Data Bank,* URL http://www.expasy.ch/cgi-bin/lists?/glucosid.txt
25. Zhang, S.; Lao, G.; Wilson, D. B. *Biochemistry* **1995**, *34*, 3386-3395.
26. Meinke, A.; Gilkes, N. R.; Kwan, E.; Kilburn, D. G.; Warren, R. A.; Miller, R. C., Jr. *Mol. Microbiol.* **1994**, *12*, 413-422.

27. Tomme, P.; Warren, R. A. J.; Miller, R. C., Jr.; Kilburn, D. G.; Gilkes, N. R. In *Enzymatic Degradation of Insoluble Carbohydrates*; Saddler, J. N.; Penner, M. H., Ed.; ACS Symposium Series; American Chemical Society: Washington, DC, 1995; Vol. 618, pp 142-163.

28. Gilkes, N. R.; Henrissat, B.; Kilburn, D. G.; Miller, R. C., Jr.; Warren, R. A. J. *Microbiol. Rev.* **1991**, *55*, 303-315.

29. Hall, J.; Gilbert, H. J. *Mol. Gen. Genet.* **1988**, *213*, 112-117.

30. Luthi, E.; Love, D. R.; McAnulty, J.; Wallace, C.; Caughey, P. A.; Saul, D.; Berquist, P. L. *Appl. Environ. Microbiol.* **1990**, *56*, 1017-1024.

31. Utte, A.; Eddy, C. K.; Keshav, K. F.; Ingram, L. O. *Appl. Environ. Microbiol.* **1991**, *57*, 1227-1234.

32. Sakka, K.; Yoshika, K.; Kojima, Y.; Karita, S.; Ohmiya, K.; Shimada, K. *Biosci. Biotechnol. Biochem.* **1993**, *57*, 268-272.

33. Whitehead, T. R. *Biochim. Biophys.* **1995**, *1244*, 239-241.

34. Ahsan, M. M.; Kimura, T.; Karita, S.; Sakka, K.; Ohmiya, K. *J. Bacteriol.* **1996**, *178*, 5732-5740.

35. Nichols, B. L.; Eldering, J. A.; Avery, S. E.; Hahn, D.; Quaroni, A.; Sterchi, E. E. *J. Biol. Chem.* **1998**, *273*, 3076-3081.

36. Hunziker, W.; Spiess, M.; Semenza, G.; Lodish, H. F. *Cell* **1986**, *46*, 227-234.

37. Chantret, I.; Lacasa, M.; Chevalier, G.; Ruf, J.; Islam, I.; Mantei, N.; Edwards, Y.; Swallow, D.; Rousset, M. *Biochem. J.* **1992**, *285*, 915-923.

38. Chandrasena, G.; Osterholm, D. E.; Sunitha, I.; Henning, S. J. *Gene* **1994**, *150*, 355-360.

39. Hatada, Y.; Igarashi, K.; Okaki, K.; Ara, K.; Hitomi, J.; Kobayashi, T.; Kasai, S.; Watabe, T.; Ito, S. *J. Biol. Chem.* **1996**, *271*, 24075-24083.

40. Flint, H. L.; Martin, J.; McPherson, C. A.; Daniel, A. S.; Zhang, J-X. *J. Bacteriol.* **1993**, *175*, 2943-2951.

41. Heitz, T.; Segond, S.; Kauffmann, S.; Geoffoy, P.; Prasad, V.; Brunner, F.; Fritig, B.; Legrand, M. *Mol. Gen. Genet.* **1994**, *245*, 246-254.

42. Beuno, A.; Vasquez de Aldana, C. R.; Correa, J.; Del Rey, F. *Nucleic Acids Res.* **1990**, *18*, 4248.

43. Gibbs, M. D.; Saul, D. J.; Luthie, E.; Berquist, P. L. *Appl. Environ. Microbiol.* **1992**, *58*, 3864-3867.

44. Jung, E. D.; Lao, G.; Irwin, D.; Barr, B. K.; Benjamin, A.; Wilson, D. B. *Appl. Environ. Microbiol.* **1993**, *59*, 3032-3043.

45. Sakon, J.; Irwin, D.; Wilson, D. B.; Karplus, P. A. *Nat. Struct. Biol.* **1997**, *4*, 810-818.

46. Han, S. J.; Yoo, Y. J.; Kang, H. S. *J. Biol. Chem.* **1995**, *270*, 26012-26019.

47. Saul, D. J.; Williams, L. C.; Love, D. R.; Chamley, I. W.; Bergquist, P. I. *Nucleic Acids Res.* **1989**, *17*, 439.

48. Saul, D. J.; Williams, L. C.; Grayling, R. A.; Chamley, L. W.; Love, D. R.; Berquist, P. L. *Appl. Environ. Microbiol.* **1990**, *56*, 3117-3124.

49. Te'o, V. S.; Saul, D. J.; Berquist, P. L. *Appl. Microbiol. Biotechnol.* **1995**, *43*, 291-296.

50. Zverlov, V.; Mahr, S.; Riedel, K.; Bronnenmeier, K. *Microbiology* **1998**, *144*, 457-465.

Chapter 4

Production of Microbial Cellulases in Transgenic Crop Plants

B. S. Hooker[1], Z. Dai[1], D. B. Anderson[1], R. D. Quesenberry[1], M. F. Ruth[2], and S. R. Thomas[2]

[1]Pacific Northwest National Laboratory, P.O. Box 999, MSIN: K2–10, Richland, WA 99352
[2]National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO 80401

The expression of *Acidothermus cellulolyticus* endoglucanase (E1) gene and *Trichoderma reesei* cellobiohydrolase (CBH1) gene in transgenic tobacco (*Nicotiana tabaccum*) and potato (*Solanum tuberosum*) was examined in this study. CBH1 and E1 produced in transgenic tobacco plants were found to be biologically active and to accumulate in leaves at levels of up to 0.05% and 2.6% of total soluble protein, respectively. The biochemical characteristics of plant-based recombinant E1 enzyme were similar to those of natural E1 purified from bacterial culture. In addition, plant-derived E1 was resistant to plant proteolysis at different developmental stages. Transgenic plants exhibited normal growth and developmental characteristics with photosynthetic rates similar to those of untransformed SR1 tobacco plants. Based on E1 activity and E1 protein accumulation in leaf extracts, E1 expression in potato is much higher than that measured in transgenic tobacco bearing the same transgene constructs. E1 expression under the control of the RbcS-3C promoter was specifically localized in leaf tissues, while E1 was expressed in both leaf and tuber tissues under the control of the constitutive Mac promoter. This suggests dual-crop applications in which potato vines serve as enzyme production "bioreactors" while tubers are preserved for culinary applications. Economic analysis demonstrated that the cost of large-scale E1 production, based on potato vine-based expression, could be as low as $1.40/kg enzyme.

55

Cellulose, an unbranched, linear polymer of D-glucose residues linked by β-1,4 glycosidic bonds, is a critical structural component of the cell wall of higher plants. It is the most abundant organic macromolecule on earth, accounting for over one half of the carbon in the biosphere (*1*). In addition, cellulose is the main component of primary and secondary biomass wastes generated through various domestic, commercial and industrial activities, including traditional industrial activities directly associated with forestry and agriculture. Until recently, most of these cellulosic wastes were not perceived as useful materials and were disposed in landfills (*2*), but these activities are becoming more restricted. Effective, alternative means for the recycling of cellulosic residues are being explored. Conversion of the polysaccharide components of lignocellulosic residues to fuel ethanol and other feedstock chemicals provide a promising alternative to traditional landfilling and combustion methods for disposal of cellulosic wastes (*3,4*) Two features of the cellulose molecule make it extremely resistant to enzymatic hydrolysis. The first is that long, unbranched cellulose chains formed by the β-1,4-glycosidic linkage are completely insoluble in water. Second, in plant tissues, cellulose polymers adhere strongly to one another in overlapping parallel arrays by intermolecular hydrogen bonds to form cellulose microfibrils. Nevertheless, there are many fungal and bacterial species that are able to break down cellulose. These organisms produce a large ensemble of different cellulases acting in synergy on the crystalline cellulose surface to hydrolyze the β-1,4-glycosidic linkages. There are at least three major groups of cellulases involved in the hydrolysis of the glycosidic bonds of cellulose to yield a single product, glucose: β-1,4-endoglucanase (or endo β-1,4-glucanohydrolase; E.C. 3.2.1.4), β-1,4-exoglucanase (or cellobiohydrolase; E.C. 3.2.1.91), and β-D-glucosidase (E.C. 3.2.1.21) (*2*). The β-1,4-endoglucanases cleave cellulose at random along the polysaccharide chain, producing new reducing and non-reducing chain ends. Two types of β-1,4-exoglucanases generally remove cellobiosyl units from either the reducing or the non-reducing end of a cellulose strand. The β-D-glucosidase hydrolyses cellobiose units to produce two glucose molecules. The gene of cellobiohydrolase I (CBH I), has been isolated from the filamentous fungus, *Trichoderma reesei* (*5-7*). The encoding region of the mature CBHI gene consists of 1485 bp, which encodes a protein of approximately 67 kD (*5*).

In addition, the gene of endoglucanase I (E1) has been isolated from *Acidothermus cellulolyticus*, a thermotolerant, cellulolytic actinomycete bacterium (*8,9*). Native purified E1 enzyme is extremely thermostable and analysis of the amino acid sequence derived from the sequence of the E1 gene shows that it is a member of glycosyl hydrolase family 5, subfamily 1. The mature E1 protein with 521 amino acids (1563 bp) consists of an N-terminal catalytic domain, a proline/serine/threonine-rich linker region, and a C-terminal cellulose binding domain with the mature protein size of about 72kD (*10*).

Typically, industrial enzymes such as, cellulases and xylanase, are produced primarily by fermentative processes. However, the cost of commercially available cellulose degrading enzymes is prohibitively high to be considered for "low value" applications such as the conversion of lignocellulosic feedstocks to fuel ethanol. Agricultural systems provide the most cost effective means for production of biomass. Plant biomass can be produced in bulk quantities using limited and widely practiced

infrastructure. Enzymes such as those used in food, feed, or processing industries, can be produced economically in genetically engineered plants (*11*). In an effort to demonstrate technical feasibility and to reduce the cost of production of cell wall degrading enzymes, alternative production in transgenic plants is currently being studied (*12-15*).

So far, heterologous expression of the E1 gene has been studied in bacteria and other microorganisms (*16-18*). In this study we have examined the expression of CBH1 and E1 in transgenic tobacco and potato plants. In addition, we have completed a preliminary assessment of the economics of cellulase production in crop plant residues (i.e., potato vines).

Materials and Methods

Bacterial Strains, Plant Material, Plant Transformation, And Plant Growth Conditions. *Escherichia coli* strains MC1000 and JM83 (*ara, leu, lac, gal, str*) were used as the hosts for routine cloning experiments. *Agrobacterium tumefaciens* LBA-4404 containing the Ach5 chromosomal background and a disarmed helper-Ti plasmid pAL-4404 (19) was used for transformation of tobacco (*Nicotiana tobacum* L. Cv Petit Havana SR1) and potato (*Solanum tuberosum* L. Cv Desiree). Transgenic plants were obtained by the co-cultivation method (*20*) using tobacco leaf discs grown aseptically on Murashige and Skoog agar (MS) medium supplemented with 3% sucrose and appropriate levels of plant growth regulators (21). MS agar containing 2 mg L^{-1} α-naphthaleneacetic acid (NAA) and 0.5 mg L^{-1} 6-benzylaminopurine (BA) was used for callus induction and the same media containing no NAA and 0.5 mg L^{-1} BA was used for shoot induction. The regenerated shoots were maintained on MS agar for further rooting. Kanamycin was employed in media throughout the regeneration process at 100 mg L^{-1}. Kanamycin-resistant transformants were grown and selected on MS medium in a tissue-culture room and transferred successively to shooting and rooting media for organogenesis. The transformants were transferred to a growth room described below for plant materials of primary transformants and seed production. Tobacco seeds were collected from primary transformants after self-fertilization. Seeds were germinated on MS agar medium containing 100 mg L^{-1} of kanamycin, and healthy kanamycin-resistant plants were grown in a growth room under a 14 h light (25-28°C, 60% relative humidity)/10 h dark (22°C, 70% relative humidity) cycle. Irradiance, in the growth room provided by six high pressure metal halide lamps (Philips, USA), was 350 to 500 μmol quanta $m^{-2} s^{-1}$ at the plant canopy.

Recombinant DNA Techniques. Standard procedures were used for recombinant DNA manipulation (*22*). The plasmid pB210-5A containing the entire cDNA of exo-cellobiohydrolase I (CBH I) isolated from the fungus *T. reesei* cDNA library (*5*) was obtained from the National Energy Renewable Laboratory in Golden, Colorado. The 1500 bp Pst I/ Xho I fragment (containing the entire CBH I coding region) was excised from pB210-5A and treated with DNA polymerase I (Klenow enzyme) to produce blunt ends. The Klenow fragment was further inserted into the Hpa I site in

the poly-linker of binary vector pGA643 (*23*) in the sense orientation relative to the 35S cauliflower mosaic virus promoter and T5-T7 transcription terminator.

Plasmid pMPT4-5 containing a genomic clone of the endoglucanase (E1) gene isolated from *A. cellulolyticus* genome library (GenBank Accession No. U33212) was also obtained from the National Energy Renewable Laboratory in Golden, Colorado (*18*). The 1566 bp fragment (containing the mature-peptide coding region) was isolated from pPMT4-5 by PCR using the primers E1-A3 (5'-CTA ATG CAT GCG GGC GGC GGC TAT TGG CAC ACG AGC-3') and E1-B (5'-CTT AGA TCT GAG CTC TTA ACT TGC TGC GCA GGC GAC TGT CGG TG-3'). The mature E1 coding sequence was fused downstream of either the hybrid Mac promoter (*24*) or the Rubisco small subunit (RbcS-3C) promoter (*25*), in-frame with various transit peptide coding regions, as shown in Figure 1, to create four unique expression vectors. In one vector, a synthetic leader sequence derived from alfalfa mosaic virus RNA4 was inserted downstream of the RbcS-3C promoter fragment. It has reported that mRNA translation can be enhanced by this viral leader sequence (*26,27*). The E1 gene was terminated by the transcriptional termination sequences of either the T7 and T5 genes of the octopine type Ti plasmid (for Rubisco-based cassettes) or the mannopine synthetase gene (for Mac-based cassettes). Resulting binary vectors were mobilized into *A. tumefaciens* LBA 4404 by the freeze-thaw method (*28*) and further used in *Agrobacterium*-mediated leaf disc transformation.

RNA Preparation and Gel Blot Analysis. Mature leaf tissues were harvested and immediately frozen in liquid nitrogen. Total RNA was extracted from 1 to 2 g frozen leaf tissues in a prechilled mortar and pestle with about 5 volume (v/w) of GTC buffer (4 M guanidine isothiocyanate, 25 mM sodium citrate, 0.5% sarkosyl (W/V), and 0.1 M mercaptoethanol). After homogenized, 0.5 mL g^{-1} 2 M sodium acetate, pH 4.0 was added to the mixtures and mixed well. The preparation was extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and centrifuged to separate the liquid phases. The aqueous phase containing RNA was transferred to a new centrifuge tube and RNA was precipitated by the addition of an equal volume of isopropanol at -20°C. After centrifugation, the pellets were suspended in RNase-free water and further precipitated with an equal volume of 4M lithium chloride. The RNA pellets were resuspended in RNase-free H_2O, extracted once with chloroform and precipitated by addition of sodium acetate, pH 5.0, to 0.2 M and an equal volume of isopropanol. The total RNA concentration was quantified spectrophotometrically. Fifty µg of total RNA samples were used for gel blot analysis as described by Dai and An (*29*).

Enzyme Extraction and Assays. The third or fourth leaves from the shoot apex were used for enzyme extraction. Leaf samples were harvested at 2 to 3 h into the light period. Leaf tissues were sectioned into about 1 cm^2 leaf discs and pooled together. About 0.1 g of leaf discs was collected and placed in ice cold 1.5 mL microcentrifuge tubes and kept on ice. E1 enzyme in leaf tissues was extracted with a pellet pestle (Kontes Glass Co, Vineland, NJ) and 5 volumes (v/w) of ice cold extraction medium containing 80 mM MES (pH 5.5), 10 mM β-mercaptoethanol, 10 mM EDTA, 0.1% Sarcosyl, 0.1% Triton X-100, 1 mM PMSF, 10 µM Leupeptin, and 1µg mL^{-1} other

(A)
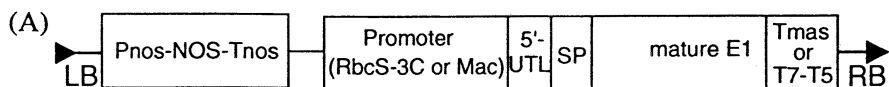
(B)

Promoter	5'-UTL	SP	cDNA	Terminator	Transformant Designation
Mac	mas	Native E1	Mature E1	Tmas	mm-E1
Mac	mas	Apoplast	Mature E1	Tmas	mm-apo
RbcS-3C	AMV	Chloroplast	Mature E1	T7-T5	ra-chl
RbcS-3C	RbcS-3C	Vacuole	Mature E1	T7-T5	rr-vac

Figure 1. A, Schematic of the t-DNA region of binary vectors used for E1 plant transformation. In all cased, neomycin phosphotransferase II (nptII) was used as the selectable marker, with corresponding nopaline synthase (nos) promoter and terminator elements. B, Table of genetic elements used for the expression of E1 in four separate expression vectors.

protease inhibitors (aprotinin, pepstin A, chymostatin). The crude extract was centrifuged at 15,000 g for 10 min at 4°C. The supernatant was transferred to a new tube and used for protein determination, enzymatic analysis, polyacrylamide gel electrophoresis, and Western blot analysis.

E1 enzyme activity was measured at 55°C and pH 5.5 in a final volume of 1 mL. The reaction mixture for E1 activity analysis contained 80 mM MES, pH 5.5, 1 mM EDTA, 1 mM DTT, and 5 to 10 μL enzyme extract. The enzymatic reaction was initiated by adding the substrate, 4-MUC, into the pre-warmed reaction mixture to a final concentration of 2 mM. Aliquots (100μL) of the reaction mixture were removed at 15, 30, and 45 minutes after the start of the reaction and transferred to 1.9 mL 0.2 M Na_2CO_3 buffer to terminate the reaction. The fluorescent reporter moiety, 4-methylumbelliferone (MU), released from 4-MUC by the action of E1 has a peak excitation of 365 nm (UV) and a peak emission of 455 nm (blue). Emission of fluorescence from the mixture was measured with a Hoefer DyNA Quant 200 Fluorometer using 365 nm and 455 nm exitation and emission filters, respectively. Enzyme activities were expressed on a leaf fresh weight or on a total leaf soluble protein basis.

SDS-PAGE and Western Blot. Soluble protein was extracted from 0.1 g leaf tissue in 0.5 mL cold extraction buffer as described above in the "Enzyme extraction and assays" section. The concentration of soluble protein was determined by the method of Bradford (*30*) with BSA as the standard. Leaf extracts were clarified by centrifugation (20,000 g for 10 min at 4°C). Electrophoretic analysis of polypeptides was performed in a 7.5 to 15% (w/v) full size linear gradient polyacrylamide gel containing 0.1% SDS, stabilized by a 5 to 17% (w/v) linear sucrose gradient (*31*) or 4 to 20% (w/v) mini precast gel from Bio-Rad Laboratories (Hercules, CA). Protein samples were prepared for electrophoresis by mixing four volumes of protein extract with one volume of protein loading buffer (250 mm Tris-HCl, pH 7.5, 10% SDS, 28% glycerol, 38% β-mercaptoethanol, and a trace amount bromphenol blue) and then heated at 90°C for 3-5 min. Denatured samples were centrifuged at 10,000g for 2 min. CBH1 and E1 protein samples were separated by electrophoresis and then electrophoretically transferred onto nitrocellulose membranes (BA-S85; Schleicher & Schuell, Keene, NH), as described previously (*31*). Protein blots were probed with CBH1- and E1-specific, affinity-purified mouse monoclonal antibody (1:250 dilution), generously provided by R. Nieves (NREL, Golden, CO). Antibody was detected by a color development kit (Bio-Rad, Hercules, CA) using a goat anti-mouse secondary antibody (IgG) conjugated with alkaline phosphatase (Pierce, Rockford, IL). Color development was performed in the dark for 30 min to 4 h. Mature CBH1 and E1 protein used as positive controls in these experiments were generously provided by W. Adney (NREL, Golden, CO).

The amount of CBH1 and E1 expressed in leaf tissues was estimated by densitometry analysis. The protein blot bands were scanned with a Hewlett Packard ScanJet 6100C Scanner (Hewlett Packard Inc, Palo Alto, CA). The imaging data were analyzed with a DENDRON 2.2 program (Solltech, Inc., Oakdale, IA). A series of known amounts of CBH1 and E1 protein was used as a standard for estimating the production of CBH1 and E1 protein in transgenic plants.

Immunocytochemistry. Young leaf tissues of transgenic E1-producing and normal tobacco plants were cut into about 1 mm pieces with a razor blade and immediately placed into a fixative containing 1.25% glutaraldehyde, 3% (w/v) paraformaldehyde, and 50 mm PIPES, pH 7.2. The tissues were fixed overnight at 4°C. Fixed leaf fragments were then washed with 50 mm PIPES buffer (pH 7.2), dehydrated in an ethanol series, and gradually infiltrated in LR white resin (Electron Microscopy Sciences, Fort Washington, PA), ending with three changes of pure resin. Samples were polymerized for 15 h at 60°C and sectioned 100 nm thickness on a Leica Ultracut microtome. Ultrathin serial sections from leaf tissues collected on nickel grids were blocked in 1% (w/v) BSA in PBS. The sections were then incubated in a drop of mouse anti-E1 monoclonal antibody diluted 1:10 in blocking buffer for 2 hr at room temperature. The sections were washed four times for 10 min in PBS. The sections were further incubated for 2 hr at room temperature with goat anti-mouse IgG conjugated to 10 nm gold particles (Sigma, St. Louis, MO) which was diluted 1:50 in blocking buffer. The sections were washed twice for 5 min in blocking buffer and twice for 5 min in PBS. Sections were stained with potassium permanganate and uranyl acetate and subsequently evaluated and photographed with a Hitachi transmission electron microscope.

Effect Of Temperature and pH On E1 Activity. For the effect of different temperatures on E1 activity, a series of gradient temperatures was generated by a RoboCycler 96 gradient temperature cycler with the hot top assembly. The reaction mixture was the same as above. The E1 extract from transgenic plants was mixed well with the reaction mixture and a 100μL of the resulting mixture was transferred into PCR microtubes (250μL) just before incubating at the selected temperature. The RoboCycler 96 gradient temperature cycler can generate 11°C gradient temperatures at each gradient temperature setting between 47 and 98°C with an error of ± 0.1°C. Four sets of gradient temperatures generated by the RoboCycler 96 gradient temperature cycler were 47°C to 58°C, 61°C to 72°C, 75°C to 86°C, and 86°C to 97°C. Three replicate samples were taken from 3 different transgenic plants for each temperature point. The samples in the PCR microtubes were incubated at the specific temperature for 15 min and immediately quenched by immersing in ice cold water. In addition, 100 μL of 0.2 M Na_2CO_3 was added and mixed by vortexing to stop the reaction completely.

A series of reaction mixtures were prepared with different pH buffers (pH 3.5, 4.0, 4.5, 4.75, 5.0, 5.25, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0). Buffers below 6.0, between 6 and 7, and above 7 were phosphate-citrate buffer, phosphate buffer, and Tris-HCl buffer, respectively. The reaction mixtures contained 80 mM buffer, 1 mM EDTA, 1 mM DTT, and 5 μL enzyme extract and 2 mM 4-MUC. All reaction mixtures were kept on ice cold H_2O until the reaction was initiated at 55°C. The 100 μL reaction mixture was removed at 10, 20, 30 min intervals and immediately quenched by mixing with 1.9 mL 0.2 M Na_2CO_3.

Effect Of Leaf Dehydration at Room Temperature. The 4th to 6th leaf counted from the top of E1-bearing plants was harvested and cut in about 1 cm^2 leaf discs. Approximately 5 to 10 leaf discs were taken per sample, and allowed to dehydrate at

ambient conditions up to 100 h. The samples were harvested at different intervals and stored at -70°C. Then all samples were extracted with enzyme extraction buffer and 5 µL extract per sample was used for E1 activity measurement. The amount of total soluble protein was determined as described above. Enzyme activity was calculated on the basis of total soluble protein in each sample or on the basis of the original expanded leaf area determined just before dehydration treatment.

Photosynthesis Measurement. Rates of CO_2 assimilation were measured on the third or fourth intact leaf from the apex of E1-bearing plants using an Analytical Development Co. (ADC) infrared gas analyzer (225-MK3; PP System, Haverhill, Mass., USA) and a Bingham Interspace (Hyde Park, Utah, USA) model BI-6-dp computer controller system (*32*). This was operated as an open system where a given partial pressure of CO_2 is passed through the sample cells (in line with the leaf enclosed in a cuvette) and the reference cell; the rate of CO_2 removal by photosynthesis was compensated for by a controlled rate of injection of CO_2 from a high CO_2 source. The leaf cuvette contained a dewpoint sensor for measuring humidity and a copper constantan thermocouple for monitoring leaf temperature. The rate of CO_2 assimilation was directly calculated from gas-exchange measurements according to von Caemmerer and Farquhar (*33*).

Results and Discussion

Exoglucanase CBH1 Expression in Tobacco. Using SDS-PAGE and protein immunoblotting, CBHI expression in selected transgenic tobacco calli and leaf tissues was examined (Figures 2 and 3). Results for leaf tissue and calli samples, shown in Figures 2 and 3, respectively, clearly indicate CBHI enzyme accumulation as contrasted with untransformed controls (SR-1 plants and NT-1 calli). By comparing the protein immunoblot staining intensity of plant produced CBHI protein to that of a known amount of *T. reesei*-derived CBHI, the maximum plant-based CBHI expression is estimated at 0.05% of total soluble protein (Figure 2). Similarly, transformed calli showed maximum levels of CBHI expression of about 0.1% total soluble protein (Figure 3). The results from Figures 2 and 3 also show that CBHI protein produced by transgenic tobacco leaves and calli possesses a similar size to that of *T. reesei* CBHI (*5*), suggesting proper transcription/translation in plant cells without extensive proteolysis.

In order further to determine correct processing of CBHI enzyme in tobacco leaf tissues and calli, the 4-methylumbelliferyl-b-D-cellobioside assay was performed on protein extracted with phosphate-citrate protein extraction buffer. Levels of cellobiohydrolysing activity measured in transgenic plants, used in earlier immunoblot assays, are shown in Figure 4. All transformed plants tested possessed at least two times or greater activity than that of control SR tobacco plants. The highest CBHI activity achieved was observed in transformant 1073-12, which reached a level of 66.1 μmol h^{-1} g^{-1} total soluble protein.

The CBHI activity in selected transgenic calli was also determined using a similar method. More than 80% of transgenic calli examined possessed higher CBHI
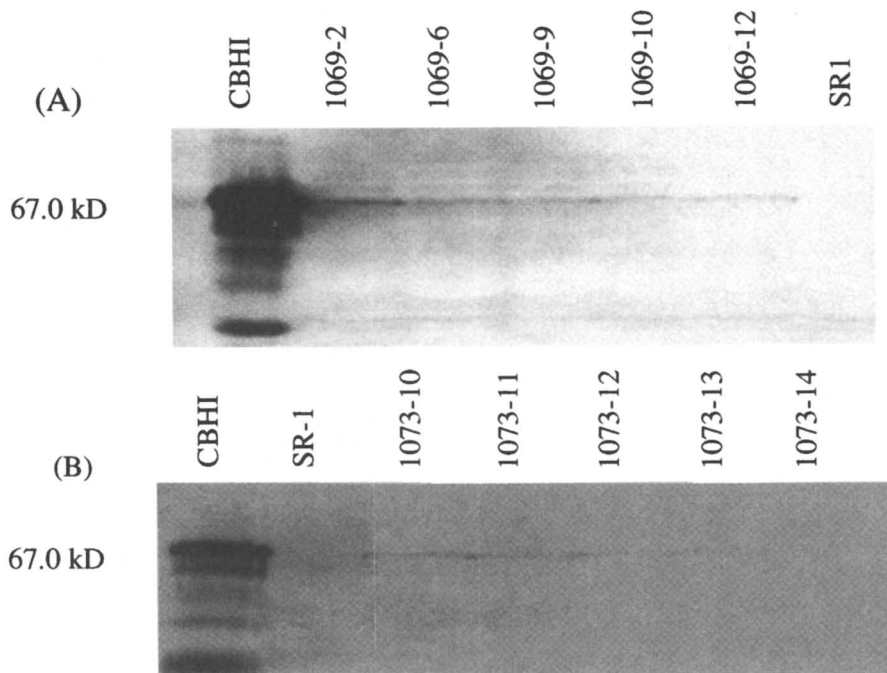
Figure 2. *Detection of CBHI protein in transgenic tobacco leaf tissues. Equal amounts of protein (50 µg) were loaded on each lane with the exception of the CBHI standard derived from T. reesei. In panel (A), 3µg of affinity-purified CBHI from T. reesei was used as a positive control (lane 1) and five independent transgenic lines originating from Agrobacterium strain 1069 were selected (lanes 2 through 6). A negative control (wild type SR-1 tobacco) is shown in lane 7. In panel (B), 0.5µg of affinity-purified CBHI from T. reesei was used as a positive control (lane 1) and another five independent lines originating from Agrobacterium strain 1073 were (lanes 3 through 7). A negative control (wild type SR-1 tobacco) is shown in lane 2. Protein blots were immuno-reacted with rabbit CBHI polyclonal antiserum. Molecular weights are indicated in kilodaltons.*
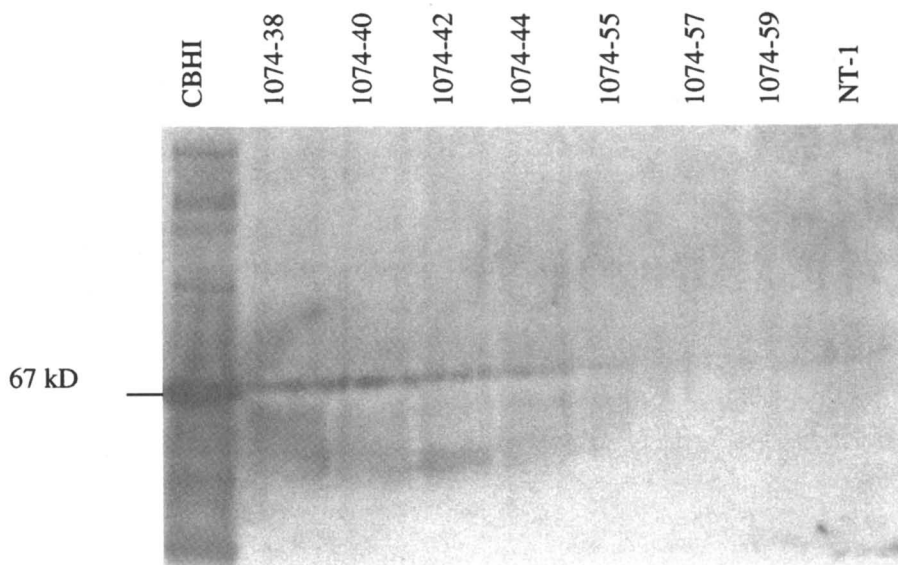
Figure 3. *Detection of CBHI protein in transgenic tobacco calli. Equal amounts of protein (30μg) were loaded on the each lane with the exception of the CBHI standard (0.5μg) derived from T. reesei. Protein blots were immuno-reacted with rabbit CBHI polyclonal antiserum. Molecular weights are indicated in kilodaltons.*
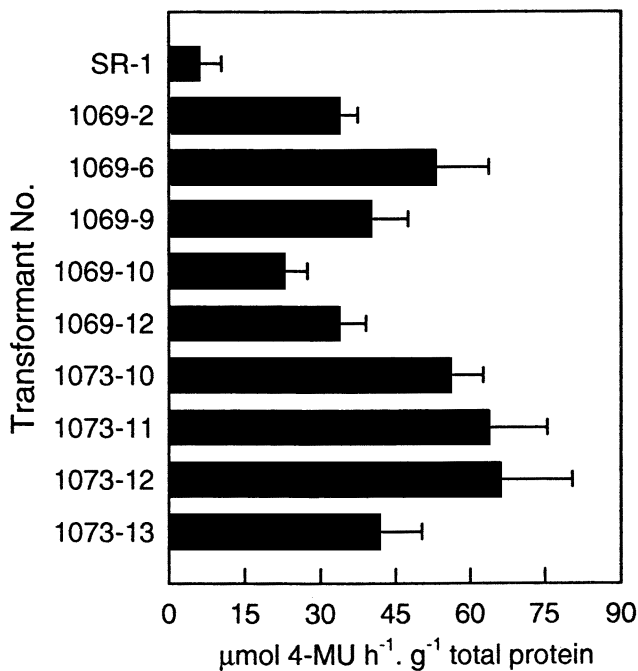
Figure 4. Measurements of CBHI activity in selected transgenic tobacco plants using the 4-methylumbelliferon assay. Each bar represents the mean ± one standard deviation of four to five replicates.

activity than that of untransformed control NT1 tobacco calli (data not shown). The CBHI activities of sixteen independent transgenic calli are plotted in Figure 5. Among all tested calli, the transformant 1074-38 possessed the highest level of CBHI activity at 83.6 μmol h^{-1} g^{-1} total soluble protein, which is approximately 15 times higher than that of control NT1 calli (5.3 μmol /h. g total proteins). The CBHI activity of most transformed calli was generally at least four times higher than that of NT1 calli.

Endoglucanase E1 Expression in Tobacco. The binary vector pZD276, used for plant transformation, included the entire mature E1 coding sequence isolated from *A. cellulolyticus* (Genbank accession No. U33212) under the control of the Rubisco small subunit promoter RbcS-3C (*25*) and transcriptional T7-T5 terminators. More than 25 independent kanamycin-resistant transgenic plants were regenerated and transferred to soil. Transgenic plants were grown for 3 to 4 weeks in the growth room, and leaf tissues were subsequently harvested for genomic DNA PCR, RNA blot, E1 activity analysis, and protein immunoblotting. All transgenic plants were first screened by PCR with primers corresponding to the E1 coding sequence (data not shown).

Results of RNA blot using a [α-^{32}P]-dCTP labeled 1.2 kb XbaI/BamH I E1 coding sequence fragment (Figure 6) show the expected size mRNA of 1.85 kb. The average E1 activity from three replicate samples for selected T1 transgenic plants ra-chl-2-1, ra-chl-2-2, ra-chl-2-3 was 2756.3, 1022.6, and 1559.1 pmol MU mg^{-1} total soluble protein min^{-1}, respectively (Figure 7A). Western blotting (Figure 7B) with monoclonal antibodies raised against *A. cellulolyticus* E1 demonstrated that the observed E1 activity was correlated with expression of the heterologous E1 enzyme in leaf tissues. Control SR1 extract shows no antibody staining material and little or no activity on 4-MUC. Western blots also show that the E1 protein extracted from transgenic plants is similar in molecular weight to E1 protein purified from a *S. lividans* expression system, suggesting that the 89 amino acids (including 57 amino acids of transit peptide and 22 amino acid of mature sequence of RbcS-2A protein, and an additional 8 amino acids comprising the cleavage site of the RbcS-2A transit peptide just before mature E1 coding sequence) of the Rubisco signal peptide had been removed to form mature E1 protein. The N-terminal sequence of the *S. lividans* E1 protein is identical to the native *A. cellulolyticus* protein (S. R. Thomas, National Renewable Energy Laboratory, unpublished). As the transgenic E1 protein has not yet been purified from leaf extracts, we have no data regarding its N-terminal sequence.

The expressed level of E1 protein in selected transgenic plants was estimated by gel electrophoresis and Western blot analysis. A series of diluted *S. lividans* E1 proteins was used as a standard. The Western blot densities of known *S. lividans* E1 proteins and E1 proteins from leaves of selected transgenic plants were estimated using DENDRON 2.2 program. Based on the densitometric determination, the amounts of E1 protein accumulated in selected transgenic plants ra-chl-2-1, ra-chl-2-2, and ra-chl-2-3 were estimated to be 1.35 %, 0.18 %, and 0.69 % of the total leaf soluble protein, respectively.
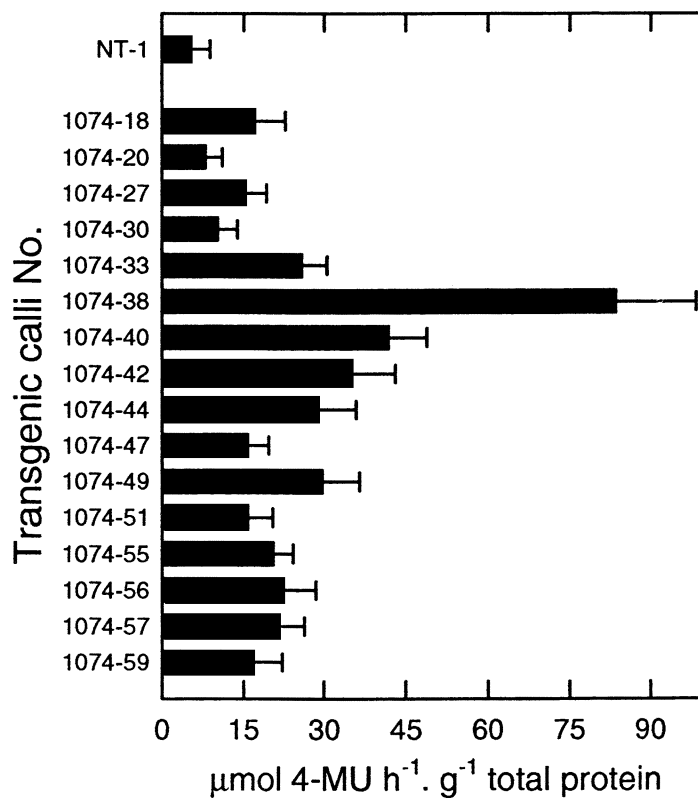
Figure 5. *Measurements of CBHI activity in selected transgenic tobacco calli using the 4-methylumbelliferon assay. Each bar represents the mean ± one standard deviation of four to five replicates.*
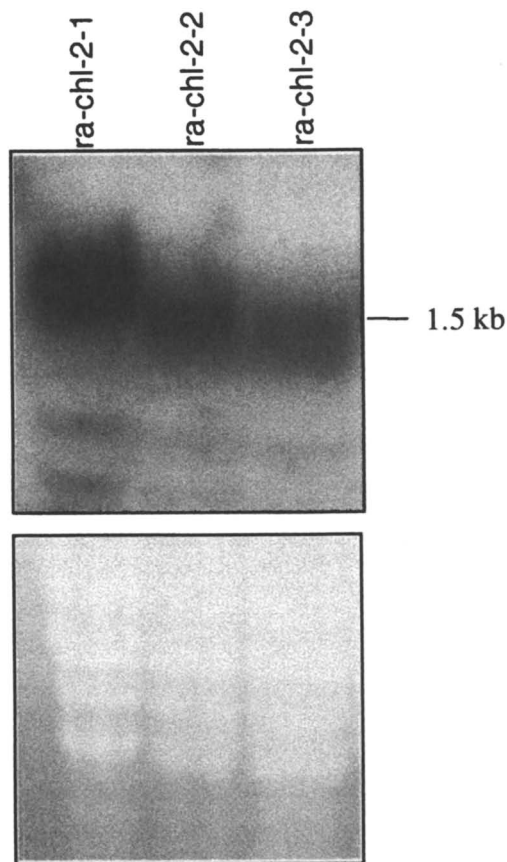
Figure 6. *Expression of E1 transgene in selected transgenic tobacco lines ra-chl-2-1, ra-chl-2-2, ra-chl-2-3. Shown are RNA blots containing (50 μg per lane) total RNA isolated from selected transgenic lines. The upper panel is the RNA blot hybridized with radioactive labeled probe prepared from the E1 coding region. The lower panel is the ethidium bromide-stained RNA bands shown as the loading control in upper panel.*
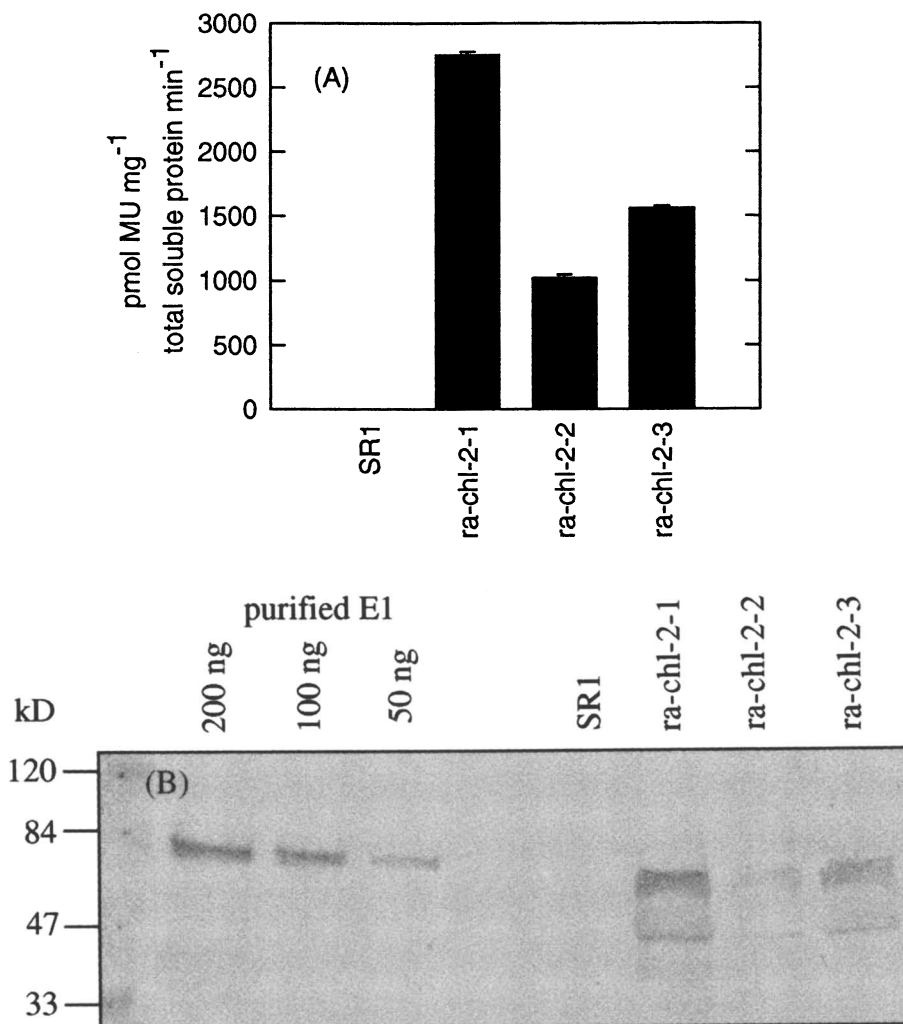
Figure 7. *A, Measurements of E1 enzyme activity in transgenic line ra-chl-2-1, ra-chl-2-2, ra-chl-2-3 by the MUC assay using the 4-methylumbelliferyl-β-D-cellobioside as a substrate. Each bar represents the mean ± the standard deviation of 3-4 replicates. B, Protein gel immunoblot analysis of E1 protein extracted from transgenic lines listed in A. The amount of protein loaded on the gel was determined before treatment. Lane 1 to 3 was the 200, 100, and 50 ng of E1 protein purified from bacterial cultures. Lane 4 was not loaded. Lane 5 contains 50 μg of total soluble leaf protein from wild-type control tobacco plants. Lanes 6 through 8 contain 40 μg per lane of total soluble leaf protein from transgenic tobacco plants.*

Immunocytochemistry. The transgenic line ra-chl-2-1 was used for immunolocalization of E1 in conjunction with the Rubisco signal peptide RbcS-2A (*25,34*). Figure 8 clearly shows immunogold particles preferentially labeling in the chloroplasts, indicating proper targeting of the E1 protein. Wild type SR-1 plants have no obvious immunogold labeling in chloroplasts and show only a low level of non-specific background signal. The result shows that the RbcS-2A signal peptide functions properly to direct E1 enzyme into chloroplasts.

Biochemical Properties of Recombinant E1 Enzyme in Leaf Extracts. Three kanamycin-resistant T2 transgenic plants (ra-chl-2-3-6, ra-chl-2-3-9, ra-chl-2-3-20), grown from T1 transgenic line ra-chl-2-3, were used to determine biochemical characteristics of recombinant E1 derived from plant hosts. The pH response of E1 enzyme from leaf extracts was first measured at 55°C at various pH values (Figure 9). Figure 9 shows the pH response curves of the three transgenic plants. As pH increased from 3.5 to 5.2, E1 enzyme activity gradually increased to an optimum. Thereafter, the E1 enzyme activity dramatically decreased as pH was further increased from 5.2 to 6.5. The three transgenic plants exhibited similar pH responses, though line ra-chl-2-3-6 differed in scale. The optimum pH for E1 enzyme in plant extracts is similar to the optimum growth condition of *A. cellulolyticus* (*9*).

The temperature response curve of E1 enzyme in plant extract was also determined with the three transgenic plants at pH 5.5. Figure 10 shows E1 enzyme activity at different temperatures. With increasing reaction temperature from 0°C to 50°C, E1 enzyme activity increased only up to 26% of maximal. Thereafter, with increasing temperature from 50°C to 81°C, E1 activity increased more dramatically. As temperature increased further from 82°C to 96°C, E1 activity decreased precipitously to only 5% of the maximum value. At 30°C (plant growth room temperature), E1 activity is less than 10% of the maximum value. The temperature response of E1 enzyme extracted from transgenic plants is similar to that of unfractionated supernatant of *A. cellulolyticus* (*9*).

Effect Of Leaf Age on E1 Enzyme Activity. The effect of leaf age on E1 enzyme stability and its activity was determined in T1 transgenic line ra-chl-2-1. The leaf tissues used for E1 enzyme activity measurement were harvested from upper leaves (leaves 3 and 4 counted from top to bottom), middle leaves (leaves 10 and 11), and lower leaves (leaves 15 and 16). E1 enzyme activity was similar between upper and middle leaves, but much higher in the lower leaf tissues when E1 enzyme activity was calculated on the basis of mg total soluble protein (Figure 11A). This increase of E1 enzyme activity in older leaf tissues is mainly due to the decrease of total protein in senescing leaves. E1 is evidently more stable than the majority of other proteins resident in plant tissues. However, E1 enzyme activity expressed on an expanded leaf area basis it remains relatively constant between different developmental stages. The degree of accumulation and stability of E1 protein was observed in protein immunoblots (Figure 11B). Figure 11B shows that E1 protein purified from *A. cellulolyticus* possessed two polypeptides. The upper polypeptide was the intact E1 protein, while the lower polypeptide was catalytic domain evidently resulting from E1 protein proteolysis. Compared with E1 protein purified from *A. cellulolyticus*, the
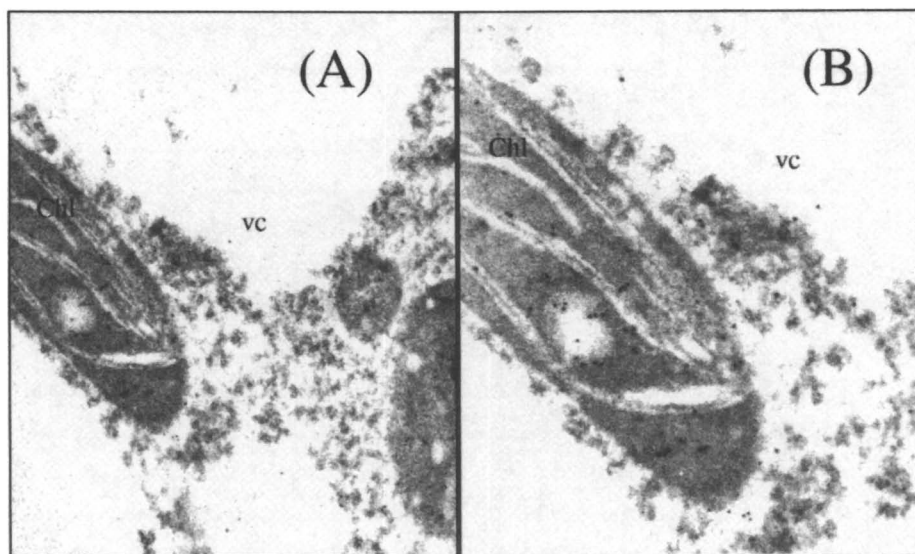
Figure 8. *Electron micrographs showing immunolocalization of E1 in transgenic line ra-chl-2-1 leaves. Affinity-purified Anti-E1 monoclonal antibody was used for immunolocalization. Immunogold particles reside in the stroma of chloroplasts. A, is* ×8500 and B is ×21,250. vc, vacuole; ch, chloroplast.
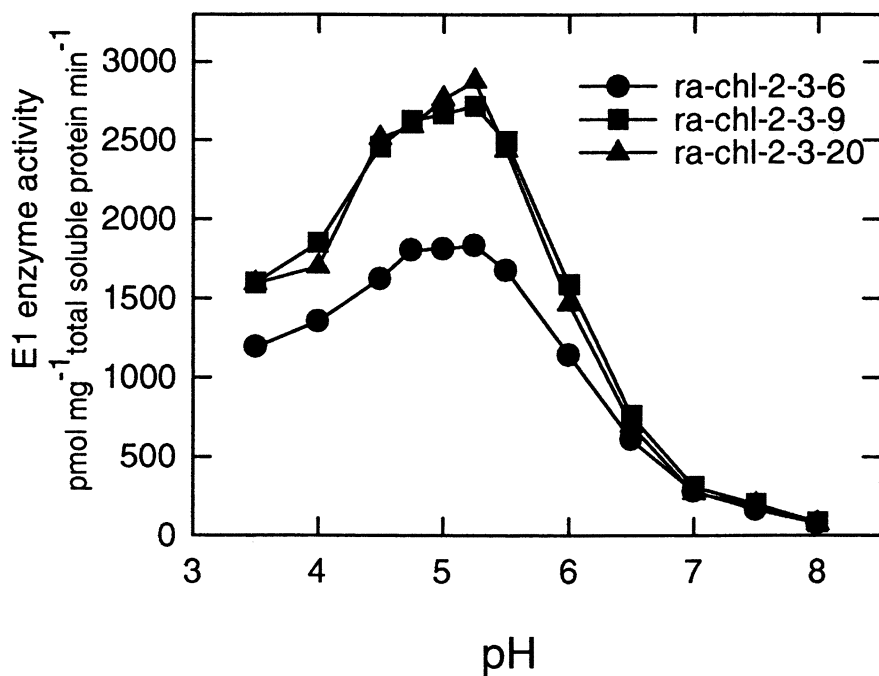
Figure 9. *The response of E1 enzyme extracted from transgenic plants to varying pH at 55°C. (A), E1 enzyme activity was determined by the MUC assay using 4-methylumbelliferyl-β-D-cellobioside as a substrate at different pH for three T2 transgenic lines ra-chl-2-3-6, ra-chl-2-3-9, ra-chl-2-3-20.*
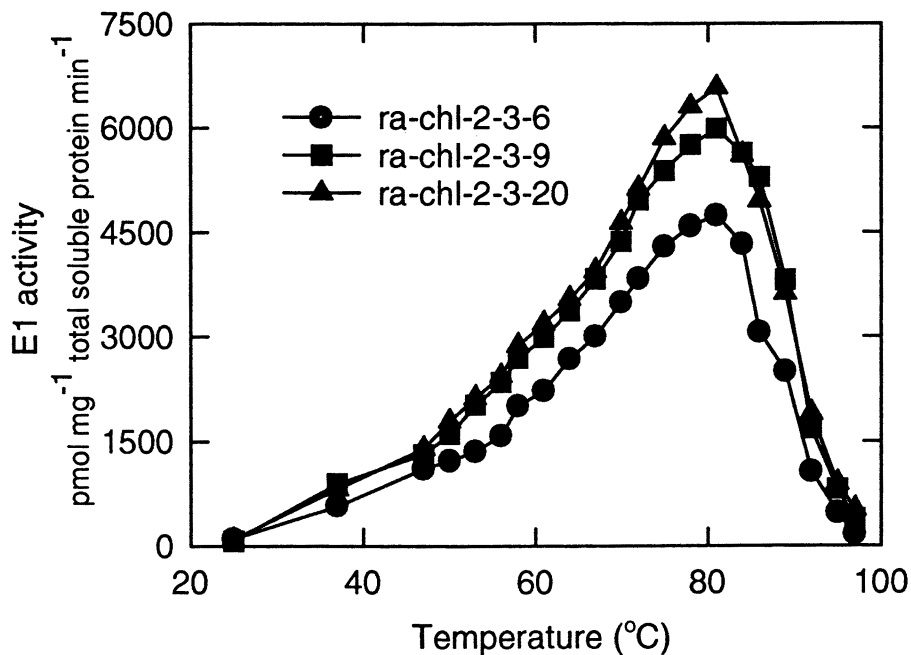
Figure 10. The response of E1 enzyme extracted from three transgenic plants to varying reaction temperature in MES buffer, pH 5.5. (A), E1 activity was determined by MUC assay using 4-methylumbelliferyl-β-D-cellobioside as a substrate at different reaction temperatures for the three T2 transgenic lines ra-chl-2-3-6, ra-chl-2-3-9, ra-chl-2-3-20.
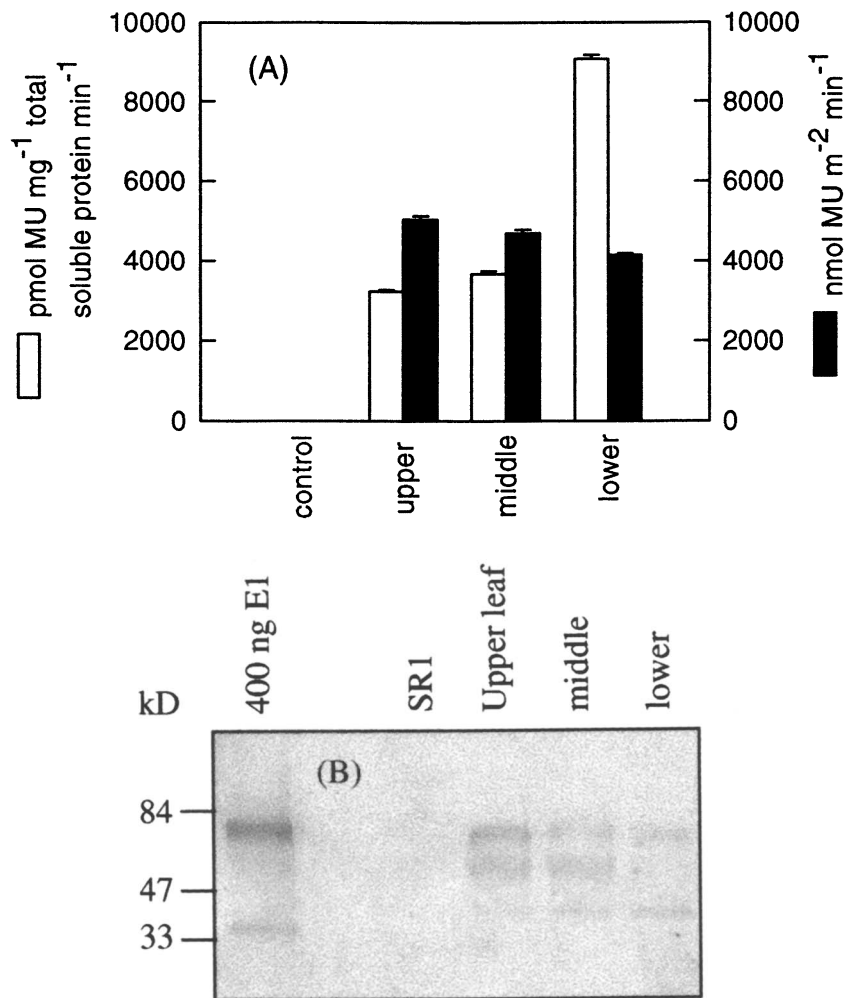
Figure 11. (A), Measurements of E1 enzyme activity in upper, middle, and lower leaves of transgenic line ra-chl-2-1 by the MUC assay using the 4-methylumbelliferyl-β-D-cellobioside as a substrate. Each bar represents the mean ± the standard deviation of 3-4 replicates. The open bars represent the E1 enzyme activity determined on the basis of mg total soluble protein. The filled bars are the E1 enzyme activity determined on the basis of initial leaf area. (B), Protein gel immunoblot analysis of E1 protein extracted from different leaf tissues listed in A. The amount of protein loaded on the gel was determined before treatment. Lane 1 was 400 ng of E1 protein purified from bacterial cultures. Lane 2 was not loaded. Lane 3 contains 50 μg total soluble leaf protein from wild type control tobacco plants. Lanes 4 through 6 contain 20 μg per lane of total soluble leaf protein from upper, middle, and lower leaf tissues of transgenic tobacco individual ra-chl-2-1.

upper leaf tissues of transgenic plants retained more intact E1 protein than that of middle and lower leaf tissues. In the lower leaf tissues, most of the E1 protein was partially degraded to a polypeptide larger in molecular weight than the E1 catalytic domain. The results from E1 enzyme activity assays and protein immunoblotting indicate that catalytic domain of E1 protein in transgenic plants is relatively resistant to proteolytic degradation *in planta*.

The Effects of Leaf Dehydration on E1 Activity. E1 enzyme stability and activity in leaf disks were determined at different stages of leaf dehydration (Figure 12). As shown in Figure 12A, E1 enzyme activity in leaf discs dramatically increased after 40 h of dehydration when expressed on a total soluble protein basis (pmol MU mg^{-1} total soluble protein min^{-1}). When leaf disks were further dehydrated up to 100 h, E1 enzyme activity increased only slightly. When expressed on a gram fresh leaf weight basis (pmol MU 0.1 g^{-1} min^{-1}), E1 enzyme activity remained fairly constant during the entire 100-hr dehydration period. Changes in E1 protein molecular weight in dehydrated leaf discs are shown in Figure 12B. After 5 h, E1 protein in leaf tissues has slightly degraded. As leaf discs were further dehydrated through 100 h, levels of E1 protein degradation products increased. The lowest molecular weight E1 fragment appearing in Figure 12B is approximately the same size as the E1 catalytic domain standard. Since levels of E1 enzyme activity did not decrease with dehydration time, it may be assumed that the smallest degradation fragment seen in Figure 12B corresponds to the E1 catalytic domain. Accumulation of the E1 catalytic domain in the dehydrated leaf disks gradually increased with prolonged leaf disk dehydration. However, more than 50% of E1 protein remains intact after 100 h dehydration. The results further demonstrate that E1 protein has a high degree of tolerance to endogenous proteolytic enzymes in leaf discs.

Photosynthetic Rates. Since the E1 enzyme was targeted to chloroplasts, CO_2 fixation in transgenic leaves may be affected. Photosynthetic rates were measured in mature 3rd or 4th leaves of transgenic line ra-chl-2-1 and untransformed control SR1 tobacco plants at 30°C leaf temperature and ambient CO_2 concentration. The photosynthetic rate of transgenic line ra-chl-2-1 was 23.8 ± 0.5 µmol m^{-2} s^{-1}, which was very similar to that of wild type SR1 (22.8 ± 0.35 µmol m^{-2} s^{-1}). The results indicate that E1 protein accumulation in chloroplasts does not affect the photosynthetic function of the plant. The transgenic plants were normal in morphology, plant growth and development, and seed setting as compared with untransformed SR1 tobacco plants. It has also been shown that expression of a thermostable xylanase from *Clostridium thermocellum* expressed at high levels in the apoplast of transgenic tobacco did not affect the growth of transgenic plants (*13*).

E1 Expression In Transgenic Potato. More than 100 independent kanamycin-resistant transgenic potato plants were obtained from transformation experiments with four E1 expression vectors: mm-E1 (with E1 native signal); ra-chl (chloroplast), rr-vac (vacuole), and mm-apo (apoplast). mm-E1 and mm-apo were under the control of constitutive Mac promoter (*24*), while ra-chl and rr-vac were driven by a leaf specific RbcS-3C promoter (*25*) Tobacco transformation using these expression vectors,
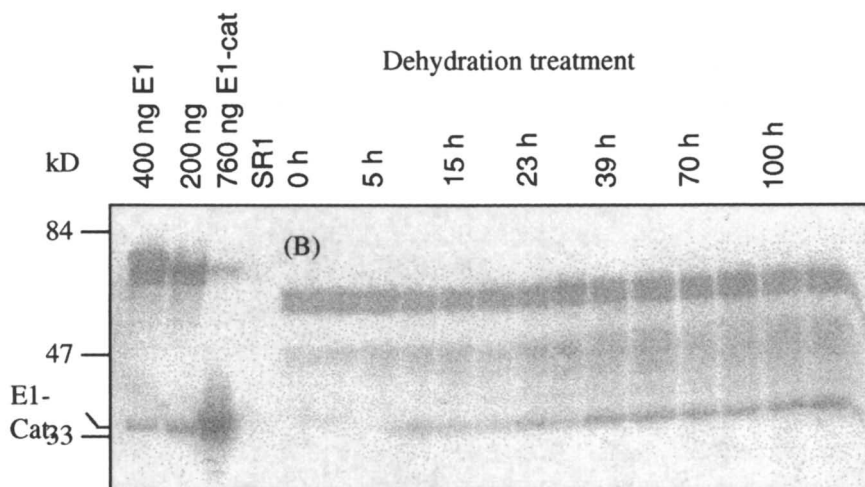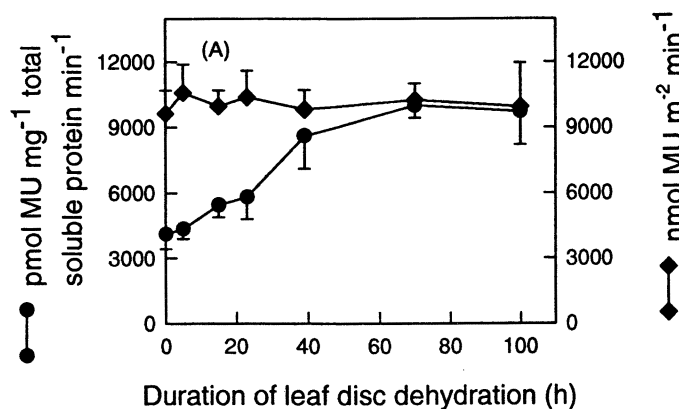
Figure 12. *The E1 enzyme activity and E1 protein of leaf discs dehydrated at room temperature at different incubation times. A, Measurements of E1 enzyme activity of leaf discs of transgenic line ra-chl-2-1 dehydrated at room temperature for different time period (0 to 100 hr), using the 4-methylumbelliferyl-β-D-cellobioside as a substrate. Each data point represents the mean ± the standard deviation of 3-4 replicates. B, Protein gel immunoblot analysis of E1 protein extracted from leaf disks dehydrated at room temperature for different time periods. The amount of protein loaded on the gel was determined before treatment. Lane 1 and 2 contain 400 and 200 ng purified E1 protein. Lane 3 contains 760 ng of the catalytic domain of E1 protein (also produced in S. lividans). Lane 4 contains 50 μg of total soluble leaf protein from a wild type control tobacco plant. Lanes 5 through 18 contain 20 μg per lane of total soluble leaf protein extracted from leaf disks of transgenic tobacco ra-chl-2-1 dehydrated at room temperature for different time periods. The label 1 and 2 are two repeat samples at the each time point.*

completed previously in our laboratory, yielded fairly high levels of E1 expression in leaf tissues (Z. Dai, unpublished). It has been demonstrated that the pathogensis related gene S (PR-S) signal peptide (*11*), chloroplast (RbcS-2A) signal peptide (*35*), and sporamin gene signal peptide (*36-38*) are able to target heterologous proteins to apoplast, chloroplast, and vacuole, respectively.

Transgenic plants were grown for 4 weeks in the growth room and healthy leaf samples were subsequently harvested for E1 enzyme extraction. E1 enzyme activity in leaf extracts of different potato transformants is shown in Figure 13. Among the four transformant types, individual transgenic plants with apoplast targeting (Figure 13B) typically yielded the highest E1 activity followed by chloroplast targeting transgenic plants (Figure 13D), transformants using the native prokaryotic E1 signal peptide (unknown localization, Figure 13A), and vacuole signal peptide (Figure 13C). More than 37% of the transgenic plants under the control of Mac promoter, mas 5'-UTL sequence, and apoplast targeting signal sequence of PR-S gene possessed over 10,000 pmol MU mg^{-1} total soluble protein min^{-1} E1 activity (Figure 13B). The highest E1 activity in these transgenic plants was 67,000 pmol MU mg^{-1} total soluble protein min^{-1}. The highest E1 enzyme activity in transgenic plant, in which the E1 gene under the control of leaf specific promoter RbcS-3C, AMV 5'-UTL sequence, and the chloroplast signal sequence of RbcS-2A gene, was 48027 pmol MU mg^{-1} total soluble protein min^{-1}, while most of individual transgenic plants had E1 enzyme activity between 3000 and 9000 pmol MU mg^{-1} total soluble protein min^{-1} (Figure 12D). E1 enzyme activity in leaf extracts of transgenic plants with E1 gene under the control of Mac promoter, mas 5'-UTL, and native prokaryotic E1 signal sequences is shown in Figure 13A. There were only 50% of transgenic plants analyzed that had E1 activity. The highest E1 activity was 3700 pmol MU mg^{-1} total soluble protein min^{-1}. E1 activity in the leaf extracts of these transgenic plants was 3 to 20 times lower than that of these transformants with apoplast or chloropast targeting. The transformant with the E1 gene under the control of leaf specific promoter RbcS-3C, RbcS-3C 5'-UTL, and vacuole signal sequence of sporamin gene possessed lowest E1 activity (Figure 13C) with the highest E1 activity of 180 pmol MU mg^{-1} total soluble protein min^{-1}. E1 enzyme activity results show that E1 expression under the control of the Mac promoter with an apoplast signal peptide and RbcS-3C promoter with AMV 5'-UTL and chloroplast signal peptide is much higher than that under the control of the Mac promoter with the native prokaryotic E1 signal peptide and RbcS-3C promoter with the vacuolar signal peptide in transgenic potato.

E1 protein accumulation in transgenic potato leaves was determined by Western blot analysis with monoclonal antibodies raised against *A. cellulolyticus* E1. E1 protein accumulation in 15 selected F0 transformants possessing higher E1 activity was examined in leaf protein extracts (Figure 14). Among the F0 transformants examined, transgenic plants (ra-chl-F series) with E1 coding sequence under the control of the RbcS-3C promoter, AMV leader and chloroplast signal sequence of RbcS-2A gene accumulated higher levels of E1 protein in leaf tissues. Western blots also show that the E1 protein extracted from transgenic plants with E1 protein targeted to chloroplasts is similar in molecular weight to E1 protein purified from a *S. lividans* expression system, suggesting that the 89 amino acids (including 57 amino acids of transit peptide and 22 amino acid of mature sequence of RbcS-2A protein,
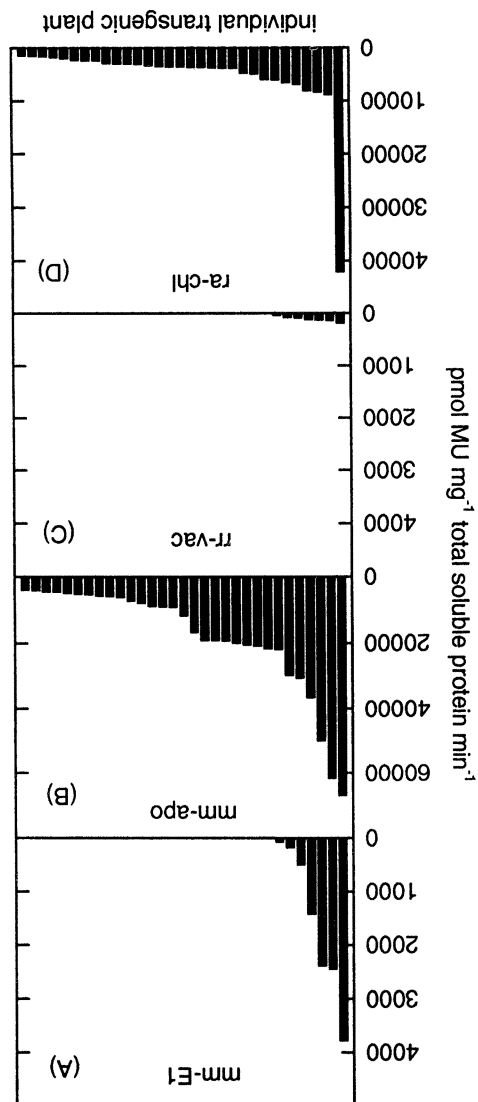
Figure 13. E1 activity of different individual transgenic plants bearing different expression cassettes. Panel (A) and (B): E1 coding sequence was under the control of Mac promoter and mas 5'-untranslation leader, native prokaryotic E1 signal peptide (mm-E1, A), and apoplast signal peptide (mm-apo, B), Panel (C) and (D): E1 coding sequence under the control of leaf specific RbcS-3C promoter, and its 5'-untranslation leader with the signal peptide sequence of a sporamin (rr-vac, C) or AMV 5'-untranslational leader with a chloroplast signal peptide (ra-chl, D).
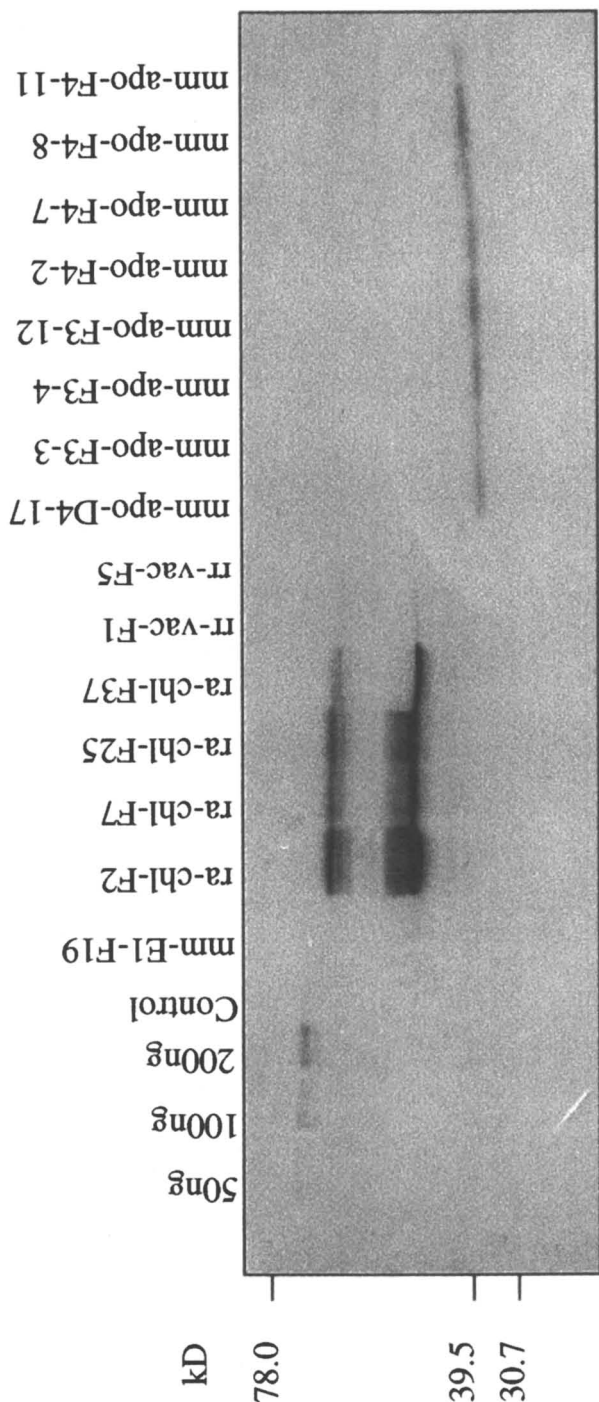
Figure 14. *Western blot for E1 protein expressed in leaf tissues of selected transgenic plants. Lane 1 is the protein marker. Lanes 2 to 4 are 50, 100, and 200ng purified E1 protein from bacterial cultures. Lane 5 is the potato control plant. Lane 6 is transgenic plant mm-E1-F19 carrying the native prokaryotic E1 signal peptide. Lanes 7 to 10 are transgenic plants ra-chl-F2, ra-chl-F7, re-chl-F-25, and ra-chl-F-37 containing the chloroplast signal peptide RbcS-2A. Lanes 11 and 12 are transgenic plants rr-vac-F-1 and rr-vac-F-5 with E1 protein targeted to the vacuole. Lanes 13 to 20 are transgenic plants mm-apo-D4-17, mm-apo-F3-3, mm-apo-F3-4, mm-apo-F3-12, mm-apo-F4-2, mm-apo-F4-7, mm-apo-F4-8, and mm-apo-F4-11 with apoplast targeting. D or F in the front of the transgenic plant number correspond to potato cv Desiree and FL1607, respectively. Forty µg protein was used for each lane from lanes 5 to 20.*

and an additional 8 amino acids comprising the cleavage site of the RbcS-2A transit peptide just before mature E1 coding sequence) of Rubisco signal peptide had been removed to form mature E1 protein. Some of the chloroplast targeted E1 protein was partially degraded, to a molecular weight being larger than that of the E1 catalytic domain (about 39 kD) (Z. Dai, unpublished). Interestingly, all of the E1 protein in transformants with apoplast targeting was partially degraded to a polypeptide similar in molecular weight to the E1 catalytic domain. However, *Aspergillus niger* phytase (*11*) and the catalytic domain of *Clostridium thermocellum* xylanase (*13*) are resistant to proteolytic degradation in the apoplast of transgenic tobacco. The results from E1 activity and E1 protein immunoblotting suggest that the E1 catalytic domain is more resistant to proteolytic degradation when targeted to the apoplast. Transgenic plants with the native prokaryotic E1 signal peptide and vacuole signal peptide showed very low E1 protein accumulation. No E1 enzyme activity and E1 protein accumulation are detected in the control potato plant.

The expressed level of E1 protein in selected transgenic plants was estimated by gel electrophoresis and Western blot analysis. A series of diluted *S. lividans* E1 proteins was used as a standard. The gel blot densities of known *S. lividans* E1 proteins and E1 proteins from leaves of selected transgenic plants were estimated using the DENDRON 2.2 program. Based on the densitometric determination, the amounts of E1 protein accumulated in selected transgenic plants ra-chl-F2, F7, F25, and F37 were estimated to be 2.6 %, 1.8 %, 1.9%, and 0.73 % of the total leaf soluble protein, respectively. E1 protein accumulated in the transformants with apoplast targeting ranged from 0.38% to 0.92% of total leaf soluble protein.

Since there are some discrepancies between the E1 activity and protein accumulation, we further examined RNA accumulation in transgenic potato plants. Total RNA was isolated from leaf tissues of selected transgenic potato plants. The transcription of the E1 gene in these selected transgenic plants was determined by RNA gel blot using a 1.2 kb Xba I/BamH I E1 coding sequence fragment labeled with [a-^{32}P]-dCTP as probe (Figure 15). Results in Figure 15 show that the selected transgenic plants properly transcribed the integrated E1 gene, yielding the predicted 1.5 kb mRNA though the mRNA accumulation varies among these selected transgenic potato plants. Obviously, the transgenic potato plants with chloroplast targeting, in which E1 protein was accumulated at higher level, possessed high E1 gene transcription as compared to transgenic potato plants with the native prokaryotic E1 signal peptide, vacuole signal peptide or apoplast signal peptide. However, E1 protein polypeptides accumulated in the apoplast possessed higher E1 activity, which may indicate that the partial degraded polypeptide of E1 protein from leaf extract may have stronger hydrolytic activity catalyzing the MUC substrate. Further experiments will be needed to address this interesting phenomenon.

In addition, the results from E1 activity, E1 protein immunobloting and RNA gel blot show that apoplast (PR-S) signal peptide enhanced E1 gene expression as compared to the E1 gene under the control of same Mac promoter, mas 5'-UTL, and mas terminator with native prokaryotic E1 signal peptide. Low E1 activity, E1 protein accumulation and no detectable RNA in the transgenic plants carrying the construct rr-vac suggests that the potato sporamin signal peptide may affect E1 gene expression. Further study will be required to determined the effects of sporamin
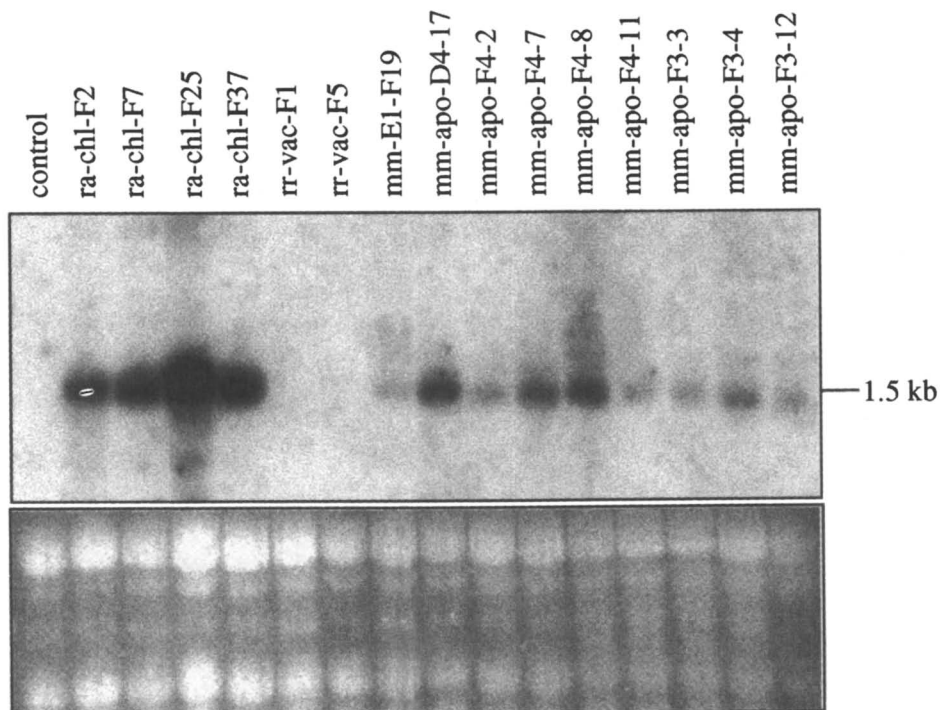
Figure 15. *Expression of the E1 transgene in selected transgenic potato lines from different transformants with higher E1 activity. Shown are RNA blots containing (20 μg per lane) total RNA isolated from selected transgenic lines. Upper panel is the RNA blot hybridized with radioactive labeled probe prepared from the E1 coding region. The lower panel is the ethidium bromide-stained RNA bands shown as the loading control in upper panel.*

signal peptide on transgene expression in transgenic potato. It has been demonstrated that the E1 gene under the control of RbcS-3C promoter and T7-T5 terminator with chloroplast signal peptide had higher E1 gene transcription and E1 protein accumulation in transgenic tobacco. (Z. Dai, unpublished).

E1 Protein Accumulation in Tuber Tissues of Transgenic Potato. E1 enzyme activity and protein accumulation were also examined in tuber tissues of transgenic potato under the control of the constitutive Mac promoter and leaf specific RbcS-3C promoter. E1 enzyme activity was detected in both leaves and tubers of transgenic plants carrying E1 gene under the control of the Mac promoter (mm-E1-F19; mm-apo-D4-17; mm-apo-F3-3; mm-apo-F4-2; mm-apo-F4-8; and mm-apo-F4-11) (Figure 16). Overall, E1 enzyme activity is lower in tuber extracts than those in the leaf extracts on the basis of total soluble protein (Figure 16A and B) or of fresh weight (Figure 16C and D). However, transgenic plants (ra-chl-F2, 7, 25, 37, and 41) with the E1 gene under the control of leaf specific RbcS-3C promoter, with E1 activity ranging from 7000 to 46000 pmol mg^{-1} total soluble protein min^{-1}, or from 1700 to 20000 pmol MU g^{-1} fr wt min^{-1}, possessed no detectable E1 activity in tuber extracts.

In order to determine E1 protein accumulation in tubers, protein extracts from tuber tissues of selected transgenic plants were evaluated using protein immunoblotting. The immunoblot gel in Figure 17 shows that E1 protein accumulated in the tuber tissues of transgenic plants bearing the Mac construct (though the transgenic potato plant mm-E1-F19 possessed low E1 protein accumulation). No E1 protein was observed in tuber extract of transgenic plant ra-chl-F2 that was under the control of leaf specific RbcS-3C promoter. All of selected transgenic plants (mm-E1-F19; mm-apo-F3-3; mm-apo-F4-2; mm-apo-F4-8, and mm-apo-F4-11), except ra-chl-F2 accumulated partially degraded polypeptide of E1 protein in potato tubers, similar to that in leaf tissues (Figure 14). We also found that the degradation pattern of E1 protein in transgenic line mm-E1-F19 was similar to that of those potato transgenic lines with E1 protein targeted to the apoplast. This suggests that the native prokaryotic E1 signal peptide may serve as an apoplast signal.

Transcription of the E1 gene in tubers was also evaluated by RNA blotting using a 1.2 kb Xba I/BamH I E1 coding sequence fragment labeled with the [α-^{32}P]-dCTP as a probe (Figure 18). The total RNA from leaf tissues of the selected transgenic plant was used as a positive control. Results in Figure 18 shows that the selected transgenic plants mm-apo-F4-2 and mm-apo-F4-8, in which the E1 gene was under the control of Mac promoter, properly transcribed the integrated E1 gene in tuber tissues, yielding the predicted 1.5 kb mRNA. The transcription level of the E1 gene in tuber tissues was much less than that in leaf tissues. However, the transgenic plant ra-chl-F25, with E1 expression under the control of the RbcS-3C promoter, possesed no detectable E1 transcription. Although significantly lower levels E1 gene transcription was found in tuber tissues of transgenic plants mm-apo-F4-2 and mm-apo-F4-8, E1 enzyme activity and protein accumulation were not significantly different between the tuber tissues and leaf tissues. These results suggest that E1 protein, especially its catalytic domain, was very stable in both leaf and tuber tissues.
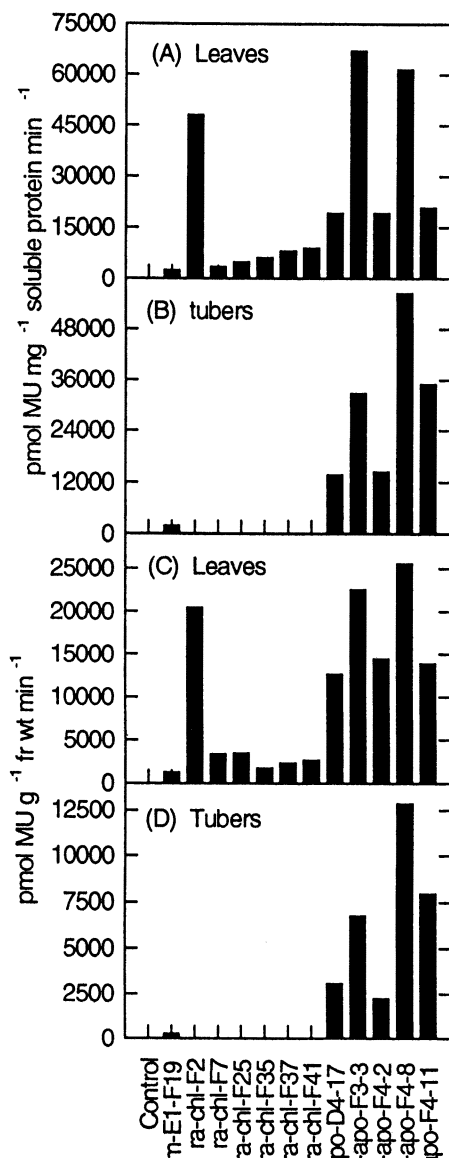
Figure 16. *E1 activity in leaf and tuber tissues of selected individual transgenic plants bearing different expression cassettes. Panels (A) and (B) are E1 activities in leaf and tuber tissues, respectively, expressed as pmol MU mg^{-1} total soluble protein min^{-1}. Panels (C) and (D) are E1 activities in leaf and tuber tissues, respectively, expressed as pmol MU g^{-1} fresh weight min^{-1}.*
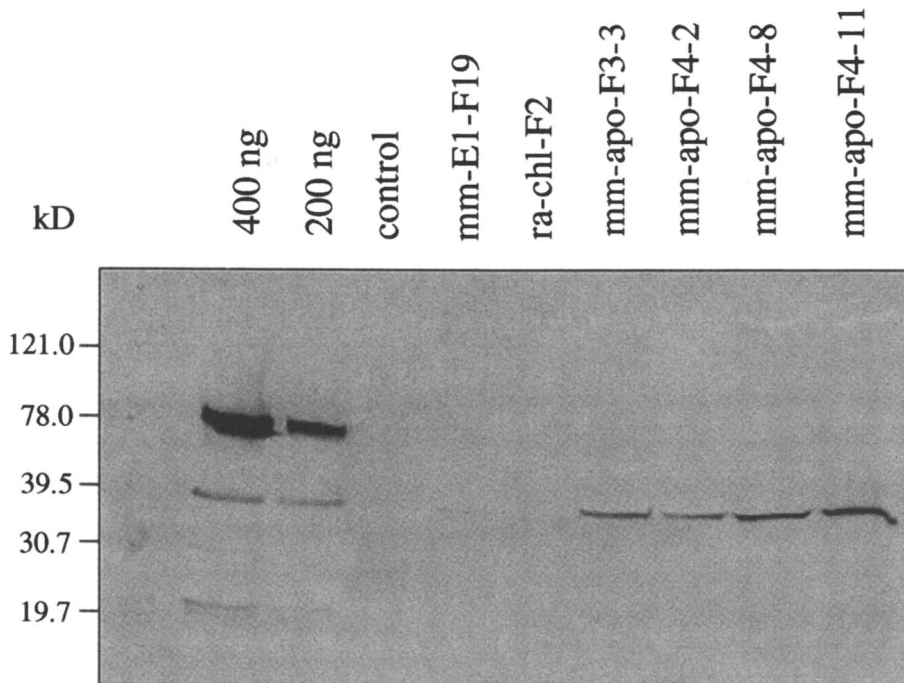
Figure 17. *The Western blot for E1 protein expressed in tuber tissues of selected transgenic plants. Lane 1 is the protein marker. Lanes 2 and 3 are 388 and 194 ng purified E1 protein from bacterial cultures. Lane 4 is potato control plant. Lanes 5 to 10 are transgenic plants mm-E1-F19, ra-chl-F2, mm-apo-F3-3, mm-apo-F4-2, mm-apo-F4-8, and mm-apo-F4-11. Twenty µg total soluble protein from tuber tissues was used for each lane from 4 to 10.*
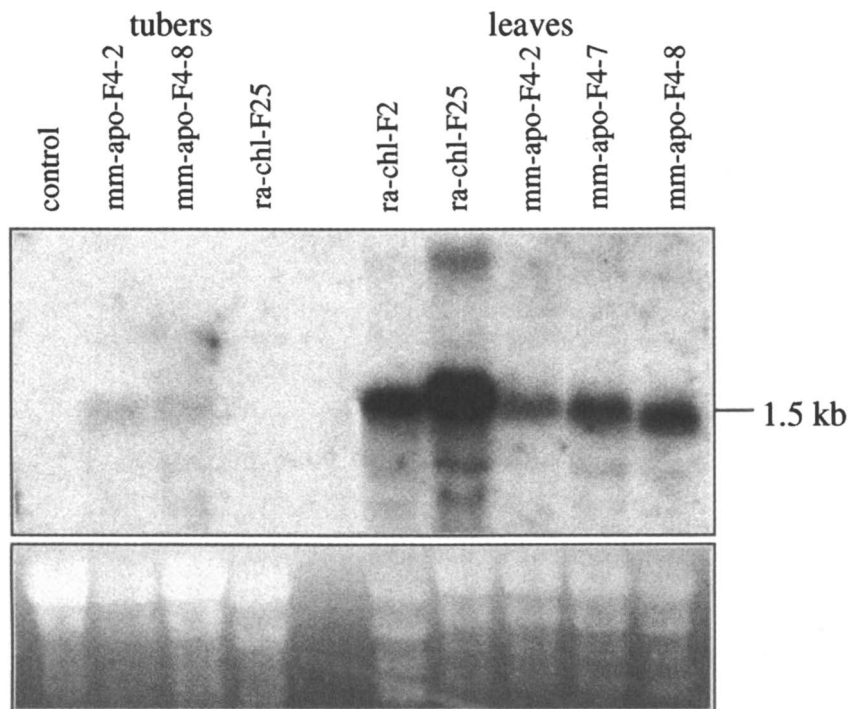
Figure 18. *Expression of the E1 gene in tuber and leaf tissues of selected transgenic potato plants. Twenty μg of total RNA was isolated from tuber tissues or leaf tissues of selected transgenic plants. Upper panel is the RNA blot hybridized with radioactive labeled probe prepared from the E1 coding region. The lower panel is the ethidium bromide-stained RNA bands shown as the loading control in the upper panel.*

Economic Analysis of Cellulase Production in Potato Vines (2005 Technology).
Approximately, 600,000 of the 1,363,000 acres (44%) employed for potato production
in the United States are in the Pacific Northwest, which was used as the model
growing area for this analysis. Potato foliage yield is about 40 tons/acre with disease
resistant cultivars and 13 tons/acre with non-resistant cultivars. Since genetically
engineered potato cultivars carrying *Verticellium* and late blight resistant genes have
already been developed by third parties, it was assumed for this analysis that most of
the potatoes grown by 2005 will be disease resistant. Thus, the higher yield numbers
were employed here. The foliage is 10% of total solids at harvest and 15.5% protein
within the solids. It was assumed that E1- and CBH1-bearing plants could be
optimized such that 10-15% of the total protein would be active cellulase enzyme
without harming the value of the tubers (traditional potato crop). This level of
production in the plant is regarded as moderately ambitious, but definitely achievable
based on previous effort (*11*). The economic analysis accounted for all costs
associated with potato cultivation, foliage harvest, transportation, storage, and
processing to gain a final, high purity cellulase enzyme product. The potato vines are
"greenchopped," collected in windrows between potato hills and left to field dry.
Greenchopping will actually induce "skin-set" and eventual ripening in potato tubers.
After drying, vines are collected, coarse chopped and transported to cellulase
processing facilities (Figure 19). Processing includes the addition of make-up water
to elute the soluble protein fraction from dry vine biomass, heat denaturation of non-
cellulolytic proteins, centrifugation/dewatering, microfiltration and ultrafiltration

Major uncertainties in this analysis are the farmer profit per acre of land, the
fraction of cellulase lost in processing, and the grade of byproduct cattle feed
($0.33/kg protein for poor, $0.47/kg protein for fair). Anecdotal evidence indicates
that the profit required by the farmer may range from $150 to $200/acre. If potato
tops are valued as fair grade animal feed the farmer can make $206/acre profit by
selling the tops directly to animal feed lots so the required profit will probably be
$200/acre or more. Depending on the efficiency of elution of proteins from dried
potato foliage and recoveries in downstream steps, overall cellulase recovery could
range as widely as 25-75%. Lacking any definitive process information in this regard,
we have chosen a middle of the road to worst case scenario here (25-50%). Fair grade
feed product is assigned based on a total protein content of 15% while poor grade
product dips down to as low as 10% protein.

This analysis is based on reasonable assumptions and yields plausible, bracketed
numbers for the cost of cellulases produced in potato tops. In the best case scenario
(required farmer profit is $150/acre, enzyme recovery is 50%, expression level is 15%
and fair grade byproduct is assumed) enzyme cost decreases to $1.40/kg protein
($0.78/$10^6$ FPU, $0.056/gal ethanol). When the required farmer profit is $200/acre,
the cost of enzyme for that case is increased to $2.59/kg protein ($1.44/10^6 FPU,
$0.104/gal ethanol). In the worst case scenario (farmer profit is $200/acre the enzyme
yield is 25%, expression level is 10% and poor grade byproduct is assumed) enzyme
cost increases to $8.28/kg protein ($0.46/10^5 FPU, $0.329/gal ethanol). A range of
costs for cellulase enzyme production is shown in Figure 20.
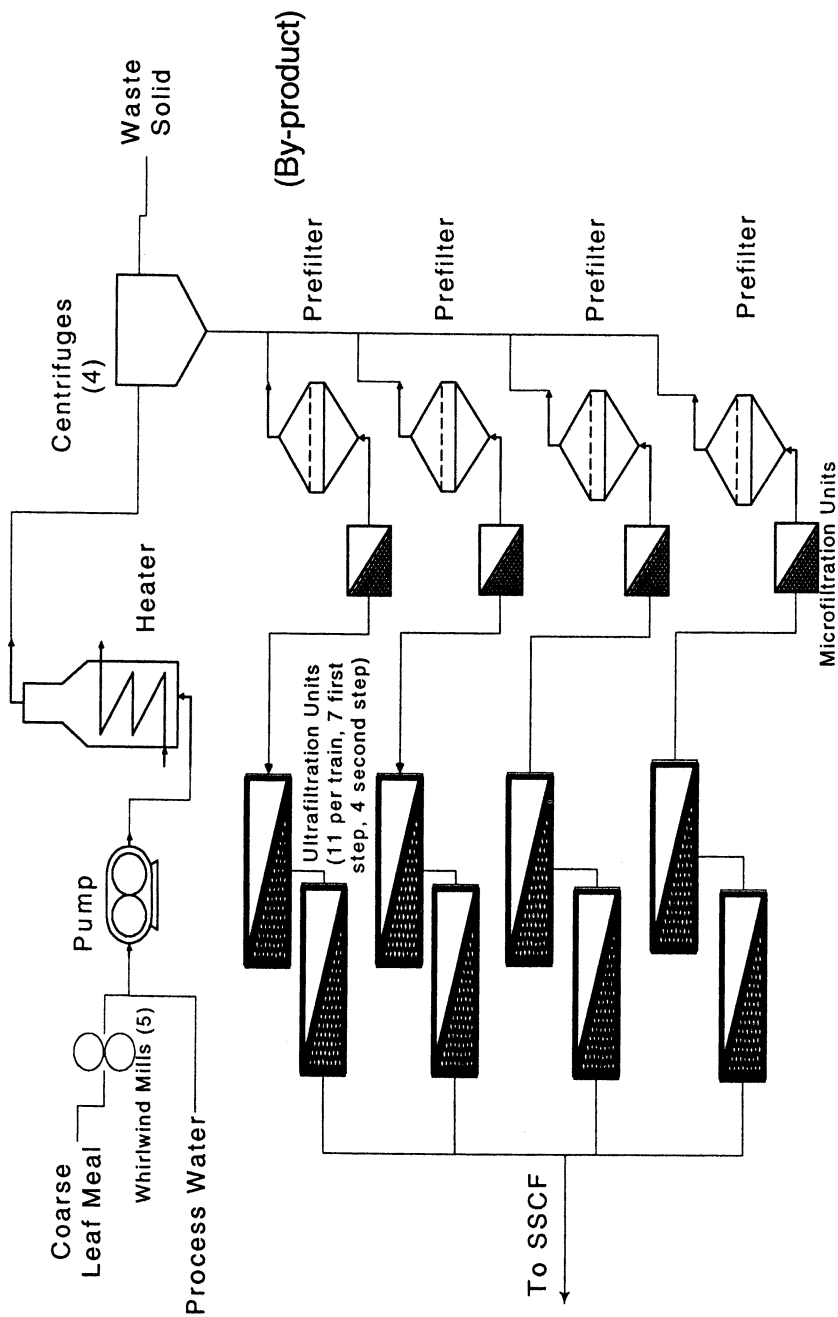
Figure 19. *Cellulase processing facility – chopped potato vines are converted to thermostable cellulase enzyme and cattle feed byproduct.*
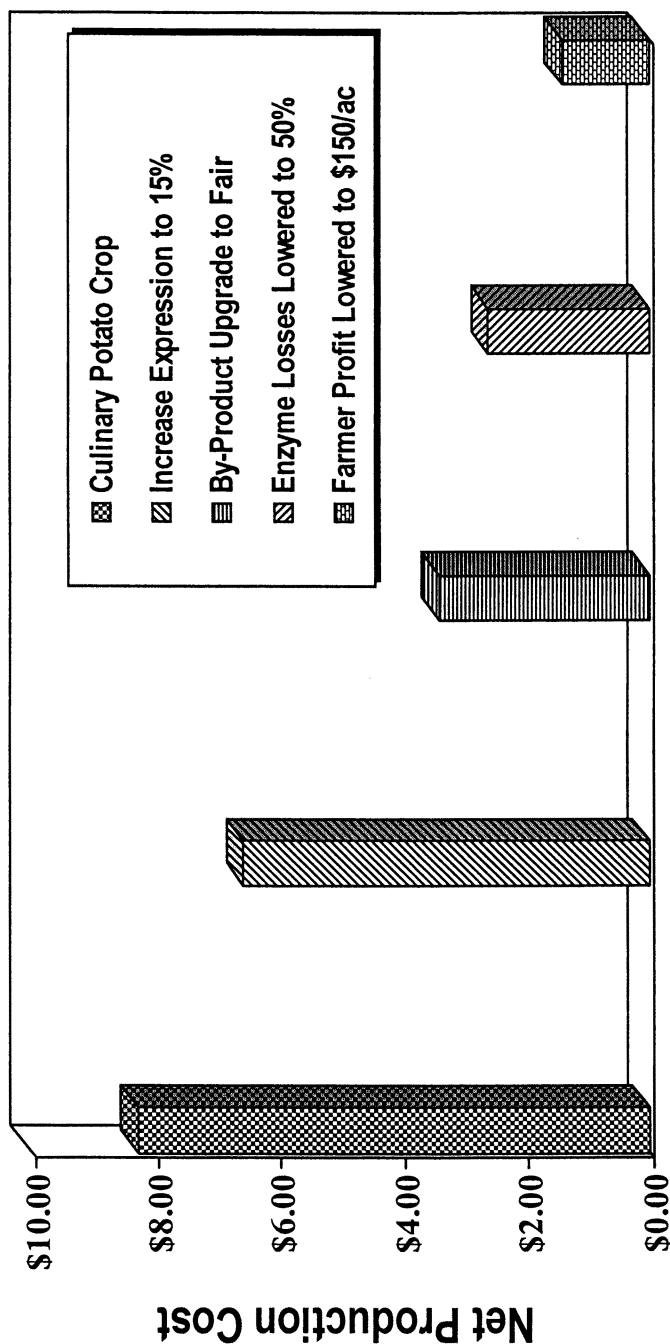
Figure 20. *Detailed cost of cellulase production (field to purified enzyme). Worst case scenario – 10% cellulase expression, 25% recovery of enzyme, poor grade feed byproduct, and $200/acre farmer profit.*

References

1. Voet, D.; Voet, J. G. *Biochemistry;* John Wiley & Sons, New York, 1990, p 254.
2. Duff, S. J.; Murray, W. D. *Bioresource Technol.* **1996,** *55,* 1-33.
3. Gilbert, H. J.; Hazlewood, G. P. *J. Gen. Microbiol.* **1993,** *139,* 187-194.
4. Lynd, L. R.; Cushman, J. H.; Nichols R. J.; Wyman, C. E. *Science* **1991,** *251,* 1318-1322
5. Shoemaker, S.; Watt, K.; Tsitovsky, G.; Cox, R. *Bio/Technology* **1983,** *1,* 687-690.
6. Shoemaker, S.; Schweickart, V.; Landner, M.; Gelfand, D.; Kwok, S.; Myambo, K.; Innis, M. *Bio/Technology* **1983,** *1,* 691-696.
7. Terri, T.; Salovuori, I.; Knowles, J. *Bio/Technology* **1983,** *1,* 697-702.
8. Mohagheghi, A.; Grohmann, K.; Himmel, M.; Leighton, L.; Updegraff, D. *International J. System. Bacteriol.* **1986,** *36,* 435-443
9. Tucker, M. P.; Mohagheghi, A.; Grohman, K.; Himmel, M. E. *Bio/Technology* **1989,** *7,* 817-820
10. Sakon, J.; Adney, W. S.; Himmel, M. E.; Thomas, S. R.; Karplus, A. *Biochem.* **1996,** *35,* 10648-10660
11. Verwoerd, T.C.; van Paridon, P.A.; van Ooyen, A. J. J.; van Lent, J. W. M.; Hoekema, A.; Pen, J. *Plant Physiol.* **1994,** *109,* 1199-1203
12. Dai, Z.; Hooker, B. S.; Quesenberry, R. D.; Gao, J. *Applied Biochem. Biotech.* **1999,** *77-79,* 689-699.
13. Herbers, K.; Wilke, I.; Sonnewald, U. *Bio/Technology* **1995,** *13,* 63-66
14. Jensen, L.G.; Olsen, O.; Kops, O.; Wolf, N.; Thomsen, K.K.; Wettstein, D. V. *Proc. Natl. Acad. Sci. USA* **1996,** *93,* 3487-3491
15. Liu, J. H.; Selinger, B.; Cheng, K. J.; Beauchemin, K. A.; Moloney, M. M. *Mol. Breeding* **1997,** *3,* 463-470.
16. Adney, W. S.; Thomas, S. R.; Himmel; M. E.; Baker, J. O.; Chou, Y. C. U.S. Patent 5,712,142, 1998.
17. Lastick, S. M.; Tucker, M. P.; Grohmann, K. U.S. Patent 5,514,584, 1996.
18. Thomas, S. R.; Laymon, R. A.; Himmel, M. E. U.S. Patent 5,536,655, 1996.
19. Hoekema, A.; Hirsch, P. R.; Hooykaas, P. J. J.; Schilperoort, R. A. *Nature* **1983,** *303,* 179-181.
20. An, G.; Ebert, P. R.; Ha, S.-B. In *Plant Molecular Biology Manual;* Gelvin, S. B.; Schilperoort, R. A.; Eds.; Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988; pp 1-19.
21. Murashige, T.; Skoog, F. *Physiol. Plant.* **1962,** *15,* 473-497
22. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual,* 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

23. An, G. In *Methods in Molecular Biology;* Gartland, K. M. A.; Davey, M. R. Eds.; Humana Press, Totowa, NJ, 1995; Vol. 44, pp 47-58.

24. Comai, L.; Moran, P.; Maslyar, D. *Plant Mol. Biol.* **1990,** *15,* 373-381.

25. Sugita, M.; Manzara, T.; Pichersky, E.; Cashmore, A.; Gruissem, W. *Mol. Gen. Genet.* **1987,** *209,* 247-256

26. Brederode, F. T.; Koper-Zwarthoff, E. C.; Bol, J. F. *Nucleic Acids Res.* **1980,** *8,* 2213-2223.

27. Jobling, S. A.; Gehrke, L. *Nature* **1987,** *325,* 622-625.

28. Holsters, M.; de Waele, D.; Depicker, A.; Messens, E.; Van Montagu, M.; Schell, J. *Mol. Gen. Genet.* **1978,** *163,* 181-187.

29. Dai, Z.; An, G. *Plant Physiol.* **1995,** *109,* 1191-1197

30. Bradford, M. M. *Anal. Biochem.* **1976,** *72,* 248-254.

31. Dai, Z.; Ku, M. S. B.; Zhang, D.; Edwards, G. E. *Planta* **1994,** *192,* 287-294.

32. Dai, Z.; Edwards, G. E.; Ku, M. S. B. *Plant Physiol.* **1992,** *99,* 1426-1434.

33. Van Caemmerer, S. V.; Farquhar, G. D. *Planta* **1981,** *153,* 376-387

34. Pichersky, E.; Bernatzky, R.; Cashmore, A. R.; Gruissem, W. *Proc. Natl. Acad. Sci. USA* **1986,** *82,* 3880-3884

35. Dai, Z.; Hooker, B. S.; Quesenberry, R. D.; Thomas, S. R. *Transgenic Research* **2000,** In Press.

36. Matsuoka, K.; Bassham, D. C.; Raikhel, N. V.; Nakamura, K. *J. Cell. Biol.* **1995,** *130,* 1307-1318.

37. Matsuoka, K.; Nakamura, K. *Proc. Natl. Acad. Sci. USA* **1991,** *88,* 834-838.

38. Nakamura, K.; Matsuoka, K.; Mukumoto, F.; Watanabe, N. *J. Exp. Bot.* **1993,** *44 (suppl),* 331-338.

Chapter 5

Hydrolysis of Mannans by *Cellulomons fimi*

D. Stoll, H. Stålbrand[1], and R. A. J. Warren

Department of Microbiology & Immunology and Protein Engineering Network of Centers of Excellence, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada
[1]**Present address: Department of Biochemistry, Center for Chemistry, Lund University, S–221 00 Lund, Sweden**

Cellulomonas fimi produces a single extracellular mannanase, Man26A, and a single intracellular β-mannosidase, Man2A. The enzymes belong to families 26 and 2 of retaining glycosyl hydrolases, respectively. Man26A is a modular protein, comprising an N-terminal catalytic module, followed by a mannan-binding module, then an S-layer homology module, and finally two repeated modules of unknown function at the C-terminus. Man2A is not modular. Although the mannan-binding module binds to soluble mannans, such as locust bean gum and azo-carob galactomannan, but not to insoluble ivory nut mannan, it enhances the activity of the catalytic module on the insoluble but not on the soluble mannans. Man26A and Man2A together are sufficient to hydrolyze ivory nut mannan to mannose.

Mannans are components of the cell walls of many plants; they may be present in others as storage polysaccharides. They are linear polymers of mannose, or of mannose and glucose (glucomannans), linked β-1,4. Galactosyl substituents may be linked α-1,6 to some of the mannose residues; the 2- or 3-hydroxyls of some of the mannose residues may be acetylated.

The complete hydrolysis of a particular mannan requires some combination of mannanase(s), β-mannosidase, β-glucosidase, α-galactosidase and acetyl esterase. Plants and microorganisms produce mannanases, in some instances more than one (*1-15*). Mannanases (EC 3.2.1.78) hydrolyze β-1,4-mannosidic linkages within the backbones of mannans, galactomannans and glucomannans (*1-9*). Those for which the amino acid sequences are known are in families 5 and 26 of retaining glycosyl hydrolases (*3, 4, 7, 8, 10-14*), with the possible exception of a modular mannanase from *Thermoanaerobacterium polysaccharolyticum* (*15*). β-mannosidases are produced by animals, plants and microorganisms. Some remove mannose residues from the non-reducing ends of oligosaccharides, others process the glycans of

glycoproteins (*9, 16-21*). Those for which the amino acid sequences are known are in families 1 and 2 of retaining glycosyl hydrolases (*17, 18, 20, 21*). Mannanases and β-mannosidases are potentially of use in pulp processing and the conversion of plant materials to useful chemicals.

The aerobic, mesophilic bacterium *Cellulomonas fimi* produces extracellular enzymes that hydrolyze starch, cellulose, chitin and xylans but enzymes from this organism that hydrolyze mannans were not described previously. The genes encoding a mannanase and a β-mannosidase from *C. fimi* have been cloned and sequenced, and the enzymes they encode have been partially characterized (*22*).

A Mannanase from *C. fimi*

C. fimi produces three polypeptides with mannanase activity when grown in the presence of locust bean gum (LBG), a galactomannan. Two major polypeptides of ~75 and ~30 kDa, and a minor polypeptide of ~100 kDa are detectable by zymogram with azo-carob galactomannan (ACG) as substrate in a non-reducing gel with SDS (*22*). The N-terminal amino acid sequence of the 75 kDa polypeptide is APADET. A lambda ZAPII library of *C. fimi* DNA was screened for mannanase-positive clones on plates containing ACG. Thirteen positive clones were obtained. Restriction mapping and partial sequencing showed them all to have DNA sequence in common. Phagemid pCMAN30 carries an open reading frame of 3030 bases, encoding a polypeptide of 1010 amino acids (GenBank AF126471). Amino acids 50-55 in the deduced amino acid sequence are APADET, matching the N-terminal sequence of the mannanase of 75 kDa in the culture supernatant from *C. fimi*. Ala50 in the sequence is designated residue 1 in the mature polypeptide, which is 961 amino acids long, with a calculated molecular weight of 107,033. The N-terminal amino acid sequence of Man26A produced by *E. coli* is APAPAAP, corresponding to residues –14 to –8 relative to the N-terminus of the enzyme produced by *C. fimi*. Either *C. fimi* and *E. coli* use different leader peptide processing sites in pro-Man26A, or *C. fimi* processes the N-terminus further after removing the leader peptide (*22*). Cellobiohydrolase Cel48A (formerly CbhB) of *C. fimi* is also processed differently by *C. fimi* and *E. coli* (*23*).

Alignment of its amino acid sequence with those of proteins in the data base reveals a modular structure for the mannanase (Fig. 1). Its N-terminal amino acid sequence places the mannanase in family 26 of glycosyl hydrolases, which contains mannanases and glucanases. Therefore, the enzyme is designated mannanase 26A (Man26A; see ref. *24*). Amino acids 1-398 share the greatest identity with the non-modular mannanase ManA from *Pseudomonas fluorescens* subsp. *cellulosa*: 28% identical and 13% similar residues (*4, 8, 22*). Amino acids 637-830 of mature Man26A share 23-30 % sequence identity with S-layer homology (SLH) modules, which typically comprise three repeats of about 60 amino acids each (*22, 25*). There are three repeats of about 60 amino acids in this region of Man26A that are 20 % identical: 647-707, 708-765 and 766-825 (*22*). Although a significant fraction of the mannanase activity in cultures of *C. fimi* appeared to be cell-associated, Man26A

produced in *E. coli* did not bind to a peptidoglycan fraction prepared from *C, fimi* (*22*). Amino acids 450-636 and 831-961 of Man26A do not share sequence identity with any proteins in the data base (*22*). The function of the C-terminal amino acids (831-961) is unknown. Amino acids 450-636 contain a mannan-binding module (see below).

⇓   ⇓   ⇓

Catalytic module..Mannan-binding module..SLH module..Unidentified module
1   398 450   636 637   830 831   961
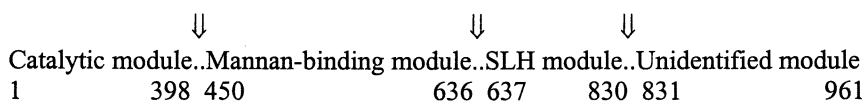
Figure 1. Modules in Man26A. The numbers indicate the position of amino acids in the sequence of the protein. ⇓ *indicates approximate site of cleavage by the protease from C. fimi.*

Other mannanases are also modular proteins. Mannanase ManA from *Caldicellulosiruptor saccharolyticus* comprises an N-terminal family 5 mannanase catalytic module, followed by repeated cellulose-binding modules, and a C-terminal catalytic module with endoglucanase activity (*12*). Mannanase ManA from *T. polysaccharolyticum* is somewhat similar to Man26A; it comprises an N-terminal catalytic module, possibly in family 5 of glycosyl hydrolases, followed by repeated cellulose-binding modules and a C-terminal SLH module (*15*).

Proteolysis of Man26A

The N-terminal amino acid sequences of Man26A produced by *E. coli* and of the 75 kDa mannanase produced by *C. fimi* indicate that the latter may arise by proteolysis of Man26A. *C. fimi* produces an extracellular protease(s) that cleaves its modular polysaccharidases between the modules (*26*). The protease releases several discrete fragments from Man26A produced by *E. coli*, cleaving the enzyme between the modules, thereby confirming the modular structure of the enzyme (Fig. 2). A zymogram of the digest gives a pattern of active bands identical to that for the culture supernatant from *C. fimi* (*22*). Man26A appears to be the only mannanase produced by *C. fimi*. It is converted to active fragments by proteolysis.

Mannanase Man26A has a mannan-binding module

Man26A does not bind to ivory nut mannan (INM), an insoluble, unsubstituted polymer of mannose. Affinity gel electrophoresis shows that it does bind to soluble mannans, such as LBG and ACG. The K_d for binding to LBG is 0.2 μM. Analysis of proteolytic fragments of Man26A, and subcloning and expression of fragments of *man*26A show that the binding is mediated by a sequence within amino acids 438-639 of the enzyme.The module does not align with any polypeptides in the data base. This is the first mannan-binding module to be reported. A striking property of the
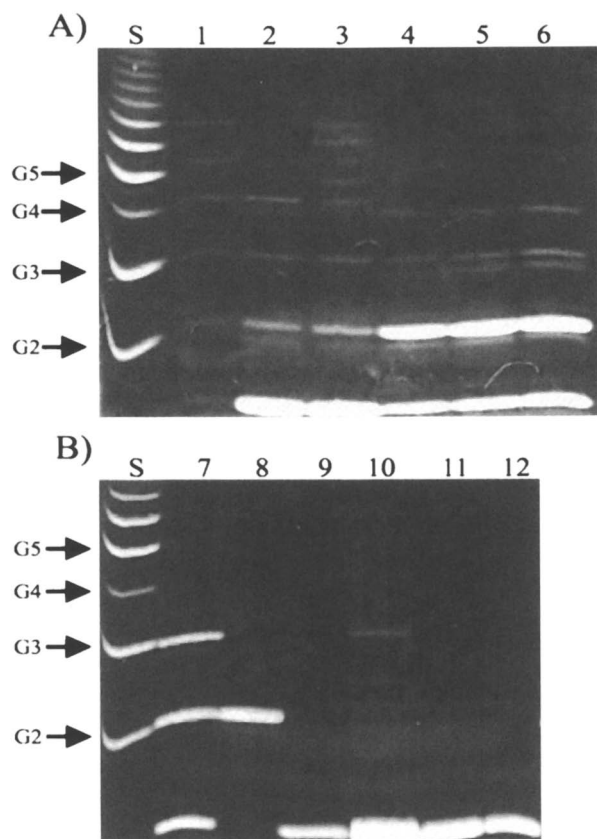
Figure 2. FACE(Fluorophore-assisted carbohydrate electrophoresis) analysis of the products relesade from mannooligosaccharides by Man26A and Man2A. Reaction mixtures contained 0.5 nmol of each enzyme and 1 mM substrate in 100 μL. Incubation was for 1 h at 37 ℃. Lanes: S, α-1,4-glucooloigosaccharide standards as indicated; 1, mannooligosaccharide standards (M1-M6); 2 and 3, Man2A with M4 and M6, respectively; 4-9, Man26A with M4, M5, M6, M3, M2 and mannose, respectively;10-12, Man2A with M3, M2 and mannose respectively.

binding module is that it enhances the activity of the catalytic module on the insoluble but not on the soluble substrates, despite binding to the latter but not to the former. A cellulose- binding module of endoglucanase Cel9A (formerly E4; see ref.*24*) behaves in a similar manner: it enhances the activity of the catalytic module on insoluble but not on soluble cellulose, despite not binding to the insoluble substrate (*27*).

A β-mannosidase from *C. fimi*

The N-terminal amino acid sequence of a β-mannosidase partially purified from cell extracts of *C. fimi* is MITQDLYD. Two mannosidase-positive clones were obtained by screening phagemids derived from the lambda ZAPII library of *C. fimi* DNA, using methylumbelliferyl β-D-mannoside as substrate. Restriction mapping and partial sequencing showed them to have a sequence of *C. fimi* DNA in common. Phagemid PCManII contains an ORF of 2,526 bases, encoding a polypeptide of 842 amino acids with a calculated molecular weight of 94,960. The N-terminal amino acid sequence deduced for the polypeptide matches that of the enzyme partially purified from *C. fimi* (GenBank126472).

Alignment of its amino acid sequence places the mannosidase in a sub-family of family 2 of glycosyl hydrolases. Therefore, it is designated mannosidase 2A (Man2A. It is not a modular protein. Besides Man2A and β-mannosidase MANB from *Aspergillus aculeatus* (*21*), the sub-family contains currently only mammalian mannosidases that are involved in the processing of glycans on proteoglycans (*17, 18, 20*).

Hydrolysis of mannooligosaccharides and mannans by the enzymes from *C. fimi*

Man2A hydrolyzes mannooligosaccharides completely to mannose (Fig. 2). Man26A hydrolyzes mannotetraose, mannopentaose and mannohexaose completely, releasing mannose and mannobiose as the major products; it hydrolyzes mannotriose poorly; it does not hydrolyze mannobiose (Fig. 2).

Man2A and Man26A are retaining enzymes, capable of transglycosylation to form β-1,4 and other β-1 bonds. The traces of longer oligosaccharides in the digests are probably transglycosylation products.

Man2A appeared to be more active on INM than on LBG; it released mannose from from both substrates (Fig. 3). Man26A released a complex mixture of oligosaccharides from LBG, some of which presumably carried galactosyl substituents, but mostly mannose, mannobiose and mannotriose from INM (Fig. 3). Except for the absence of mannose, Man2A and Man26A together released a similar mixture of products from LBG as Man26A alone (Fig. 3). They released only mannose from INM (Fig. 3).
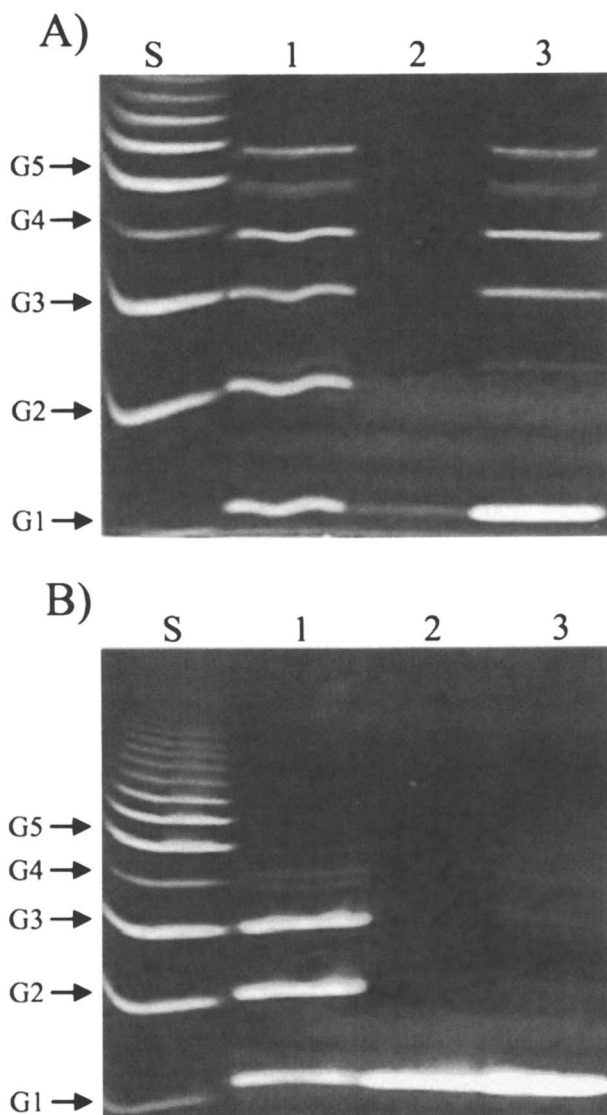
Figure 3. *FACE analysis of the products released from locust bean gum (panel A) and ivory nut mannan (panel B) by Man26A and Man2A. The reaction conditions and incubation were as for Fig. 2, except for the substrate concentration of 1 mg mL^{-1}. Lanes: S, α-1,4-glucooligosaccharide standards as indicated; 1, Man26A; 2, Man2A; 3, Man26A plus Man2A.*

Acknowledgments

References

1. McLeary, B. V.; Taravel, F. R. *Carbohyd. Res.* **1983,** *118,* 91.
2. Stålbrand, H.;Siika-Aho, M.; Tenkanen, M.; Viikari, L. *J. Biotech.* **1993,** *29,* 229.
3. Arcand, N.; Kluepfel, D.; Paradis, F. W.;Morosoli, R.; Sharek, F. *Biochem. J.* **1993,** *290,* 857.
4. Braithwaite, K. L.; Black, G. W.; Hazlewood, G. P.; Ali, B. R.; Gilbert, H. J. *Biochem. J.* **1995,** *305,* 1005.
5. Tamaru, Y.; Araki, T.; Amagoi, H.; Mori, H.; Morishita, T. *Apl. Environ. Microbiol.* **1995,** *61,* 4454.
6. Harjunpa, V. A.; Teleman, A.; Siika-Aho, M.; Drakenberg, T. *Eur. J. Biochem.* **1995,** *234,* 278.
7. Morris, D. D.; Reeves, R. A.; Gibbs, M. D.; Saul, D. J.; Bergquist, P. *Appl. Environ. Microbiol.* **1995,**.*61,* 2262.
8. Bolam, D. N.; Hughes, N.; Virden,R.; Lakey, J. H.; Hazlewood, G. P.; Henrisat, B.; Braithwaite, K. L.; Gilbert, H. J. *Biochemistry* **1996,** *35,* 16195
9. Duffaud, G. D.; McCutcheon, C. M.; Leduc, P.; Parker, K. N.; Kelly, R. M. *Appl. Environ. Microbiol.* **1997,** *63,* 169.
10. Hilge, M.; Gloor, S. M.; Rypniewski, W.; Sauer, O.; Heightman, T. D.; Zimmerman, W.; Winterhalter, K.; Piontek, K. *Structure* **1998,** *6,*1433.
11. Stålbrand, H.; Saloheimo, A.; Vehmaanpera, J.; Henrissat, B.; Penttila, M. *Appl. Environ. Microbiol.* **1995,** *61,* 1090.
12. Gibbs, M. D.; Saul, D. J.; Lüthi, E.; Bergquist, P. L. *Appl. Environ. Microbiol.* **1992,** *58,* 3864.
13. Millward-Sadler, S. J.; Hall, J.; Black, G.W.; Hazlewood, G. P.; Gilbert, H.J. *FEMS Microbil. Lett.* **1996,** *141,* 183.
14. Ethier, N.; Talbot, G.; Sygusch, J. *Appl. Environ. Microbiol.* **1998,** *64,* 4428.
15. Cann, I. K. O.; Kocherginskaya, S.; King, M. R.; White, B. R.; Mackie, R. I. *J. Bacteriol.* **1999,** *181,* 1643.
16. Neustroev, K. N.; Krylov, A. S.; Firsov, L. M.; Abrosinka, O.N.; Khorlin, A. Y. *Biokhimiya* **1991,** *56,* 1406.
17. Chen, H.; Leipprandt, J. R.; Travis, C. E.; Sopher, B. L.; Jones, M. Z.; Cavanagh, K. T.; Friderici, K. H. *J. Bilo. Chem.* **1995,** *270,* 3841.
18. Leipprandt, J. R.; Kraemer, S. A.; Haithcock, B. E.; Chen, H.; Dyme, J. L.; Friderici, K. H.; Jones, M. Z. *Genomics* **1996,** *37,* 51.
19. Bauer, M. W.; Bylina, E. J.; Swanson, R. V.; Kelly, R. M. *J. Biol. Chem.* **1996,** *271,* 23749.

20.   Alkhayat, A. H.; Kraemer, S. A.; Leipprandt, J. R.; Macek, M.; Kleijer, W. J.; Friderici, K. H. *Hum. Mol. Genet.* **1998,** *7,* 75.

21.   Takada, G.; Kawaguchi, T.; Kaga,T.; Sumitani, J-i.; Arai, M. *Biosci. Biotechnol Biochem.* **1999,** *63,* 209.

22.   Stoll, D.; Stålbrand, H.; Warren, R. A. J. *Appl. Environ. Microbiol.* **1999,** *65,* 2598.

23.   Shen, H.; Gilkes, N. R.; Kilburn, D. G.; Miller, R. C., Jr.; Warren, R. A. J. *Biochem. J.* **1995,** *311,* 67.

24.   Henrissat, B.; Teeri, T. T.; Warren, R. A. J. *FEBS Lett.* **1998,** *425,* 352.

25.   Olabarria, G.; Carrascosa, J.; Pedro, M. A.; Berenguer, J. *J Bacteriol.* **1996,** *178,* 4765.

26.   Sandercock, L.; Meinke, A.; Warren, R. A. J. *FEMS Microbiol. Lett* . **1996,** *143,* 7.

27.   Irwin, D.; Shin, D.-H.; Zhanf, S.; Barr, B. K.; Sakon, K.; Karplus, P. A.; Wilson, D. B. *J. Bacteriol.* **1998,** *180,* 1709.

Chapter 6

An Overview of Factors Influencing the Enzymatic Hydrolysis of Lignocellulosic Feedstocks

Ali R. Esteghlalian[1], Vinit Srivastava[1,2], Neil Gilkes[2], David J. Gregg[1], and John N. Saddler[1,3]

[1]Chair of Forest Products Biotechnology, Forest Sciences Centre and [2]Department of Microbiology and Immunology, The University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada
[3]Corresponding author

Lignocellulosic biomass conversion to ethanol promises to provide an environmentally benign alternative fuel that can reduce the consumption of gasoline, thereby reducing our dependence on a non-renewable energy source and improving the urban air quality. The process of microbial degradation of biomass is not understood at the molecular level and yet it is clear that many microorganisms have evolved diverse mechanisms to accomplish this task. We review current concepts regarding enzymatic hydrolysis of bioethanol feedstocks and point to technical challenges requiring further R&D.

Combustion of gasoline in road vehicle engines generates large amounts of greenhouse gases, ozone, and volatile organic compounds, all of which have been shown to have a discernible impact on global climate and human health. In addition to providing a cleaner fuel, biomass-to-ethanol schemes can circumvent the adverse impacts of traditional waste management practices, such as field burning or wood waste incineration and stockpiling that can result in air pollution and groundwater contamination.

Many species of trees and herbaceous plants have been considered as possible sources of biomass for bioconversion to simple sugars and subsequently to ethanol. Herbaceous biomass includes agricultural residues derived from wheat, barley, rice and corn, as well as grasses (e.g., lovegrass and switchgrass) that can be grown as bioenergy crops.

Wood from both hardwood (deciduous) and softwood (coniferous) trees are another potential feedstock for bio-ethanol production. Short-rotation hardwood trees, such as poplar and willow, have been considered as likely feedstock candidates due to their fast growth and high yield (5-20 dry ton biomass/acre/year). Use of wood and forestry residues for bio-ethanol production has received considerable attention in the USA, Canada and Scandinavia because their vast evergreen forests are viewed as a renewable resource. In addition, wood residues generated by forest products industries can serve as a sustainable source of feedstock for a bioconversion process.

Although the major components of various forms of biomass are the same (cellulose, hemicellulose and lignin), each type of feedstock has different processing requirements and potential ethanol yield as a result of chemical and structural variations. In general, softwoods are more difficult to process than hardwoods and herbaceous crops as they contain higher quantities of guaiacyl lignin which is generally more resistant to chemical modification than the syringyl based lignin primarily found in hardwoods and grasses (1). Although cellulose in all types of biomass is similar, there are major differences in their hemicellulose composition. For example, arabinoxylan is the principal type of hemicellulose found in hardwoods and herbaceous plants, whereas galactoglucomannans constitute the major hemicellulose fraction in softwoods. The difference in hemicellulose chemistry has important implications for the down-stream processes such as fermentation.

The major steps in a potential biomass-to-ethanol process are feedstock preparation (cleaning and size reduction), pretreatment, enzymatic hydrolysis and ethanologenic fermentation. The pretreatment step is an acid or alkali hydrolysis process that removes, or otherwise alters, the hemicellulose and lignin fractions producing a highly digestible solid rich in cellulose; and a soluble hemicellulose fraction known as the prehydrolysate. The solubilization of the hemicellulose renders the remaining cellulose more porous, and therefore, more accessible to cellulase enzymes. At the same time, disruption of the lignin-cellulose association increases the exposed surface area of the cellulose structure, again increasing the efficiency of enzyme action (1). It has been shown that the prehydrolysate contains monomeric and oligomeric pentoses and hexoses that differ with the feedstock, and that typically, further chemical treatment is needed to neutralize and detoxify the hemicellulosic fraction prior to fermentation.

Hydrolysis of cellulose to glucose can be achieved using either inorganic acids or cellulolytic enzymes. Despite its relatively slow rate, enzymatic hydrolysis is a viable process due to several major advantages. Enzymatic hydrolysis avoids use of corrosive chemicals and requires less costly equipment. Unlike acid hydrolysis, it does not result in a significant loss of sugars and consequent reduction of actual ethanol yield, and it also minimizes the formation of toxic by-products that inhibit the growth of microorganisms during fermentation.

Techno-economic models show that the enzymatic hydrolysis step constitutes a large portion of the overall cost of biomass to ethanol process (2). The most significant feature of the hydrolysis reaction is the biphasic behavior that involves an initial fast reaction followed by a slower asymptotic phase. With an enzyme loading of about 20 Filter Paper Units (FPU)/g cellulose, up to 55 % of the original cellulose is hydrolyzed within 24 hours (fast), while hydrolysis of the remaining cellulose can

take up to three more days (slow). Such a long residence time adds to the operating cost of the process and lowers its economic feasibility.

The substrate physical form and chemical structure as well as enzyme related functions, such as specific activity, enzyme composition and kinetics, all contribute to the inefficiency of the hydrolysis reaction. Factors that limit the rate and extent of lignocellulose hydrolysis are summarized in Table 1. Although 'fine' characteristics of the cellulosic substrates (degree of polymerization [DP], crystallinity index [CrI], lignin content, and particle size) play a major role in determining the rate and extent of hydrolysis, "gross" substrate characteristics are also important (4, 5). While cellulose accessibility is generally considered to be the major determinant of hydrolytic efficiency, the enzyme-related factors, such as entrapment of enzymes by lignin or loss of synergism due to separation of large and small enzyme components can significantly contribute to this end (4).

Table 1. Substrate and enzyme related factors that influence the enzymatic hydrolysis of cellulose in lignocellulosic feedstocks

Enzyme related factors	*Substrate related factors*
- Reaction heterogeneity (soluble enzyme vs. insoluble substrate) - Irreversible binding of enzymes onto lignin - Gradual loss of synergism in cellulase mixture - Substrate dependence of synergism and binding (specificity) of enzyme components - End-product inhibition - Thermal inactivation of enzyme	- Cellulose crystallinity (CrI) - Cellulose degree of polymerization (DP) - Feedstock particle size - Lignin barrier (content and distribution) - Substrate's available surface area (pore volume) - Cell wall thickness (coarseness)

Structure and Accessibility of Lignocellulosic Substrates

The fiber is the simplest unit of the macro-structure in lignocellulosic substrates. Each fiber is an elongated plant cell containing three concentric cell wall layers; primary (outermost), secondary (middle) and warty (innermost). The secondary wall consists of three layers, S_1, S_2 and S_3: the S_2 layer is the thickest and contains about 30-150 sub-layers (3) and it generally defines the overall properties of the fiber. A highly lignified layer known as the middle lamella acts as the interface between two adjacent fibers. The main constituents of the cell wall, (cellulose, hemicellulose and lignin), along with other non-structural compounds, are intimately associated to create the structural framework of the plant.

The smallest unit of fiber organization are elementary fibrils consisting of bundles of 30-40 cellulose chains held together by inter- and intra-molecular hydrogen bonds and with diameters of 1-10 nm. Two or more elementary fibrils, with hemicelluloses intimately associated by non-covalent interactions, form microfibrils of 10-30 nm diameter with intermicrofibrillar spaces of about 10 nm. Microfibrils associate into larger bundles (macrofibrils) that form the three layers in the fiber wall.

Barriers to enzyme penetration are present at the microfibril, fibril and fiber level (5). At the microfibril level, cellulose is primarily in a crystalline form. Native cellulose occurs as cellulose I in which individual chains have a parallel organization. A proportion of the cellulose, with the extent dependent on prior treatment, has a paracrystalline or "amorphous" structure. The DP of cellulose chains in lignocellulose varies widely (15-15,000), depending on source (7). The CrI and DP have been considered as important factors in determining the hydrolysis rates of relatively refined cellulosic substrates, but data from several independent investigations indicate that these parameters alone do not explain the recalcitrance of lignocellulosic substrates (5-7). For example, while pretreatment of lignocellulosics improves hydrolyzability, it often increases the CrI of the cellulose fraction (5, 7). Moreover, there is a strong negative correlation between specific surface area and CrI: this suggests that the hydrolyzability of cellulose is largely determined by the accessibility of cellulose to cellulases with crystallinity becoming less of an important factor if the accessibility is sufficiently increased (4).

At the fibril level, the intimate association of lignin and hemicellulose creates very narrow interfibrillar spaces that impede enzyme accessibility. In general, fiber capillaries are of two different categories. The "gross" capillaries, such as the cell lumen, pit aperture and pit-membrane pores, are relatively large (0.02-10 microns) (7). Cell-wall capillaries are much smaller and include naturally occurring intermicrofibrillar cavities and those formed during pretreatment by the dissolution of hemicellulose and lignin. Gross capillaries, usually, constitute only a small fraction of total available surface and are easily accessible by enzymes (7) but smaller capillaries remain inaccessible unless enlarged by swelling, steam explosion or delignification pretreatments.

There appears to be some correlation between available surface area (pore volume) and substrate digestibility (8-14). The available surface area can be measured by solute exclusion techniques using a series of dextran spheres with diameters ranging from 0.8 to 160 nm. It is considered that the critical pore size corresponds to dextran probe of 5.1 nm diameter, which is equal to the average diameter of a hydrated cellulase molecule (11). However, although porosity is related to enzyme accessibility, pretreatment (steam explosion or acid hydrolysis) or pulping procedures that remove the majority of hemicelluloses leave behind lignin that restricts the accessibility of cellulose to enzymes. For example, swelling of Kraft pulp by sulfonation, without delignification, did not result in improved hydrolysis, although it increased the population of 5.1 nm pores. In contrast, delignification of pulp enhanced both its porosity and hydrolyzability (10). These data show the constraints imposed by lignin, even in a highly swollen substrate.

It has been shown that lignin can impede the hydrolytic reaction by irreversibly binding to enzymes, and effectively decreasing enzyme concentration thereby

lowering the hydrolytic efficiency (8, 10). At the fiber level, coarseness (average fiber weight:length ratio), which is a measure of fiber wall thickness, is closely related to specific surface area (available surface area per unit weight of a fiber). There is an obvious correlation between coarseness and hydrolysis rate (15). It has also been suggested that there are different mechanisms involved in the hydrolysis of coarse and fine fibers: hydrolysis of coarse fibers is initiated by 'peeling' into smaller fibers, while the finer particles are simply 'eroded' (12).

Pretreatment of Lignocellulosics to Enhance Enzymatic Hydrolysis

A chemical or physical pretreatment that hydrolyzes the hemicellulosic polysaccharides, renders the pretreated feedstock more amenable to enzymatic digestion. In fact, a lower hemicellulose content in the feedstock is directly related to cellulose digestibility (16-18). Formation of free chain ends is another important effect of pretreatment, as the free ends are a prerequisite for the enzymatic attack by exo-acting enzymes (19). Conditions must be optimized to minimize the formation of sugar degradation products (e.g., hydroxymethyl furfural, furfural and acetic acid) that inhibit microbial pentose and hexose fermentation (17). Dilute-sulfuric acid hydrolysis and acid-catalyzed steam explosion have been shown to be among the most promising pretreatment processes identified so far (17). Both operations require high temperatures (170-250°C), low acid concentrations (1-5% w/w), and short residence times (a few seconds to 5 min). Using various acid-catalyzed pretreatment schemes, 80-90% hemicellulose sugar recovery yields have been achieved (17, 18).

Acid pretreatment removes easily hydrolyzable carbohydrates and increases pore volume and enzyme accessibility. Combination of steam-explosion with acid treatment also reduces particle size, fibrillates cellulose microfibrils and expands the lignocellulosic matrix. Interestingly, the explosion step in steam-explosion seems to be unnecessary as improved hydrolysis cab still be obtained when pressure is reduced slowly, rather than suddenly as is achieved during steam explosion (20).

Although it was found that sequential steam-explosion and alkali-peroxide washing of aspen wood (hardwood) improved the rate and yield of hydrolysis by about 10% (1), the same procedure carried out with softwoods reduced the hydrolyzability of these pretreated substrates (8). It is probable that the difference in the behavior of hardwoods and softwoods can be explained by the differences in cell morphology. The highly lignified primary wall of fibers appears to constrain the secondary wall that constitutes the bulk of the fiber. Delignification removes this constraint allowing the underlying layers to expand (a phenomenon known as "ballooning" (3), thereby increasing enzyme accessibility. The reduced digestibility of softwoods after alkali-peroxide washing has been attributed to a partial collapse of pore structure (16). Another possible explanation is that alkali treatment causes extensive delignification of the heavily lignified S_2 layer and that under the mild conditions of alkali treatment, this lignin can be reprecipitated on cellulose microfibrils, presenting a new barrier to hydrolysis (8). Lignin composition also effects cellulose digestibility. The large vessel elements of hardwoods, which should be the most accessible hardwood cell type, are actually most resistant to enzymatic attack, probably due to their higher guaiacyl to syringyl lignin ratio (1, 8).

In contrast to steam-explosion and alkali-peroxide washing, sequential steam-explosion and water washing enhances the hydrolyzability of both softwoods and hardwoods (*1, 8*). The wash contains mostly sugars derived from hemicellulose and usually requires some further processing to increase sugar concentration and remove the inhibitory products prior to fermentation. Also, washing must be done at low temperature to avoid degradation of monomeric sugars by acid carried over from the pretreatment stage.

Steam explosion procedures require optimization for each type of lignocellulosic feedstock. Very severe reaction conditions can result in the removal of almost all of the hemicellulose and provide highly digestible solids. However, it also promotes sugar degradation and may solubilize some part of the cellulose fraction. Too mild conditions, on the other hand, produce a low yield of oligomeric hemicellulose-derived sugars that need further hydrolysis before fermentation, and a cellulose fraction that is still resistant to hydrolysis (*21*). Therefore, compromise conditions have to be defined to produce adequate levels of hemicellulose solubilization and acceptable levels of soluble sugar degradation. These types of optimal conditions have been reported for steam explosion of softwoods (*21*).

The Mechanism of Cellulose Hydrolysis

Even if the complications of gross substrate characteristics are ignored, there are significant gaps in our understanding of enzymatic cellulose hydrolysis. Microorganisms appear to have evolved a number of different strategies to hydrolyze cellulose in biomass to simple soluble sugars. Here, we briefly discuss cellulolytic systems from bacteria and fungi that use cellulosic biomass as a source of energy. Cellulolytic microorganisms, such as *Azoarcus* spp. that produce cellulases and related glycosyl hydrolases to mediate infection (*22*) are not considered.

Typically, cellulolytic microorganisms produce an array of β-1,4-glucanases during growth on cellulosic substrates. These include endoglucanases that attack β-1,4-glucan chains randomly and exoglucanases, usually cellobiohydrolases, with a strong preference for acting at chain ends (*5, 23*). Despite intensive research in recent years, we still lack a comprehensive understanding of the mechanisms that allow these enzymes to act in a concerted fashion to efficiently hydrolyze cellulosic biomass. Most accounts of enzymatic cellulose hydrolysis continue to feature a simple model in which endoglucanases, acting at easily accessible sites, create new sites for the attack of crystalline regions by exo-acting cellobiohydrolases. The soluble sugars produced by this synergistic interaction are then hydrolyzed to glucose by β-glucosidase (*24, 25*). However, the simplicity of this model may be contrasted with the complexity of cellulase systems that typically contain at least two cellobiohydrolases and multiple endoglucanases. The cellulolytic bacterium *Clostridium thermocellum* produces as many as 15 β-1,4-glucanases (*26*). This apparent redundancy of enzyme function probably indicates that important features of enzymatic cellulose hydrolysis remain to be explained, and that we have probably underestimated the structural complexity of cellulosic substrates.

In general, aerobic fungi (e.g., *Trichoderma reesei*, *Phanerochaete chrysosporium* and *Poria placenta*) and aerobic bacteria (e.g., *Cellulomonas fimi* and *Thermomonospora fusca*) secrete cellulases individually in the external medium: these systems are described as "cell-free" or "non-complexed" (*24, 25*). The component enzymes usually contain a catalytic domain joined to an independent cellulose-binding domain (CBD) that mediates adsorption to the substrate (Figure 1).

Close association of the cellulases with cellulose results in the formation of multiple, non-covalent and thermodynamically favorable bonds that stabilize the protein on cellulose and mediate its penetration into the intermicrofibrillar space. The presence of the large enzyme within such narrow spaces causes an increase in the mechanical pressure exerted on the cavity walls thereby swelling the cellulose structure and accommodating more and more water molecules between the microfibrils. Water breaks the hydrogen bonds and helps disassociate the individual microfibrils. Enzymes thus adsorbed prevent the solvated chains and free chain ends from being reannealed (*27*).

Some of the bacterial cellulases contain additional structural and functional modules and a few contain two catalytic domains. On the other hand, cellulase systems from anaerobes are organized into very high molecular weight complexes called cellulosomes (*23*). In addition to their catalytic domain, such enzymes contain a "dockerin" domain. This small terminal domain interacts strongly with one of several "cohesin" domains, on a large structural molecule called a scaffoldin, to produce a cellulosome complex. In this case, a CBD on the scaffoldin molecule is responsible for substrate adsorption. Originally discovered in *C. thermocellum*, evidence for similar complexes has since been found in several other anaerobic bacteria and more recently in anaerobic fungi. Cellulosome-like complexes appear to be a general feature of anaerobic cellulase systems (*23, 28*).

The availability of an increasingly large database of catalytic domain amino acid sequences following the application of molecular biology to cellulase research led to their classification into structurally related families; β-1,4-glucanases are currently classified into at least 13 families (*29*). The concept that sequence similarity within families reflects an underlying similarity in catalytic mechanism and three-dimensional fold has been confirmed by X-ray crystallographic analysis of several representative enzymes (*30*). More recently, it was recognized that some families can be grouped into "clans", which are groups of families that lack common sequence similarities but share a common global fold in which the positions of key catalytic residues are conserved (*23, 31*).

The distinction between randomly acting endoglucanases and exo-acting cellobiohydrolases, long recognized on the basis of biochemical analyses, is now strongly supported by high-resolution X-ray crystallographic data (*32, 33*). Cellobiohydrolases have generally been shown to have tunnel-shaped active sites that restrict access to the chain ends of long cellulose molecules, whereas endoglucanases generally have open, cleft-shaped active sites that offer relatively unrestricted access to substrate. While the difference between endoglucanases and cellobiohydrolases may not be absolute (the loops enclosing the tunnel in cellobiohydrolases may open infrequently to permit endo attack) the distinction remains useful and the physiological significance of endo-activity by cellobiohydrolases is questionable.

Catalytic Domain

Linker

Binding Domain

ENZYME

H_2O

CBD

Small cellulosic particles

1

2

3

- Adsorption of CBD at cellulose micro-cracks

- Penetration of CBD into the interfibrillar space

- Insertion of mechanical pressure on pore walls

- Penetration of H_2O molecules into the interfibrillar space and cleavage of hydrogen bonds

- Formation of free chain ends
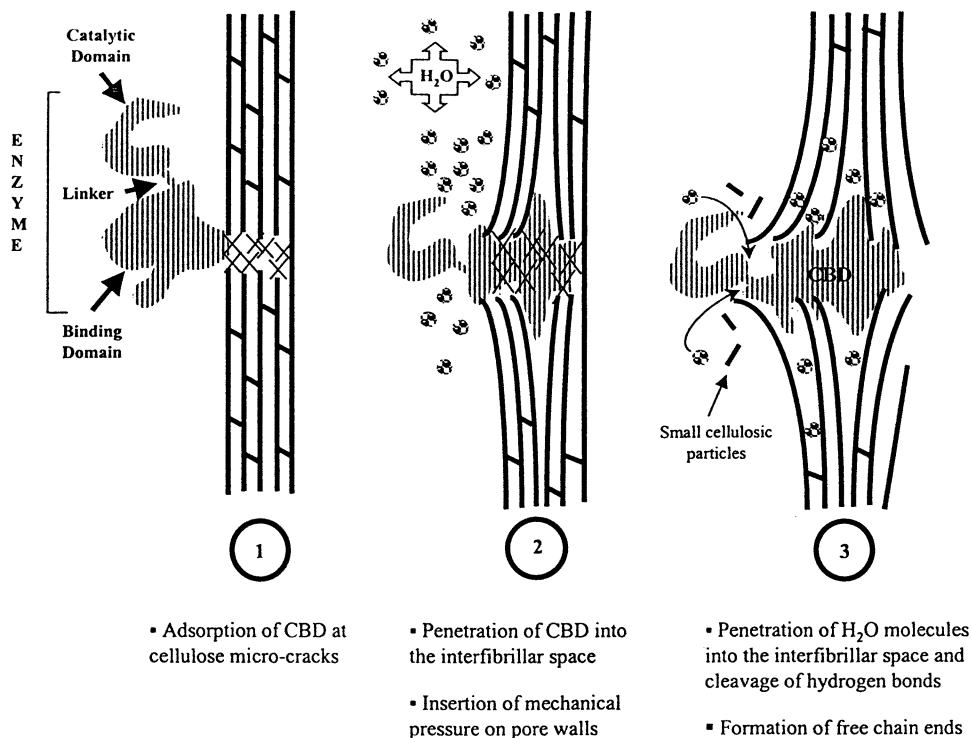
Figure 1. A schematic representation of enzymatic dispersion of cellulose mediated by the initial adsorption of the cellulose-binding domain (CBD) at amorphous regions. Penetration of enzyme's CBD followed by rush of water molecules increases the pressure on interfibrillar cell walls and disrupts the ordered structure. Disassociation of cellulose chains creates free chain ends that will be acted upon by exoglucanases.

As noted above, simple models of cellulose hydrolysis fail to explain the necessity for rather complex cellulase systems comprised of multiple enzymes with similar specificities. The occurrence of two cellobiohydrolases in several well-characterized aerobic cellulase systems is one instance of enzyme multiplicity, but recent data offer a partial explanation for this apparent redundancy. The general assumption that all cellobiohydrolases attack from the non-reducing end was questioned following the demonstration of two enzymes in *Aspergillus aculeatus* that attack in opposite directions (*34*). There is now evidence that several non-complexed systems contain a similar pair of cellobiohydrolases with opposite directions of attack. For example, it has been shown that a family 6 cellobiohydrolase can attack from the non-reducing end (in both bacterial and fungal systems) and a family 48 (bacterial) or family 7 cellobiohydrolase (fungal) can attack from the non-reducing end (*35, 36*).

The production of two cellobiohydrolases in these systems implies that cellulose has structural features related to hydrolysis that we have yet to understand. Further levels of complexity are indicated by the white-rot fungus *P. chrysosporium* which produces one family 6 cellobiohydrolase and six family 7 cellobiohydrolase isozymes (Cel7A-Cel7B). The expression of the family 7 isozymes appears to be regulated by the substrate because distinctly different levels of transcription of the cel7 genes are seen during growth on defined medium supplemented with glucose or cellulose or on wood chips (*37*). If this is the case, then *P. chrysosporium* offers a new approach to understanding the process of enzymatic cellulose hydrolysis and a potential "rational" approach to strain improvement for specific applications.

Analysis of hydrolysis of a model cellulose substrate derived from *Cladophora* suggests that cellulases can discriminate between the two distinct crystalline phases (called Iα and Iβ) that occur in highly crystalline native cellulose. Other experiments using model substrates synthesized by the bacterium *Acetobacter xylinum* indicate that hemicelluloses can profoundly influence the way that cellulose is assembled (*38*). It seems likely that this *Acetobacter* model reflects the nature of cellulose formation in higher plants where hemicellulose occurs in very close association with cellulose microfibrils, and that the key to understanding the complexity of cellulase systems at the microfibril level lies in a more detailed consideration of the substrate (*39*).

Early work on cellulose hydrolysis by fungi led to speculation that cellulase systems contain an independently-acting, non-hydrolytic factor, termed C_1, that renders crystalline cellulose more accessible to subsequent attack by hydrolytic enzymes (collectively called C_x) through disruption of non-covalent bonds (*40*). With the realization that many cellulases have a modular organization, it was suggested that CBDs may correspond to the C_1 component and evidence for non-hydrolytic disruption of ramie and cotton fibers was obtained using an isolated CBD derived from *C. fimi* endoglucanase CenA (*41, 42*). However, these experiments provided no indication of disruption of a highly crystalline cellulose from *A. xylinum* and CBDs from other cellulases appear unable to disrupt even cotton fibers (*43*). Consequently, the significance of fiber disruption by the CenA CBD is still unresolved. Removal of CBD from a range of modular cellulases has been shown to reduce activity on insoluble substrates but not on soluble substrate analogues, an effect that can be attributed to CBD-mediated concentration of enzyme on the cellulose surface. Targeting of enzymes to particular regions of cellulose structure is

another possible function for CBDs. In this context, it is interesting that other hydrolases involved in biomass degradation, notably xylanases, contain CBDs (*25*). Presumably in these instances, the CBD targets the catalytic domain to regions of the plant cell wall where cellulose and hemicellulose are intimately associated.

Recently, a 50 kDa protein called swollenin was reported in *T. reesei* (*44*). Incubation of cotton fibers with a crude swollenin preparation appears to result in an opening and swelling of fiber structure without hydrolysis. The protein is comprised of an N-terminal CBD joined by a linker to a region with sequence similarity to expansins (molecules implicated in cell wall extension in higher plants). Similarly, an 11 kDa "fibril-forming protein", also from *T. reesei*, was reported to cause non-hydrolytic disruption of filter paper (*45*). Although there is presently no convincing evidence for their role in cellulose hydrolysis, the presence of a non-hydrolytic factor remains intriguing and continues to provide a possible avenue towards more efficient biomass hydrolysis schemes.

Conclusion

Improvements in enzymatic hydrolysis of lignocellulosics require further investigations in both fundamental and applied dimensions. While Nature seems to have provided the most effective and efficient cellulolytic enzymes, there is still much to be done to make the best use of these biological catalysts. Enzyme production may always be a major part of the operating cost in a bioconversion process, however, advances in reactor design and development of novel schemes for recycle and reuse of enzyme can enhance the efficacy of enzymes (g glucose produced/g enzyme used/hour reaction time). It is probable that increased understanding of the mechanisms of enzyme action and elucidating the function of the various enzyme domains in combination with results of future studies on chemical and structural changes of substrate during various operations, e.g., pretreatment, delignification and hydrolysis, will all contribute to making the biomass-to-ethanol process an economically viable endeavor.

Literature Cited

1. Ramos, L. P.; Breuil, C.; Saddler, J. N. *Applied Biochem. and Biotech.* 1992, *34/35*, 37-47.
2. Gregg, D. J.; Boussaid, A.; Saddler, J. N. *Bioresource Technol.* 1998, *63*, 7-12.
3. Sjostrom, E. *Wood Chemistry*. Academic Press: San Diego, CA, 1993.
4. Converse, A. O. In *Bioconversion of Forest and Agricultural Plant Residues;* Saddler, J. N.; Ed.; Biotechnology in Agriculture No. 9; C.A.B. International: Wallingford, UK, 1993, pp 93-106.
5. Mansfield, S. D.; Mooney, T.; Saddler, J. N. *Biotechnol. Prog.* 1999, *15*, 804-816.
6. Jorgensen, O. B.; Cowan, D. In *Enzyme Systems for Lignocellulose Degradation*; Coughlan, M. P.; Ed.; Elsevier Applied Science: London, 1989, pp 347-356

7. Cowling, E.; In *Cellulose as a Chemical and Energy Resource;* Wilke, C. R.; Ed.; Biotechnology and Bioengineering Symposium No. 5; Wiley: New York, 1975, pp 163-181.

8. Ramos, L. P.; Saddler, J. N. *Applied Biochem. Biotechnol.* 1994, *45/46,* 193-207.

9. Sinitsyn, A. P.; Gusakov, A. V.; Vlasenko, E. Y. *Applied Biochem. Biotechnol.* 1991, *30,* 43-59.

10. Mooney, T.; Mansfield, S.D.; Tuohy, M.; Saddler, J. N. *Bioresource Technol.* 1998, *64,* 113-119.

11. Stone, J., Scallan; A., Donefer, E.; Ahlgren, E. *Advances in Chemistry Series.* 1969, *95,* 219-241.

12. Allan, G. G.; Ko, Y. C.; Ritzenthaler, P. *Tappi* 1991. *3,* 205-212.

13. Peters, L. E.; Walker, L. P.; Wilson, D. B.; Irwin, D. C. *Bioresource Technol.* 1991, *35,* 313-319.

14. Grethlein, H. E. *Bio/Technology* 1985, *3,* 155-160.

15. Mooney, T.; Mansfield, S. D.; Beatson, R.; Saddler, J. N. *Enzyme and Microbial Technology. in press.*

16. McMillan, J. D. Report No. NREL/TP-421-4978. National Renewable Energy Laboratory: Golden, CO, 1992.

17. Saddler, J. N.; Ramos, L. P.; Breuil, C. In *Bioconversion of Forest and Agricultural Plant Residues;* Saddler, J. N.; Ed.; Biotechnology in Agriculture No. 9; C.A.B. International: Wallingford, UK, 1993, pp 73-92.

18. Esteghlalian, A.; Hashimoto, A.; Fenske, J.; Penner, M. *Bioresource Technol.* 1997, *59,* 129-136.

19. Nieves, R.; Ellis, R.; Todd, R.; Johnson, J. A.; Grohmann, K.; Himmel, M. *Applied and Env. Microb.* 1991, Vol. 57, *11,* 3163-3170.

20. Brownell, H. H.; Yu, E. K. C.; Saddler, J. N. *Biotechnol. Bioeng.* 1986, *28,* 792-801.

21. Boussaid, A.; Esteghlalian, A.; Gregg, D. J.; Lee, K.; Saddler, J. N. *Applied Biochem. Biotechnol.* 1999. *in press.*

22. Reinhold-Hurek, B.; Hurek T.; Claeyssens, M.; van Montagu, M. *J. Bacteriol.* 1993, *175,* 7056-7065.

23. Bayer, E. A.; Chanzy, H.; Lamed, R.; Shoham, Y. *Curr. Op. Struct. Biol.* 1998, *8,* 548-557.

24. Béguin, P.; Aubert. J. P. *FEMS Microbiol. Rev.* 1994, *13,* 25-58.

25. Tomme, P.; Warren, R. A. J.; Gilkes, N. R. *Adv. Microb. Physiol.* 1995, *37,* 1-81.

26. Hazlewood, G. P.; Romaniec, M. P. M.; Davidson, K.; Grépinet, O.; Béguin, P.; Millet, J.; Raynaud, O.; Aubert, J.P. *FEMS Microbiol. Lett.* 1988, *51,* 231-236.

27. Klyosov, A. *Biochemcistry* 1990, 29 (47), 10577-10585.

28. Bayer, E. A.; Shimon, L. J. W.; Shoham, Y.; Lamed, R. *J. Struct. Biol.* 1998, *124,* 221-234.

29. Coutinho, P. M.; Henrissat, B. Carbohydrate-Active Enzymes. URL http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html., 1999.

30. Gebler, J.; Gilkes, N. R.; Claeyssens, M.; Wilson, D. B.; Béguin, P.; Wakarchuk, W. W.; Kilburn, D. G.; Miller, R. C. Jr.; Warren, R. A. J.; Withers, S. G. *J. Biol. Chem.* 1992, *267,* 12559-12561.

31. Henrissat, B.; Davies G. *Curr. Opin. Struct. Biol.* 1997, *7*, 637-644.
32. Divne, C.; Ståhlberg, J.; Reinikainen, T.; Ruohonen, L.; Petterson, G.; Knowles, J. K. C.; Teeri, T. T.; Jones, T. A. *Science* 1994, *265*, 524-528.
33. Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. *Science* 1990, *249*, 380-386.
34. Arai, M.; Sakamoto, R.; Murao, S. *Agric. Biol. Chem.* 1989, *53,* 1411-1412.
35. Gilkes, N. R.; Kwan, E.; Kilburn, D. G.; Miller, R. C.; Warren, R. A. J. *J. Biotechnol.* 1997, *57*, 83-90.
36. Imai, T.; Boisset, C.; Samejima, M.; Igarashi, K.; Sugiyama, J. *FEBS Lett.* 1998, *432*, 113-116.
37. Vallim, M. A.; Janse, B. J. H.; Gaskell, J.; Pizziranikleiner, A. A.; Cullen, D. *Applied Environ. Microbiol.* 1998, *64*, 1924-1928.
38. Atalla, R. H.; Hackney, J. M.; Uhlin, I.; Thompson, N. S. *Int. J. Biol. Macromol.* 1993, *15*, 109-112.
39. Atalla, R. H. In *Comprehensive Natural Products Chemistry;* Pinto, B. M.; Ed.; Pergamon Press: Amsterdam, Netherlands, 1999; Vol. 3, pp 529-598.
40. Reese, E. T.; Siu, R. G. H.; Levinson, H. S. *J. Bacteriol.* 1950, *59*, 485-497.
41. Din, N.; Gilkes, N. R.; Miller Jr., R. C.; Warren, R. A. J.; Kilburn, D. G. *Proc. Nat. Acad. Sci.* 1994, *91*, 1383-11387.
42. Din, N.; Gilkes, N. R.; Tekant, B.; Miller, R. C. J.; Warren, R. A. J.; Kilburn, D. G. *Bio/Technology.* 1991, *9*, 1096-1099.
43. Gill, J.; Rixon, J. E.; Bolam, D. N.; McQueen-Mason, S.; Simpson, P. J.; Williamson, M. P.; Hazlewood, G. P.; Gilbert, H. J. *Biochem. J.* 1999, *342*, 473-480.
44. Swanson, B. A.; Ward, M.; Pentilla, M.; Saloheimo, M. International Patent WO 99/02693, 1999.
45. Banka, R. R.; Mishra, S.; Ghose, T. K. *World J. Microbiol. and Biotechnol.* 1998, *14*, 551-558.

Chapter 7

Molecular Mechanics Studies of Cellulases

Rocio Palma[1], Pierfrancesco Zuccato[1], Michael E. Himmel[2],
Guyan Liang[1,3], and John W. Brady[1]

[1]Department of Food Science, Stocking Hall, Cornell University,
Ithaca, NY 14853
[2]National Renewable Energy Laboratory, 1617 Cole Boulevard,
Golden, CO 80401
[3]Computer Aided Drug Design, Rhone-Poulenc Rorer, 500 Arcola Road,
SW8, Collegeville, PA 19426–0995

Molecular Mechanics (MM) simulations are ideally suited for studying the properties of aqueous solutions of biological molecules, which are difficult to probe experimentally. To date, there have been relatively few MM calculations of cellulases, although there are a number of questions which could perhaps be answered with modeling studies. Several examples of MM calculations of cellulase systems are given to illustrate the potential applications in understanding and modifying cellulase activity. These include the calculation of the change in substrate binding affinity for the Tyr240Phe point mutant of the E1 cellulase from *Acidothermus cellulolyticus*, and the calculation of the potential of mean force for the binding of a glucose molecule to methane, as a model for the sidechain of alanine.

Great strides have been achieved in recent years in our understanding of the structure and function of various cellulases, but many questions remain unanswered, particularly regarding how cellulases might be engineered for greater efficiency in industrial applications. In a number of these cases, basic experimental limitations inhibit progress, or the difficulty of the experiments makes them inefficient or impractical. However, in many instances, Molecular Mechanics (MM) computer simulations (*1, 2*) could potentially be used to model these systems, either as an efficient guide for planning experiments, or to provide information which cannot be currently obtained experimentally. These simulations allow the direct examination of the detailed atomic interactions and processes, which give rise to macroscopic thermodynamic properties. Unfortunately, while the MM simulation of proteins is now a well-developed field, there have been relatively few modeling studies of cellulase systems. Simulations have perhaps been inhibited by the large size of many of the proteins involved, the relative absence, until recent years, of good crystallographic

determinations of the protein structures, poor knowledge of the potential energy functions for carbohydrates, and the possible role of solvent, which further increases the cost of such simulations. The stage is now set for the application of molecular modeling techniques to the study of cellulases. We will cite a few simple examples here of how modeling calculations can be of help in understanding cellulase structure and function.

Molecular Mechanics Calculations

Molecular Mechanics calculations use empirical energy functions to describe the way in which the energy of atomic and molecular systems changes with atomic coordinates, and predict observable properties based on this knowledge of energetics. There are a number of types of MM calculations. The most familiar and common types are conformational energy and substrate docking calculations based on numerical algorithms which attempt to minimize the energy as a function of conformation. Molecular Dynamics (MD) simulations model the atomic motions that arise from the forces acting on the atoms due to the empirical energy function, and Monte Carlo (MC) calculations use the energy function to construct ensembles of configurations from which average properties can be calculated. Other types of MM studies include normal mode calculations to predict vibrational spectra and reaction coordinate analysis. More recently, MM calculations have been coupled with quantum mechanical calculations (QM/MM simulations) to provide theoretical models for chemical reactions.

The fundamental requirement of molecular mechanics studies of any type is a complete description of the variation of the total potential energy of the system as a function of the molecular coordinates (*3*). For macromolecules and condensed phases the accurate calculation of this quantum mechanical energy is not possible, and it is thus common to employ analytic, semi-empirical energy functions which have theoretically reasonable functional forms and which have been parameterized to the results of experiment and simple calculations. These semi-empirical potential energy functions vary in their details and, to some extent, even in their functional forms from one set to the next; most, however, represent the intramolecular potential energy as a sum of electrostatic and van der Waals interactions between non-bonded atoms and terms for hindered rotation about molecular bonds, with bond stretching and bond bending forces derived from quadratic restoring potentials,

$$V(\mathbf{q}) = \Sigma k_b(b-b_0)^2 + \Sigma k_\theta(\theta - \theta_0)^2 + \Sigma k_\phi[1+\cos(n\phi- \delta)] +$$

$$\Sigma A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6} +q_iq_j/r_{ij} \tag{1}$$

Special functions are sometimes included to account for hydrogen bonding. More sophisticated potential functions account for anharmonicity and the coupling of various internal deformations, as in a Urey-Bradley type force field (*4*). The

parameters in these potential energy surfaces, such as atomic partial charges, bond lengths and angles, and force constants, are obtained from simple quantum mechanical calculations and experimental data, such as crystal structures and vibrational spectra. Molecular mechanics potential energy functions have been developed to describe a variety of systems, such as various small molecules (including the important case of water (5, 6)), simple organic compounds, and biopolymers (3, 7, 8). The simulations described here employed the CHARMM23 force field for the protein atoms (9) and a newly-developed CHARMM-type force field for carbohydrates which is described in detail in the Appendix.

MD calculations (1, 2) involve the numerical integration of Newton's equations of motion for all atoms in a system as they move in response to the forces acting on them. These forces are computed directly from the derivatives of the empirical potential energy function. In such calculations, physical observables (properties) are calculated as time averages over the various states that arise during the course of the simulation, since in principle the frequency of occurrence of each state and the time spent in each state will, for simulations of sufficient length, converge to the value determined by the Boltzmann distribution. Because atomic motions are simulated directly, entropic effects that do not appear in conformational energy minimizations are implicitly included in MD simulations, and the complete histories of the time sequence of events produced by MD calculations can provide rate information that cannot be obtained from Monte Carlo calculations. For a sufficiently realistic potential energy function, such simulations provide information of unparalleled detail about the microscopic behavior of a system, which often cannot be obtained by any currently available experimental methods.

In MD simulations, complete phase space points are saved at frequent intervals, and the properties of interest are calculated by averaging over the post-equilibration phase of the trajectory, since the ergodic hypothesis allows the calculation of thermodynamic properties from time-averages of sufficient length. Many interesting properties of aqueous systems converge slowly on the MD timescale, and care must be taken to ensure, if possible, that simulations have proceeded far enough such that the averaging has converged to the thermodynamic limit. The long simulation times needed to ensure the statistical validity of calculated properties for systems containing slowly reorienting water molecules and large solutes undergoing slow conformational change is one of the primary limitations on MD calculations. This problem is particularly troublesome in modeling large protein systems, and cellulases interacting with crystalline substrates are at the large end of the size spectrum for which modeling is currently practical.

Conventional Molecular Mechanics simulations are usually limited to non-chemical events since the force fields used cannot easily represent the major redistributions of electronic structure which accompany the breaking or formation of chemical bonds. This is a major handicap in using such calculations to study enzymatic mechanisms. These inherently quantum mechanical processes require a solution of the Schrödinger equation for the system during the course of the reaction, which is clearly impossible for enzymatic systems. However, exciting progress has been made in recent years in the development of hybrid quantum

mechanical/molecular mechanical (QM/MM) simulation techniques which treat the small group of atoms directly involved in the chemical changes quantum mechanically, coupled to a molecular mechanical simulation of the remainder of the system as an influencing environment (*10, 11*). Such QM/MM calculations can be used to calculate the energetics along proposed alternate reaction paths and allow the most energetically favorable to be identified as the most probable mechanism. These simulation techniques have been successfully applied to a number of enzymatic reactions in recent years (*12-18*). While no such calculations have yet been reported for cellulases, the possibility of using QM/MM calculations to explore cellulose mechanistic details is exciting.

Simulations of the E2 Cellulase from *Thermomonospora fusca*

The most common types of MM calculations are graphics-based docking, energy minimizations, and related methods to study protein conformations and the binding of ligands to proteins. We have conducted a series of such studies of the Cel6A cellulase from *Thermomonospora fusca* (previously referred to as E2) and its substrates to explore how oligosaccharides bind in the active-site cleft of the protein, which residues are involved, and what the structure of the complex suggests about the enzymatic mechanism. Several different types of calculations were used, including conformational energy minimizations, adiabatic mapping for the substrate disaccharide linkages, and molecular dynamics simulations. The crystal structure of E2 has been determined to high resolution (*19*). Unfortunately, in these experiments, as is often the case with enzymes, it was not possible to co-crystallize the protein with its substrate, although it was possible to produce a co-crystal with cellobiose, an inhibitor of the enzyme, in the active site.

In order to provide a picture of how the full substrate fits into the binding region of E2, Molecular Mechanics calculations were used to model a cellotetraose molecule into the binding cleft of the protein (*20*) using our previously reported carbohydrate force field (*21*) and a newly developed, improved parameter set (see Appendix). Cellulose is believed to have a relatively flat, ribbon-like structure. Because of the inevitable curvature of any extended binding cleft on the surface of a small globular protein, the portion of the substrate bound to the active site must bend away from this flat structure to some extent. Using the positions of the disaccharide rings to position the first two residues of a tetrasaccharide in the binding cleft, we used energy minimization to place the remaining two rings. In the crystal structure of the complex with the disaccharide, the observed glycosidic linkage angles tilt the second ring at an unusual angle relative to the first. In cellulose, the successive sugar rings are almost parallel to one another, producing the flat ribbon-like structure. The reason for the tilt of this linkage angle in the complex is the topology of the binding pocket occupied by this dimer in the active site cleft. The binding cleft has been identified as consisting of at least four binding sites, generally designated as either A-D or -2 to +2, with each site occupied by a single sugar residue. In the model complex of the tetrasaccharide substrate docked in the binding cleft, the "A" residue sits comfortably in a deep recess

or pocket at the beginning of the cleft, while the protein surface bulges out at the position of the "B" residue, forcing it to tilt in order to fit properly. This bulge is caused by the presence of the Lys 259 residue, which is hydrogen bonded to both the first and second sugar residues. However, another deep recess exists in the "C" position, allowing the third residue of the oligomer to twist back in the opposite direction. This ligand alignment is very similar to that predicted for the homologous endocellulase CBH II from *Trichoderma reesei* (*22*), and the placement of the substrate shares several features with other carbohydrate-binding proteins (*23-28*). In particular, the binding groove is lined with several hydrophobic residues, including three Trp residues, which complement the hydrophobic "tops" and "bottoms" of the glucose rings of the polymer which result from having all of the hydroxyl groups equatorial. These groups not only provide a contribution to the thermodynamic driving force for the molecule to leave solution, where these hydrophobic faces, consisting of axial aliphatic protons, are removed from contact with water, but also presumably help to position the substrate properly in the catalytic site.

The modeling of the substrate into the binding site was followed by MD simulations of the complex (*20*), undertaken in part to allow the protein and substrate to further relax and possibly find lower-energy conformations by thermally-induced crossings of energy barriers, and in part to determine the types and extent of the conformational motions of the protein in the complex and which residues these motions brought into the vicinity of the scissile bond. In the MD simulations of the protein-ligand complex, the conformation of the oligosaccharide was stable over the time-period simulated, while somewhat larger fluctuations and shifts were observed in the conformations of some of the protein groups.
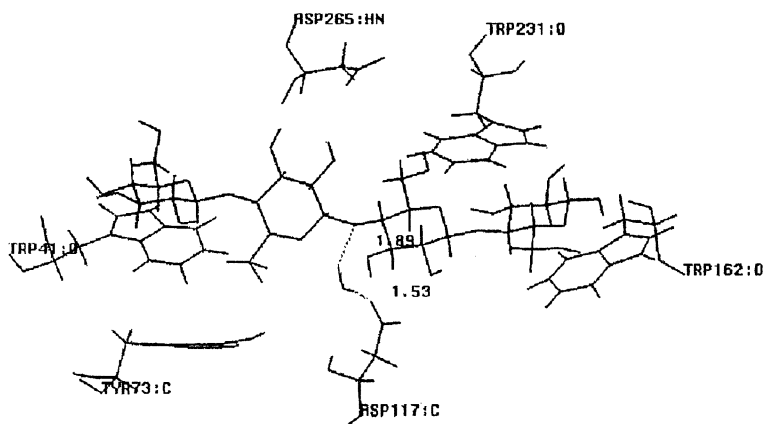
Figure 1. "Snapshot" from an MD simulation of the binding site of E2 modeled with a tetrasaccharide substrate, illustrating how a water molecule can bridge between the catalytic acid residue Asp 117 and the scissile bond, in a way that would facilitate the transfer of a proton between the Asp and the linkage oxygen. Note the position of Tyr 73 and the three Trp residues (41, 231, and 162).

One of the unresolved questions concerning the functioning of this enzyme is that while the active site binding cleft contains at least two Asp residues which could serve in the roles of the general acid and base in the glycosidic hydrolysis, neither of these groups is closer than 5 Å to the scissile bond in the crystal structure with the tetrasaccharide modeled into the cleft. It is assumed that water molecules could serve as the intermediaries in this reaction, by accepting a proton from the acid and transferring it to the glycosidic oxygen, but such water molecules were not observed in the crystal structure, which of course did not contain the substrate either. However, the MD simulations can explicitly include water molecules, and when such simulations were performed, it was found that water molecules did indeed simultaneously hydrogen bond to both the putative catalytic acid residue Asp 117 and to the glycosidic oxygen atom of the scissile bond, as shown in Figure 1, which illustrates a snapshot from an MD simulation, displaying only the relevant water molecule, the substrate, and a few of the more important protein residues in the binding groove. The general conformation of these groups and the approximate positioning of a bridging water molecule did not depend upon whether Asp 117 was protonated of not in the simulation.

An interesting feature of the predicted structure of the enzyme-substrate complex is the Tyr 73 residue located just "below" the cleavage site (see Figure 1). This is a conserved residue across a series of homologous cellulases (29, 30), and the corresponding residue in CBH II (22), for which the crystal structure is known and for which MM calculations have been used to position the substrate, has the same unusual tilted conformation seen in E2. Even more interestingly, during the course of MD simulations of the complex, the proton of the hydroxyl group of this residue came into direct van der Waals contact with the oxygen atom of the scissile glycosidic bond. Based on this observation, a mutant of the protein in which this residue was replaced by phenylalanine was prepared and characterized by D.B. Wilson and co-workers. This mutant was found to have only 10% of the activity of the wild-type protein. However, if this residue played a catalytic role, the expected reduction in activity should have been much greater, suggesting that perhaps the role might be more indirect, or might involve binding affinity instead. The function of this residue requires more study, since in the complexes generated thus far, this side chain does have significant binding interactions with the ligand, but is not positioned optimally for its hydrophobic aromatic ring to be serving as a complementary surface for the hydrophobic undersurface of the second sugar ring of the oligomer.

Additional modeling was performed for the cleavage product bound to the protein (20), which is expected to consist, for this cellotetraose substrate, of two cellobiose molecules bound to the protein, one in the "A" and "B" sites, and one in the "C" and "D" sites. The surface topology of the binding cleft in the crystal structure does not exactly match the spacing of the four sugar rings in the cellulose structure, and relaxing the requirement of positioning all four constrained by chemical connectivity produces shifts in the position of the second product cellobiose in sites C and D which has several possible positions. It is possible that this mismatch in reactant and binding pocket helps in small part to drive the reaction to products.

Calculation of the Potential of Mean Force for Sugar Binding

Another use for MM calculations in the study of cellulases would be to determine the effect of specific mutations on the binding affinity for cellulosic substrates. As a first approximation, it might be useful to determine the increment in binding energy that results from a sugar interacting with a single side chain type, regardless of its local environment in the active site. One approach to this problem could be to use MD simulations to calculate the potential of mean force for the binding of a glucose molecule to various amino acid sidechains in aqueous solution. For example, one could attempt to compute the free energy change for a glucose molecule approaching a benzene ring, as a model for the binding to the sidechain of phenylalanine in a protein, or approaching a methyl group, as a model for the sidechain of alanine. The difference between these two would then be a measure of the difference in affinity of these two sidechains for the sugar molecule. In practice such free energy functions are difficult to calculate, but special techniques have been developed to overcome these problems. For example one could use umbrella sampling free energy calculations to constrain a glucose and a benzene ring at various distances so as to collect statistically meaningful probability density data for the construction of the free energy as a function of **r**. At each separation distance, very long MD simulations are necessary to average over all of the possible water arrangements consistent with the given separation distance.

The potential of mean force $W(\xi)$ as a function of some internal reaction coordinate ξ is given by the equation

$$W(\xi) = -kT \ln P(\xi) \qquad (2)$$

where $P(\xi)$ is the probability density for ξ. This potential of mean force is the work function of the system, or the change in free energy between two different values of ξ, and in an aqueous solution includes the average contribution of all of the solvent water molecules. Ordinary molecular dynamics procedures will not adequately sample the high energy barrier regions between the stable conformations to allow the calculation of $W(\xi)$. This problem can be overcome by augmenting the true intramolecular potential $V(\xi)$ for the system with an "umbrella potential" $U(\xi)$ which makes transition conformations more probable (*31, 32*). Averages of some property A for the system's true potential can be obtained from the dynamics calculations with this augmented potential through the equation

$$\langle A \rangle = \langle A/\exp(-\beta U(\xi)) \rangle \, \langle 1/\exp(-\beta U(\xi)) \rangle^{-1} \qquad (3)$$

where the averaging is over the augmented system ensemble and β is $1/kT$. $\langle A \rangle$ could be the probability density $P(\xi)$ needed to calculate the work function, for example. In the case of glucose binding to a sidechain, ξ is the separation distance between the two molecules, **r**, and a series of calculations for different distances can be performed with

overlapping quadratic localizing potentials centered on each successive value of **r**. It should be noted that while these dynamics trajectories can be used to evaluate the potential of mean force, the specific motions of each individual trajectory in such a simulation will not be directly physically meaningful, due to the dynamical effects of the artificial potential.

We have completed such calculations for benzene, as a model for phenylalanine, and for methane, as a model for alanine, allowing an evaluation of the solution contribution to the binding free energy change for mutating a phenylalanine to alanine in a protein. The benzene calculation has been previously reported (*33*), and the results for the methane calculation are shown below. At each separation distance, very long MD simulations were used to average over all of the possible water arrangements consistent with the given separation distance; the vacuum and solution calculations for the methane case are shown in Figures 2 and 3.
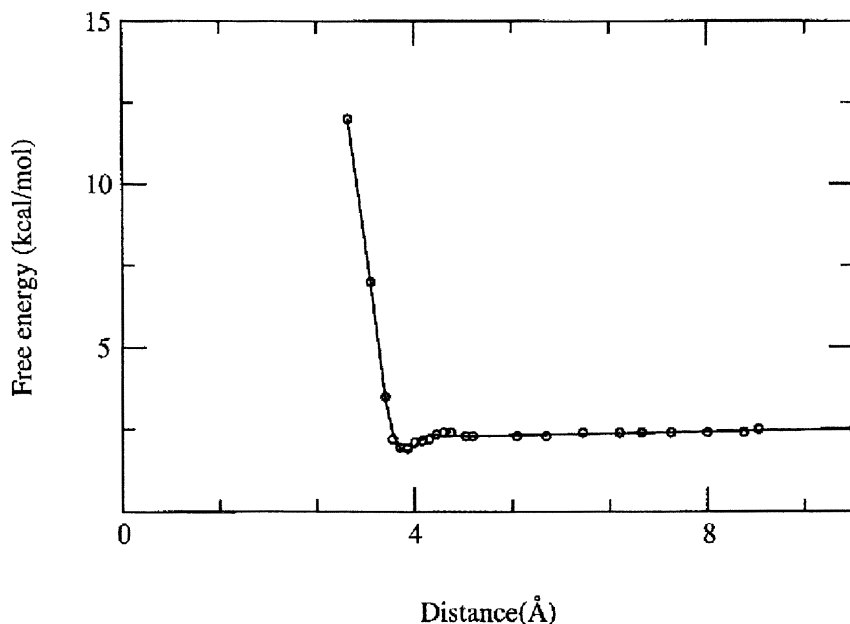
*Figure 2. The potential of mean force for a single molecule of methane interacting with a single molecule of α-D-glucopyranose in vacuum. The free energy for bringing the two molecules from infinity to a distance **r** is plotted as a function of **r**.*

The curves in Figs. 2 and 3 show the change in free energy for bringing a sugar molecule to a given separation distance **r** from the side chain functional group. The barrier splitting the primary contact minimum of the solution case in Figure 3 from the secondary, solvent-separated minimum around 8 Å can be thought of as the energy

needed to "squeeze" out the last solvation layer between the two molecules as they approach to contact. By taking the difference of these two curves, the change for the mutation can be estimated. In both cases the net binding affinity of the two molecules for glucose in water is calculated to be about 1 kcal/mol (*33*). Perhaps surprisingly, the difference in the lowest energy complex for these two curves was quite small, indicating that this mutation might have only a small impact in a protein binding site. Apparently the alanyl-like functional group is sufficiently large to provide a hydrophobic surface to match against the hydrophobic undersurface of the sugar, without precluding water molecules from approaching the peripheral hydroxyl groups to make hydrogen bonds.
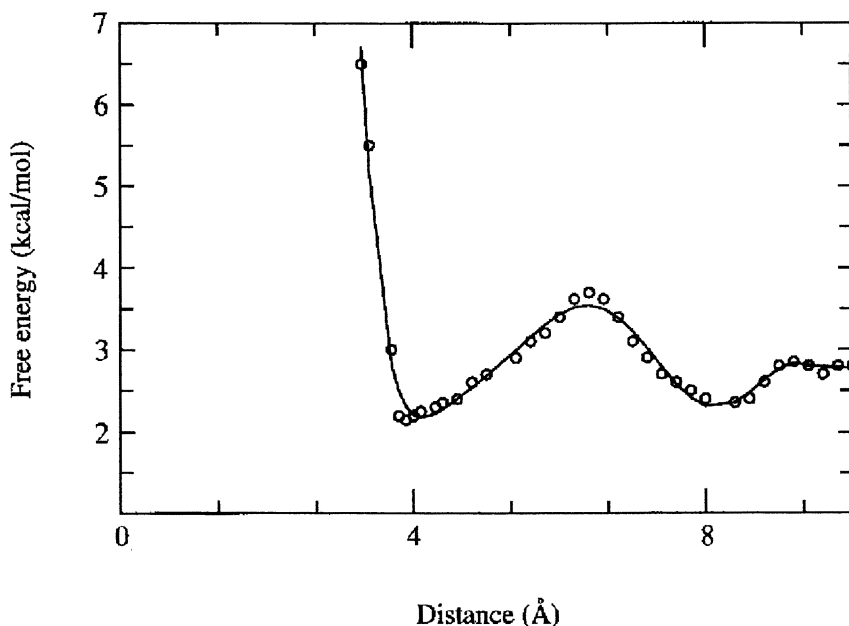
*Figure 3. The potential of mean force for a single molecule of methane interacting with a single molecule of α-D-glucopyranose in aqueous solution. The free energy for bringing the two molecules from infinity to a distance **r** is plotted as a function of **r**, where the methane is constrained to remain centered over the sugar ring.*

Point Mutations in Cellulases

A common problem in bioengineering is to find point mutants of proteins in which the mutation alters the functionality, as for example changing the denaturation or "melting" temperature, or as already mentioned, the substrate binding affinity. A

major focus of cellulase research is to find variants of the enzyme with higher activity for industrial purposes. This objective includes many types of effects, such as changing the efficiency of the actual catalysis due to the placement of the catalytic residues, changing the ease of disruption of the crystalline substrate, or altering the binding affinity of the protein for the substrate or for the products. The changes in substrate binding affinity which result from point mutations can be calculated from MD or MC simulations employing a procedure sometimes referred to as "computational alchemy" (*1, 2, 34, 35, 36*). In this type of calculation, the differences in binding energies for the different substrates or protein mutants are calculated based upon a hypothetical "mutation" of one sugar into another or one amino acid side chain into another. In these free energy simulations, a coupling parameter approach is used to transform one sugar or amino acid into another through a series of calculations for different values of the parameter λ, using an equation of the form

$$E_\lambda = \lambda E_1 + (1-\lambda)E_0 \tag{4}$$

In these simulations, both sugar molecules, or both the wild-type and mutant side chains, are simultaneously present, with the degree to which each contributes to the total energy determined by λ, which transforms one molecule into the other as λ varies from 0 to 1. While these simulations do not correspond to any physical system, they define a path between two physical end points as a series of steps for which the free energy difference can be calculated from a perturbation equation of the form

$$G_1 - G_0 = -kT \ln <\exp[-\beta(E_1-E_0)]>_0 \tag{5}$$

Since free energy is a state function, the total difference between the two physical end points can be calculated from the total difference for the individual steps

$$\Delta G = \sum_{i=1}^{N} \Delta G_i \tag{6}$$

The relative binding free energy difference $\Delta\Delta G_b$ for two ligands to the same binding site in a protein, or the relative binding free energy difference for a single ligand to a wild-type protein and to a mutant of that protein can be estimated from such free energy perturbation calculations by employing an appropriate thermodynamic cycle. The calculation of the difference in binding affinities of the wild-type protein P_1 and the mutant protein P_2 for a ligand must take into account the fact that the two proteins do not have the same free energy in aqueous solution, as well as the difficulty of directly estimating the free energy change associated with the substantial, non-perturbative changes involved with taking a ligand out of solution and putting it in the binding site, desolvating both in the process. The total difference in

binding affinities of the two proteins, $\Delta\Delta G_b$, can be calculated from a thermodynamic cycle (*37, 38*) of the type

$$
\begin{array}{ccc}
& \Delta G_1 & \\
P_1 + L & \rightarrow & P_1L \\
\downarrow \Delta G_3 & & \downarrow \Delta G_4 \\
P_2 + L & \rightarrow & P_2L \\
& \Delta G_2 &
\end{array}
\qquad (7)
$$
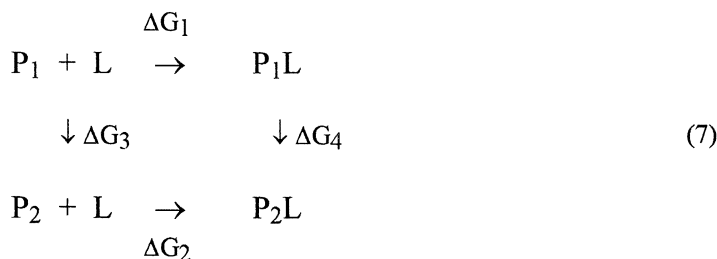
In this scheme, ΔG_1 and ΔG_2 are the free energy differences for the binding of the sugar ligand to the wild-type and mutant proteins, respectively, and ΔG_3 and ΔG_4 are the free energy changes for the non-physical process of transforming the wild-type protein into the mutant in solution and in the ligand complex, respectively. Since free energy is a state function, the difference in the binding energies of the two proteins can be written as

$$
\Delta\Delta G_b = \Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3 \qquad (8)
$$

While it is difficult to obtain the quantities ΔG_1 and ΔG_2 from MD calculations (see previous section), estimation of the free energy difference associated with the change in protein structure in going from one amino acid side-chain to another, both in solution and in the protein binding site, is tractable using the perturbative mutation pathway described above, thus allowing the calculation of $\Delta\Delta G_b$.

As an example of this method, we have used free energy perturbation calculations to estimate the difference in binding energy for a tetrasaccharide substrate in the binding site of the E1 cellulase from *Acidothermus cellulolyticus* (*39*) between the wild-type and a point mutant. In this calculation the Tyr 240 residue was mutated to Phe. In the crystal structure this residue is hydrogen bonded to the substrate, and thus changing it to a nonpolar Phe would be expected to have a significant effect on the binding affinity. Figure 4 shows the results for this Tyr240Phe mutation. The actual binding energy difference must take account of the role of water in the binding, as the substrate and protein dehydrate partially upon binding. Accordingly, one must calculate this binding energy using a thermodynamic cycle, and estimate the difference in interaction energy with water between the mutant and wild type proteins, in order to subtract from the number computed for the substrate complex. The curve indicated by open squares is the change for the protein in aqueous solution. From these calculations it would seem that the Phe mutant binds the tetrasaccharide less tightly by about 2.1 kcal/mol (this number would be consistent with the loss of a hydrogen bond). This calculation used a starting structure for the tetrasaccharide in which the sugar ring at

the −1 position had distorted to a twist-boat conformation. As the simulation progressed and the tyrosine residue was gradually transformed into a phenylalanine, this distortion in the sugar ring gradually relaxed until it was in an (albeit flattened) chair for the mutant case. Unfortunately, this mutant has not yet been produced in the laboratory for an experimental determination of this binding affinity.
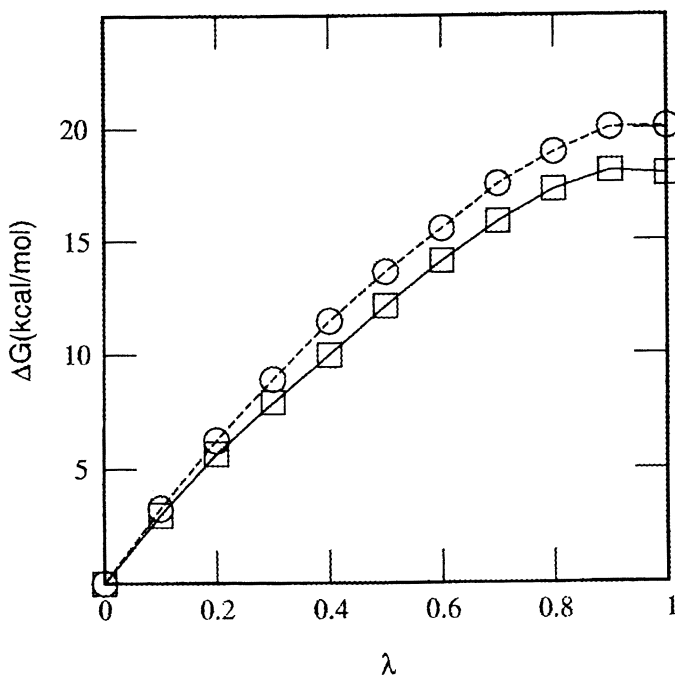
Figure 4. The free energy change in the computer "mutation" of Tyr 240 into Phe in E1 as a function of the scaling variable λ, used in the mixed energy function $E_\lambda = \lambda E_m + (1-\lambda)E_{wt}$. The function indicated by squares is for the mutation in the solvated protein without substrate, while that indicated by circles is for the mutation with a tetrasaccharide substrate. In the simulation with substrate the sugar in the −1 position is distorted into a half–chair ring conformation for small λ values, but this stress gradually relaxed as the residue mutated to a phenylalanine.

Conclusions

Molecular Mechanics simulations of cellulase systems have until recently been relatively rare, although they are now very commonly applied to other biophysical systems with great success. Cellulases are moderately large proteins, which makes simulating them expensive, and the insolubility of their cellulose substrates makes modeling the protein-substrate interactions is even more difficult. In deciding whether

124

a particular question is a good candidate for numerical simulation, the basic limitations on size and timescale in these simulations must be carefully considered. However, many aspects of cellulase systems that are presently poorly understood could benefit from computational modeling. These include using free energy simulations to predict the effects of point mutations before they are produced in the laboratory, determining mechanistic details through modeling enzyme-substrate interactions, simulating the role of the linker segment in cellulases with separate catalytic and binding domains, and determining the role of water in the binding of cellulases to microcrystalline substrate. Properly applied, such simulations could be a useful guide to experimental studies of these systems, and help in the practical goal of producing cellulases with increased activity for industrial purposes.

Acknowledgments

The authors thank D.B. Wilson, J.S. Taylor, and B. Teo for many helpful discussions. This work was supported by subcontract XCO-8-17101-01 from the National Renewable Energy Laboratory.

Appendix

The results of Molecular Mechanics simulations in general can only represent physical systems accurately if the force fields used to model them are sufficiently realistic as to include the important features of the system. A large number of force fields have been developed for modeling proteins, including the widely used CHARMM (*40*), AMBER (*41*), and GROMOS (*42*) energy functions, and these have been extensively validated against know experimental data. There are fewer force fields available for carbohydrates, but in recent years several have been proposed, and these have had very good success in modeling carbohydrate properties. However, these first-generation carbohydrate force fields also suffered from several known shortcomings. The most important of these is that they tend to give inaccurate rotameric distributions for the exocyclic primary alcohols of pyranoid hexoses, due to an inadequate treatment of gauche effects in -O-C-C-O- bond topologies such as occur in sugars or ethylene glycol.

We have developed a revision of our previously quite successful CHARMM-type force field for carbohydrates (*21*) which corrects the -O-C-C-O- terms and also revises the barrier for hydroxyl rotations, which may have been close to 10% too low, based on experimental and *ab initio* estimates of this barrier in methanol and on the rotation rates observed in MD simulations of sugar crystals (*43*), known from neutron diffraction to have little disorder in their hydroxyl proton positions. *Ab initio* 6-31G* calculations of the rotational energy profile for the C-C bond in ethylene glycol, with electron correlation treated at the MP2 level, and with the two hydroxyl groups kept in the *trans* conformation, allowed the evaluation of the "gauche effect" contribution to this rotation and by extension to that of the exocyclic primary alcohol in sugars, without the complication of intramolecular hydrogen bonding (*44*). Such calculations

In Glycosyl Hydrolases for Biomass Conversion; Himmel, M., et al.;
ACS Symposium Series; American Chemical Society: Washington, DC, 2000.

found a substantial stabilization of gauche conformations, as well as higher barriers to rotation than in previous MM energy functions.

This revised force field is intended for use with the CHARMM23 parameters for proteins, and uses the new conventions adopted in that version of the general CHARMM force field (*9*). In particular, in this new model, all possible torsion angles for each chemical bond appear in the energy function, rather that just one per bond as in earlier CHARMM functions. This change allows the intrinsic energy term for each bond to be asymmetric to account for the different substituents which may be bound to the central atoms. In this new force field, each intrinsic torsional term in Equation (1) was treated as a three-term expansion consisting of terms with periodicities n = 1, 2, and 3, designated as V1, V2, and V3 in Table III. For all three terms, the phase factor δ was taken to be 0. Atomic partial charges and van der Waals parameters were adapted from those given in the CHARMM23 parameter set for similar functional groups. The partial charges used for α-D-glucopyranose are listed in Table V; the same anomeric charges were used for both the α and β anomers of the molecule.

The new force field was developed as before through a non-linear least-squares optimization of known vibrational frequencies and structural properties (*45*) in both the sugars and in a number of small-molecule homologs of sugar functions groups, subject to constraints on properties known from other sources such as van der Waals parameters, atomic partial charges, and rotational barriers. The overall effect of these revisions is to produce a sugar molecule which is "stiffer" than with our previous force field. In particular, hydroxyl groups spend much longer times between rotations, and rotations of the exocyclic hydroxymethyl become extremely unlikely in simulation times of less than 1 ns. These changes necessitate much longer simulation times to achieve thermodynamic convergence, since it is known that the intramolecular energy can change significantly with hydroxyl rotations (*46*), and a representative ensemble of these different states must be sampled in order to produce meaningful results.

Table I. Bond Stretching Parameters for Carbohydrates

TYPE	k_b (kcal/mol/Å^2)	b_0 (Å)
HOS-OHS	546.4212	0.9595
HAS-CTS	335.6034	1.1000
HAS-CBS	335.6034	1.1052
CTS-OHS	384.0792	1.4066
CBS-OHS	384.0792	1.3932
CTS-CTS	325.5297	1.5066
CBS-CTS	325.5297	1.5074
CTS-OES	385.3133	1.4165
CBS-OES	385.3133	1.4202

Table II. Angle Bending Parameters for Carbohydrates

TYPE	k_θ (kcal/mol/rad^2)	θ_0 (degrees)
HAS-CTS-CTS	42.9062	109.7502
HAS-CBS-CTS	42.9062	109.7502
HAS-CTS-CBS	42.9062	109.7502
OHS-CTS-CTS	112.2085	107.6019
OHS-CBS-CTS	112.2085	107.6019
OHS-CTS-CBS	112.2085	107.6019
HOS-OHS-CTS	57.5478	109.1722
HOS-OHS-CBS	57.5478	109.1722
HAS-CTS-HAS	36.8220	106.1784
HAS-CTS-OHS	52.5070	109.3850
HAS-CBS-OHS	52.5070	109.3850
HT-OT-HT	55.0000	104.5200
HAS-CTS-OES	62.2500	105.4025
HAS-CBS-OES	62.2500	105.4025
CTS-CTS-CTS	167.3535	110.6156
CBS-CTS-CTS	167.3535	110.6156
CTS-CTS-OES	169.0275	108.3759
CBS-CTS-OES	169.0275	108.3759
CTS-CBS-OES	169.0257	108.3759
CTS-OES-CTS	92.5892	111.5092
CBS-OES-CTS	92.5892	111.5092
OES-CTS-OHS	74.2586	115.7322
OES-CBS-OHS	74.2586	110.3385
OES-CTS-OES	37.4370	112.1882
OES-CBS-OES	37.4370	106.9160

Table III. Torsional Force Constants k_ϕ (kcal/mol) for
Carbohydrates. V1, V2, and V3 refer to periodicities n of 1,2, and 3.
The energy for each bond is the sum of all three.

TYPE	V1	V2	V3
CTS-CTS-CTS-CTS	-1.0683	-0.5605	0.1955
CBS-CTS-CTS-CTS	-1.0683	-0.5605	0.1955
CTS-CTS-CTS-OES	-1.2007	-0.3145	-0.0618
CBS-CTS-CTS-OES	-1.2007	-0.3145	-0.0618
CTS-CTS-CBS-OES	-1.2007	-0.3145	-0.0618
CTS-CTS-CTS-OHS	-1.9139	-0.3739	-0.0340
CTS-CTS-CBS-OHS	-1.9139	-0.3739	-0.0340
CBS-CTS-CTS-OHS	-1.9139	-0.3739	-0.0340
CTS-OES-CTS-CTS	-0.8477	-0.3018	0.3763
CBS-OES-CTS-CTS	-0.8477	-0.3018	0.3763
CTS-OES-CBS-CTS	-0.8477	-0.3018	0.3763
HAS-CTS-CTS-CTS	0.0000	0.0000	0.1441
HAS-CBS-CTS-CTS	0.0000	0.0000	0.1441
HAS-CTS-CTS-CBS	0.0000	0.0000	0.1441
HAS-CTS-CTS-HAS	0.0000	0.0000	0.1595
HAS-CBS-CTS-HAS	0.0000	0.0000	0.1595
HAS-CTS-OES-CTS	0.0000	0.0000	0.2840
HAS-CBS-OES-CTS	0.0000	0.0000	0.2840
HAS-CTS-OES-CBS	0.0000	0.0000	0.2840
HOS-OHS-CTS-CTS	1.0504	0.1336	0.3274
HOS-OHS-CBS-CTS	1.0504	0.1336	0.3274
HOS-OHS-CTS-CBS	1.0504	0.1336	0.3274
HOS-OHS-CBS-HAS	0.0000	0.0000	0.1694
HOS-OHS-CTS-HAS	0.0000	0.0000	0.1694
OES-CTS-CTS-HAS	0.0000	0.0000	0.1686
OES-CBS-CTS-HAS	0.0000	0.0000	0.1686
OES-CTS-CBS-HAS	0.0000	0.0000	0.1686
OES-CTS-CTS-OES	-2.6785	0.7851	0.2552
OES-CBS-CTS-OES	-2.6785	0.7851	0.2552
OES-CTS-CTS-OHS	-3.7993	0.5688	0.4204
OES-CBS-CTS-OHS	-3.7993	0.5688	0.4204
OES-CTS-CBS-OHS	-3.7993	0.5688	0.4204
OES-CTS-OES-CTS	0.1948	0.9778	0.9115
OES-CBS-OES-CTS	0.1948	0.9778	0.9115
OHS-CTS-OES-CTS	1.9193	1.0102	0.7294
OHS-CBS-OES-CTS	1.9193	1.0102	0.7294
OES-CTS-OHS-HOS	1.2936	1.3295	0.4323
OES-CBS-OHS-HOS	1.2936	1.3295	0.4323
OHS-CTS-CTS-HAS	0.0000	0.0000	0.1472
OHS-CBS-CTS-HAS	0.0000	0.0000	0.1472
OHS-CTS-CBS-CTS	0.0000	0.0000	0.1472
OHS-CTS-CTS-OHS	-4.9362	0.2907	0.4638
OHS-CBS-CTS-OHS	-4.9362	0.2907	0.4638

Table IV. Lennard-Jones Parameters for Carbohydrates

TYPE	ε (kcal/mol)	σ (Å)
HOS	-0.0460	0.2245
HAS	-0.0220	1.3200
CTS*	-0.0200	2.2750
CBS*	-0.0200	2.2750
OHS	-0.1521	1.7700
OES	-0.1521	1.7700

* For 1-4 interactions involving these atoms, different values of σ and ε are used; in these cases, ε = -0.0100 and σ = 1.9000.

Table V. Atomic Partial Charges for Glucose

ATOM	TYPE	CHARGE
C1	CTS	0.200
H1	HAS	0.090
O1	OHS	-0.660
HO1	HOS	0.430
C5	CTS	0.250
H5	HAS	0.090
O5	OES	-0.400
C2	CTS	0.140
H2	HAS	0.090
O2	OHS	-0.660
HO2	HOS	0.430
C3	CTS	0.140
H3	HAS	0.090
O3	OHS	-0.660
HO3	HOS	0.430
C4	CTS	0.140
H4	HAS	0.090
O4	OHS	-0.660
HO4	HOS	0.430
C6	CTS	0.050
H61	HAS	0.090
O6	OHS	-0.660
HO6	HOS	0.430

References

1. McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, MA, 1987.
2. Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Prigogine, I,; Rice, S.A., Eds.; Advances in Chemical Physics Vol. LXXI; Wiley-Interscience: New York, NY, 1988.
3. Burkert, U.; Allinger, N. L. *Molecular Mechanics*; ACS Monograph Vol. 177; American Chemical Society: Washington, DC, 1982.
4. Urey, H. C.; Bradley, C.A. *Phys. Rev.* **1931**, *38*, 1969-1978.
5. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. in *Intermolecular Forces*; Pullman, B., Ed., Reidel Publishing: Dordrecht, Holland, 1981; pp. 331-342.
6. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.
7. van Gunsteren, W. F.; Berendsen, H. J. C.; Hermans, J.; Hol, W. G. J.; Postma, J. P. M. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 4315-4319.
8. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.
9. MacKerell, A. D. et al., *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
10. Singh, U. C.; Kollman, P.A. *J. Comput. Chem.* **1986**, *7*, 718-730.
11. Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700-733.
12. Alagona, G.; Desmeules, P.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1984**, *106*, 3623-3632.
13. Bash, P. A. et al., *Biochemistry* **1991**, *30*, 5826-5832.
14. Lyne, P. D.; Mulholland, A. J.; Richards, W. G. *J. Am. Chem. Soc.* **1995**, *117*, 11345-11350.
15. Cunningham, M. A.; Ho, L. L.; Nguyen, D. T.; Gillilan, R. E.; Bash, P. A. *Biochemistry* **1997**, *36*, 4800-4816.
16. Antonczak, S.; Monard, G.; Ruiz-Lopez, M. F.; Rivail, J. –L. *J. Am. Chem. Soc.* **1998**, *120*, 8825-8833.
17. Bash, P. A.; Ho, L. L.; MacKerell, A. D.; Levine, D.; Hallstrom, P. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3698-3703.
18. Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 3448-3457.
19. Spezio, M.; Wilson, D. B.; Karplus, P. A. *Biochemistry* **1993**, *32*, 9906-9916.
20. Taylor, J. S.; Teo, B.; Wilson, D. B.; Brady, J. W. *Protein Engng.* **1995**, *8*, 1145-1152.
21. Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207-221.
22. Rouvinen, J. R.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. *Science* **1990**, *249*, 380-386.
23. Vyas, N. K.; Vyas, M. N.; Quiocho, F. A. *Science* **1988**, *242*, 1290-1295.
24. Quiocho, F. A.; Vyas, N. K. *Nature* **1984**, *310*, 381-386.
25. Quiocho, F. A. *Ann. Rev. Biochem.* **1986**, *55*, 287-315.

26. Vyas, N. K.; Vyas, M. N.; Quiocho, F. A. *J. Biol. Chem.* **1991**, *266*, 5226-5237.
27. Quiocho, F. A.; Wilson, D. K.; Vyas, N. K. *Nature* **1989**, *340*, 404-407.
28. Quiocho, F. A. *Pure & Appl. Chem.* **1989**, *61*, 1293-1306.
29. Henrissat, B. *Biochem. J.* **1991**, *280*, 309-316.
30. Zhang, S.; Lao, G.; Wilson, D. B. *Biochemistry* **1995**, *34*, 3386-3395.
31. Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199.
32. Valleau, J. P.; Torrie, G. M. in *Statistical Mechanics. Part A: Equilibrium Techniques*; Berne, B. J., Ed., Plenum Press: New York, NY, 1977; pp. 169-194.
33. Palma, R.; Himmel, M.; Brady, J. W. *J. Phys. Chem. B*. Submitted.
34. Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050-3054.
35. Postma, J. P. M.; Berendsen, H. J. C.; Haak, J. R. *Faraday Symp. Chem. Soc.* **1982**, *17*, 55-67.
36. Tembe, B. L; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281-283.
37. Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830-3832.
38. Wong, C. F.; McCammon, J. A. *Israel Journal of Chemistry* **1986**, *27*, 211-215.
39. Sakon, J.; Adney, W. S.; Himmel, M. E.; Thomas, S. R.; Karplus, P. A. *Biochemistry* **1996**, *35*, 10648-10660.
40. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187-217.
41. Pearlman, D. A. et al., *AMBER*; University of California: San Francisco, CA, 1991.
42. van Gunsteren, W. F.; Berendsen, H. J. C. *GROMOS*; University of Groningen, Groningen, The Netherlands, 1986.
43. Kouwijzer, M. L. C. E.; van Eijck, B. P.; Kroes, S. J.; Kroon, J. *J. Comput. Chem.* **1993**, *14*, 1281-1289.
44. Murcko, M. A.; DiPaola, R. A. *J. Am. Chem. Soc.* **1992**, *114*, 10010-10018.
45. Liang, D.; Fox, P. C.; Bowen, J. P. *J. Comput. Chem.* **1996**, *17*, 940-953.
46. Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745-5753.

Absorption of Endoglucanase I
and Cellobiohydrolase I of *Trichoderma reesei*
during Hydrolysis of Microcrystalline Cellulose

H. Ding[1], E. Vlasenko[2], C. Shoemaker[1], and S. Shoemaker[2]

[1]Department of Food Science and Technology and [2]California Institute of Food and Agricultural Research, University of California, Davis, CA 95616

Endoglucanase I and cellobiohydrolase I (EG I and CBH I) were purified from a commercial *Trichoderma* cellulase preparation. Adsorption of EG I and CBH I, alone and in combination, was investigated using microcrystalline cellulose. Capillary zone electrophoresis was developed as a sensitive (picomolar range) and reliable method for quantifying individual adsorbed enzymes. Adsorption kinetics as well as adsorption isotherms were evaluated for EG I and CBH I. The adsorption data followed the Langmuir adsorption equation. Adsorption parameters, A_{max} and K_p, were calculated for EG I and CBH I. CBH I showed higher binding affinity to Avicel, compared to EG I. The hydrolysis kinetic parameters, V_{max} and K_m, were determined for EG I and CBH I, and found to correlate with adsorption parameters.

Cellulose is the most abundant natural polymer on earth. The disposal of cellulose-containing waste materials, such as agricultural residues, waste wood, and municipal solid wastes, poses a serious problem. Considerable efforts have been directed toward developing bioprocesses to convert cellulosic wastes to ethanol and other chemicals (*1-6*). The key and rate-limiting step in cellulose-to-ethanol processes is enzymatic depolymerization of cellulose (*2,7*). Despite the tremendous efforts of researchers in this area, the mechanism of cellulose hydrolysis is not entirely understood.

The adsorption of cellulase is very important because adsorption is the prerequisite step in the enzymatic hydrolysis of cellulose. A good understanding of the adsorption of cellulases onto cellulose will provide further insights into the overall reaction mechanism and the observed synergism of cellulase components (*8-11*). Also, adsorption of cellulases may be utilized as a method of enzyme recycling (*12-14*).

131

Previously reported studies examining adsorption of cellulase mostly have used the cellulase system of *Trichoderma* (*8,9,12,14,15-21*). Adsorption studies based on a mixture of cellulase components do not give information on the adsorption of any particular component. Some studies using purified cellulase components have been reported (*10,11, 22,23,24*). However, these studies relied on total protein (*10,11*) and cellulase activity measurements (*10,24*) and did not allow specific measurement of cellulase components in their reconstituted mixtures. Thus, it was not possible to compare the adsorption of cellulases as single components and in reconstituted mixtures. Recently, fast protein liquid chromatography (FPLC) was suggested for quantitative determination of cellulase components in mixtures (*25,26*). There are only a few reports in the literature using FPLC to quantify the adsorption of cellulases during hydrolysis of lignocellulose (*16*) and microcrystalline cellulose (*22*). More recently, capillary zone electrophoresis (CZE) was reported to be a powerful analytical method to fractionate and quantify biomolecules such as proteins, DNA and polysaccharides (*27-31*). CZE is electrophoresis in free homogeneous solution. It separates substances based on their physico-chemical properties. The sensitivity of CZE with fluorescent detection is greater than most HPLC methods (*30,31*).

In this paper, purified EG I and CBH I from *Trichoderma reesei* cellulases were used for adsorption studies. The enzymes were studied alone and in reconstituted mixtures. Capillary zone electrophoresis was used to determine the amount of enzyme remaining in solution (free enzyme) so as to calculate the amount of enzyme adsorbed onto microcrystalline cellulose (bound enzyme). Adsorption parameters were obtained using the Langmuir adsorption equation. Hydrolysis kinetic parameters were also determined for EG I and CBH I based on the Michaelis-Menten equation, and the two sets of parameters were compared with each other.

Materials and Methods

Substrate. Avicel® PH-101 was purchased from FMC Corporation.

Enzyme Purification. EG I and CBH I were purified from Spezyme® CP cellulase (Genencor International, Inc.). The purification procedures are given in Figure 1. Isoelectric focusing (PhastGel IEF 4-6.5, Pharmacia) with Coomassie blue staining gave a single band for CBH I, and a predominant single band and a slight contaminating band for EG I. The isoelectric points of the purified EG I and CBH I, based on PI standards, were found to be 4.58 and 4.18, respectively. SDS gel electrophoresis (PhastGel Homogeneous 7.5, Pharmacia) with silver staining gave single bands for both EG I and CBH I. The molecular weight of the purified EG I and CBH I, based on molecular weight standards, were found to be 56,000 and 61,000 Da, respectively. These values were close to what has been previously reported (*32*). The purified cellulase components showed single peaks on FPLC Mono Q column (Pharmacia) at pH 6.5, as well as using capillary zone electrophoresis at pH 3.0 on a μSIL-FC capillary (J & W Scientific). Purified EG I and CBH I were concentrated by ultrafiltration, and the buffer was changed to 50 mM sodium acetate buffer, pH 5.0.

Protein contents were measured by the Lowry method (*33*). Enzyme activities were measured using carboxymethyl cellulose, Avicel, and H_3PO_4-swollen cellulose as substrates (Table 1).

Table 1. Activities of EG I and CBH I purified from *Trichoderma reesei* cellulase

Enzyme	Specific Activity (U/mg) towards		
	CM-cellulose	Avicel	H_3PO_4-swollen cellulose
EG I	74.6	0.12	24.3
CBH I	0.07	0.11	0.21

Specific activities were measured at 50 °C, pH 5.0 using the substrates, CM-cellulose, Avicel and H_3PO_4-swollen cellulose.
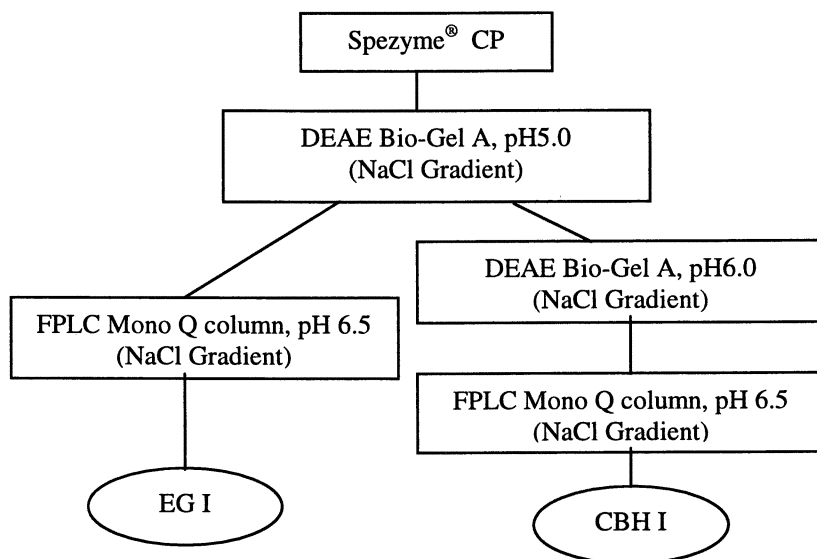
Figure 1. Flow chart of the purification of EG I and CBH I from a commercial cellulase preparation.

Quantitative Determination of EG I and CBH I. EG I and CBH I in solution were quantitatively determined by capillary zone electrophoresis using a 50 μm* 25 cm μSIL-FC capillary in a BioFocus®-3000 system (BioRad). Cartridge temperature was set at 20°C. Sodium acetate buffer (50 mM, pH 4.0) was used as the running buffer. Running voltage was set at 25 kV. Sample injection was performed at 15 psi*sec for

all samples. Ultraviolet (UV) absorbance of the eluent was followed at 200 nm. The separation of EG I from CBH I is shown in the electropherogram (Figure 2). Calibration curves based on peak areas were used to determine enzyme concentration in the injected sample.
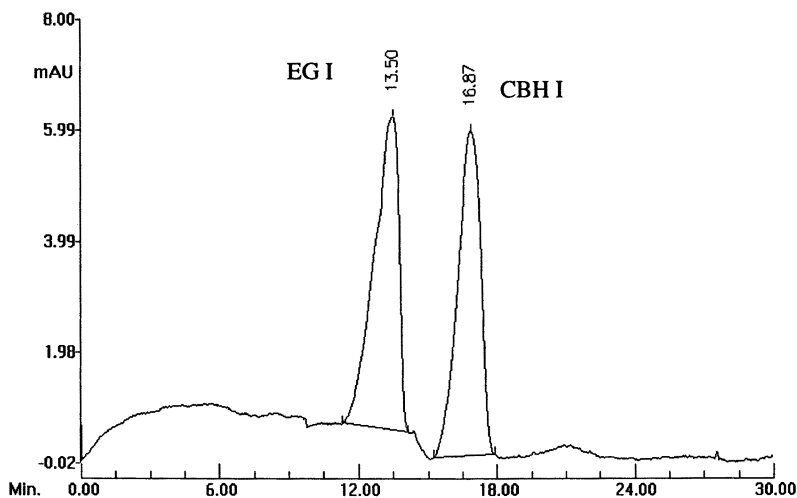
Figure 2. Electropherogram of sample containing EG I and CBH I. Sample was applied on a 50 μm* 25 cm μSIL-FC capillary, using 50 mM, pH 4.0 sodium acetate buffer as the running buffer.

Adsorption Kinetics.   Avicel PH-101 was used as the cellulose adsorbent. Avicel (25 mg) was suspended in 50 mM sodium acetate buffer, pH 5.0, and preincubated at 25 °C for 30 minutes. After preincubation, a specified amount of EG I and/ or CBH I was added to bring the total reaction volume to 5 mL. The enzyme-to-substrate ratio was set at 0.27 μmol/g for both EG I and CBH I (based on the calculated molecular weight for EG I and CBH I of 56,000 g/mol and 61,000 g/mol, respectively). Experiments were performed with EG I and CBH I, alone and in equimolar mixtures. The reaction mixture was contained in a 10 mL glass bottle with a screw cap, and stirred on a magnetic stirring plate. Samples were taken throughout a 60-minute incubation and centrifuged for 2 minutes at 10,000 rpm. The amount of free enzyme in the supernatant was determined by capillary zone electrophoresis. Adsorbed enzyme was calculated by subtracting the free enzyme remaining after time (t_x) from the initial enzyme (t_0).

Adsorption Isotherms. Avicel PH-101 was used in the adsorption isotherm studies. Avicel concentration was set at 5 mg/mL (as described in adsorption kinetic studies), and the enzyme-to-substrate ratio was varied from 0.27 μmol/g to 1.34 μmol/g (0.27, 0.40, 0.50, 0.60, 0.80, 1.34 μmol/g). The incubation time was set at 30 minutes in order to attain adsorption equilibrium. Experiments were performed with EG I and CBH I alone and in equal molar mixtures.

Hydrolysis Kinetics. Avicel PH-101 was used as the substrate for hydrolysis kinetics studies. Avicel was incubated at 25 °C in 50 mM sodium acetate buffer, pH 5.0, for 30 minutes, then a specified amount of EG I or CBH I was added to start the reaction. Substrate concentration varied in a range of 0.1 to 5.0 mg/ mL. Enzyme concentration was 10 ug/ mL for EG I, and 50 ug/ mL for CBH I. Total reaction volume was 20 mL. The reaction mixtures were contained in 25-mL flasks, and stirred on a magnetic stirring plate. Samples were taken over a 30-minute reaction time, reducing sugars produced were measured by the BCA method (*34*), and the initial rates were calculated.

Results and Discussion

Analytical Method. Capillary zone electrophoresis was found to be a reliable method to quantify the amount of enzymes. At the conditions established EG I was separated from CBH I completely (Figure 2), so that it was possible to quantify EG I and CBH I in mixtures. Purified EG I and CBH I were used as the standards to construct the calibration curves based on peak areas. Using the calibration curves, the enzyme concentration in the samples, ranging of 10 to 200 μg/mL, was determined. It was found that the standard deviation calculated from three injections was less than 6%.

Capillary zone electrophoresis is defined as electrophoresis in free homogeneous solution. In CZE no interactions between the substances and a gel or a micellar pseudo-stationary phase are present. Thus, CZE must be considered as the simplest electrophoretic technique since the substances are simply separated on the basis of the applied electric field and as a function of the physico-chemical properties of the substance, such as their charge-to-radius ratio and their relationship with the composition of the solution, i.e., the ionic strength and the pH (27).

CZE offers some advantages when compared with chromatographic techniques: the sample required is very small (injection volume less than 1 μL), and the separation is highly efficient, rapid and quantitative (27). Capillary zone electrophoresis gives a more direct separation and analysis when compared to chromatography techniques because it does not involve the chromatographic process of adsorbing proteins on the column followed by elution with a sodium chloride gradient or pH gradient. Using UV detection, the sensitivity of capillary zone electrophoresis is similar to that of FPLC and other chromatographic methods (picomolar range). The sensitivity of capillary zone electrophoresis can be further increased using fluorescence detection.

Adsorption Kinetics. Adsorption of EG I and CBH I on Avicel was determined by capillary zone electrophoresis during the 60-minute incubation at 25 °C. The enzyme to substrate ratio was set at 0.27 μmol/g for both EG I and CBH I. EG I and CBH I were used alone and in reconstituted mixtures. When these enzymes were used in mixtures, a fixed 1:1 molar ratio was used in order to compare the results obtained using the enzymes alone and in reconstituted mixtures. We have not compared these conditions with conditions present in *Trichoderma* cellulase systems where CBH I and EG I are about 60% and 15% of total protein by weight, respectively.

In earlier adsorption studies, the required time to reach equilibrium varied from 2 minutes (*35*) to 2 hours (*36*). The required time depends on the characteristics of the substrate, the origin of the enzyme and the enzyme to substrate ratio. In this study, 30 minutes was found to be adequate to reach adsorption equilibrium in all cases.

Figure 3 shows the adsorption process for EG I and CBH I, alone and in equimolar mixture. When used alone, about 92% of the initially added CBH I was adsorbed on Avicel at equilibrium, while about 75% of the total EG I was adsorbed. CBH I demonstrated a higher binding capacity on Avicel, compared with EG I. Comparing the adsorption of CBH I when used alone and when used together with equimolar EGI, the amount of adsorption was similar. This pattern also was observed for EG I, alone and in combination with equimolar CBH I (Figure 3).

There are various statements in the literature describing the interactions between enzyme components during adsorption on crystalline cellulose. Tomme et al. (*10*) reported synergistic adsorption of CBH I and CBH II on Avicel. They suggested that a "loose complex" between CBH I and CBH II occurs in solution prior to adsorption, thus more adsorption occurs when both enzymes are present. In contrast, Ryu et al. (*8*) and Kyriacou et al.(*23*) found competition between different cellulase components for adsorption onto cellulose. In the present study, the results in Figure 3 did not support the "loose complex" between EG I and CBH I, neither did it show competition between EG I and CBH I during adsorption. Instead, the results in Figure 3 suggest that there are some specific binding sites on Avicel for EG I and CBH I, respectively. At a low enzyme to substrate ratio, both EG I and CBH I mostly bound to their specific binding sites, so that no matter whether they were used alone or in an equimolar mixture, the amount of adsorption was similar. However, it should be noted that this phenomenon was found using a low enzyme to substrate ratio.

Adsorption Isotherms. In adsorption isotherm studies, EG I and CBH I were again used alone as well as in equimolar mixtures. The enzyme to substrate ratio was varied from 0.27 μmol enzyme / g Avicel to 1.34 μmol enzyme / g Avicel. The reaction mixtures were incubated at 25 °C for 30 minutes to ensure equilibrium conditions for enzyme adsorption under all conditions of this study.

Table 2 shows the conversion of the substrate during the 30-minute incubation at 25 °C. When EG I and CBH I were used alone, the conversion was less than 2% at the highest enzyme to substrate ratio. When EG I and CBH I were used together in the equimolar mixture, 4.6% substrate was converted with the highest enzyme to substrate ratio. These data reflect conditions where there was no significant degradation or

structural change on the substrate, conditions that are very important to interpreting adsorption data.
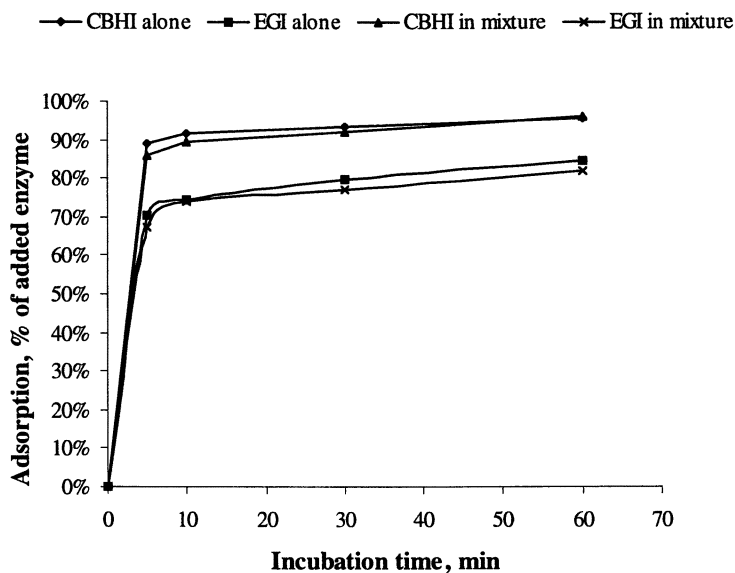
Figure 3. Adsorption kinetics of EG I and CBH I from Trichoderma reesei cellulases on Avicel PH-101. Incubation temperature was 25 °C. Enzyme to substrate ratio was 0.27 μmol/g for both EG I and CBH I. Enzymes were studied alone and in equimolar mixture. In the mixture, both EG I and CBH I were used at 0.27 μmol/g.

The adsorption isotherms of EG I and CBH I when used alone and in equimolar mixtures are shown in Figure 4. A predominance in the adsorption of CBH I was observed at all enzyme to substrate ratios. The amount of adsorption increased with the increase of enzyme to substrate ratio, until the substrate was gradually saturated by the enzymes. Comparing the adsorption of CBH I, alone and together with EG I, the amount of adsorption was similar at low enzyme to substrate ratios. However, at high enzyme to substrate ratios, more CBH I was adsorbed when used alone than when used together with EG I. In addition, the adsorption of EG I alone was found to be higher than EG I in the mixture at high enzyme to substrate ratios (Figure 4). These data provide evidence in support of competition between EG I and CBH I for binding sites on crystalline cellulose under condition of high enzyme to substrate ratios.

Based on the data from Figure 4, the sum of the adsorbed amounts of EG I alone and CBH I alone was calculated and showed in Figure 5 as the theoretical value of total adsorption, assuming no competition for binding sites between EG I and CBH I. The sum of the adsorbed amounts of EG I in mixture and CBH I in mixture also was

calculated and shown in Figure 5 as the measured value of total adsorption when these two enzymes were presented as an equimolar mixture. The comparison of the two lines in Figure 5 with low enzyme to substrate ratios (less than 0.5 μmol/g) indicates that the competition between EG I and CBH I is not obvious. However, when enzyme to substrate ratio increases, the competition increases. These results suggest that there are specific binding sites, as well as common binding sites on Avicel. When the enzyme to substrate ratio is low, enzymes mostly adsorb on their specific binding sites; when enzyme to substrate ratio increases, the enzymes start to bind on common sites, and the competition between enzymes for binding sites starts to show up. The higher the enzyme to substrate ratio is, the greater the competition.

Table 2. Conversion of the Substrate during Adsorption Processes

Enzymes	Enzyme Loading, umol/g	%Conversion (25°C, 30min)
EG I	0.27	0.63
	0.34	0.69
	0.40	0.72
	0.50	0.91
	0.80	1.28
	1.34	1.56
CBH I	0.27	0.33
	0.34	0.35
	0.40	0.41
	0.50	0.55
	0.80	0.92
	1.34	1.15
Mixture of EG I + CBH I	0.27+0.27	1.35
	0.34+0.34	1.68
	0.40+0.40	1.74
	0.50+0.50	2.41
	0.80+0.80	3.44
	1.34+1.34	4.64

The adsorption of cellulases on cellulose was reported to follow the Langmuir adsorption equation (9,11):

$$A = A_{max} K_p E / (1 + K_p E) \qquad (i)$$

where A is the bound enzyme, E is the free enzyme in solution at adsorption equilibrium, A_{max} is the maximun amount of enzyme that can be adsorbed onto a unit weight of cellulose, and K_p is the adsorption equilibrium constant. K_p gives the binding affinity of the enzyme on cellulose. Equation (i) can be rearranged as follows:

$$1/A = 1/ A_{max} K_p E + 1/ A_{max} \qquad \text{(ii)}$$

Equation (ii) was used to calculate the adsorption parameters, A_{max} and K_p, in the range where cellulose surface was not saturated by enzymes. Figure 6 shows the Langmuir plots for the adsorption of EG I, CBH I and the equimolar mixture of EG I and CBH I. A good fit was found between the adsorption data and the Langmuir adsorption equation.

Hydrolysis Kinetics. Adsorption of cellulases on crystalline cellulose is the prerequisite and integral step for hydrolytic action. Thus, it is interesting to compare the adsorption parameters, A_{max} and K_p, based on the Langmuir adsorption equation, with hydrolysis parameters, V_{max} and K_m, based on the Michaelis-Menten equation.

In the hydrolysis kinetic studies, the reaction conditions were set to be similar to those of the adsorption studies so as to compare hydrolysis kinetic parameters with adsorption parameters. Avicel PH-101 was used as a model microcrystalline cellulosic substrate, and the reaction temperature was set at 25 °C. Substrate concentration
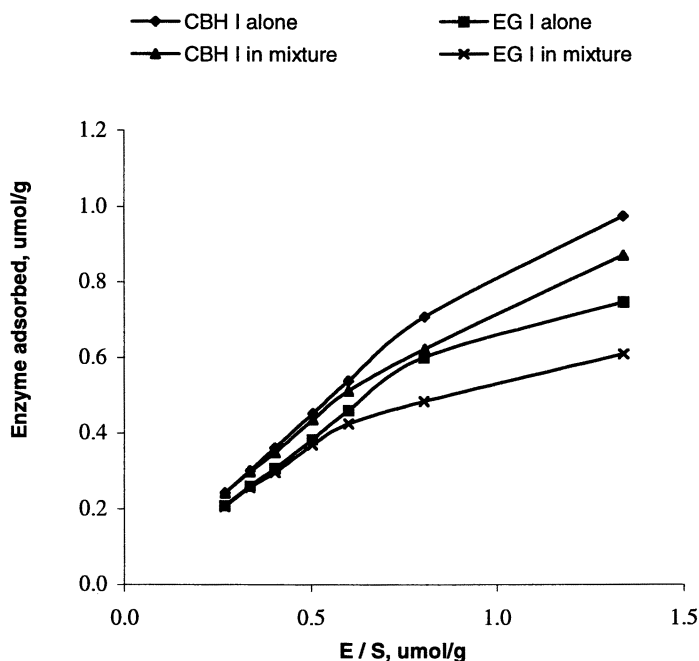
Figure 4. The adsorption isotherms of EG I and CBH I when used alone and when used in equimolar mixtures.

▲ EGI+CBHI, calculated
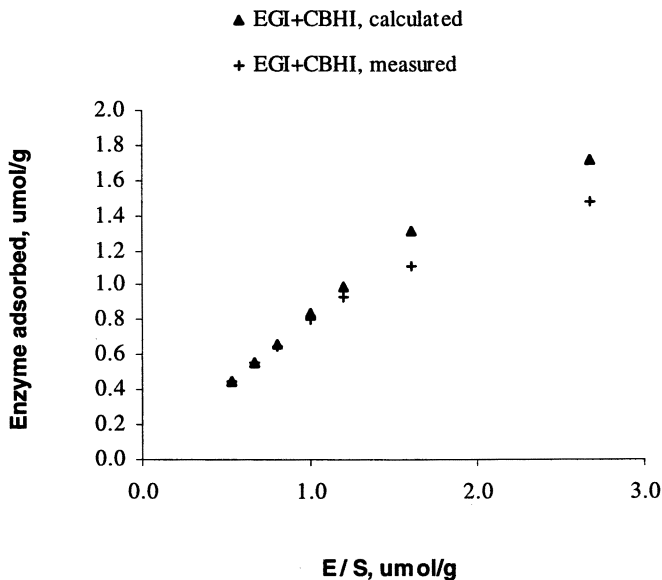
+ EGI+CBHI, measured

Figure 5. *The total adsorption of the equimolar mixtures of EG I and CBH Iat adsorption equilibrium. (EG I + CBH I, calculated) was determined as the sum of the adsorption of (EG I alone) and (CBH I alone) from Figure 4. (EG I + CBH I, measured) was determined as the sum of the adsorption of (EG I in mixture) and (CBH I in mixture) from Figure 4.*

varied in a range of 0.1 to 5.0 mg/ mL. Enzyme concentration was 10 ug/ mL for EG I, and 50 ug/ mL for CBH I. Reducing sugar produced during the 30-minute incubation was measured by the BCA method and the initial hydrolysis rate was determined as μmol/min/mg protein. It was found that under the conditions of this study, the reducing sugar produced per reaction time showed a linear relationship in all cases. The hydrolysis kinetic parameters were calculated using the Lineweaver-Burk plot, and the results are listed in Table 3, together with the adsorption parameters.

The absolute value of the parameters is not comparable, because of the differences in techniques used for adsorption studies and hydrolysis kinetic studies. However, it is interesting to note that K_p and K_m, as indicators of the affinity of the enzyme to the substrate, gave similar correlations. The K_p value was higher for CBH I, compared to EG I, indicating that CBH I has a higher affinity to Avicel. The K_m value also was higher for EG I, indicating that CBH I has a higher affinity to Avicel.

■ EGI alone   ◆ CBHI alone   ▲ EGI+CBHI
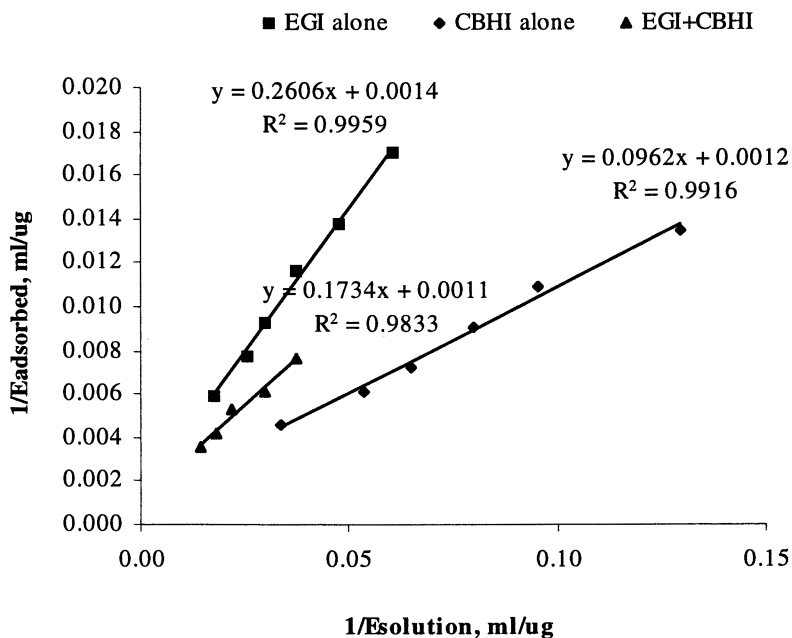
Figure 6. Langmuir plots for the adsorption of EG I, CBH I and equimolar mixture of EG I and CBH I.

Table 3. Adsorption Parameters and Hydrolysis Kinetic Parameters

Enzyme	Adsorption Parameters		Hydrolysis Kinetic Parameters	
	A_{max} mg E/g S	K_p L/g	V_{max} $\mu M/min$	K_m g/L
EG I	142.8	5.37	5.39	8.48
CBH I	166.7	10.40	2.11	1.82

Conclusions

Purified EG I and CBH I from *Trichoderma reesei* cellulases were used for adsorption studies on the microcrystalline cellulose, Avicel. Capillary zone electrophoresis was found to be an accurate method for determining the amount of free enzyme in solution, and was used to calculate the amount of adsorbed enzyme. Capillary zone electrophoresis was found to resolve EG I from CBH I down to pico molar range and could be used to accurately measure the adsorption of EG I and CBH I in a reconstituted mixture. CBH I was found to have a higher affinity to Avicel than was EG I. The adsorption data support the hypothesis that there are specific binding sites, as well as common binding sites for EG I and CBH I on cellulose. When the enzyme to substrate ratio was low, EG I and CBH I mostly bound to their specific binding sites and adsorption of one enzyme, alone or in combination with the other enzyme was similar. At high enzyme to substrate ratios, EG I and CBH I showed competition for common binding sites, and adsorption was lower when both enzymes were present, compared to when only one enzyme was present. The adsorption isotherm of EG I and CBH I followed the Langmuir adsorption equation. The adsorption parameters calculated by the Langmuir adsorption equation were found to correlate to the hydrolysis kinetic parameters calculated by the Michaelis-Menten equation.

Literature Cited

1. Gusakov, A.V.; Sinitsyn, A.P.; Manenkova, J.A.; Protas, O.V. *Appl. Biochem. Biotechnol.* **1992**, *34/35*, 625-637.
2. Beguin, P.; Aubert, J.-P. *FEMS Microbiol. Reviews* **1994**, *13*, 25-58.
3. Leschine, S.B. *Annu. Rev. Microbiol.* **1995**, *49*, 399-426.
4. Ingram, L.O.; Doran, J.B. *FEMS Microbiol. Reviews* **1995**, *16*, 235-241.
5. Duff, S.J.B.; Murray, W.D. *Biorresource Technol.* **1996**, *55*, 1-33.
6. Olsson, L.; Hahn-Hagerdal, B. *Enzyme Microb. Technol.* **1996**, *18*, 312-331.
7. Klyosov, A.A. *Biocheistry* **1990**, *29*, 10577-10585.
8. Ryu, D.D.Y.; Kim, C.; Mandels, M. *Biotechnol. Bioeng.* **1984**, *26*, 488-496.
9. Ooshima, H.; Burns, D.S.; Coverse, A.O. *Biotechnol. Bioeng.* **1990**, *36*, 446-452.
10. Tomme, P.; Heriban, V.; Claeyssens, M. *Biotechnol. Lett.* **1990**, *12*, 525-530.
11. Kim, D.W.; Jeong, Y.K.; Lee, J.K. *Enzyme Microb. Technol.* **1994**, *16*, 649-658.
12. Galbe, M.; Eklund, R.; Zacchi, G. *Appl. Biochem. Biotechnol.* **1990**, *24/25*, 87-102.
13. Tjerneld, F.; Persson, I.; Albertsson, P.-A .; Hahn- Hagerdal, B. *Biotechnol. Bioeng.* **1985**, *27*, 1036-1043.
14. Bader, J.; Bellgardt, K.-H.; Singh, A.; Kumar, P.K.R.; Schugerl, K. *Bioprocess Eng.* **1992**, *7*, 235-240.
15. Hogan, C.M.; Mes- Hartree, M.; Saddler, J.N.; Kushner, D.J. *Appl. Microbiol. Biotechnol.* **1990**, *32*, 614-620.
16. Yu, A.H.C.; Lee, D.; Saddler, J.N. *Biotechnol. Techniques* **1993**, *7*, 713-718.

17. Yu, A.H.C.; Lee, D.; Saddler, J.N. *Appl. Biochem. Biotechnol.* **1995**, *21*, 203-216.
18. Lee, N.E.; Woodward, J. *J. Biotechnol.* **1989**, *11*, 75-82.
19. Nidetzky, B.; Steiner, W. *Biotechnol. Bioeng.* **1993**, *42*, 469-479.
20. Bernardez, T.D.; Lyford, K.; Hogsett, D.A.; Lynd, L.R. *Biotechnol. Bioeng.* **1993**, *42*, 899-907.
21. Ishihara, M.; Uemura, S.; Hayashi, N.; Jellison, J.; Shimizu, K. *J. Ferment. Biioeng.* **1991**, *72*, 96-100.
22. Medve, J.; Stahlberg, J.; Tjerneld, F. *Biotechnol. Bioeng.* **1994**, *44*, 1064-1073.
23. Kyriacou, A.; Neufeld, R.J. *Biotechnol. Bioeng.* **1989**, *33*, 631-637.
24. Chernoglazov, V.M.; Ermolova, O.V.; Klyosov, A.A. *Enzyme Microb. Technol.* **1988**, *10*, 503-507.
25. Ellouz, S.; Durand, H.; Tiraby, G. *J. Chromatogr.* **1987**, *396*, 307-317.
26. Hayn, M.; Esterbauer, H. *J. Chromatogr.* **1985**, *329*, 379-387.
27. *Capillary Electrophoresis in Analytical Biotechnology*; Righetti, P.R.; CRC Series in Analytical Bioctechnology; CRC Press, Inc., Boca Raton, FL, **1996**; 239-276.
28. Chiesa, C.; O'Neill, R.A. *Electrophor.* **1994**, *15*, 1132-1140.
29. Chen, S.-M.; Wiktorowicz, J.E. *Anal. Biochem.* **1992**, *206*, 84-90.
30. Chen, F.-T.A.; Evangelista, R.A. *Anal. Biochem.* **1995**, *230*, 273-280.
31. Evangelista, R.A.; Liu, M.-S.; Chen, F.-T.A. *Anal. Chem.* **1995**, *67*, 2239-2245.
32. Shoemaker, S.; Watt, K.; Tsitovsky, G.; Cox, R. *Bio/Technol.* **1983**, *Oct.* 687-690.
33. Lowry, O.H.; Rosebrough, N.J.; Farr, A.E.; Randall, R.J. *J. Biol. Chem.* **1951**, *193*, 265-273.
34. Johnston, D.B.; Shoemaker, S.P.; Smith, G.M.; Whitaker, J.R. *J. Food Biochem.* **1998**, *22*, 301-319.
35. Woodward, J.; Hayes, M.K.; Lee, N.E. *Bio/Technol.* **1988**, *6*, 301-304.
36. Stahlberg, J.; Johansson, G.; Pettersson, G. *Bio/Technol.* **1991**, *9*, 286-290.

Chapter 9

Assessing the Efficacy of Cellulase Enzyme Preparations under Simultaneous Saccharification and Fermentation Processing Conditions

J. D. McMillan, N. Dowe, A. Mohagheghi, and M. M. Newman

Biotechnology Center for Fuels and Chemicals, National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO 80401–3393

A series of 2-level factorial cellulose conversion experiments was conducted to compare the efficacy of a reference commercial cellulase preparation and preparations produced in-house at NREL using a cellulosic substrate (Solka-floc). Experiments were performed in shake flasks with cellulose conversion of water-washed dilute-acid pretreated yellow poplar solids carried out using simultaneous saccharification and fermentation (SSF). Results demonstrate that significant differences exist in the efficacy of the two preparations, with the in-house preparation generally exhibiting superior performance. Non-linear regression of enzyme loading results show that to achieve above 80% cellulose conversion level in a 7-day SSF requires loading the commercial enzyme at a level above approximately 20 FPU/g cellulose, whereas enzyme loadings above about 10 FPU/g are sufficient using the in-house preparation. Studies carried out to assess the impact of β-gluocosidase (β-g) supplementation showed that increasing the β-g:FPU ratios substantially increased saccharification rates in both the commercial and NREL-produced preparations, but only increased absolute cellulose conversion levels substantially at low enzyme loading. Only a very small beneficial effect was observed in using unprocessed cellulase production whole broth rather than filtered cellulase broth. A small effect of solids loading was also observed, with cellulose conversion slightly decreasing at higher solids loading levels. Overall, results support the hypothesis that more effective cellulase enzyme preparations for complete cellulose hydrolysis are produced when cellulase production is induced in the presence of cellulose.

144

Introduction

Cellulase enzymes mediate the hydrolysis of crystalline cellulose to fermentable glucose. Not surprisingly, the production and utilization of cellulase enzymes represent critical cost factors in the development of commodity biomass-to-ethanol ("bioethanol") production processes. Our group at the National Renewable Energy Laboratory (NREL) is chartered with developing integrated and cost-effective enzymatic cellulose hydrolysis-based bioethanol process technology, and we have been focussing on the development of technology based on simultaneous saccharification and cofermentation (SSCF) of dilute-acid pretreated biomass (*1*). We refer to this as cofermentation rather than simply fermentation to emphasize that in SSCF all of the hemicellulosic sugars (e.g., xylose, arabinose, galactose and mannose) present in pretreated biomass hydrolyzates – not just the glucose released upon enzymatic hydrolysis of cellulose – are available for fermentative conversion to ethanol. In contrast, ethanol is only produced by fermentation of glucose and in some cases cellobiose in the better known simultaneous saccharification and fermentation (SSF) process.

Most recently, our efforts to develop integrated conversion technology have focussed on the key issues of enzyme production and enzyme usage. The general objectives are to minimize the cost of cellulase production by maximizing cellulase production volumetric productivity (*2*) and to reduce the cost of cellulase usage in the enzymatic hydrolysis step by identifying approaches that enable lower cellulase enzyme loadings to be used for cellulose hydrolysis. Here we describe some of our recent work exploring the efficacy of cellulase enzyme preparations on the enzymatic saccharification of dilute-acid pretreated hardwood cellulose (lignocellulose).

Previous work suggests that cellulase enzymes produced in the presence of cellulose exhibit higher efficacy than those produced using soluble sugar substrates (*3*). Thus, our purpose here was to carry out a fairly detailed comparison of two enzyme preparations, one produced using cellulose substrate and the other a commercial preparation presumably produced using soluble sugar substrate(s), to assess the influence of a variety of cellulase enzyme preparation quality factors on preparation efficacy. In addition to comparing the preparations directly, we examined the impact of spiking the preparations with β-glucosidase. Supplementation of the enzyme preparations with β-glucosidase was investigated because β-glucosidase is a component of the cellulase enzyme complex that is considered essential to minimize product inhibition of endoglucanase by cellobiose during cellulose hydrolysis. In addition to comparing the two types of preparations, we also investigated the impact of using unfiltered cellulase production broths. This was done because an integrated enzyme-based bioethanol process in which enzyme production occurs on site is expected to use unprocessed cellulase production broths ("whole broth") rather than clarified or filtered preparations.

Experimental Approach

A series of experiments were designed to better understand the factors determining cellulose conversion performance under SSCF type conditions. In order to facilitate our ability to perform factorial experiments involving multiple run conditions, we first sought to develop a shake flask screening protocol. Initial attempts to characterize cellulose conversion using batch saccharification, i.e., without simultaneous fermentation present but otherwise mimicking the initial conditions encountered in an SSCF, showed that significant inhibition of the enzyme system occurred. Inhibition resulted first from soluble sugars initially present in the system because real or synthetic hemicellulose hydrolyzate liquors were used as a background for carrying out the enzymatic cellulose hydrolysis reaction. Even when enzymatic hydrolysis was carried out in batch mode using washed pretreated lignocellulosic solids, i.e., with soluble sugars initially absent, inhibition of the reaction occurred as the sugars released upon cellulose hydrolysis accumulated in the system. Hence, we ultimately settled on using a shake flask screening methodology that is based on SSF. Details of the SSF protocol are described in the Materials and Methods section.

Our initial experiments focused on characterizing and quantifying the differences in cellulose conversion performance between two enzyme preparations. The first was an enzyme preparation produced in-house using a cellulosic substrate (Solka-floc). The second was a reference commercial preparation that we have used in much of our previous work. A comparative 2^3 (3 factors, 2 levels) factorial experiment with replicated centerpoints was run to compare the impact of enzyme source (in-house versus commercial), enzyme loading (10 filter paper units per gram (FPU/g) versus 25 FPU/g), and cellulose loading (2% w/w versus 6% w/w cellulose, corresponding to 3.3% w/w versus 10% w/w total pretreated yellow poplar hardwood sawdust (YP) solids, respectively). After establishing the performance of both enzyme preparations within this enzyme loading and solids loading design space, experiments were run to explore the impact of enzyme quality factors on cellulose conversion. This included examining the effect of supplementing each of the preparations with high levels of β-glucosidase, as well as comparing the performance of whole (unprocessed) broth versus filtered in-house enzyme preparations. These experiments were also run within the previously established enzyme loading and solids loading experimental design space in order to facilitate comparative statistical analysis.

In addition to the factorial SSF experiments designed to assess the influence of selected enzyme quality factors, SSF experiments were also run to quantify cellulose conversion performance as a function of enzyme loading. Again, detailed performance curves were developed for both the in-house and commercial enzyme preparations. The results of these experiments identify a probable range of enzyme loadings where cellulose conversion levels approaching or exceeding techno-economic performance objectives can be achieved. In the future, we will carry out a series of integrated simultaneous saccharification and cofermentation (SSCF)

experiments at selected enzyme loading levels in this range to confirm that the SSF screening methodology is indeed providing an acceptable predictor of SSCF performance.

Materials and Methods

Pretreated Lignocellulosic Solids. Pretreated yellow poplar hardwood sawdust (YP) solids were produced in NREL's Sunds hydrolyzer using a protocol similar to that described previously (*4, 5*). Pretreatment was carried out at a total solids level of approximately 20% w/w and at a temperature of 175°C, a sulfuric acid concentration of 0.8% w/w, and a residence time of 10 minutes. The pretreated solids were separated from the liquor using a Bock centrifuge (Bock Engineered Products, Toledo, OH). The dewatered solids were 47.5% w/w solids and contained 60.8% w/w cellulose (dry basis).

The dewatered pretreated YP solids were extensively washed with water prior to performing cellulose digestibility tests in order to remove entrained soluble sugars. Aliquots of approximately 50 g pretreated and dewatered solids were combined in a large centrifuge bottle, mixed well with 100 mL of water and centrifuged. The supernatant was then decanted and replaced with another aliquot of water. This washing step was repeated, typically for a total of 6 water wash cycles, until the concentration of glucose in the supernatant fell below 0.1 g/L. Pelleted washed solids contained approximately 30% w/w solids.

Cellulases. Two cellulase enzyme preparations were used, a commercial cellulase preparation ("CPN") from Iogen Corporation (Ottawa, Ontario, Canada) and a preparation produced at NREL ("in-house") by *Trichoderma reesei* strain L27 using insoluble cellulose (Solka-floc, a pure microcrystalline cellulose powder). The commercial preparation has been used for several years in integrated process development work at NREL and has become a *de facto* reference cellulase preparation against which in-house and other enzyme preparations are compared. The production protocol used to produce the three in-house *T. reesei* cellulase preparations or enzyme lots was similar to those described previously (*6,7*).

Novozyme 188, a commercial β-glucosidase (β-g) preparation (Novo Nordisk A/S, Bagsvaerd, Denmark), was added to the commercial cellulase preparation as well as to one of the in-house preparations to assess the impact of increasing β-g levels in the in-house and commercial preparations. In-house enzyme preparations were also examined in both filtered and whole broth forms. When whole broth was used, the enzyme was loaded based on measured filter paper activity titer (FPU/mL) of the supernatant and the volume that was added to achieve a target loading was adjusted to account for insoluble solids present in the whole broth. Table 1 summarizes the protein concentrations and enzyme activities in the six cellulase preparations investigated in these experiments.

Table 1. Summary of protein concentrations and enzyme activities in the cellulase preparations used in this work.

Enzyme Preparation Description	Total Soluble Protein	High MW Protein (>30 kDa)	Filter Paper Activity	β-g Activity	Estimated Cellulase Specific Activity	β-g:FPU Ratio
	(g/L)	(g/L)	(FPU/mL)	(IU/mL)	(FPU/g High MW Protein)	(IU/FPU)
Commercial (CPN)	205	158	55	147	348	2.7
Commercial β-g supplemented	189	147	43	202	293	4.7
In-house Lot 1 (Exp.38/F3)	40	32	11	17	344	1.6
In-house Lot 1 β-g supplemented	39	34	11	50	324	4.5
In-house Lot 2 (Exp.43/F4)	18	14	7	7	500	1.0
In-house Lot 3 (Exp.44/F1)	18	14	7	10	500	1.4

Microorganism. *Saccharomyces cerevisiae* D₅A (ATCC 200062) yeast was used as the fermentative microorganism in the simultaneous saccharification and fermentation (SSF) cellulose digestibility testing experiments. Stock cultures were stored at −70°C in 20% v/v glycerol solutions.

Inoculum Preparation. One milliliter of frozen stock *S. cerevisiae* D₅A culture was transferred into 200 mL of sterile YPD liquid medium (composed of 1% w/v yeast extract, 2% w/v peptone, and 5% w/v glucose) in a sterile 1-L Erlenmeyer flask. The flask was capped with a Morton closure and then the culture was incubated at 37°C and 150 rpm for 14-18 hours.

Enzymatic Cellulose Digestibility Testing Protocol. NREL's standard SSF-based cellulose digestibility testing protocol was carried out in 250-mL baffled shake flasks with water traps (*8*). Each shake flask was prepared to contain the following components: water-washed wood solids in an amount appropriate for the target cellulose loading for the particular experimental condition (2, 4, 6, 8, or 10% w/w); enzyme solution in an amount appropriate for the enzyme loading being tested (5, 7.5, 10, 12.5, 15, 17.5, 20, or 25 FPU/g cellulose); 5% w/w 1M citrate buffer at pH 4.8; 1% w/w yeast extract and 2% w/w peptone (YEP); yeast inoculum (sufficient to achieve an initial turbidity of 0.5 at 600 nm); and sufficient make up water to bring the total mass to 100 g.

After the flasks were filled with wet wood solids and water, they were capped with empty water traps, weighed, and then autoclaved for 30 minutes. After thoroughly cooling, the flasks were weighed again and water was added to make up for evaporative losses due to autoclaving. The buffer and YEP were then added and the SSF experiment was initiated by adding the cellulase enzyme and yeast inoculum and filling the water trap. A control containing no solids was typically run to determine the amount of ethanol produced from the enzyme and medium components alone. Flasks were incubated for 7 days in an Innova 4000 shaker incubator (New Brunswick Scientific, New Brunswick, NJ) operating at a temperature of 32 °C and a rotation rate of 100-200 rpm depending upon the experiment. The flasks were initiated at pH 5.2.

Analytical Procedures. Time course samples were analyzed for pH, soluble sugars and fermentation products. The concentrations of biomass sugars (cellobiose, glucose, galactose, mannose, xylose, and arabinose) and fermentation products (ethanol, acetic acid, glycerol, and lactic acid) were determined by High Performance Liquid Chromatography (HPLC) using Hewlett Packard model HP1090 HPLCs and standard methodologies based on the use of Biorad HPX-87H and HPX-87P columns (*9*). Mixed component concentration verification (CV) standards were periodically run with the HPLC samples to verify calibration accuracy, and analyses were repeated if the reported concentrations of CV standards deviated from their actual values by greater than ± 2.5%. Independent ethanol concentration measurements by gas chromatography were also made using a Hewlett Packard model HP5890 series

II gas chromatograph running an internal standard-based analysis method. Routine quantitation of glucose concentrations in wash supernatants and fermentation broths was performed using YSI enzymatic glucose analyzers (Yellow Springs Instrument Co., Yellow Springs, OH).

The filter paper activity titers (FPU/mL) of the various enzyme preparations were measured using a standard NREL protocol based on the IUPAC cellulase activity assay (*10, 11, 12*). The BCA protein assay was used to determine protein concentrations in the cellulase preparations (*13*) before and after filtration through 30 kDa molecular weight cut-off (MWCO) filters. Low molecular weight protein that passed through the 30kDa MWCO filter was not considered to be cellulase enzyme, but rather protein from the fermentation medium. Consequently, the concentration of the higher molecular weight cellulase enzyme proteins was estimated by subtracting the concentration of low molecular weight protein from the total protein concentration, i.e., to determine the concentration of protein >30kDa. The β-glucosidase activity of the cellulase preparations was measured using the PNPG assay (*14*). The percent dry weight of the wet solids was determined gravimetrically using an infra-red balance/drying oven. The cellulose content of solids, which is necessary to calculate how much enzyme solution to add to achieve a target cellulase loading, was measured using a standard NREL laboratory analytical procedure (*9, 15, 16*).

Calculations.

Cellulose Conversion. The percent cellulose converted to soluble sugars was determined based on the amount of ethanol produced. The concentration of ethanol produced from the enzyme and medium components alone was measured in control flasks included in each experiment. The net ethanol produced from cellulose in each flask was calculated by subtracting the concentration of ethanol produced in the enzyme and medium control flasks from the total concentration of ethanol produced in each flask. The amount of glucose required to enable this net amount of ethanol to be produced was then calculated assuming a maximum stoichiometric ethanol yield of 0.51 g ethanol per g consumed glucose. The glucose concentration calculated from the amount of ethanol produced was converted to a w/w basis assuming a density of 1.0 g/cm^3. The amount of cellulose required to produce this amount of glucose, i.e., the estimated amount of hydrolyzed cellulose, was calculated assuming 1.1 g glucose formed per g hydrolyzed cellulose. The percent cellulose conversion is defined as the ratio of hydrolyzed cellulose to initial cellulose multiplied by 100 percent.

Statistical Analysis. Analysis of variance (ANOVA) calculations were performed on different combinations of the individual and combined experimental results, as described in the results section, using Design-Expert 6 software (Stat-Ease, Inc. Minneapolis, MN). Factor effects calculated to be present at the 95% confidence level (p = 0.05) are defined as statistically significant.

Results

Comparative Performance of In-house and Commercial Enzyme Preparations.
A 2-level factorial shake flask SSF experiment was carried out to compare the
performance of the in-house and commercial enzyme preparations on the
saccharification of pretreated yellow poplar lignocellulose. The factors varied were
enzyme loading and solids loading. The low level of enzyme loading investigated
was 10 FPU/g and the high loading level was 25 FPU/g. The low and high levels of
solids loading tested were 3.3% w/w (2% w/w cellulose) and 10% w/w (6% w/w
cellulose), respectively. Centerpoint replicates were also run at an enzyme loading of
17.5 FPU/g and a solids loading of 6.7% w/w (4% w/w cellulose). Variable results
were obtained at the low solids level and these results are not presented here.

Figure 1 shows the extent of cellulose conversion at 4, 6 and 7 days plotted for
each condition tested. The error bars shown in Figure 1 represent plus and minus
one standard deviation from the average for replicated test conditions; this
convention is also in Figures 2, 3 and 4. As Figure 1 illustrates, better performance
was achieved at all of the tested conditions using the in-house cellulase enzyme
preparation. Whereas a cellulose conversion level of approximately 80% was
achieved after 7 days using the in-house preparation at the low enzyme loading level
of 10 FPU/g, the 7-day cellulose conversion level at the same conditions was below
45% using the commercial preparation. The differences between the performance
achieved by the two preparations were much smaller but still evident at the high
enzyme loading level of 25 FPU/g. In this case, 7-day cellulose conversion yields
were approximately 90% using the in-house preparation but below 85% using the
commercial preparation. The higher efficacy of the in-house preparation was also
observed in the replicated centerpoint tests carried out at a cellulose loading level of
4% w/w (6.7% w/w insoluble pretreated solids) and an enzyme loading of 17.5
FPU/g. A 7-day cellulose conversion level of 85% was achieved using the in-house
preparation, compared to a level of 71% using the commercial preparation.

The in-house preparation enabled somewhat higher initial cellulose conversion
rates to be achieved. The markedly higher performance levels obtained after 4 days
using the in-house preparation reflect this. While slower, however, the commercial
preparation was effective in sustaining cellulose conversion, and increases in
cellulose conversion levels from 4 to 6 days were generally more pronounced using
this preparation.

Analysis of variance (ANOVA) of the 7-day results indicates that enzyme
loading, enzyme source, and their two-way interaction are all statistically significant
factors affecting the overall cellulose conversion performance. Similar findings were
made by ANOVA of the 4- and 6-day results, with the exception that the two-way
interaction was not statistically significant at 4 days. Also, enzyme source was the
effect with the largest magnitude at 4 days, whereas enzyme loading exerted the
largest effect, followed by enzyme source, at 6 and 7 days.
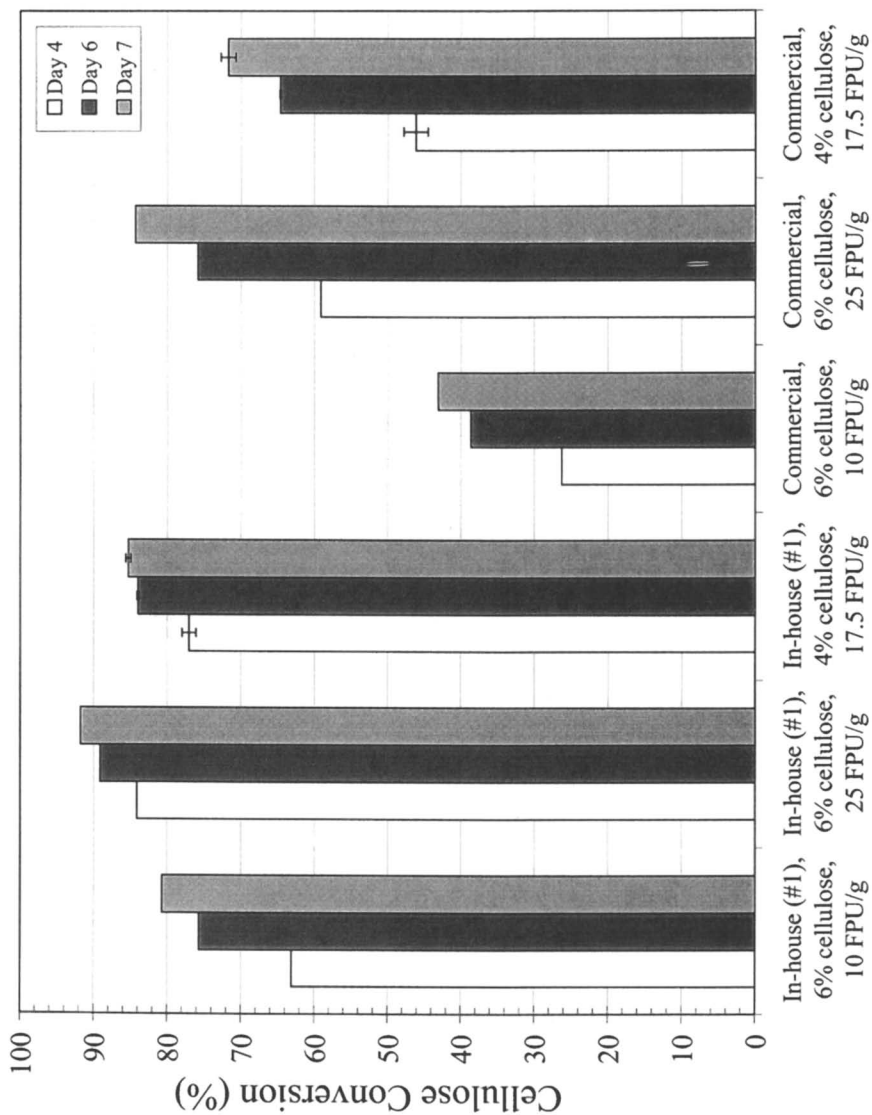
Figure 1. *Comparison of in-house (NREL-produced) and commercial cellulase preparations.*

Impact of β-glucosidase Supplementation. The impact of supplementing the enzyme preparations with β-glucosidase was explored within the same solids loading-enzyme loading design space used to compare the performance of the unsupplemented preparations. Figure 2 compares cellulose conversion results obtained for the β-glucosidase supplemented preparations. As this figure shows, the addition of β-glucosidase significantly improved the rates and in some cases increased the extents of cellulose conversion achieved by the two preparations. Comparison of Figures 1 and 2 shows that conversion levels after 4 days are substantially higher using the β-glucosidase-supplemented preparations, particularly in the cases where low conversion levels were obtained using the unsupplemented preparations. As a comparison of the centerpoint results depicted in Figure 2 shows, the benefit of β-glucosidase addition to increasing enzyme efficacy was most pronounced for the commercial preparation. In terms of absolute performance levels, however, the β-glucosidase-supplemented in-house preparation performed better than the β-glucosidase-supplemented commercial preparation. Whereas 7-day cellulose conversion levels above 86% were achieved at all conditions tested using the supplemented in-house preparation, 7-day conversion levels above 85% were only achieved at the centerpoint and high enzyme loading conditions using the supplemented commercial preparation.

The results presented in Figures 1 and 2 were combined and an ANOVA was performed on the pooled 2^3 factorial data set (three factors – enzyme source, enzyme loading, and β-glucosidase-supplementation – each present at two levels). Results of this analysis confirmed that the effect of β-glucosidase supplementation was statistically significant and demonstrated that the magnitude of this effect is generally smaller than that of enzyme loading and enzyme source.

The notably larger impacts of β-glucosidase supplementation on initial cellulose conversion rates and on the performance of the commercial enzyme preparation are also supported by ANOVA. The influence of enzyme source on the magnitude of the supplementation effect is reflected by a statistically significant two-way interaction between enzyme source and β-glucosidase supplementation. The impact on conversion rates is evident in two ways. First, at 4 and 6 days the magnitude of the supplementation effect is larger than any of the two way interactions, whereas by 7 days it has become smaller than the two-way interaction between enzyme loading and enzyme source. Second, the magnitude of the two-way interaction between enzyme source and β-glucosidase supplementation is the largest at 4 days.

Effect of Solids Loading. Although the 2% cellulose results are not presented here, statistical analysis of the results of the full factorial comparative experiments suggested that the solids loading level might be influencing cellulose conversion performance. Therefore, in an effort to better characterize the solids loading effect, a follow up SSF shake flask experiment was conducted using the commercial enzyme preparation. A higher range of solids loadings, 6%-10% w/w cellulose (corresponding to 10%-16.7% w/w insoluble pretreated solids), was investigated in this experiment. Figure 3 summarizes the combined experimental results obtained at
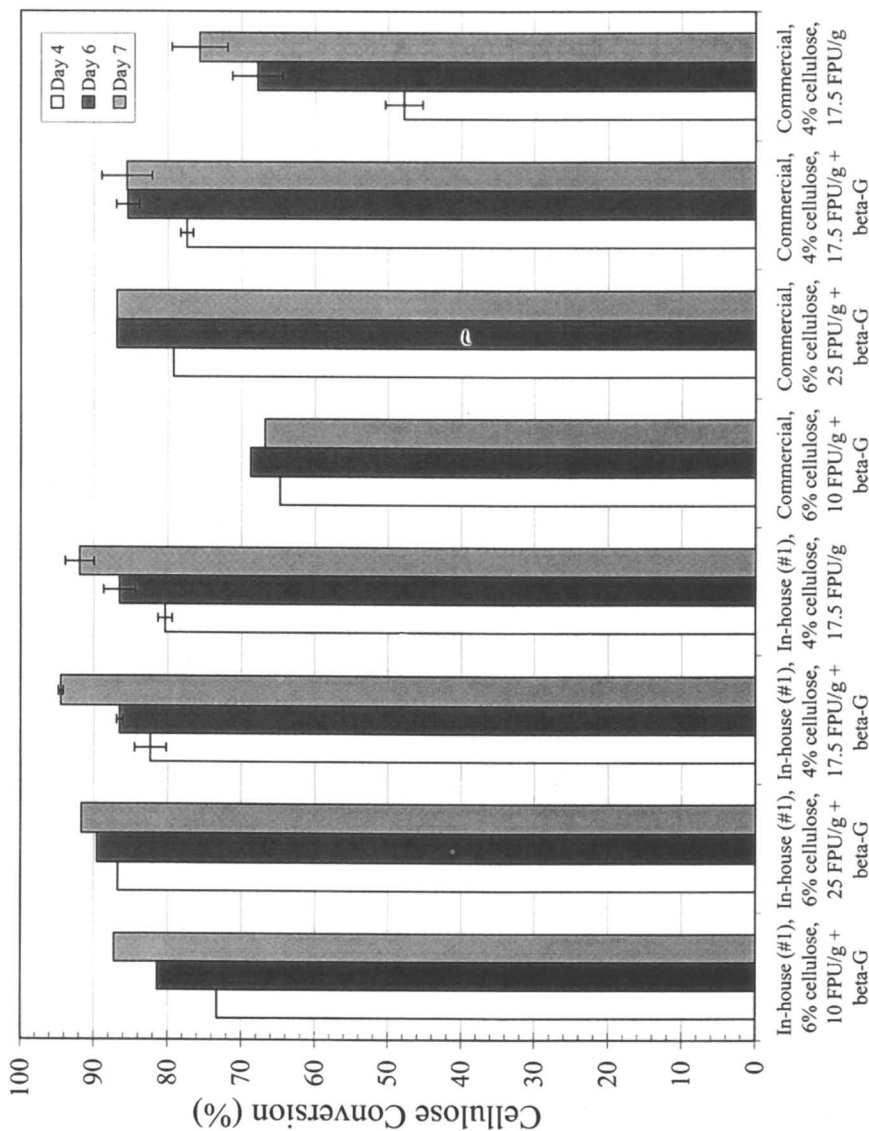
Figure 2. Impact of β-glucosidase supplementation on performance.
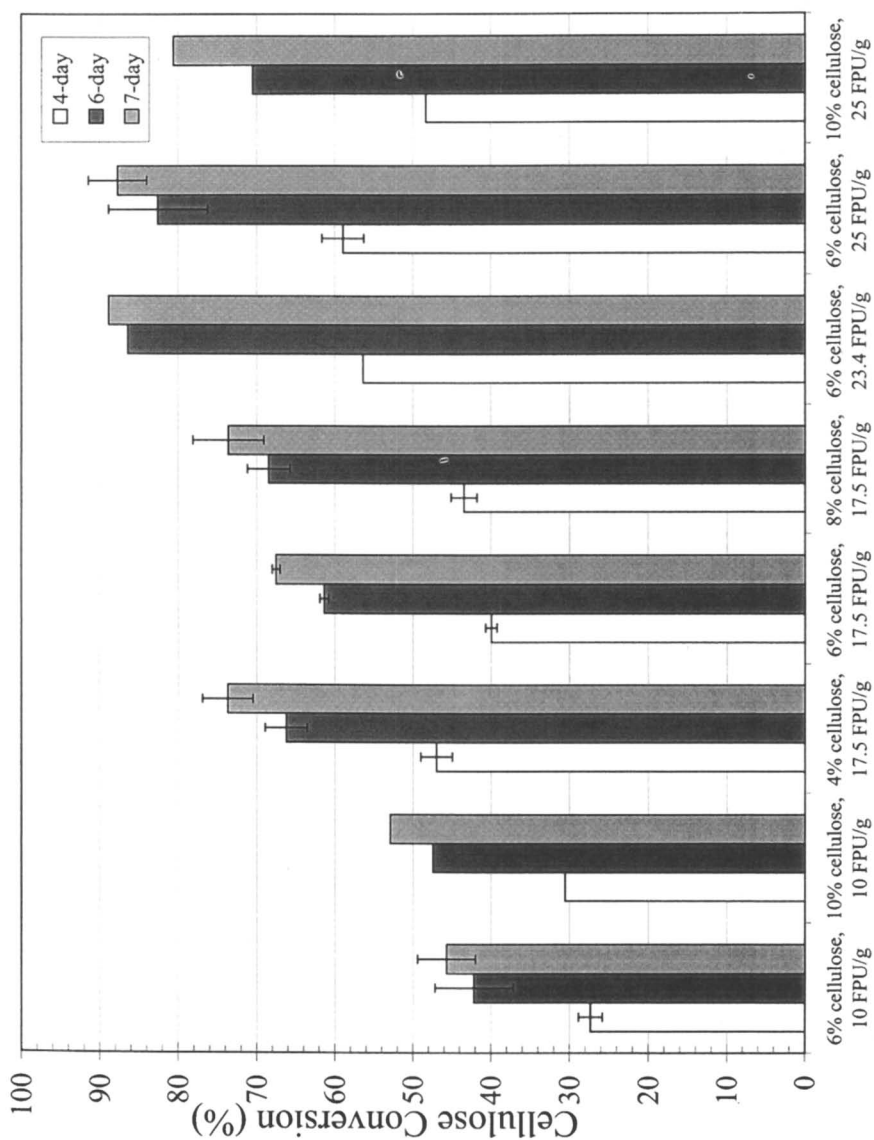
Figure 3. *Impact of cellulose loading level on the performance of the commercial cellulase preparation.*

the different cellulose loadings investigated. Although a large effect of cellulase enzyme loading is evident, with cellulose conversion increasing at increased enzyme loading, at a particular enzyme loading level the extent of cellulose conversion remains essentially unchanged at higher cellulose loading levels. For example, 7-day cellulose conversion levels remain at approximately 50% when the cellulose loading level is increased from 6% w/w to 10% w/w at the low enzyme loading level of 10 FPU/g. Similarly, 7-day cellulose conversion levels remain at approximately 70% when the cellulose loading level is increased from 4% w/w to 8% w/w at the centerpoint cellulase loading level of 17.5 FPU/g. However, a modest decrease in the extent of cellulose conversion achieved in 7 days is observed at the highest enzyme loading level of 25 FPU/g when the cellulose loading level is increased from 6% w/w to 10% w/w, with performance falling from 85-90% to 80%. Results of an ANOVA of this data show that solids loading does not have a statistically significant main effect on cellulose conversion performance. Although there is a significant two-way interaction between solids loading and enzyme loading, the magnitude of this effect is very small.

Comparison of Whole Broth and Filtered Preparations. Three independent cellulase production runs were performed using 5% cellulose (Solka-floc) to produce in-house cellulase enzyme preparations (lots) for comparing the performance of whole broth and filtered cellulase preparations. Preparations were once again compared within the baseline solids loading–enzyme loading design space. Figure 4 summarizes the results of these comparative experiments. As this figure shows, relatively little variability in performance was observed between the different lots of the in-house produced cellulase enzyme, with the exception that the lot 2 preparation generally achieved poorer performance than either lots 1 or 3, especially at the low enzyme loading level of 10 FPU/g. The differences between the SSF cellulose conversion performance achieved using filtered or whole cellulase production broth were smaller, with only a few exceptions, than the differences observed using the different lots. Figure 4 shows that upper range of cellulose conversion extents achieved at the 10 FPU/g cellulase enzyme loading level is approximately 80%. Cellulose conversion levels increase at higher cellulase loading levels, reaching 80%-90% at an enzyme loading of 17.5 FPU/g and 85%-93% at a loading of 25 FPU/g. ANOVA was performed on the combined data set. Results indicate that filtering the enzyme whole broth has no statistically significant effect on cellulose conversion performance. However, a significant two-way interaction between enzyme loading and enzyme production lot is present, reflecting the dissimilar performance achieved by the lot 2 preparation.

Performance as a Function of Enzyme Loading. The results of the factorial experiments described above indicated that there were substantial differences between the commercial and in-house cellulase preparations in terms of the enzyme loadings required to achieve high cellulose conversion levels. Additional SSF experiments were therefore performed to provide more detailed information about

how cellulose conversion levels varied as a function of enzyme loading for the two preparations. All of these experiments were performed at a fixed cellulose loading of 6% w/w (10% w/w insoluble pretreated solids). Figure 5 summarizes the 7-day cellulose conversion results achieved using enzyme loadings of 5-25 FPU/g. Also shown in this figure are results of replicated factorial centerpoint runs performed at 4% w/w cellulose loading and an enzyme loading of 17.5 FPU/g, as well as the results obtained using β-glucosidase-supplemented preparations and whole broth (unfiltered) in-house preparations. The solid lines shown in Figure 5 are best-fit second order polynomials fitted to the 6% w/w cellulose conversion data for either the filtered in-house preparation or the commercial preparation.

Figure 5 clearly illustrates that dramatic differences exist between the efficacy of the two cellulase preparations. The best fit polynomial lines indicate that an 80% cellulose conversion level (depicted by the dashed line) requires using the commercial enzyme at a loading of approximately 23 FPU/g cellulose, whereas a loading of only about 14 FPU/g is required using the in-house preparation; these are perhaps conservative estimates, as some of the data points lie above the best fit lines. Figure 5 also shows that conversion levels achieved by the whole broth in-house preparations tend to be on the high side of the range of performance levels achieved at a given enzyme loading. The beneficial positive impact of supplementing the preparations with β-glucosidase can also be seen in Figure 5. Cellulose conversion levels increase by 10-20% over the levels achieved using the unsupplemented preparations when enzyme is loaded at the low enzyme loading level of 10 FPU/g.

As illustrated by the results depicted in Figure 5, the cellulose hydrolysis performance levels of the commercial and in-house cellulase preparations under SSF conditions are similar at high cellulase loading levels, but the in-house preparations reproducibly exhibit superior efficacy at enzyme loadings below 20 FPU/g. This difference between the two types of preparations is also evident when their performance is compared on a protein loading rather than an activity loading basis. Figure 6 shows the same cellulose conversion results as Figure 5, but in this case plotted as a function of high molecular weight (MW) protein loading per gram cellulose. High MW protein is defined as protein retained by a 30 kDa MWCO filter. As Figure 6 illustrates, with the in-house preparations SSF cellulose conversion levels of 80% or higher are achieved using high MW protein loadings above 35 mg high MW protein/g cellulose. In contrast, protein loadings above 65 mg high MW protein/g cellulose are required to achieve this level of performance using the commercial preparation.

Discussion

The results of the shake flask studies comparing the NREL-produced ("in-house") and reference commercial cellulase preparations are intriguing. Most significantly, they show that substantial differences in enzyme efficacy under SSF processing conditions can exist between cellulase enzyme preparations that have been assayed
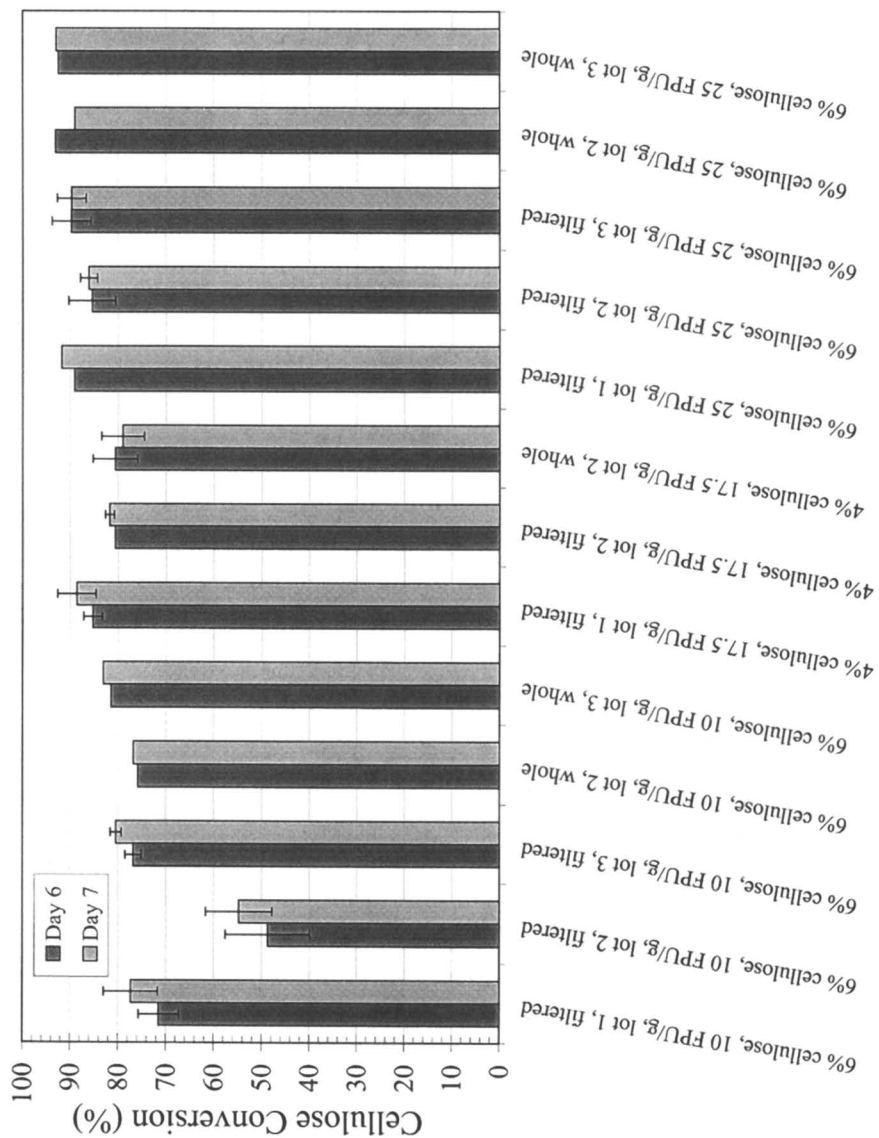
Figure 4. Comparative performance of whole broth and filtered in-house cellulase preparations.
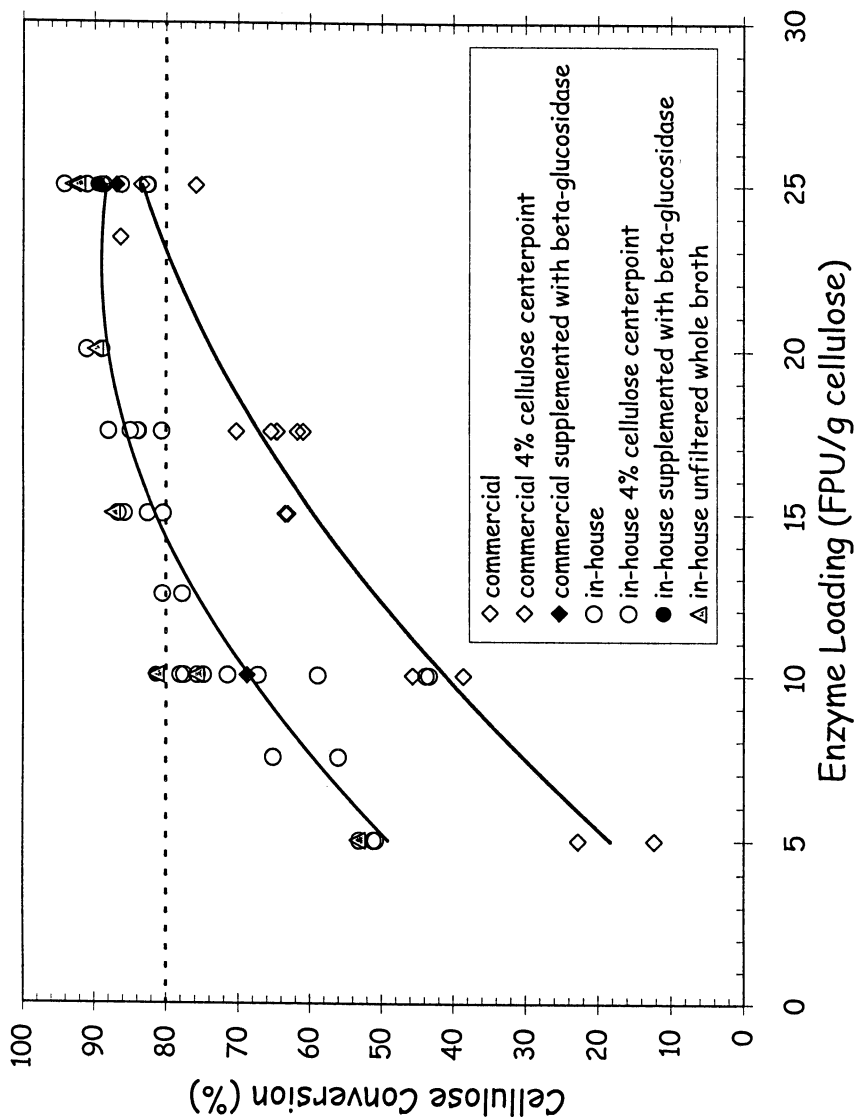
Figure 5. *7-day performance as a function of enzyme loading (6% w/w cellulose data).*

for cellulolytic activity and assessed on equivalent filter paper activity or protein loading bases. The results clearly demonstrate that within the investigated design space the in-house cellulase preparations exhibit substantially higher cellulose hydrolyzing capabilities in SSF than the commercial cellulase preparation. More specifically, the results demonstrate that 7-day cellulose conversion levels above 80% can be achieved at enzyme loadings below 17.5 FPU/g under SSF conditions using unprocessed in-house enzyme preparations. The SSF results also suggest that conversion levels above 80% can be achieved at enzyme loadings as low as 10 FPU/g, and perhaps lower, using in-house preparations enriched in β-glucosidase.

While the differences between the efficacy of the in-house and commercial preparations are striking, further work will need to be performed to understand why the cellulase preparations produced in-house are performing so much better than the reference commercial preparation. One reason may be that higher efficacy cellulase enzyme complexes are synthesized when enzyme production is induced in the presence of cellulose. We and our co-workers at NREL have some evidence that this occurs (*1, 3*), but additional studies are needed to more rigorously test this hypothesis. There are several approaches that could be used to explore this possibility. For example, cellulase preparations produced by other *Trichoderma reesei* strains growing on cellulosic substrates could be tested to determine if they also exhibit high efficacy. Ultimately, the efficacy of cellulase preparations produced by particular strains using soluble sugar substrates (e.g., lactose) should be compared with that of preparations produced using cellulosic substrates. An alternative potential cause for the observed difference between the two preparations could be that the reference commercial cellulase preparation is suboptimal and for whatever reason(s) (e.g., degradation during long-term storage) exhibits poor efficacy for complete saccharification of lignocellulosic substrates. Testing several lots of the reference commercial enzyme preparation as well as other commercial preparations recommended for complete saccharification applications by reputable enzyme manufacturers would help to evaluate this possibility. The lot to lot variability of the commercial preparation is unknown, as we have used a single lot in all of the work carried out to date. Developing a more fundamental understanding of the particular components of the cellulase enzyme complex that are different between higher and lower efficacy preparations will require application of biochemical methods such as 2-dimensional gel electrophoresis that can be used to identify and quantify the different proteins present in cellulase preparations.

The effect of supplementing the in-house and commercial preparations with β-glucosidase was investigated to explore the impact of enzyme quality characteristics beyond filter paper activity on saccharification efficiency. (While supplementing with β-glucosidase is not likely to be economically feasible based on NREL's current estimates of commercial enzyme costs and on-site enzyme production economics, it is possible that more optimal cellulase preparations can be produced in the future. A better understanding of what enzyme quality characteristics are necessary to minimize enzyme loading requirements is needed to guide such efforts.) Results show that β-glucosidase supplementation makes very little difference in enzyme

efficacy at high enzyme loadings (25 FPU/g), but significantly increases the efficacy of both cellulase preparations at low enzyme loadings (10 FPU/g). The magnitude of the beneficial effect of β-glucosidase supplementation at low enzyme loading is largest for the commercial preparation. This is somewhat perplexing since, as Table 1 shows, spiking the commercial preparation with β-glucosidase only increased its β-glucosidase:FPU activity ratio by 74% (from 2.7 to 4.7), whereas spiking the in-house preparation increased its β-glucosidase:FPU activity ratio by 181% (from 1.6 to 4.5). The fact that the magnitude of the beneficial impact of β-glucosidase supplementation doesn't correlate with the relative changes in the β-glucosidase:FPU ratios of the two preparations may indicate that some cellulase component other than β-glucosidase is actually limiting the efficacy of the preparations. For instance, it has been suggested that these results might be caused by low level contamination of the β-glucosidase preparation with endoglucanase(s). Based on this speculation, a detailed examination to assess the purity of the β-glucosidase preparation used in these experiments is warranted.

Regardless of the ultimate cause of the observed β-glucosidase effect, there clearly is a large benefit to supplementing cellulase preparations with β-glucosidase to increase their cellulolytic efficacy at low FPU loadings. This finding convincingly illustrates the critical role played by enzyme quality factors beyond FPU activity measurements alone in determining enzyme efficacy in hydrolyzing pretreated biomass substrates. Further work with a variety of different cellulase preparations is needed to more clearly elucidate the role of cellulase preparation quality factors beyond FPU and β-glucosidase:FPU ratio on saccharification efficiency.

The impact of post-production processing of cellulase preparations, i.e., centrifugation and filtration, on their efficacy was studied because, especially for the case of on-site enzyme production, it is desirable to minimize downstream purification and concentration of cellulase preparations and thereby reduce overall bioethanol production costs. The results summarized in Figures 4, 5 and 6 suggest that there is at most only a modest benefit to using unprocessed whole broth cellulase preparations. Moreover, this finding is at least partially confounded by the fact that whole broth preparations and filtered preparations were loaded on an equivalent free liquid basis. Thus a greater total volume of enzyme broth slurry was added when whole broth enzyme preparations were used, since they contained low levels of residual insoluble solids that were not present in the corresponding filtered preparations. Carry over of unconverted cellulose in the whole broth preparations may have contributed to greater ethanol production and therefore to higher calculated cellulose conversion levels as compared to SSFs run with filtered preparations.

Although the results of the whole broth studies suggest only a small benefit in terms of enzyme efficacy, there are a variety of other anticipated advantages to using unprocessed cellulase whole broth preparations for large-scale bioethanol production. For example, while it is generally recognized that cellulase enzymes are synthesized intracellularly and then secreted into the fermentation broth, it has been reported that as much as 50% of the β-g activity and 7% of the endoglucanase activity remain
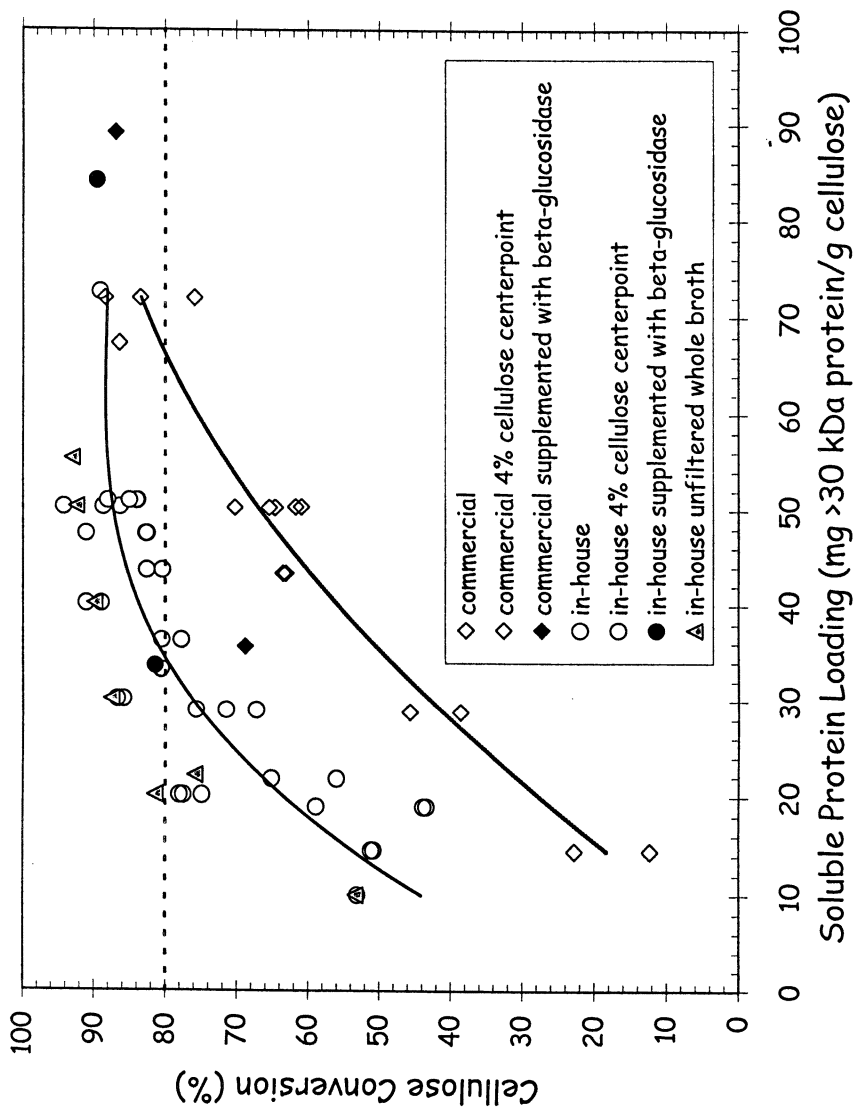
Figure 6. 7-day performance as a function of high MW protein loading (6% w/w cellulose data).

bound to the fungal cells (*17*). This suggests that it may be possible to significantly reduce enzyme loading requirements in an ethanol production process by using cellulase preparations that still contain fungal cells and cell debris. This opportunity does not typically exist with commercial cellulase enzyme preparations that are supplied as concentrated filtrates. The idea of using cellulase production whole broth in an SSF process to improve process economics was first reported by Takagi et al. (*18*). These investigators reported a 25% improvement in SSF ethanol yield when whole broth was used instead of culture filtrate. Work previously carried out at NREL observed a 8-25% increase in SSF ethanol yields when *Trichoderma reesei* strain L27 whole enzyme production broth was used in an SSF process (*7*). These findings suggest that additional efforts are warranted to investigate the potential to further reduce enzyme loading requirements by using cellulase whole broth. One possibility for the differences observed between the present work and previous reports may be that cellulase production broths were harvested at different times. Further investigation is needed to characterize the influence of harvest time on the amount of cell-bound cellulase component activity, and thus to better understand the potential benefit of using unprocessed cellulase preparations rather than filtered preparations.

In addition to intrinsic variability in the filter paper assay, small differences in harvest time may partially explain why we observed unexpectedly high lot-to-lot variability in the efficacy of the three in-house preparations used for these experiments. In particular, there was a clear difference in the performance of lot 2 compared to lots 1 and 3, with the former showing markedly lower efficacy at low enzyme loadings. As Table 1 shows, lots 2 and 3 exhibited similar FPU activity titers and high MW protein concentrations. However, lot 3, which had a higher β-glucosidase titer and therefore an increased β-glucosidase:FPU activity ratio, clearly performed better than lot 2. While lot 1 had higher FPU and β-glucosidase titers than either lots 2 or 3, its β-glucosidase:FPU activity ratio was high and similar to that for lot 3. These observations suggest that β-glucosidase:FPU activity ratio is an important measurable quality factor that can be used to correlate, at least in part, the relative efficacy of different preparations and production lots.

The lot-to-lot differences exhibited by the in-house preparations confounded the ability to discern whether or not unprocessed cellulase preparations enable better SSF cellulose conversion to be achieved than filtered preparations. In particular, the largest differences between whole broth and filtered cellulase preparations were observed using lot 2 enzyme. For reasons that remain unclear, the lot 2 cellulase preparation had substantially more unconverted cellulose (not quantified, but apparent by visual observation) than either lots 1 or 3. Thus, additional cellulose (and cellulase enzymes adsorbed on this cellulose) was added to SSF flasks receiving lot 2 whole broth. This unquantified additional amount of cellulose was not considered in the calculations to estimate cellulose conversion levels. Comparing the performance of a greater number of in-house produced preparations is needed to better quantify the impact of processing by filtration.

The effect of solids loading was also explored in the SSF shake flask experiments but the results were equivocal. The much larger effect of enzyme loading within the experimental domain investigated made it relatively difficult to observe any substantial solids loading effect. While the results clearly indicate that solids loading has a small but statistically significant effect, the impact of solids loading within a cellulose loading range of 4% w/w to 10% w/w (6.7% w/w to 16.7% w/w insoluble lignocellulosic solids) is small enough that further efforts to more accurately quantify the magnitude and other features of this effect are probably not warranted.

Conclusions

Significant differences exist in the efficacy of a reference commercial cellulase preparation and cellulase preparations produced at NREL using cellulosic substrates. Results of comparative shake flask SSF experiments show that the in-house preparations exhibit markedly superior performance at enzyme loadings below 17.5 FPU/g cellulose or 50 mg high MW protein/g cellulose. Results also show that increasing the β-glucosidase:FPU activity ratio in either type of preparation substantially increases cellulose saccharification rates and modestly increases the extent of cellulose conversion. These findings suggests that one strategy for achieving further reductions in the cellulase loading levels required to achieve extensive cellulose conversion (and thereby enable high-yield bioethanol production) is to develop enzyme production protocols that maximize the β-glucosidase:FPU activity ratio in addition to achieving high FPU productivity, yield, and titer.

Although it was anticipated based on the literature that there would be a beneficial impact to using unprocessed cellulase preparations rather than filtered preparations, only a very small positive effect of using whole broth to increase cellulose conversion performance was observed. The difficulty in observing this effect was at least partially the result of confounding caused by unanticipated lot-to-lot variations in the quality characteristics of the in-house-produced enzyme preparations. The reasons underlying the lot-to-lot variability are unclear but may indicate differences in the time at which enzyme broths were harvested. A solids loading effect was also observed in the comparative SSF tests, but the magnitude of this effect was very small.

The results support but do not prove the hypothesis that higher efficacy cellulase enzyme preparations are produced when cellulase production is induced in the presence of cellulose, e.g., by using a cellulosic substrate. A comparison of cellulase enzyme preparations produced by a particular strain on both soluble sugar and insoluble cellulosic substrates is needed to rigorously test this hypothesis. As it stands, our in-house cellulase enzymes produced using cellulose substrate exhibited an approximately twofold higher efficacy than the reference commercial preparation on both FPU activity and high MW protein bases. Whereas greater than 80% cellulose conversion was only achieved in 7 days using the commercial preparation

at enzyme loadings above approximately 20 FPU/g cellulose or 60 mg high MW protein/g cellulose, this level of conversion was achievable using the in-house preparations at loadings above 10 FPU/g cellulose or 30 mg high MW protein/g cellulose.

Acknowledgments

References

1. McMillan, J.D.; Newman, M.M.; Templeton, D.W.; Mohagheghi, A. *Appl. Biochem. Biotechnol.* **1999**, 77/79, 649-665.
2. Hayward, T.K.; Hamilton, J.; Tholudur, A.; McMillan, J.D. *Appl. Biochem. Biotechnol.* **2000**, *in press.*
3. Baker, J.O.; Vinzant, T.B.; Ehrman, C.I.; Adney, W.S.; Himmel, M.E. *Appl. Biochem. Biotechnol.* **1997**, 63/65, 585-595.
4. Nguyen, Q.A.; Dickow, J.H.; Duff, B.W.; Farmer, J.D.; Glassner, D.A.; Ibsen, K.N.; Ruth, M.F.; Schell, D.J.; Thompson, I.B.; Tucker, M.P. *Bioresource Technol.* **1996**, 58, 189-196.
5. Tucker, M.P.; Farmer, J.D.; Keller, F.A.; Schell, D.J.; Nguyen, Q.A. *Appl. Biochem. Biotechnol.* **1998**, 70/72, 25-35.
6. Mohagheghi, A.; Grohmann, K.; Wyman, C.E. *Appl. Biochem. Biotechnol.* **1988**, 17, 263-277.
7. Schell, D. J.; Hinman, N. D.; Wyman, C. E.; Werdene, P. J.; *Appl. Biochem. Biotechnol.* **1990**, 24/25, 287-296.
8. Hayward, T.K.; Combs, N.S.; Schmidt, S.L.; Philippidis, G.P. SSF Experimental Protocols: Lignocellulosic Biomass Hydrolysis and Fermentation. Laboratory Analytical Procedure #008, Biofuels Ethanol Project, National Renewable Energy Laboratory, Golden, CO, May 31, 1995.
9. Ehrman, C. I. In *Handbook on Bioethanol: Production and Utilization;* Wyman, C.E., Ed.; Taylor and Francis: Washington, DC, 1996; pp 395-415.
10. Ghose, T. K. *Pure & Appl. Chem.* **1987,** 59, 257-268.
11. Miller, G.L. *Anal. Chem.* **1959,** 31, 426-428.
12. Adney, B.; Baker, J. Measurement of Cellulase Activities. Laboratory Analytical Procedure #006, Biofuels Ethanol Project, National Renewable Energy Laboratory, Golden, CO, July 18, 1996.

13. Smith, P.K.; Krohn, R.I.; Hermanson, G.T.; Mallia, A.K.; Gartner,F.P.; Provenzano, M.D.; Fujimoto, E.K.; Gorke, N.M.; Olson, B.J.; Klenk, D.C. *Anal. Biochem.* **1985**, 150, 76-85.

14. Wood, T.M. *Biochem. J.* **1971,** 121, 353-362.

15. Karr, W.E.; Cool, L.G.; Merriman, M.M.; Brink, D.L. *J. Wood Chem. Technol.* **1991,** 11, 447-463.

16. Ruiz, R.; Ehrman, T. Determination of Carbohydrates in Biomass by High Performance Liquid Chromatography. Laboratory Analytical Procedure #002, Biofuels Ethanol Project, National Renewable Energy Laboratory, Golden, CO, February 14, 1996.

17. Acebal, C.; Castillon, M.P.; Estrada, P. *Biotech. Appl. Biochem.* **1988,** 10, 1-5.

18. Takagi, M.; Abe, S.; Suzuki, S.; Emert, G.H.; Yata, N. Proc. Bioconv. Symp. **1977,** IIT, Delhi, India, p 551.

Chapter 10

Phage Display of Cellulose Binding Domains for Biotechnological Application

Itai Benhar[1], Aviva Tamarkin[1], Lea Marash[1], Yevgeny Berdichevsky[1], Sima Yaron[2], Yuval Shoham[2], Raphael Lamed[1], and Edward A. Bayer[3,4]

[1]Department of Molecular Microbiology and Biotechnology, The George S. Wise Faculty of Life Sciences, Tel-Aviv University, Ramat Aviv, Israel
[2]Department of Food Engineering and Biotechnology, Technion IIT, Haifa, Israel
[3]Department of Biological Chemistry, The Weizmann Institute of Sciences, Rehovot, Israel
[4]Corresponding author

In recent years, cellulose-binding domains (CBDs) derived from the cellulolytic systems of cellulose-degrading microorganisms have become a focal point of attention for a wide range of biotechnological applications. The low cost and availability of cellulose matrices have rendered CBDs attractive as affinity tags for the purification and immobilization of a plethora of proteins. Intensive studies of cellulose degradation pathways and the identification of components of the cellulose-degrading machinery have contributed significantly to our understanding of the structure and function of CBDs. The time is now ripe to utilize engineered CBDs to improve existing applications and to devise novel ones.

Here we describe our recent results of experiments where the *Clostridium thermocellum* scaffoldin CBD was genetically engineered for such purposes. We describe the development of a novel phage display system where the *C. thermocellum* CBD is displayed as a fusion protein with single-chain antibodies. Our system is a filamentous (M13) phage display system that enables the efficient isolation and

168   © 2001 American Chemical Society

characterization of single-chain antibodies and related ligand-binding proteins. Furthermore, the system sets the stage for the utilization of *in vitro* evolutionary approaches for the production of CBD derivatives having novel binding and elution properties. We further describe the incorporation of the CBD into a recombinant protein-expression system for the production of target proteins by an efficient cellulose-assisted refolding method. In this procedure, denatured proteins are immobilized onto cellulose before being refolded, thus circumventing problems of their aggregation during refolding - the most formidable problem in such protocols. Taken together, the data presented here provides a streamlined process for the isolation, characterization and large-scale production of recombinant proteins. Similar approaches should be appropriate for future CBD-based technologies, where *C. thermocellum* as well as other CBDs may be exploited to their full potential.

Cellulose Binding Domains

CBDs are found in nature either in free cellulases or in cellulosomes. In their simplest form, free cellulases contain a single catalytic module, that hydrolyzes the β-1,4 glucosidic bond of the cellulose substrate, connected via a flexible linker segment to a CBD. Cellulosomes are multienzyme complexes devoted to the efficient degradation of cellulose and hemicellulose by a highly specialized class of cellulolytic microorganisms (*1-5*). The cellulosome concept was first described in the anaerobic cellulolytic bacterium, *C. thermocellum*, the cellulase system of which was shown to comprise a discrete multifunctional multienzyme complex, which appeared to account for the efficient solubilization of insoluble cellulose by this organism (*1, 3*)

The hydrolytic components of the *C. thermocellum* cellulosome are organized around a large, noncatalytic glycoprotein that acts both as a scaffolding component and a cellulose-binding domain (CBD). This noncatalytic subunit has been termed CipA (for cellulosome integrating protein) or scaffoldin. Catalytic subunits of the cellulosome bear conserved, noncatalytic domains, termed dockerin domains, which bind to complementary modules of CipA, termed cohesin domains. CipA includes nine cohesin domains, a CBD, and its own specialized dockerin domain (*2, 3, 5*).

A common feature of all CBDs is that they lack hydrolytic activity, and their function is confined to mediate the association of the hydrolytic enzymes onto

the cellulose substrate. CBDs have been classified into over a dozen different families, based on size and amino acid sequence homology. Fungal and bacterial CBDs appear to be different, the bacterial form being much larger (*7, 8*). Although CBDs from different families differ in size and sequence, comparisons of recently obtained 3D structures revealed that their cellulose-binding faces are quite similar, suggesting a common binding mechanism. As parts of cellulolytic complexes, CBDs may be found as N-terminal or internal components of scaffoldins. CBDs retain their cellulose-binding properties when fused to heterologous proteins. As such, they appear to be promising candidates for application as affinity tags, where they are genetically fused or chemically linked to other proteins. Indeed, CBDs have been applied as affinity tags for protein purification and for enzyme immobilization (*2, 7, 9-15*). CBDs bind to, and can be eluted from cellulose under relatively mild conditions and specific reagents are usually not required. All the experiments described in this chapter were conducted with standard laboratory equipment and reagents that do not require any special safety measures in handling or disposal.

The cellulase system of *C. thermocellum* contains both cellulosomal and non-cellulosomal (free) cellulases and xylanases. In this chapter, CBD refers to the internal cellulose-binding domain of the *C. thermocellum* scaffoldin (the product of the *cipA* gene). This CBD (a Family-IIIa clostridial CBD) was cloned (*16*), expressed and purified from *E. coli* (*17*), and its three-dimensional structure has been solved at high-resolution (*18*). Although this CBD is efficiently expressed in *E. coli*, and is easily purified from overexpressing cultures, it has not yet been applied in recombinant systems for the production of CBD-containing fusion proteins. In this chapter we describe recent applications where the *C. thermocellum* CBD has been applied. In the first part, we describe phage display of the CBD and, in the second part, an efficient protein refolding system based on the unique properties of this particular CBD. Two aspects related to CBD are presented: the application of the CBD as a tool for antibody engineering and approaches for the engineering of the CBD itself that provide the basis for future versatile biotechnological applications.

Phage Display of the *C. thermocellum* CBD

In the last decade, the display of ligands such as peptides and proteins on the surface of bacteriophage has revolutionized the field of ligand isolation. This approach allows for repertoire selection, whereby ligands of a desirable property are isolated from vast collections of variants, based on their unique binding properties. The power of selection makes the isolation of rare variants a feasible task, where rational design, followed by screening of designed variants, is limited in the number of variants that can be tested (*19-23*). The peptides or proteins are expressed on the surface of the phage particle as fusion proteins

with a phage coat protein. As such, they can be isolated along with their respective coding genes by affinity selection on an immobilized target. Recently, a small (36-residue) fungal CBD has been displayed on filamentous phage as a pIII fusion, where it was used as a scaffold for the engineering of novel binding properties of knottins (*24*). In that work, mutagenesis followed by phage-selection was used to replace the cellulose-binding ability of the *Trichoderma reesei* CBD with the ability to specifically bind the enzyme alkaline phosphatase. Since CBDs are the focus of intense studies aimed at the isolation of variants with altered binding properties, the full potential of phage display for the isolation of such CBD variants has yet to be exploited.

We have recently developed a phage-display system, wherein single-chain antibodies (scFvs) are being displayed on the pIII minor coat protein of the filamentous phage M13 (23) while expressed as in-frame fusion proteins with a modified *C. thermocellum* CBD (*25*). This system serves two purposes. First, the CBD is utilized as an affinity tag, thus allowing rapid phage capture and concentration from crude culture supernatants. The CBD also serves as an antigen for immunological detection of displaying phages. Second, this novel phage display system allows the isolation of CBD variants with altered cellulose-binding properties. In addition, the system can be used for the identification of CBD variants where properties unrelated to cellulose binding can be addressed, such as expression levels and secretion efficiency in bacterial expression systems.

Antibody phage display is now considered as an alternative to hybridoma technology (*26*), traditionally used for the production of monoclonal antibodies. It is now possible to mimic the key features of the *in vivo* antibody production and antigen-driven affinity maturation processes, by expressing antibody gene repertoires on the surface of bacteriophage. As a result, high affinity antibodies can be made without prior immunization, and their binding properties can be further manipulated *in vitro* (*27-31*). Seemingly a technology of such high potential, the general application of the antibody-phage display technology is still limited, owing to several technical difficulties. One of the obstacles to the general application of the technology is the accumulation of "insert loss" clones in the library, i.e., phagemid clones from which the cloned scFv DNA-encoding portion has been partially or completely deleted. *E. coli* cells carrying such clones (thus, not expressing a scFv) have a growth-rate advantage over the scFv-expressing clones. During growth, the deleted clones gradually dominate the population, sometimes rendering the library useless. Another difficulty arises when phage technology is applied for cloning hybridoma-derived antibody genes for phage display. The conventional approach employs RT-PCR, where oligonucleotide primers are used to amplify variable regions of the antibody. However, hybridoma cell lines contain significant levels of an aberrant PCR-amplifiable kappa chain-encoding mRNA. These mRNAs contain a 4-bp deletion (*32, 33*), and so, in addition to having irrelevant binding specificity, they can contain premature stop codons. Expressing an affinity tag at the 3' end

of the scFv and its use for phage capture should eliminate approximately two thirds of frame-shifted clones and all of the prematurely terminated clones. This should allow the direct isolation of authentic hybridoma-derived scFvs. Moreover, such a tag would eliminate most of the "insert-loss" clones described above, as about 70% of them would also be frame-shifted and would thus fail to encode the in-frame affinity tag.

Most of the vectors used for antibody-phage display code for a small peptide tag, linked to the 3' end of the scFv, allowing for its detection by various immunological assays. However, these tags are not applied for the actual capture and purification of antibody displaying phage particles. Phage particles are usually partially purified from culture supernatants of infected *E. coli* by polyethylene glycol (PEG) precipitation, a time consuming, and inefficient process. Moreover, for display we use the common type of an antibody phage display system, which is the 3+3 format. In that system, the antibody-pIII coding sequence is carried on a phagemid and phage structural proteins are provided by superinfection of phagemid-bearing *E. coli* with a helper phage (*23, 31, 34, 35*) (phagemid "rescue"). This results in the dominance of "empty" (non-displaying) particles in the population, since the wild type (WT) minor coat protein provided by the helper phage is either more efficiently produced, or better incorporated into the virions. Incorporation of an affinity tag that co-expresses with the displayed antibody should therefore have the advantage of enabling selective concentration of displaying phage particles only.

Our antibody phage-display system was designed to address these issues. When the scFvs are expressed as in-frame single-chain fusions with the CBD, clones that contain deletions or premature stop codons do not display the CBD on the phage surface. As a result, such clones are efficiently eliminated when the entire phage population is concentrated by binding to and elution from a cellulose matrix. The outline of our method is schematically shown in Figure 1. We have successfully applied our method for the construction of several libraries and the subsequent isolation of scFvs (*25*), as well as hybridoma-derived antibodies (*36*).

When the choice of CBD as a candidate for display was made, some factors that may contribute to successful display had to be considered. Although many proteins and peptides have been displayed on phage, attempts at protein display on phage are sometimes unsuccessful for reasons that are not completely understood. Since filamentous phage assembly takes place in the bacterial periplasm, to be displayed, a candidate protein should be suitable for expression and secretion in *E. coli*, and diverse proteins make varying demands on the bacterial export apparatus (*37*). Many proteins are aggregated upon expression in *E. coli* and thus cannot be secreted. The addition of a bacterial signal peptide coding sequence at the 5' end of the protein coding sequence does not guarantee its secretion. In some cases, the protein is secreted but then aggregates in the bacterial periplasm, and as such cannot be assembled into phage particles. Finally, a careful choice of the display format is necessary, since several phage
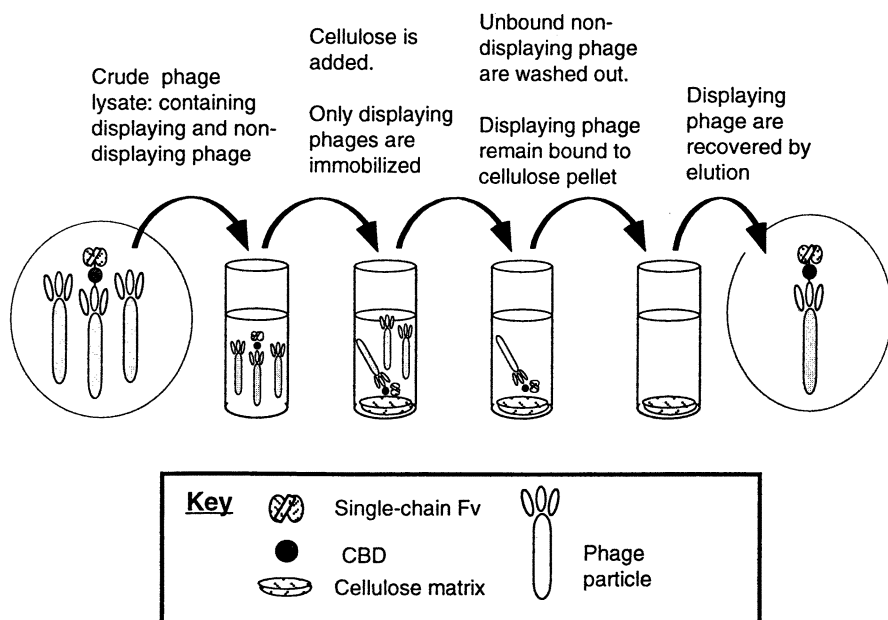
Figure 1. Schematic representation of the phage system used to display scFv-CBD fusion proteins.

systems are available. Some are more suitable for the display of short peptides and would function poorly or not at all in the display of larger proteins (*23, 31, 38-43*). In fact, an *ab initio* approach is necessary to test the suitability of a particular protein for phage display. For example, although the CBD is expressed at high levels in *E. coli* systems (*17*), we found that it is poorly secreted. An encouraging finding was the report of Francisco and co workers (*44, 45*), who successfully displayed the CBD of *Cellulomonas fimi* (a Family-II CBD) on the surface of *E. coli* cells. Indeed, our attempt to display the *C. thermocellum* CBD was successful, as evident from the binding of displaying phage to cellulose. However, we found that about 1% of the phage particles in the population actually display CBD, as opposed to about 10% of the particles displaying scFvs using a similar display format (*25*). The low level of displayed scFv-CBD indicated that the protein is either poorly secreted or poorly incorporated into phage particles.

To improve display efficiency of the CBD, we opted for an *in vitro* evolutionary approach made possible by the phage display method. CBD coding DNA was subjected to error-prone PCR. The mutagenic conditions we used ensure similar levels of substitution frequencies, including AT⇒GC and GC⇒AT transitions as well as AT⇒TA transversions (*46*). The mutated CBD-coding DNA fragments were re-cloned into the phage display vector, and displaying phage were selected on cellulose. This way, mutants that lost the ability to bind cellulose were eliminated. The cellulose-binding clones were then tested for display efficiency. Since our phage display a scFv-CBD fusion, display efficiency can be estimated by comparing the intensity of ELISA signals, wherein binding of equal numbers of phage to the scFv's cognate antigen is tested. Several of the clones showed improved display, and we chose two for further analysis. Their improved display is shown in Figure 2A. DNA sequencing was employed to identify the mutations that are responsible for improved display of CBD. One mutation was G⇒A that changed a **GGG** glycine codon to an **AGG** arginine codon at position 76 of CBD. The second mutation was T⇒C that changed a **TTT** phenylalanine codon to a **CTT** leucine codon at position 78 of CBD. These positions are distant from the cellulose-binding face of CBD, as can be seen in Figure 3. Indeed, neither mutation had an effect on the cellulose-binding properties of the resultant displaying phage. In investigating the cause of the improved display, we compared the secretion efficiency of the WT or mutated scFv-CBD fusion proteins. An ELISA was also used in this experiment, where the signal intensity is proportional to the scFv-CBD content of the culture medium. Both mutants were found to be more efficiently secreted than the WT (Figure 2B), suggesting that WT CBD display is inefficient because it is poorly secreted. Phage displaying the mutated CBDs, reach the 10% efficiency of scFv display alone. We are now applying the same basic approach to isolate CBD variants with altered elution properties (pH, buffer composition, competitive elution, etc.).

A. Phage (starting at 10⁹ cfu/well)   B. Culture medium
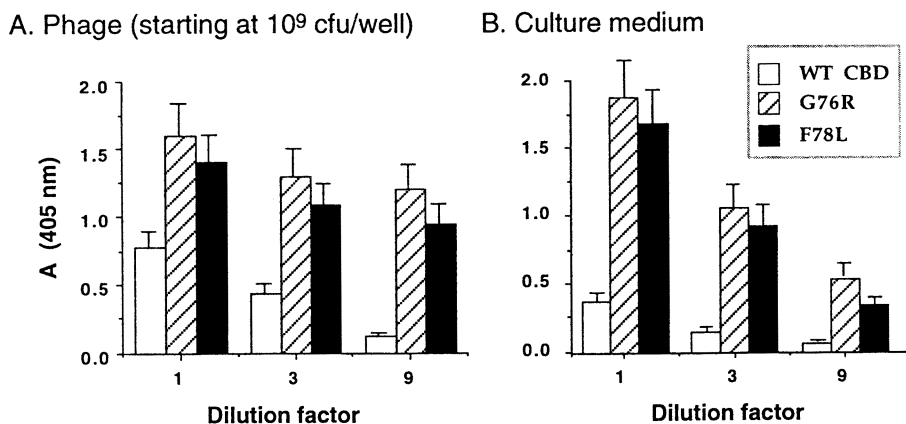
*Figure 2. Antigen-binding ELISA of: A. Phage displaying an
anti-β-galactosidase scFv fused to WT or mutated CBD. B. scFv-CBD present in
the culture medium of cells carrying the same phagemid vectors.
β-Galactosidase binding by tested fractions was detected with an anti-CBD
polyclonal rabbit antiserum followed by HRP/anti-rabbit conjugate and
development with a chromogenic substrate. Error bars represent the standard
deviation of the data.*

Binding Properties of CBD Mutants

The suitability of our phage display system for the analysis of the cellulose-binding properties of CBD mutants was also investigated. In these experiments, rational design was applied to identify CBD residues that may be involved in cellulose binding. The design was based on our solution of the 3D structure of CBD (*18*), and the proposed mechanism by which CBDs bind cellulosics (*3, 8, 47*). The 3D structure of the CBD reveals that it forms a 9-stranded ß-barrel with a calcium-binding site. Conserved surface-exposed residues map to two defined surfaces of the molecule. A plane of aromatic and polar residues, which are proposed to directly interact with cellulose, dominates one surface. The model suggests that the dominant interaction with cellulose involves a linear planar strip of residues on the surface of the CBD molecule that aligns with the glucose rings along one of the chains of the cellulose substrate. This planar strip (Figure 3) consists of three aromatic residues (Trp118, His57, Tyr67), and a salt-bridge formed by an aspartate (Asp56) and an arginine residue (Arg112). Two neighboring groups of polar residues are proposed to anchor the CBD molecule to the two successive chains of crystalline cellulose (*18*). One of these groups includes Asn10, Asn16 and Gln110; the other consists of Ser12 and Ser133. Thus, this surface of the CBD is considered to interact with three adjacent chains of the cellulose substrate.

   To assess the contribution of the above-mentioned residues to cellulose binding, two analyses were performed. First, we mutated these residues into alanine, either individually or in combination. CBD derivatives carrying these mutations were produced in *E. coli*, and their cellulose-binding capacity and dissociation constants were analyzed with microcrystalline cellulose (Avicel). Mutating the "anchoring" residues had little effect on the relative cellulose-binding affinity, except when the three central residues were mutated simultaneously (Table I). On the other hand, single mutations of any of the planar strip residues significantly decreased the affinity for cellulose; mutation of two residues simultaneously reduced the affinity by approximately 20-25 fold. These results suggested that individual anchoring residues have a less significant contribution to cellulose binding than do the individual planar strip residues. This is in agreement with other studies where replacing or chemically modifying residues located at the planar strip resulted in dramatic changes in the cellulose-binding properties of CBDs (*48-53*).

   We then analyzed these mutations in our scFv-CBD phage display system. The WT CBD, cloned into our vector, was replaced with mutation-carrying DNA fragments corresponding to the mutated CBDs listed in Table I. Interestingly, several of the mutants failed to be displayed in our system. The mutants that were successfully displayed were analyzed by binding to and elution from cellulose under various conditions. The results of the analysis are qualitative in nature, and can be interpreted on a scale demonstrating the relative binding strength in comparison to WT CBD. In these experiments, equal
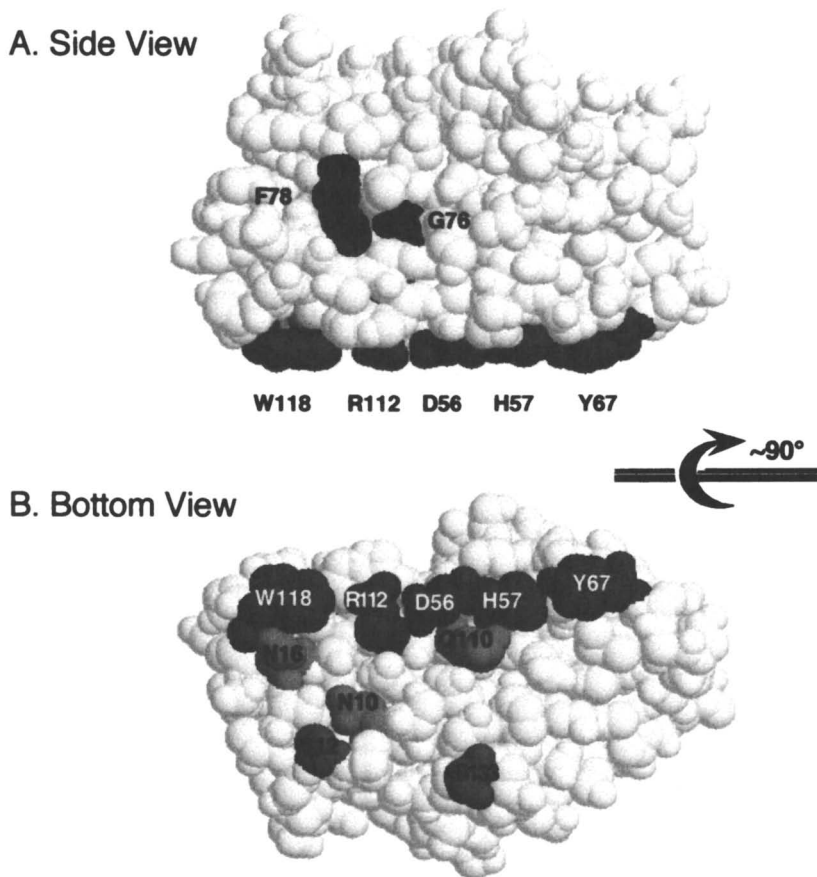
Figure 3. Structural model of the C. thermocellum CBD, based on PDB accession no. 1NBC (18). Residues suggested as being involved in cellulose binding are marked: A. Side view of the molecule, showing Gly76 and Phe78. B. Bottom view (~90° rotation of A along the horizontal axis) showing the planar strip residues: W118, R112, D56, H57, and Y67. Anchoring residues, N16, N10, Q110, S12 and S133.

Table I. Binding capacity and dissociation constants of mutated CBDs

	Mutant	Pc_{max} [mg/g Avicel]	k_d [µM]
	CBD-W.T.	17	1.2
A	N10A	14	1.0
	N16A	15	1.2
	Q110A	19	1.6
	N10A-N16A-Q110A	18	10.4
	S12A	21	1.8
	S133A	17	1.8
	S12A-S133A	23	2.7
B	D56A	21	3.4
	H57A	15	5.0
	Y67A	14	4.0
	R112A	15	14.0
	W118A	13	13.0
	D56A-H57A	15	32.0
	D56A-W118A	15	26.0

A: "Anchoring" residues; **B**: Planar strip residues.
Binding capacity and dissociation constants were calculated as follows: Purified mutated CBDs (30-800µg) were added to microcrystalline cellulose (Avicel). The assay tubes were mixed at 21°C for 1h. After centrifugation, the amount of protein that remained in the supernatant (Free CBD) was determined at A_{280}. The amount of adsorbed CBD was calculated by subtracting the amount of free CBD from the total added to the assay tube.

numbers of phage were adsorbed to cellulose, washed and eluted with a high pH buffer. The phage fractions were tested for antigen binding in an ELISA. As shown in Figure 4, weaker binding is indicated by the presence of antigen-binding phage in the cellulose-unbound fraction, and by more efficient elution from cellulose (a stronger ELISA signal in the eluate fraction). As shown in Table II, the binding strength of the mutant CBD-displaying phage correlated well with the relative binding affinities of the corresponding CBD mutants. Mutation of anchoring residues of the two scFv-CBD-displaying phage had no apparent effect on their cellulose-binding properties. Most of the anchoring-residue mutants, however, failed to be displayed, and the two that were displayed, did so poorly. In contrast to the anchoring residues, mutations located at the planar strip had a more profound effect on binding (Figure 4), and their effects on binding correlated with the dissociation constants of the original corresponding mutants. Most of the planar strip mutants were successfully displayed, with some being more efficiently displayed than WT CBD is. The results from the binding analysis of phage-displayed CBDs laid the groundwork for ongoing research, which is focused on using *in vitro* evolutionary approaches for the isolation of mutated CBD variants having altered binding properties.

Cellulose-Assisted Refolding of ScFv-CBD Fusion Proteins

Many recombinant proteins, particularly proteins of diagnostic and therapeutic potential (such as antibodies, lymphokines, receptors, enzymes and enzyme-inhibitors) have been produced from transformed host cells containing recombinant DNA. Often, the heterologous protein produced by the host cells precipitates within the cell to form refractile (inclusion) bodies. For isolation in the native (biologically active) state, the protein should be separated from cell debris, solubilized with a suitable chaotropic agent (such as urea or guanidinium hydrochloride) and then refolded by gradual removal of the denaturant (*54*). When the native state of the protein is dependent on the formation of intramolecular disulfide bonds, the protein is reduced while in the denatured state by addition of reducing agents (*55*). The formation of disulfide bonds of the protein during the refolding process is initiated by the addition of oxidizing agents in the refolding buffer either during or subsequent to removal of denaturant from the denatured protein.

Denaturation followed by refolding of recombinant proteins from inclusion bodies is frequently the only means by which an *E. coli*-expressed recombinant protein may be recovered in its native form. However, refolding is an empirical process that has to be optimized for each particular protein of interest (*54*). Under conventional folding conditions, the yield of renatured protein is often exceedingly low, due to side reactions such as aggregation or formation of thermodynamically stable, but non-native, folding intermediates. Furthermore,
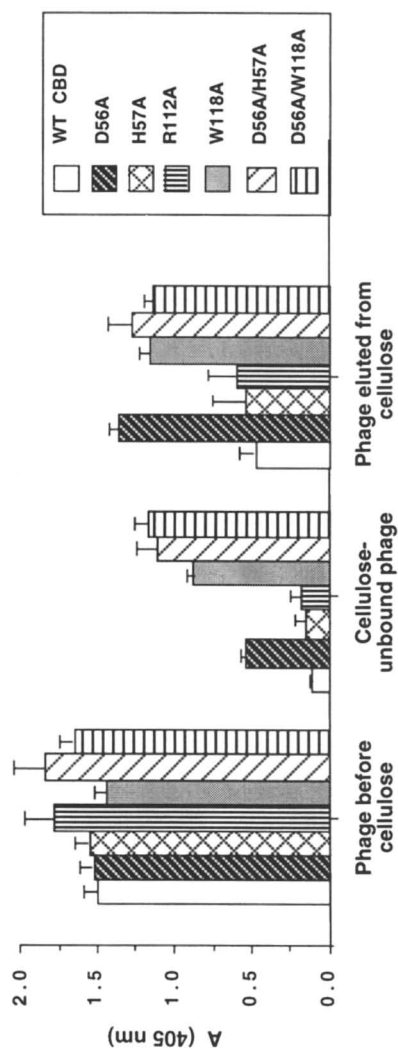
Figure 4. Antigen-binding ELISA of phage displaying an anti-β-galactosidase scFv fused to WT or mutated CBD derivatives (planar strip mutations). Phage (10^9 cfu) were first bound to microcrystalline cellulose, and the bound phage were separated from unbound phage by centrifugation. Cellulose pellets were washed with phosphate buffered saline, and bound phage were released (eluted) with 20 mM NaOH/100 mM NaCl and immediately neutralized. All fractions were maintained at an equal volume. β-Galactosidase binding by the phage in tested fractions was detected with an anti-CBD polyclonal rabbit antiserum followed by HRP/anti rabbit conjugate and development with a chromogenic substrate. Error bars represent the standard deviation of the data.

refolding should be performed at extremely low protein concentrations, owing to the kinetic competition between folding and aggregation processes (56). Aggregation during refolding may be prevented by immobilization of the denatured protein onto a solid phase, prior to initiation of refolding. Some proteins may be charged under refolding conditions, so that they may be immobilized onto ion-exchange matrices in a low-ionic strength denaturant. Refolding is then induced by gradual removal of denaturant, and increasing the salt concentration (57) facilitates release from the matrix. This approach, while suitable for charged proteins, is not adequate for proteins for which the net charge under pH conditions that are compatible with refolding, is too small to allow efficient immobilization onto an ion-exchange matrix. Furthermore, the protein may be immobilized by multiple interactions with the matrix, which may hinder its ability to fold properly. When immobilization of the denatured protein by itself is not an option, it is possible to engineer an affinity tag, linked to either the *N*- or *C*-terminus of the protein. Small affinity tags, such as a hexa-histidine or hexa-arginine tag, have been applied for such purposes. Histidine-tagged proteins may be immobilized and refolded on a metal-carrying support (58), while arginine-tagged proteins may be immobilized and refolded on a support carrying polyanionic groups (59). One must assume that the folding of the immobilized protein is not perturbed by its proximity to the matrix and that inclusion of a short peptide tag at its terminus does not divert the tagged protein into a non productive folding pathway.

Refolding of an immobilized protein with fewer physical constraints may be achieved by its immobilization through a polypeptide fusion partner. In folding of fusion proteins, when a flexible peptide linker joins the fusion partners, the fusion partners fold independently of each other (60). However, most known protein tags (such as glutathione-S-transferase, maltose-binding protein, staphylococcal protein A, and most bacterial and fungal cellulose and chitin-binding domains) will not be useful for the purpose of immobilization of a denatured fusion protein onto an affinity support. This because the protein affinity tags themselves are not necessarily functional under denaturing conditions. A possible solution would be to use as an affinity tag, a protein whose compact and stable structure retards its unfolding under conditions where its fusion partner is completely denatured. Such proteins are rare, and one example, as presented here, is the *C. thermocellum* CBD. A unique property of this particular CBD is that it maintains its specific cellulose-binding properties under conditions where most proteins (including other CBDs) are denatured and nonfunctional. We found that the CBD, alone, or when fused to single-chain antibodies, will bind reversibly to cellulose in up to 6 M urea. This property makes the CBD an ideal candidate for an affinity tag to be applied for matrix-assisted (in this case, cellulose-assisted) refolding of fusion proteins. When a protein-CBD fusion is denatured and allowed to bind cellulose, it is immobilized onto the cellulose matrix while the protein fused to CBD is still in the unfolded state. Gradual removal of the denaturant then facilitates the refolding of the protein into the native state. Finally, the refolded protein can be

separated from the cellulose matrix by applying appropriate elution conditions. When necessary, the protein can be separated from the CBD by proteolytic cleavage and chromatographic or cellulose-based separation. In addition, after being refolded while bound to the cellulose matrix, the protein-CBD fusion may by recovered by digestion of the cellulose matrix by a cellulolytic enzyme (61). The matrix-assisted refolding scheme is shown in Figure 5.

We have cloned several scFvs into expression vectors where the C-terminus of the scFv was fused through a short tether comprised of Ala-Ala-Ala-Gly-Gly and the N-terminus of the CBD. Next, we introduced by PCR the sequence encoding for Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys-Leu between the short tether and the N-terminus of the CBD cassette. This sequence comprises a FLAG epitope (62)), followed by an enterokinase cleavage site. A complete expression cassette was then assembled and cloned into a T7 promoter controlled pET vector for intracellular expression. A scheme of the constructed expression vector is shown in Figure 6. For bacterial expression, the plasmids were introduced into *E. coli* BL21(DE3) cells carrying the T7 RNA polymerase gene on its chromosome under the control of an IPTG-inducible *lac*UV5 promoter. Upon expression, the recombinant scFv-CBD fusion proteins accumulated as insoluble inclusion bodies. The inclusion bodies were solubilized in a 6 M urea-containing buffer and the urea-soluble proteins were mixed with a slurry of crystalline cellulose. The suspension was stirred for 1 h at room temperature to allow for binding to cellulose, and then transferred into dialysis tubes and dialyzed for 24-72 hours to enable refolding. Subsequently, the cellulose composite was recovered, washed to remove contaminants, and the cellulose-bound scFv-CBD was then eluted with a high pH buffer (100 mM NaOH/100 mM NaCl). The eluted proteins were neutralized and stored as frozen fractions. These harsh elution conditions were tolerated by the scFvs we produced. However, other proteins may be less tolerant to such conditions. To make the system more "user friendly", we incorporated the CBD D56A-W118A double mutant described above into the expression vector shown in Figure 6. As shown in Tables I and II, this mutation did not affect the binding capacity of the CBD mutant, but significantly decreased its cellulose-binding affinity. Indeed, by incorporation of this mutation into our expression vectors we can now elute scFv-CBD fusion proteins using milder conditions (10 mM NaOH or 1% triethylamine). Such conditions are acceptable for numerous protocols for the elution of proteins from affinity columns. We are now in pursuit of additional CBD mutants with varying elution profiles (pH, ionic strength, competitive elution with oligosaccharides etc.). Such mutants will not only broaden the applicability of our system for the recovery of a multitude of fusion partners, but will also serve as tools for efficient exploitation of CBDs in other avenues of biotechnology (2).

The cellulose-assisted refolding protocol is simple and based on cheap, non-hazardous materials. Scale-up should be feasible with little difficulty, and we have prepared batches of >50 mg purified proteins from 1 liter shake-flask cultures of several different scFvs. In all cases, the yields were at least 3 fold better, and in most cases 10 fold better than those obtained by standard refolding

Overview of cellulose-assisted refolding protocol

1. IPTG-induced *E. coli* : Preparation of inclusion bodies

2. Solubilization (reduction), 8 M urea

3. Dilution to 4-6 M urea and binding to cellulose

4. Chaotroph removal and dialysis (pH 9.5, redox)

5. Recovery of refolded protein:

Alkaline pH
elution + dialysis

Specific proteolytic
cleavage *in cellulo*

Cellulase-induced
matrix digestion

(Additional polishing steps)
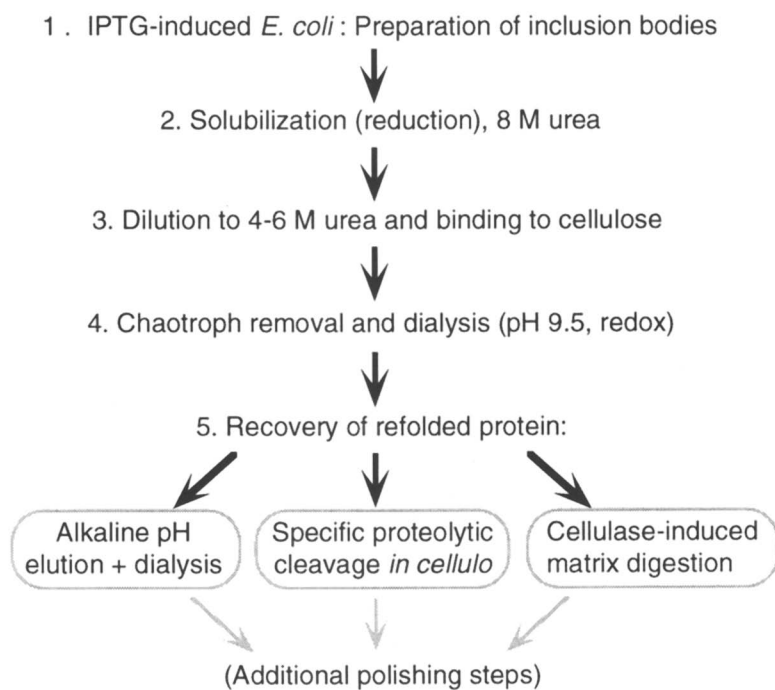
Figure 5. Schematic representation of the cellulose-assisted refolding mothod.
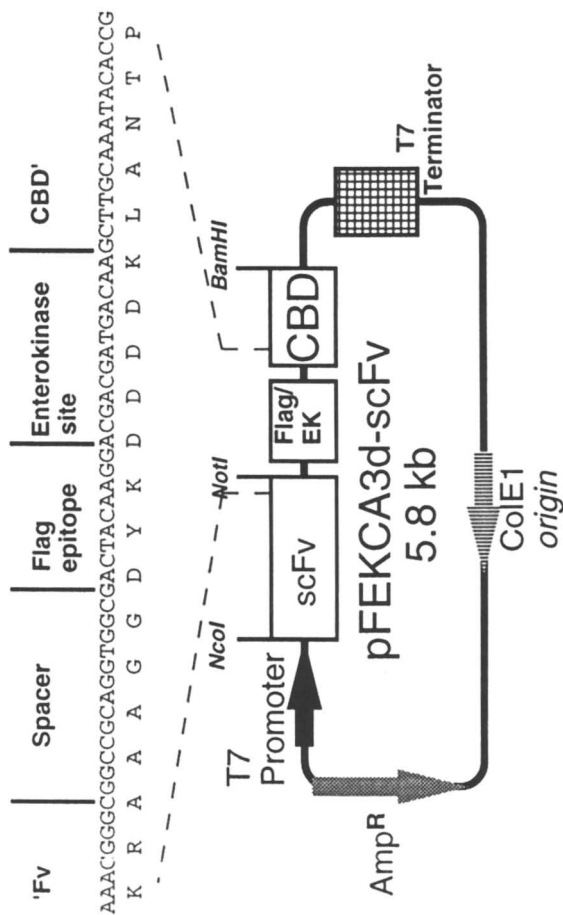
Figure 6. The pasmid vector pFEKCA3d-scFv used for expression of scFv-CBD fusion proteins. The nucleotide and amino acid sequence of the scFv-Flag-enterokinase-CBD region is shown at the top.

Table II. Binding of mutant CBD-displaying phage to cellulose

	Mutant	Relative affinity[1]	Binding of phage and display efficiency[2]	
	CBD-W.T.	1.00	++++	
A	N10A	0.83	++++	↓
	N16A	1.00	ND	
	Q110A	0.75	++++	↓
	N10A-N16A-Q110A	0.12	ND	
	S12A	0.67	ND	
	S133A	0.67	ND	
	S12A-S133A	0.44	ND	
B	D56A	0.35	+++	
	H57A	0.24	+++	
	Y67A	0.30	ND	↑
	R112A	0.09	++	↑
	W118A	0.09	++	
	D56A-H57A	0.04	+	↑
	D56A-W118A	0.05	+	

A: "Anchoring" residues; **B**: Planar strip residues.
[1]: Binding of mutated CBD proteins. Results calculated based on relative binding constants taken from Table I.
[2]: Binding of mutated CBD displaying phage. Results taken from Figure 4.
ND: Not Displayed; CBD mutant that could not be displayed on phage.
↓   Displayed less efficiently than W.T. CBD
↑   Displayed better than W.T. CBD.

protocols. The specific binding activities were similar to those of the same scFv-CBD fusions that have been isolated from bacterial periplasmic fractions of cells carrying phage-display vectors. In all cases, about ten-fold more scFv-CBD could be isolated from similar culture volumes by using intracellular expression and cellulose-assisted refolding compared to secretion (*63*). Our refolding method is based on the reagent system described by Buchner et al. (*55*) with some modifications. In particular, and in striking contrast to published protocols, we can refold the protein at a high concentration. Most of the published refolding protocols are based on diluting the denatured protein to very low concentrations (100 µg/ml or less), to minimize aggregation. Since aggregation is prevented by immobilizing the denatured protein prior to removal of denaturant, large quantities of protein can be handled using modest reagent volumes. This property is shared by other refolding methods that rely on immobilization of denatured proteins (*57-59*). However, these methods may not be generally suitable for every protein of interest. For example, 8 different scFvs previously isolated in our lab fail to bind ion exchange resins, so the method described by Creighton (*57*) would not be applicable for their production by refolding.

In initial attempts to recover several scFv-CBD fusions by cellulose-assisted refolding, the proteins we recovered were not of sufficient purity. Apparently, the crystalline cellulose we used (Sigmacell 20) is not a highly selective matrix, and binds other contaminating proteins in addition to the CBD fusion. Other available cellulose matrices for the isolation of CBD fusion proteins are commercially available (*12*), some of which may perform better in the removal of contaminating proteins from the inclusion bodies. In our case, we could obtain refolded protein of satisfactory homogeneity only when the inclusion bodies were highly enriched in the scFv-CBD fusion protein. Since there is large batch-to-batch variability in the quality of the inclusion bodies (mostly dependent on the induction level that varies for different scFvs), additional purification steps may be required to recover a highly homogeneous protein preparation. Our solution to that particular problem involved the addition of the His tag linked to the *N*-terminus of the fusion protein. Thus it was possible to purify the protein in the denatured state on a Ni-NTA resin before the refolding step. Metal-chelate affinity chromatography is highly efficient as a single-step purification method (*64, 65*), and in our case little additional purification was necessary after it was applied to enrich the scFv-CBD fraction in the inclusion bodies. When we liberated the scFv from CBD by specific proteolytic digestion at the engineered enterokinase site, the quality of the inclusion bodies was of lesser concern, as the contaminants were not liberated from the cellulose matrix under these conditions. On a larger scale, the protease will have to be removed after digestion is completed. We believe that our method will prove to be of general applicability for the recovery of proteins other than single-chain antibodies. Downstream purification steps can be designed, based on the properties of the specific fusion partner.

Literature Cited

1. Bayer, E. A.; Setter, E.; Lamed, R. *J. Bacteriol.* **1985** *163,* 552-529.
2. Bayer, E. A.; Morag, E.; Lamed, R. *Trends Biotechnol.* **1994** *12,* 379-386.
3. Bayer, E. A.; Shimon, L. J.; Shoham, Y.; Lamed, R. *J. Struct. Biol.* **1998** *124,* 221-234.
4. Doi, R. H.; Park, J. S.; Liu, C. C.; Malburg, L. M.; Tamaru, Y.; Ichiishi, A.; Ibrahim, A. *Extremophiles* **1998** *2,* 53-60.
5. Beguin, P.; Lemaire, M. *Crit. Rev. Biochem. Mol. Biol.* **1996** *31,* 201-236.
6. Lamed, R.; Setter, E.; Bayer, E. A. *J. Bacteriol.* **1983** *156,* 828-836.
7. Ong, E.; Greenwood, J. M.; Gilkes, N. R.; Kilburn, D. G.; Miller , R. C. J.; Warren, R. A. J. *Trends Biotechnol.* **1989** *7,* 239-243.
8. Tomme, P.; Driver, D. P.; Amandoron, E. A.; Miller, R. C., Jr.; Antony, R.; Warren, J.; Kilburn, D. G. *J. Bacteriol.* **1995** *177,* 4356-4363.
9. Ong, E.; Alimonti, J. B.; Greenwood, J. M.; Miller, R. C., Jr.; Warren, R. A.; Kilburn, D. G. *Bioseparation* **1995** *5,* 95-104.
10. Assouline, Z.; Shen, H.; Kilburn, D. G.; Warren, R. A. J. *Protein Eng.* **1993** *6,* 787-792.
11. Le, K. D.; Gilkes, N. R.; Kilburn, D. G.; Miller, R. C., Jr.; Saddler, J. N.; Warren, R. A. *Enzyme. Microb. Technol.* **1994** *16,* 496-500.
12. Shoseyov, O.; Karmely, Y. US patent application No. 460,458. **1995.**
13. Shpigel, E.; Elias, D.; Cohen, I. R.; Shoseyov, O. *Protein. Expr. Purif.* **1998** *14,* 185-191.
14. Sakka, K.; Karita, S.; Kimura, T.; Ohmiya, K. *Ann. NY. Acad. Sci.* **1998** *864,* 485-488.
15. Rechter, M.; Lider, O.; Cahalon, L.; Baharav, E.; Dekel, M.; Seigel, D.; Vlodavsky, I.; Aingorn, H.; Cohen, I. R.; Shoseyov, O. *Biochem. Biophys. Res. Commun.* **1999** *255,* 657-562.
16. Poole, D. M.; Morag, E.; Lamed, R.; Bayer, E. A.; Hazlewood, G. P.; Gilbert, H. J. *FEMS Microbiol. Lett.* **1992** *78,* 181-186.
17. Morag, E.; Lapidot, A.; Govorko, D.; Lamed, R.; Wilchek, M.; Bayer, E. A.; Shoham, Y. *Appl. Environ. Microbiol.* **1995** *61,* 1980-1986.
18. Tormo, J.; Lamed, R.; Chirino, A. J.; Morag, E.; Bayer, E. A.; Shoham, Y.; Steitz, T. A. *Embo J.* **1996** *15,* 5739-5751.
19. Smith, G. P. *Science* (Washington DC) **1985** *228,* 1315-1317.
20. Smith, G. P.; Scott J. K. *Methods Enzymol.* **1993** *217,* 228-257.
21. Scott, J. F.; Smith G. P. *Science* (Washington DC) **1990** *249,* 386-390.
22. Hill, H. R.; Stockley P. G. *Mol. Microbiol.* **1996** *20,* 685-692.
23. Hoogenboom, H. R.; Griffiths, A. D.; Johnson, K. S.; Chiswell, D. J.; Hudson, P.; Winter, G. *Nucleic Acids Res.* **1991** *19,* 4133-4137.
24. Smith, G. P.; Patel, S. U.; Windass, J. D.; Thornton, J. M.; Winter, G.; Griffiths, A. D. *J. Mol. Biol.* **1998** *277,* 317-332.
25. Berdichevsky, Y.; Ben-Zeev, E.; Lamed, R.; Benhar, I. *J. Immunol. Methods* **1999** *228,* 151-162.

26. Köhler, G.; Milstein, C. *Nature* (London) **1975** *256,* 495-497.

27. Parmley, S. F.; Smith, G. P. *Gene* **1988** *73,* 305-318.

28. Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G. *Nature* (London) **1991** *352,* 624-628.

29. Lerner, R. A.; Kang, A. S.; Bain, J. D.; Burton, D. R.; Barbas, C. F., 3rd. *Science* (Washington DC) **1992** *258,* 1313-1314.

30. Griffiths, A. D.; Williams. S. C.; Hartley, O.; Tomlinson, I. M.; Waterhouse, P.; Crosby, W. L.; Kontermann, R. E.; Jones, P.; Low, N. M.; Allison, T.; Prospero, T. D.; Hoogenboom, H. R.; Nissim, A.; Cox, J. P. L.; Harrison, J. L.; Zaccolo, M.; Gherardi, E.; Winter, G. *Embo J.* **1994** *13,* 3245-3260.

31. Hoogenboom, H. R.; de Bruine, A. P.; Hufton, S. E.; Hoet, R. M.; Arends, J. W.; Roovers, R. C. *Immunotechnology* **1998** *4,* 1-20.

32. Carroll, W. L.; Mendel, E.; Levi, S. *Mol. Immunol.* **1988** *25,* 991-995.

33. Duan, L.; Pomeranz, R. J. *Nucleic Acids Res.* **1994** *22,* 5433-5438.

34. Barbas, C. F. 3d; Bain, J. D, Hoekstra D. M.; Lerner, R. A. *Proc. Natl. Acad. Sci.* (USA) **1992** *89,* 4457-4461.

35. de Haard, H.; Henderikx, P.; Hoogenboom, H. R. *Adv. Drug Delivery Rev.* **1998** *31,* 5-31.

36. Frenkel, D.; Solomon, B.; Benhar, I. *J. Neuroimmunology* **2000** *In press.*

37. Pugsley, A. P. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **1993** *57,* 50-108.

38. Makowski, L. *Gene* **1993** *128,* 5-11.

39. Felici, F.; Luzzago, A.; Monaci, P.; Nicosia, A.; Sollazzo, M.; Traboni, C. *Biotechnol. Annu. Rev.* **1995** *1,* 149-183.

40. Nieba, L.; Honegger, A.; Krebber, C.; Plückthun, A. *Protein Eng.* **1997** *10,* 435-444.

41. Bothmann, H.; Plückthun, A. *Nat. Biotechnol.***1998** *16,* 376-380.

42. Wilson, D. R.; Finlay, B. B. *Can. J. Microbiol.* **1998** *44,* 313-29.

43. Rodi, D. J.; Makowski, L. *Curr. Opin. Biotechnol.* **1999** *10,* 87-93.

44. Francisco, J. A.; Stathopoulos, C.; Warren, R. A.; Kilburn, D. G.; Georgiou, G. *Bio/Technology* **1993** *11,* 491-495.

45. Francisco, J. A.; Georgiou, G. *Ann. NY. Acad. Sci.* **1994** 745, 372-382.

46. Fromant, M.; Blanquet, S.; Plateau, P. *Anal. Biochem.* **1995** *224,* 347-353.

47. Bayer, E. A.; Morag, E.; Shoham, Y.; Tormo, J.; Lamed, R. In *Bacterial adhesion: molecular and ecological diversity;* Fletcher, M., Ed., Wiley-Liss, Inc., New York. **1996** 155-182.

48. Linder, M.; Mattinen, M. L.; Kontteli, M.; Lindeberg, G.; Stahlberg, J.; Drakenberg, T.; Reinikainen, T.; Pettersson, G.; Annila, A. *Protein Sci.* **1995** *4,* 1056-1064.

49. Linder, M.; Lindeberg, G.; Reinikainen, T.; Teeri, T. T.; Pettersson, G. *Febs. Lett.* **1995** *372,* 96-98.

50. Linder, M.; Nevanen, T.; Teeri, T. T. *Febs. Lett.* **1999** *447,* 13-16.

51. Macarron, R.; Henrissat, B.; Claeyssens, M. *Biochim. Biophys. Acta* **1995** *1245,* 187-190.

52. Bray, M. R.; Johnson, P. E.; Gilkes, N. R.; McIntosh, L. P.; Kilburn, D. G.; Warren, R. A. *Protein Sci.* **1996** *5,* 2311-2318.

53. Nagy, T.; Simpson, P.; Williamson, M. P.; Hazlewood, G. P.; Gilbert, H. J.; Orosz, L. *Febs. Lett.* **1998** *429,* 312-316.

54. Lilie, H.; Schwarz, E.; Rudolph, R. *Curr. Opin. Biotechnol.* **1998** *9,* 497-501.

55. Buchner, J.; Pastan, I.; Brinkmann, U. *Anal. Biochem.* **1992** *205,* 263-270.

56. Kiefhaber, T.; Rudolph, R.; Kohler, H. H.; Buchner, J. *Biotechnology* (NY) **1991** *9,* 825-829.

57. Creighton, T. E. in *UCLA symposia on molecular and cellular biology new series;* Oxender, D. L., Ed.; **1985** Vol 39 pp. 249-258.

58. Glansbeek H. L.; van Beuningen, H. M.; Vitters, E. L.; van der Kraan, P. M.; van den Berg, W. B. *Protein Expr. Purif.* **1998** *12,* 201-207.

59. Stempfer, G.; Holl-Neugebauer, B.; Rudolph, R. *Nat. Biotechnol.* **1996** *14,* 329-334.

60. Garel, J-R. in *Protein folding;* Creighton, T. E., Ed; W. H. Freeman and Company, New York NY, **1992** pp 405-454.

61. Morag, E.; Bayer, E. A.; Lamed, R. *Enzyme Microb. Technol.* **1992** *14,* 289-292.

62. Brizzard, B. L.; Chubet, R. G.; Vizard, D. L. *Biotechniques* **1994** *16,* 730-735.

63. Berdichevsky, Y.; Lamed, R.; Frenkel, D.; Gophna, U.; Bayer, E.; Yaron, S.; Shoham, Y.; Benhar, I. *Protein Express. Purif.* **1999** *17,* 249-259.

64. Hochuli, E. *Genet. Eng.* (NY) **1990** *12,* 87-98.

65. Crowe, J.; Dobeli, H.; Gentz, R.; Hochuli, E.; Stuber, D.; Henco, K. *Methods Mol. Biol.* **1994** *31,* 371-387.

Chapter 11

Solid-State Enzymes for Fiber Hydrolysis

George Szakács[1], Katalin Urbánszki[1], and Robert P. Tengerdy[2]

[1]Department of Agricultural Chemical Technology,
Technical University of Budapest, 1111 Budapest, Gellért tér 4, Hungary
[2]Colorado State University, Department of Microbiology,
Fort Collins, CO 80523–1677

Lignocellulose degrading enzymes were produced by solid substrate fermentation (SSF) with *Trichoderma* and *Gliocladium* spp. on various agricultural by-products, such as spent brewing grain, corn fiber and wheat straw/wheat bran mixture. In a 10 day SSF, the hypercellulolytic mutant *Trichoderma reesei* Rut C30 produced 31 Filter Paper Unit (FPU)/g dry weight (DW) cellulase activity and 650 International Unit (IU)/g DW xylanase activity on corn fiber (corn seed hull), 55 FPU/g DW cellulase and 2400 IU/g DW xylanase activity on spent brewing grain. In a much faster 3 day fermentation, the wild strain *Trichoderma hamatum* TUB F-105 produced 7 FPU/g DW cellulase activity and 7600 IU/g DW xylanase activity on extracted (starch-removed) corn fiber. *Gliocladium sp.* TUB F-498 secreted 11 FPU/g DW cellulase activity in a 6 day SSF on a wheat straw-wheat bran 9:1 mixture. The hydrolyzing potential of SSF enzymes on native autologous substrates is substrate specific and does not always coincide with their filter paper activity. The F-498 and F-105 enzymes, at 5.0 FPU/g cellulose enzyme load hydrolyzed corn fiber and spent brewing grain more efficiently than the Rut C30 enzyme, although the Rut C30 enzyme hydrolyzed pure cellulose (Solka Floc) faster and more efficiently. The F-498 enzyme required lower enzyme loads for the hydrolysis of corn fiber and spent brewing grain than Rut C30. Due to adequate hydrolyzing potential, much reduced production cost and direct applicability, SSF enzymes are excellent candidates for enzymatic agrobiotechnological processes.

Enzymes are increasingly used in a wide range of agrobiotechnological processes, mainly in enzyme assisted ensiling, in bioprocessing of crops and crop residues, in fiber bioprocessing and in feed supplements. Enzymes are also involved directly or

190 © 2001 American Chemical Society

indirectly in soil fertility improvement through directed composting, bioremediation and post harvest residue decomposition (1-5) (Table I).

The current market value of industrial enzymes is about $ 1.6 billion, with a forecast of $ 3 billion by 2008. Of this market approx. 50 % is used in agrobiotechnology (including food and feed supplements) with an expected annual growth of 15-20 % (3,4,5).

The increased demand for enzymes stimulates improved and more economic enzyme production. The main avenues for improvement are genetic manipulations of microbes, transgenic enzymes in plants, and improvements in fermentation technology. The present work only concerns improvements in fermentation technology.

Enzymes are currently produced mostly by submerged fermentation (SF), a relatively expensive process even with currently used high producing, genetically engineered strains, and even with *on site* production (6,7,8). The price of commercially available enzymes is usually too high for agrobiotechnological applications (c.f. cost analysis in Table II). An alternative technology for enzyme production is solid substrate fermentation (SSF). With recent advances in SSF technology and the availability of substrate and process selected strains, SSF is becoming a viable alternative for enzyme production (8,9).

Although SSF technology has been used traditionally in Oriental countries for enzyme production, and a healthy industry grew up around the Koji process for the food and beverage industries, SSF technology is still not widely used in Western countries. The reasons are high labor intensity, lower yields, and difficulty with scaleup and process controls.

The authors of this paper are convinced that SSF technology is not a general substitute or alternative of SF technology, but a special application for targeted agrobiotechnological processes, where the fermented substrate (usually an agricultural by-product) is the enzyme source. The premise of the present work is that SSF should be used for enzyme production only if the *in situ* or *on site* produced enzyme enriched substrate may be used directly, without downstream processing in an agrobiotechnological process.

A comparison of the yield and productivity of cellulase production by SF vs SSF is presented in Table III. Although the yield and productivity is higher in SF, the unit production cost is much lower in SSF, even in comparison with *on site* SF, due to the much lower fermentation costs. The lower cost of enzymes translates into lower application costs in agrobiotechnological applications, as illustrated in Table IV. One example is bioethanol production by simultaneous saccharification and co-fermentation (SSCF) (21,23), the other is bioprocessing of sweet sorghum by enzyme assisted ensiling (19,20). The much lower application costs may break the economic barrier for enzyme use in many agrobiotechnological processes.

In this paper the production of lignocellulolytic SSF enzymes on various agricultural byproducts is presented together with the application of such enzymes for fiber hydrolysis.

Materials and Methods

Cultures. *Trichoderma hamatum* TUB F-105 (ATCC 62392) was isolated from decaying reed in Hungary by G. Szakács and identified by L. Vajna, Budapest;

Table I. Potential Use of Crude SSF Enzymes in Biotechnological Processes

Process	Enzyme	References
Enzyme assisted ensiling	Fungal cellulases and hemicellulases	19, 20
Bioprocessing of crops and crop residues (e.g., biofuels production)	Fungal cellulases and hemicellulases	21, 22, 23
Fiber processing (e.g., retting)	Fungal pectinases, cellulases and hemicellulases	24, 25
Feed supplement	Amylases, proteases, phytases, lipases, cellulases, hemicellulases	3, 26, 27
Bioprocessing of pulp and paper (biobleaching, biopulping, deinking)	Xylanases, cellulases, ligninases	28, 29, 30
Directed composting	Hydrolytic enzymes	31, 32
Soil bioremediation	Laccases, ligninases	33, 34
Post harvest residue decomposition	*Trichoderma harzianum* cellulase	35
Biopesticide	*Trichoderma harzianum* cellulase and chitinase	36

Table II. Estimated Economy of SF vs SSF for Cellulase Production

1. Reference: Celluclast 1.5 L (Novo-Nordisk A/S, Copenhagen, Denmark)

Bulk price: $ 13/liter; Activity: 8×10^4 FPU/liter

Unit price: $ 13/8 x 10^4 = $ 1.6 x 10^{-4}/FPU

2. Crude SF enzyme:

Average production: 20×10^6 FPU/m^3

Fermentation cost: $ 200/m^3

Unit cost: 200/20 x 10^6 = $ 10 x 10^{-6}/FPU

3. Crude SSF enzyme:

Average production: 50×10^6 FPU/MT

Fermentation cost*: $ 150/MT

Unit cost: 150/50 x 10^6 = $ 3.0 x 10^{-6}/FPU

* Breakdown of SSF costs:

    Feedstock (SBG) $ 75/MT

    Substrate preparation: $ 20/MT

    Additives: $ 5/MT

    Operational costs: $ 50/MT

SF: Submerged Fermentation; SSF: Solid Substrate Fermentation; MT: Metric Ton (dry)
FPU: Filter Paper Unit; SBG: Spent Brewing Grain.

Table III. Comparison of SF vs SSF for Cellulase Production
(*Trichoderma reesei* Rut C30)

SF: Average production level in 5 % slurry is 10-20 FPU/ml which corresponds to a yield of 200-400 FPU/g and a volumetric productivity of 160-200 FPU/l.h
SSF: Average production level is 20-200 FPU/g, which corresponds to 1-10 FPU/ml in a 5 % slurry and a volumetric productivity of 2-20 FPU/l.h

SF: Submerged Fermentation; SSF: Solid Substrate Fermentation;
FPU: Filter Paper Unit.

Table IV. Estimated Enzyme Cost in Biotechnology Applications

1. SSCF for lignocellulose hydrolysis:
Recommended level of cellulase used: 7.0 FPU/g cellulose
Cost with Celluclast 1.5 L (Novo-Nordisk):
7.0×10^3 FPU/kg x $ 1.6 x 10^{-4}/FPU = $ 1.12/kg cellulose
Cost with crude SF enzyme:
7.0×10^3 FPU/kg x $ 10 x 10^{-6}/FPU = $ 0.07/kg cellulose
Cost with crude SSF enzyme:
7.0×10^3 FPU/kg x $ 3.0 x 10^{-6}/FPU = $ 0.02/kg cellulose

2. Bioprocessing of sweet sorghum:
Recommended level of enzymes used:
0.025 % Celluclast 1.5 L[*] at $ 13/liter;   0.25 x 13 = $ 3.25/MT
0.025 % Viscozyme 120 L at $ 27/liter; 0.25 x 27 = $ 6.75/MT
Total treatment cost = $ 10.00/MT
SSF enzyme at 4 kg/MT[**] at $ 0.05/kg[***];   4 x 0.05 = $ 0.20/MT
[***] *in situ* production reduces enzyme costs to $ 50/MT

SSCF: Simultaneous Saccharification and Co-Fermentation; FPU: Filter Paper Unit
SF: Submerged Fermentation; SSF: Solid Substrate Fermentation; MT: Metric Ton (dry)
[*] This represents the addition of 200,000 FPU/MT at 80 FPU/ml activity
[**] This represents an equivalent 200,000 FPU/MT addition at 50 FPU/g activity
[***] *In situ* production eliminates feedstock cost and substrate preparation plus additive costs

Gliocladium sp. TUB F-498 was isolated from soil, Germany; *Trichoderma reesei* Rut C30 was kindly provided by D.Eveleigh, Rutgers University. The strains were routinely maintained and sporulated on potato-dextrose-agar (PDA) medium.

Substrates. Untreated corn fiber (CF) and extracted (starch-removed) corn fiber (corn hull) were gifts from M.K.Dowd, USDA, SRRL, New Orleans. Spent brewing grains (SBG) were received from the New Belgian Brewery, Ft.Collins, Colorado. Wheat straw (WS) and wheat bran were from local sources in Hungary.

Solid Substrate Fermentation (SSF). SSF was carried out in 500-ml cotton plugged Erlenmeyer flasks. Five g dry substrate was supplemented with either tap water or salt solution. The quantity of tap water or salt solution was 10, 15 or 20 ml, depending on the substrate and/or the fungus. The optimal moisture content and presence/absence of salt additives were determined for each substrate-fungus combinations in preliminary SSF experiments (data not shown). The composition of salt solution was: 5 g/l KH_2PO_4; 5 g/l $(NH_4)_2SO_4$; 1 g/l $MgSO_4x7H_2O$; 1 g/l NaCl; 5.0 mg/l $FeSO_4x7H_2O$; 1.6 mg/l $MnSO_4$; 3.45 mg/l $ZnSO_4x7H_2O$; 2.0 mg/l $CoCl_2x6H_2O$. The prepared substrate was sterilized at 121°C for 30 min, then inoculated with a dense spore suspension of the test fungi to a final concentration of 10^7 spores/g DW. The spore suspension was prepared by washing spores from the surface of 10 days old sporulating PDA plates of the respective fungi with 0.1 % Tween-80 containing water. The inoculated flasks were incubated at 25°C, in a stationary state.

Enzymatic hydrolysis with the crude SSF enzymes. Autologous SSF enzyme sources were used in all hydrolysis experiments on native untreated substrates, i.e. enzyme produced on CF was used for CF hydrolysis, SBG enzyme for SBG hydrolysis. Solka Floc was hydrolyzed with enzymes produced on CF. Five g finely ground (< 0.3 mm) dry substrate was suspended in 50 ml 0.05 M citrate buffer, pH 5.2 and pasteurized at 100°C for 15 min. The crude wet SSF enzyme, or Celluclast 1.5 L (Novo-Nordisk) was added to the cooled suspension to a final activity of 5.0 FPU/g cellulose, and the reaction mixture was incubated at 50°C for 12-48 h with 100 rpm shaking. Hydrolysis efficiency is expressed as the hydrolyzed portion, in percent, of total carbohydrates in the substrate.

Analytical Procedures. Enzyme activities were determined from the culture extract of SSF samples: 5 g dry weight (DW) fermented substrate was extracted with 95 ml water, containing 0.1 % Tween-80, by shaking for 60 min at room temperature. From the centrifuged extract, filter paper activity for cellulase was determined by standard IUPAC method (*10*). Beta-1,4-endoglucanase activity was determined according to Bailey and Nevalainen (*11*). Xylanase activity was assayed according to Bailey et al. (*12*). Beta-glucosidase was determined following Kubicek (*13*). Reducing sugar was determined by the dinitrosalicylic acid method (*14*). Detailed sugar analysis was performed by HPLC (Biotronic 2000, Biotronic Wissenschaftliche Gerate GmbH, Frankfurt/Main, Germany). Clarified extracts were injected onto a 250x4 mm Aminex HPX-87 C-type column and eluted with a mixture of

acetonitrile:water = 80:20. Each variable was tested in four reps (four flasks), and each assay was done in duplicate. The means of these tests are shown in the results.

Results and Discussion

1. Lignocellulolytic enzyme production

The time course of enzyme production by *Trichoderma reesei* Rut C30 on different substrates is shown in Figs 1 and 2, and Table V. On SBG both cellulase and xylanase were produced faster than on CF and WS. All four enzymes shown in Table V had peak activities on day 8 on SBG, but on day 10 on WS and day 12 on CF. The substrate specificity for different enzymes is evident; xylanase is produced more efficiently on WS, while more cellulase, endoglucanase and beta-glucosidase is produced on SBG. The kinetics of the production of these four enzymes by Rut C30 on SBG is shown in Fig.3. The comparison of enzyme production by the reference hypercellulolytic mutant *Trichoderma reesei* Rut C30 and wild strains of a *Trichoderma* and *Gliocladium sp.* is shown in Table VI. While Rut C30 is the best producer of all enzymes tested, the wild strains have a much shorter fermentation time to reach maximum enzyme production. Other authors have found a similar range of cellulase production in SSF on various substrates and by various fungi (*15,16,17*)

2. Hydrolysis efficiency of SSF enzymes

The hydrolysis efficiency of SSF enzymes appears to be substrate specific and does not always coincide with the filter paper activity of the enzyme complex. The TUB F-498 and F-105 enzymes produced on corn fiber are more efficient for corn fiber hydrolysis than the Rut C30 enzyme, although the Rut C30 enzyme is more efficient for the hydrolysis of pure cellulose (Solka Floc)(Table VII).
The distribution of reducing sugars generated during hydrolysis is also different. The Rut C30 enzyme complex produces more arabinose and xylose but less glucose on corn fiber and spent brewing grain than the F-498 enzyme complex (Table VIII).
The dependence of lignocellulose hydrolysis on enzyme load by different SSF enzymes on different substrates is shown in Fig.4. The F-498 and F-105 enzymes are much more efficient at low enzyme loads on corn fiber, but not on Solka Floc. The Rut C30 enzyme is less efficient on corn fiber, but very efficient on spent brewing grain and Solka Floc, even at low enzyme loads. The Celluclast 1.5 L is less efficient on both SBG and CF, but good on Solka Floc.
These results demonstrate the substrate specificity of natural unprocessed enzyme complexes, such as the crude SSF enzymes. It is important, therefore, that in screening for efficient lignocellulolytic enzymes, selection should be made for substrate specificity.
An important warning is that the generally accepted filter paper activity is only a crude approximation of the actual hydrolyzing potential of an enzyme complex. A comparison of several commercial cellulase preparations supports this statement
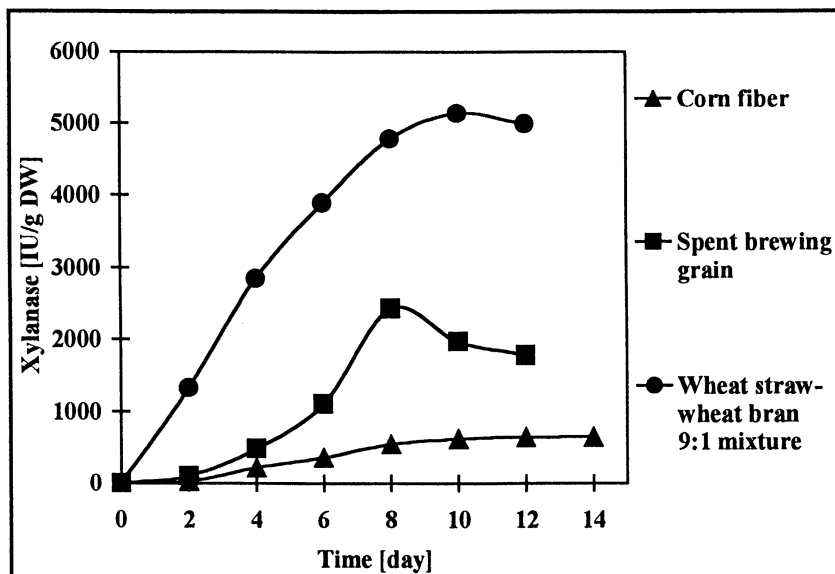
Figure 1. Time course of xylanase production on different lignocellulose containing materials in solid substrate fermentation with Trichoderma reesei Rut C30
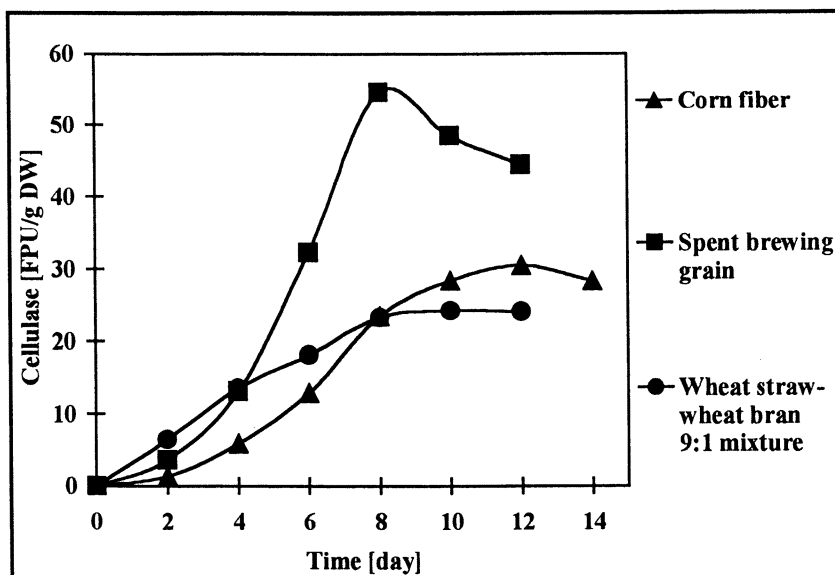
Figure 2. Time course of cellulase production on different lignocellulose containing materials in solid substrate fermentation with Trichoderma reesei Rut C30

Table V. Enzyme Production on Different Lignocellulose Containing Materials by *Trichoderma reesei* Rut C30 in Solid Substrate Fermentation[a]

Substrate	Cellulase [FPU/g DW]	Xylanase [IU/g DW]	β-glucosidase [IU/g DW]	1,4-β-endoglucanase [EGU/g DW]
Corn fiber	30.3	650	12.3	8713
Spent brewing grain	54.6	2415	23.2	18565
Wheat straw wheat bran 9:1 mixture	24.0	5133	6.8	6767

[a]Fermentation time: 12 days on corn fiber, 8 days on spent brewing grain, 10 days on wheat straw-wheat bran 9:1 mixture. Medium composition: 5g substrate + 10 ml salt solution (spent brewing grain) or 20 ml salt solution (wheat straw-wheat bran 9:1 mixture) or 15 ml tap-water (corn fiber). Initial pH= 4.5 (corn fiber) or 6.0 (spent brewing grain and wheat straw-wheat bran 9:1 mixture).

Table VI. Enzyme Production on Spent Brewing Grain by *Trichoderma reesei* Rut C30, *Trichoderma hamatum* TUB F-105 and *Gliocladium sp.* TUB F-498 in Solid Substrate Fermentation[a]

Strain	Cellulase [FPU/g DW]	Xylanase [IU/g DW]	β-glucosidase [IU/g DW]	1,4-β-endoglucanase [EGU/g DW]
Rut C30	54.6	2415	23.2	18565
TUB F-498	5.6	546	13.4	1425
TUB F-105	4.4	673	8.7	1183

[a]Fermentation time: Rut C30 8 days, TUB F-498 4 days, TUB F-105 4 days. Medium composition: 5g spent brewing grain + 10 ml salt solution (Rut C30) or 15 ml salt solution (TUB F-105, TUB F-498). Initial pH= 6.0 (Rut C30) or 5.0 (TUB F-105, TUB F-498). NH_4NO_3 was substituted for $(NH_4)_2SO_4$ in the salt solution for strain TUB F-105

Table VII. Efficiency of Enzymatic Hydrolysis of Cellulose Containing Materials and Solka Floc SW200 Cellulose Pulp with Celluclast 1.5L and Crude Autologous Solid Substrate Fermentation (SSF) Enzymes[a]

| Enzyme | Hydrolysis Efficiency[b] | | | | | |
| | Solka Floc SW200 | | Corn Fiber | | Spent Brewing Grain | |
	12 h	48 h	12 h	48 h	12 h	48 h
Celluclast 1.5L	26.4	47.1	4.5	5.6	12.1	16.1
Rut C30 SSF enzyme	30.6	49.9	8.6	12.5	11.6	14.3
TUB F-498 SSF enzyme	16.9	24.4	24.9	36.1	25.3	30.1
TUB F-105 SSF enzyme	19.7	22.7	26.6	30.6	23.9	25.2

[a]The substrates were hydrolyzed at 10 % dry matter content, pH= 5.2, 50 °C for 12 and 48 hour at 5.0 FPU/g cellulose enzyme load with 100 rpm shaking.
[b]Percent of total hydrolyzable carbohydrates

Table VIII. Distribution of Reducing Sugars in Corn Fiber and Spent Brewing Grain Hydrolysates Prepared with *Gliocladium sp.* TUB F-498 and *Trichoderma reesei* Rut C30 Autologous SSF Enzymes[a]

| Reducing sugar composition | Corn Fiber[b] | | Spent Brewing Grain[b] | |
	Rut C30	TUB F-498	Rut C30	TUB F-498
Fructose	0.7	0	0	0
Maltose	0	0	1.6	7.5
Arabinose	17.4	5.6	12.9	11.3
Xylose	10.0	1.0	24.2	11.9
Glucose	63.4	85.8	49.5	59.0
Non-identified	8.5	7.6	11.8	10.3

[a]The substrates were hydrolyzed at 10 % dry matter content, pH= 5.2, 50 °C for 12 hour at 5.0 FPU/g cellulose enzyme load with 100 rpm shaking.
[b]Percent of total reducing sugars
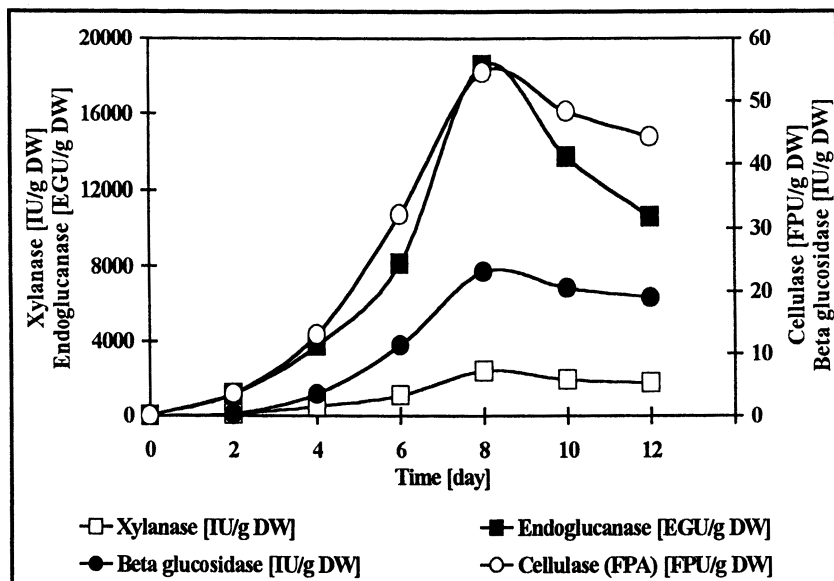
Figure 3. The kinetics of production of lignocellulose degrading enzymes on spent brewing grain by Trichoderma reesei RUT C30
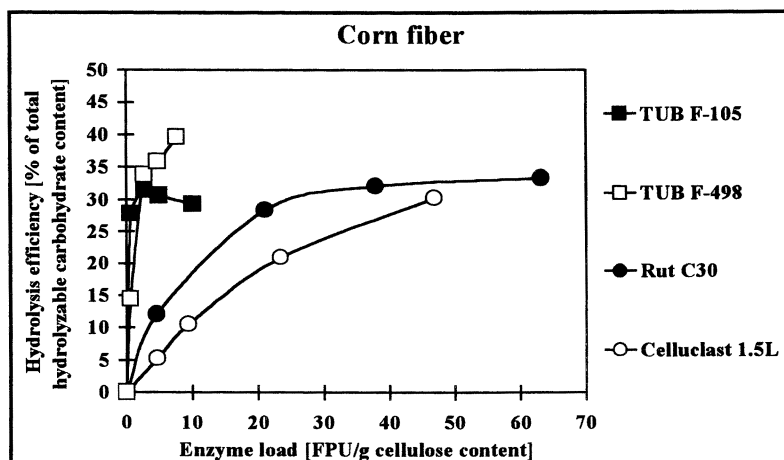
Figure 4. Dependence of hydrolysis on enzyme load and enzyme source. Enzyme sources: Celluclast1.5L (Novo-Nordisk); crude autologous SSF enzymes of Rut C30 (Trichoderma reesei), TUB F-498 (Gliocladium sp.) and TUB F-105 (Trichoderma hamatum)
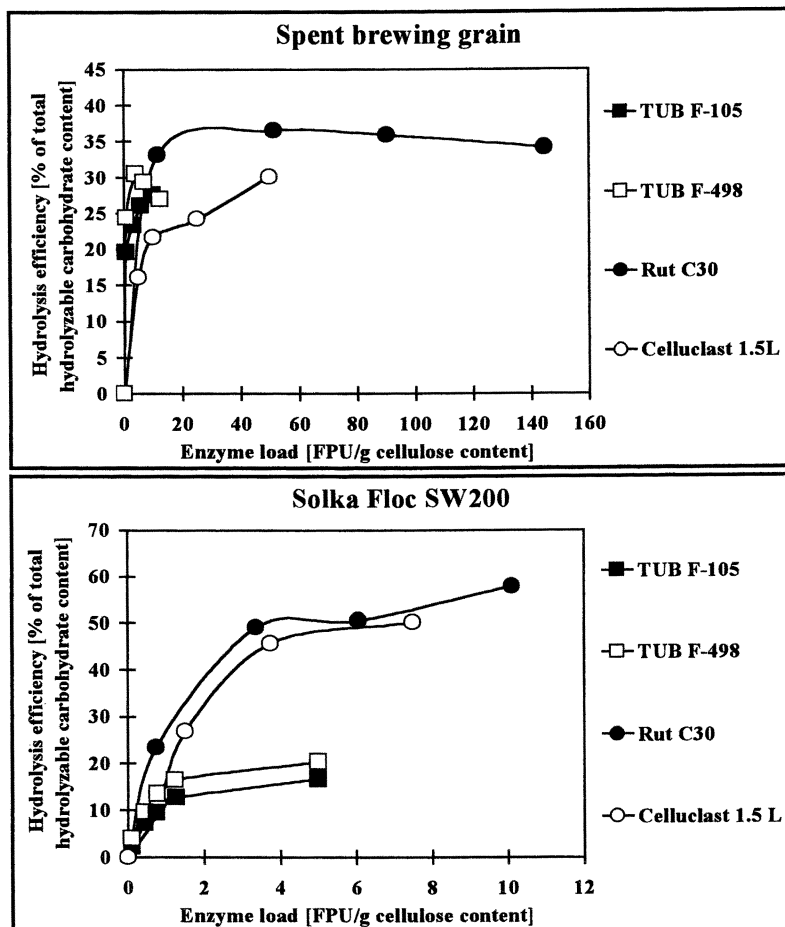
Figure 4. *Continued.*

(*18*). This may be particularly true for crude SSF enzymes, because their action depends on complex synergistic interactions with the substrates, such as affinity of a cellulase component for a substrate, component stereospecificity, ratio of enzyme components, adsorption-desorption on a substrate, etc.

In conclusion, SSF enzymes may be produced efficiently on agricultural byproducts and may be used advantageously for the hydrolysis of natural substrates, according to their substrate specificity. The much reduced production cost of SSF enzymes and their direct applicability in many agrobiotechnological processes make them excellent candidates for wider use in such processes.

Acknowledgement

Literature Cited

1. Tengerdy, R.P.; Szakács, G. *J.Biotechnol.* **1998**, *66*, 91-99.
2. Lonsane, B.K.; Ghildyal, N.P. In *Solid Substrate Cultivation*; Doelle, H.W., Mitchell, D.A., Rolz, C.E., Eds.; Elsevier Applied Science: London and New York, 1992; pp. 191-208.
3. McCoy, M. *Chem.& Eng.News*, **1998**, May 4, 29-30.
4. Wrotnowski, C. *Genet.Eng.News*, **1997**, Feb.1, 14,30.
5. Duchiron, F. *Agro-Food-Ind Hi-Tech.*, **1997**, *8*, 2-4.
6. Persson, I.; Tjerneld, F.; Hahn-Hagerdal, B. *Proc.Biochem.* **1991**, *26*, 65-74.
7. Haltrich, D.; Nidetzky, B.; Kulbe, K.D.; Steiner, W.; Zupancic, S. *Biores.Technol.* **1996**, *58*, 137-161.
8. Tengerdy, R.P. *J.Sci.&Ind.Res.* **1996**, *55*, 313-316.
9. Durand, A.; Almanza, S.; Renaud, R.; Maratray, J. *Agro-Food-Ind. Hi-Tech.*, **1997**, *8*, 39-42.
10. Ghose, T.K. *Pure & Appl.Chem.* **1987**, *59*, 257-268.
11. Bailey, M.J.; Nevalainen, K.M.H. *Enzyme Microb.Technol.* **1981**, *3*, 153-157.
12. Bailey, M.J.; Biely, P.; Poutanen, K. *J.Biotechnol.* **1992**, *23*, 257-270.
13. Kubicek, C.P. *Arch.Microbiol.* **1982**, *132*, 349-354.
14. Miller, G.L. *Anal.Chem.* **1959**, *31*, 426-428.
15. Chahal, D.S. *Appl.Environ.Microbiol.* **1985**, *49*, 205-210.
16. Awafo, V.A.; Chahal, D.S.; Simpson, B.K.; Le, G.B. *Appl. Biochem.Biotechnol.* **1996**, *57/58*, 461-470.
17. Roussos, S.; Raimbault, M.; Saucedo-Castaneda, G.; Viniegra-Gonzalez, G.; Lonsane, B.K. *Micol.Neotrop.Apl.* **1991**, *4*, 19-40.
18. Nieves, R.A.; Ehrman, C.I.; Adney, W.S.; Elander, R.T.; Himmel, M.E. *World J.Microbiol.Biotechnol.* **1998**, *14*, 301-304.
19. Tengerdy, R.P.; Szakács, G.; Sipöcz, J. *Appl.Biochem.Biotechnol.* **1996**, *57/58*, 563-569.

20. Schmidt, J.; Sipöcz, J.; Kaszás, I.; Szakács, G.; Gyepes, Á.; Tengerdy, R.P. *Biores.Technol.* **1997**, *60*, 9-13.

21. Duff, S.J.B.; Murray, W.D. *Biores.Technol.* **1996**, *55*, 1-33.

22. Gulati, M.; Kohlmann, K.; Ladisch, M.R.; Hespell, R.; Bothast, R.J. *Biores. Technol.* **1996**, *58*, 253-264.

23. Ruth, M.F.; Howard, J.A.; Nikolov, Z.; Hooker, B.S.; Himmel, M.E.; Thomas, S.R. *21st Symposium on Biotechnology for Fuels and Chemicals*, Ft.Collins, Colorado, 1999. Abstract book, poster No. 3-49.

24. Akin, D.E.; Morrison III, W.H.; Gamble, G.R.; Rigsby, L.L.; Henriksson, G.; Eriksson, K.-E.K. *Textile Res.J.* **1997**, *67*, 279-287.

25. Uludag, S.; Loha, V.; Prokop, A.; Tanner, R.D. *17th Symposium on Biotechnology for Fuels and Chemicals*, Vail, Colorado, 1995, Abstract book, poster No. 54.

26. Anon. *Agro-Food. Ind. Hi-Tech*, **1999**, *10*, 45.

27. Lyons, T.P.; Jaques, K.A. (Eds), *Proc. Alltech's 14th Annual Symposium (Biotechnology in the Feed Industry)*, Nottingham University Press, UK, 1998.

28. Kirk, T.K.; Jeffries, T.W. In *Enzymes for Pulp and Paper Processing*; Jeffries, T.W. and Viikari, L., Eds., ACS Symposium Series 655, American Chemical Society, Washington, 1996; pp.2-14.

29. Viikari, L.; Suurnakki, A.; Buchert, J. **Ibid.** pp. 15-24.

30. Scott, G.M.; Akhtar, M.; Lentz, M.J.; Kirk, T.K.; Swaney, R. *TAPPI J.* **1998**, *81*, 220-225.

31. Finstein, M.S.; Miller, F.C.; Strom, P.F.; MacGregor, S.T.; Psarianos, K.M. *Bio/Technol.* **1983**, *1*, 347-353.

32. Doelle, H.W.; Mitchell, D.A.; Rolz, C. (Eds.) *Solid Substrate Cultivation.* Elsevier, New York, 1992.

33. Atlas, R.M. *Chem.&Eng.News*, **1995**, Apr.3, 32-42.

34. Chaudhry, G.R. (Ed.) *Biological Degradation and Bioremediation of Toxic Chemicals*, Dioscoroides Press, Portland, Oregon, 1994

35. Lynch, J.M. *Soil Biotechnology. Microbiological Factors in Crop Productivity.* Blackwell Scientific Publ., Oxford, 1983.

36. Harman, G.E.; Kubicek, C.P. (Eds.) *Trichoderma & Gliocladium. Vol.2. Enzymes, Biological Control and Commercial Applications.* Taylor & Francis Ltd., London, 1998.

Chapter 12

Thermostable Cellulase and Xylanase from *Thermoascus aurantiacus* and Their Potential Industrial Applications

M. K. Bhat[1], N. J. Parry[1,3], S. Kalogiannis[1,4], D. E. Beever[2], and E. Owen[2]

[1]Food Materials Science Division, Institute of Food Research,
Norwich Research Park, Colney, Norwich NR4 7UA, United Kingdom
[2]Department of Agriculture, University of Reading, Earley Gate,
P.O. Box 236, Reading RG6 6AT, United Kingdom
[3]Current address: Unilever Research, Colworth Laboratory,
Colworth House, Bedford MK44 1LQ, United Kingdom
[4]Current address: Hellenic Sugar Industry, Sindos, Thessalonika, Greece

Thermostable cellulase and xylanase have potential applications in food, animal feed and chemical industries. A major endoglucanase (Endo-I), a major β-glucosidase (β-Glu-I), two major exoglucanases (Exo-HA and Exo-LA) and a major xylanase (Xln-I) were purified and characterised from the thermophilic fungus, *Thermoascus aurantiacus*. Endo-I, Exo-HA and Exo-LA were optimally active at 70°C, at pH 3.5 and between pH 4.4-5.2 respectively, whereas β-Glu-I and Xln-I showed optimal activity at 80°C and between pH 4.0-5.0. All five enzymes showed remarkable stability up to 60°C and between pH 4.0-6.0 for several hours. β-Glu-I was specific for β-linked glycosides, disaccharides and oligosaccharides. The Endo-I, Exo-HA and Exo-LA were specific for oligosaccharides and polysaccharides with $\beta(1\rightarrow4)$ glycosidic bonds. The Xln-I showed higher affinity towards substituted xylans and was active towards xylo-oligosaccharides (Xyl_n). β-Glu-I cleaved one glucose unit at a time from the non-reducing end of 4-nitrophenyl cello-oligosaccharides ($4\text{-}NPGlc_n$). Endo-I hydrolysed Glc_n, 4-methylumbelliferyl cello-oligosaccharides ($MeUmbGlc_n$), carboxymethyl cellulose (CM-cellulose) and H_3PO_4-swollen cellulose in a random fashion. The Exo-HA and Exo-LA preferentially attacked the second glycosidic bond adjacent to 4-methylumbelliferone (MeUmb) of $MeUmbGlc_n$. Xln-I cleaved xylans and Xyl_n in an endo-fashion. β-Glu-I and Xln-I from *T. aurantiacus* catalysed the synthesis of various alkyl- and aryl-glycosides. These results and the potential industrial applications of all five enzymes from *T. aurantiacus* are discussed.

204

Plant cell-wall polysaccharides which include mainly cellulose and hemicellulose serve as a major source of renewable energy worldwide. Research during the sixties and seventies indicated that cellulose and hemicellulose could be converted to soluble sugars by a group of enzymes called cellulases and hemicellulases (*1-4*). Nevertheless, the subsequent research revealed that enzymatic conversion of lignocellulosic materials to soluble sugars on an industrial scale was rather difficult and uneconomical (*2*). Alternatively, cellulases and hemicellulases are widely used in food, animal feed, textile, paper and pulp as well as in many other industrial applications (*5,6*). Such applications, however, prefer thermostable cellulases and hemicellulases which are active and stable over a wide range of pH and temperatures. The other advantages of using thermostable cellulases and hemicellulases include (a) the repeated use of these enzymes, (b) rapid completion of the desired reaction, and (c) reduced risk of microbial contamination during processing. We have therefore concentrated our research on cellulases and hemicellulases from the thermophilic fungi and bacteria.

Thermoascus aurantiacus is a thermophilic fungus which produces relatively high levels of cellulase and hemicellulase components when grown on lignocellulosic carbon sources such as corn cob and cereal straw (*7-9*). Also, the crude cellulase and hemicellulase preparations from this fungus were found to be stable over a wide range of pH and temperatures (*10*). Cellulase and hemicellulase components from different strains of *T. aurantiacus* have been purified and characterised, but only to a limited extent (*11-14*). The present paper describes the detailed purification and characterisation of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T. aurantiacus* and their potential industrial applications.

Materials and Methods

Materials

T. aurantiacus IMI 216529 obtained from the International Mycological Institute, Surrey, UK was maintained in potato dextrose agar medium. All chemicals used were of analytical grade and purchased either from Sigma or BDH.

Methods

Production of Cellulase and Hemicellulase by *T. aurantiacus*

The fungus was cultured in a 300 litre fermenter using Mandels medium (*15*). Initially, a 750 ml pre-culture was grown in Mandels medium without wheat bran for 30 h at 50°C and used to prepare a 10 litre inoculum. This inoculum was prepared using Mandels medium containing glucose (30 g l⁻¹), yeast extract (2.5 g l⁻¹) and wheat bran (5 g l⁻¹), and grown for 48 h at 50°C. This was used to inoculate the fermenter containing 300 litres of Mandels medium supplemented with paper paste (20 g l⁻¹), wheat bran (20 g l⁻¹) and yeast extract (5 g l⁻¹) at pH 5.0. The fermentation was carried

out for six days at 50°C with agitation speed of 400 rpm and air flow of 10 litres min[-1], and harvested by filtration using glass fibre. The filtrate was concentrated and fractionated into cellulase and hemicellulase rich fractions by ultra-filtration (200,000 and 15,000 kDa molecular mass cut off membranes), and precipitated using 50% (v/v) acetone.

Assay of endoglucanase, exoglucanase, β-glucosidase and xylanase activities

Endoglucanase: This was assayed using CM-cellulose (medium viscosity), according to the method of Wood and Bhat (*16*). A 2 ml reaction mixture containing 1ml of 1% (w/v) CM-cellulose, 0.5 ml of 0.2 M sodium acetate buffer, pH 5.0 and 0.45 ml distilled water was incubated at 50°C for 5 min before adding either 0.05 ml culture filtrate or 0.5 μg pure enzyme. The reaction was continued for further 15 min at the same temperature and the reducing sugars released were determined using Somogyi-Nelson reagent (*17*).

Exoglucanase: This was assayed using Avicel PH101, according to the method of wood and Bhat (*16*). A 2 ml reaction mixture containing 1 ml of 1% (w/v) Avicel PH101 suspension, 0.5 ml of 0.2 M sodium acetate buffer pH 5.0, 0.4 ml of distilled water and either 0.1 ml of culture filtrate or 2 μg pure enzyme was incubated at 50°C for 2 h. The reaction was terminated by placing the tubes in a boiling bath for 10 min and the reducing sugars present in the supernatant were determined using Somogyi-Nelson reagent (*17*).

β-*Glucosidase:* This was assayed by microtitre plate method using 4-nitrophenyl-β-D-glucoside (4-NPGlc) as a substrate. A reaction mixture containing 25 μl of culture filtrate or 20 ng of pure enzyme, 25 μl of 0.2 M sodium acetate buffer, pH 5.0 and 50 μl of distilled water was incubated at 50°C for 10 min. The reaction was initiated by adding 25 μl of 4-NPGlc and incubated for further 15 min at 50°C, and terminated by adding 100 μl glycine buffer, pH 10.8. The colour developed was measured at 405 nm.

Xylanase: This was assayed using birchwood xylan, according to the method of Bailey *et al* (*18*). A 1.8 ml of 1% (w/v) birchwood xylan prepared in 0.05 M sodium citrate buffer, pH 5.3 was incubated at 50°C for 5 min. A 0.2 ml of culture filtrate or 2 μg of pure enzyme was added and incubated for 5 min before terminating the reaction by adding 3 ml of dinitrosalicylic acid (DNS) reagent. The reducing sugars released were determined by DNS method (*19*). In all cases, the unit of enzyme activity was expressed as μmoles of product released min[-1]. mg protein[-1].

Purification of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T.aurantiacus*

Fifteen grams (by weight) of dried cellulase rich fraction from *T. aurantiacus* was dissolved in 100 ml distilled water and subjected to 80% $(NH_4)_2SO_4$ saturation. The precipitated protein was dissolved in 50 ml of 0.01 M sodium acetate buffer, pH 5.0, filtered through a 0.2 μM PVDF membrane and desalted on a Bio-Gel P6 DG column (2.6x100 cm). Active fractions were pooled, concentrated and equilibrated with 0.05 M sodium acetate buffer, pH 5.0 using an Amicon cell fitted with an Omega membrane (10 kDa, molecular weight cut off). The concentrated sample (35 ml) was

fractionated on a DEAE-Sepharose (fast flow) column (XK2.6x90 cm) equilibrated with 0.05 M sodium acetate buffer, pH 5.0. A major xylanase and a minor β-glucosidase were eluted in the buffer wash. The bound endoglucanase, exoglucanase and β-glucosidase activities were eluted using a NaCl gradient (0-0.5 M). The fractions containing these three activities were pooled, concentrated and desalted. This sample (25 ml) was separated on an Ultrogel ACA 44 column (XK 1.6x100 cm), at a flow rate of 1 ml min^{-1}. The β-Glu-I, completely separated from endo- and exoglucanase activities, was further purified by chromatofocusing on a Mono-P column over the pH range 3.4-5.5 and using Polybuffer 74 (Pharmacia). The endo- and exoglucanases were separated and purified using p-aminobenzyl-1-thio-β-D-cellobioside coupled to Sepharose as an affinity matrix (20). The endo- and exoglucanases were eluted using 0.1 M sodium acetate buffer pH 5.0, containing 0.1 M glucose and 0.1 M cellobiose, respectively. Also, the affinity chromatography separated major and minor endoglucanases as well as high and low affinity exoglucanases. The final purification of Endo-I, Exo-HA and Exo-LA was carried out using a Mono-Q column, equilibrated with Tris HCl buffer, pH 8.0 (Scheme 1).

The hemicellulase rich fraction (1.5 g by weight) from *T. aurantiacus* was dissolved in 30 ml of 0.05 M bicine buffer, pH 8.7 and desalted on a Bio-Gel P6 DG column (5x27 cm), equilibrated with the same buffer. Active fractions were pooled and applied on to a Q-Sepharose fast flow column (20x3.4 cm), equilibrated with 0.05 M bicine buffer, pH 8.7, and eluted using the same buffer at a flow rate of 600 ml h^{-1}. The Xln-I was eluted in the buffer wash, while a minor xylanase (Xln-II), β-xylosidase and acetylesterase with β-xylosidase activity bound to the Q-Sepharose column were eluted using 0 - 0.4 M NaCl gradient. All these enzymes were subsequently purified at pH 3.5 using a Hi trap SP column, equilibrated with 20 mM sodium acetate buffer, pH 3.5 (Scheme 2).

Characterisation of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I

Determination of molecular mass and isoelectric point (pI): The purity of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I was confirmed by SDS-PAGE according to the method of Lugtenberg *et al.*(21) and using BioRad mini-protean II electrophoresis unit. The molecular mass of individual proteins was determined using the plot of relative mobility of standard protein markers versus their log molecular mass. The standard protein markers used in the order of increasing molecular mass include: lysozyme (14.5 kDa), trypsin inhibitor (21.5 kDa), carbonic anhydrase (31.0 kDa), lactate dehydrogenase (36.5 kDa), glutamic dehydrogenase (55.4 kDa), bovine serum albumin (66.3 kDa), phosphorylase B (97.4 kDa), β-galactosidase (116.3 kDa) and myosin (200 kDa).

For the determination of pI of purified enzymes, the isoelectric focusing (IEF) was performed according to the instructions of Novex '96 electrophoresis catalogue, using a Novex X-mini IEF cell system and Novex pre-cast gels (5%) containing 2% ampholytes covering a pH range 3.0-7.0. The pI values of individual proteins were determined using the plot of relative mobility of standard protein markers versus their

Cellulase rich fraction

Ammonium sulphate precipitation (80%)

Desalting (Biogel P6 DG column)

DEAE Sepharose column chromatography, pH 5.0

Xylanase   **β-Glucosidase (minor)**   **Endoglucanase/Exoglucanase/ β-Glucosidase (major)**

Ultragel AcA 44 column chromatography

β-Glucosidase   **Exoglucanase/Endoglucanase**

Mono P-Chromatofocusing   *Affinity chromatography*

β-Glucosidase (β-Glu-I-major-Pure)

28kD protein   **Endoglucanase**   **low affinity**   **high affinity**

Mono Q, pH 8.0   *Mono Q, pH 8.0*   *Mono Q, pH 8.0*   *Mono Q, pH 8.0*

28kD protein   **Endoglucanase (Endo-I-major-pure)**   **Exoglucanase HA (Exo-HA-pure)**

Minor   **Exoglucanase LA (Exo-LA-major-pure)**

Scheme 1. Schematic representation of the purification of cellulolytic enzymes from T. aurantiacus.

Hemicellulase rich fraction

Q-Sepharose, pH 8.7

Pool A - Xylanase-I (major-Xln-I) pure   **Pool B** *xylanase-II (minor-Xln-II)*   **Pool C** *xylosidase / acetylesterase*   **Pool D** *acetylesterase / xylosidase*

HI Trap SP pH 3.5   HI Trap SP pH 3.5   HI Trap SP pH 3.5

xylanase-II (minor-Xln-II) pure   **Pool C1** *xylosidase*   **Pool D1** *acetylesterase / xylosidase*
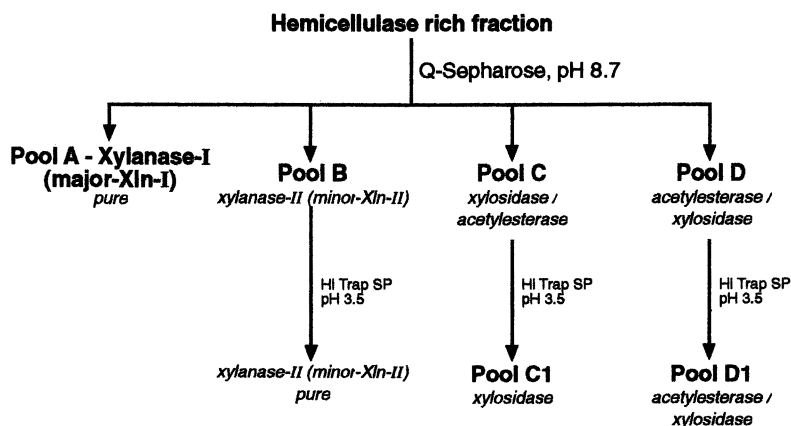
Scheme 2. Schematic representation of the purification of xylanolytic enzymes from T. aurantiacus.

pI. The standard protein markers used in the order of increasing pI include: pepsinogen (pI, 2.80), amyloglucosidase (pI, 3.50), glucose oxidase (pI, 4.15), lactoglobulin (pI, 5.20), bovine carbonic anhydrase (pI, 5.85) and human carbonic anhydrase (pI, 6.55).

Determination of pH and temperature optima and stability: The pH and temperature optima of Endo-I, β-Glu-I, Exo-HA and Exo-LA were determined by measuring the activity as described above and using a citrate/phosphate buffer (pH 2.8-6.8) at temperatures ranging from 30-90°C. The stability of these enzymes was determined by measuring the residual activity for 48 h at different temperatures (50-80°C) as a function of pH (2.8-6.8). The pH and temperature optima of Xln-I were determined by measuring the activity in the pH region 2.2-10.2 and in the temperature range 50-90°C. The stability of Xln-I was determined at different pH (3.0-9.8) and temperatures (50-80°C) for six days.

Determination of carbohydrate content: This was determined as described (*22*), using mannose as the standard. A 1 ml enzyme solution containing 10 μg pure protein was mixed with 25 μl of 80% (w/v) phenol, followed by the rapid addition of 2.5 ml of H_2SO_4. The sample was left at room temperature for 10 min, and at 30°C for a further 10 min, before reading the absorbance at 490 nm.

Determination of amino acid composition: A 250 μg of pure protein was hydrolysed with 1 ml of 6 M HCl containing 0.1 % (v/v) phenol *in vacuo* at 110°C for 24, 48 and 72 h. Amino acids were quantitatively determined using a Beckman 6300 amino acid analyser. Cysteine and methionine were determined after oxidation with performic acid as cysteic acid and methionine sulphone, respectively.

Determination of thermostability of Endo-I by differential scanning calorimetry (DSC): A 1.5 mg of Endo-I was dissolved in 25 ml of 0.05 M glycine buffer, pH 2.8 and concentrated to 2 ml using an Amicon cell (10 ml) fitted with an Omega membrane (10 kDa, molecular weight cut off). The enzyme solution was placed in the DSC cell and the temperature increased to 100°C in order to follow the thermodynamics of protein unfolding. Scans were performed in triplicates using a Microcal MC-2 Scanning Calorimeter with the scan rate of $60°h^{-1}$. Data were collected and analysed using Microcal Origin Software. The peak areas were integrated and a thermodynamic model fitted. The energetic parameters were determined according to the method of Goldenberg and Creighton (*23*).

Determination of activity towards 4-nitrophenyl glycosides (4-NPGly), disaccharides, oligosaccharides and polymers: The activity of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I towards 4-NPGly was determined by measuring the 4-NP released at 405 nm as described in β-glucosidase assay, while the activity towards disaccharides, oligosaccharides and polysaccharides was determined by measuring either the glucose (*16*) or reducing sugars (*17,19*) released.

Hydrolysis of model substrates and analysis of hydrolysis products by HPLC: Model substrates, Glc_n (n=2-6), and $4\text{-}NPGlc_n$, (n= 2-5) were used to study the mode of action of β-Glu-I, while $MeUmbGlc_n$, (n=2-5) and Xyl_n, (n=3-6) were used to determine the mode of action of Endo-I, Exo-HA, Exo-LA and Xln-I, respectively. A 0.3 ml reaction mixture containing 0.2 ml of either Glc_n (100 μg), $4\text{-}NPGlc_n$ (100 μg),

MeUmbGlc$_n$ (100 μM) or Xyl$_n$ (1-10 mM), 0.05 ml of 25 mM buffer (citrate phosphate, pH 3.5 in case of Endo-I; sodium acetate, pH 5.0 in case of β-Glu-I, Exo-LA and Xln-I as well as sodium acetate, pH 4.7 in case of Exo-HA) and 0.05 ml of enzyme (0.02-2 μg) was incubated at 50°C for 2 h. A 25 μl sample was withdrawn at different time intervals and mixed with 25 μl acetonitrile/water (70:30). Hydrolysis products released from Glc$_n$, 4-NPGlc$_n$ and Xyl$_n$ were analysed using an NH$_2$-Spherisorb column (0.25x10 cm) fitted to a Gynkotech HPLC system with Evaporative light scattering detector (ELSD), and acetonitrile/water (70:30) as a solvent at a flow rate of 0.25 ml min^{-1}. Hydrolysis products released from MeUmbGlc$_n$ were analysed using BioRad Biosil Polyol column (25x4.6 cm) fitted to a Walters HPLC system with UV detector, and acetonitrile/water (70:30) as a solvent at a flow rate of 1 ml min^{-1}. Kinetics of bond cleavage and bond cleavage frequencies were determined using reciprocal plots on Enzfitter programme and according to the method of Walker and Dewar (*24*), respectively.

Transferase activity of β-Glu-I *and Xln-I:* This was assayed using glucose, cellobiose and xylan as donors and methanol, ethanol, propanol and butanol as acceptors. In case of β-Glu-I, a 300 μl reaction volume containing 50 μl (100 ng) enzyme, 50 μl (100 μg) cellobiose or glucose, 75 μl of alcohol (25%, v/v) and 125 μl distilled water was incubated at 50°C for 24 h. A 25 μl sample was withdrawn at different time intervals, mixed with 25 μl of acetonitrile:water (70:30) and analysed using an NH$_2$-Spherisorb column (0.25x10 cm) fitted to a Gynkotech HPLC system with ELSD and acetonitrile/water (70:30) as a solvent at a flow rate of 0.25 ml min^{-1}. In case of Xln-I, a 400 μl reaction mixture containing 200 μl of 2% (w/v) birchwood xylan in 0.05 M citrate phosphate buffer pH 5.3, 1.6 g of Xln-I and 200 μl of alcohol was incubated at 50°C for 2 h. A 25 μl sample was withdrawn at different time intervals and analysed by HPLC as described above.

Results and Discussion

Purification of Cellulase and Hemicellulase Components from *T.aurantiacus*

The protocols used for the purification of cellulase and hemicellulase components from *T. aurantiacus* are shown in schemes 1 and 2. Using ion-exchange, gel filtration and affinity chromatography techniques, Endo-I, β-Glu-I, Exo-HA and Exo-LA were purified from the cellulase rich fraction of *T. aurantiacus*. In addition, a minor endoglucanase, a minor β-glucosidase and a minor exoglucanase were purified in small quantities. Exo-HA and Exo-LA were designated as high and low affinity exoglucanases based on their affinity towards p-aminobenzyl-1-thio-β-D-cellobioside. Xln-I, Xln-II, β-xylosidase and an acetyl esterase with xylosidase activity were isolated and purified from the hemicellulase rich fraction of *T.aurantiacus*. The present paper describes the detailed biochemical properties of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T. aurantiacus*.

Physico-Chemical Properties of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I

Some of the physico-chemical properties of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T.aurantiacus* are summarised in Tables I and II. The molecular masses of Endo-I and Xln-I were 34 and 32 kDa, respectively, while that of β-Glu-I was 120 kDa. Both Exo-HA and Exo-LA, however, showed identical molecular mass between 52-55 kDa. Endo-I, β-Glu-I, Exo-HA and Exo-LA had acidic pI values, while Xln-I had a neutral pI (Table 1). All the above enzymes were optimally active in acidic pH and at temperature optima of 70 or 80°C (Table 1 and Figure 1 A-E). Exo-HA, Exo-LA and β-Glu-I were highly glycosylated with carbohydrate content of 10-19.6%, while the Endo-I was less glycosylated (Table 1). Interestingly, Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T. aurantiacus* were resistant against protease from *Aspergillus sojae*. Furthermore, the determination of amino acid composition revealed that these enzyme components from *T. aurantiacus* were rich in acidic amino acids as well as glycine and alanine (Table 2).

Table I. Physico-Chemical Properties of Endo-I, Exo-HA, Exo-LA, β-Glu-I and Xyl-I from *Thermoascus aurantiacus*

Property	Endo-I	Exo-HA	Exo-LA	β-Glu-I	Xln-I
M_r value (kDa)	34	52 - 55	52 - 55	120	32
pI	3.5 - 3.7	3.6 - 3.8	3.6 - 3.8	4.1 - 4.5	7.0
pH optimum	3.8 - 4.0	4.6 - 6.0	5.0 - 5.6	4.4 - 5.6	4.4 - 5.0
Temp. optimum	70°C	70°C	70°C	80°C	80°C
Sugar content (%)	2.7	15.5	10.5	19.6	ND

ND - not determined

Table II. Amino acid composition (Mol %) of Endo-I, Exo-HA, Exo-LA, β-Glu-I and Xln-I from *Thermoascus aurantiacus*

Amino acid	Endo-I	Exo-HA	Exo-LA	β-Glu-I	Xln-I
Asp	15.4	14.2	14.0	13.4	13.8
Thr	7.9	10.2	10.1	6.0	7.2
Ser	8.0	7.9	8.0	6.9	5.4
Glu	9.0	12.2	8.2	8.2	9.5
Pro	5.6	4.8	6.0	6.2	4.7
Gly	10.1	11.3	12.0	11.8	7.4
Ala	9.2	7.6	7.6	10.2	11.6
Cys	0.4	1.6	4.0	0.7	0.7
Val	4.4	5.8	5.7	6.4	8.6
Met	2.2	0.8	0.4	1.1	0.9
Ile	6.6	3.8	4.0	3.6	5.4
Leu	5.6	5.2	4.8	8.1	7.5
Tyr	4.4	4.0	4.0	3.8	3.0
Phe	4.6	3.8	4.0	4.0	3.0
Lys	3.6	3.6	3.2	4.6	1.3
His	0.8	1.4	1.9	1.4	4.3
Arg	2.2	1.8	2.1	3.6	3.2
Trp	0.0	0.0	0.0	0.0	2.5
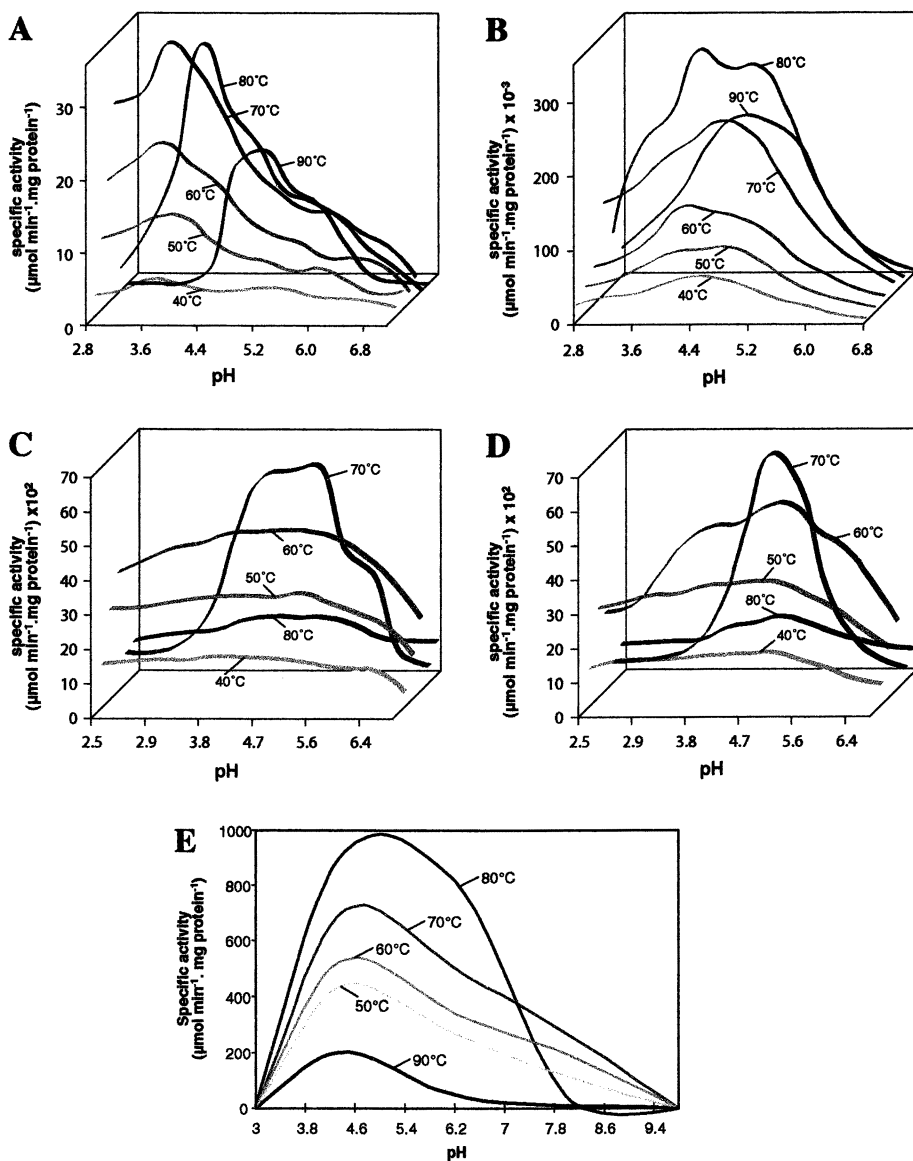
Figure 1. Activity of Endo-I (A), β-Glu-I (B), Exo-HA (C), Exo-LA (D) and Xln (E) from T. aurantiacus as a function of pH and temperature. Figure 1E reproduced with permission from reference 8. Copyright 1996 Med.Fac Landbouww. Universitiet, Gent, Belgium.

The stability of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T. aurantiacus* was determined at different temperatures and pH as a function of time. The Endo-I was completely stable up to 60°C, at pH 5.2 for 48 h (Figure 2A). At 70°C and at pH 5.2, the Endo-I was 60% stable for 48 h, but at 80°C and at pH 5.2, the enzyme was

unstable (Figure 2A). Nevertheless, the Endo-I was completely stable between pH 2.8-6.8 at 50°C for 48 h. Determination of thermostability by DSC revealed that the Endo-I was in a fully folded state up to 70°C and completely unfolded above 80°C (Figure 2B). The close examination of thermal unfolding suggested that the Endo-I unfolded as a two domain protein, one at 75°C and the other at 80°C (Figure 2B).
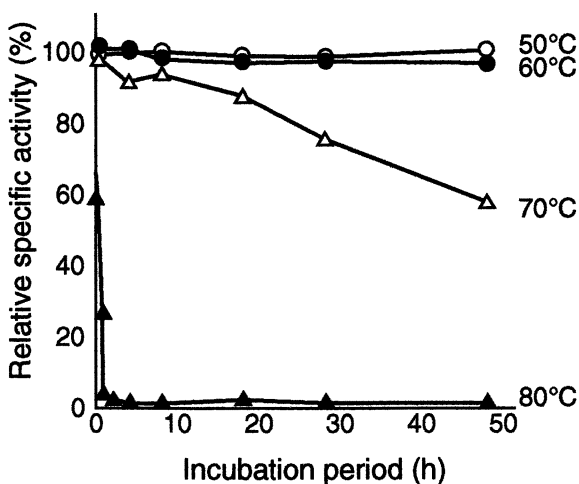
Figure 2A. Stability of Endo-I from T. aurantiacus at pH 5.2.

Tm1   76.61
Hv1   3.95E4
H1   8.48E4
Tm2   81.00
H2   4.98E4
Hv2   1.56E5
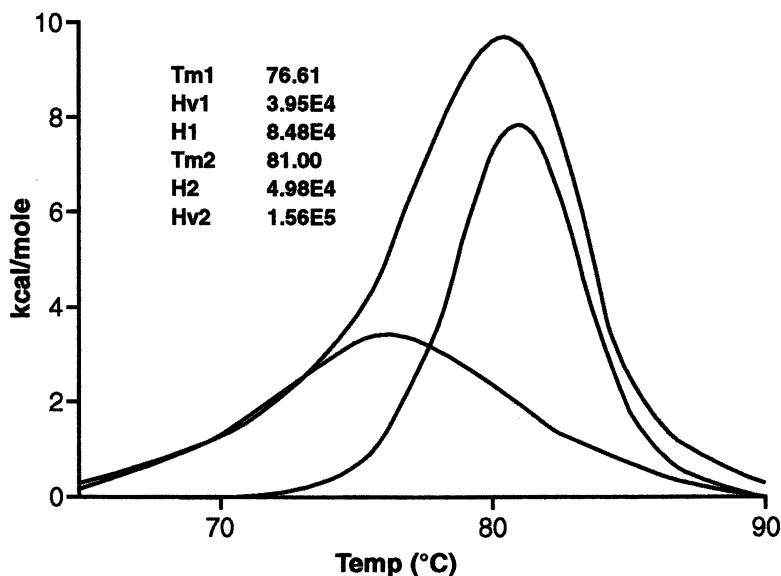
Figure 2B. Thermal unfolding of Endo-I from T. aurantiacus.

β-Glu-I was completely stable up to 70°C at pH 5.2 for 48 h (Figure 3A), but was less stable above and below pH 5.0 and above 60°C. Exo-HA and Exo-LA from *T.aurantiacus* were remarkably stable up to 60°C at pH 5.0 for at least 48 h. Besides, these two enzymes were less stable above 60°C and at pH values below 4.0 and above 6.0. Xln-I, on the other hand, was completely stable up to 80°C and between pH 4.4 - 6.2 for at least six days (Figure 3B). Also, the Xln-I was 90% stable at 60°C between pH 4.4 - 8.0, and 50% stable at 60°C at pH 3.0 and 9.8 for 2 days and 2 h, respectively. These results showed that the Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T.aurantiacus* were remarkably stable and could be ideal for processing lignocellulosic materials.
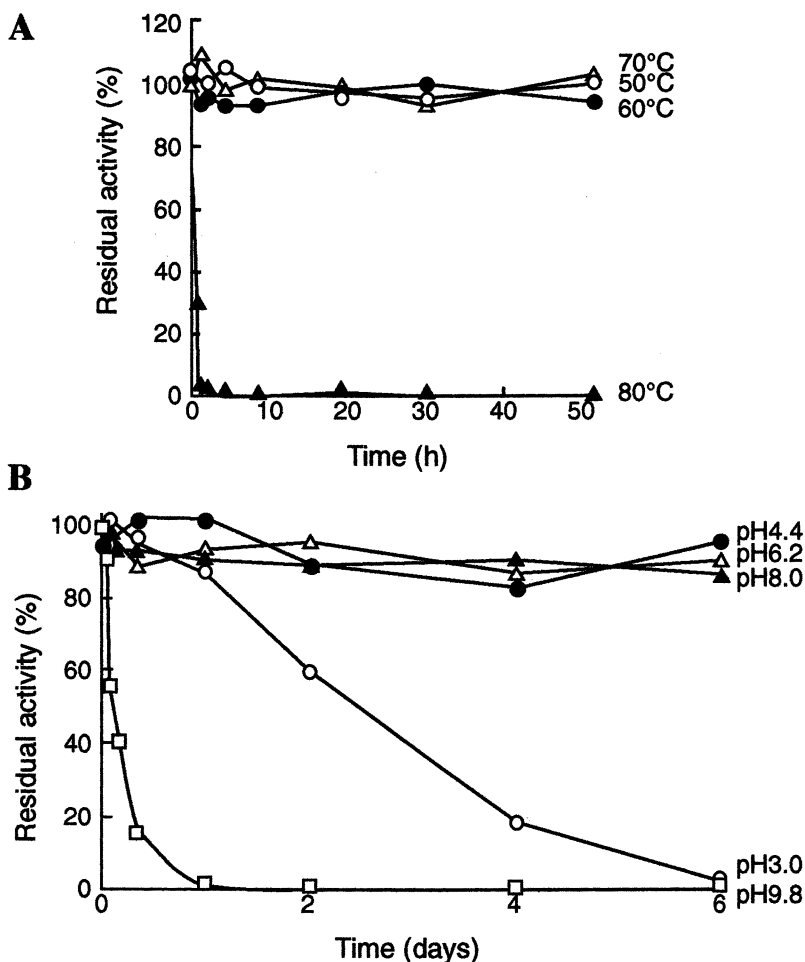
Figure 3. *Stability of β-Glu-I and Xln-1 from T. aurantiacus at pH 5.2 (A) and 60°C (B) respectively.*

Substrate specificity of Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I

The Endo-I showed the highest activity towards polysaccharides such as CM-cellulose, barley β-glucan and lichenan. However, the enzyme hydrolysed substituted and unsubstituted Glc_n, neosugar (mixture of 1-kestose, nystose and 1^F-β-fructo furanosyl-nystose), xylan and Avicel to a limited extent. The Endo-I was inactive on CM-pachyman, laminarin, β (1->3) linked glycosides and different cyclodextrins. These results demonstrated that the Endo-I from *T. aurantiacus* was specific for the substrates with β (1->4) glycosidic bond.

The β-Glu-I from *T. aurantiacus* showed the highest activity towards 4-NPGlc. Nevertheless, the enzyme was active towards β-linked glycosides (e.g. 2-NPGlc, 2-NPXyl, n-octyl β-D-glucoside), disaccharides (e.g. Glc_2, β-trehalose, sophorose, β-gentiobiose, 4-$NPGlc_2$) and oligosaccharides (e.g. Glc_n, n=2-6; 4-$NPGlc_n$, n=2-5 and $MeUmbGlc_n$, n=2-5). Besides, β-Glu-I hydrolysed laminarin and CM-cellulose to a limited extent, and was inactive on Avicel, xylan, lichenan, α (1->4), α (1->6) and α (1->1) linked glucopyranosyl compounds.

Both Exo-HA and Exo-LA were highly specific for 4-$NPGlc_2$ and 4-NPLac. These enzymes were active towards substituted and unsubstituted Glc_n (n=2-6), Avicel, H_3PO_4-swollen cellulose, barley glucan, lichenan and to a limited extent on CM-cellulose. Interestingly, the Exo-HA showed a higher activity towards all the above polysaccharides compared to Exo-LA. Nevertheless, both Exo-HA and Exo-LA were inactive towards laminarin, xylan, CM-pachyman, neosugar and cyclodextrins.

Xln-I from *T.aurantiacus* showed higher activity on birchwood xylan (633 μmol min^{-1} mg $protein^{-1}$) than on oat spelt xylan (561 μmol min^{-1} mg $protein^{-1}$). However, the affinity of this enzyme towards oat spelt xylan (K_m, 2.3 mg ml^{-1}) was higher compared to that towards birchwood xylan (K_m, 3.3 mg ml^{-1}). The enzyme was also active towards Xyl_n (n= 2-6), 4-$NPGlc_2$, 4-NPXyl, $MeUmbXyl_2$ and MeUmbXylGlc. Interestingly, the enzyme showed higher activity towards 4-NP-α-L-arabinopyranoside and 4-NP-α-L-arabinofuranoside than that towards 4-NPXyl. The efficient hydrolysis of 4-NPXyl and 4-NPXylGlc by Xln-I suggested that this enzyme belongs to family 10 of glycosyl hydrolases.

Mode of action of Endo-I, Exo-HA, Exo-LA, β-Glu-I and Xln-I from *T. aurantiacus*

The mode of action of Endo-I, Exo-HA and Exo-LA was studied using $MeUmbGlc_n$, while the mode of action of β-Glu-I and Xln-I was studied using unsubstituted and 4-NP substituted Glc_n and Xyl_n, respectively. In all cases, the hydrolysis products were analysed by HPLC and the initial product ratios and the reaction kinetics were used to determine the mode of action of individual enzymes on above model substrates.

The mode of action, initial product ratios and the reaction kinetics of Endo-I on $MeUmbGlc_n$ are summarised in Figure 4. Endo-I preferentially cleaved the glycosidic bond adjacent to MeUmb of MeUmbglycosides up to $MeUmbGlc_3$ (Figure 4). Also, the turnover number of the Endo-I towards the preferential site of $MeUmbGlc_3$ was ten times higher than that towards the preferential site of $MeUmbGlc_2$. The Endo-I

preferentially cleaved the internal glycosidic bonds of MeUmbGlc$_4$ and MeUmbGlc$_5$. Moreover, the mode of action of Endo-I on MeUmbGlc$_5$ was random and suggested that it was a true endoglucanase, similar to endoglucanases characterised from other fungi (25,26).

k_{cat} (MeUmb) = 4.5

k_{cat} (MeUmb) = 7.7

k_{cat} (MeUmb) = 77.8

k_{cat} (MeUmb) = 117.4
k_{cat} (MeUmbGlc) = 151.8

k_{cat} (MeUmb) = 40.5
k_{cat} (MeUmbGlc) = 48.4
k_{cat} (MeUmbGlc$_2$) = 40.5
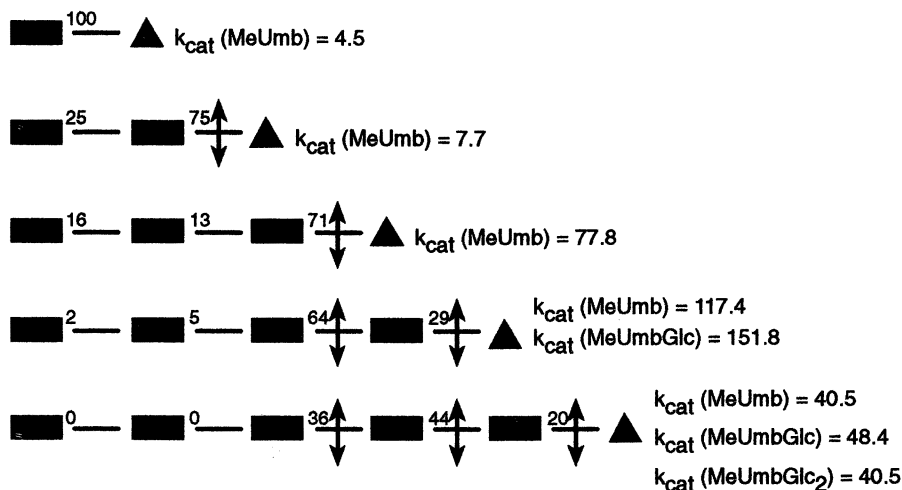
Figure 4. Bond cleavage frequencies and turnover numbers (sec^{-1}) for the hydrolysis of MeUmbGlc$_n$ by Endo-I from T. aurantiacus.

The β-Glu-I rapidly hydrolysed Glc$_2$ to Glc. The hydrolysis of Glc$_3$ by β-Glu-I initially resulted in the formation of Glc and Glc$_2$, and the concentration of Glc$_2$ decreased subsequently as it was hydrolysed to Glc. Glc$_4$ was initially hydrolysed to Glc, Glc$_2$ and Glc$_3$ by β-Glu-I, but Glc$_2$ and Glc$_3$ were subsequently hydrolysed to Glc. Similarly, the enzyme released Glc, Glc$_2$, Glc$_3$ and Glc$_4$ from Glc$_5$, and these oligosaccharides were subsequently hydrolysed to Glc. These results suggested that β-Glu-I from T. aurantiacus hydrolysed Glc$_n$ in a sequential manner, but it was not possible to establish the actual mode of action of β-Glu-I using Glc$_n$. Indeed, the use of 4-NP substituted Glc$_n$ and subsequent analysis of hydrolysis products by HPLC enabled us to demonstrate the precise mode of action of β-Glu-I. Thus, the data revealed that the β-Glu-I from T. aurantiacus released one Glc unit at a time from the non-reducing end of respective 4-NPGlc$_n$ (Figure 5). The rate of hydrolysis of 4-NPGlc$_n$ however decreased with increasing chain length (Figure 5). Also, the enzyme was specific for the terminal glycosidc bond of Glc$_n$ and hydrolysed Glc$_2$ and Glc$_3$ more rapidly than higher Glc$_n$.

the rate of terminal
glycosidic bond cleavage
decreases as the chain
length increases
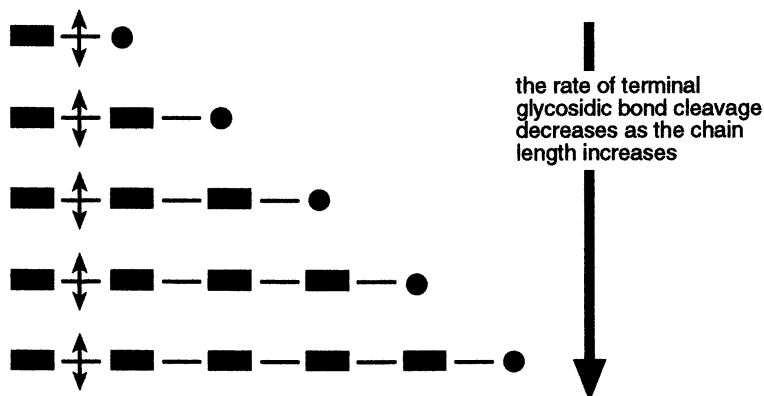
Figure 5. Mode of action of β-Glu-I from T. aurantiacus on NPGlc$_n$.

The mode of action, initial product ratios and the reaction kinetics of Exo-HA and Exo-LA on MeUmbGlc$_n$ are summarised in Figures 6A and B. Both Exo-HA and Exo-LA preferentially cleaved the second glycosidic bond from the MeUmb residue of MeUmbGlc$_2$ and MeUmbGlc$_3$. However, when MeUmbGlc$_4$ was used as a substrate, both enzymes preferentially cleaved the second glycosidic bond adjacent to MeUmb. Both Exo-HA and Exo-LA hydrolysed MeUmbGlc$_5$ rapidly and it was not possible to determine the kinetic constants for the cleavage of individual glycosidic bonds of this substrate. Overall, both Exo-HA and Exo-LA showed an identical mode of action on MeUmbGlc$_n$, but with small differences in their turnover number towards these substrates (Figures 6A and B). Also, the mode of action of these two enzymes on MeUmbGlc$_2$ and MeUmbGlc$_3$ together with their limited activity on CM-cellulose confirmed that both enzymes were typical exoglucanases.

A

B

Figure 6. Bond cleavage frequencies and turnover numbers (sec^{-1}) for the hydrolysis of MeUmbGlc$_n$ by Exo-HA (A) and Exo-LA (B) from T. aurantiacus.
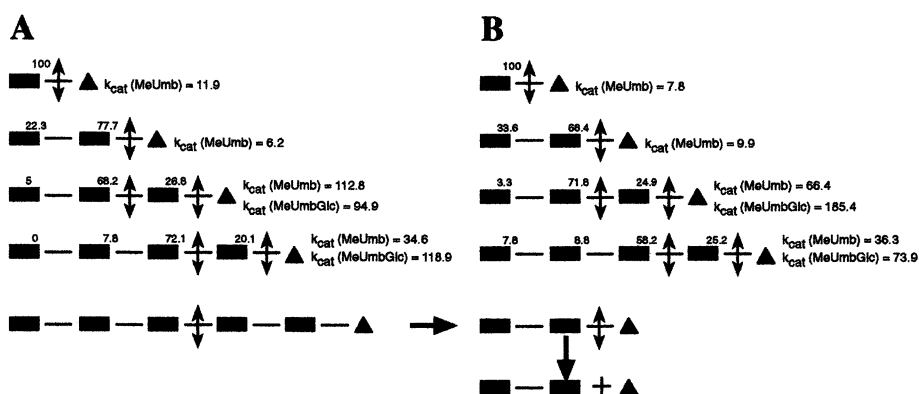
The mode of action of Xln-I was studied using Xyl_3, Xyl_4, Xyl_5 and Xyl_6. The enzyme released equimolar Xyl and Xyl_2 from Xyl_3 up to 2 mM. With higher concentrations of Xyl_3, the ratio of Xyl_2 to Xyl was higher than one. This suggested that, at concentrations higher than 2 mM of Xyl_3, the Xln-I from *T. aurantiacus* catalysed both hydrolysis and transglycosylation reactions. The enzyme released Xyl, Xyl_2 and Xyl_3 from Xyl_4, where Xyl_2 was the main product (Figure 7). However, the concentration of Xyl_3 was noticeably higher than Xyl, when the concentration of Xyl_4 was higher than 2 mM (Figure 7). These results confirmed that Xln-I from *T. aurantiacus* catalysed the hydrolysis and transglycosylation reactions simultaneously.

The Xln-I released mainly Xyl_2 and Xyl_3 from Xyl_5. Similarly, the enzyme released mainly Xyl_3 and equimolar Xyl_2 and Xyl_4 from Xyl_6. These results demonstrated that Xln-I from *T. aurantiacus* cleaved the internal glycosidic bonds of higher Xyl_n and behaved as a typical endo-xylanase.
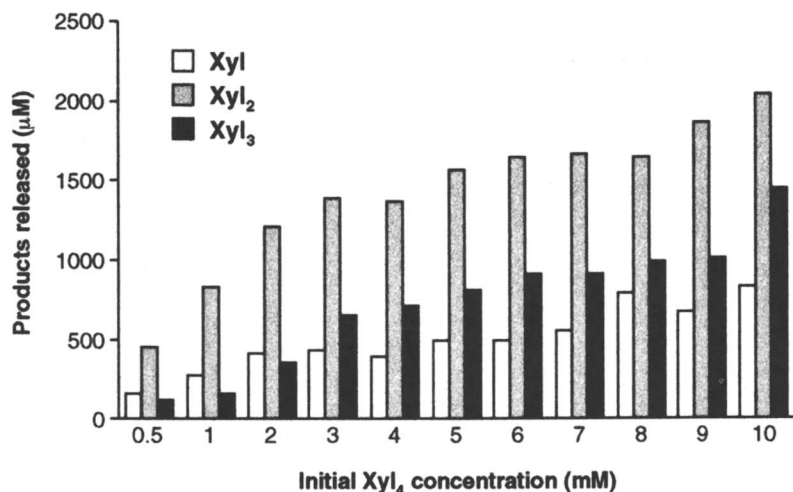
Figure 7. Hydrolysis of Xyl_4 by Xln-I from T. aurantiacus.

Synthesis of Alkyl- and Aryl-glycosides by Transferase Reaction

Alkyl-glycosides and alkyl-polysaccharides have potential applications in detergent and other industries (*27,28*). Many glycosidases catalyse the synthesis of desired alkyl- and aryl-glycosides as well as substituted polysaccharides by transferase reaction (*29*). Both β-Glu-I and Xln-I from *T. aurantiacus* catalysed the synthesis of respective alkyl-glycosides in the presence of methanol, ethanol, propanol and butanol. Furthermore, the Xln-I from *T. aurantiacus* catalysed the synthesis of alkyl- and aryl- xylo-oligosaccharides in the presence of xylan and respective alcohol. These results suggested that the β-Glu-I and Xln-I from *T. aurantiacus* could be useful for the synthesis of alkyl- and aryl-glycosides of commercial interest.

Potential Industrial Applications of Thermostable Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I from *T. aurantiacus*

Thermostable cellulase and xylanase have potential applications in food, animal feed, textile, paper and pulp as well as in chemical industries *(30,31)*. Based on stability, substrate specificity and mode of action, the Endo-I from *T. aurantiacus* is likely to be useful in the production of substituted and unsubstituted Glc_n, filtration of beer, pre-treatment of animal feeds as well as in laundry and biostoning. The ability of β-Glu-I to hydrolyse a wide range of β-linked disaccharides, oligosaccharides, aryl- and alkyl-glycosides suggested that it could be useful in improving the aroma of wines as well as in removing the end-product inhibition during cellulose hydrolysis. Both Xln-I and β-Glu-I would be useful for the production of desired alkyl- and aryl-glycosides as well as transglycosylated products. Also, Xln-I is expected to be useful in the production of xyl_n from agricultural wastes, in improving the nutritional quality of animal feeds as well as in processing paper pulp.

Conclusions

Endo-I, β-Glu-I, Exo-HA, Exo-LA and Xln-I were purified in mg quantities from the culture filtrate of *T. aurantiacus* using relatively simple protocols. All five enzymes characterised in the present study were optimally active between 70 - 80°C and stable over a wide range of pH and temperatures. Substrate specificity and mode of action studies revealed that the Endo-I, β-Glu-I, Exo-HA and Exo-LA from *T. aurantiacus* were highly specific for β-linked glycosidic bonds. In contrast, the Xln-I from *T. aurantiacus* showed a broad substrate specificity and was more active on 4-NP-α-L-arabinopyranoside and 4-NP-α-L-arabinofuranoside than on 4-NPXyl. Furthermore, the β-Glu-I and Xln-I from *T. aurantiacus* catalysed the synthesis of alkyl- and aryl-glycosides of commercial interest. Thus, the major cellulase and xylanase components characterised from *T. aurantiacus* appeared to possess a wide range of potential industrial applications.

Acknowledgements

References

1. Reese, E.T. *Biotechnol. Bioeng. Symp.* **1976**, 6, 9-20.
2. Ryu, D.D.; Mandels, M.; *Enzyme Microb. Technol.* **1980**, 2, 91-101.
3. Coughlan, M.P. *Biochem. Soc. Trans.* **1985**, 405-406.
4. Mandels, M. *Biochem. Soc. Trans.* **1985**, 13, 414-415.
5. Beguin, P; Aubert, J.-P. *FEMS Microbiol. Rev.* **1994**, 13, 25-58.
6. Bhat, M.K.; Bhat, S. *Biotechnol. Adv.* **1997**, 15, 583-620.
7. Parry, N.J.; Kalogiannis, S.; Owen, E.; Beever, D.E.; Bhat, M.K. Poster presented at "PIRA International Conference on Non-Wood Fibres for Industry", Silsoe Research Institute, 1994.
8. Kalogiannis, S.; Owen, E.; Beever, D.E.; Bhat, M.K. *Med. Fac. Landbouww.Univ. Gent.* **1995**, 60/4a, 1995-1999.
9. Kalogiannis, S.; Owen, E.; Beever, D.E.; Claeyssens, M.; Nerinckx, W.; Bhat, M.K. *SCI Lecture papers* **1996**, 78, 1-11.
10. Bhat, M.K.; Parry, N.J.; Kalogiannis, S., Beever, D.E.; Owen, E.; Nerinckx, W.; Claeyssens,M. In: *Carbohydrases from T.reesei and other microorganisms*, Claeyssens, M., Nerinckx, W. & Piens, K. Eds., Royal Society of Chemistry, Cambridge, UK, **1998**, pp. 102- 112.
11. Shepherd, M.G.; Tong, C.C; Cole, A.L. *Biochem. J.* **1981**, 193, 67-74.
12. Tan, L.U.L.; Mayers, P.; Saddler, J.N. *Can. J. Microbiol.* **1987**, 33, 689-692.
13. Khandke, K.M.; Vithayathil, P.J.; Murthy, S.K. *Archiv. Biochem. Biophys.* **1989a**, 274, 491-500.
14. Khandke, K.M.; Vithayathil, P.J.; Murthy, S.K. *Archiv. Biochem. Biophys.* **1989b**, 274, 501-510.
15. Mandels, M.; Sternberg, D. *J. Ferment. Technol.* **1976**, 54, 267-286.
16. Wood, T.M.; Bhat, M.K. In *Methods Enzymol.*, Wood, W.A.; Kellogg, S.T. Eds., Academic Press, London, UK., **1988**, pp. 87-112.
17. Nelson, N. *J. Biol. Chem.* **1952**, 153, 376-380.
18. Bailey, M.J.; Biely, P.; Poutanen, K. *J. Biotechnol.* **1992**, 23, 257-270.
19. Miller, G.K.; Dean, J.; Blum, R. *Arch. Biochem. Biophys.* **1960**, 91, 21-26.
20. van Tilbeurgh, H.; Bhikhabhai, R.; Pettersson, L.G. *FEBS Lett.* **1984**, 169, 215-218.
21. Lugtenberg, B.; Meijers, J.; Peters, R.; Van Der Hoek, P.; Van Alphen, L. *FEBS Lett.* **1975**, 58, 254-258.
22. Dubois, M.; Gilles, K.; Hamilton, J.K.; Rebers, P.A.; Smith, F. *Anal. Chem.* **1956**, 28, 350-356.
23. Goldenberg, D.P.; Creighton, T.E. *Biopolymers*, **1985**, 24, 167-182.
24. Walker, G.J.; Dewar, M.D. *Carbohydr. Res.* **1978**, 39, 303-315.
25. Bhat, M.K.; Hay, A.J.; Claeyssens, M.; Wood, T.M. *Biochem. J.,* **1990**, 266, 371-378.
26. Christakopoulos, P.; Kekos, D.; Macris, B.J.; Claeyssens, M.; Bhat, M.K. *J. Biotechnol.,* **1995**, 39, 85-93.

27   Shinoyama, H.; Takei, K.; Ando, A.; Fujii, T.; Sasaki, M.; Doi, Y.; Yasui, T. *Agric. Biol. Chem.* **1991**, 55, 1679-1681.

28   Paik, Y.H.; Swift, G. *Chem. Ind.* **1995**, 55-59.

29.   Ichikawa, Y.; Look, G.C.; Wong, C.-H. *Anal. Biochem.* **1992**, 202, 215-238.

30.   Zeikus, J.G.; Lee, C.; Lee, Y.-E.; Saha, B.C. In Enzymes in Biomass Conversion, Leatham, G.F.; Himmel, M.E. Eds. ACS Symp. Amercian Chem. Soc., Washington, DC, **1991**, 460, 36-52.

31.   Sudgen, C.; Bhat, M.K. *World J. Microbiol. Biotechnol.,* **1994**, 10, 444-451.

Chapter 13

Two Novel Alkalotolerant Dextranases from *Streptomyces anulatus*

Stephen R. Decker, William S. Adney, Todd B. Vinzant, and Michael E. Himmel

Biotechnology Center for Fuels and Chemicals, National Renewable Energy Laboratory, Golden, CO 80401

Two alkalotolerant dextranases were isolated during a directed screening operation. Soil, water, and biomass samples were collected from the Pawnee National Grasslands in Northeastern Colorado and screened on blue dextran plates for activity. A single colony type was observed that consistently produced large clear halos in the blue dextran indicator, even at pH 8.0. The organism was monocultured and identified by 16S rRNA sequencing as *Streptomyces anulatus*. Growth on dextran yielded two proteins active in hydrolyzing dextran, Dex1 and Dex2. Dex1 has a molecular weight of 63.3 kDa, pH range of pH 5.0 to 9.5 and a temperature optimum of 40°C. Dex2 has a molecular weight of 81.8 kDa, with a pH range of pH 5.0 to 9.5 and a temperature optimum of 50°C. The enzymes retain >50% activity from pH 5.3 to 9.3 after three hours, with 100% retention from 6.0 to 8.5 after three hours. Internal protein sequences from Dex1 indicates a high homology to two dextranases from *Arthrobacter* spp., with a homology of 64.3% to each dextranase. Enzyme production occurs optimally at pH 8.0 in submerged fermentation. Also, three internal peptides coding Dex1 have been determined by Edman degradation sequencing;...NWDNWNAWGPGGNPDPG..., ...GGGPNRAIHTEPRNS..., and ...EIIYFRPGTY... Because extended activity lifetime at broad pH is important for enzymes used in biofilm biomass reduction and control, dextranases with these characteristics are of keen interest to industry.

Dextran is a mainly $\alpha(1\to6)$ glucose polymer containing some minor secondary $\alpha(1\to2)$, $\alpha(1\to3)$, and $\alpha(1\to4)$ linkages. Enzymes degrading these polysaccharides are known as (EC 3.2.1.11) α-1,6-glucan 6-glucanohydrolase (α-1,6 glucanase or

dextranase). Dextrans are produced as an extracellular carbohydrate storage compound mainly by the bacteria *Leuconostoc mesenteroides*, some *Lactobacillus* and some *Streptococcus* species (*1*). Dextran is currently used in such widely divergent products as food additives, cosmetics, blood plasma expanders, and as a chromatography support matrix (*2,3*). Dextranases are important commercial enzymes with applications in the food, chemical, and detergent industries, with fungi being the primary production host (*2*). There is evidence that dextranases can actively prevent the formation of dental caries through the hydrolysis of the α-1,6 linkages in the streptococcal adherence polymer (*4*). Dextranase has also been shown to increase antibiotic efficacy against streptococcal endocarditis (*1*). One major industrial application of dextranases is the reduction of sliming in sugar production. Dextran production by *Leuconostoc* species during processing of sugar cane causes multiple problems, including fouling of filters, reduction in yield, poor crystallization, and low quality of the final product (*5*). Additional applications include cleaning of biofilm biomass from many surfaces, such as contact lenses, surgical implants and food processing equipment and modification of dextran for food additives and support applications in cosmetics and chromatography media. It is reasonable to expect that formulation of cleaning products to contain dextranase enzymes could result in enhanced product performance on these surfaces. For effective performance in cleaners, dextranases must be able to resist the chemical and physical requirements of use. It was thus important to identify new enzymes that could withstand excursions to high pH (at least pH 8.0) and elevated temperature (at least 60°C) for a period of time consistent with exposure expected for liquid cleaner use (at least 15 min).

Currently, there are 29 U.S. Patents pertaining to dextranases. An examination of the literature by NREL disclosed that 5 dextranase enzymes have received U.S. patent protection; (Kofod et al., 5,770,406; Compana et al., 5,637,491; Matsushiro et al., 5,306,639; Day et al., 4,820,640; Tosoni et al. 4,490,470) and that only 6 microbial dextranases are known in the scientific community (SWISSPROT Data Base). The most common application cited for commercial use of dextranases is in drug delivery systems and oral hygiene and dentifrice composition. Other applications include detergent compositions designed to clean surfaces and dextranases used in processing sugar solutions or acting as antimicrobial agents. Although several patents cover some production, purification, or modification of dextranases, none of the dextranase patents include protection of specific dextranases as unique compositions of matter. To our knowledge, no thermal tolerant and alkalophilic dextranases have been reported.

Materials and Methods

Microorganism. *Streptomyces anulatus* was isolated on blue dextran indicator plates from soil sample taken from a chicken coop situated on an abandoned homestead site in the Pawnee National Grasslands. The microorganism was cultured at pH 7.1, 28°C and was identified by 16S rRNA sequence analysis as *Streptomyces anulatus* at MIDI Labs (Newark, DE). Enzyme production occurs optimally at pH 8.0 in submerged fermentation (Figure 1) with dextranase activity closely following cell mass production (Figure 2).
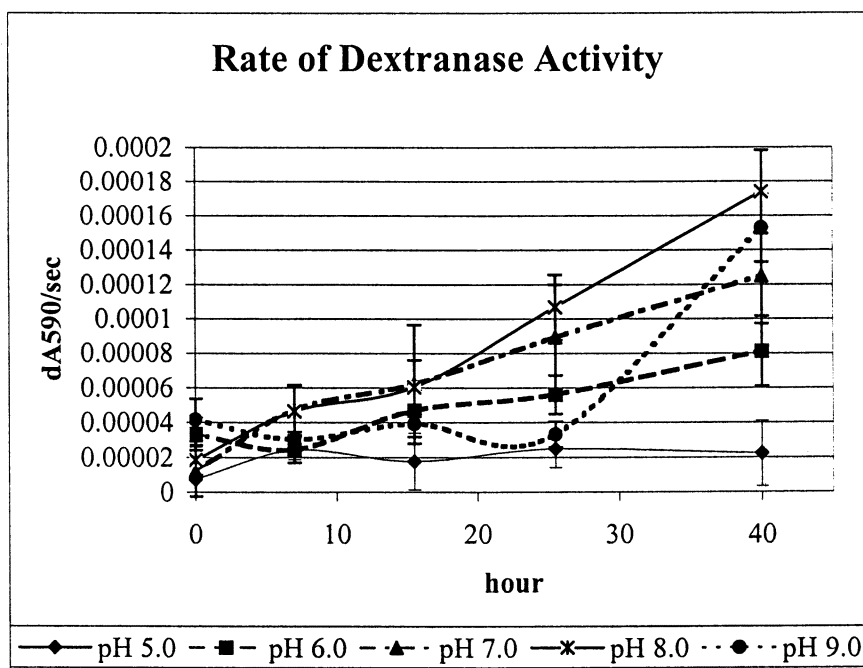
Figure 1. *Dextranase production was measured by plotting the rate of activity of culture broth against time of culture for each of 5 pH cultures. Dextranase activity was determined using a Cary 3 dual beam thermostated spectrophotometer with a stirring multi-cell attachment. Each time point sample was run in triplicate for each pH against a blank with no enzyme. The substrate was 0.025% (w/v) AZCL-Dextran (Megazyme, Wicklow, Ireland). The activity was monitored for dye release at 590nm over 10 minutes.*

Dextranase production 100L
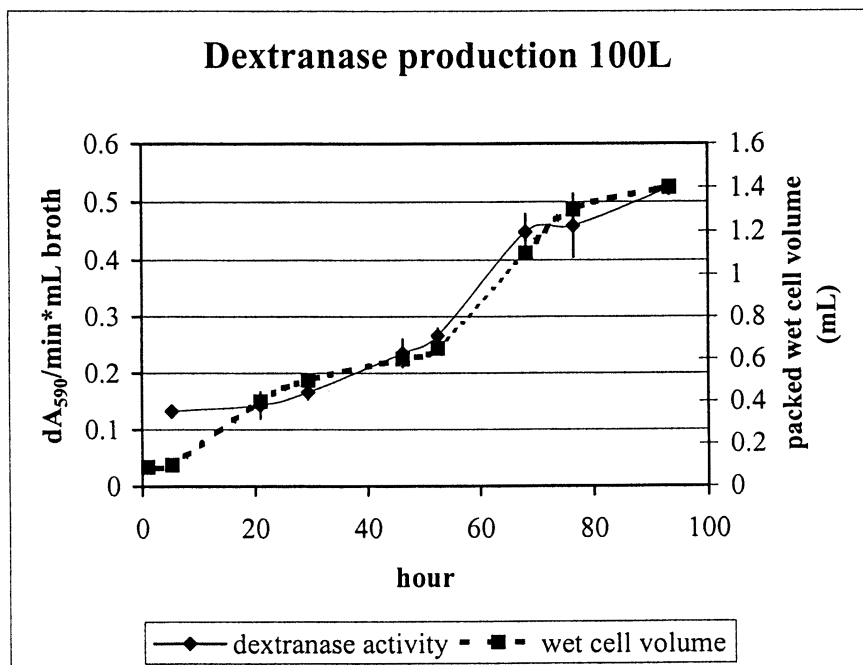
Figure 2. Dextranase production was monitored over the course of using a Cary 3 dual beam thermostated spectrophotometer with a stirring multi-cell attachment. Each sample was run in triplicate for enzyme activity against a blank with no enzyme. The substrate was 0.025% (w/v) AZCL-Dextran (Megazyme, Wicklow, Ireland). The activity was monitored for dye release at 590nm over 10 minutes. Cell mass was measured as packed cell volume of 15mL broth samples centrifuged and measured in graduated 15 mL conical tubes.

Media and Culture Conditions. The basic media consisted of: $(NH_4)_2SO_4$ 11.7 g/L, $MgSO_4 \cdot 7H_2O$ 0.6 g/L, K_2HPO_4 8.7 g/L, NaCl 0.6 g/L, $CaCl_2 \cdot 2H_2O$ 0.05g/L, yeast extract 5.0 g/L, tryptone 5.0 g/L, and trace elements ($FeSO_4 \cdot 7H_2O$ 0.5 g/L, $MnSO_4 \cdot H_2O$ 0.16 g/L, $ZnSO_4 \cdot 7H_2O$ 0.14 g/L, $CoCl_2 \cdot 6H_2O$ 0.37g/L) 1.0 mL/L. Dextran (5.0 g/L) was used as a carbon source/inducer and Blue Dextran (0.05 g/L) was incorporated as an enzyme activity indicator in the initial isolation and screening plates (6). Fermentation optimization for pH was carried out in New Brunswick BioFloIII 1.0 L fermenters. The optimal pH for enzyme production was determined over the range of pH 5.0 to 9.0 and measured by plotting the rate of enzyme activity produced against time of culture growth for each pH measured. Optimum enzyme production temperature was determined by looking at cell growth rates over a temperature range of 20-37°C. For enzyme production, the organism was cultured in 100 mL liquid medium and sequentially transferred to 400 mL, 10 L and finally to a 100 L culture in a 150 L capacity production fermenter (New Brunswick ML4100) which was run at 28°C, 200 rpm, with dO_2 maintained via airflow regulation, and pH control at 7.6. Dextran was fed at 42 and 68 h (100 g and 200 g, respectively), with the culture harvested at 96h.

Activity Measurements and Enzyme Characterization. Enzyme activity was measured using an insoluble dyed, cross-linked dextran (AZCL-Dextran, Megazyme, Inc.) and measuring A_{590} in a Cary3 UV/Vis thermostated spectrophotometer using a 3 mL stirred cell containing 0.25 g/L AZCL-Dextran in 20 mM Tris pH 8.0 at 40°C. The rate of activity was measured over ten minutes as the rate of increase in A_{590} over time. For enzyme pH and temperature optima, all conditions remained the same except for the parameter being measured. To follow activity, 150 μL of 0.3% AZCL-Dextran in 20 mM Tris pH 8.0 was incubated with 50 μL of sample at 40°C in a microtiter plate well. Dye release was determined by visual examination. The molecular weights of the proteins were determined by SDS-PAGE using a Novex system 14% tris-glycine 1.0 mm gel run at 20 mA constant, 110 min under reducing conditions. Sigma molecular weight markers were used as standards in the molecular weight determination. Gels were stained with a Novex Colloidal Coomassie staining kit. Amino acid sequences for Dex1 were determined at the University of Virginia Biomolecular Research Facility. The protein was alkylated and then digested with lysyl peptidase. The peptides were separated on a C18 HPLC column and then sequenced by Edman degradation.

Enzyme purification. One hundred liters of culture broth from *Streptomyces anulatus* was clarified using a Cepa continuous flow centrifuge (ZF41) and concentrated (100 L→13.3 L) with an Amicon DC30 diafiltration unit equipped with a 10 kDa molecular weight cut off hollow fiber cartridge. After further clarification by centrifugation, 50 mL of the concentrated culture broth was diluted with 3 volumes of 20 mM Bis-Tris-Propane, pH 9.0 and loaded onto a 6 mL Pharmacia ReSourceQ column. The unbound flow through was collected and brought to 130 mS/cm with $(NH_4)_2SO_4$ and loaded onto a 20 mL Pharmacia Phenyl Sepharose column and eluted with a complex decreasing $(NH_4)_2SO_4$ gradient (1.0→0.0 M) in 20 mM Tris buffer pH

8.2 over 20 column volumes. Dex1 and Dex2 eluted as distinct peaks at 0.7 M and 0.4 M ammonium sulfate, respectively. The two active peaks were concentrated and further purified on a Pharmacia SuperDex200 size exclusion column.

Results and Discussion

Screening of soil/biomass/water samples from the Pawnee National Grasslands sites gave a multitude of colony types on the variety of media employed, but only a single dextran-hydrolyzing colony type on Blue Dextran plates (Figure 3). *Streptomyces anulatus* grew as the sole organism on one of the Blue Dextran plates inoculated with soil from the homestead chicken coop. It sporulated readily, forming soft, white, powdery spores, and appeared to be a streptomycete under macro- and microscopic examination. The 16s RNA sequence data obtained by MIDI Labs confirmed this. The organism showed 100% homology in the first ~500 bps of the 16s RNA to *Streptomyces anulatus*. The ability of the isolated organism to actively exclude any other organism from the plate has been attributed to the production of puromycin by *S. anulatus*.

The organism was originally isolated at room temperature. Growth at 28°, 30°, and 37°C was also monitored, with growth and enzyme production peaking at 28° or 30°C. The culture pH optimum for enzyme production (Figure 1) was determined to be pH 8.0. Lower pH values yielded significantly less enzyme activity as demonstrated by the comparison of enzyme rates from each pH over time. Interestingly, the culture grown at pH 9.0 showed a long lag phase in getting started, but presented a rapid rise in enzyme production once the cells had adapted to the high pH and started growing. The correlation between cell mass and enzyme production was also indicated in the 100 L culture where enzyme production closely followed biomass generation (Figure 2).

Purification of the enzymes by column chromatography presented a difficult challenge, as the proteins did not bind to cation exchange media at pH 5.0. Also, the proteins did not readily bind to anion exchange columns, except at pH 9.0 under extremely low salt concentrations. Fortunately, the anion exchange column bound most of the other proteins and allowed the unbound dextranases to be separated on a phenyl sepharose HIC column. Final purification was achieved by size-exclusion on a SuperDex200 column. This was carried out only for protein characterization, as the application of dextranases to a dextran-based column was deemed "risky". The SuperDex purified proteins were used to determine molecular weight by SDS-PAGE (Figure 4). For all other enzyme characterization, the HIC fractions were used. The molecular weights were determined to be 63.3 and 81.8 kDa for Dex1 and Dex2 respectively.

The pH optimum for enzyme activity was broad for both enzymes (Figure 5). For Dex1, the active range was about pH 5.1 to 10.1. The range of Dex2 was slightly lower, from pH 4.8 to 9.5. This activity range was measured over 1.5 hours, so it indicates the initial stability and activity of the enzymes. Long-term stability was not measured. The broad range exhibited by both enzymes translates to broad temperature optima, with Dex1 and Dex2 having pH optima of pH 7.0-8.0 and pH 6.0
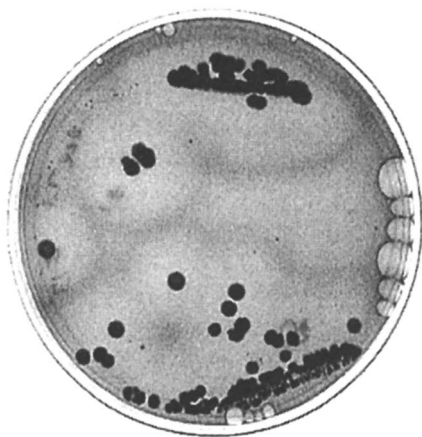
Figure 3. *Blue Dextran agar showing dye build up at the edges of clearing zones around Streptomyces anulatus colonies. Note the absence of any other colony type on this initial isolation plate.*
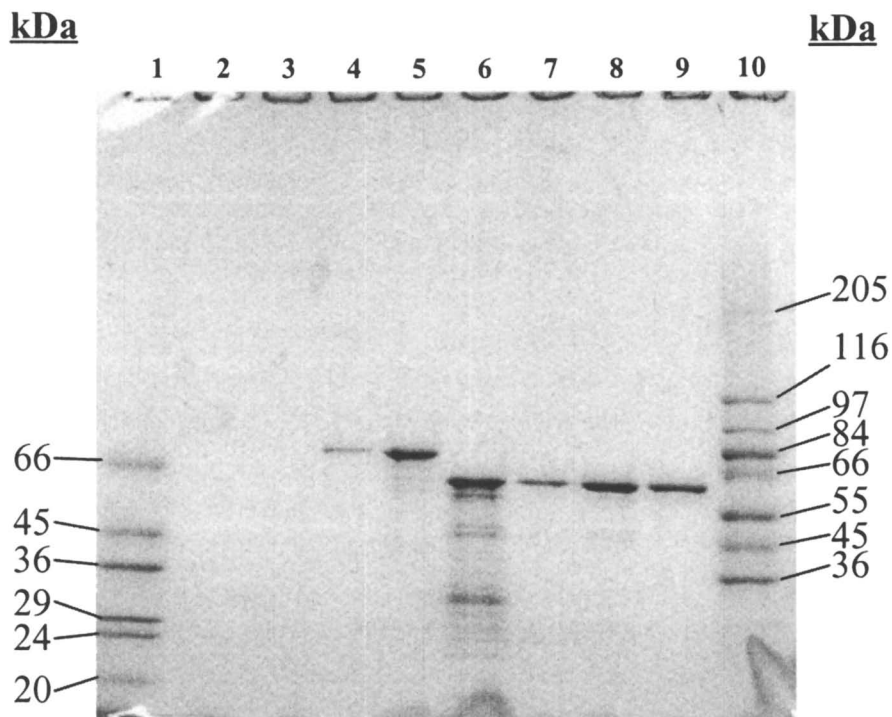
Figure 4. The molecular weights of Dex1 and Dex2 were determined on a Novex 14% Tris-Glycine gel run at 20mA constant for 110 minutes. The gel was stained with Novex Colloidal Coomassie Blue. Lane 1 Sigma Low Range molecular weight markers, Lanes 2-3 empty, Lane 4 and 5 Dex2, Lane 6 partially purified Dex1, Lane 7-9 Dex1, Lane 10 Sigma High Range molecular weight markers.
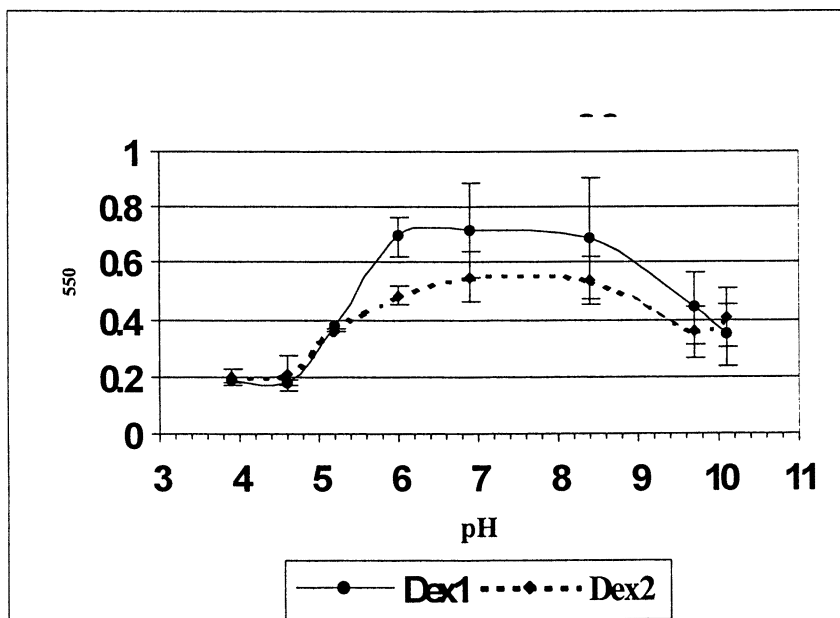
Figure 5. *pH optima and range for Dex1 and Dex2 were determined using a Cary 3 dual beam thermostated spectrophotometer with a stirring multi-cell attachment. Each pH was run in duplicate for each enzyme prep (Dex1 or Dex2) against a blank with no enzyme. The substrate was 0.025% (w/v) AZCL-Dextran (Megazyme, Wicklow, Ireland). The activity was monitored for dye release at 590nm over 10 minutes. The pH range covered 3.9 to 10.1 using a Tris-Acetate-Citrate buffer system adjusted to constant ionic strength with NaCl. Rate of dye release (A_{590}) was plotted against pH.*

to 8.0 respectively. Similarly, temperature optima of the enzymes were also determined (Figure 6). Dex1 has a temperature activity range of 30°-50°C and Dex2 ranges over 35°-75°C. Again, these activities are for initial activity and temperature stability was not measured. The temperature optima for Dex1 and Dex2 were 40° and 50°C, respectively. These numbers are encouraging, as most dextranases exhibit low pH and moderate temperature optima. The pH data is especially important for applications requiring activity at alkaline pH, such as detergent formulations. These data are summarized in Table 1.

Tolerance for alkaline pH is an important characteristic for enzymes incorporated into detergent formulations. Although the pI for the dextranases has not been determined, the binding characteristics on ion exchange chromatography media and the elevated pH tolerance and optima indicate an alkaline pI. This is supported by preliminary IEF data suggesting alkaline isoelectric points for both proteins.

From the protein sequence data, Dex1 apparently belongs to Family 49 of the glycosyl hydrolase families (Table 2). Other members of this group include dextranases from *Penicillium minioluteum*, *Arthrobacter globiformis*, and another *Arthrobacter* species, as well as an isopullulanase from *Aspergillus niger* and a dextran 1,6-a-isomaltotriosidase from *Brevibacterium fuscum* var. *dextranlyticum* (*15*). Pairwise alignment of the Dex1 protein sequence fragments with these known dextranases indicated a 64.3% identity to the two *Arthrobacter* dextranases, placing Dex1 one in Family 49 (*15*). Interestingly, the pairwise percent identity with the *Penicillium* dextranase was only 28.6%, even though both proteins are apparently in the same glycosyl hydrolase family. At first glance, this is unexpected since these families are based on sequence similarities. However, the enzymes involved are from very different sources (bacterial vs. fungal) and have probably been grouped based on highly conserved sequences, such as the catalytic domain. Similarly, homology between the *Arthrobacter* dextranases and the *Penicillium* dextranase is also low, around 35% (*15*).

Conclusions

Two distinct dextranases, Dex1 and Dex2, were isolated from *Streptomyces anulatus*. Dex1 has an optimum pH range of pH 7.0 to 8.0 and a temperature optimum of 40°C. Dex2 had a broader optimum pH range of pH 6.0 to 8.0 and a higher temperature optimum of 50°C. Molecular weights of 63.3 and 81.8 kDa were determined for Dex1 and Dex2, respectively. Enzyme production occurred optimally at 28°C and pH 8.0 and was closely correlated with biomass production. The two proteins, Dex1 and Dex2, are active in depolymerizing dextran over broad pH and temperature ranges, indicating potential usefulness in detergent assisted cleaning and biofilm removal. Although several dextranases are currently known, none have the broad, alkaline pH range of *Streptomyces anulatus* Dex1 and Dex2 (Table 3).

Initial Rates of Dex1 and Dex2
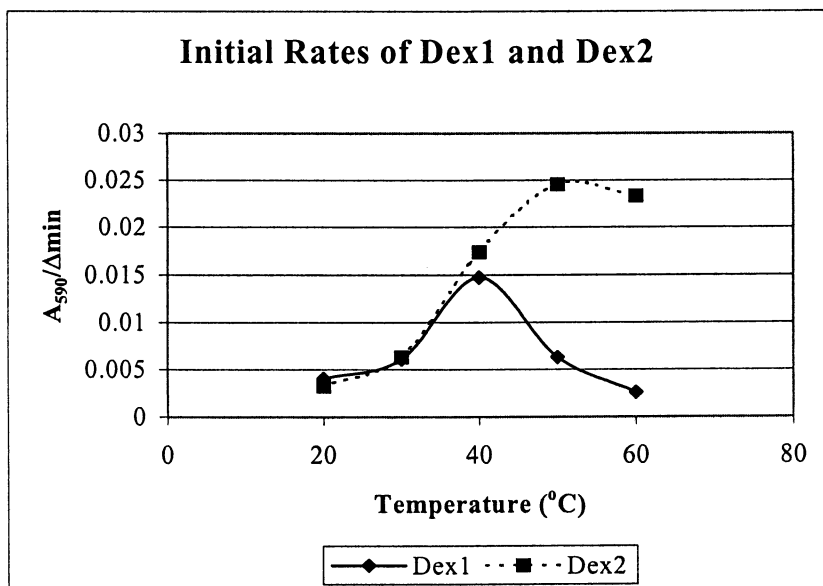
Figure 6. Temperature optima and range for Dex1 and Dex2 were determined using a Cary 3 dual beam thermostated spectrophotometer with a stirring multi-cell attachment. Each temperature was run in duplicate for each enzyme prep (Dex1 or Dex2) against a blank with no enzyme. The substrate was 0.025% (w/v) AZCL-Dextran (Megazyme, Wicklow, Ireland). The activity was monitored for dye release at 590nm over 10 minutes. The temperature range covered 20-60°C. Rate of dye release (A_{590}) was plotted against temperature.

Table 1. Temperature and pH characteristics of Dex1 and Dex2

Enzyme	T°_{opt}	pH_{opt}	T° range[1]	pH range[2]
Dex1	~40°C	7.0-8.0	30-50°C	5.1-10.1
Dex2	~50°C	6.0-8.0	35-75+°C	4.8-9.5+

[1]Range is defined as lower and upper temperature giving 50% of the optimal activity.
[2]Range is defined as lower and upper pH giving 50% of optimal activity.

Table 2. Comparison of Dex1 sequences with Family 49 dextranases

Source	Fragment 2	Fragment 3
A. globiformis	190* EAEGNRPIHTEPRNS	247 EIIYFRPGTY
Arthrobacter. spp.	190 EAAGNRPIHTEPRNS	247 EIIYFRPGTY
P. minioluteum	182 VTSGGSVVGVEPTNA	232 SILYFPPGVY
Dex1	GGGPNRAIHTEPRNS	EIIYFRPGTY

Source	Fragment 1
A. globiformis	629 NWDSWNAWKSAP-----
Arthrobacter. spp.	629 NWENWNAWKSAP-----
P. minioluteum	596 NIDGSYWGEWQIS----
Dex1	NWDNWNAWGPGGNPDPG

*Numbers indicate amino acid position in start of sequence.

Table 3. Previously Described Dextranases

Source Microorganism	Enzyme MW	Optima		Enzyme pI (theoretical)	Reference
		Temp	pH		
Arthrobacter globiformis	68,069			5.19	DEXT_ART GO***
Arthrobacter sp. (Str. CB-8)*	66,645			5.24	DEXT_ART XT***
Penicillium minioluteum **	62,510			5.03	DEXT_PEN MI***
Streptococcus downei	140,089			4.35	DEXT_STR DO***
Streptococcus mutans	94,538			4.94	DEXT_STR MU***
Streptococcus salivarius	84,087			4.52	DEXT_STR SL ***
Penicillium notatum D-1	55,800	50	4.8	4.9 (experimental)	(2)
Penicillium notatum D-2	50,100			4.75 (experimental)	(2)
Lipomyces starkeyi		55	5.5		(7)
Penicillium lilacinum	26,500				(8)
Thermoanaerobacter spp. AB11Ad		70	5.5		(9)
Rt364	140,000	80	5.5		(10)
Streptococcus suis	62,000				(11)
Sporothrix schenckii	79,000		5.0		(12)
Streptococcus sobrinus (*E. coli* expressed)	160,000 - 260,000	39	5.3	3.56, 3.6, 3.7 (experimental)	(13)
Chaetomium gracile		55	~5.5		(9)
Chaetomium spirale			~5.0		(14)
Gibberella fujikuroi			~5.0		(14)
Humicola grisea			~5.0		(14)
Sporotrichum asteroides			~5.0		(14)
Penicillium roquefortii			~5.0		(14)
Aspergillus fumigatus			~5.0		(14)
Streptomyces cinnamonensis			~5.0		(14)
Trichoderma harzianum	ca. 50,000	30-40	5.0	5.60 (experimental)	U.S. Patent #5,770,406

U.S. Patents *# 5,306,639 and ** #5,637,491.   *** SWISS-PROT Data Base

Acknowledgements

References

1. Płeszczyńska, M.; Szczodrak, J.; Rogalski, J.; and Fierdurek J. *Mycol. Res.* **1997**, *101*, 69-72.
2. Płeszczyńska, M.; Rogalski, J.; Szczodrak, J.; and Fierdurek J. *Mycol. Res.* **1996**, *100*, 681-686.
3. Iwai, A.; Ito, H.; Mizuno, T.; Mori, H.; Matsui, H.; Honma, M.; Okada, G.; and Chiba, S. *J. Bacteriol.* **1994**, *176*, 7730-7734.
4. Ohnishi, Y.; Kubo, S.; Ono, Y.; Nozaki, M.; Gonda, Y.; Okano, H.; Matsuya, T.; Matsishiro, A.; and Morita, T. *Gene*, **1995**, *156*, 93-96.
5. Garcia, B.; Margolles, E.; Roca, H.; Mateu, D.; Raices, M.; Gonzoles, M. E.; Herrera, L.; and Delgado, J. *FEMS Microbiol. Lett.* **1996**, 143, 175-183.
6. Wynter, C. *J. Gen. Appl. Microbiol.* **1996**, *42*, 213-223.
7. Kim, D.; and Day, D. F. *Lett. Appl. Microbiol.* **1995**, *20*, 268-270.
8. Das, D. K.; and Dutta, S. K. *Internat. J. Biochem. Cell Biol.* **1996**, *28*, 107-113.
9. Wynter, C.; Patel, B. K. C.; Bain, P, de Jersey, J.; Hamilton, S.; and Inkerman, P.A. *FEMS Microbiol. Lett.* **1996**, *140*, 271-276.
10. Wynter, C. V. A.; Chang, M.; de Jersey, J.; Patel, B.; Inkerman, P.A.; and Hamilton, S. *Enz. Microbial Tech.* **1997**, *20*, 242-247.
11. Serhir, B.; Dugourd, D.; Jacques, M.; Higgens, R.; and Harel, J. *Gene.* **1997**, *190*, 257-261.
12. Arnold, W. N.; Nguyen, T. B. P.; and Mann, L. C. *Arch. Microbiol.* **1998**, *170*, 91-98.
13. Wanda, S. Y.; and Curtiss, R. III.; *J. Bacteriol.* **1994**, *176*, 3839-3850.
14. Hattori, A.; and Ishibashi, K. *Agric. Biol. Chem.* **1981**, *10*, 2347-2349.
15. Coutinho, P.M. & Henrissat, B. (1999) Carbohydrate-Active Enzymes server at URL: http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html

Chapter 14

Microbial Mannanases: Substrates, Production, and Applications

G. M. Gübitz[1], A. Sachslehner[2], and D. Haltrich[2]

[1]Department of Environmental Biotechnology, Graz University
of Technology, Petersgasse 12, A–8010 Graz, Austria
[2]Department of Food Technology, University of Agricultural Sciences Vienna,
Muthgasse 18, A–1190 Wien, Austria

Mannans and xylans are the main components of woods beside cellulose and lignin and thus they belong to the most abundant polymers in the biosphere. Biodegradation of β-mannans occurs by the synergistic action of endo-1,4-β-mannanases, β-D-mannosidases, β-D-glucosidases, α-D-galactosidases and acetyl mannan esterases. Mannan degrading enzymes are widespread in nature and play an important role in the hydrolysis of hemicellulose containing biomass. They are produced by various microorganisms, including bacteria, yeasts and fungi, and also occur in germinating seeds of terrestrial plants, marine algae and animals. Mannanase activities secreted by microorganisms into the extracellular environment vary between 0.02 IU ml^{-1} and 800 IU ml^{-1}. Mannanases cleave polymeric mannans as well as mannooligosaccharides, usually with a degree of polymerization greater than three to yield mannobiose, mannotriose and various mixed oligosaccharides. The activities of the individual enzymes can vary depending on the structure and substitution pattern of the heteromannans and many mannanases show transglycosylation activity. Mannan degrading enzymes have found applications including the food, feed, pharmaceutical, and pulp/paper industry as well as for oil and gas well stimulation.

The estimate for the world market of enzymes in 2005 is in the range of 1.7 and 2 billion US$ (*1*). At least 75% of all industrial enzymes are hydrolases, which are used for the degradation of natural substances. Carbohydrases including cellulases and hemicellulases represent the second largest group of commercially produced

enzymes (*1*). The most abundant source of carbohydrates in the world is derived from plant biomass, which represents the lignocellulosic materials that comprise the cell wall of all plants. Lignocellulose consists of three main polymeric constituents: cellulose (β-1,4-glucan), lignin (a complex polyphenolic polymer), and hemicellulose. Hemicelluloses are classified according to the sugars composing the main chain in D-xylans, D-mannans, D-galactans, and L-arabans. In recent years there has been considerable interest in the utilization of plant material as a renewable source for fermentable sugars and the subsequent conversion of these sugars into liquid fuels, solvents, other chemicals, or single-cell protein (*1a*). Since hemicelluloses comprise up to 35% of the dry weight of plant materials, either from primary sources or from agricultural, forest, industrial as well as from domestic wastes and residues, they are among the most abundant organic materials in existence.

Traditionally, the application of hemicellulases has been viewed from the perspective of the bioconversion of lignocellulosic materials. Since it was found in the mid-1980s that xylanases can improve the bleachability of pulps, the use of enzymes in the pulp and paper industry has also grown rapidly. Additionally, the application of mannanases and xylanases in the food and feed industry facilitating food processing and improving nutrition value received a lot of interest (*2*). This review focuses on the production and applications of mannan degrading enzymes, a class of useful enzymes which has been relatively neglected until the last few years compared to cellulases and xylanases.

The Substrates: Mannans

Structure and Occurrence. β-1,4-Mannans are substituted heteropolysaccharides and are widespread in hardwood and softwood, as well as in plant seeds and tubers. Depending on the monomers in the backbone and the side chain constituents they can be divided into pure mannans, glucomannans, galactomannans and galactoglucomannans.

Pure Mannans. Pure mannans are defined as polysaccharides, which contain less than 10% non-mannose sugar residues (*3*). They are commonly found in the endosperm of palm seeds, in the seeds of umbelliferous species and in green coffee beans. Cell wall mannans in the seed endosperms of the Palmae and Umbelliferae are storage polysaccharides, which disappear during germination. The best characterized pure mannans of palm seeds are those from date (*Phoenix dactylifera*) and ivory nut (*Phytelephas macrocarpa*). Both seeds contain mannans differing in their solubility in cuprammonium solution and in aqueous alkali, in their degree of polymerization, and in the content of galactose and glucose (*3*). Coffee mannan has a relatively low molecular weight (7 kDa) and contains on the average only one galactose residue attached at the C-6 position per 100 mannose residues (*4*).

Glucomannans. Glucomannans occur in hardwoods, as storage polysaccharides in seeds of certain species of Liliaceae and Iridaceae, and in tubers of various species of Amorphophallus. Structurally they differ from the pure mannans in that a large proportion of the D-mannose residues in the β-1,4-linked backbone are replaced by D-glucose residues. Hardwoods consist of 3-5% glucomannan with a glucose to mannose ratio of 1:2. In some species the ratio 1:1 seems to predominate. Galactosyl substituents are not present in hardwood mannans (5). The storage glucomannans of seeds are formed of a backbone consisting of randomly distributed D-mannose and D-glucose residues and small amounts (3-6%) of D-galactose residues α-1,6 linked to the main chain (3). Glucomannan from *Amorphophallus konjac* is composed of glucose and mannose subunits at a molar ratio of 0.66 substituted with acetyl groups, which are located on average, on every tenth sugar unit, and has a molecular weight of about 300 kDa (6).

Galactomannans. Galactomannans occur predominantly in the endosperm of leguminous seeds, and products from several species are important as industrial raw materials. Galactomannans have also been found in immature palm seeds and in some non-leguminous species (3). Seed galactomannans are formed of a backbone consisting of mannose residues, which carry numerous α-1,6-D-galactopyranosyl side chains. The degree of galactose substitution varies in a wide range from about 20% to nearly 100%, which is taxonomically significant within the Leguminosae. The deposited galactomannans are mobilized during germination and the embryo consumes the breakdown products. Galactomannans are hydrophilic and are usually obtained from crushed seeds or endosperms with hot water extraction (3).

Locust bean gum (LBG, carob gum) is obtained from the ground endosperm of *Ceratonia siliqua*, a leguminous tree that grows in the Mediterranean countries. The LBG molecule has an average molecular weight of 350 kDa and consists of a straight mannose chain, which carries galactose residues at about every fourth mannose monomer. The seeds of *Cyamopsis tetragonoloba* contain a galactomannan (guar gum) with a molecular weight of 800-1000 kDa and a mannose to galactose ratio of 1.6:1 (6). Like LBG, guar gum is an industrial raw material of which thousands of tons are produced annually. In addition, the galactomannan of the tara bush (*Caesalpinia spinosa*) with a galactose to mannose ratio of 1:3 also has commercial importance (6).

Galactoglucomannans. Hetero-1,4-β-D-mannans consisting of mannose and glucose in the main chain are most abundant in woods and annual plants (112). They are the major hemicelluloses in softwoods where galactoglucomannans account for approx. 15-20% of the total dry weight. The predominant (water-soluble) polysaccharide consisting of mannose:glucose:galactose residues in the ratio of 3:1:1 also exists. A less soluble polysaccharide consists of residues of mannose:glucose:galactose in the ratio of 3:1:0.1. Both mannans are composed of a main chain of randomly distributed β-1,4-linked D-glucopyranose and mannopyranose residues. Some of the hexose units carry a terminal residue of α-D-

galactopyranose attached to C-6. Between 5.9% to 8.8% of the mannose and glucose residues can be partially substituted by acetyl groups in C-2 and in C-3 positions (*5,5a*). The structure of an acetylated galactoglucomannan from softwood is shown in Figure 1. In the seed endosperms of some plants like *Cercis siliquaestrum* and *Bauhinia sp.* galactoglucomannans are also known to serve as storage polysaccharides (*7*).

Application of Mannans. Based on their rheological properties, different mannans have a long list of applications in the food and animal feed industry as well as in the non-food industry. As an example, they are used for oil drilling muds, textile printing, warp sizing, paper coating, carpet printing, tertiary oil recovering, and mining flotation (*8*). The consumption of guar gum in the United States was about 65,000 tons per year in the mid-eighties of which approximately 75% were used in the oil services business (*8*). Guar gum produces the highest viscosity among the natural, commercial gums, with a 1% solution having a viscosity of approx. 6,000-10,000 mPa.s (*9*). LBG, guar gum, and tara gum are used as emulsifiers of salts in bakery and baking mixes, as firming agents in fats and oil, and as thickeners and stabilizers in soups, sauces, jams, jellies, toppings, syrups, vegetable juices and many other food categories (*10*). LBG and tara gum act synergistically with kappa-carrageenan and xanthan gum, forming gels that are elastic, very cohesive, and relatively free from syneresis. Glucomannan from *Amorphophallus konjac* is a soluble dietary fiber and has been used as a source of food since ancient time in China and Japan. Despite their inclusion into many food preparations, little is known about the digestion of these mannans within the human digestive system. Certain human intestinal tract-inhabiting bacteria such as *Aerobacter mannanolyticus* are known to produce enzymes capable of degrading konjac glucomannan (*11*). Recently, an anaerobic human intestinal bacterium of the *Clostridium butyricum-Clostridium beijerinckii* group capable of degrading konjac glucomannan was isolated (*12*). Several studies have indicated that glucomannan has the ability to lower serum cholesterol and may lower serum triglyceride and bile acid level as well. This hypocholesterolaemic activity of mannans was shown to be associated with certain structural and physical features (high molecular weight and viscosity, water solubility). Furthermore, the activity was lost when the carbohydrate was subjected to acid or enzymatic hydrolysis (*11*). Glucomannan may also have an influence on glucose tolerance and glucose absorption.

Galactomannans are used in industry as biodegradable drilling fluids. Lubricity under high shear, a typical feature of the high-molecular-weight polysaccharides which behave strongly non-Newtonian, reduces pump power required. Additionally, galactomannans are applied for the stimulation of wells. Oil and gas wells produce additional petroleum product after proper stimulation (*137*). In the paper making process, mannans act as an inter-fiber adhesive and strength building agent. Bonding strength increases by 10-20% by the addition of guar gum. Sheet uniformity is increased and printability is improved *(for a review see 6)*. Mannans are also used as a binding agent for balling, pelletizing and briqetting of fine materials to produce

α-D-Gal*p*   α-D-Gal*p*
1   1
↓   ↓
6   6

→ 4)-D-Glc*p*-β-(1→ 4)-D-Man*p*-β-(1→4)-D-Man*p*-β-(1→4)-D-Glc*p*-β-(1→ 4)-D-Man*p*-β-(1→

                    2   3
                    ↑   ↑
               Acetyl   Acetyl

Figure 1. Structure of galactoglucomannan from softwood.

large, dust free agglomerates (*13*). The water proofing, thickening and foam stabilizing properties of guar gum are required in the preparation of explosives (*13*).

Mannanolytic Enzymes

Biodegradation of β-mannans occurs by the synergistic action of endo-1,4-β-mannanases, β-D-mannosidases, β-D-glucosidases, α-D-galactosidase and acetyl mannan esterases (Figure 2). Mannan degrading enzymes are widespread in nature and play an important role in the hydrolysis of hemicellulose.

Mannanases: Production. Endomannanases are produced by various microorganisms, including bacteria, yeasts and fungi, and also occur in germinating seeds of terrestrial plants, marine algae and animals (*14*). Tables 1 and 2 summarize the production of fungal and bacterial mannanases. Mannanase activities measured in the culture fluids of these microorganisms varied greatly, ranging from 0.016 IU ml^{-1} for *Aspergillus tamarii* (*15*) to 786 IU ml^{-1} for *Sclerotium rolfsii* (*16*). In accordance with reports on other endoglycanases involved in the degradation of lignocellulose, mannanase production is inducible in microbial cells. In addition, their expression is typically controlled by carbon repression when more readily metabolizable carbon sources (e.g., glucose) are present in the culture medium together with an inducing substrate (*17*). In agreement with this mechanism of induction and carbon control, appropriate inducing carbon sources (e.g., various mannans or mannan-rich lignocellulosic material) have been reported to be employed for an efficient production of mannanases by different organisms. All bacterial strains listed in table 2 produced highest mannanase activity when mannans were added to the culture media, while some fungi produced considerable amounts of mannanase even when grown on various cellulosic substrates as sole carbon source. However, it cannot be excluded that these complex media contained traces of mannans. Since the various mannans are large polymers that cannot penetrate the plasma membrane, it can be anticipated that mannanases are induced by some mannan fragments formed by the action of very low levels of extracellular hydrolases that are produced even under noninducing conditions (*18*). Soluble, low-molecular weight sugars that have been identified to provoke β-mannanase formation in various microorganisms include mannose (*19*), mannobiose and mannotriose (*18*), or cellobiose as well as its positional isomers sophorose, laminaribiose and gentiobiose (*20*). Furthermore, the synthetic glycoside methyl β-D-mannopyranoside has been identified as a nonmetabolizable, gratuitous inducer of mannanase activity in the yeast *Aureobasidium pullulans* (*18*). While the presence of suitable inducing compounds seems to be a prerequisite for mannanase synthesis in most microorganisms studied to date, and the absence of repressing carbon sources is generally not sufficient to trigger endoglycanase synthesis, the basidiomycete *Sclerotium (Athelia) rolfsii* produces significant mannanase activities under noninduced, derepressed conditions (i.e., when glucose, which as a typical repressing

α-D-Gal *p*
1
↓
6
→ 4)-D-Glc *p*-β-(1 → 4)-D-Man *p*-β-(1 → 4)-D-Man *p*-β-(1 → 4)-D-Man *p*-β-(1 → 4)-D-Man *p*-β-(1 → 4)-D-Glc *p*-β-(1 → 4)-D-Man *p*-β-(1

α-D-Gal *p*
1
↓
6

2
↑
Acetyl

3
↑
Acetyl

D-Man *p*-β-(1 → 4)-D-Man *p*-β-(1 →

Endo-β-(1→4)-D-Mannanase (EC 3.2.1.78)

β-Mannosidase (EC 3.2.1.25)

Acetylesterase (EC 3.1.1.6) or Acetylglucomannanesterase

α-Galactosidase (EC 3.2.1.22)
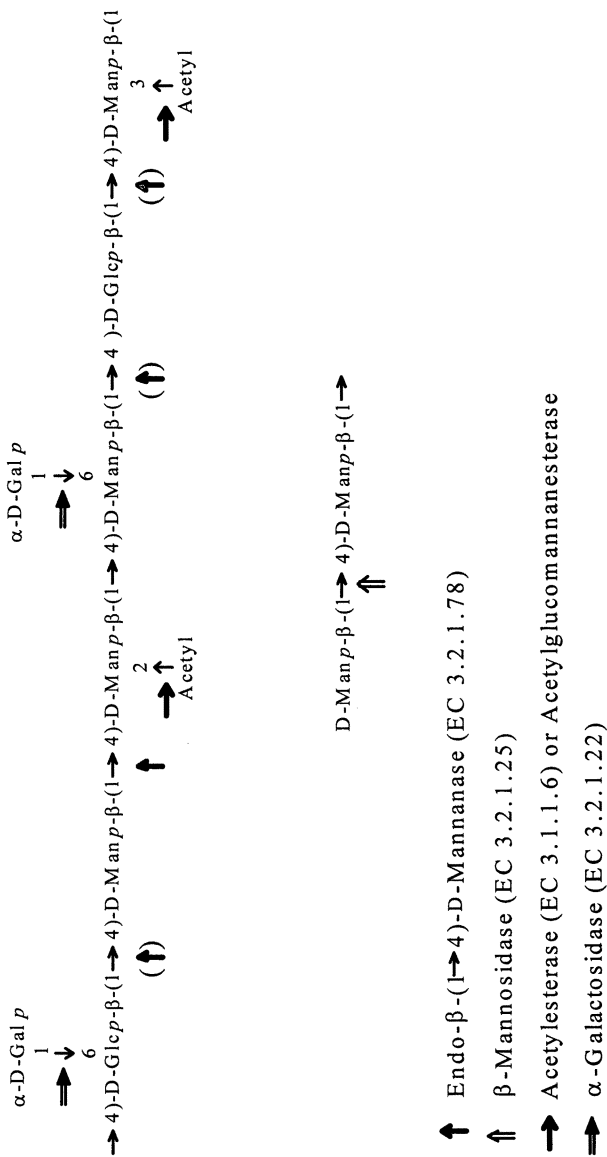
Figure 2. A hypothetical softwood mannan and the sites of possible attacks by mannanolytic enzymes.

Table 1. Production of mannanases by fungi and yeasts.

Source Microorganism	Cultivation conditions		Activity		Ref.
	Substrate, g l^{-1}		^{a}IU ml^{-1}	^{b}IU l^{-1} h^{-1}	
Aspergillus awamori VTT-D-75028	LBG^{d}, 10 wheat bran, 30	28°C	2.04 SN	SF n.a.	(65)
Aspergillus giganteus QM 7974	mannan, 5	29°C 14-21 d	12.9 DNS	SF n.a.	(66)
Aspergillus luchuensis QM 813	mannan, 5	29°C 14-21 d	2.4 DNS	SF n.a.	(66)
Aspergillus nidulans v. echinulatus QM 1909	mannan, 5	29°C 14-21 d	1.8 DNS	SF n.a.	(66)
Aspergillus niger NRRL 337	LBG, 10	45°C 48 h	5.2 DNS	SF 108	(19)
Aspergillus niger ATCC 46890	LBG, 20	28°C 168 h	5.4 DNS	SF 32.1	(67)
Aspergillus tamarii IP 1017-10	white-clover gal.-mannan, 1	29°C 144 h	0.016 SN	SF 0.11	(15)
Epicoccum yuccae QM 284e	mannan, 5	29°C 14-21 d	1.1 DNS	SF n.a.	(66)
Fusicoccum sp. QM 2045	mannan, 5	29°C 14-21 d	6.9 DNS	SF n.a.	(66)
Glionatix trabeum NRCC 5907	wheat bran, 10	30°C 168 h	0.19 SN	SF 1.1	(68)
Haematosterium sanguinolentum NRCC 5902	wheat bran, 10	30°C 168 h	0.39 SN	SF 2.3	(68)
Lenzites saepiaria NRCC 5910	wheat bran, 10	30°C 168 h	0.52 SN	SF 3.1	(68)
Memnoniella echinata QM 6859	mannan, 5	29°C 14-21 d	0.92 DNS	SF n.a.	(66)
Paecilomyces varioti QM 3408	mannan, 5	29°C 14-21 d	12.2 DNS	SF n.a.	(66)
Penicillium expansum QM 6758	mannan, 5	29°C 14-21 d	1.66 DNS	SF n.a.	(66)
Penicillium funiculosum QM 474	mannan, 5	29°C 14-21 d	3.86 DNS	SF n.a.	(66)
Penicillium isariiforme QM 1897	mannan, 5	29°C 14-21 d	2.12 DNS	SF n.a.	(66)
Penicillium kloeckeri NRRL 1017	gal.-mannan, 10	26-28°C 168 h	55.0 DNS	SF 327	(69)

Table 1 continued. Production of mannanases by fungi and yeasts.

Source Microorganism	Cultivation conditions		Activity		Ref.
	Substrate g l^{-1}		[a]IU ml^{-1}	[b]IU l^{-1}h^{-1}	
Penicillium ochrochloron QM 477	mannan, 5	29°C 14-21 d	1.9 DNS	SF n.a.	(66)
Penicillium piscarium QM 471	mannan, 5	29°C 14-21 d	1.9 DNS	SF n.a.	(66)
Penicillium purpurogenum 618	copra meal, 40	35°C 124 h	9.5 SN	SF 76.6	(70)
Penicillium verruculosum QM 7999	mannan, 5	29°C 14-21 d	4.6 DNS	SF n.a.	(66)
Penicillium wortmanni QM 1859	mannan, 5	29°C 14-21 d	29.4 DNS	SF n.a.	(66)
Polyporus versicolor NRCC 5909	avicel, 10	30°C 192 h	1.92 SN	SF 10.0	(68)
Poria placenta NRCC 5909	wheat bran, 10	30°C 168 h	0.33 SN	SF 2.0	(68)
Sclerotium rolfsii ATCC 200224	cellulose, 37.6	30°C 190 h	150 DNS	SF 789	(71)
Sclerotium rolfsii ATCC 200224	KGM[e], 2.4, α-cell[c], 37.6	30°C 336 h	202 DNS	SF 601	(72)
Sclerotium rolfsii CBS 191.62	α-cell[c], 42.6	30°C 210 h	786 DNS	15l-F 3740	(16)
Sclerotium rolfsii CBS 191.62	glucose, 42.6	30°C 240 h	240 DNS	15l-F 1000	(21)
Schizophyllum commune NRCC 5911	avicel, 10	30°C 192 h	1.32 DNS	SF 6.9	(68)
Schizophyllum commune BT 2115	avicel, 73.4	30°C 200 h	23.3 DNS	10l-F 117	(73)
Schizophyllum commune BT 2115	LBG, 2.4, avicel, 37.6	30°C 168 h	32.8 DNS	SF 195	(72)
Sporotrichum cellulo-philum ATCC 20493	LBG, 10	45°C 48 h	4.8 DNS	SF 100	(19)
Sporotrichum cellulo-philum ATCC 20493	LBG, 5	45°C 36 h	6.0 DNS	15l-F 167	(75)
Talaromyces byssochla-mydoides NRRL 3658	LBG, 10	45°C 48 h	52.4 DNS	SF 1092	(19)
Talaromyces emersonii NRRL 3221	LBG, 10	45°C 48 h	43.1 DNS	SF 898	(19)
Talaromyces emersonii ATCC 18080	LBG, 10	45°C 48 h	0.6 DNS	SF 12.5	(19)

Table 1 continued. Production of mannanases by fungi and yeasts.

Source Microorganism	Cultivation conditions		Activity		Ref.
	Substrate g l⁻¹		[a]IU ml⁻¹	[b]IU l⁻¹ h⁻¹	
Thermoascucs auran-tiacus ATCC 26904	LBG, 10	45°C 48 h	0.1 DNS	SF 2.1	(*19*)
Thermomyces lanuginosus DSM 5826	KGM[e], 9.4, xylan, 21.6	50°C 168h	1.8 DNS	SF 10.7	(*72*)
Thermomyces lanuginosus IMI158749	LBG, 20	50°C 96 h	30.0 SN	SF 313	(*76*)
Thielavia terrestris NRRL 8126	LBG, 10	45°C 48 h	12.5 DNS	SF 260	(*19*)
Thielavia terrestris ATCC 26917	LBG, 10	45°C 48 h	2.8 DNS	SF 58	(*19*)
Trichocladium candense NRCC 5903	wheat bran, 10	30°C 168 h	0.29 SN	SF 1.7	(*68*)
Trichoderma harzianum E 58	avicel, 10	28°C 192 h	0.66 SN	SF 3.4	(*27*)
Trichoderma reesei RUT C-30	cellulose	30°C 100 h	0.34 SN	SF 3.4	(*77*)
Trichoderma reesei QM 9414	LBG, 10, wheat bran, 30	28°C	1.44 SN	SF n.a.	(*65*)
Trichosporon cutaneum JCM 2947	konjac powder, 10	25°C 48 h	0.12 SN	SF 2.5	(*78*)
Trichurus spiralis QM 834	mannan, 5	29°C 14-21 d	6.4 SN	SF n.a.	(*66*)

[a]Maximum activity, quantification of reducing sugars with dinitrosalicylic acid (DNS), Somogy-Nelson (SN), or determination of mannanase activity by hydrolysis of Azo-carob galactomannan (azo-LBG).
[b]Productivity obtained with shake flask cultures (SF) or fermentation (F).
[c]α-cell: α-cellulose.
[d]LBG: locust bean gum.
[e]KGM: konjac glucomannan.
n.d.: not detectable. n.a.: not available.

Table 2. Production of mannanases by bacteria.

Source Microorganism	Cultivation conditions		Activity		Ref.
	Substrate, g l⁻¹		[a]IU ml⁻¹	[b]IU l⁻¹ h⁻¹	
Bacillus pumilis ATCC 72	LBG[d], 5	45°C 22 h	34.0 DNS	14l-F 1545	*(79)*
Bacillus sp. W-2	KGM[e], 10	37°C 24 h	4.57 SN	SF 190	*(80)*
Bacillus sp. AM-001	copra meal, 10	37°C 42 h	30.1 SN	SF 717	*(81)*
Bacillus sp.	copra meal, 10	30°C 48 h	1.6 SN	SF 33.3	*(82)*
Bacillus stearo-thermophilus ATCC 266	LBG, 10	3-1 48 h	0.9 DNS	31-F 18.8	*(48)*
Bacillus subtilis 168	palm seed powder, 10	37°C 24 h	102.0 DNS	SF 4250	*(22)*
Bacillus subtilis 5H	LBG, 10	45°C 24 h	7.8 SN	SF 325	*(83)*
Bacillus subtilis ATCC 12711	LBG, 5 wheat bran, 30	24 h	16.0 SN	F 667	*(65)*
Bacillus subtilis NM-39	LBG, 10	24 h	0.007 SN	18l-F 0.30	*(84)*
Bacillus subtilis KK 01	copra meal, 10	30°C 48 h	2.0 SN	SF 41.7	*(55)*
Clostridium butyricum-Clostridium beijerinckii	KGM, 2	37°C 24 h	0.097 SN	4.0	*(12)*
Enterococcus casseliflavus FL 2121	KGM, 10	37°C 24 h	0.61 DNS	SF 25.4	*(85)*
Flavobacterium sp.	KGM 10	4°C 144 h	3.8 SN	SF 26.4	*(86)*
Rhodothermus marinus ITI 376	LBG, 9.86	61°C 96 h	29.6 DNS	SF 308.3	*(87)*
Streptomyces olivochro-mogenes ATCC 21713	LBG, 10 wheat bran, 30	28°C 24 h	2.7 SN	SF 112.5	*(65)*
Thermotoga neapolitana 5068	guar gum, 0.3 lactose, 5	85°C	0.26 azoLBG	8l-F n.a.	*(88)*
Vibrio sp. MA-138	konjac powder, 5	25°C 24 h	0.30 SN	SF 12.5	*(89)*

[a]Maximum activity, quantification of reducing sugars with dinitrosalicylic acid (DNS), Somogy-Nelson (SN), or determination of mannanase activity by hydrolysis of Azo-carob galactomannan (azo-LBG). [b]Productivity obtained with shake flask cultures (SF) or fermentation (F). [c]α-cell: α-cellulose. [d]LBG: locust bean gum. [e]KGM: konjac glucomannan. n.d.: not detectable. n.a.: not available.

substrate was used as the sole carbon source for growth, had been consumed from the medium (*16*). Since mannanase formation in *S. rolfsii* was only repressed when glucose was present above a certain critical concentration, slow feeding of this substrate resulted in a considerable increase (more than 90%) of mannanase formation compared to a batch cultivation on glucose (*16*).

In general, much higher mannanase activities were found in the culture filtrates of fungi than in those from bacteria. However, considering longer incubation times for fungi, similar overall mannanase productivities were reached with both bacteria and fungi. Highest mannanase productivities were reported for *Bacillus subtilis* grown in shake flasks for one day [4250 IU ml^{-1} h^{-1}, (*22*)] and *S. rolfsii* cultivated in a 15 l fermenter for 8 days [3740 IU ml^{-1} h^{-1}, (*16*)].

Mannanase Multiplicity. Fungal hemicellulases are often secreted as multiple enzymes into the extracellular environment. Although very little is known about the origin of mannanase multiplicity, it can be suggested that the function of multiple enzymes is similar to those of cellulases and xylanases. This multiplicity might result from the requirement to bind to, and degrade substrates of varying complexity, as was proposed by Johnson and Ross (*24*). It is not clear yet whether the apparent isoenzymes are the products of multiple genes or the result of the modification of a single enzyme (*25*). Mannanase heterogeneity may arise from post-translational modifications, such as differential glycosylation or proteolysis or both. It is possible that some of the mannanases detected in culture filtrates are precursors or degradative products of other mannanases. Some organisms secrete proteases at some point during growth which may cause extracellular post-secretory modifications. This increasing multiplicity as a function of culture age has been demonstrated for cellulases in the case of *Trichoderma reesei* or *Schizophyllum commune* (*26*). Isoforms of mannanase activity have been reported for *Trichoderma harzianum* (*27*), *Sclerotium rolfsii* (*28*), *Polyporus versicolor* (*24*), *Trichoderma reesei* (*29*), and *Thielavia terrestris* (*30*). A *T. reesei* β-mannanase gene was cloned and expressed in *Saccharomyces cerevisiae* (*31*) and two mannanase isoforms have been detected in the resulting culture fluid of the yeast indicating that one mannanase gene encodes the two mannanases. The slight difference in molecular mass between the two forms could be accounted for by differences in glycosylation. In culture filtrates of *Cellulomonas fimi* grown on locust bean gum, three proteins with endomannanase activity were detected by zymogram analysis (*32*). When cloning the mannanase gene *man26a* in *E. coli*, the active enzyme produced correlated to the largest mannanase (approx. 100 kDa) found in the wild-type enzyme preparation. However, after treatment with a *C. fimi* protease several active fragments were obtained which correlated to those found in culture supernatants of *C. fimi*. Hence it was suggested that Man26A is the only mannanase produced by *C. fimi* and that all multiple enzymes arise by proteolysis of this protein in the cultures (*32*).

Substrate Specificities of Mannanases. Endo-β-mannanases (β-1,4-D-mannan mannanohydrolases, EC 3.2.1.78) cleave polymeric mannans as well as mannooligosaccharides, usually with a degree of polymerization greater than three. The main hydrolysis products from galactomannans and glucomannans are mannobiose, mannotriose and various mixed oligosaccharides. Some endomannanases also cleave β-1,4 linkages between mannose and glucose in glucomannans depending on the glucose to mannose ratio. For example, salep glucomannan (glucose to mannose ratio 2:8) is hydrolyzed more extensively than konjac glucomannan (*7*). Properties of purified mannanases and their main hydrolysis products of various substrates are illustrated in tables 3 and 4. The degree of substitution and the distribution of the side groups influence mannanase activity (*7*). Thus, the degrees of hydrolysis of lucerne (48% galactose), guar (38% G), tara (28% G) and carob (24% G) galactomannans by a purified *Aspergillus niger* mannanase are 1, 5, 20, and 22%, respectively (*7a*). The galactosyl side groups of galactomannans as well as the acetyl groups have been shown to restrict hydrolysis of the main chain. A combined action of mannanase with α-galactosidase and acetylesterase on acetyl galactoglucomannan increased the amount of mannobiose and mannotriose released (*5*).

Besides their ability to hydrolyze mannans various β-mannanases show transglycosylation activity when incubated with mannooligosaccharides (*28, 33, 113, 114*). During hydrolysis, oligosaccharides can serve as acceptors instead of water (*115*). These enzymes belong to the class of glycosidases hydrolyzing the glycosidic bond with retention of anomeric configuration (*116*).

Recently there has been a considerable amount of work carried out to try to classify or group various hemicellulases and cellulases based on the degree of homology of the amino acid sequence of the various catalytic and binding domains (*34*). Several fungal carbohydrolases have been grouped with enzyme families which are more closely related to known bacterial enzymes than they are to each other. Hydrophobic cluster analysis and comparison of the primary structures of several mannanases, the sequences of which have been determined, have been placed these into two distinct families, namely families 5 and 26 of the glycosyl hydrolases, both of which are in the clan GH-A of retaining enzymes (*34-36*). Consequently, carbohydrolases from different microbial origins have shown similar substrate specificities, both on isolated and synthetic oligo- and polysaccharides and on their natural substrates, such as wood. In contrast, some closely related endoglucanases have shown quite different substrate specificities (*117*). However, the individual enzymes that have been classified using model substrates including the standard assay procedures have often shown quite different effects when they were applied to more natural substrates (*118*). It is recognized that a better understanding of the mechanism of hydrolysis of different substrates by different bacterial and fungal mannanases is required so that we can take full advantage of these highly specific catalysts.

Table 3. Endo-β-mannanases characterized from fungi.

Source Microorganism	pI	MM[a] kDa	pH optimum	Substrate / Mannan	Main hydrolysis products[d]	Ref.
Aspergillus niger	n.d.	42[b]	3.0; 3.8	guar gum	M_2, M_3, M_xGy	(*91*)
Aspergillus niger	3.7	40	3.5	ivory nut	M_2, M_3	
Aspergillus tamarii IP 1017-10	n.d.	53	4.5	locust bean	M_2, M_3, GM_2	(*15*)
Chrysosporium lignorum	4.11	n.d.	n.d.		n.d.	(*92*)
Fomes annosus	3.9-4.2	n.d.	n.d.		n.d.	(*92*)
Penicillium kloeckeri NRRL 1017	3.1-3.7 3.2-3.3	48-58 25-32	4.5-5.0 4.5-5.0	locust bean	M_2, M_3, M_xGy M_2, M_3, M_xGy	(*93*)
Penicillium purpurogenum 618	4.1	57	5.0	copra	M, M_xGy	(*94*)
Polyporus versicolor	4.6 3.75	45 34	n.d. n.d.		n.d. n.d.	(*24*)
Sclerotium rolfsii ATTC 200224	3.5 3.2	61.2 41.9	2.9 3.3	ivory nut	M_2, M_3 M_2, M_3	(*41*)
Sclerotium rolfsii CBS 191.62	2.75	46.5	3.0-3.5	ivory nut	M_2, M_3	(*95*)
Sporotrichum cellulophilum ATCC 20493	n.d. n.d.	108[c] 32.3[c]	5.5 6.0		n.d. n.d.	(*75*)
Stereum sanguinolentum	3.58	n.d.	n.d.		n.d.	(*92*)
Thielavia terrestris NRRL 8126	n.d n.d	52 30	4.5 5.5	coffee bean	n.d. n.d.	(*30*)
Thielavia terrestris NRRL 8126	n.d n.d	55 89	5.0 5.5	coffee bean	n.d. n.d.	(*30*)
Trichoderma harzianum E 58[f]	6.55	n.d.	n.d.		n.d.	(*27*)
Trichoderma reesei RUT C 30	5.4 4.6	53 51	3.5-4.0 3.5-4.0	ivory nut	M_2, M_3 M_2, M_3	(*29*)
Trichoderma reesei RUT C 30	5.2	46	5.0	locust bean	M_2, M_3 M_xGy	(*77*)
Tyromyces palustris	3.45	61[b]	3.5-3.8	larchwood	M, glucose, M_2, M_3, oligomers	(*97*)

[a]SDS-PAGE; [b]ultracentrifugation, [c]gel filtration; ; [d]M, mannose; G, galactose;
[f]not purified; n.d., not determined.

Table 4. Endo-β-mannanases characterized from bacteria.

Source Microorganism	pI	MM^a kDa	pH optimum	Substrate / Mannan	Main hydrolysis products[d]	Ref.
Aeromonas sp.	5.6	24^b	5.5	coffee	M_2, M_3, M_xGy	(*89*)
Bacillus	n.d.	55	5.5-6.9		n.d.	(*79*)
pumilis ATCC 72	n.d.	37	6.0			
Bacillus sp. AM-001	5.9	58	9		n.d.	(*98*)
	5.7	59	9		n.d.	
	5.1	42	8.5		n.d.	
Bacillus sp.	n.d.	58.5	9		n.d.	(*81*)
	n.d.	59.5	9		n.d.	
	n.d.	42	8.5		n.d.	
B. stearothermophilus ATCC 266	n.d.	162	5.5-7.5		n.d.	(*48*)
Bacillus subtilis 5H	n.d.	37	7.0	LBG konjac	M_2 (M_3, M_1) oligosaccharides Glc-Man disaccharides	(*99*)
Bacillus subtilis K-50	5.24	22^b	6	konjac coffee bean	M_2, M_3, oligosaccharides M_2, M_3, M_4, M_xGy	(*100*)
Bacillus subtilis KU-1	4.5	40	7.0	konjac	heterooligo-saccharides	(*101*)
Bacillus subtilis TX1	5.1	37	5-6	carob	M_2, M_3, M_xGy	(*102*)
Cellulomonas fimi	n.d.	100	5.5		n.d.	(*32*)
Enterococcus casseliflavus FL 2121	n.d / n.d.	142 / 137	6 / 6	konjac	oligosaccharides / n.d.	(*85*)
Pseudomonas sp. PT5	n.d.	34	7.5		n.d.	(*103*)
Pseudomonas fluorescens subsp. *cellulosa*	n.d.	46	7.0	carob M_3–M_6	M_2, M_3, M_1 M_2, M_1	(*104*)
Streptomyces lividans IAF 36	3.5	36	6.7	M_6	M_2, M_3, M_1	(*105*)
Streptomyces sp. 17	3.65	42.9	6.8	konjac	M_1-M_3, oligosaccharides	(*106*)
Vibrio sp. MA-138	3.6	41	7.5		n.d.	(*107*)

[a]SDS-PAGE; [b]ultracentrifugation, [c]gel filtration; ; [d]M, mannose; G, galactose; [f]not purified; n.d., not determined.

β-**Mannosidase.** β-Mannosidases (β-D-mannoside mannohydrolase, EC 3.2.1.25) catalyze the hydrolysis of terminal, non-reducing β-D-mannose residues in mannans, heteromannans and mannooligosaccharides. Some β-mannosidases also cleave the D-Man*p*-1,4-β-D-Glu*p* linkages in glucomannans. β-Mannosidases have been reported to occur in a wide range of plant and animal tissues, and in many microorganisms (*138*). Table 5 shows some properties of purified β-mannosidases from different microorganisms. However, reports on substrate specificities of β-mannosidases are limited. The β-mannosidase of *Tremella fuciformis* showed its highest activity towards mannobiose and the rate of hydrolysis of mannooligosaccharides decreased with increasing chain length (*37*), the β-mannosidase of *Aspergillus niger* showed moderate activity towards mannotriose and mannobiose but not towards mannan (*38*), and the β-mannosidase of *Bacillus sp.* demonstrated highest activity towards mannotetraose (*39*). Holazo et al. (*40*) reported on β-mannosidases from *Aspergillus niger* which also show transmannosylation activity enabling these enzymes to transfer mannose residues to primary and secondary alcohols. Addition of β-mannosidases to purified mannanases showed a pronounced synergistic effect on the degree of hydrolysis from heteromannans to monomeric sugars (*41*).

β-**Glucosidase.** β-Glucosidase (β-D-glucoside glucohydrolase, EC 3.2.1.21) catalyses the hydrolysis of terminal, non-reducing β-D-glucose residues from β-D-glucosides, including cellobiose and cellooligosaccharides. In some cases, mixed oligosaccharides consisting of mannose and glucose serve as substrates as well. In the enzymatic conversion of cellulose, sufficient levels of β-glucosidase are required in order to limit the inhibitory effect of cellobiose on the cellulases. A commercial, purified β-glucosidase from almonds has been shown to hydrolyze terminal glucose residues from β-mannanase-produced glucomannan oligosaccharides (*42*).

α-**Galactosidase.** α-Galactosidases (α-D-galactoside galactohydrolase, EC 3.2.1.22) catalyze the hydrolysis of terminally linked galactosidic residues present in simple galactose containing oligosaccharides, as well as in more complex polysaccharides such as galactomannans and galactoglucomannans. A synergistic effect between α-galactosidases and endomannanases during the degradation of guar gum has been observed (*119*). α-Galactosidases are widely distributed in microorganisms, plants, and animals. Microorganisms have the advantage of higher production yields, and among them, fungal galactosidases are the most suitable for technological applications. This is mainly due to their extracellular localization, acidic pH optima, and broad stability profiles. α-Galactosidases have been purified and characterized from many fungi, such as *Penicillium purpurogenum* (*43*), *Penicillium ochrochloron* (*44*), *Trichoderma reesei* (*45*), *Aspergillus niger* (*46*), *Mortierella vinacea* (*47*), and from the bacteria *Bacillus stearothermophilus* (*48*) and *Thermotoga neapolitana* (*49*). The ability to hydrolyse galactose groups from polymeric and oligomeric substrates varies between different α-galactosidases. The activity of an α-galactosidase from *Penicillium simplicissimum* decreased with increasing chain

Table 5. β-Mannosidases characterized from fungi, yeasts and bacteria [adapted from Viikari et al. (*90*)].

Source Microorganism	pI	MM (kDa)	pH optimum	Ref.
Thielavia terrestris	n.d.	72[a]	4.5-5.5	(*30*)
Polyporus sulfureus	n.d.	64[b]	2.4-3.4	(*108*)
Tremella fuciformis	n.d.	140[b]	5.0	(*37*)
Aspergillus niger	4.7	130[b]	3.5	(*109*)
Aspergillus niger	n.d.	120[b]	3.5-4.0	(*38*)
Aspergillus niger	5.0	264[c]	2.5-5	(*110*)
Trichosporon cutaneum	n.d.	193[b]	6.5	(*78*)
Sclerotium rolfsii	4.5	57.5[a]	2.5	(*41*)
Bacillus sp.	5.5	94[a]	6.0	(*39*)
Pyrococcus furiosus	6.9	220[a]	7.4	(*111*)

[a]SDS-PAGE; [b]gel filtration; [c]nonreducing SDS-PAGE

length of the substrate (*50*), whereas the hydrolysis of galactose oligomers by an α-galactosidase from *Aspergillus niger* becomes more efficient when the chain length of the oligosaccharide increases (*46*). Two α-galactosidases from *Mortierella vinacea* showed great differences with respect to the removal of galactose residues from galactomannans. α-Galactosidase I could not act on galactomannans, while α-galactosidase II could hydrolyse α-galactosidic linkages of galactomannans, and was able to liberate 65% and 85% of the total α-galactosyl residues from guar gum and locust bean gum, respectively (*47*). *Aspergillus niger* α-galactosidase III was able to remove galactose from galactomannobiose and from the undegraded galactomannans, whereas α-galactosidase I and II did not show any activity on the various galactomannans (*15*).

Acetyl Esterases. Acetyl esterases or acetyl galactoglucomannanesterases (AGME) liberate acetic acid from acetylated galactoglucomannans, which occur in annual plants and softwoods. Hitherto only few AGMEs, the presence of which was first reported 1991, have been purified and characterized to date. A purified acetyl esterase of *Trichoderma reesei* has been shown to liberate acetic acid from short oligomers of glucomannan, and in combination with mannanase, also from galactoglucomannan (*51*). An AGME from *Aspergillus oryzae* was shown to be able to act on polymeric substrates and its activity was enhanced by addition of mannanase (*52*). The specific activity of the enzyme was 10-times higher for acetylated mannan than for acetylated xylan.

Application of Mannan Degrading Enzymes

Mannan degrading enzymes have many uses in the food, feed, pharmaceutical, and pulp/paper industry as well as for oil and gas well stimulation. For these applications commercial mannanase preparations are produced and sold by several companies. Additionally, mannanase activity is also prevalent as minor activity in many hemicellulase, cellulase and pectinase preparations. Mannanases can be applied in the processing of instant coffee, in which extracts of coffee beans are concentrated by evaporation and then dried by freeze- or spray-drying. Since almost half of the green coffee bean dry weight is made of polysaccharides with mannan as the major component (*4*), the viscosity of the extract can be drastically reduced by cleaving the mannan constituent. This reduction of viscosity facilitates the production of instant coffee by improving the effectiveness with which the extracts can be concentrated and hence lowering cost processes such as evaporation (*2*). The cleavage of the mannan content in the cell wall of fruits can improve juice recovery. In processing of clear pineapple juice the presence of galactomannans was found to decrease the ultrafiltration capabilities, as well as to increase pulp suspension and foaming properties of the juice (*53*).

The hydrolysis of coconut residual cake by mannanases can increase the commercial value of coconut products and minimize pollution. The defatted coconut

cake contains a large amount of mannan which could potentially be used for the production of mannooligosaccharides which can find application in the pharmaceutical industry and as growth factor for *Bifidobacterium sp.* and *Lactobacillus sp.* (*55*). Via hydrolysis various mannooligosaccharides up to DP 9 including 6^3-α-D-galactosyl-β-D-mannotriose were isolated from *Ceratonia siliqua* galactomannan using a mannanase from *Aspergillus niger* (*113*). Other authors used an endomannanase from *Penicillium purpurogenum* to obtain galactosylmannooligosaccharides from copra galactomannan (*120*). β-1,4-Mannotriose was produced efficiently using the *Penicillium* derived enzymes and removal of mannobiose and mannose by yeast fermentation (*121*). From softwood galactoglucomannan 6^2-α-D-galactosyl-β-D-mannobiose was isolated using a mannanase from *Tyromyces palustris* (*122*). An endo-1,4-β-mannanase from *Sclerotium rolfsii* has been shown to produce 6^1-α-D-galactosyl-β-D-mannotriose both from monomers and via hydrolysis from polysaccharides (*123*).

Mannanases have applications in the cleavage of the galactomannans in feedstuffs, especially from legume seeds such as soybean, alfalfa (*Medicago sativa*), and guar meal. Guar meal, a by-product in the production of guar gum, is a good source of essential amino acids and has a potential use as a protein supplement in poultry diets (*56*). The galactomannan content causes growth depression by reduction of nitrogen retention, fat absorption, and metabolizable energy when fed to chicks. The addition of mannanases to the feedstuff was found to inactivate the growth repressing properties of this galactomannan (*57*).

Another potential use of mannan degrading enzymes is for the hydrolysis of galactomannan-based, water-soluble polymers used in hydraulic fracturing of oil and gas wells. In hydraulic fracturing applications, the polymer solution that is added to the wellbore contains particles that are used to hold crevices generated by applying high levels of hydrostatic pressure to the flooded well. To allow gas or oil to flow to the wellbore, the viscosity of the fracturing fluid must be subsequently reduced or broken in situ, either by chemical oxidation or enzymatic hydrolysis of the polymer structure. The low viscosity fluid can then be pumped out, after which oil or gas production can proceed (*49*).

In the last decade the use of hemicellulases in the pulp and paper industry gained a lot of interest. Mannanases can be applied in combination with xylanases for prebleaching of softwood kraft pulps (*60*). The main goal in the enzymatic bleaching of kraft pulps has been to reduce the consumption of chlorine chemicals. However, enzymes can also be used successfully for increasing the brightness of pulp, which is of key importance in the development of totally chlorine-free bleaching sequences (*58*). The addition of an hemicellulolytic enzymatic step to any conventional chemical bleaching sequence results in a higher final brightness value of the pulp. The hemicellulase treatment is an indirect bleaching method, rendering the fibers more accessible to bleaching chemicals and leading to more efficient delignification (*58*). Hitherto, only a few mannanases have been reported to increase pulp bleachability. A mannanase from *Trichoderma reesei* was effective in enhancing the bleachability of softwood pulps, whereas the effects of *Aspergillus*

niger and *Caldocellulosiruptor saccharolyticus* mannanases were less efficient (*59*). The effect of mannanases on the bleachability of pulps strongly depends on the cooking method and the bleaching sequences used. Two purified mannanases of *Sclerotium rolfsii* have been shown to increase the brightness of softwood kraft pulp and decreased the kappa number (*60*). Several reports in the literature have concluded that there is no correlation between the ability of mannanases to release saccharides from purified mannans and their action on pulps. Similarly, endoglucanases and hemicellulases have been used for the deinking of waste paper. However, compared to xylanases and endoglucanases, mannanases seem to play only a minor role in this process (*124*).

Together with xylanases, mannanases have been shown to remove hemicellulose from dissolving pulps. Dissolving pulps are purified celluloses used for making viscose rayons, cellulose esters and cellulose ethers. They are derived from prehydrolyzed kraft pulps or acid sulfite pulps (*2*). In the viscose process, high amounts of hemicellulose can affect the viscose filterability and the xanthanation of cellulose. High hemicellulose concentrations can also adversely influence the strength of the end product. In the acetate process, high amounts of hemicellulose can affect the solubility of the resulting cellulose acetate, while even relatively small concentrations of glucomannans have been found to unfavorably influence acetate filterability, solution haze and to cause acetate false body effect. The amount of mannan removed by the enzymatic treatment from bleached pulp varies between 8 and 20% (*61*). A mannanase from *Sclerotium rolfsii* and a xylanase from *Thermomyces lanuginosus* acted synergistically on the pulp to solubilize 50% more mannan and 11% more xylan than did the individual enzymes. Solubilization of mannan and xylan was further improved by the addition of purified endoglucanases (*125*).

α-Galactosidases have several applications in processing foods containing galactooligo-saccharides. The raffinose-family oligosaccharides formed in various legumes induce flatulence in humans and many animals as their digestive tracks lack the enzymes needed to cleave the α-galactosyl linkages present in these carbohydrates. The application of an α-galactosidase to hydrolyze these flatulence-causing sugars in legume foods (e. g. soybean milk) can improve their nutritional value (*62*). In the sugar industry, raffinose is known to inhibit sucrose crystallization from beet sugar syrups. By removing raffinose in beet molasses using enzymatic treatment with α-galactosidase, it is possible to enhance the crystallization efficiency of beet sugar. For this purpose, α-galactosidase of *Mortierella vinacea* has been immobilized in Japan for the treatment of hundreds of tons of molasse per day (*63*). Furthermore, the hydrolysis of raffinose in molasse can improve its use as substrate in alcohol fermentation processes increasing the yield (*63*).

The modification of galactomannans with α-galactosidases can improve their rheological and gelling properties. Galactomannans such as locust bean gum (LBG) form gels and act as promoters of gelling when mixed with other polysaccharides such as agar, carrageenan, and xanthan. Since LBG is rather expensive, guar gum, which is more easily available but has poorer gelling properties than LBG, can be

modified using α-galactosidase. Removal of a quantitative proportion of galactose moieties from guar gum to generate a modified galactomannan with less than 30% galactose residues yields a polysaccharide with functional properties comparable with those of LBG (44). In addition, α-galactosidases can be used in medical or pharmaceutical applications (64).

In the last few years oligosaccharides became of interest in the pharmaceutical and in the food industry. Oligosaccharides, especially heterooligosaccharides, play an important role in glycoproteins and glycolipids. They are part of membrane-bound proteins, where they serve as recognition-signals for hormones, toxins, antibodies, viruses and bacteria. New techniques involving enzymes are currently being investigated to synthesize oligosaccharide-based compounds with therapeutic or diagnostic properties (i.e. cardiotonic and inflammatory agents, receptors for pathogens, antibodies for diagnostic-kits) (126-128).

Oligosaccharides can be specifically synthesized with glycosidases (EC 3.2) via reversed hydrolysis or via transglycosylation reactions of unprotected carbohydrates under mild conditions (128). In most cases the kinetically controlled transglycosylation reaction, where a glycosyl moiety from a glycosyl derivative is transferred to an acceptor molecule, is characterized by higher reaction rates and yields (129). Several β-mannosidases (130-132), α-galactosidases (133, 134) and β-glucosidases (135, 136) catalyze the synthesis of oligosaccharides from monomers.

Acknowledgements

References

1.   Godfrey, T.; and West, S. In *Industrial Enzymology;* Godfrey, T.; West, S., Ed.; Macmillan Press: London, 1996, pp. 1-8.
1a.   Saddler, J. N. (Ed.) *Bioconversion of Forest and Agricultural Plant Residues.* C.A.B. International: Wallingford, UK, 1993.
2.   Wong, K. K. Y.; and Saddler, J. N. In *Hemicellulose and Hemicellulases;* Coughlan, M. P.; and Hazlewood, G. P., Eds.; Portland Press: London and Chapel Hill, 1993, pp. 127-143.
3.   Meier, H.; and Reid, J. S. G. In *Plant Carbohydrates I. Intracellular Carbohydrates;* Loewus, F. A.; and Tanner, W., Eds.; Springer: Berlin, 1982,. Vol 13A, pp. 418-471.
4.   Bradbury, A. G.; and Halliday, D. J. *J. Agric. Food Chem.* **1990**, *38*, 389-392.
5.   Puls, J.; and Schuseil, J. In *Hemicellulose and Hemicellulases;* Coughlan, M. P.; and Hazlewood, G. P., Eds.; Portland Press: London and Chapel Hill, 1993, pp. 1-27.

5a. Eriksson, K.-E. L.; Blanchette, R. A.; and Ander, P. *Microbial and Enzymatic Degradation of Wood and Wood Components.* Springer: Berlin, 1990.

6. Tombs, M.; and Harding, S. E. *An Introduction to Polysaccharide Biotechnology,* Taylor&Francis: London (1998).

7. McCleary, B. V. *ACS Symp. Ser.* **1991**, *460*, 437-449.

7a. McCleary, B.; and Matheson, N.K. *Carbohydr. Res.* **1983**, *119*, 191-219.

8. Yalpani, M.; and Sandford, P. A. In *Industrial Polysaccharides: Genetic Engineering, Structure/Property Relations and Applications;* Yalpani, M., Ed.; Progress in Biotechnology; Elsevier Science Publishers: Amsterdam, 1987, Vol. 3; pp. 311-335.

9. Whistler, R. L.; and BeMiller, J. N. *Carbohydrate Chemistry for Food Scientists,* Eagan Press: St. Paul (1997).

10. Reid, J. S. G.; and Edwards, M. E. In *Food Polysaccharides and Their Applications;* Stephen, A. M., Ed.; Marcel Dekker: New York, 1995, pp. 155-186.

11. Dekker, R. F. H. In *Polysaccharides in Food;* Blanshard, J. M. V.; and Mitchell, J. R., Ed.; Butterworth: London, 1979, pp. 93-108.

12. Nakajima, N.; and Matsuura, Y. *Biosci. Biotech. Biochem.* **1997**, *61*, 1739-1742.

13. Lapasin, R.; and Pricl, S. *Rheology of Industrial Polysaccharides: Theory and Applications,* Blackie Academic and Professional: London (1995).

14. Dekker, R. F. H.; and Richards, G. N. *Adv. Carbohydr. Chem. Biochem.* **1976**, *32*, 277-352.

15. Civas, A.; Eberhard, R.; Le Dizet, P.; and Petek, F. *Biochem. J.* **1984**, *219*, 857-863.

16. Großwindhager, C.; Sachslehner, A.; Nidetzky, B.; and Haltrich, D. *J. Biotechnol.* **1999**, *67*, 189–203.

17. Biely, P. In *Hemicellulose and Hemicellulases;* Coughlan, M. P.; and Hazlewood, G. P., Ed.; Portland Press: London and Chapel Hill, 1993, pp. 29-51.

18. Kremnicky, L.; and Biely, P. *Arch. Microbiol.* **1997**, *167*, 350-355.

19. Araujo, A.; and Ward, O. W. *J. Ind. Microbiol.* **1990**, *6*, 171-178.

20. Sachslehner, A.; Nidetzky, B.; Kulbe, K. D.; and Haltrich, D. *Appl. Environ. Microbiol.* **1998**, *64*, 594-600.

21. Sachslehner, A.; Haltrich, D.; Gübitz, G.; Nidetzky, B.; and Kulbe, K. D., *Appl. Biochem. Biotechnol.* **1998**, *70–71*, 939–953.

22. El-Helow, E. R.; Sabry, S. A.; and Khattab, A. A. *Antonie van Leeuwenhoek.* **1997**, *71*, 189-193.

23. Sachslehner, A.; Haltrich, D.; Nidetzky, B.; and Kulbe, K. D. *Appl. Biochem. Biotechnol.* **1997**, *63-65*, 189-201.

24. Johnson, K. G.; and Ross, N. W. *Enzyme Microb. Technol.* **1990**, *12*, 960-964.

25. Wong, K. K. Y.; Tan, L. U. L.; and Saddler, J. N. *Microbiol. Rev.* **1988**, *52*, 305-317.

26. Willick, G. E.; and Seligy, V. E. *Eur. J. Biochem.* **1985**, *151*, 89-96.

27. Torrie, J. P.; Senior, D. J.; and Saddler, J. N. *Appl. Microbiol. Biotechnol.* **1990**, *34*, 303-307.
28. Gübitz, G. M.; Hayn, M.; Urbanz, G.; and Steiner, W. *J. Biotechnol.* **1996**, *45*, 165-172.
29. Stålbrand, H.; Siika-aho, M.; Tenkanen, M.; and Viikari, L. *J. Biotechnol.* **1993**, *29*, 229-242.
30. Araujo, A.; and Ward, O. W. *J. Ind. Microbiol.* **1990**, *6*, 269-274.
31. Stålbrand, H.; Salheimo, A.; Vehmaanperä, J.; Henrissat, B.; and Penttilä, M. *Appl. Environ. Microbiol.* **1995**, *61*, 1090-1097.
32. Stoll, D.; Stålbrand, H.; and Warren, R. A. *Appl. Environ. Microbiol.* **1999**, *65*, 2598–2605.
33. Harjunpää, V.; Helin, J.; Koivula, A.; Siika-aho, M.; and Drakenberg, T. *FEBS Lett.* **1999**, *443*, 149–153.
34. Henrissat, B.; and Bairoch, A. *Biochem. J.* **1993**, *293*, 781-788.
35. Henrissat, B.; and Bairoch, A. *Biochem. J.* **1996**, *316*, 695-696.
36. Henrissat, B.; Callebaut, I.; Fabrega, S.; Lehn, P.; Mornon, J. P.; and Davies, G. *Proc. Natl. Acad. Sci.* **1995**, *92*, 7090-7094.
37. Sone, Y.; and Misake, A. *J. Biochem.* **1978**, *83*, 1135-1144.
38. Elbein, A. D.; Adya, S.; and Lee, Y. C. *J. Biol. Chem.* **1977**, *252*, 2026-2031.
39. Akino, T.; Nakamura, N.; and Horikoshi, K. *Agric. Biol. Chem.* **1988**, *52*, 1459-1464.
40. Holazo, A.; Shinoyama, H.; Kamiyama, Y.; and Yasui, T. *Biosci. Biotech. Biochem.* **1992**, *56*, 822-824.
41. Gübitz, G. M.; Hayn, M.; Sommerauer, M.; and Steiner, W. *Bioresource Technol.* **1996**, *58*, 127-135.
42. McDougall, G. J. *Carbohydr. Res.* **1993**, *241*, 227-236.
43. Shibuya, H.; Kobayashi, H.; Park, G. G.; Komatsu, Y.; Sato, T.; Kaneko, R.; Nagasaki, H.; Yoshida, S.; Kasamo, K.; and Kusakabe, I. *Biosci. Biotech. Biochem.* **1995**, *59*, 2333-2335.
44. Dey, P. M.; Patel, S.; and Brownleader, M. D. *Biotechnol. Appl. Biochem.* **1993**, *17*, 361-371.
45. Zeilinger, S.; Kristufek, D.; Arisan-Atac, I.; Hodits, R.; and Kubicek, C. P., *Appl. Environ. Microbiol.* **1993**, *59*, 1347-1353.
46. Manzanares, P.; de Graaff, L. H.; and Visser, J. *Enzyme Microb. Technol.* **1998**, *22*, 383–390.
47. Shibuya, H.; Kobayashi, H.; Sato, T.; Kim, W.-S.; Yoshida, S.; Kaneko, S.; Kasamo, K.; and Kusakabe, I. *Biosci. Biotech. Biochem.* **1997**, *61*, 592-598.
48. Talbot, G.; and Sygusch, J. *Appl. Environ. Microbiol.* **1990**, *56*, 3505-3510.
49. McCutchen, C. M.; Duffaud, G. D.; Leduc, P.; Petersen, A. R. H.; Tayal, A.; Khan, S. A.; and Kelly, R. M. *Biotechnol. Bioeng.* **1996**, *52*, 332-339.
50. Luonteri, E.; Tenkanen, M.; and Viikari, L. *Enzyme Microb. Technol.* **1998**, *22*, 192-198.
51. Tenkanen, M.; Puls, J.; Rättö, M.; and Viikari, L. *Appl. Microbiol. Biotechnol.* **1993**, *39*, 159-165.

52. Tenkanen, M.; Thornton, J.; and Viikari, L. *J. Biotechnol.* **1995**, *42*, 197-206.

53. Grassin, C.; and Fauquembergue, P. In *Industrial Enzymology;* Godfrey, T.; and West, S., Eds.; Macmillan Press Ltd: London, 1996, pp. 225-264.

54. Chenchin, K. L.; and Yamamoto, H. Y. *J. Food Sci.* **1978**, *43*, 1261-1263.

55. Hossain, M. Z.; Abe, J.-I.; and Hizukuri, S. *Enzyme Microb. Technol.* **1996**, *18*, 95-98.

56. Patel, M. B.; and McGinnis, J. *Poultry Sci.* **1985**, *64*, 1148-1156.

57. Ray, S.; Pubols, M. H.; and McGinnis, J. *Poultry Sci.* **1982**, *61*, 488-494.

58. Viikari, L.; Suurnäkki, A.; and Buchert, J. *ACS Symp. Ser.* **1996**, *655*, 15-24.

59. Suurnäkki, A.; Clark, T. A.; Allison, R. W.; Buchert, J.; and Viikari, L. In *Biotechnology in the Pulp and Paper Industry.;* Srebotnik, E.; and Messner, K., Eds.; Facultas-Universitätsverlag: Vienna, ;1996, pp. 69–74.

60. Gübitz, G. M.; Schnitzhofer, W.; Balakrishnan, H.; and Steiner, W. *J. Biotechnol.* **1996**, *50*, 181-188.

61. Gübitz, G. M.; Lischnig, T.; Stebbing, D.; and Saddler, J. N. *Biotechnol. Lett.* **1997**, *19*, 491-495.

62. Cruz, R.; and Park, Y. K. *J. Food Sci.* **1982**, *47*, 1973-1975.

63. Buchholz, K.; and Kasche, V. *Biokatalysatoren und Enzymtechnologie*, VCH: Weinheim (1997).

64. Müller, G.; and Köhler, A. *Biol. Rdsch.* **1985**, *23*, 345-366.

65. Rättö, M.; and Poutanen, K. *Biotechnol. Lett.* **1988**, *10*, 661-664.

66. Reese, E. T.; and Shibata, Y. *Can. J. Microbiol.* **1965**, *11*, 167-183.

67. Ademark, P.; Varga, A.; Medve, J.; Harjunpää, V.; Drakenberg, T.; Tjerneld, F.; and Stålbrand, H. *J. Biotechnol.* **1998**, *63*, 199-210.

68. Johnson, K. G. *World J. Microbiol. Biotechnol.* **1990**, *6*, 209-217.

69. Farrell, R. L.; Biely, P.; and McKay, D. L. In *Biotechnology in the Pulp and Paper Industry.;* Srebotnik, E.; and Messner, K., Eds.; Facultas-Universitätsverlag: Vienna, 1996, pp 485-489.

70. Kusakabe, I.; Zama, M.; Park, G. G.; Tubaki, K.; and Murakami, K. *Agric. Biol. Chem.* **1987**, *51*, 2825-2826.

71. Haltrich, D.; Laussamayer, B.; Steiner, W.; Nidetzky, B.; and Kulbe, K. D. *Bioresource Technol.* **1994**, *50*, 43-50.

72. Gübitz, G. M.; and Steiner, W. *ACS Symp. Ser.* **1995**, *618*, 319-331.

73. Haltrich, D.; Preiß, M.; and Steiner, W. *Enzyme Microb. Technol.* **1993**, *15*, 854-860.

74. Haltrich, D.; and Steiner, W. *Enzyme Microb. Technol.* **1994**, *16*, 229-235.

75. Araujo, A.; and Ward, O. W. *J. Ind. Microbiol.* **1991**, *8*, 229-236.

76. Puchart, V.; Katapodis, P.; Biely, P.; Kremnicky, L.; Christakopoulos, P.; Vrsanska, M.; Kekos, D.; Macris, B. J.; and Bhat, M. K. *Enzyme Microb. Technol.* **1999**, *24*, 355-361.

77. Arisan-Atac, I.; Hodits, R.; Kristufek, D.; and Kubicek, C. P. *Appl. Microbiol. Biotechnol.* **1993**, *39*, 58-62.

78. Oda, Y.; and Tonomura, K. *Lett. Appl. Microbiol.* **1996**, *22*, 173-178.

79. Araujo, A.; and Ward, O.P. *Appl. Environ. Microbiol.* **1990**, *56*, 1954-1956.

80. Ooi, T.; and Kikuchi, D. *World J. Microbiol. Biotechnol.* **1995**, *11*, 310-314.

81. Akino, T.; Nakamura, N.; and Horikoshi, K. *Appl. Microbiol. Biotechnol.* **1987**, *26*, 323-327.

82. Abe, J.-I.; Hossain, M. Z.; and Hizukuri, S. *J. Ferment. Bioeng.* **1994**, *78*, 259-261.

83. Khanongnuch, C.; Lumyong, S.; Ooi, T.; and Kinoshita, S. *Biotechnol. Lett.* **1999**, *21*, 61-63.

84. Mendoza, N. S.; Arai, M.; Kawaguchi, T.; Yoshida, T.; and Joson, L. M. *World J. Microbiol. Biotechnol.* **1994**, *10*, 551-555.

85. Oda,Y.; Komaki,T.; and Tonomura,K. *J. Ferment. Bioeng.* **1993**, *76*, 14-18.

86. Zakaria, M. M.; Ashiuchi, M.; Yamamoto, S.; and Yagi, T. *Biosci. Biotech. Biochem.* **1998**, *62*, 655-660.

87. Gomes, J.; and Steiner, W. *Biotechnol. Lett.* **1998**, *20*, 729–733.

88. Duffaud, G. D.; McCutchen, C. M.; Leduc, P.; Parker, K. N.; and Kelly, R. M. *Appl. Environ. Microbiol.* **1997**, *63*, 169-177.

89. Tamaru, Y.; Araki, T.; Amagoi, H.; Mori, H.; and Morishita, T. *Appl. Environ. Microbiol.* **1995**, *61*, 4454-4458.

90. Viikari, L.; Tenkanen, M.; Buchert, J.; Rättö, M.; Bailey, M.; Siika-aho, M.; and Linko, M. In *Bioconversion of Forest and Agricultural Plant Residues;* Saddler, J. N., Ed.; C.A.B. International: Wallingford, 1993, Vol. 9, pp. 131-182.

91. Eriksson, K. E.; and Winell, M. *Acta Chem. Scand.* **1968**, *22*, 1924-1934.

92. Ahlgren, E.; and Eriksson, K. E. *Acta Chem. Scand.* **1967**, *21*, 1193-1200.

93. Biely, P.; McKay, D. L.; and Farrell, R. L. In *Biotechnology in the Pulp and Paper Industry.;* Srebotnik, E.; and Messner, K., Eds.; Facultas-Universitätsverlag: Vienna, 1996, pp. 479-484.

94. Park, G. G.; Kusakabe, I.; Komatsu, Y.; Kobayashi, H.; Yasui, T.; and Murakami, K. *Agric. Biol. Chem.* **1987**, *51*, 2709-2716.

95. Sachslehner, A.; and Haltrich, D. *FEMS Microbiol. Lett.* **1999**, *177*, 47–55.

96. Großwindhager, C.; Sachslehner, A.; Fontana, J. D.; and Haltrich, D. In *7th International Conference on Biotechnology in the Pulp and Paper Industry*; Canadian Pulp and Paper Association: Vancouver, Canada, 1998; pp 27–30.

97. Ishihara, M.; and Shimizu, K. *Mokuzai Gakkaishi* **1980**, *26*, 811-818.

98. Akino, T.; Nakamura, N.; and Horikoshi, K. *Agric. Biol. Chem.* **1988**, *52*, 773-779.

99. Khanongnuch, C.; Asada, K.; Tsuruga, H.; Ooi, T.; Kinoshita, S.; and Lumyong, S. *J. Ferment. Bioeng.* **1998**, *86*, 461-466.

100. Emi, S.; Fukumoto, J.; and Yamamoto, T. *Agric. Biol. Chem.* **1972**, *36*, 991-1001.

101. Zakaria, M.M.; Yamamoto, S.; and Yagi, T. *FEMS Microbiol. Lett.* **1998**, *158*, 25-31.

102. McCleary, B. *Phytochemistry.* **1979**, *18*, 757-763.

103. Yamaura, I.; Matsumoto, T.; Funatsu, M.; and Funatsu, Y. *Agric. Biol. Chem.* **1990**, *54*, 2425-2427.

104. Braithwaite, K. L.; Black, G. W.; Hazlewood, G. P.; Ali, B. R.; and Gilbert, H. J. *Biochem. J.* **1995**, *305*, 1005-1010.
105. Arcand, N.; Kluepfel, D.; Paradis, F. W.; Morosoli, R.; and Shareck, F. *Biochem. J.* **1993**, *290*, 857-863.
106. Takahashi, R.; Kusakabe, I.; Kobayashi, H.; Murakami, K.; Maekawa, A.; and Suzuki, T. *Agric. Biol. Chem.* **1984**, *48*, 2189-2195.
107. Tamaru, Y.; Araki, T.; Morishita, T.; Kimura, T.; Sakka, K.; and Ohmiya, K. *J. Ferment. Bioeng.* **1997**, *83*, 201-205.
108. Wan, C. C.; Muldrey, L. E.; Li, S. C.; and Li, Y. T. *J. Biol. Chem.* **1976**, *251*, 4384-4388.
109. Bouquelet, S.; Spik, G.; and Montreuil, J. *Biochim. Biophys. Acta.* **1978**, *522*, 521-530.
110. Ademark, P.; Lundqvist, J.; Hägglund, P.; Tenkanen, M.; Torto, N.; Tjerneld, F.; and Stålbrand, H. *J. Biotechnol.* **1999**, *75*, 281-189.
111. Bauer, M. W.; Bylina, E. J.; Swanson, R. V.; and Kelly, R. M. *J. Biol. Chem.* **1997**, *271*, 23749-23755.
112. Sjöström, E. *Wood chemistry*; Academic Press: San Diego, **1993**.
113. Mccleary, B.V. and Matheson, N.K. *Carbohydrate Res.* **1983**, *119*, 191-219.
114. Biely, P.; McKay, D.L. and Farell, R.L. *Sixth International Conference on Biotechnology in the Pulp and Paper Industry, Vienna, Austria, 11-15 June 1995*, **1995**, *Abstract OG1-221*.
115. Sinnot, M.L. *Chem. Rev.* **1990**, *90*, 1171-1202.
116. Mccarter, J.D. and Withers, S.G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885-892.
117. Tomme, P.; Warren, A.J. and Gilkes, N.R. *Adv. Microb. Physio.l* **1995**, *37*, 1-87.
118. Gübitz, G.M.; Haltrich, D.; Latal, B. and Steiner, W. *Appl. Microbiol. Biotechnol.* **1997**, *47*, 658-662.
119. Tayal, A.; Pai, V.B. and Khan, S.A. *Macromol.* **2000**, *32*, 5567-5574.
120. Park, G.G. and Hak, N. *J. Microbiol. Biotechnol.* **1992**, *2*, 204-208.
121. Park, G.G.; Kusakabe, I.; Yasui, T. and Murakami, K. *Japan J. Trop. Agric.* **1988**, *32*, 208-214.
122. Shimizu, K. and Ishihara, M. *Agric. Biol. Chem.* **1983**, *47*, 949-955.
123. Gübitz, G.M.; Laussamayer, B.; Schubert-Zsilavecz, M. and Steiner, W. *Enzyme Microb. Technol.* **2000**, *26*, 15-21.
124. Gübitz, G.M.; Mansfield, S.D. and Saddler, J.N. *J. Biotechnol.* **1998**, *65*, 209-215.
125. Mansfield, S.D.; Saddler, J.N. and Gübitz, G.M. *Enzyme Microb. Technol.* **1998**, *23*, 133-140.
126. Nilsson, K.G.I. *ACS Symp. Ser.* **1991**, *466*, 51-78.
127. Nilsson, K.G. *Tibtech* **1988**, *6*, 256-264.
128. Ichikawa, Y.; Look, G.C. and Wong, C.H. *Anal. Biochem.* **1992**, *202*, 215-238.
129. Monsan, P. and Paul, F. *Fems Microbiol Rev.* **1995**, *16*, 187-192.
130. Fujimoto, H.; Isomura, M. and Ajisaka, K. *Biosci. Biotechnol Biochem.* **1997**, *61*, 164-165.

131. Itoh, H. and Kamiyama, Y. *J. Ferment. Bioeng.* **1995**, *80*, 510-512.
132. Taubken, N.; Sauerbrei, B. and Thiem, J. *J. Carbohydr. Chem.* **1993**, *12*, 651-667.
133. Savelev, A.N.; Ibatyllin, F.M.; Eneyskaya, E.V.; Kachurin, A.M. and Neustroev, K.N. *Carbohydr. Res* **1996**, *296*, 261-273.
134. Ajisaka, K. and Fujimoto, H. *Carbohydr. Res.* **1989**, *185*, 139-146.
135. Christakopoulos, P.; Goodenough, P.W.; Kekos, D.; Macris, B.J.; Claeyssens, M. and Bhat, M.K. Oxysporum. *Eur. J. Biochem.* **1994**, *224*, 379-385.
136. Christakopoulos, P.; Kekos, D.; Macris, B.J.; Goodenough, P.W. and Bhat, M.K. *Biotechnol. Lett.* **1994**, *16*, 587-592.
137. Koedritz, L.F.; Harvey, A.H.; and Honarpour, M. *Introduction to Petroleum Reservoir Analysis.* Gulf Publishing Company: Houston, 1989.
138. Dey, P.M. *Adv. Carbohydr. Chem. Biochem.* **1978**, *35*, 341-376.

Author Index

Subject Index