UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Contribution of Oxidative Enzymes to the Degradation of Cellulose by Filamentous Fungi

Permalink

Author

Beeson, William Templeton

Publication Date

2011

Peer reviewed|Thesis/dissertation

Contribution of Oxidative Enzymes to the Degradation of Cellulose by Filamentous Fungi

by

William Templeton Beeson IV

A dissertation in submitted in partial satisfaction of the requirements for the degree

of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professors Jamie H. D. Cate & Michael A. Marletta, Co-Chairs

Professor Michelle C. Y. Chang

Professor Russell E. Vance

Fall 2011

Abstract

Contribution of Oxidative Enzymes to the Degradation of Cellulose by Filamentous Fungi

by

William Templeton Beeson IV

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professors Jamie H. D. Cate and Michael A. Marletta, Co-Chairs

The cost of cellulose degrading enzymes is still a major barrier to the economical production of liquid fuels from lignocellulose. Fungi play a central role in the degradation of plant biomass in terrestrial environments. They use a wide variety of secreted enzymes to break down biopolymers present in the plant cell wall. Here I report on the mechanisms of cellulose degradation used by *Neurospora crassa*, a model genetic organism, and *Myceliopthora thermophila*, a thermophilic fungus. Enzymes important for the degradation process were identified using a combination of transcriptomics, proteomics, genetics, and biochemistry. A key result from this work is that, in *N. crassa* and many other fungi, oxidative enzymes play a critical role in the depolymerization of cellulose. In contrast to the accepted models of oxidative cellulose degradation, via non-specific hydroxyl radical species, I found that in *N. crassa*, an oxidoreductase and several copper oxidases bind to cellulose and hydroxylate the substrate at specific positions leading to cleavage of the glycosidic bonds. This mode of action is orthogonal to that of traditional hydrolases and if used optimally in conjunction with other cellulases, could reduce the required enzyme loading for lignocellulose saccharification by 2-4 fold.

Dedication

This dissertation is dedicated to my parents, Trey and Linda Beeson, who have always supported and encouraged me in my studies. This dissertation is also dedicated to Dr. James Horvath and Dr. Nigel Richards, of the University of Florida, who rekindled my interest and love of science during my freshmen and sophomore years of college. This dissertation is also dedicated to my wife, Kristin Mimbs, who has endured life as a "lab widow" for the past four years while I performed the research described in the following pages.

Contents

Chapter 1: Introduction to biofuels and enzymatic cellulose degradation

1.1 Biomass as a renewable feedstock for liquid fuel production

Increasing energy demand from developing nations and limited petroleum resources in politically unstable parts of the world are driving a world-wide search for new sources of liquid fuels. Biomass has long been regarded as a source of renewable sugars that could be fermented by microorganisms to make biofuels (*1*). The key barrier to adopting biomass derived fuels has been the high costs of conversion to monosaccharides (*2*). Enzymatic depolymerization is too slow and chemical depolymerization forms too many unwanted side products. The key applied goal of my thesis research was to identify new ways to enzymatically depolymerize lignocellulosic biomass so that low cost cellulosic feedstocks could be used for biofuel production.

Lignocellulosic biomass represents a large, untapped, and renewable resource for the production of fuels and chemicals. In 2005, the US Department of Energy released the *Billion-Ton Study* which reported on the potential biomass resources available in the United States (*3*). The purpose of the analysis was to determine if 30% of liquid transportation fuels could be displaced by fuels sourced from renewable biomass. Using conservative estimates of biomass availability, the report concluded that with minimal changes to agricultural practices a 1.3 billion ton annual supply of biomass could be produced in the U.S.

There are three categories of biomass resources produced at large enough scale to be considered for high volume fuel production: forestland, agricultural, and secondary residues and waste. Each type of biomass is sourced from a different place and would require substantially different conversion technologies for biofuel production. A large amount of research and development has focused on the conversion of existing agricultural wastes, including corn stover and sugarcane bagasse as well as potential dedicated energy crops, like switchgrass, energy cane, and *Miscanthus* (*4, 5*). Dedicated energy crops provide several environmental and potentially economic advantages as compared to agricultural waste. The productivity of dedicated energy crops already exceeds that of the major cereal crops which have been domesticated and improved by humans for several thousand years. The inputs for the production of cereal crops, including tillage, fertilizer, and irrigation are also very energy intensive.

Miscanthus X giganteus is a C4 perennial grass that is being developed as a potential bioenergy crop (Figure 1) (*4*). A key advantage of *Miscanthus* compared to other potential bioenergy crops like sugar cane and energy cane is cold tolerance (*6*). *Miscanthus* can be grown throughout most of the Midwestern and Southeastern portions of the U.S., while sugarcane and energy cane are limited to locations with near tropical climates (Florida, Louisiana, Hawaii). After the initial establishment, *Miscanthus*

can be harvested for several years without replanting (7). At the end of the growing season the plant senesces and most of the nutrients are transported to the root system where they can be reused during the next year's growth. In the U.S., 30 tons or more of dry biomass from *Miscanthus X giganteus* has been produced per hectare on cropland in field trials (8). Currently, the establishment costs for *Miscanthus* are high because it has to be propagated as rhizomes and mechanized farm equipment has not yet been developed for efficient planting. Recent estimates for the break even cost to produce *Miscanthus* are as low as $46 per dry ton in Missouri (9). These costs are comparable to energy cane production costs in Brazil, which are estimated to be as low as $34 per dry ton. Future improvements in agronomical practices, traditional plant breeding, and genetic tools will further lower the costs to produce biomass in the U.S.

Figure 1. *Miscanthus X giganteus* growing at the University of Illinois. Photo taken early in the growing season. Will Beeson is pictured on the right (ht. = 6 ft. 5 in.).

First generation biofuels have been produced from starch, plant oil, or sucrose. In 2010, 13.2 billion gallons of ethanol were produced in the United States with most of the production coming from the conversion of corn starch (9). Although the raw feedstock costs for first generation biofuels are higher, the conversion cost is much lower than for lignocellulosic material. Corn starch is composed principally of α-1-4 linked glucans. Two main enzymes are used to depolymerize corn starch, α-amylase and glucoamylase. The α-amylase works as an endoglucanase, rapidly decreasing the viscosity and degree of polymerization of the starch. Amylases used for corn starch hydrolysis have been engineered to work at temperatures above 90 °C where the corn starch is gelatinized. Glucoamylase is then added and hydrolyzes the oligosaccharides to glucose which is typically simultaneously fermented by yeast.

Second generation biofuels will be produced from lignocellulosic materials. Lignocellulose is derived from the cell walls of plants and is principally composed of cellulose, hemicellulose, and lignin (*3*). The relative amounts of the plant cell wall constituents varies from plant to plant, but for perennial grasses, like *Miscanthus*, cellulose constitutes 43 percent, hemicellulose 21 percent, and lignin 21 percent of the of the total dry biomass (Bauer, S., unpublished). To understand the challenges associated with depolymerization of the plant cell wall, a more detailed analysis of the composition of the cell wall is necessary.

1.2 Plant cell wall composition and structure

Plant cell walls are complex structures containing cellulose, hemicellulose, lignin, pectins, and proteins (*10*). The organization of plant cell walls can be thought of as a multicomponent gel matrix consisting of heterogeneous non-cellulose polysaccharides reinforced by crystalline cellulose microfibrils and structural proteins (Figure 2). The various polymer components of the wall may be linked via covalent or non-covalent interactions. The plant cell wall can be further divided into a primary cell wall and a secondary cell wall. The primary cell wall is produced first and is composed mostly of cellulose, hemicellulose, and pectins. The secondary cell wall is produced after the cell has stopped expanding and contains more cellulose and lignin. The secondary cell wall waterproofs the cell and provides extra rigidity. A key function of the cell wall is to protect the plant cell from microbial and enzymatic assault. The lignin and cellulose polymers are especially resistant to enzymatic degradation.
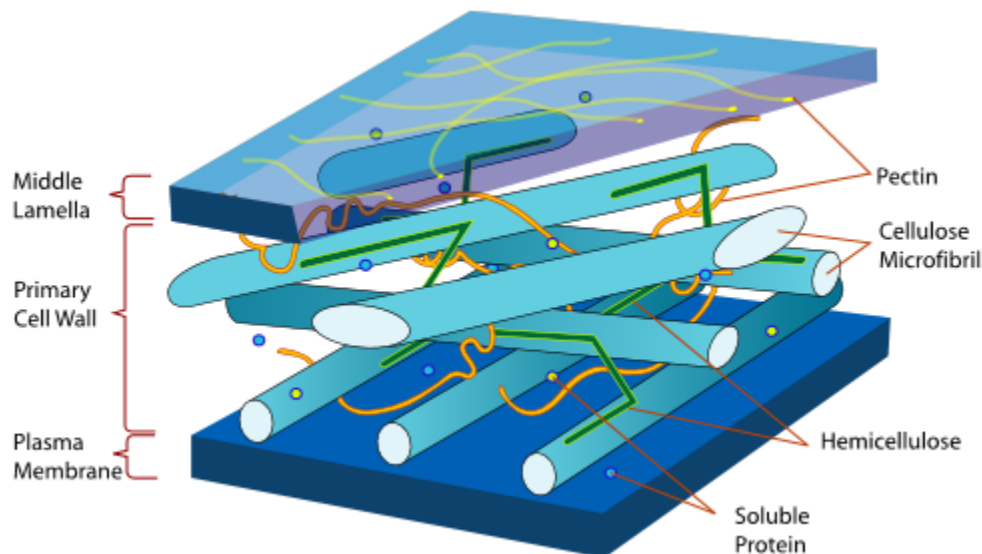

Figure 2. Schematic of the plant cell wall (*11*).

Cellulose is the most abundant polysaccharide on earth and an essential structural component of the plant cell wall that must be utilized for economical biofuel

production. Cellulose is a linear homopolymer of β-1-4-linked β-D-glucopyranose units. The pyranose rings are known to be in the chair conformation $^4C_1$, with the hydroxyl groups in equatorial orientation. Each glucopyranose unit is inverted 180 degrees relative to the previous monomer unit facilitating inter and intra-chain hydrogen bonding. The average degree of polymerization of native cellulose in primary cell walls is around 6,000 and up to 14,000 in secondary walls (*12*). In native cellulose, the chains are all oriented in the same direction and form microfibrils, probably consisting of 24-36 individual chains (*13*). The microfibrils are thought to be formed by the action of large, multi-subunit cellulose synthases present in the plasma membranes of plant cells that coordinately secrete growing cellulose chains into the cell wall (*14*).

Hemicelluloses are an important group of polysaccharides present in the plant cell wall that will also need to be utilized for next generation biofuel production. In contrast to cellulose, hemicellulose is a general designation that refers to the amorphous and branched polysaccharides of mostly arabinose and xylose, which are linked through β-1-4 glycosidic bonds (*15*). Both arabinose and xylose are pentose sugars lacking the C6 hydroxymethyl group and thus do not form the strong inter and intra-chain hydrogen bonding networks that cellulose does. Covalent crosslinks between hemicellulose and lignin are known (*16*). Hemicellulose is also thought to coat the cellulose microfibrils and is probably within van der Waals distance of cellulose chains. Hemicellulose can be completely hydrolyzed by dilute acid pretreatments; however, it is resistant to base pretreatments (*17*).

Lignin is the third major component of the plant cell wall and gives the cell wall strength and rigidity. Lignin is enriched in woody biomass and in contrast to cellulose and hemicellulose is a polyphenolic material that contains no carbohydrate component (*15*). The structure of lignin is poorly understood; however, it is known to be composed principally of three monomer units with varying degrees of methylation: p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol (*18*). The monomer units are bonded to one another through carbon-carbon and carbon-oxygen bonds, possibly due to untemplated free radical polymerization (*18*); however, very little is known about lignin formation in the cell wall. The average molecular weight of lignin is estimated to be higher than 10,000 (*19*). The amount of lignin in the cell wall is thought to be strongly correlated to the recalcitrance of the biomass to hydrolysis by enzymes (*20*). Removal or modification of lignin in cellulosic biomass may reduce conversion costs.

1.3 Organisms capable of cellulose degradation

More than 10^{11} metric tons of cellulose are produced per year by plants (*21*). Cellulose is an important source of carbon and energy for many organisms. The most commonly known cellulose degraders are termites and ruminant animals. Despite the common public association of cellulose degradation with animal species, both termites

and ruminant animals have large consortia of gut microorganisms that are responsible for the hydrolysis of cellulose (*22, 23*). Indeed, the vast majority of cellulose utilization is microbial (*24*).

Microbial cellulose degradation occurs both aerobically and anaerobically. Anaerobic cellulose degradation has been estimated to account for approximately 10% of total cellulose degradation (*25*). Cellulose-degrading anaerobes are found in the guts of animals, bogs, waste water treatment facilities, and in hot springs. They include many species of *Clostridia*, *Fibrobacter*, and poorly studied species of anaerobic fungi. Anaerobic cellulose degrading organisms are thriftier than aerobes in their secretion of hydrolytic enzymes, probably because much less energy can be extracted from glucose in the absence of oxygen. Aerobic microorganisms are responsible for most of the cellulose degradation in the environment (*25*). Aerobic cellulose degrading microorganisms include fungi and bacteria which live in soils, leaf litters, and compost piles.

In terrestrial environments fungi are major contributors to carbon cycling. Fungi are a large group of eukaryotic organisms including molds, mushrooms, and yeasts (*26*). Fungi are used to produce important food products like bread, beer, soy sauce, wine, and truffles. Most fungi are filamentous multi-cellular organisms. All fungi are heterotrophic, getting energy and carbon for growth from the consumption of organic compounds in their environment. Since they lack digestive systems, fungi must secrete enzymes to convert large insoluble nutrient sources into water soluble metabolites that can be imported by transporters in the cell membrane. Hence, many fungi have been exploited by the industrial biotechnology industry for the production of proteases, cellulases, pectinases, hemicellulases, and several other classes of enzymes (*27*). Fungi are one of the most promising sources of enzymes for production of cellulosic biofuels because they have been engineered for many years to secrete very high concentrations of protein. Several reports in the literature claim protein production levels higher than 100 grams per liter from fungal fermentation broths (*28*).

1.4 Hydrolytic cellulose degradation

Cellulose is highly crystalline and completely insoluble in water and most organic solvents (*12*). Both of these properties make cellulose resistant to enzymatic and chemical hydrolysis. Enzymatic hydrolysis of cellulose is catalyzed by cellulases. Cellulases are the third largest industrial enzyme product worldwide. They are used in textile processing, laundry detergents, juice extraction, paper recycling, and as animal feed additives (*29*). If a large cellulosic biofuels industry takes hold, they will become the largest produced industrial enzyme. Most commercial cellulases are produced in fungal hosts, like *Trichoderma reesei*, *Humicola insolens*, and *Myceliophthora thermophila* (*28*). Even though the specific activity of cellulases is very slow when

compared to other enzymes, their overall rate enhancement is dramatic. The half-life of crystalline cellulose in water at pH 7.0 is estimated to be approximately 100 million years (*15*). In sharp contrast to native cellulose, regenerated or amorphous cellulose can be hydrolyzed rapidly by cellulases, similar to the rates of starch hydrolysis by amylases (*30*). Accessibility to the substrate is thus key for the enzymes' hydrolytic activity.
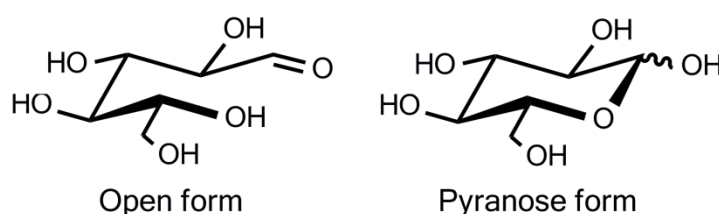
Figure 3. Structures of the open and closed form of D-glucose.

Before discussing mechanistic issues of cellulose depolymerization, a more thorough discussion of the chemistry of carbohydrates is warranted. The simplest repeating unit of cellulose is the monosaccharide glucose. Glucose can exist in an open form with a terminal aldehyde functional group or in a cyclic 6-membered ring hemiacetal form where the C5 hydroxyl group has added to the C1 aldehyde (Figure 3). The cyclic six membered ring form of glucose is commonly referred to as a pyranose. In aqueous solution the pyranose and open chain form of glucose are in dynamic equilibrium, with the pyranose accounting for greater than 99% of species in solution (*31*). The cyclic form of glucose can also exist in an alpha or beta form, depending on the configuration of the C1 hydroxyl group. If the C1 hydroxyl group is axial, the glucose is designated alpha, if equatorial it is beta. The conformation of the C1 position has large implications for the properties of glucose polymers. Glucose linked through α-1-4-glycosidic bonds forms the polymer amylose, while β-1-4-linked glucose polymers are known as cellulose. The axial linkages in amylose do not allow for the strong inter and intra-chain hydrogen bonding that makes cellulose so recalcitrant. Starches can be partially solubilized in water by heating to 100 °C (*32*), while cellulose remains completely insoluble and largely unchanged.

Figure 4. Acid catalyzed hydrolysis of cellulose.

The rate of chemical hydrolysis of glycosidic bonds is enhanced by general acid catalysis (Figure 4). Protonation of the glycosidic oxygen makes the reducing end carbohydrate moiety a better leaving group. The adjacent carbohydrate is eliminated and a transient cyclic oxonium ion is formed. Water then adds to the oxonium ion completing the hydrolysis reaction. Amorphous cellulose is hydrolyzed one to two orders of magnitude more rapidly than crystalline cellulose (*33*). Dilute acid treatment of cellulose incre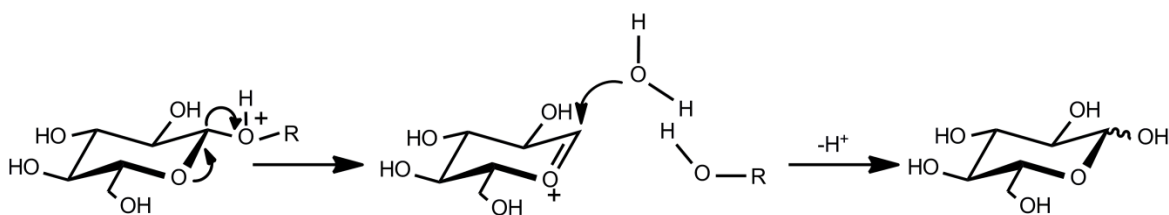ases the crystallinity over time because amorphous regions are more susceptible to hydrolysis. A possible explanation for the difference in reactivity is that hydrolysis of crystalline internal regions of cellulose will lead to "snap back" reactions where the glycosidic bond is reformed rather than being hydrolyzed by free water.

As of 2011, there are ten well studied families of cellulases (*15, 34*). The family designation is based on sequence homology, the fold of the protein, and the catalytic mechanism. Crystal structures have been solved for several different cellulases and there are least seven distinct protein folds known for cellulases (*34*). The diversity of cellulases found in nature may be due to the variations in the natural substrate, which is thought to be embedded in the complex milieu of the plant cell wall or because cellulose hydrolysis is still under positive selection.

There are three functional classes of cellulases produced by filamentous fungi: endoglucanases, reducing end exoglucanases, and non-reducing end exoglucanases (Figure 5) (*35*). For complete hydrolysis of cellulose to glucose, an additional type of enzyme, β-glucosidase may also be produced. Most fungi produce an intracellular and extracellular form of β-glucosidase. A trend that has emerged from the many cellulase crystal structures is that exoglucanases contain active sites present in tunnels (*36*), while endoglucanase active sites are present in solvent exposed clefts (*37*). This structural variation suggests that endoglucanases, as their name implies, access and cleave disordered or solvent exposed internal regions of the cellulose chains. Exoglucanases work from the ends of cellulose chains and thread the polysaccharide through their active site tunnels. Once an exoglucanase has engaged a chain it is hypothesized that it will process along it for several successive cuts.
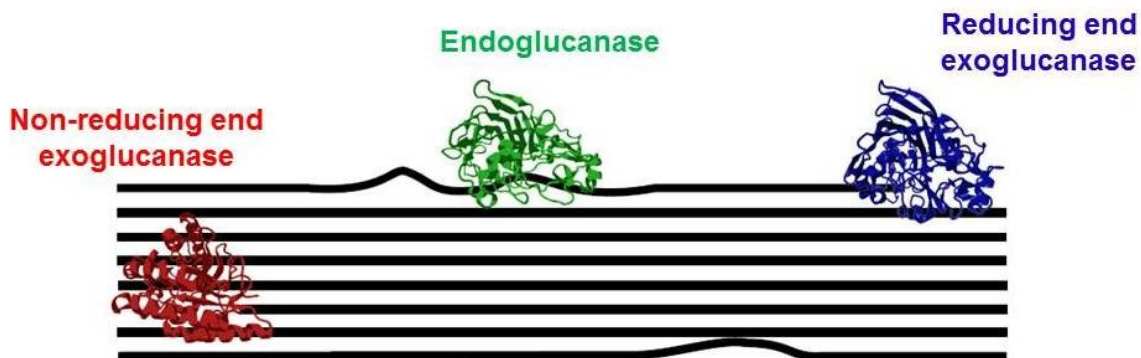

Figure 5. Schematic showing endoglucanases (*38*) and exoglucanases (*36*) acting on cellulose.

Further experimental data to support these models have been provided over the years. Endoglucanases rapidly decrease the viscosity of solutions of carboxymethyl cellulose, a soluble cellulose derivative, while exoglucanases have little effect on the viscosity (*15*). Viscosity is directly proportional the number average degree of polymerization of the cellulose and the drastic reduction in viscosity is consistent with a reduction in the length of the cellulose. When incubated with insoluble forms of cellulose, endoglucanases produce many more insoluble reducing ends than do exoglucanases. Processivity has been a more difficult property to study, but recent work using single molecule imaging techniques have confirmed that exoglucanases will process along an individual chain for many successive cuts (*39, 40*).

An additional feature of many cellulases is the presence of cellulose binding modules (CBMs) (*34*). Typically, CBMs are present at the N- or C- terminus of a cellulase, linked to the catalytic domain by a short flexible linker region. In fungi, the CBM is a short, 30 amino acid domain containing three conserved aromatic residues (*41*). The aromatic residues are spaced and oriented so they form a flat plane that can hydrogen bond and form stacking interactions with the pyranose rings present on the surface of crystalline cellulose. The presence of the CBM increases the local concentration of the catalytic domain on the surface of the cellulose. Because catalytic domains have only weak affinity for cellulose, the removal of the CBM will generally drastically decrease cellulase activity on crystalline cellulose. Interestingly, the activity on amorphous cellulose has been reported to increase when the CBM is removed (*42*). There are about as many naturally occurring cellulases that do not contain a CBM as those that do. Clearly there is a tradeoff associated with the enhanced binding that may not always be beneficial. Often CBMs are also attached to catalytic domains that have no hydrolytic activity on cellulose, including xylanases, xyloglucanases, cellobiose dehydrogenases, and cutinases (*43*). The attachment of CBMs to these diverse catalytic domains further supports the model of cellulose fibrils being embedded in a complex matrix of other polysaccharides in the plant cell wall.

It was observed in the 1950s that no single cellulase could completely hydrolyze crystalline cellulose (*44*). Since then synergy between cellulases has been studied extensively and several important conclusions on the topic have been reached (*35, 45*). The simplest model of synergy is based on the notion that endoglucanases will make cuts in internal regions of cellulose chains which generates new chain ends for action of exoglucanases. There is no evidence that endoglucanase-exoglucanase synergy requires multi-protein complexation and in general any endoglucanase will show synergy with any exoglucanase (*15*). Exoglucanases can also show synergy with one another, so long as they target opposite ends of the cellulose chains. Synergy effects are largest in the digestion of crystalline cellulose and most amorphous or soluble substrates show only very small effects. As more genomes have been sequenced, an

emerging trend is that many species of fungi secrete several different endoglucanases. Whether these endoglucanases have different substrate specificities or enhanced synergism in the presence of each other is an area of active research.

1.5 Oxidative cellulose degradation

Hydrolytic cellulose degradation has been extensively studied over the last 60 years and is a relatively well understood biological process (*46*). Recent genome sequencing projects of diverse species of fungi have suggested that some fungi may be using other mechanisms to degrade cellulose. The genome sequence of *Postia placenta*, a brown-rot fungus known to degrade cellulose, lacks genes encoding exoglucanases, or any proteins with a cellulose binding module (*47*). The dogma in the field is that for efficient crystalline cellulose degradation there must be both endoglucanases and exoglucanases. During growth on cellulose *Postia placenta* upregulates the expression of genes encoding iron reductases, quinone reductase, and multiple diverse oxidases. It was speculated that *Postia placenta* might be using oxidative mechanisms to degrade cellulose, although direct biochemical evidence in support of this hypothesis has not been reported (*47*).

The most prominent model for oxidative degradation of cellulose is based on the Fenton reaction (Figure 6) (*48*). The Fenton reaction was developed by Henry Fenton in the late 1800s as a reagent to destroy tartaric acid (*49*). Later, the mechanism of action was better understood by Haber and Weiss and the reaction was developed as a way to oxidize toxic organic compounds (*50*). In the Fenton reaction, an aqueous solution of hydrogen peroxide is mixed with ferrous iron leading to the formation of hydroxyl radicals. More than 20 years ago it was reported that several species of fungi produced hydrogen peroxide during growth on cellulosic substrates (*51, 52*). The availability of hydrogen peroxide and iron complexes in wood led researchers to propose an oxidative degradation pathway based on the Fenton reaction.

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + HO^- + HO^\bullet$$
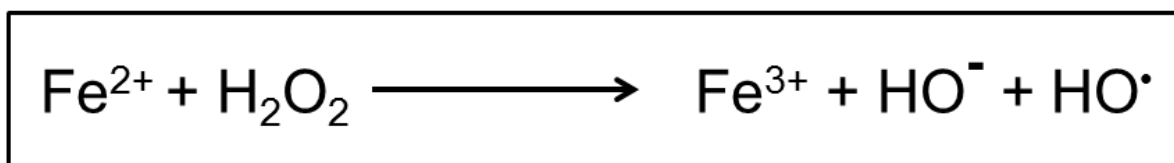
Figure 6. The Fenton reaction.

In the Fenton model, oxidases and reductases secreted by fungi during growth on cellulose participate in redox cycling reactions with molecular oxygen and transition metals which lead ultimately to the production of hydroxyl radicals which non-specifically degrade plant cell wall polymers (*53*). Hydroxyl radicals have been reported to abstract hydrogen atoms from cellulose and other polysaccharides with rate constants near 10^9

9

$M^{-1}s^{-1}$ (*54*). After hydrogen atom abstraction, carbon centered radicals will be formed which can then rapidly react with molecular oxygen to give peroxyl radical species. If the radical is formed on C1 or C4, an elimination reaction may take place which results in glycosidic bond cleavage and ejection of superoxide (*55*). In the most well developed Fenton model, an extracellular heme-flavoprotein, cellobiose dehydrogenase, which is produced by many fungi during growth on cellulose, is thought to play a key role (*48, 56*).
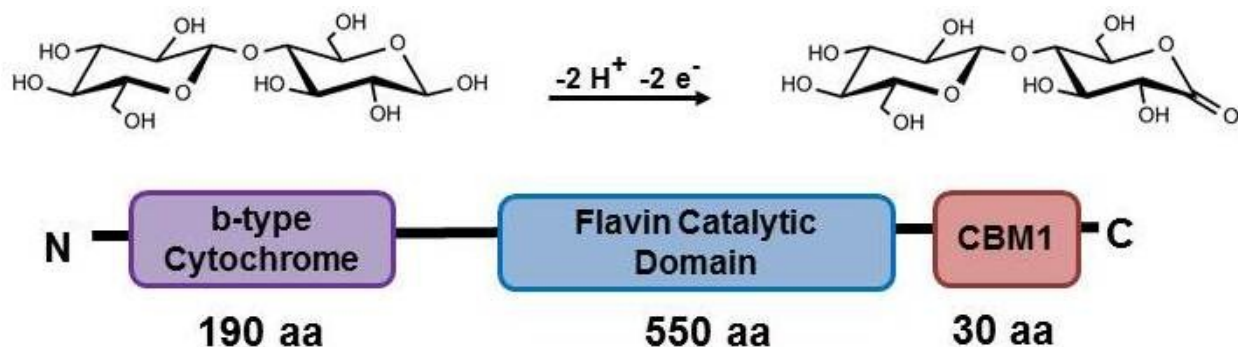
Figure 7. Domain architecture and reaction catalyzed by cellobiose dehydrogenase.

Cellobiose dehydrogenase (CDH) is the only known example of an extracellular heme-flavoprotein and is produced only by filamentous fungi during growth on cellulose. CDH is a multi-domain enzyme consisting of an N-terminal cytochrome domain, a flavin catalytic domain, and in some species a C-terminal cellulose binding module (Figure 7) (*57*). It catalyzes the 2-electron oxidation of the reducing end of cellobiose and longer chain cellodextrins to the corresponding aldonolactones. Glucose is a very poor substrate for CDH; the apparent second order rate constant for cellobiose is 87,000 times higher than that for glucose (*57*). Substrate oxidation takes place in the flavin domain and electrons are then transferred to a heme prosthetic group bound in the cytochrome domain. The heme iron is complexed by absolutely conserved methionine and histidine residues which cannot be displaced by diatomic ligands or azide, suggesting a potential role in outer-sphere electron transfer reactions (*58*). To regenerate the enzyme for subsequent turnover, the electrons would then need to be passed on to an exogenous electron acceptor. CDH has very low reactivity with molecular oxygen and it has been speculated that the natural electron acceptor might be ferric-oxalate complexes present in wood or quinones secreted by the fungus (*57*). If the electrons were transferred to ferric complexes, CDH would be playing a critical role in the generation of ferrous iron for Fenton chemistry. Many experiments have been performed that show CDH is able to transfer electrons to quinones and metal ions at more than 20-fold the rate of reduction of molecular oxygen.

From a purely chemical standpoint, the scope of the reactions proposed in the Fenton chemistry model of oxidative cellulose degradation is reasonable, but from a biological perspective, there are potentially many pitfalls to uncontrolled generation of hydroxyl radicals. Hydroxyl radicals are some of the most reactive chemical species known, and it is unclear how they would be targeted to the plant cell wall and kept away from the penetrating hyphal tips of the fungus. Furthermore, the presence of free metal ions has been reported to be highly inhibitory to cellulase activity *in vitro (59)*. Although there are conflicting reports (*60*), under most reaction conditions the addition of CDH to mixtures of free cellulases has no effect on cellulose hydrolysis or is inhibitory (Beeson W.T., unpublished results).

Clearly, CDH is produced in the presence of many cellulases and it would be counter-productive for the fungus to secrete an enzyme that reduces the activity of cellulases it relies upon for food. The genome of the model filamentous fungus, *Neurospora crassa,* contains two genes encoding predicted CDHs and many genes encoding predicted cellulases and other carbohydrate active enzymes (*61*). A central theme of the work presented in this thesis is aimed at using a combination of genetic and biochemical experiments to understand how *N. crassa* degrades plant cell wall material and what role cellobiose dehydrogenase plays in the process. Chapter one describes a systems biology analysis of plant cell wall utilization in *N. crassa* where many genes and proteins were identified as likely to be involved in cell wall degradation. Chapter two reports on the first biochemical characterization of a fungal extracellular aldonolactonase induced during growth on cellulose. These aldonolactonases are highly conserved in cellulolytic fungi and may play a secondary role in oxidative cellulose degradation pathways. The final four chapters describe genetic and biochemical experiments performed to determine the contribution of oxidative enzymes to cellulose degradation in *N. crassa.* These results set the foundation for future work to develop more cost efficient enzyme systems for lignocellulosic biomass conversion. The addition of oxidative enzymes to commercial cellulase mixtures could lower the amount enzyme required for biofuel production by as much as 2-4 fold.

Chapter 2: Systems analysis of plant cell wall degradation by *Neurospora crassa*

2.1 Abstract

The filamentous fungus *Neurospora crassa* is a model laboratory organism, but in nature is commonly found growing on dead plant material, particularly grasses. Using functional genomics resources available for *N. crassa*, which include a near-full genome deletion strain set and whole genome microarrays, we undertook a system-wide analysis of plant cell wall and cellulose degradation. We identified approximately 770 genes that showed expression differences when *N. crassa* was cultured on ground *Miscanthus* stems as a sole carbon source. An overlap set of 114 genes was identified from expression analysis of *N. crassa* grown on pure cellulose. Functional annotation of up-regulated genes showed enrichment for proteins predicted to be involved in plant cell wall degradation, but also many genes encoding proteins of unknown function. As a complement to expression data, the secretome associated with *N. crassa* growth on *Miscanthus* and cellulose was determined using a shotgun proteomics approach. Over 50 proteins were identified, including 10 of the 23 predicted *N. crassa* cellulases. Strains containing deletions in genes encoding 16 proteins detected in both the microarray and mass spectrometry experiments were analyzed for phenotypic changes during growth on crystalline cellulose and for cellulase activity. While growth of some of the deletion strains on cellulose was severely diminished, other deletion strains produced higher levels of extracellular proteins that showed increased cellulase activity. These results show that the powerful tools available in *N. crassa* allow for a comprehensive system level understanding of plant cell wall degradation mechanisms used by a ubiquitous filamentous fungus.

2.2 Introduction

Neurospora crassa is a well-known model organism that has been used for over 90 years to study genetics, biochemistry and fungal biology (*62*). Many *N. crassa* isolates have been recovered from sugar cane, which is closely related to *Miscanthus*, an attractive crop for biofuel production (*63-65*). Although it was shown to degrade cellulose more than 30 years ago (*66, 67*), relatively little has been reported on plant biomass utilization by *N. crassa*. The *N. crassa* genome is predicted to contain twice as many cellulases as *H. jecorina* (*68*), as well as many hemicellulases and other enzymes involved in plant biomass degradation. Genetic and molecular tools to manipulate *N. crassa* are extensive (*62*) as are genomic resources, including whole genome microarrays and a near full genome deletion strain set (*69*). *N. crassa* is the only example of a model organism that also happens to be a proficient degrader of plant cell walls.

In this study, we exploit functional genomic resources to perform a systems analysis of the *N. crassa* transcriptome associated with complex plant biomass and pure cellulose utilization. In addition, the secretome of *N. crassa* grown under identical conditions was analyzed using a shotgun proteomics approach. We evaluated strains containing deletions in genes encoding proteins identified from overlapping transcriptome and secretome datasets for their ability to utilize cellulose and for cellulase activity. From this analysis, we identified known proteins involved in plant cell wall degradation, but also proteins of unknown function that affect cellulose degradation and cellulase activity. Taken together, these data begin to unravel the functionally distinct strategies used by *N. crassa* to degrade plant cell walls and highlight how a systems biology approach using genomic resources is a powerful tool to identify novel and industrially important components associated with plant cell wall degradation.

2.3 Results
2.3.1 Transcriptome analysis of *N. crassa* grown on *Miscanthus* and Avicel

Growth and cellulase activity of *N. crassa* (FGSC 2489) cultured on minimal medium with crystalline cellulose (Avicel) as the sole carbon source was similar to that of *H. jecorina* (QM9414) (Fig. 1); *N. crassa* completely degraded Avicel in ~4 days. *N. crassa* also grew rapidly on ground *Miscanthus* stems, suggesting functional cellulase and hemicellulase degradative capacity. To determine the transcriptome associated with plant cell wall deconstruction, we used full genome microarrays (*70-72*) to monitor gene expression profiles during growth of *N. crassa* on ground *Miscanthus* stems. RNA sampled from *N. crassa* grown for 16 hours of growth on sucrose was compared to RNA from *N. crassa* grown on *Miscanthus* medium at 16 hours, 40 hours, 5 days and 10 days (Fig. 2).
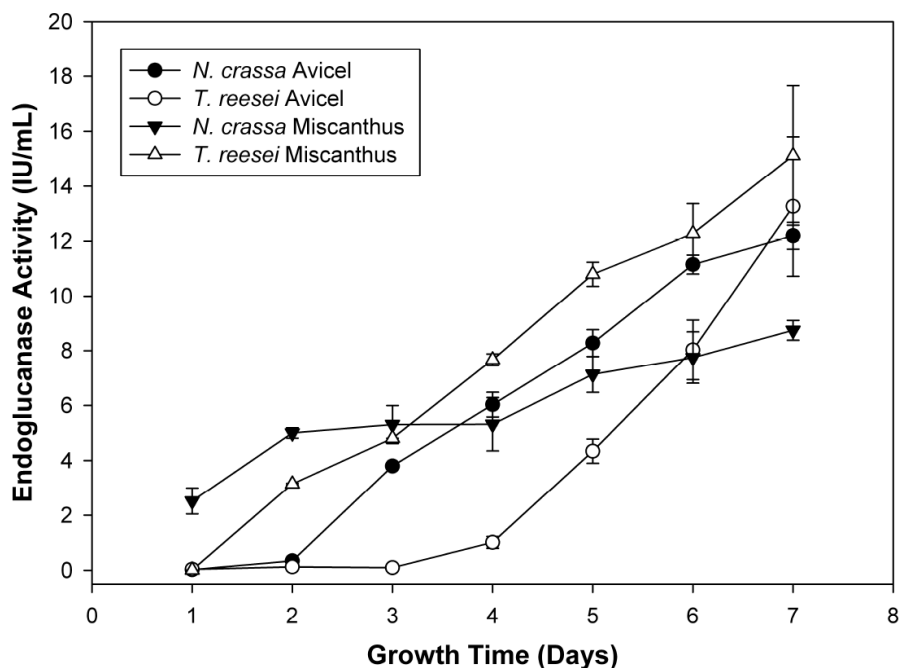
Figure 1. Analysis *N. crassa* FGSC2489 and *T. reesei* QM9414 endoglucanse activity when grown on *Miscanthus* and Avicel as a sole carbon source. Endoglucanase activity in culture filtrates of *N. crassa* WT strain FGSC2489 and *T. reesei* QM9414. *N. crassa* was grown on Vogel's minimal medium containing 2% of either Avicel or *Miscanthus* powder as a sole carbon source at 25℃ *T. reesei* strain was inoculated in MA medium with either 1% Avicel or *Miscanthus* powder as sole carbon source at 25 ℃ Both strains were inoculated with the same amount of conidia (1 x 10^6 /ml in 100ml culture). The endoglucanase activity in the cultures at different time points were measured at pH 4.5 using Azo-CM-cellulose as a substrate according to the manufacturer's instructions (Megazyme, Ireland).

A total of 769 *N. crassa* genes showed a statistically significant difference in relative expression level among the four *Miscanthus* samples as compared to the sucrose sample. Hierarchical clustering showed that these genes fell into three main clusters (Fig. 2A). The first cluster of genes (C1; 300 genes) showed the highest expression levels in minimal medium with sucrose. Functional category (FunCat) analysis (*73*) of these genes showed an enrichment for ribosomal proteins and other functional categories associated with primary metabolism. The second cluster (C2) included 327 genes that showed the highest expression levels in *Miscanthus* cultures at later time points (40 hrs to 10 days; Fig. 2A). Within this group were 89 genes that showed a high relative expression level in *Miscanthus* cultures at all time points. FunCat analysis (*73*) of the remaining 238 genes showed one functional category (C-compound and carbohydrate metabolism) was slightly enriched.
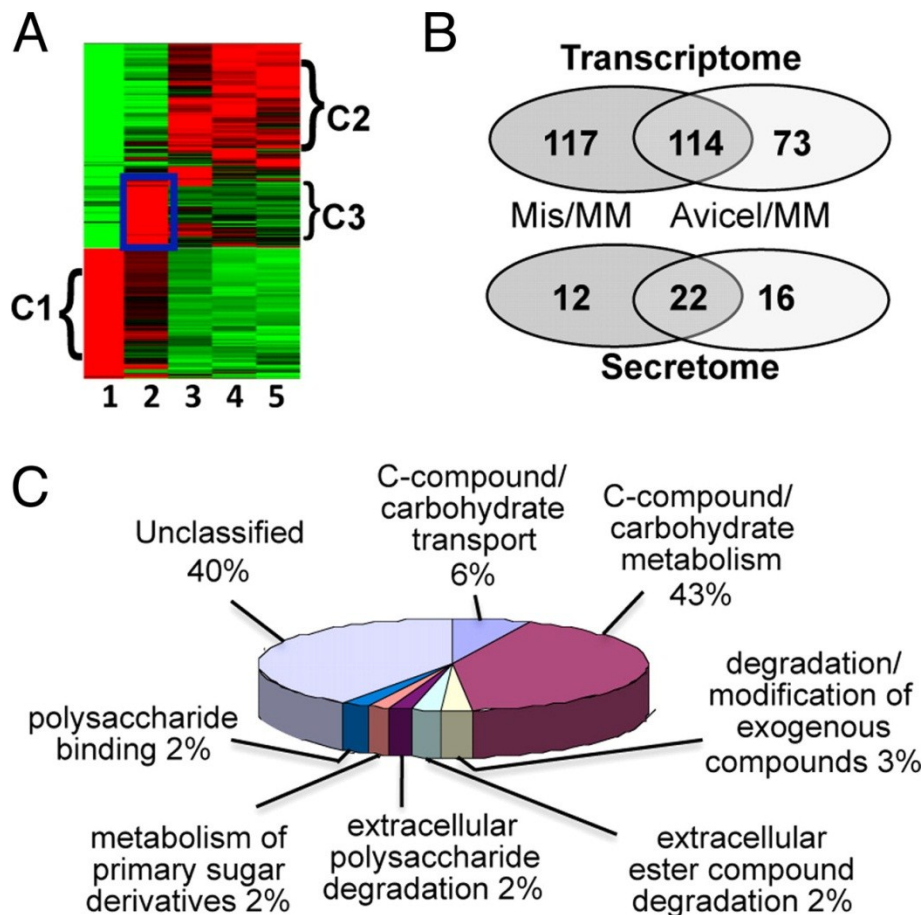
Figure 2. Transcriptional profiling of *N. crassa* grown on *Miscanthus* and Avicel. (*A*) Hierarchical clustering analysis of 769 genes showing expression differences in *Miscanthus* culture. Red indicates higher relative expression and green indicates lower relative expression. Lane 1: A 16 h *N. crassa* culture grown in sucrose minimal medium. Lane 2: A 16 h culture with *Miscanthus* as a sole carbon source. Lanes 3–5: Expression profiles from cultures grown on *Miscanthus* for 40 h, 5 days, and 10 days. The C3 cluster showed increased expression levels of most of the cellulase and hemicellulase genes (boxed). (*B*) Overlap in expression profiles between the *N. crassa Miscanthus* versus Avicel grown cultures (*Top*). Overlap of proteins in culture filtrates detected by tandem mass spectrometry (*Bottom*). (*C*) Functional category analysis (16) of the 231 genes that showed a significant enrichment ($P < 0.001$) in relative expression levels in *Miscanthus* cultures.

A third cluster of 142 genes showed the highest relative expression level after 16 hours of growth of *N. crassa* on *Miscanthus* (C3, Fig. 2A). FunCat analysis (*73*) of these 142 genes plus the 89 genes that showed high expression levels in *Miscanthus* cultures at all time points (C3+ cluster; total 231 genes) showed an enrichment for proteins involved with carbon metabolism, including predicted cellulases and hemicellulases (Fig. 2C). Of the 23 predicted cellulase genes in the *N. crassa* genome, 18 showed significant increases in expression levels during growth on *Miscanthus* (Table 1), particularly at the 16 hour time point (Fig. 3). Five genes showed an increase in expression level over 200-fold (*cbh-1* (CBH(I); NCU07340, *gh6-2* (CBH(II)-like gene;

NCU09680), *gh6-3* (NCU07190) and two GH61 genes (*gh61-4;* NCU01050 and NCU07898)).
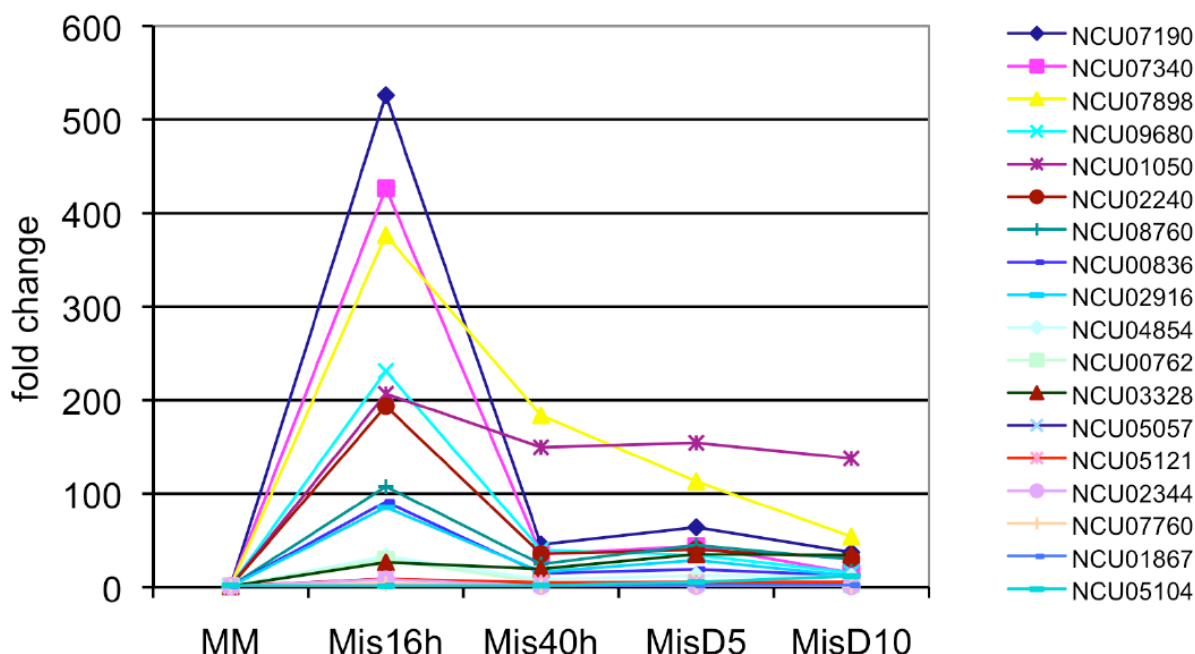
Figure 3. Fold change induction of cellulases during growth on *Miscanthus* relative to sucrose. Most cellulases show highest induction at 16 hours, then a lower, but steady level of induction at later time points.

Plant cell walls are complex structures composed of cellulose microfibrils, hemicellulose, lignin, pectin, cutin, and protein. Thus, we compared expression profiles of *N. crassa* grown on *Miscanthus* to expression profiles of *N. crassa* grown on Avicel, a pure form of crystalline cellulose. Over 187 genes showed a significant increase in relative expression level during growth of *N. crassa* on Avicel. Of these genes, 114 overlapped with the 231 genes in the C3+ cluster (Fig. 2B). FunCat analysis of the 114-

Table 1. Predicted cellulases genes in *Neurospora crassa*

Gene	GH family	CBM1	SP	MS	EL *Miscanthus*	EL Avicel
NCU00762	5	Yes	Yes	Both	29.6	31.5
NCU03996	6	No	No	ND	ND	ND
NCU07190	6	No	Yes	Both	526.0	119
NCU09680	6	Yes	Yes	Both	230.9	251.3
NCU04854	7	No	Yes	ND	32.9	10.8
NCU05057	7	No	Yes	Both	8.7	7.4
NCU05104	7	No	Yes	ND	11.6	NC[7]
NCU07340	7	Yes	Yes	Both	426.4	382.2

Gene	GH family	CBM1	SP	MS	EL *Miscanthus*	EL Avicel
NCU05121	45	Yes	Yes	avi	8.6	17.2
NCU00836	61	Yes	Yes	ND	91.2	31
NCU01050	61	No	Yes	Both	206.7	382.1
NCU01867	61	Yes	Yes	ND	2.2	NC
NCU02240	61	Yes	Yes	avi	193.5	84
NCU02344	61	No	Yes	ND	8.1	4.1
NCU02916	61	Yes	Yes	ND	85.2	17.7
NCU03000	61	No	Yes	ND	NC	ND
NCU03328	61	No	Yes	ND	26.4	23.8
NCU05969	61	No	Yes	ND	ND	12.7
NCU07520	61	No	Yes	ND	ND	ND
NCU07760	61	Yes	Yes	ND	3.7	NC
NCU07898	61	No	Yes	Both	376.3	230
NCU07974	61	No	Yes	ND	NC	NC
NCU08760	61	Yes	Yes	Both	107.5	44.7

* GH, glycoside hydrolase; CBM1, carbohydrate binding module; SP, signal peptide prediction; MS, mass spectrometry analysis; EL, relative expression level; ND, not detected; NC, no change.

overlap gene set showed a clear enrichment for genes predicted to be involved in carbon metabolism. Within this gene set, there was a further enrichment for secreted proteins (53 of the 114 gene products). Of the 53 genes, 32 encode predicted proteins with annotation suggesting a role in plant cell wall degradation, while 16 encode putative or hypothetical proteins. The remaining 61 genes encode predicted intracellular proteins, including ten predicted major facilitator superfamily transporters (NCU00801, NCU00988, NCU01231, NCU04963, NCU05519, NCU05853, NCU05897, NCU06138, NCU08114 and NCU10021) and 23 putative or hypothetical proteins.

Of the 117 genes within the *Miscanthus*-specific cluster (Fig. 2B), 37 encoded proteins predicted to be secreted. Nine predicted hemicellulases or enzymes related to the degradation of hemicellulose were identified (NCU00710, NCU04265, NCU04870, NCU05751, NCU05965, NCU09170, NCU09775, NCU09923 and NCU09976). The remaining 80 *Miscanthus*-specific genes encode predicted intracellular proteins, including genes involved in the metabolism of pentose sugars (for example, NCU00891, xylitol dehydrogenase and NCU00643, a predicted arabinitol dehydrogenase), a predicted sugar transporter (NCU01132), and 48 proteins of unknown function.

2.3.2 Secretome analysis of *N. crassa* grown on *Miscanthus* and Avicel

Lignocellulose degradation by fungi requires the secretion of proteins associated with depolymerization of cell wall constituents (*74*). To compare with transcriptional profiling data, we analyzed the secretome of *N. crassa* using a shotgun proteomics approach (Fig. 1B). Supernatants from three and seven day old *Miscanthus* and Avicel cultures were digested with trypsin and analyzed by liquid chromatography nano-electrospray ionization tandem mass spectrometry (MS); datasets from the 3 and 7 day samples showed no significant difference. Secreted proteins that bound to phosphoric acid swollen cellulose (PASC) were enriched and also analyzed by MS.

A total of 50 proteins were identified with confidence by tandem MS (Table 2). There were 34 proteins detected in the *Miscanthus* grown *N. crassa* cultures, while 38 proteins were identified from Avicel grown cultures; twenty-two proteins were detected in both samples. Of these 22 proteins, 21 were predicted to be secreted based on computational analyses and 19 showed increased expression levels in both the *Miscanthus* and Avicel grown cultures. The overlap dataset included eight of the 23 predicted cellulases in *N. crassa* (Table 1). There were also five predicted hemicellulases, a predicted β-glucosidase (*gh3-4;* NCU04952), five proteins with predicted activity on carbohydrates, and two proteins with unknown function (NCU07143 and NCU05137) (Table 2).

There were 16 proteins identified with confidence in the Avicel culture only and 14 of these were predicted to be secreted (Table 2) including two predicted cellulases (*gh61-1*; NCU02240 and *gh45-1*; NCU05121), one xylanase (*gh11-1*; NCU02855), one predicted protease (NCU04205), three other proteins with predicted activity on carbohydrates (NCU08909, NCU05974 and *gh30-1* (NCU04395)), three *Neurospora*-specific proteins of unknown function, and four conserved hypothetical proteins, including one protein with a cellulose binding domain (NCU09764). Twelve proteins were specific for culture filtrates of *Miscanthus* cultures and seven of these had predicted secretion signals (Table 2). Three of the five predicted intracellular proteins were conserved hypothetical proteins. The remaining two included a predicted glyoxal oxidase (NCU09267, identified from the *N. crassa Miscanthus* transcriptome) and a nucleoside diphosphate kinase (*ndk-1;* NCU04202, not identified in the *N. crassa* transcriptome). The seven proteins predicted to be secreted included three predicted esterases (NCU04870, NCU05159, and NCU08785), two predicted xylanases (GH51; NCU02343 and GH54; NCU09775), a predicted β-xylosidase (*gh3-7*; NCU09923) and a conserved hypothetical protein (NCU05751).

Many plant cell wall degrading enzymes contain a cellulose-binding module (CBM), which aids in attachment of the enzyme to the substrate (*75*). Within the *N. crassa* genome, 19 genes are predicted to encode proteins with a CBM1 domain (*76*).

Sixteen of these genes showed an increase in relative gene expression in *Miscanthus*-grown cultures. From the 50 proteins identified by MS, 11 contain a CBM1 domain. We used PASC to enrich for proteins that bind to cellulose (see Materials and Methods). Nine proteins that bound to PASC from the supernatant of *Miscanthus*-grown *N. crassa* cultures and eight proteins from Avicel supernatant were identified by MS; seven proteins were identified in both. These included NCU00206, a predicted cellobiose dehydrogenase; *gh5-1* (NCU00762), a predicted endoglucanase; NCU05955, a predicted GH74 xyloglucanase; *gh11-2* (NCU07225), a predicted endoxylanase; *cbh-1* (NCU07340); *gh61-5* (NCU08760), a predicted endoglucanase, and *gh6-2* (NCU09680), a predicted cellobiohydrolase 2 precursor.

Table 2. Proteins identified by LC-MS/MS in culture filtrates of *Neurospora crassa* grown on Avicel or *Miscanthus*.

GENE ID	ANNOTATION	CULTURE
NCU00206	cellobiose dehydrogenase	BOTH
NCU00762	endoglucanase II (GH5)	BOTH
NCU01050	GH61 protein	BOTH
NCU04952	β-glucosidase (GH3)	BOTH
NCU05057	endoglucanase I (GH7)	BOTH
NCU05137	hypothetical protein	BOTH
NCU05924	endoxylanase (GH10)	BOTH
NCU05955	xyloglucanase (GH74)	BOTH
NCU07143	hypothetical protein	BOTH
NCU07190	endoglucanase 6 (GH6)	BOTH
NCU07225	endoxylanase (GH11)	BOTH
NCU07326	arabinofuranosidase (GH43)	BOTH
NCU07340	cellobiohydrolase I (GH7)	BOTH
NCU07898	GH61 protein	BOTH
NCU08189	endoxylanase (GH10)	BOTH
NCU08398	aldose-1-epimerase	BOTH
NCU08412	conserved hypothetical protein	BOTH
NCU08760	GH61 protein	BOTH
NCU09024	conserved hypothetical protein	BOTH
NCU09175	conserved hypothetical protein	BOTH
NCU09491	feruloyl esterase B	BOTH
NCU09680	cellobiohydrolase II (GH6)	BOTH
NCU00798	hypothetical protein	AV
NCU01595	SOF1	AV

19

Table 2 continued

GENE ID	ANNOTATION	CULTURE
NCU02240	GH61 protein	AV
NCU02696	DEAD DEAH box RNA helicase	AV
NCU02855	endoxylanase (GH11)	AV
NCU04205	peptidase A4 family	AV
NCU04395	endo-1,6-b-d-glucanase (GH30)	AV
NCU05121	endoglucanase V (GH45)	AV
NCU05134	hypothetical protein	AV
NCU05852	hypothetical protein	AV
NCU05974	similar to Mwg1	AV
NCU08171	predicted protein	AV
NCU08909	beta-1,3 glucanosyltransferase	AV
NCU08936	hypothetical protein	AV
NCU09046	hypothetical protein	AV
NCU09764	hypothetical protein	AV
NCU01651	hypothetical protein	MIS
NCU02343	α-l-arabinofuranosidase	MIS
NCU04202	nucleoside diphosphate kinase	MIS
NCU04870	acetyl xylan esterase	MIS
NCU05159	acetylxylan esterase	MIS
NCU05751	hypothetical protein	MIS
NCU06239	hypothetical protein	MIS
NCU08785	hypothetical protein	MIS
NCU09267	hypothetical protein	MIS
NCU09708	hypothetical protein	MIS
NCU09775	α-L-arabinofuranosidase	MIS
NCU09923	β-xylosidase (GH3)	MIS

2.3.3 Characterization of extracellular proteins and cellulase activity in strains containing deletions in genes identified in the overlap of the transcriptome and secretome datasets.

Of the 22 extracellular proteins detected in both the *Miscanthus* and Avicel grown cultures, homokaryotic strains containing deletions in 16 genes are available (*69*). None of these 16 deletion strains have been previously characterized with respect to plant cell wall or cellulose degradation by *N. crassa*. The 16 deletion strains were grown in media containing sucrose or Avicel as a carbon source. All strains showed a wild-type

phenotype on sucrose medium. The total secreted protein, endoglucanase activity, β-glucosidase activity, and aggregate Avicelase activity of Avicel-grown culture filtrates was measured after seven days and compared to the wild-type strain from which all mutants were derived (Fig. 4). SDS-PAGE of unconcentrated culture supernatants assessed the relative abundance of secreted proteins.

There were growth deficiencies on Avicel for strains containing deletions of two predicted exoglucanases, *Δcbh-1* (NCU07340) and *Δgh6-2* (NCU09680), plus a predicted β-glucosidase, *Δgh3-4* (NCU04952). The phenotype of the *cbh-1* mutant was the most severe; after seven days in culture much of the Avicel remained, while in the wild-type strain all of the Avicel was degraded. SDS-PAGE of extracellular proteins from 10 of the 16 deletion strains showed an altered protein profile with loss of a single band, allowing assignment of a particular protein band to a predicted gene (Fig. 4A, boxes). These included NCU00762 (*gh5-1),* NCU04952 (*gh3-4*), NCU05057 (*gh7-1*), NCU05137*,* NCU05924 *(gh10-1),* NCU05955, NCU07190 (*gh6-3*), NCU07326, NCU07340 (*cbh-1*) and NCU09680 (*gh6-2*).

For the majority of the deletion strains, the total secreted protein, endoglucanase, β-glucosidase, and Avicelase activities of the culture supernatants were similar to wild type (WT) (Fig. 4B, C). Deviations from this trend were seen with *Δgh5-1* (NCU00762), *Δgh3-4* (NCU04952), ΔNCU05137, *Δcbh-1* (NCU07340) and *Δgh6-2* (NCU09680). In *Δgh5-1* (NCU00762), *Δgh3-4* (NCU04952), and *Δcbh-1* (NCU07340) Avicelase, endoglucanase or β-glucosidase activities were lower than the corresponding WT activities. In particular, the deletion of NCU04952 eliminated all β-glucosidase activity from the culture supernatant, as evidenced by PNPGase activity and by higher levels of cellobiose and lower levels of glucose in the Avicelase enzyme assays (Fig. 4B, C). Unlike a WT strain, culture filtrates from ΔNCU04952 were completely unable to hydrolyze cellobiose to glucose, consistent with loss of β-glucosidase activity. Despite lowering endoglucanase activity, the culture filtrate from *Δgh5-1* (NCU00762) showed no significant deficiency in Avicelase activity relative to the WT strain (Fig. 4C). As expected, mutations in *cbh-1* (NCU07340) resulted in lower endoglucanase and Avicelase activity. A strain containing a deletion of NCU09680, encoding a CBH(II)-like protein (*gh6-2*), also showed reduced cellobiose accumulation (Fig. 4C).

Mutations in three strains resulted in an increased level of secreted proteins, especially CBH(I) (Fig. 4A); *Δgh3-4* (NCU04952), *Δgh7-1* (NCU05057) and a hypothetical protein gene, ΔNCU05137. The ΔNCU05137 mutant also showed increased endoglucanase, β-glucosidase and Avicelase activity (Fig. 4B, C). NCU05137 is highly conserved in the genomes of a number of filamentous ascomycete fungi, including other cellulolytic fungi, but notably does not have an ortholog in *H. jecorina*. It is possible that the increase in CBH(I) levels observed in *Δgh3-4*, *Δgh7-1* and ΔNCU05137 could be due to either increased secretion, protein stability or feedback
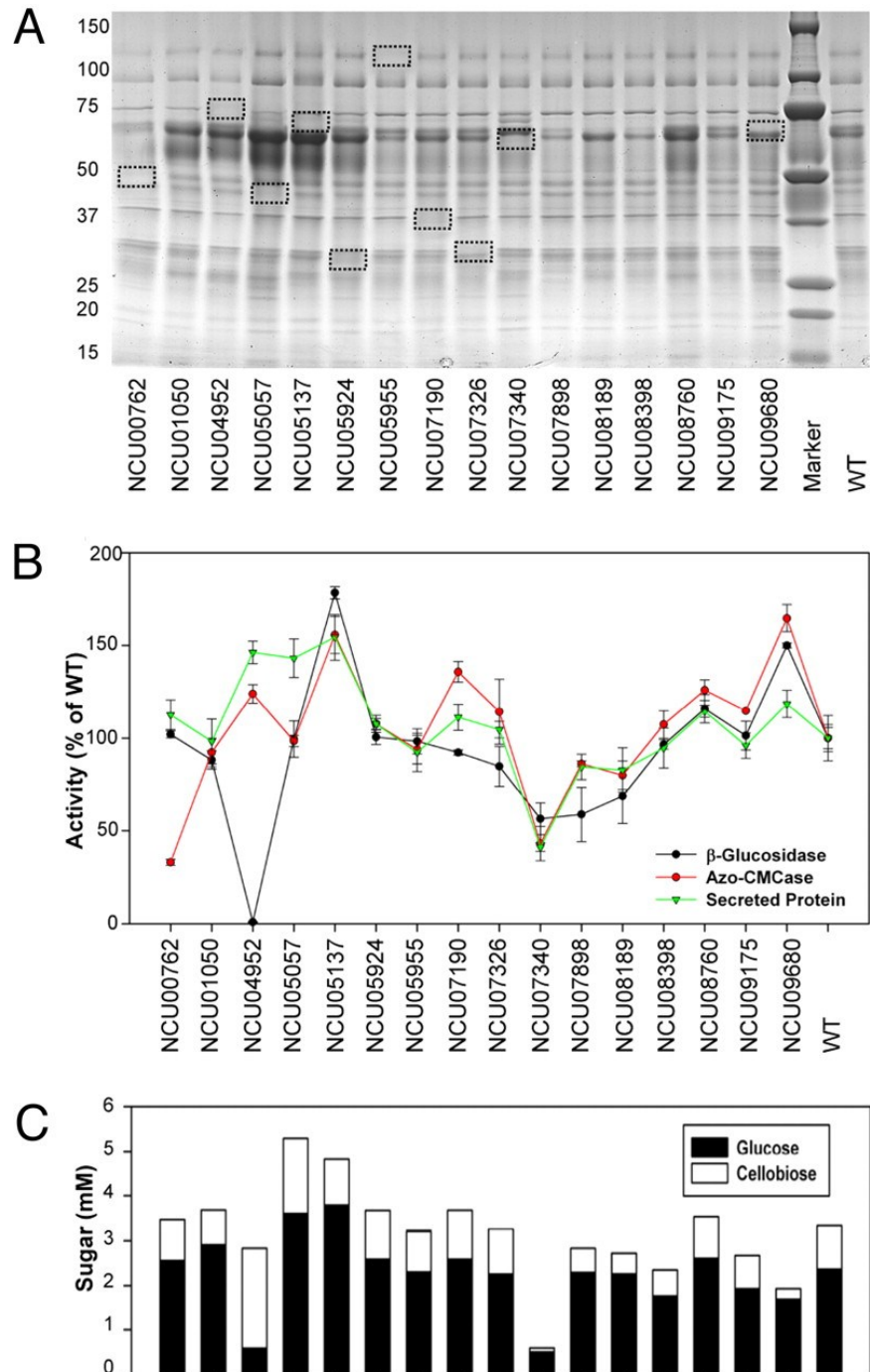
Figure 4. Protein profile and enzymatic activity of culture supernatants from strains containing deletions of genes encoding secreted proteins identified by MS. (*A*) SDS/PAGE of proteins present in the culture filtrates of 16 deletion strains as compared to WT when grown on Avicel for 7 days. Missing protein bands that correspond to the deleted genes are marked with boxes. (*B*) Total secreted protein, azo-CMCase, and β-glucosidase activity assays performed on 16 deletion strains and the WT parental strain (FGSC 2489) using the same sample from *A*. (*C*) Cellulase activity of the culture filtrates from the 16 deletion strains using the same samples as in *A*. Bars, standard deviation. Glucose (black) and cellobiose (white) were measured after 8 h of incubation at 40 °C. Bars, standard deviation.

Figure 5. Time Course Analysis of Secreted Protein in WT and ΔNCU04952, ΔNCU05137.
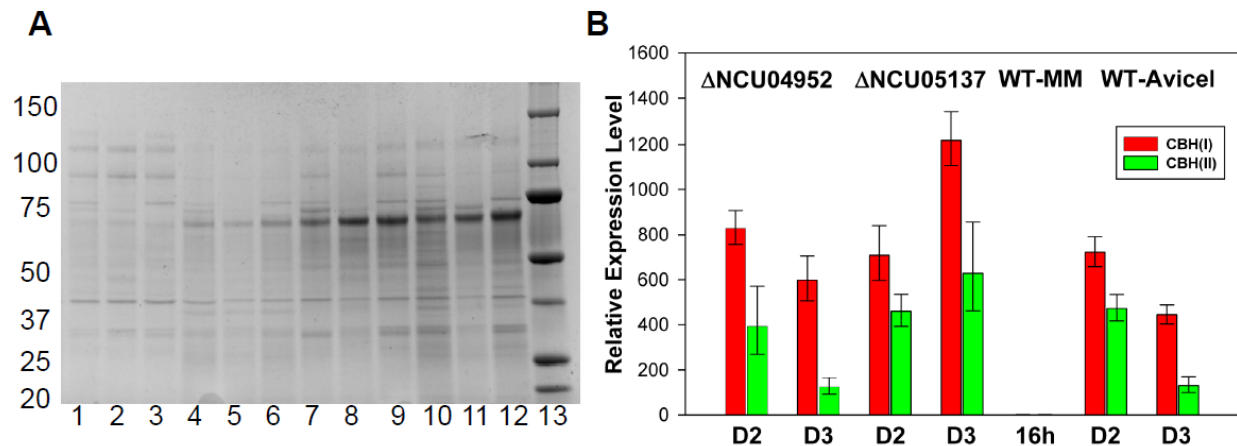(A.) SDS-PAGE of total secreted proteins in WT, ΔNCU04952, and ΔNCU05137. The cultures were grown on Avicel and harvested at 30 hours, 48 hours, and 72 hours. Lanes 1-3, 20x concentrated culture filtrates after 30 hours of growth on Avicel from WT, ΔNCU04952, and ΔNCU05137. Lanes 4-6, unconcentrated culture filtrates after two days of growth from WT, ΔNCU04952, and ΔNCU05137. Lane 7-9, unconcentrated culture filtrates after three days of growth from WT and ΔNCU04952 and ΔNCU05137. (B) RT-PCR of CBHI and CBHII in the WT, ΔNCU04952, and ΔNCU05137 during growth on Avicel (see materials and methods).

that results in increased expression of *cbh-1.* To differentiate these possibilities, we compared the profile of extracellular proteins produced by ΔNCU05137 and *Δgh3-4* (NCU04952) with gene expression levels of *cbh-1* (NCU07340) and *gh6-2* (CBH(II); NCU09680) as assayed by quantitative RT-PCR (Fig. 5). The ΔNCU05137 and *Δgh3-4* strains showed a higher level of CBH(I) protein as early as two days in an Avicel-grown culture (Fig. 5a) and higher expression levels of both *cbh-1* and *gh6-2* at three days of growth, while expression of both of these genes decreased significantly in the WT strain (Fig. 5b). Sustained expression of *cbh-1* and *gh6-2* genes in the ΔNCU05137 and *Δgh3-4* mutants could be responsible for the observed increase in CBH(I) and CBH(II) protein levels.

2.4 Discussion

Degradation of plant biomass requires production of many different enzymes, which are regulated by the type and complexity of the available plant material (Fig. 6) (77). Here, we report on the systematic analyses of plant cell wall degradation by a cellulolytic fungus, which includes transcriptome, secretome and mutant analyses. Our profiling data shows that *N. crassa* coordinately expresses a host of extracellular and intracellular proteins when challenged by growth on *Miscanthus* or Avicel (Fig. 6). The most highly expressed genes encode proteins predicted to be involved in the metabolism of plant cell wall polysaccharides, many of which were identified by MS analyses. The genomes of filamentous fungi have a large number of predicted glycosyl

hydrolases (~200) with varying numbers of predicted cellulases, from 10 in *H. jecorina* (*68*) to 60 in *Podospora anserina* (*78*). A comparison between our results and a cDNA expression/Northern analysis of 8 endoglucanases and 7 β-glucosidases in *H. jecorina* (*79*) showed complete overlap with our profiling data, with the exception of one ortholog of a β-glucosidase (cel-3e=NCU05577). However, a recent transcriptome/secretome study on the white rot basidiomycete fungus, *Phanerochaete chrysosporium* (*80*) showed little overlap. These data suggest that different fungi may utilize different gene sets for plant cell wall degradation. One aspect all of these studies have in common is the high number of uncharacterized genes/proteins associated with cellulose degradation. Using the functional genomic tools available with *N. crassa*, we can address both function and redundancy of plant cell wall degrading enzyme systems to create optimal enzyme mixtures for industrial production of liquid fuels from lignocellulose biomass.

N. crassa cellobiohydrolase(I) (CBHI) is the most highly produced extracellular protein during growth on Avicel or *Miscanthus*, and deletion of this gene causes the most severe growth deficiencies on cellulosic substrates. By contrast, in *H. jecorina*, deletion of *cbhII* caused the most severe phenotype (*81-83*). In *N. crassa*, deletion of cellobiohydrolase(II) also causes growth deficiencies on cellulosic substrates, but to a much lesser extent than CBH(I), suggesting that exoglucanase activity in *N. crassa* is predominantly from CBH(I) and that cellulases and other CBHs do not compensate for the loss of CBH(I). The three most highly produced endoglucanases are proteins encoded by NCU05057, NCU00762, and NCU07190. These proteins have homology to endoglucanases EG1, EG2, and EG6, respectively. Deletion of these genes did not affect growth on Avicel, although differences in secreted protein levels and endoglucanase activity were observed. Unexpectedly, in ΔNCU05057, extracellular protein levels were much higher, especially CBH(I), suggesting increased production of other cellulases or that the products of NCU05057 catalysis may repress cellulase production to maintain the WT growth phenotype on crystalline cellulose. We conclude that no one endoglucanase in *N. crassa* is required for growth on crystalline cellulose and that different endoglucanases have overlapping substrate specificities.

The glycoside hydrolase family 61 enzymes are greatly expanded in *N. crassa* compared to *H. jecorina* (*68*). These enzymes have poorly defined biological function, but their general conservation and abundance in cellulolytic fungi suggests an important role in plant cell wall metabolism. Genes for 10 of the 14 GH61 enzymes were identified in the *N. crassa* transcriptome, suggesting that these enzymes are utilized during growth on cellulosic biomass. The four GH61 deletion strains tested showed only small differences compared to wild type in the secreted protein levels, endoglucanase, and total cellulase activities. However, analyses of additional GH61 mutants and the capacity to create strains containing multiple mutations in *N. crassa* via sexual crosses will address redundancy and expedite functional analysis of this family.
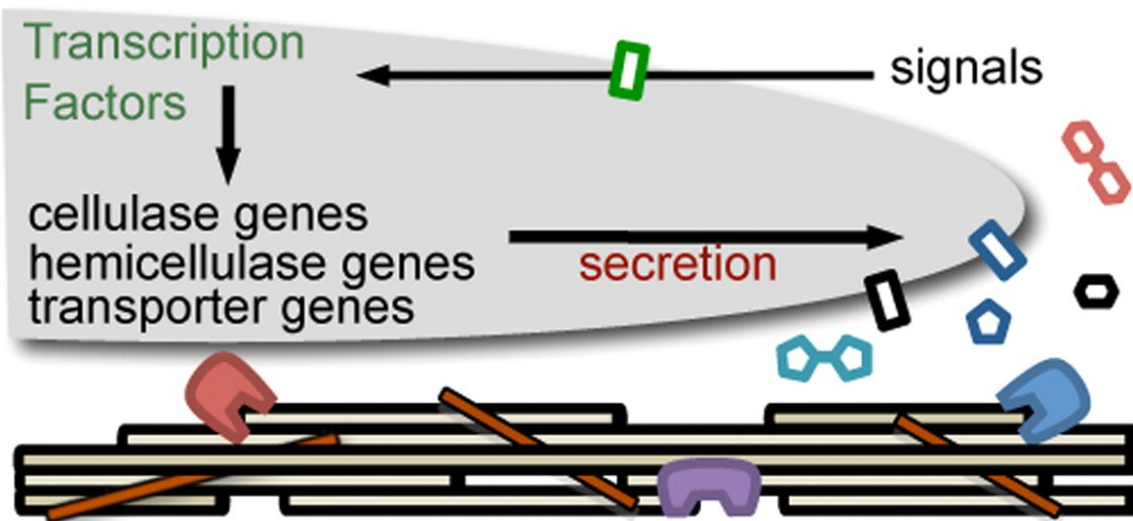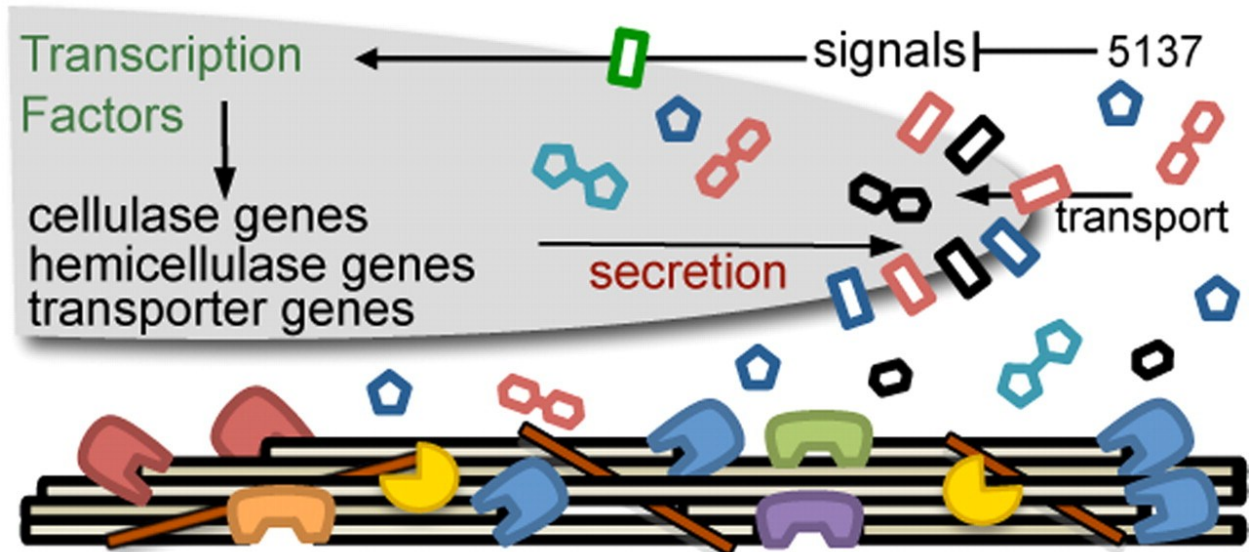
Figure 6. Model of plant cell wall deconstruction in *N. crassa*. Induction: Extracellular enzymes expressed at low levels generate metabolites that signal *N. crassa* to dramatically increase the expression level of genes encoding plant cell wall degrading enzymes. Utilization: Extracellular enzymes and transporters specific for translocation of cell wall degradation products enable *N. crassa* to use plant cell material for growth. Some extracellular proteins may generate metabolites that modulate gene expression of cellulases and hemicellulases during the utilization phase. Double red hexagon (cellobiose), double teal pentagon (xylobiose), black hexagon (glucose), and blue pentagon (xylose). Blue, CBH(I); red, CBH(II); purple, EG2; green, EG1; orange, EG6; and yellow, xylanase. Additional cellulolytic enzymes not shown. Thickness of arrows indicates relative strength of response.

In addition to predicted cellulase genes, genes encoding hemicellulases, carbohydrate esterases, β-glucosidases, β-xylosidases and other proteins predicted to have activity on carbohydrates were identified in the *N. crassa* transcriptome from both *Miscanthus* and Avicel. Expression of hemicellulase genes even when *N. crassa* was grown on bacterial cellulose indicates that cellulose is the primary inducer of genes encoding plant cell wall degrading enzymes. However, genes encoding some hemicellulases and carbohydrate esterases were only expressed during growth on *Miscanthus*. Similarly, in other cellulolytic fungi such as *H. jecorina* and *Aspergillus niger*, genes encoding some cellulases and hemicellulases are coordinately regulated, while others are differentially regulated (*84*). As expected, deletions of non-cellulase genes had little effect on growth on Avicel or cellulase activity, with the exception of NCU05137 and *gh3-4*. The ΔNCU05137 strain secreted more protein, had higher cellulase activity and showed higher expresssion of *cbh-1* (CBH(I)) and *gh6-2* (CBH(II)) than WT. NCU05137 encodes a secreted protein that lacks homology to proteins of known function, but is highly conserved in other cellulolytic fungi. NCU05137 also has distant homologs of unknown function in a number of bacterial species. We hypothesize that the NCU05137 protein may affect signaling processes associated with the regulation of cellulase gene expression *N. crassa* (Fig. 6). Similarly, mutations in *gh3-4* (NCU04952) also increased CBH(I) activity. Deletion of *Δgh3-4* completely removed PNPGase activity and resulted in cellobiose accumulation in *in vitro* cellulase assays. These data suggest that NCU04952 encodes the primary extracellular β-glucosidase in *N. crassa*.

Extracellular degradation of plant cell walls results in the formation of soluble carbohydrates that are subsequently transported into the cell (Fig. 6). In this study, we identified 10 genes encoding permeases/transporters whose expression increased significantly when *N. crassa* was grown on *Miscanthus* or Avicel. The major degradation products by cellulases and hemicellulases *in vitro* are cellobiose, glucose, xylobiose, and xylose. Some of these transporters may be functionally redundant or capable of transporting oligosaccharides. Construction of downstream processing strains capable of transporting oligosaccharides by heterologous expression of *N. crassa* transporters may improve industrial fermentation of biomass hydrolysis products. None of these transporters or what they may transport has been characterized at the molecular or functional level in any filamentous fungus.

Many genes that increased in expression level during growth on *Miscanthus* and Avicel encode proteins of unknown function and are conserved in other cellulolytic fungi. By assessing the phenotype of only 16 strains, we identified a mutation in a protein of unknown function that significantly affected cellulase activity. The well understood genetics and availability of functional genomic resources make *N. crassa* an ideal model organism to assess biological function of these proteins, examine regulatory aspects of

cellulase and hemicellulase production, and dissect redundancies and synergies between extracellular enzymes involved in the degradation of plant cell walls.

2.5 Materials and Methods

Strains. The sequenced *N. crassa* strain (FGSC 2489) was used for transcriptional profiling. Gene deletion strains were generated from the *N. crassa* functional genomics project (*69*); all mutations were verified by PCR.. *H. jecorina* QM9414 was a gift from Dr. Monika Schmoll (Vienna University of Technology). Strains were grown on Vogel's salts (*85*) with 2% (w/v) carbon source (*Miscanthus*, sucrose or Avicel (Sigma)). Growth of *N. crassa* on Avicel and *Miscanthus* was monitored by macroscopic and microscopic visual assessment. Deletion strains were obtained from the Fungal Genetics Stock Center (FGSC), including FGSC 16747 (ΔNCU00762), FGSC 16543 (ΔNCU01050), FGSC 13732 (ΔNCU04952), FGSC 13343 (ΔNCU05057), FGSC 11682 (ΔNCU05137), FGSC 15626 (ΔNCU05924), FGSC 13535 (ΔNCU05955), FGSC 19315 (ΔNCU07190), FGSC 19534 (ΔNCU07326), FGSC 15630 (ΔNCU07340), FGSC 19600 (ΔNCU07898), FGSC 19861 (ΔNCU08189), FGSC 20310 (ΔNCU08398), FGSC 15664 (ΔNCU08760), FGSC 11750 (ΔNCU09175), FGSC 15633 (ΔNCU09680). Stem and leaf tissues from *Miscanthus gigantaeus* were harvested from field-grown plants at the University of Illinois, Urbana-Champaign (2007) at the end of the growing season, air dried, and milled to a particle size of 100 μm.

Enzyme activity measurements. Total extracellular protein content was determined by a Bio-Rad *DC* Protein Assay kit (Bio-Rad). Endoglucanase activity in culture supernatants was measured with an azo-CMC kit (Megazyme SCMCL). β-glucosidase activity was measured by mixing 10x diluted culture supernatant with 500 uM 4-nitrophenyl β-D-glucopyranoside in 50 mM sodium acetate buffer, pH 5.0, for 10 minutes at 40C. The reaction was quenched with 5% w/v sodium carbonate and the absorbance at 400 nm was measured. Avicelase activity was measured by mixing 2X-diluted culture supernatant with 50 mM sodium acetate pH 5.0 and 5 mg/mL Avicel at 40C. Supernatants were analyzed for glucose content using a coupled enzyme assay with glucose oxidase/peroxidase. 50 uL of the avicelase reaction was transferred to 150 uL of glucose detection reagent containing 100 mM sodium acetate pH 5.0, 10 U/mL horseradish peroxidase, 10 U/mL glucose oxidase, and 1 mM o-dianisidine. After 30 minutes absorption was measured at 540 nm. Cellobiose concentrations in the avicelase reaction mixture were determined using a coupled enzyme assay with cellobiose dehydrogenase (CDH) from *Sporotrichum thermophile*. CDH was isolated from *S. thermophile* similar to previous reports (*86*). 50 uL of the Avicelase reaction was transferred to 250 uL of cellobiose detection reagent containing 125 mM sodium

27

acetate pH 5.0, 250 uM dichlorophenol indophenol, and 0.03 mg/mL CDH. After 10 minutes absorption was measured at 530 nm.

Transcriptome analysis. *N. crassa* strain FGSC 2489 was inoculated onto slants of Vogel's minimal medium (MM) with 1.5% agar and 2% sucrose and grown for 10 days at 25°C. A conidial suspension was inoculated into 50 ml of Vogel's liquid medium (*85*) in 250 ml Pyrex flasks supplemented with either 2% Avicel, *Miscanthus* powder or sucrose as sole carbon sources at a final concentration of 10^6 conidia per ml. Cultures were grown under controlled temperature (25°C) with constant shaking and light (300 lux). Mycelia was harvested by vacuum filtration using Whatman paper #1, frozen immediately in liquid nitrogen and stored at 80°C. Two independent biological duplicates (flasks) were evaluated for each time point.

RNA isolation and cDNA labeling. Total RNA from frozen samples was isolated using Zirconia/silica beads (0.2 g, 0.5-mm diameter; Biospec) and a MiniBeadBeater (Biospec) with 1 ml of TRIzol reagent (Invitrogen Life Technologies) according to manufacturer's instructions. Total RNA (100 μg) was further purified using RNAeasy kit (Qiagen). RNA integrity was checked by Nanodrop and agarose gel electrophoresis. cDNA synthesis and labeling followed the protocol of ChipShot indirect labeling kit from Promega (Catalog #Z4000). The dyes Cy3 and Cy5 (Amersham; catalog no. RPN5661) were incorporated into cDNA by adding Cy3 or Cy5 monofunctional *N*-hydroxysuccinimide ester dye to the cDNA solution for 1 h at 22°C. The cDNA was subsequently cleaned by using a ChipShot membrane column and dried under vacuum.

Hybridization and image acquisition. A set of 10,910 70-mers representing predicted genes in *N. crassa* were printed onto gamma amino propyl silane slides (Corning) at the University of California—San Francisco Center for Advanced Technology (http://cat.UCSF.edu/). *N. crassa* microarray slides are available from the Fungal Genetics Stock Center (http://www.fgsc.net/). Information on the oligonucleotide gene set is available at the *Neurospora* functional genomics database (http://www.yale.edu/townsend/Links/ffdatabase/introduction.htm). Pre-hybridization and hybridization of slides used the Pronto kit (Corning; catalog no. 40076) according to manufacturer's instructions. Microarray slides were added to a pre-soak solution that was pre-warmed to 42°C and incubated for 20 min. Slides were subsequently washed (Pronto kit) and transferred to pre-warmed pre-hybridization solution (42°C) for 15 min, and rewashed. Labeled cDNA was resuspended in 30 μl of hybridization solution (Pronto kit). The suspension was heated at 95°C for 5 min and then pipetted into the space between the microarray slide and a LifterSlip cover glass (Erie Scientific, Portsmouth, NH). Hybridization was performed at 42°C for 16 hr. Unbound DNA was washed off according to manufacturer's instructions (Pronto kit).

Hybridized slides were scanned using Genepix 4000B. Four independent hybridizations were performed per sample. Genepix6 software was used for alignment of hybridization spots and initial analysis (.GPR files). Hybridized spots with a mean fluorescence intensity for at least one of the Cy3 and Cy5 dyes that was greater than the mean background intensity plus three standard deviations were scored for further analysis. For quality control, slides were used where the ratio of Cy5:Cy3 signal was approximately one. Global normalization was performed to each GPR file using in-house excel-based Macro (*72*). The Macro is available at http://glasslab.weebly.com/links.html and Yale Microarray Experimental Design site.

Experimental design and data analysis. A closed-circuit design for microarray comparisons was used because they are statistically robust and improve resolution in identifying differentially regulated genes compared to designs for microarrays that use a universal reference (*87-89*). The gene relative expression data was estimated by BAGEL software (*90*). BAGEL explores the ratio for all samples using a Markov Chain Monte Carlo (MCMC) approach in a Bayesian framework, and infers relative expression levels and statistical significance from the parameter values it samples. BAGEL also provides gene-by-gene 95% credible intervals, values in which 95% of the samples from the chain are bounded. The 769 *N. crassa* genes indentified in this study that were identified as showing a significant difference in expression between *Miscanthus* and minimal sucrose medium passed two filters. The first filter was that the 95% credible interval of a gene expressed under *Miscanthus* versus sucrose conditions showed no overlap and the closest boundary value between the two samples showed at least 1.5 fold difference in expression level. The second filter used was the average expression of a gene in *Miscanthus* versus sucrose grown cultures was required to have at least a 2-fold change at one of four time points. These data were obtained by in-house PERL script "strict-errobar.pl" which is available in http://glasslab.weebly.com/links.html. Clustering analysis performed by HCE3 (Hierarchical Clustering Explorer 3.0), using default parameter in which a Euclidean Distance model (*91*) is used to join genes and clusters successively. Distances between clusters represent the average distances between genes in the clusters.

The MIPs functional catalog (FunCat) was used to associate functional annotation (*73*) with *Neurospora* genes. The overrepresentation of *Neurospora* gene groups in functional categories relative to the whole genome was determined using hypergeometric distribution for *P* value calculation.

Signal peptides were predicted using the N-terminal 70 amino acid region of each predicted protein with the signalP3 program (http://www.cbs.dtu.dk/services/SignalP-3.0/).

Protein gel electrophoresis. Except where otherwise noted, unconcentrated culture supernatants were treated with 5x SDS loading dye and boiled for 5 minutes before loading onto Criterion 4-15% Tris-HCl polyacrylamide gels. Coomassie dye was used for staining.

Preparation of tryptic peptides for secretome analysis. Culture supernatants were concentrated with 10 kDa MWCO PES spin concentrators. Cellulose binding proteins were isolated from the culture supernatant by addition of phosphoric acid swollen cellulose (PASC). 5 mL of a suspension of 10 mg/mL PASC was added to 10 mL of culture supernatant. After incubation at 4C for 5 minutes, the mixture was centrifuged and the pelleted PASC was then washed with 20 pellet volumes of 100 mM sodium acetate pH 5.0. The supernatant after treatment with PASC was saved as the unbound fraction and concentrated. 36 mg of urea, 5 uL of 1M Tris PH 8.5, and 5 uL of 100 mM DTT were then added to 100 uL of concentrated culture supernatant or protein-bound PASC and the mixture was heated at 60C for 1 hour. After heating 700 uL of 25 mM ammonium bicarbonate and 140 uL of methanol were added to the solution followed by treatement with 50 uL of 100 ug/mL trypsin in 50 mM sodium acetate pH 5.0. For the PASC bound proteins, the PASC was removed by centrifugation after heating, and the supernatant was then treated with trypsin. The trypsin was left to react overnight at 37C. After digestion the volume was reduced by speedvac and washed with MilliQ water three times. Residual salts in the sample were removed by using OMIX microextraction pipette tips according to the manufacturer's instructions.

Liquid chromatography of tryptic peptides. Trypsin-digested proteins were analyzed using a tandem mass spectrometer that was connected in-line with ultraperformance liquid chromatography (UPLC). Peptides were separated using a nanoAcquity UPLC (Waters, Milford, MA) equipped with C_{18} trapping (180 μm × 20 mm) and analytical (100 μm × 100 mm) columns and a 10 μL sample loop. Solvent A was 0.1% formic acid/99.9% water and solvent B was 0.1% formic acid/99.9% acetonitrile (v/v). Sample solutions contained in 0.3 mL polypropylene snap-top vials sealed with septa caps (Wheaton Science, Millville, NJ) were loaded into the nanoAcquity autosampler prior to analysis. Following sample injection (2 μL, partial loop), trapping was performed for 5 min with 100% A at a flow rate of 3 μL/min. The injection needle was washed with 750 μL each of solvents A and B after injection to avoid cross-contamination between samples. The elution program consisted of a linear gradient from 25% to 30% B over 55 min, a linear gradient to 40% B over 20 min, a linear gradient to 95% B over 0.33 min, isocratic conditions at 95% B for 11.67 min, a linear gradient to 1% B over 0.33 min, and isocratic conditions at 1% B for 11.67 min, at a flow rate of 500 nL/min. The analytical column and sample compartment were maintained at 35 ºC and 8 ºC, respectively.

Mass spectrometry. The column was connected to a NanoEase nanoelectrospray ionization (nanoESI) emitter mounted in the nanoflow ion source of a quadrupole time-of-flight mass spectrometer (Q-Tof Premier, Waters). The nanoESI source parameters were as follows: nanoESI capillary voltage 2.3 kV, nebulizing gas (nitrogen) pressure 0.15 mbar, sample cone voltage 30 V, extraction cone voltage 5 V, ion guide voltage 3 V, and source block temperature 80 °C. No cone gas was used. The collision cell contained argon gas at a pressure of 8×10^{-3} mbar. The Tof analyzer was operated in "V" mode. Under these conditions, a mass resolving power[1] of 1.0×10^4 (measured at m/z = 771) was routinely achieved, which is sufficient to resolve the isotopic distributions of the singly and multiply charged peptide ions measured in this study. Thus, an ion's mass and charge could be determined independently, i.e., the ion charge was determined from the reciprocal of the spacing between adjacent isotope peaks in the m/z spectrum. External mass calibration was performed immediately prior to analysis, using solutions of sodium formate. Survey scans were acquired in the positive ion mode over the range m/z = 450-1800 using a 0.95 s scan integration and a 0.05 s interscan delay. In the data-dependent mode, up to five precursor ions exceeding an intensity threshold of 35 counts/second (cps) were selected from each survey scan for tandem mass spectrometry (MS/MS) analysis. Real-time deisotoping and charge state recognition were used to select 2+, 3+, 4+, 5+, and 6+ charge state precursor ions for MS/MS. Collision energies for collisionally activated dissociation (CAD) were automatically selected based on the mass and charge state of a given precursor ion. MS/MS spectra were acquired over the range m/z = 50-2500 using a 0.95 s scan integration and a 0.05 s interscan delay. Ions were fragmented to achieve a minimum total ion current (TIC) of 30,000 cps in the cumulative MS/MS spectrum for a maximum of 3 s. To avoid the occurrence of redundant MS/MS measurements, real time exclusion was used to preclude re-selection of previously analyzed precursor ions over an exclusion width of ± 0.25 m/z unit for a period of 180 s.

Mass spectrometry data analysis. The data resulting from LC-MS/MS analysis of trypsin-digested proteins were processed using ProteinLynx Global Server software (version 2.3, Waters), which performed background subtraction (threshold 35% and fifth order polynomial), smoothing (Savitzky-Golay[2] 10 times, over three channels), and centroiding (top 80% of each peak and minimum peak width at half height four channels) of the mass spectra and MS/MS spectra. The processed data were searched against the *N. crassa* database (Broad Institute)[3] using the following criteria: tryptic fragments with up to five missed cleavages, precursor ion mass tolerance 50 ppm, fragment ion mass tolerance 0.1 Da, and the following variable post-translational modifications: carbamylation of N-terminus and Lys side chains, Met oxidation, and Ser/Thr dehydration. The identification of at least three consecutive fragment ions from

the same series, i.e., b or y-type fragment ions, was required for assignment of a peptide to an MS/MS spectrum. The MS/MS spectra were manually inspected to verify the presence of the fragment ions that uniquely identify the peptides.

Quantitative RT-PCR. The RT-PCR was performed in a ABI7300 with reagents from Qiagen (SYBR-green RT-PCR kit (cat#204243)) or Invitrogen (EXPRESS One-step SYBR GreenER Kits (cat#11790-200)). The primers for CBHI (NCU07340) were: forward 5'-ATCTGGGAAGCGAACAAAG-3' and reverse 5'-TAGCGGTCGTCGGAATAG-3'. The primers for CBHII (NCU09680) were: forward 5'-CCCATCACCACTACTACC-3' and reverse 5'-CCAGCCCTGAACACCAAG-3'. The primers for NCU02855 were: forward 5'-CTT CAA CAA CGG CTT CTA-3' and reverse 5'-TGC CAG TTC ACA GAG TAA-3'. The primers for NCU05924 were: forward 5'-AGG GGC CAA TAT AAC TTT-3'and reverse 5'-GAG TGC CAG AGG AAG GTA TG-3'. The primers for NCU08189 were: forward 5'-AGC ATG TCG CTG CCA TTA-3' and reverse 5'-GCC ATT CTT ATA CGC CTC CT-3'. Actin was used as a control for normalization. The primers for actin were: forward 5'-TGA TCT TAC CGA CTA CCT -3' and reverse 5'-CAG AGC TTC TCC TTG ATG -3'). Quantitative R-PCR was performed according to (*92*).

Complementation and segregation analysis. The genotype of all deletion strains was confirmed by using a gene-specific primer and a common primer for the hygromycin (hph) cassette. The forward primer for hph was 5'-TGC AAT AGG TCA GGC TCT-3. Reverse primers were: NCU00762: 5'-GGA AGA ACA CTG TCT TGA GG-3'; NCU01050: 5'-TTG CAC TCC ATC TCT ACA CC-3'; NCU04952: 5'-AAC ACA CAC ACA CAC ACT GG-3'; NCU05057: 5'-GGA AGG AGT TCC TCA AGT GG-3'; NCU05137: 5'-CTT ATC TCC AGT CAC GAT CC-3'; NCU05924: 5'-CTG ATC TAT CGG TGT TCA GC-3'; NCU05955: 5'-CCT TCC AAT CTG TCT CTT GC-3'; NCU07190: 5'-CTG GTA TGT AGT CCT GAA GC-3'; NCU07326: 5'-GTG GTG GTA CAA GTA GAT GG-3'; NCU07340: 5'-CGT CTG ATT TGT CCA GTA CC-3'; NCU07898: 5'- GCC CAC AGT ACC TAC CCT GT-3'; NCU08189: 5'-TCT ATT CTC GAT TCC CGT CC-3'; NCU08398: 5'-TAC CAA GGT ACA CCC TAT CG-3'; NCU08760: 5'-AAA CTG TCT CTG TGG TCT GC-3'; NCU09175: 5'-GGT AAG AGA GAG ATC GAT GG-3'; NCU09680: 5'-AAC TTA CCA CTC ACT CCT CC-3'.
A *his-3* plasmid (*93*) was used target NCU05137-GFP and NCU00762-GFP to the *his-3* locus of FGSC 6103 (*his-3*) via transformation. Transformants that were His+ and GFP positive were crossed to ΔNCU05137 or ΔNCU00762 and progeny were screened for hph resistance and GFP fluorescence. Positive transformants were grown in Vogel's (*85*) plus 2% Avicel for 4 days; an analysis of total secreted protein, cellulase activity and protein profile by SDS-PAGE of secreted proteins of complemented strains was identical to wild-type. The phenotype of the remaining mutants was assessed by

segregation analysis. Crosses analyzed were *ΔNCU04952 a* (FGSC 13731) X *mus-52*::*bar his-3 A*; *ΔNCU09680 a* (FGSC 13731) X *mus-52*::*bar his-3 A*; *ΔNCU07340 A* (FGSC 15633) X *mus-52*::*bar his-3 a* (FGSC 9539). Twenty-four germinated ascospore progeny inoculated onto VMM slants with histidine for 5 days and then conidia from each progeny were screened for hygromycin resistance, histidine auxotrophy and cellulase phenotype via analysis of secreted proteins and cellulase activity using assay described above. In all cases, hygromycin resistance co-segregated with cellulase phenotype.

*Footnote to chapter 2

This chapter, presented here with few modifications from its original format, represents a peer-reviewed paper published in *The Proceedings of the National Academies of Sciences of the United States of America* (2009), 106, 22157-22162. The authors are Chaoguang Tian, William T. Beeson, Anthony T. Iavarone, Jianping Sun, Michael A. Marletta, Jamie H.D. Cate, and N. Louise Glass.

Chapter 3: Extracellular aldonolactonases from *Myceliophthora thermophila*

3.1 Abstract

Fungi secrete many different enzymes to deconstruct lignocellulosic biomass, including several families of hydrolases, oxidative enzymes, and many uncharacterized proteins. Here we describe the isolation, characterization, and primary sequence analysis of an extracellular aldonolactonase from the thermophilic fungus *Myceliophthora thermophila* (syn. *Sporotrichum thermophile*). The lactonase is a 48 kDa glycoprotein with a broad pH optimum. The enzyme catalyzes the hydrolysis of glucono-δ-lactone and cellobiono-δ-lactone with an apparent second order rate constant, k_{cat}/K_M, of ~ 1×10^6 M^{-1} s^{-1} at pH 5.0 and 25 °C, but is unable to hydrolyze xylono-γ-lactone or arabino-γ-lactone. Sequence analyses of the lactonase show that it has distant homology to cis-carboxy-muconate lactonizing enzymes as well as 6-phosphogluconolactonases present in some bacteria. The *M. thermophila* genome contains two predicted extracellular lactonases and expression of both genes is induced by the presence of pure cellulose. Homologues of the *M. thermophila* lactonase, which are also predicted to be extracellular, are present in nearly all known cellulolytic ascomycetes.

3.2 Introduction

The thermophilic fungus, *Myceliophthora thermophila* (syn. *Sporotrichum thermophile),* very rapidly degrades cellulose and metabolizes powdered cellulose at nearly the same rate as glucose (*94*). The thermostability of the hydrolytic enzymes from this organism provides some practical advantages over those from the mesophilic fungus, *Hypocrea jecorina* (syn. *Trichoderma reesei*), which has traditionally been used for production of biomass degrading enzymes. During growth on cellulosic substrates, *M. thermophila* secretes cellulases, hemicellulases (*95*), oxidative enzymes (*96*), and many proteins of unknown function.
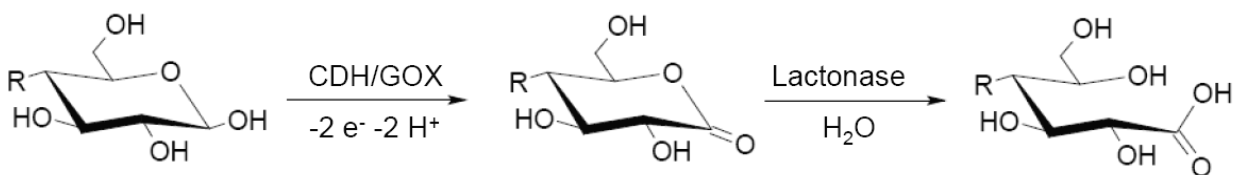

Figure 1. Extracellular sugar oxidation and lactone hydrolysis. Reducing sugars are initially converted to sugar lactones by extracellular oxidizing enzymes, including glucose oxidase(GOX) and cellobiose dehydrogenase(CDH). Conversion to the corresponding sugar acid is catalyzed by extracellular lactonases.

Cellobiose dehydrogenase (CDH) is an extracellular hemo-flavoprotein that is produced in large amounts by *M. thermophila* during growth on cellulose (*96*). CDH is produced by many cellulolytic fungi, although the biological function of the enzyme is still unclear (*97*). CDH oxidizes the reducing end of cellobiose and longer cellodextrins to the corresponding aldonolactones (Figure 1). Sugar lactones have been shown to inhibit many different types of glycosyl hydrolases (*98*). For example, glucono-δ-lactone is a potent inhibitor of β-glucosidases (*99*). The product of CDH catalysis, cellobiono-δ-lactone, has been shown to be a strong inducer of cellulase in the filamentous fungus *Hypocrea jecorina* (syn. *Trichoderma reesei*) (*100*). However, the mechanism of cellulase induction and the metabolism of cellobiono-δ-lactone by fungi have not been studied in detail.

Sugar lactones are unstable in aqueous solution and undergo hydrolysis to form the corresponding aldonic acids. The extent and rate of uncatalyzed hydrolysis are dependent on the specific lactone, pH and temperature. The equilibrium constant between glucono-δ-lactone and gluconic acid is 7.7, favoring gluconic acid, and the half-life of glucono-δ-lactone in water at room temperature and pH 5.0 is approximately one hour (*101*). Despite this lack of stability, enzymes have evolved to catalyze the hydrolysis of sugar lactones.

Enzymatic hydrolysis of sugar lactones has been studied most thoroughly in the context of the pentose phosphate pathway (*102-105*). In the pentose phosphate pathway, glucose-6-phosphate is converted to 6-phosphogluconolactone by glucose-6-phosphate dehydrogenase. The lactone is then hydrolyzed by 6-phosphogluconolactonase (PGL) to generate 6-phosphogluconate and finally converted to ribulose-5-phosphate. In glycolysis-deficient strains of *Escherichia coli*, *pgl* deletion leads to severe inhibition of growth on glucose (*106*), clearly demonstrating that spontaneous hydrolysis of 6-phosphogluconolactone is insufficient *in vivo*.

There have been reports of lactonase activity secreted into the culture filtrates of diverse fungi (*107, 108*). A lactonase from *Aspergillus niger* was purified by Bruchmann et al. (*107*) and was shown to be an important part of the fungal cellulolytic system. They speculated that the principal function of the lactonase was to relieve inhibition of cellulases and β-glucosidases by glucono-δ-lactone and cellobiono-δ-lactone. However, the gene responsible for extracellular lactonase activity in fungi and details of its kinetic properties remains unknown.

Here we report the purification, identification, and characterization, of an enzyme responsible for extracellular lactonase activity in the thermophilic ascomycete, *Myceliophthora thermophila*. Orthologues of the lactonase in *M. thermophila* are widespread among ascomycetes, but may be absent in basidiomycetes.

3.3 Results

3.3.1 Purification and properties of *M. thermophila* lactonase.

Extracellular lactonase was purified from the culture broth of *M. thermophila* grown on rich media containing Avicel as the carbon source in five steps (Table 1). The lactonase was purified 25-fold with an overall yield of 25%. Two key purification steps were the Q HP ion-exchange chromatography and the hydrophobic interaction chromatography (Figure 2). SDS-PAGE of the purified lactonase suggested a molecular mass of approximately 48 kDa. LC-MS analyses of the intact lactonase showed several species centered around 44 kDa and differing in mass by approximately 162 daltons, which corresponds to the mass of a hexose subunit ($C_6H_{10}O_5$; Figure 3). This result suggested the presence of protein glycosylation, a common post-translational modification in extracellular proteins secreted by fungi. Native protein LC-MS analyses of the lactonase in 10 mM ammonium acetate suggested the protein is a monomer in solution, consistent with gel filtration retention times (data not shown).

Table 1. Purification of *M. thermophila* extracellular lactonase.

Step	Volume (mL)	Protein (mg)	Total Activity$^\alpha$ (U)	Sp. Act. (U/mg)	Purification (fold)	Recovery (%)
Crude Concentrate	175	162.1	101,140	620	1	100
Avicel Binding	120	68.2	60,060	880	1.4	60
Q HP	10	14.6	39,510	2,710	4.3	39
MonoQ	6	3.9	24,930	6,430	10.3	25
Resource PHE	9	1.6	25,160	15,360	24.6	25

$^\alpha$ Activity was measured by adding diluted enzyme mixture with 50 mM glucono-δ-lactone and quenching the reaction after 1 minute. One unit of activity is defined as the hydrolysis of 1 μmol of glucono-δ-lactone per minute at 25 ºC.

All enzyme activity assays were conducted at 25 ºC due to the instability of the sugar lactones being assayed. The enzyme is stable for several days at 48 ºC, as this is the growth temperature for the organism. The apparent second-order rate constant, k_{cat}/K_M, for the lactonase with glucono-δ-lactone as a substrate at pH 5.0 and 25 ºC is ~ 6 x 10^5 M^{-1} s^{-1}. This activity is approximately 20-fold higher than reported for lactonases isolated from other filamentous fungi (*107*). The lactonase reaction velocity shows a broad pH optimum, with only a 3-fold change in activity from pH 3.0 to pH 7.0 (data not shown). Twelve hours of incubation with 5.0 mM EDTA at 4 ºC had no effect on hydrolysis of glucono-δ-lactone by the lactonase.

3.3.2 Substrate specificity of *M. thermophila* extracellular lactonase.

Pentose and hexose sugar lactones were synthesized as described under Materials and Methods. The glucono, cellobiono, and lactono δ-lactones were hydrolyzed completely by the lactonase to the corresponding aldonic acids at pH 5.0. The initial rates of hydrolysis
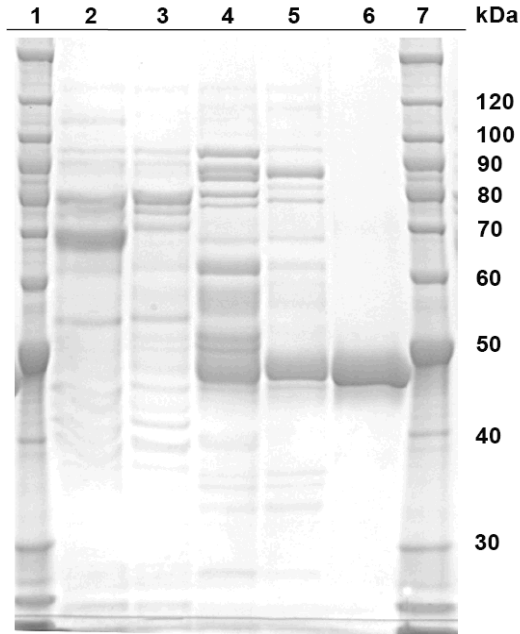
Figure 2. SDS-PAGE of fractions from the purification of *M. thermophila* extracellular lactonase. Fractions were denatured by boiling in SDS buffer containing DTT. Lanes 1 and 7, Benchmark protein ladder; lane 2, crude supernatant concentrate; lane 3, after removal of Avicel binding proteins; lane 4, QHP ion-exchange; lane 5, MonoQ ion-exchange; lane 6, Resource PHE hydrophobic interaction chromatography.

Figure 3. Deconvoluted mass spectrum of purified, intact *M. thermophila* extracellular lactonase. Mass difference between glycoforms is approximately 162 daltons. Mass difference for double peaks is approximately 20 daltons.

39

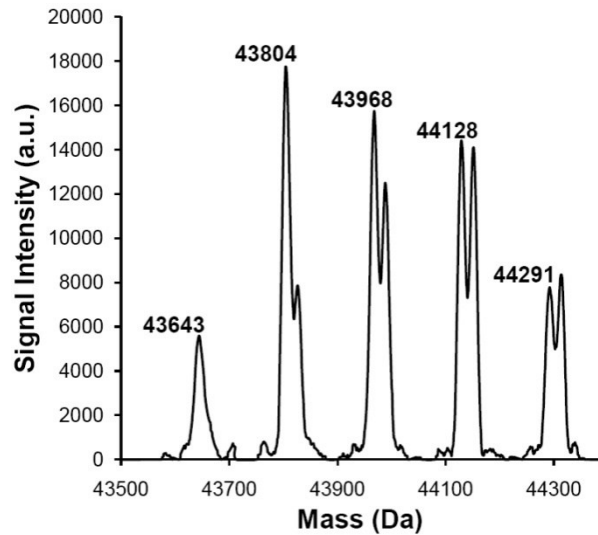were also similar (Table 2). The k_{cat}/K_M for the three substrates varied by less than 3 fold, suggesting that the reducing-end glucose moiety of each substrate is what principally interacts with the lactonase. Xylonolactone and arabinolactone, which are present as γ-lactones in aqueous solution (*98*), were not hydrolyzed by the enzyme. Glucono-δ-lactone can be partially converted to glucono-γ-lactone by heating in water above 100 ºC. When a supersaturated aqueous solution of glucono-δ-lactone was autoclaved for 15 minutes and then provided as a substrate, a portion of the lactone was rapidly hydrolyzed, while the remaining lactone was hydrolyzed at a much slower rate (data not shown). This result is consistent with rapid, enzyme catalyzed hydrolysis of the δ-lactone ring, and slow, uncatalyzed hydrolysis of the γ-lactone ring.

Table 2. Steady-state kinetic constants for various aldonolactones.

Substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (10^5 M^{-1}s^{-1})
D-Glucono-δ-lactone	20.3 ± 4.1	11,700 ± 1000	5.8 ± 1.3
D-Cellobiono-δ-lactone	21.6 ± 1.8	29,100 ±1,100	13.5 ± 1.2
D-Lactono-δ-lactone	5.9 ± 2.4	7,400 ± 600	12.5 ± 5.2

3.3.3 Amino acid sequence of extracellular lactonase.

The purified lactonase was digested with trypsin and the tryptic peptides were analyzed using liquid chromatography and tandem mass spectrometry. A database search against all possible tryptic fragments in the *M. thermophila* proteome (Joint Genome Institute, *Myceliophthora thermophila* v1.0) revealed seven peptides matching Spoth1|109678, and one peptide matching Spoth1|103702 (Figure 4). The signal intensity for the Spoth1|103702 peptide was 20- to 100-fold lower than the signal intensity for peptides from Spoth1|109678, indicating that it is a trace impurity of the protein preparation. Based on the number of peptides detected and relative signal intensity for those peptides, the lactonase enzyme was identified as Spoth1|109678.

The gene encoding Spoth1|109678 is 1,491 base pairs and contains three introns. The open reading frame is 1,206 base pairs encoding a 401 amino acid polypeptide with a calculated mass of 42,509 Da. The protein is predicted to contain a 20 amino acid N-terminal signal peptide (*109*), however Edman degradation sequencing of the mature protein to confirm this prediction was unsuccessful. There are also three predicted N-linked glycosylation sites (http://www.cbs.dtu.dk/services/NetNGlyc/) and no peptides were detected by MS in regions with predicted glycosylation.

MRTSYGVAFALSAGFRLATAAPVCGGGSASDLLWVTTYPAGEGAQGKLLTLKLDGSKLEV

VAESDTCGPYPSWLTQAGDVLYCVDEAWGGDHGTLHSLKINDDHSFTNLSQHETVGGPVS

TVIYGKDGLGLAVADYAGGGIDTFNIADPAAIKLIKSLVYPAPTDGLPDPQNSARPHEAI

LDPTGEFLVFPDLGADQIRVLKVDKETLEYVEKPSYTDFDRGTGPRHGAFFKSGDKTFFY

LVGELSNLLQGFSVAYNDDDSLTFTRIHNSTTHGDDKPLPEDTAAAELWIAPGSNFLTLS

SRFESSLEYTVANGTKVPSDPLITFSIDKETGALTHVQSAPAGGINPRHFSFNSDGTRVA

SALQSDGRVVVFERDPSTGKIGKATAEGDVEGMPNFATFKQ

Figure 4. Amino acid sequence of the extracellular lactonase Spoth1|109678. Underlined; signal peptide, Green; predicted N-linked glycosylation site, Salmon; peptide detected by LC-MS/MS.

3.3.4. Sequence alignments and phylogenetic relationships of the lactonase.

The lactonase amino acid sequence was analyzed for conserved domains using the Pfam database (*110*). Immediately after the predicted signal peptide is a 370 amino acid domain classified as a 3-carboxy-cis,cis-muconate lactonizing enzyme (CMLE). The overall sequence identity of the *M. thermophila* lactonase with *Neurospora crassa* cis-carboxy-muconate lactonizing enzyme (*111, 112*) is only 13%; however, all of the predicted active site residues in *N. crassa* CMLE (*111*) are conserved in the lactonase (Figure 5). The cDNA sequence for lactonase 1 has been deposited at Genbank (HQ437658).

Predicted homologues of the lactonase were retrieved from NCBI and JGI based on amino acid sequences showing significant similarity by BLAST. Homologues of the extracellular lactonase are present in many filamentous ascomycetes, and most are predicted to be secreted. The *M. thermophila* genome contains an additional protein with sequence similarity to the lactonase characterized in this work. Some basidiomycetes have distant homologues, which are not predicted to be secreted. Distant homologues, showing 25 to 30% sequence identity, are present in many species of bacteria. The closest biochemically characterized bacterial proteins with homology to the *M. thermophila* lactonase were shown to have 6-phosphogluconolactonase activity in *Escherichia coli* (*104, 105*) and multifunctional lactonase activity in *Pseudomonas aeruginosa* (*113*). No homologues were found in plants or animals with sequenced genomes.

```
MT_1      1  MRTSYGVAFALSAGFRLATAAPVCGGGSASDLLWVTTYPA----GEGAQGKLLTLKLD--
PA_1      1  MRATHGIFAALSAGLSLSSAAA----CSKQGLLLLVSSYPFESSPGEIVKGGVTTLKLG--
TR        1  MPSRRNLQKSLVSTLL--CAGNL----ASASLVYVSSYS----------QTVTTLNYTHG
AN        1  MGRHF-ALSWLATCL------PL-----ATATNLYATHYD----------GSVYSLSLQRS
MT_2      1  MSPVTQLLAAAALALA-----PG----ALGAHLIASHFS----------GTVYSLSFTSS
EC_pgl    1  MSLKQTVYIASP-------------------------ES---------QQIHVWNLNH-
NC_cmle   1  MPLHHLMIGTWT---------------------------PP---------GAIFTVQFDDE

MT_1     55  --G--SKLE-VVAESDTCGPYP-SWLTQA--GDVLYCVD-EAWG-GDHGTLHSLKINDDH
PA_1     55  --N--KGLEQVGEISSICGTNP-SWQTLVGG-DQYYCIN-ENFD-DGPGAFTSAKVNTDG
TR       45  QNTGIQKLN-PVAVSQGCSDNP-SWLTLDAPDSILYCIN-EGLN-TPNGSLTAFKTSASG
AN       40  -D-DTYSLS-IASSLKTCGSMP-SWLTFDSASRILYCSD-ESGDASTNGSLTTLAADEDG
MT_2     42  SN-STGTLS-VTSETDGCGATP-AWLQLYSDTGKVYCFD-ESWL-G-SGSSAEYEIADDG
EC_pgl   25  ----EGALT-LTQVVDVPGQVQ-P-MVVSPDKRYLYVG----V-RPEFRVLAYRIAPDDG
NC_cmle  26  K--L--TCK-LIKRTEIPQDEPISWMTFDHERKNIYGAAMKK--------WSSFAVKSPT

MT_1    105  SFTNLSQHETVGGPVSTVIYG-KDGL-GLAVA----------------DYAGGGIDTFNIA
PA_1    107  TLAFVGNSSTPGGPVHIALFG-ENGE-RAITS---------------NFASSSLDVFNIE
TR      101  SLQQLGQSSTPNGPVSGVVFG-NNRH-GLAVA---------------HYGGSAFTTWDVS
AN       95  TLTEIATAAAPGGGVNSIFYSGDDGTQYLAIA--------------HYGGSAMSTFRLP
MT_2     96  SLTLTGTLQTTGNSVHGALYGGADGKGFVATV---------------EYTPSTLTTYKMP
EC_pgl   73  ALTFAAESALPGSLTHI-STD-HQGQ-FVFVG---------------SYNAGNVSVTRLE
NC_cmle  73  EI--VHEASHPIGGHPRANDA-DTNTRAIFLLAAKQPPYAVYANPFYKFAGY-GNVFSVS

MT_1    148  DPAAIKLIKSLVYPAPT-DGLPDPQNSARPHEAILDPTGEFLVFPDLGADQIRVLKVDKE
PA_1    150  DPAKLQSLDNKPFP----PRADNETITSRPHQAVVDPTGGFVVIPDLSVDVLHIFSIDQT
TR      144  NPNSLKLLQTKTFKLTGPPSRPDRDQDAPHPHEAVLDPTGKFLLVPDLGMDLIHLYSFDPN
AN      140  LNQGDENLQVFRYTLSET-KQNPQQNAPHPHQVLLDPTGSFILVPDLGADLVRVYAIDKS
MT_2    141  FGGGQR-LALEKFTMEGQ-GPNPRQDVPHPHEAQVDPTGNYMVVPDLGADLLRVFRINAE
EC_pgl  115  DGLPVGV-------------VDVVEGLDGCHSANISPDNRTLWVPALKQDRICLFTVSDD
NC_cmle 129  ET----------GKLEKNVQNYEYQENTGIHGMVFDPTETYLYSADLTANKLWTHRKLA-

MT_1    207  TLEYVEKPSYTDFDRGTGPRHGAFFKSCD-------------KTFFYLVGELSNLLQGFS
PA_1    206  ALTLTEL-PAHPFGNGTGPRHAAFLKSGD-------------KTFLYVIAEKKVSILGFE
TR      204  TLALKDI-TPLSVKPGSGPRHITFVVKGS-------------KTFAYLVTELGNTIIGYD
AN      199  SGELDGTCPDLSYPEGSGPRHGLFWQTEHSTLTRQLKTRQQDSQVLYIVDELDGHFKGFT
MT_2    199  TGELTA-CSEGQAGPGDGPRHIVFWKNAEG------------LQKAYVVNELGNSVSAWD
EC_pgl  162  GHLVAQDPAEVTTVEGAGPRHMVFHPN--------------EQYAYCVNELNSSVDVWE
NC_cmle 178  SGEVELVGSVDAPDPGDHPRWVAMHPTGN--------------YLYALMEAGNRICEYV

MT_1    254  VAYNDD---------DSLTFTRIHNSTTHCDDKPLP---ED-TAAAELWIAPGSNFLTLS
PA_1    252  VSYGTN---------SLTLSEKFNIRTDGSENAPA---EG-SSGAEITISPNNKFLTVS
TR      250  VTYPNG---------QIKLTEIFNIPSHCAGPAEP---SS-YAASEVVVSPDTNYLIVS
AN      259  VSYTSD---------GCLGFDEFQSFEPYTDGQ-LP---DG-ATPSEIRQA--GNSLYVS
MT_2    246  VEYPEDDDDDEAAAGGCLALNKTQTLSTYEPGTSGG---PT-TKAAEIRVV--GNFLYAS
EC_pgl  207  LKDPHG---------NIECVQTLDM--MPEN--FS---DT-RWAADTHITPDGRHLYAC
NC_cmle 223  IDPATHM---------PVYTHHSFPLIPPGIPDRDPETGKGLYRADVCALTFSGKYMFAS

MT_1    301  SRFESSLEYT-VANG-TKVP-SDPLITFSIDKETCALTHVQS--APAG-GINPRHFSFNS
PA_1    298  TRNETTLEYTSVADG-TKIP-SDALNTFSIDPATGELTHVQS--APAG-GSFPRHFSFNK
TR      296  SRAENSTSIPDFDDPSRIIP-SDPLINFKINPTTGQLQLLQV--VPAG-GQFPRQFSINK
AN      303  IRSDQGF----------AP-NDSMANLDHS-SNNTITLNEL--TSSY-GTVPRTFVINQ
MT_2    300  NRADETF----------GPGQDSIATYTIDEQTGELAWLGA--ANSY-SYYPRTFEFNR
EC_pgl  249  DRTASL-----------------ITVFSVSEDGSVLSKEGF--QP-T-ETQPRGFNVDH
NC_cmle 274  SRANKF-E-----------L-QGYIAGFKLRDCGSIEKQLFLSPTPTSGGHSNAVSPCPW

MT_1    355  DGTRVASALQSDGRVVVFERDPSTGKIGKATAEGDV--------EGMPNFATF-KQ
PA_1    353  DGSLVAVACGGENRVNVFERDVGTGMIGKAVGERVL-------TTQVNHVIF-KE
TR      352  EGNLLAVGLQNDGRVVFVDRCPETGLLGGFVAYADI-------EGQITAAIFDQK
AN      347  AGDLVAIGDQSSSNVAIVARDPQTGKLGDEVANVQIGEPGVVGTSTGLSSVVW-DE
MT_2    346  DGTLVAVGGQTSSNVAIIARDPDTGKLGNLVANLEVGRKGRAGQEDGLSAVVW-VQ
EC_pgl  287  SGKYLIAAGQKSHHISVYEIVGEQGLLHEKGRYAVG-QGP---MWVVVNAHEG-GS
NC_cmle 321  SDEWMAITDDQEGWLEIYRWKDE--PLH-RVARVRIPEPGF------GMNAIW-YD
```

Figure 5. Multiple sequence alignment of proteins showing sequence similarity to *S.thermophile* lactonase 1. Full circle indicates a residue in the predicted active site of NC_cmle. MT_1; *Myceliophthora thermophila* lactonase 1, PA; *Podospora anserina* lactonase 1, TR *Trichoderma reesei,* AN; *Aspergillus niger*, MT_2; *Myceliophthora thermophila* lactonase 2, EC_pgl; *Escherichia coli* 6-phosphogluconolactonase, NC_cmle; *Neurospora crassa* cis-carboxy-muconate-lactonizing enzyme
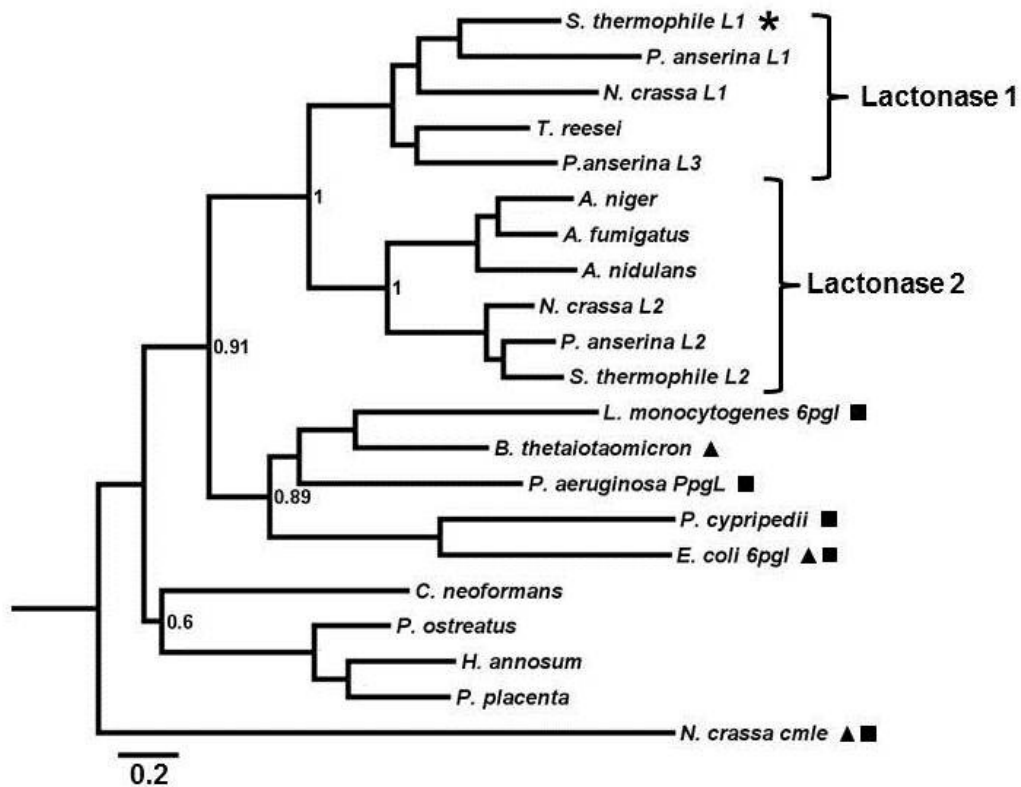
Figure 6. Phylogenetic tree of representative sequences showing sequence similarity with *M. thermophila* lactonase 1. Bootstrap values from 100 iterations are listed. (*) *M. thermophila* lactonase 1, (■) protein that has been characterized biochemically, (▲) protein for which an x-ray crystal structure has been solved. Full amino acid sequences are available in supplemental materials.

A sequence alignment and maximum likelihood phylogeny of representative homologues shows two distinct clades within the ascomycetes (Figure 6). Many sordariomycetes have two proteins with homology to the *M. thermophila* lactonase, represented as group 1 and group 2 lactonases in Figure 5. The lactonase characterized in this study is part of the group 1 lactonases, and will be referred to as lactonase 1. Group 1 lactonases are less conserved than group 2 lactonases, and the sequence identity between the two groups is 25-30%. The aspergilli only possess a single copy of a group 2 lactonase. Expression of the group 2 lactonase in *N. crassa* was shown to be highly upregulated in response to growth on pure cellulose or ground *Miscanthus* stems and was identified in the secretome under both conditions (*114*).

3.3.5 Induction of extracellular lactonases by cellulose.

Lactonase activity in the culture broth of *M. thermophila* was approximately 4-fold higher on Avicel compared to glucose after four days of growth (Figure 7a). Knowing the nucleotide sequence, we investigated the expression levels of Spoth1|109678

(lactonase 1) and Spoth1|89286 (lactonase 2), a group 2 lactonase, the two lactonases present in the *M. thermophila* genome, using mRNA-Seq (*115*) under the same growth conditions (Figure 7b). Lactonase 1 showed a relatively high basal expression level on glucose, but was upregulated six fold during growth on Avicel. Lactonase 2 was expressed at extremely low levels on glucose, and strongly induced by Avicel, although the absolute level of expression was substantially lower than that of lactonase 1.
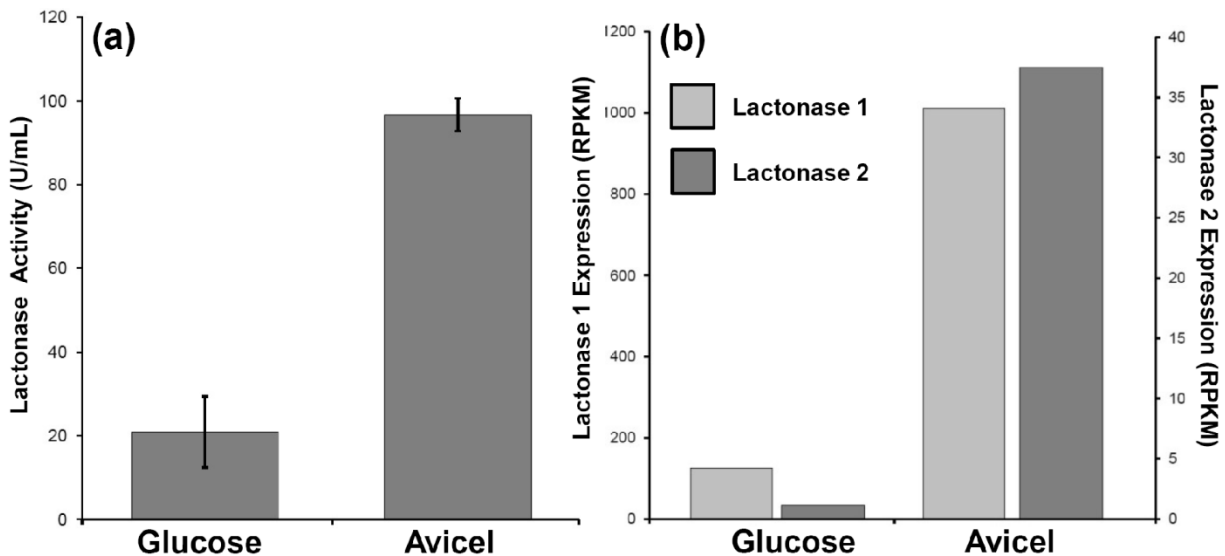

Figure 6. Induction of extracellular lactonase by *M. thermophila* in response to Avicel. (a) Specific activity of culture broth after 4 days of growth on Vogel's salts supplemented with 2% glucose or Avicel. Activity represented in arbitrary units (a.u.). (b) RNA-seq expression profiling of lactonase 1 (Spoth1|109678) and lactonase 2 (Spoth1|89286) after 20 hours of growth on glucose or Avicel. Expression is normalized based on the number of mapped reads per kilobase transcript per million mapped reads (RPKM). Absolute expression of lactonase 1 is approximately 20 fold higher than lactonase 2. Light gray; lactonase 1, black; lactonase 2.

3.4 Discussion

Two evolutionarily distinct groups of enzymes have previously been described with 6-phosphogluconolactonase (PGL) activity. PGL activity in eukaryotes and some bacteria is catalyzed by enzymes with homology to the *E.coli* glucosamine-6-phosphate isomerase, NagB (*102*). Recently, the gene responsible for PGL activity was determined in *E. coli (104, 105)*. *E. coli* PGL has no sequence similarity with the NagB-like PGLs, and 13% sequence identity with *N. crassa* cis-carboxy-muconate-lactonizing enzyme (CMLE), an enzyme involved in the β-ketoadipate pathway (*112*). Homologues of *E. coli* PGL are ubiquitous throughout the bacteria and fungi. In *M. thermophila* and other ascomycetes, the NagB-like PGL is probably responsible for PGL activity in the pentose phosphate pathway because it is predicted to be intracellular and shows ~40% sequence identity to human and yeast PGL. Interestingly, in addition to the NagB-like PGL, many ascomycetes have distant homologues of *E. coli* PGL that are predicted to

44

be extracellular proteins and probably have functions unrelated to that of PGL in the pentose phosphate pathway.

In this work we identified a gene, with low homology to *E.coli* PGL, that encodes an enzyme responsible for extracellular lactonase activity in *M. thermophila*. Lactonase 1, encoded by Spoth1|109678 is a glycoprotein and has a broad pH optimum, similar to the pH profile reported for *N. crassa* CMLE. The enzyme displays high activity with hexose δ-lactone substrates, but no detectable activity with pentose γ-lactones. Notably, there is only a threefold difference in k_{cat}/K_M values for the δ-lactones of glucose, cellobiose, and lactose, suggesting that the enzyme is only interacting with the glucose moiety. This is in stark contrast to the substrate specificity of cellobiose dehydrogenase and many cellulases, which often have much higher affinity for cellobiose over glucose (*97, 116*). Structural studies on *N. crassa* CMLE and *E. coli* PGL have shown that the active sites of these enzymes are highly conserved and probably on or near the surface of the protein (*111*). A surface exposed active site is consistent with our kinetic results and may explain the small differences in substrate specificity for mono and disaccharides. The complete lack of activity on the pentose lactones is probably due to the smaller 5-membered ring structure and different geometry of the pentose γ-lactones. Cellobiose dehydrogenase and glucose oxidase both have very low activities on pentose sugars, so pentose lactones may be uncommon under physiological conditions.

During growth on lignocellulosic biomass, *M. thermophila* induces the expression of two *E.coli* PGL-like lactonases. A recent, genome wide expression profiling study of *N. crassa* identified NCU07143, a protein with sequence homology to the *M. thermophila* lactonases, as being strongly upregulated during growth on lignocellulose and pure cellulose (*114*) NCU07143 was also detected in the secretome of *N. crassa* while growing on cellulose and lignocellulose, confirming that it is also an extracellular enzyme. *M. thermophila* lactonase 1 and lactonase 2 share approximately 33% sequence identity. Despite the low sequence identity, all of the predicted active site residues are conserved in both lactonases and they are more closely related to one another than any predicted bacterial lactonases. It is unclear what the functional differences between the two *M. thermophila* lactonases may be, but orthologues of lactonase 2 are more widespread and have a higher degree of sequence conservation than lactonase 1, suggesting a potentially more essential biological role. The direct orthologue of lactonase 1 in *N. crassa*, NCU07703, was not expressed under any condition tested by Tian et al (*114*). The induction characteristics of lactonase 2, and its orthologue in *N. crassa*, NCU07143, are very similar to the expression profile for cellulases and other enzymes known to be essential for plant cell wall metabolism.

Genes with sequence similarity to the *N. crassa* and *M. thermophila* lactonases are also present in *Aspergillus niger* and *Trichoderma reesei*, and are probably what was

isolated in Bruchmann's studies on the role of lactonases in cellulose degradation (*107*). Extracellular lactonase-like genes appear to be absent in yeasts, but widespread in filamentous ascomycetes. All cellulolytic ascomycetes with sequenced genomes have genes with homology to *M. thermophila* lactonase 1 or lactonase 2. BLAST queries against the genomes of currently sequenced basidiomycetes only return proteins with very low homology, none of which are predicted by SignalP to be extracellular. The genome of the basidiomycete, *Phanerochaete chrysosporium* (*117*), contains two proteins, Pchr1|3055 and Pchr1|7029, with 15-19% sequence identity to *M. thermophila* lactonase 1, but neither are predicted to be secreted. These proteins may function similarly to ascomycete lactonases, but experimental characterization of the localization and substrate specificities are necessary. The ascomycete lactonases may be working synergistically with cellobiose dehydrogenases (CDH) or glucose oxidases (GOX), although the genes do not appear to be co-conserved since the filamentous ascomycete, *Trichoderma reesei*, has a gene for a lactonase, but does not possess genes for CDH or GOX.

A periplasmic gluconolactonase (PpgL) with low sequence homology to the *M. thermophila* lactonases was recently characterized in the bacterium *Pseudomonas aeruginosa (113)*. Like in fungi, *P. aeruginosa* contains a NagB-like intracellular PGL that provides lactonase activity in the pentose phosphate pathway *(103)*. Deletion of PpgL caused severe growth phenotypes on gluconate, 2-ketogluconate, and mannitol as well as a decrease in pigment formation. *P. aeruginosa* is not cellulolytic, so the PpgL is not involved in any cell wall degradation processes, highlighting the diverse role of lactonases in metabolism. Further studies will need to be conducted in fungi, particularly gene deletion experiments, to clarify the biological role of the extracellular lactonases.

3.5 Materials and Methods

Chemicals. With the exception of the sugar lactones, all chemicals were of reagent grade quality and purchased from commercial vendors. The cellulose used for growth studies was Avicel PH-101 purchased from Fluka. The sugar lactones were synthesized by a modified Frush and Isbell procedure (*118*). In 100 mL of ice-cold water, 0.015 mole of sugar, 0.03 mole of calcium carbonate, and 0.02 mole of bromine were mixed and stored in the dark for 24 hours at room temperature. Residual bromine was removed by purging the solution with nitrogen gas for 1 hour. Excess silver carbonate was then added to the solution and the precipitate removed by filtration. The filtrate was then applied to an amberlite IR-120(H+) resin. The column was washed with 5 column volumes of water and the eluate was collected. The eluate was then concentrated on a rotary evaporator at 55 °C to dryness. Purity was confirmed by HPLC.

Strains and growth conditions. *M. thermophila* ATCC strain 42464 was obtained from the American Type Culture Collection. *M. thermophila* was maintained on Vogel's salts agar (*119*) containing 2% cellobiose. Conidia were isolated from agar plates grown for 7 days at 48 °C. For induction studies fresh *M. thermophila* conidia were inoculated into liquid culture containing Vogel's salts supplemented with either 2% glucose or Avicel and grown at 48 °C and 200 RPM shaking.

For lactonase purification, *M. thermophila* conidia were inoculated into a complex media containing 1.0 g/L casamino acids, 1.0 g/L yeast extract, 0.5 g/L potassium chloride, 0.2 g/L magnesium sulfate heptahydrate, 1.0 g/L potassium dihydrogen phosphate, and trace elements solution. The culture was grown at 48 °C with shaking at 200 rpm for 24 hours and then the mycelia was harvested by filtration and used to inoculate fresh complex media, containing 5 g/L Avicel to induce expression of the lactonase.

Enzyme assays. Lactonase activity was measured as described by Hestrin (*120*). Freshly dissolved sugar lactone (100 µL) was immediately mixed with 100 mM buffered solutions containing enzyme (100 µL) for 1 minute at 25 °C. The reaction was quenched by the addition of alkaline hydroxylamine (2.0 M hydroxylamine hydrochloride, 1.5 M sodium hydroxide, 40 µL) for 1 minute. Hydrochloric acid was then added to acidify the solution and stabilize the hydroxamate (4.0 M HCl, 20 µL). Color was developed by the addition of ferric chloride ($FeCl_3$, 0.5 M, 20 µL) and absorbance at 540 nm was used for quantification. For all experiments, control measurements were made where no enzyme was added. The background hydrolysis was subtracted from the total hydrolysis in the presence of enzyme. This typically amounted to less than 10% of the total lactone.

Protein purification. After 48 hours of growth on complex media containing Avicel, *M. thermophila* mycelia and residual Avicel were removed from the culture by filtering over 0.2 micron polyethersulfone (PES) filters. The filtered culture broth was then concentrated and buffer exchanged using tangential flow filtration with a 10 kDa MWCO PES membrane (Millipore, Billerica, MA) into 25 mM HEPES buffer pH 7.4. The concentrated culture broth was then mixed with excess Avicel to remove cellulose-binding proteins and the Avicel particles were separated by vacuum filtration over a glass microfiber filter. The filtrate was then loaded onto a 5 mL QHP column (GE Healthcare) at 5 mL/min. Lactonase was eluted from the Q HP column with a linear gradient from 0 to 0.5 M NaCl over 15 column volumes. Fractions with the highest lactonase activity were pooled and buffer exchanged into 25 mM HEPES buffer pH 7.4. The protein was then loaded onto a MonoQ 10/100 GL column (GE Healthcare) and eluted with a linear gradient from 0 to 0.5 M NaCl over 8 column volumes. Fractions containing lactonase activity were pooled and adjusted to 1.5 M ammonium sulfate, 25 mM HEPES pH 7.4. The sample was then loaded onto a 1.0 mL Resource PHE

column. Lactonase does not bind the PHE column and flows through as a pure species. The purified lactonase was concentrated and buffer exchanged with 10 kDa MWCO spin concentrators (Millipore) and stored at -80 ºC. Lactonase activity is stable for several months under these conditions.

Mass spectrometry peptide fingerprinting. Thirty-six milligrams of urea, 5 µL of 100 mM DTT, and 5 µL of 1.0 M Tris pH 8.5, were added to a 100 µL aqueous solution of 10 µM lactonase and heated at 70 °C for 1 hour. After heat denaturation, 700 µL of 25 mM ammonium bicarbonate and 140 µL of methanol were added to the solution followed by treatment with 50 µL of 100 µg/mL trypsin in 50 mM sodium acetate, pH 5.0. The protein was digested with trypsin overnight at 37 ºC. After digestion the volume was reduced by Speedvac and washed with Milli-Q water three times. Residual salts in the sample were removed by OMIX microextraction pipette tips according to the manufacturer's instructions (Varian, Palo Alto, CA). Peptides from the trypsin digest were analyzed using a tandem mass spectrometer that was connected in-line with ultraperformance liquid chromatography as described elsewhere (*114*).

LC-MS analysis of proteins. Protein samples were analyzed using an Agilent 1200 series liquid chromatograph (LC; Santa Clara, CA) that was connected in-line with an LTQ Orbitrap XL hybrid mass spectrometer equipped with an Ion Max electrospray ionization source (ESI; Thermo Fisher Scientific, Waltham, MA).

The LC was equipped with C_8 guard (Poroshell 300SB-C8, 5 µm, 12.5 × 2.1 mm, Agilent) and analytical (75 × 0.5 mm) columns and a 100 µL sample loop. Solvent A was 0.1% formic acid/99.9% water and solvent B was 0.1% formic acid/99.9% acetonitrile (v/v). Sample solutions contained in autosampler vials sealed with rubber septa caps were loaded into the Agilent 1200 autosampler compartment prior to analysis. For each sample, approximately 100 picomoles of protein analyte was injected onto the column. Following sample injection, analyte trapping was performed for 5 min with 99.5% A at a flow rate of 90 µL/min. The elution program consisted of a linear gradient from 30% to 95% B over 24.5 min, isocratic conditions at 95% B for 5 min, a linear gradient to 0.5% B over 0.5 min, and then isocratic conditions at 0.5% B for 9.5 min, at a flow rate of 90 µL/min. The column and sample compartments were maintained at 35 ºC and 10 ºC, respectively. Solvent (Milli-Q water) blanks were run between samples, and the autosampler injection needle was rinsed with Milli-Q water after each sample injection, to avoid cross-contamination between samples.

The connections between the LC column exit and the ESI probe of the mass spectrometer were made using PEEK tubing (0.005" i.d. × 1/16" o.d., Western Analytical, Lake Elsinore, CA). External mass calibration was performed prior to analysis using the standard LTQ calibration mixture containing caffeine, the peptide MRFA, and Ultramark 1621 dissolved in 51% acetonitrile/25% methanol/23% water/1%

acetic acid solution (v/v). The ESI source parameters were as follows: ion transfer capillary temperature 275 ºC, normalized sheath gas (nitrogen) flow rate 25%, ESI voltage 2.5 kV, ion transfer capillary voltage 33 V, and tube lens voltage 125 V. Mass spectra were recorded in the positive ion mode over the range m/z = 500 to 2000 using the Orbitrap mass analyzer, in profile format, with a full MS automatic gain control target setting of 5×10^5 charges and a resolution setting of 6×10^4 (at m/z = 400, FWHM). Raw mass spectra were processed using Xcalibur software (version 4.1, Thermo) and measured charge state distributions were deconvoluted using ProMass software (version 2.5 SR-1, Novatia, Monmouth Junction, NJ), using default "large protein" parameters and a background subtraction factor of 1.5.

Nanoelectrospray ionization mass spectrometry of native proteins. Mass spectra of native proteins were acquired using a quadrupole time-of-flight (Q-Tof) mass spectrometer equipped with a Z-spray electrospray ionization (ESI) source (Q-Tof Premier, Waters, Milford, MA). Ions were formed from aqueous solutions containing 10 μM analyte protein and 10 mM ammonium acetate, using positive-ion nanoelectrospray ionization (nanoESI). NanoESI emitters were made from borosilicate glass capillary tubes (1.0 mm o.d./0.78 mm i.d, Sutter Instruments, Novato, CA) that were pulled to a tip with an inner diameter of approximately 5 to 20 μm using a Flaming/Brown micropipette puller (Model P-87, Sutter). Approximately 10 μL of the sample solution was added into a nanoESI emitter using a 10 μL syringe (Hamilton, Reno, NV). The electrospray was initiated by gradually increasing the DC potential applied to a platinum wire (0.127 mm diameter, Aldrich, Milwaukee, WI), which was inserted into the nanoESI emitter to within approximately 2 mm of the tip, until the onset of mass spectral signal. No back pressure was used for nanoESI. Instrument parameters during data collection were as follows: nanoESI voltage 1.8 kV, sampling cone voltage 30 V, extraction cone and ion guide voltages both 4.0 V, source block temperature 80 ºC, accelerating voltage into the argon-filled cell 2 V, ion transfer stage pressure 6×10^{-3} mbar, argon-filled cell pressure 8×10^{-3} mbar, Tof analyzer pressure 8×10^{-7} mbar. The pressure in the first pumping stage was increased to 7.4 mbar by adjusting a Speedivalve (BOC Edwards) to favor the preservation of non-covalent complexes in gas phase. No cone gas flow was used. The Tof analyzer was operated in "V" mode. External mass calibration of the TOF analyzer was performed immediately prior to measurements. Mass spectra were processed using MassLynx software (version 4.1, Waters).

Multiple sequence alignments and phylogenetics. Homologues of the lactonase were found with a BLAST (*121*) query of the lactonase sequence against a database of predicted fungal proteins from finished and ongoing fungal genome projects. Bacterial sequences were obtained by BLAST against the NCBI database. Multiple sequence alignments were done locally using T-COFFEE (*122*). A maximum likelihood phylogeny

was determined using PhyML version 3.0 with 100 bootstraps through the Phylogeny.fr webserver (*123*).

mRNA-Seq expression profiling. Messenger RNA was isolated from *M. thermophila* after 20 hours of growth and Illumina cDNA libraries were generated following standard protocols according to the manufacturer's instructions (*124*). Sequencing was performed on the Illumina Genome Analyzer II. Read lengths were trimmed to 31 nt and mapped against all predicted transcripts in the *M. thermophila* genome using MAQ software (*125*). Expression was normalized by counting the number of reads mapped per kilobase of exon model divided by the total number of mapped reads in the whole dataset, RPKM (*115*).

3.6 Acknowledgements

*Footnote to chapter 3

This chapter, presented here with few modifications from its original format, represents a peer-reviewed paper published in *Applied and Environmental Microbiology* (2011), 77, 650-656.   The authors are William T. Beeson, Anthony T. Iavarone, Corinne D. Hausmann, Jamie H.D. Cate and Michael A. Marletta.

Chapter 4: Contribution of cellobiose dehydrogenase to the degradation of cellulose

4.1 Abstract

Enzymatic hydrolysis of lignocellulose by cellulases is a slow and expensive process which must be improved for the economical production of cellulosic biofuels. In addition to cellulases, many fungi secrete cellobiose dehydrogenases during growth on lignocellulose. Here we report on the deletion of *cdh-1,* the gene encoding the major cellobiose dehydrogenase (CDH) of *Neurospora crassa*. Deletion of *cdh-1* reduced cellulase activity in the culture filtrate of the Δ*cdh-1* strain to levels 37-49% lower than in the wild-type strain. Addition of purified CDHs to the Δ*cdh-1* strain culture filtrate resulted in a 1.6 to 2.0-fold stimulation in cellulase activity, which restored activity to wild-type levels, while addition of CDH to a mixture of purified cellulases showed no stimulatory effect. Domain truncations of a full length CDH showed that the heme domain is required for the observed stimulatory effects on cellulase activity. The stimulatory effect of CDH requires the presence of molecular oxygen and likely other metalloproteins. The discovery that cellobiose dehydrogenase plays an integral role in degradation of pure cellulose is a significant shift from previous models centered on the degradation of cellulose by mixtures of glycosyl hydrolases.

4.2 Introduction

Fungi play a central role in the degradation of plant biomass in terrestrial environments (*126*). They use a wide variety of secreted enzymes to break down biopolymers present in the plant cell wall, including cellulases, hemicellulases, pectinases, proteases, and oxidative enzymes (*79, 127-129*). Many of these fungal enzymes are produced industrially and used in diverse applications (*130*). Currently, the cost of cellulose degrading enzymes is still a major barrier to the economical production of liquid fuels from lignocellulose (*131*). Considerable effort has focused recently on the discovery and development of cellulases with better properties, but improvement of the catalytic activity of cellulases has proved to be a slow and challenging process.

The filamentous ascomycete, *Neurospora crassa,* is a well-known genetically tractable organism that proficiently degrades plant cell walls (*128, 132*). *N. crassa* secretes more than 30 different enzymes during growth on pure cellulose. Many of these enzymes are similar to cellulases and hemicellulases that have been extensively studied in other organisms, but there are also a significant number of enzymes with poorly defined or no known biological functions. In addition to hydrolytic enzymes, *N. crassa* also produces large amounts of cellobiose dehydrogenase (CDH) (*133-135*). CDH-1 is the major oxidoreductase secreted by *N. crassa* during growth on crystalline cellulose, and accounts for approximately 1-3% of the total secreted protein, suggesting a potential role in the degradation process.

CDH was first described over 35 years ago and since then has been identified in phylogenetically diverse fungi (*108*), but is notably absent from the industrial workhorse *Hypocrea jecorina (anamorph Trichoderma reesei)*. The CDH from the white-rot fungus, *Phanerochaete chrysosporium,* has been extensively studied biochemically and by x-ray crystallography (*58, 136, 137*). CDH catalyzes the oxidation of cellobiose and longer cellodextrins to the corresponding 1-5-δ-lactones. All CDH enzymes contain an N-terminal heme domain and a C-terminal flavin domain. The flavin domain is part of the larger glucose-methanol-choline (GMC) oxidoreductase superfamily (*138*), which is widespread throughout all domains of life, while orthologs of the heme domain are only found in fungi (*139*). Oxidation of cellobiose takes place in the flavin domain with subsequent electron transfer to the heme domain. The biological electron acceptor for CDH is unknown, but the reduced heme is able to reduce a wide variety of substrates including quinones, metal ions, and organic dyes. Reduced CDH can also react with molecular oxygen, but at a 10-20 fold slower rate than organic dyes and metal ions (*140*).

Several hypotheses have been proposed for the function of CDH including the oxidative degradation of cellulose or lignin via hydroxyl radical production (*53, 137, 141, 142*), reduction of feedback inhibition of cellulases by cellobiose (*143*) or support of ligninolytic enzymes by quinone and/or manganese reduction (*144*). The most

prominent hypothesis is that involving the generation of hydroxyl radicals formed via oxidation of cellobiose followed by heme domain reduction of an extracellular ferric complex, most likely Fe(III)-oxalate (*137*). Ferrous iron can then take part in a Fenton reaction in the presence of hydrogen peroxide produced by CDH or another oxidase. While the scope of these reactions is possible, control of hydroxyl radical reactivity would be particularly challenging.

In this study, we have determined the contribution of CDH to cellulose degradation in *N. crassa* using a combination of genetic and biochemical techniques. CDH enhances the degradation of cellulose by the cocktail of enzymes secreted by *Neurospora* nearly 2-fold, but this enhancement is not dependent on low molecular weight ferric ion complexes. This enhancement of activity is also not seen when CDH is added to purified *N. crassa* cellulases. Taken together, our results show that in *N. crassa*, CDH is a central component in cellulose degradation and works in conjunction with other factors to facilitate the degradation of crystalline cellulose by cellulases.

4.3 Results

4.3.1 Production of a strain of *N. crassa* containing a deletion of NCU00206, *cdh-1*

The *Neurospora* functional genomics project has generated knockout strains for most of the genes in the *N. crassa* genome using targeted gene replacement through homologous recombination (*145*). A heterokaryon strain of Δ*cdh-1* is available through the Fungal Genetic Stock Center (FGSC), but despite numerous attempts, a homokaryon strain could not be generated due to an ascospore-lethal linked mutation. To obtain a clean deletion of *cdh-1,* a *N. crassa* strain deficient in non-homologous end joining recombination was transformed with a cassette provided by the *Neurospora* functional genomics project (*146*). Heterokaryon transformants showing hygromycin resistance were genotyped using PCR to confirm the deletion of *cdh-1*. Transformants were crossed with wild-type *N. crassa* and several hygromycin resistant progeny were then screened for the production of CDH during growth on cellulose. The strains that showed the best growth on Avicel®, a microcrystalline cellulose, and that were also deficient in CDH activity in the culture filtrate were genotyped. Multiple homokaryon strains in which *cdh-1* was deleted were confirmed by PCR. Growth of the Δ*cdh-1* strains in liquid culture on Vogel's salts supplemented with 2% sucrose was identical to that of wild-type. There was only a slight growth defect on cellulose. Both the wild-type and Δ*cdh-1* strains completely degraded all the cellulose in the culture after 6-7 days of growth, as determined by light microscopy. The proteins present in the culture filtrate were analyzed by SDS-PAGE (Fig. 1A) and the extracellular proteins secreted by the Δ*cdh-1* strains were very similar to those of the wild-type, with the exception of the loss of the CDH-1 band between 100 and 120 kDa. The total secreted protein in the Δ*cdh-1* strains varied from ~40% lower than the wild-type strain to equal to the wild-type strain for the different transformants. CDH activity in the culture filtrate of the Δ*cdh-1* strains

was 800 ± 300-fold lower than in the wild-type culture filtrates (Fig. 1B). Cellulase-specific activities of the Δ*cdh-1* strains and the wild-type were then compared. The endoglucanase activity and cellobiohydrolase activity, as measured by the azo-CMC and MULAC assays, respectively, were similar for the wild-type and Δ*cdh-1* strains when equal levels of total protein were added. Cellulase activity was 37-49% lower in the Δ*cdh-1* strain culture filtrates than in those from wild-type when added on an equal protein basis (Fig. 1C). HPLC analysis of hydrolysis products after 24 hours of reaction time showed that in the Δ*cdh-1* strain culture filtrate, glucose (>90%) was the main sugar produced followed by cellobiose. Glucose remained the dominant product (>80%) in assays of wild-type culture filtrate, followed by cellobiose, cellobionic acid and trace amounts of gluconic acid. No additional peaks were present in the chromatograms.
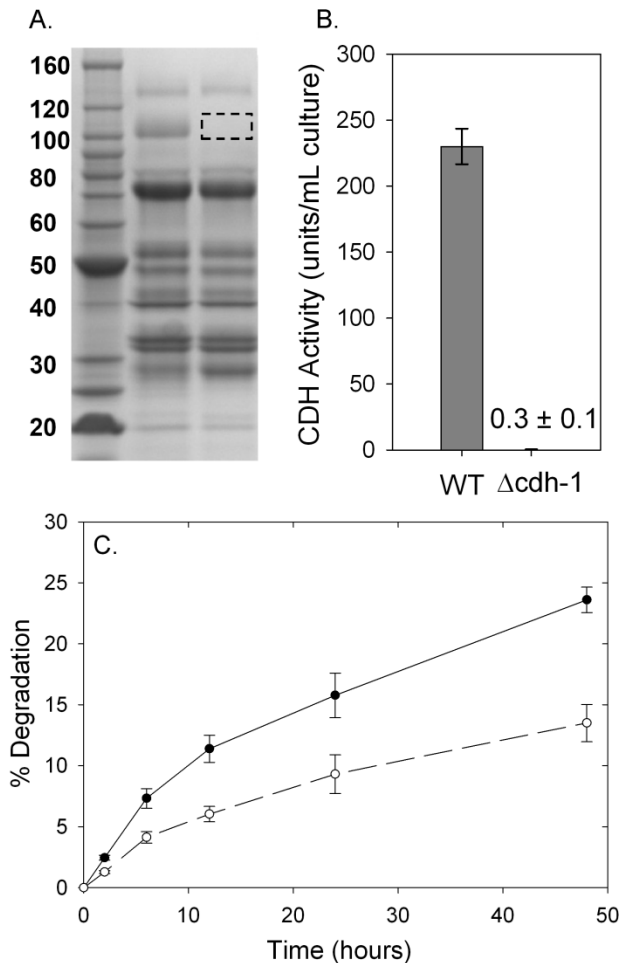

Figure 1. Deletion of *N. crassa* CDH-1. (A) SDS-PAGE of proteins present in the culture filtrate of the wild-type and the Δ*cdh-1* strain of *N. crassa* after 7 days of growth on Avicel®. Missing protein band that corresponds to CDH-1 is marked by a box. (B) CDH activity in the culture filtrate of the wild-type and Δ*cdh-1* cultures as measured by the cellobiose-dependent reduction of DCPIP. Values are the mean of three biological replicates. Error bars are the SD between these replicates. (C) Cellulase activity of the

wild-type (●) and Δ*cdh-1* (○) culture filtrates. Values are the mean of three biological replicates performed in technical triplicate. Error bars are the SD between these replicates.
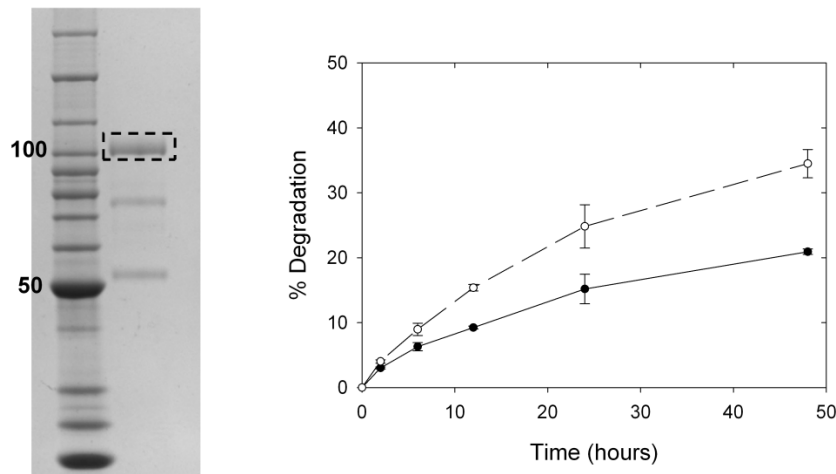

Figure 2. Stimulation of cellulose degradation by the addition of partially purified *N. crassa* CDH-1 to the Δ*cdh-1* culture filtrate. (A) SDS-PAGE of partially purified *N. crassa* CDH-1. (B) Cellulase activity of the Δ*cdh-1* culture filtrate. (○) Represent experiments where 400 μg *N. crassa* CDH-1 per gram of cellulose was added. (●) Represent experiments where no exogenous CDH was added. Values are the mean of three replicates. Error bars are the SD between these replicates.

4.3.2 Stimulation of cellulose degradation by CDH

To more directly assess the contribution of CDH-1 to the degradation of cellulose, *in vitro* complementation assays were undertaken using purified CDHs. CDH-1 is difficult to purify from *N. crassa* culture filtrates, and only a partially purified preparation of *N. crassa* CDH-1 could be isolated (Fig. 2). The orthologous protein in the closely related thermophilic fungus, *Myceliophthora thermophila* (*140*), can be purified to homogeneity and was used for most of the complementation assays (Fig. 3). *M. thermophila* and *N. crassa* CDH-1 share 70% sequence identity, the same domain architecture and a C-terminal fungal cellulose binding domain. CDH-1 from *M. thermophila* had no activity on cellulose, while the partially purified *N. crassa* CDH-1 had a slight contaminating hydrolytic activity.
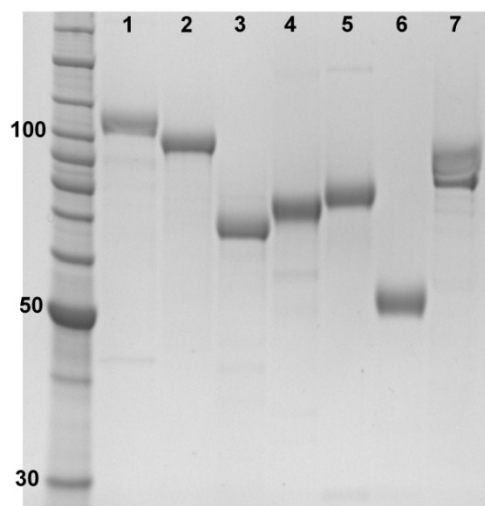
Figure 3. SDS-PAGE of purified proteins used throughout the study. All proteins were loaded at 5 µg per lane in the following order: (1) *M. thermophila* CDH-1, (2) *M. thermophila* CDH-2, (3) *M. thermophila* CDH-2 flavin, (4) *N. crassa* CBH-1, (5) *N. crassa* GH6-2, (6) *N. crassa* GH5-1, (7) *N. crassa* GH3-4.
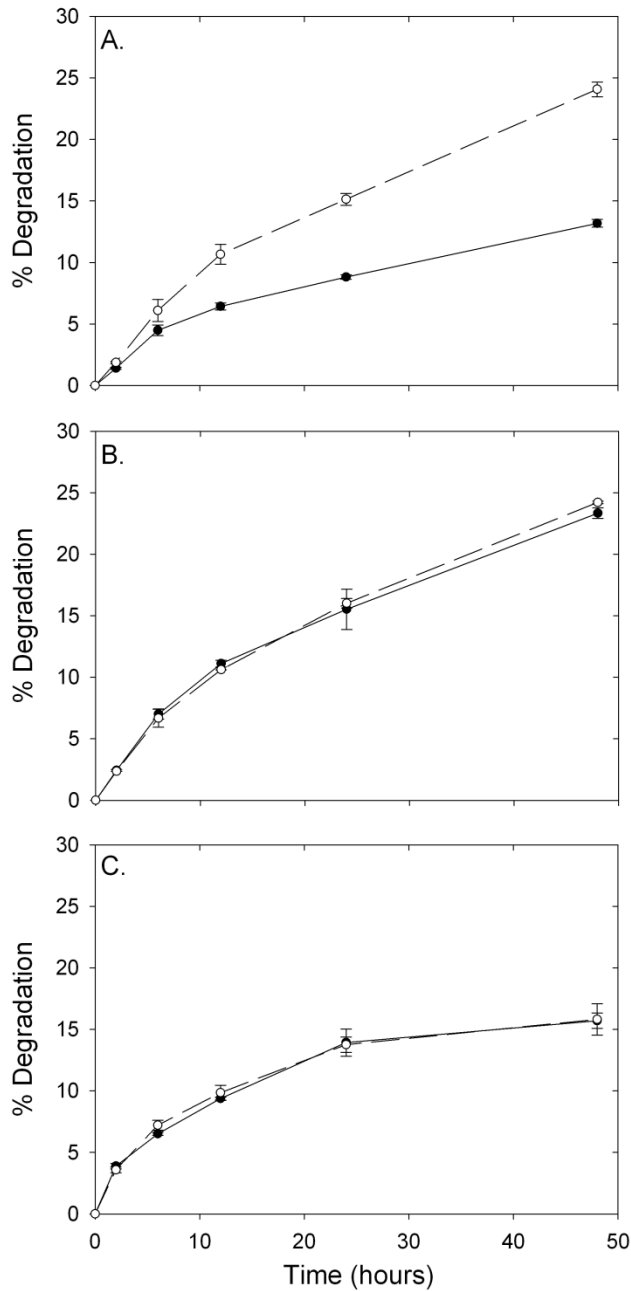
Figure 4. Stimulation of cellulose degradation by the addition of *M. thermophila* CDH-1 to the Δ*cdh-1* culture filtrate. (●) No exogenous CDH added, (○) 400 μg *M. thermophila* CDH-1 per gram of Avicel® added. Cellulase assays with or without addition of *M. thermophila* CDH-1 to (A) Δ*cdh-1 N. crassa* culture filtrate, (B) Wild-type *N. crassa* culture filtrate, or (C) a mixture of purified cellulases (CBH-1, GH6-2, GH5-1, GH3-4) from *N. crassa*. Values are the mean of three replicates. Error bars are the SD between these replicates.
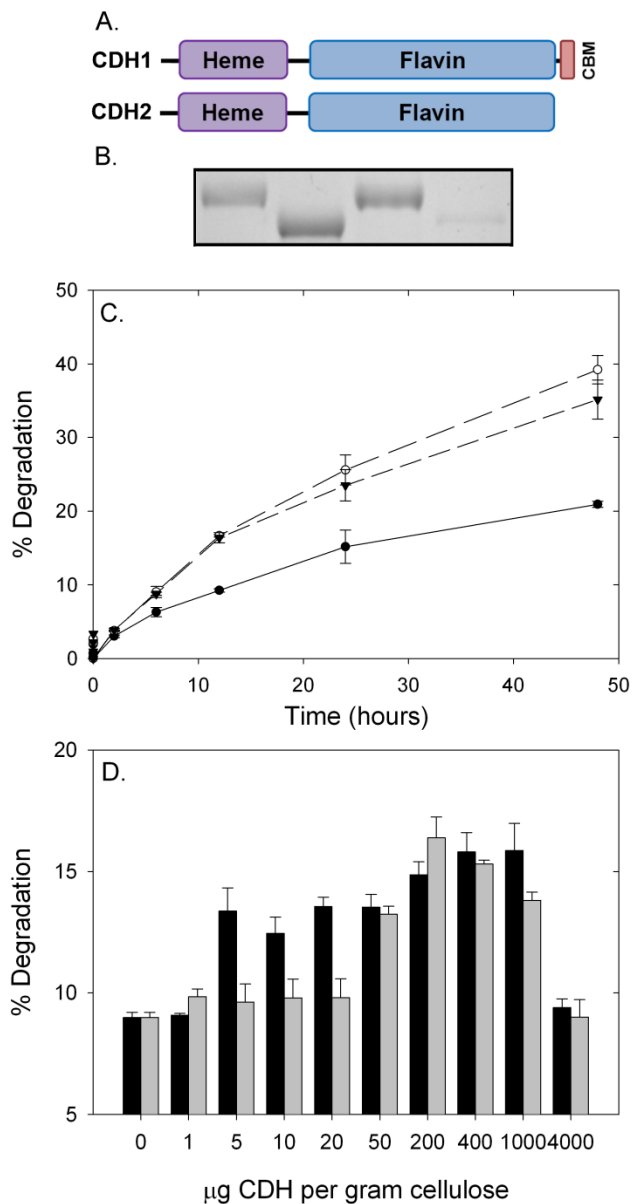
Figure 5. Stimulation of cellulose degradation by different isoforms of CDH. (A) Domain architectures of *M. thermophila* CDH-1 and CDH-2. Red C-terminal domain on CDH-1 is a fungal cellulose-binding domain (CBM1). (B) Cellulose binding assay for *M. thermophila* CDH-1 and CDH-2. Lane 1 *M. thermophila* CDH-1, Lane 2 *M. thermophila* CDH-2, Lane 3 CDH-1 bound to Avicel®, Lane 4 CDH-2 bound to Avicel®. (C) Stimulation of cellulose degrading capacity of the Δ*cdh-1* culture filtrate (●) by addition of CDH-1 (○), or CDH-2 (▼). (D) Effect of the concentration of *M. thermophila* CDH-1 (black) and *M. thermophila* CDH-2 (gray) on cellulase activity of the Δ*cdh-1* culture filtrates at 12 hours. Values are the mean of three replicates. Error bars are the SD between these replicates.

Addition of *M. thermophila* CDH-1 or partially purified *N. crassa* CDH-1 to the culture filtrate of the Δ*cdh-1* strains stimulated cellulose hydrolysis substantially (Fig. 2 Fig. 4).   The cellulase activity was 1.6- to 2.0-fold higher than the Δ*cdh-1* culture filtrate alone.   Addition of CDH-1 to wild-type culture filtrate had no stimulatory effect on cellulose hydrolysis (Fig. 4B).   Further, CDH-1 was unable to stimulate a mixture of purified cellulases (Fig. 4C) from *N. crassa* including 2 cellobiohydrolases (CBH-1 and GH6-2), an endoglucanase (GH5-1), and a ß-glucosidase (GH3-4) (Fig. 3).

M. thermophila also produces a second CDH during growth on cellulose, CDH-2, which does not contain a fungal cellulose binding module (Fig. 5A).   The cellulose binding propensity of *M. thermophila* CDH-1 and CDH-2 was determined using pull down experiments with cellulose (Fig. 5B).   *M. thermophila* CDH-1 binds strongly to cellulose, while *M. thermophila* CDH-2 has only a very weak affinity.   Aside from the different affinities for cellulose, *M. thermophila* CDH-1 and CDH-2 have very similar steady-state kinetic properties (*140*).   At a CDH loading of 0.4 mg/g cellulose, CDH-2 was able to stimulate the hydrolysis of cellulose to the same extent as CDH-1 (Fig. 5C).

To further investigate the role of the cellulose binding module on the ability of CDH to stimulate cellulose hydrolysis, a titration experiment was performed (Fig. 5D). CDH-1 was able to stimulate the activity of the Δ*cdh-1* strain culture filtrate at a 10-fold lower loading than CDH-2.   A stimulatory effect on cellulase activity in the Δ*cdh-1* culture filtrate was observed at a loading of 5 µg of CDH-1 per gram of cellulose, while 50 µg of CDH-2 was required for a similar stimulation (Fig. 5D).   At 4 mg CDH/g cellulose, both *M. thermophila* CDH-1 and CDH-2 have an inhibitory effect on cellulase activity relative to the lower loadings.

The flavin and heme domains of *M. thermophila* CDH-2 can be separated by cleavage with papain (*140*).   To determine the contribution of the heme domain to the stimulation of activity, *M. thermophila* CDH-2 was cleaved with papain and the flavin domain isolated by size exclusion chromatography (Fig. 3).   The flavin domain oxidizes cellobiose at the same rate as the full length enzyme when 2,6-dichlorophenolindophenol (DCPIP) is used as an electron acceptor, but has no activity with cytochrome c, reflecting on the importance of the heme domain for transfer to 1 electron acceptors (*140*).   The flavin domain, when added on an equal activity basis as the full length CDH-2, is unable to stimulate the hydrolysis of cellulose when added to the Δ*cdh-1* culture filtrate, despite production of cellobionic acid (Figure 6).   Even at a loading 10-fold higher than the full length CDH-2, the flavin domain is still unable to stimulate cellulose hydrolysis, suggesting that the heme domain is essential for the stimulatory effect.
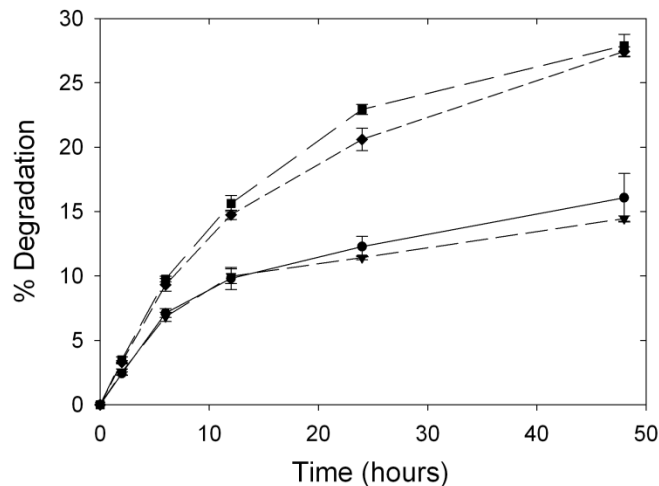
Figure 6. Stimulation of cellulose hydrolysis by domain truncations of CDH-2. Stimulation of cellulose degrading capacity of the Δ*cdh-1* culture filtrate (●) by addition of CDH-2 (■), CDH-2 flavin domain (▼), or recombinant CDH-2 heme domain (♦). Values are the mean of three replicates. Error bars are the SD between these replicates.

The heme domain of *M. thermophila* CDH-2 could not be sufficiently purified from the papain digestion of the full length protein and was thus recombinantly expressed in the yeast *Pichia pastoris*. The heme domain from CDH-2 was purified by nickel metal affinity chromatography and has the same spectral properties of the full length CDH-2 (Fig. 7). The recombinant heme domain was then tested for the ability to stimulate cellulose hydrolysis of the Δ*cdh-1* strain culture filtrate (Figure 6). Addition of the ferric heme domain at the same molar concentration as the full length CDH-2 required for maximum stimulation had no stimulatory effect. However, at a loading of 1 μM, the ferric heme domain was able to stimulate cellulase activity to nearly the same extent as the full length enzyme at 23 nM (200 μg/g cellulose) (Figure 6).

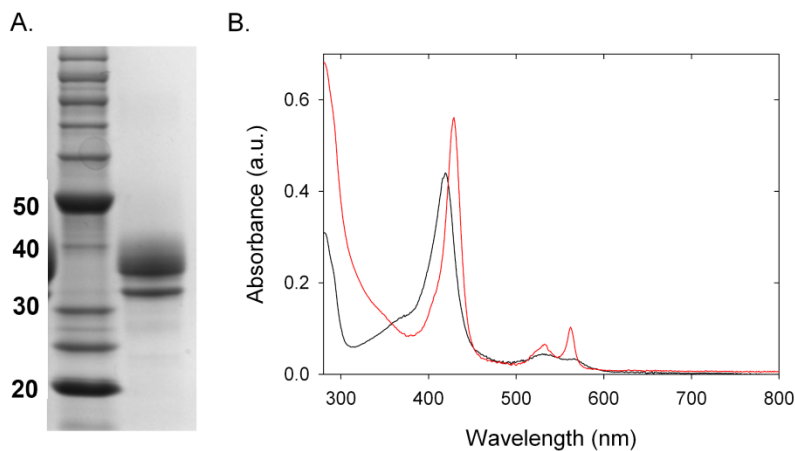

Figure 7. Purity and spectral properties of recombinant CDH-2 heme domain expressed in *Pichia pastoris*. (A) SDS-PAGE of purified recombinant CDH-2 heme domain. (B) UV-vis spectra of the oxidized (black) and reduced (red) CDH-2 heme domain.

4.3.3 Oxygen and metal ion dependence on the stimulation of cellulose degradation by CDH

A prevalent hypothesis for the biological function of CDH postulates that electrons from the heme domain of CDH are transferred to ferric complexes, quinones, molecular oxygen, or other redox mediators which lead to the production of radical species that can non-specifically degrade cellulose or lignin. Experiments were carried out to address if the stimulation of activity observed with CDH addition to the Δ*cdh-1* culture filtrate was due to a direct reaction with the cellulose or an indirect effect where metals or small molecules became reduced by CDH and subsequently contributed to the degradation.

To test for the effect of small molecules in the Δ*cdh-1* culture the culture filtrate was buffer exchanged 10,000-fold using 10,000 MWCO spin concentrators. After this treatment, CDH-1 was still able to stimulate the activity of the Δ*cdh-1* culture filtrate to the same extent. To test if there was a metal dependence for the stimulation, buffer exchanged culture filtrates from the Δ*cdh-1* cultures were incubated with 100 µM EDTA
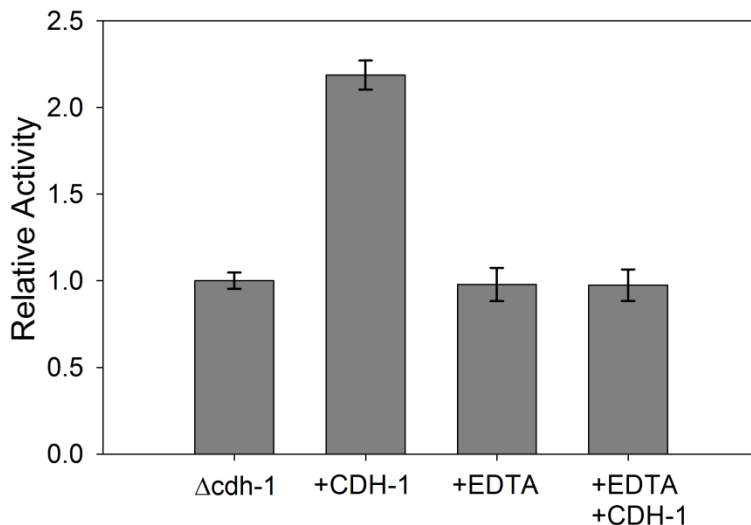

Figure 8. Contribution of metals or small molecules to the CDH dependent stimulation. (A) Cellulase activity of 10,000-fold buffer exchanged Δ*cdh-1* secretome in the presence or absence of 0.004 mg/mL CDH-1. For reactions containing EDTA, 100 µM EDTA was preincubated with the sample for 2 hours and then assayed.

for 1 hour, and then assayed for cellulase activity. EDTA had no effect on the cellulase activity of the Δ*cdh-1* culture filtrate; however, when *M. thermophila* CDH-1 was added to the EDTA treated Δ*cdh-1* culture filtrate, no stimulatory effect was observed (Fig. 8). Addition of EDTA to wild-type culture filtrate reduced cellulase activity by ~50% (data not shown). Taken together, these results suggest that there is a protein bound metal ion essential for the stimulation of cellulose degradation by CDH. Overnight incubation of *M. thermophila* CDH-1 with 1.0 mM EDTA had no effect on the oxidation of cellobiose with DCPIP or cytochrome c as electron acceptors.

Finally, the role of molecular oxygen on the stimulation of activity by CDH-1 in the Δ*cdh-1* culture filtrate was explored. Cellulase activity of the Δ*cdh-1* culture filtrates is not affected by the absence of molecular oxygen, while in wild-type culture filtrates activity is reduced by ~40%. When purified *M. thermophila* CDH-1 was added to the Δ*cdh-1* culture filtrate under anaerobic conditions no stimulatory effect on cellulase activity was observed (Fig. 9).
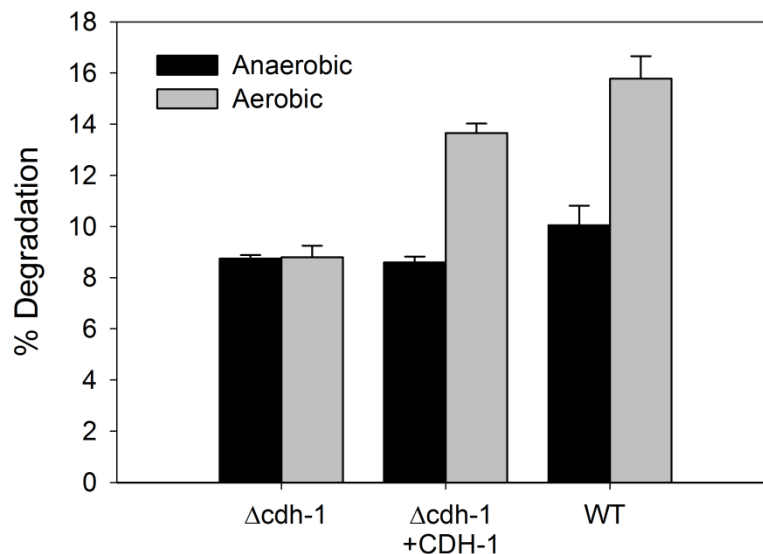

Figure 9. Oxygen dependence of the stimulation of cellulase activity by CDH. Black bars indicate experiments conducted anaerobically, gray bars indicate experiments conducted aerobically. Values are the mean of three replicates. Error bars are the SD between these replicates.

4.4 Discussion

The chemical makeup of the plant cell wall makes it a highly recalcitrant source of nutrients for animals and microbes. The structural backbone of the plant cell wall is the crystalline polysaccharide, cellulose, which typically constitutes 30-40% of the dry weight of terrestrial plants (*12*). During growth on pure cellulose, *N. crassa* increases the expression of more than 100 different genes and secretes more than 30 different proteins (*128*). Many of these proteins have high sequence homology to known hydrolytic enzymes, but there are also proteins predicted to be involved in oxidative reactions and proteins of unknown function. CDH-1 is the most abundant secreted oxidoreductase produced by *N. crassa* during growth on cellulose, constituting as much as 3% of the total secreted protein. Using a combination of genetic and biochemical techniques, we report here that CDH-1 is a major contributor to the depolymerization of cellulose by *N. crassa*. The deletion of CDH-1 lowers the aggregate cellulase activity of the culture filtrates by 37-49%. Addition of purified CDH-1 back to the Δ*cdh-1* culture filtrate stimulates the cellulase activity back to the wild-type level. A stimulatory effect was seen with an addition of CDH-1 equivalent to only 0.1% of the total protein in the culture filtrate.

In the cellulolytic Sordariales sequenced thus far, all species have two copies of CDH (*134*), one containing a C-terminal fungal cellulose binding domain (CBM), and one without a CBM. In *N. crassa,* only the CBM-containing CDH-1 is produced in significant amounts during growth on cellulose. However, in the closely related thermophilic fungus, *M. thermophila,* both CDHs are produced at high levels during growth on cellulose (*140*). Both CDHs from *M. thermophila* have been previously shown to have very similar steady-state kinetic properties with soluble substrates (*140*), and we now show that both enzymes can stimulate hydrolysis of cellulose to the same extent. A key difference between the two CDHs is that CDH-1 was functional at a 10-fold lower concentration than CDH-2. Cellulose binding modules are known to increase the local concentration of cellulases on the surface of cellulose, and the results here are consistent with the stimulatory effect being due to a reaction taking place on or near the surface of the cellulose.

The isolated and catalytically active flavin domain of CDH-2 is unable to stimulate hydrolysis of cellulose, whereas the CDH-2 heme domain was able to stimulate cellulose hydrolysis to the same extent as full length CDH-2, at a 20 to 50-fold higher concentration. The iron in the heme domain of CDH is axially ligated by absolutely conserved methionine and histidine residues. These two residues cannot be displaced by cyanide, azide when ferric, or carbon monoxide when ferrous (*147*), supporting the hypothesis that the heme domain is likely involved in outer sphere electron transfer reactions. The results presented here suggest that 1-electron transfer, mediated through the heme domain, is essential for the stimulatory effect. The stimulatory effect of the ferric heme domain was unexpected, as it would have no electron source. However, the Δcdh-1 culture solution contains a trace CDH activity, due to a very low level expression of CDH-2. Full length CDH-2 is able to reduce free CDH heme domains in the presence of cellobiose and likely accounts for the stimulation in activity by the free heme domain.

Proteins with sequence homology to the heme domain of CDH are widespread throughout the fungal kingdom. As more fungal genome sequences have become available, CDH-like heme domains have been found in proteins with diverse domain architectures. In addition to a canonical CDH, the basidiomycete, *Phanerochaete chrysosporium* also produces a secreted CDH-like heme domain with a C-terminal fungal cellulose binding domain (*148*). In the *N. crassa* genome there are three genes encoding proteins predicted to contain stand alone CDH-like heme domains with N-terminal signal peptides; however, previous transcriptional profiling of *N. crassa* did not show them upregulated during growth on cellulose. These free heme domains are conserved throughout the Sordariales with sequenced genomes and all are predicted to have the same methionine and histidine axial iron ligation.

The prevailing hypothesis for the biological function of CDH in the degradation of lignocellulose involves flavin oxidation of cellobiose with sequential electron transfer to

the heme and then heme reduction of a one electron acceptor, speculated to be an Fe(III)-oxalate complex (*137*). This ferrous iron is then proposed to take part in a Fenton reaction with peroxide generated by CDH or other extracellular oxidases to generate hydroxyl radicals which non-specifically degrade plant cell wall polysaccharides. Fenton chemistry requires ferrous iron and peroxide, and has been routinely used in DNA footprinting experiments (*149*). After a 10,000-fold buffer exchange of the Δ*cdh-1* culture filtrate, the free iron in solution is less than 1 nM; however, CDH was still able to fully stimulate cellulase activity. Treatment of the culture filtrate with EDTA completely blocks the stimulation by CDH, although EDTA should not prevent ferrous iron from participating in Fenton reactions. In fact, iron-EDTA chelates are often used to generate hydroxyl radicals (*149*). Taken together, these results do not support the involvement of Fenton chemistry in the stimulation of cellulase activity by CDH in *Neurospora*. The requirement of molecular oxygen for the stimulation in activity; however, suggests that oxidation reactions may play an important role.

Our results are consistent with the heme domain of CDH contributing to the degradation of cellulose through interaction with other metalloproteins that are larger than 10,000 daltons. Potential candidates include metal-dependent glycosyl hydrolase family 61 proteins or other metal-dependent proteins of unknown function present in the culture filtrate. GH61 family proteins were recently shown to stimulate the hydrolysis of pretreated corn stover when added to a cocktail of industrial cellulases (*150*). The stimulation by GH61 proteins is completely inhibited by addition of EDTA. This inhibition of activity by EDTA has been ascribed to the removal of an essential surface exposed divalent metal ion from the protein, and can be rescued by addition of divalent metal ions to the reaction. Interestingly, no stimulation of activity by GH61 proteins was observed on pure cellulosic substrates (*150*). CBP21, a protein structurally related to GH61 proteins, from the chitinolytic bacterium *Serratia marcescens*, was recently shown to stimulate the hydrolysis of crystalline chitin (*151, 152*), presumably by the same mechanism as GH61 proteins. CBP21 requires divalent metal ions for activity, as well as the presence of molecular oxygen (*153*). Addition of CBP21 to chitin in the presence of molecular oxygen and ascorbic acid leads to the formation of oxidized chito-oligosaccharides by an unknown mechanism. If fungal GH61 proteins stimulate hydrolysis by the same mechanism as CBP21, it is possible that CDH is functioning as an extracellular reductant. The *N. crassa* genome contains fourteen genes encoding predicted extracellular GH61 proteins, of which, at least nine are upregulated during growth on cellulose (*128*). There are also multiple hypothetical proteins secreted by *Neurospora* which could be involved in the CDH stimulated degradation pathway.

The discovery that cellobiose dehydrogenase plays a major role in pure cellulose degradation by some fungi is a significant shift from previous models centered on the degradation of cellulose by a mixture of glycoside hydrolases. The well understood

genetics and resources available in *Neurospora* should facilitate the discovery of CDH's other partners so that the full system can be reconstituted *in vitro*.

4.5 Materials and Methods

Strains. All *Neurospora crassa* strains were obtained from the Fungal Genetics Stock Center (FGSC) including wild-type *N. crassa* (FGSC 2489) and the Δ*mus-51* (FGSC 9717) used for generating gene deletions. The Δ*gh5-1* (FGSC 1647) was from the *Neurospora* functional genomics project.

Preparation of Δ*cdh-1* strains. The DNA cassette used to delete *cdh-1* was provided by the *Neurospora* functional genomics project. Details about how the cassette was generated are available online at (http://www.dartmouth.edu/~neurosporagenome/protocols.html). FGSC 9717 was grown on Vogel's minimal media (*119*) slants for 21 days at room temperature. Conidia from the FGSC 9717 slant were transformed by electroporation with 1 µg of the Δ*cdh-1* (ΔNCU00206) cassette and plated onto media containing hygromycin. Positive transformants were then crossed with wild-type *N. crassa*. Ascospores were germinated on media containing hygromycin and several hygromycin resistant transformants were harvested and screened for production of CDH. The genotypes of three transformants that showed good growth on cellulose and lacked CDH activity were confirmed by PCR using primers specific to *cdh-1* and the hygromycin resistance cassette.

Growth and harvest of *N. crassa* culture filtrates. Wild-type or Δ*cdh-1 N. crassa* was inoculated onto slants of Vogel's minimal media and grown for 3 days at 30 °C in the dark followed by 7 days at room temperature. A conidial suspension was then inoculated into 100 mL of Vogel's salts supplemented with 2 % Avicel® PH101 in a 250 mL Erlenmeyer flask. After 7 days of growth on Avicel®, cultures were filtered over a 0.2 µm polyethersulfone (PES) filter.

Protein purification. For isolation of *Myceliophthora thermophila* (ATCC 46424) CDH-1 and CDH-2, 10 day old conidia from cultures of *M. thermophila* were inoculated into 1.0 liter of Vogel's salts supplemented with 40 g/L cotton balls in 2.8 liter Fernbach flasks. After 6 days of growth at 48 °C and 200 RPM shaking, fungal biomass and residual cotton was removed by filtration using 0.2 µm PES filters. The culture filtrate was then passed over CaptoQ resin (GE Healthcare) at 20 mL/min. CDH-1 and CDH-2 were eluted from the CaptoQ resin with buffer containing 500 mM sodium chloride and 25 mM HEPES pH 7.4. The pooled CaptoQ concentrate was then desalted into 25 mM HEPES pH 7.4 using a HiPrep 26/10 desalting column (GE Healthcare). The protein was further purified on a 10/100 GL MonoQ column (GE Healthcare) and eluted with a

linear gradient from 0 to 500 mM sodium chloride. CDH-1 elutes from the column between 180 and 220 mM sodium chloride; CDH-2 elutes between 280 and 320 mM sodium chloride. CDH-1 was then further purified by binding to 1.0 grams of Avicel® in 50 mM sodium acetate buffer pH 5.0 for 5 minutes at 50 °C. The Avicel® was removed by centrifugation in 50 mL conical tubes then washed with 50 mL of acetate buffer three times. CDH-1 was eluted from the Avicel® by addition of 50 mL of 50 mM pH 11.3 phosphate buffer. The eluted CDH-1 was concentrated using 10,000 molecular weight cutoff (MWCO) PES spin concentrators and stored at -80 °C. CDH-2 was further purified by gel filtration chromatography using a 16/60 Sephacryl S200 column (GE Healthcare) with a mobile phase of 25 mM HEPES pH 7.4 and 150 mM sodium chloride. The flavin domain of CDH-2 was isolated by treatment of the full length enzyme with papain (*140*). After papain cleavage, the fragments were separated by gel filtration chromatography using a 16/60 Sephacryl S100 column (GE Healthcare) with a mobile phase of 25 mM HEPES pH 7.4 and 150 mM sodium chloride. The fractions enriched in flavin domain were then further purified using a 10/100 GL MonoQ column and eluted with a linear gradient from 0 to 500 mM sodium chloride.

N. crassa CDH-1 (NCU00206) was partially purified from wild-type *N. crassa*. Phosphoric acid swollen cellulose (PASC) was added (20 mL) to the culture filtrate at 4 °C and mixed for 10 minutes. The PASC was then removed from the culture filtrate by centrifugation for 10 minutes at 4,000 RPM. Next, 724 g/L ammonium sulfate was added to the culture filtrate and stirred at 4 °C for one hour. The precipitated proteins were then isolated by centrifugation at 8,000 RPM for 25 minutes. The pellet was resuspended in a minimal volume of water and then buffer exchanged into 25 mM Tris pH 8.5 using a HiPrep 26/10 desalting column. The buffer exchanged proteins were then loaded onto a 10/100 GL MonoQ column and eluted with a linear gradient from 0 to 500 mM sodium chloride. CDH-1 elutes from the column between 50 and 100 mM sodium chloride and is ~80% pure by SDS-PAGE. *N. crassa* CBH-1 (NCU07340), GH6-2 (NCU09680), GH5-1 (NCU00762), and GH3-4 (NCU04952) were purified as previously described (*154, 155*).

Cellulase assays. All cellulase assays were performed in triplicate with 10 mg/mL Avicel® PH101 (Sigma) in 50 mM sodium acetate pH 5.0 at 40 °C. Assays were performed in 1.7 mL microcentrifuge tubes with 1.0 mL total volume and were inverted 20 times per minute. Each assay contained 0.05 mg/mL culture filtrate or 0.05 mg/mL reconstituted cellulase mixture containing CBH-1, GH6-2, GH5-1, and GH3-4 present in a ratio of 6:2.5:1:0.5. The concentration of heme domain used in stimulatory assays was 1.0 µM as determined by absorption at 430 nm of the fully reduced protein.

Assays were centrifuged for two minutes at 4000 rpm to pellet the remaining Avicel® and 20 µL of assay mix was removed per well. Samples were incubated with 100 µL of desalted, diluted Novozyme 188 (Sigma) at 40 °C for 20 minutes to hydrolyze

cellobiose and then 10-30 µL of the Novozyme 188 treated cellulase assay supernatant was analyzed for glucose using the glucose oxidase/peroxidase assay as described previously (*128*). Percent degradation was calculated based on the amount of glucose measured relative to the maximum theoretical conversion of 10 mg/mL Avicel®.

Metal chelating reaction. To determine the requirement for small molecules and metals, wild-type or Δ*cdh-1* culture filtrate was buffer exchanged more than 10,000 fold in a 10 kDa MWCO spin concentrator (Sartorius). The culture filtrate was then incubated with 100 µM EDTA for 2 hours. EDTA or metal reconstituted culture filtrate was then assayed as described above.

CDH binding to Avicel®. Purified CDH-1 or CDH-2 (20 µg) was added to 1.0 mL of 10 mg/mL Avicel® in 50 mM sodium acetate buffer pH 5.0. After incubation for 5 minutes at room temperature, the Avicel® was spun down by centrifugation at 10,000 RPM. The supernatant was removed and the Avicel® pellet was washed with 1.0 mL of fresh 50 mM sodium acetate buffer pH 5.0. The washing step was repeated three times and then the Avicel® pellet was boiled in SDS loading buffer. After boiling, the Avicel® was spun down and the supernatant loaded onto a 10% Tris-HCl polyacrylamide gel.

Anaerobic assays. Anaerobic cellulase assays were performed as above except all assays were conducted in an anaerobic chamber (Coy) at room temperature. Buffers were sparged with nitrogen for 1 hour and culture filtrates were concentrated more than 20-fold to volumes of less than 300 µL before introduction into the anaerobic chamber. All solutions were left open in the anaerobic chamber, and buffer solutions were stirred vigorously for 72 hours before use to fully remove dissolved oxygen. Aerobic reactions were prepared in the anaerobic chamber in 3 mL reactivials and then removed from the anaerobic chamber, exposed to air, sealed, and returned to the anaerobic chamber. At specified timepoints, assays were centrifuged in the glove bag and 100 µL of assay mix was removed and analyzed by the glucose-oxidase peroxidase assay as described above.

4.6 Acknowledgements

*Footnote to chapter 4

This chapter, presented here with modifications from its original format, represents part of a peer-reviewed paper published in *ACS Chemical Biology* (2011), In press. The authors are Christopher M. Phillips, William T. Beeson, Jamie H.D. Cate and Michael A. Marletta.

Chapter 5: Identification of GH61 proteins as an integral part of the of the CDH-dependent enhancement of cellulose degradation

5.1 Abstract

Gene deletion and complementation studies demonstrated that cellobiose dehydrogenase enhances the cellulose degrading capabilities of *N. crassa* by approximately 1.6-2.0 folds. This stimulatory effect is dependent on molecular oxygen and other, unknown factors. Here, using a combination of protein purification techniques, mass spectrometry, and a novel activity screen we identify the other factors involved in the stimulation of cellulose hydrolysis by CDH as glycosyl hydrolase family 61 proteins. Taking advantage of the metal binding properties of GH61 proteins, we developed a purification strategy to isolate them free of cellulase contaminants. When added to a defined mixture of cellulases from *N. crassa*, GH61 proteins were found to enhance cellulose degradation.

5.2 Introduction

Recent transcriptomic and proteomic analyses of cellulolytic fungi have identified oxidative enzymes involved in degradation of plant biomass (*127, 156, 157*). Despite the widespread occurrence of these enzymes in fungi, the specific function of these oxidative enzymes in cellulose degradation is unknown (Fig. 1). The filamentous ascomycete, *Neurospora crassa,* is a proficient degrader of plant cell walls. During growth on cellulose *N. crassa* produces cellobiose dehydrogenase (CDH) and many other enzymes(*43, 158, 159*). CDH catalyzes the oxidation of cellobiose or longer cellodextrins to the corresponding 1-5-δ-lactones (*160*).
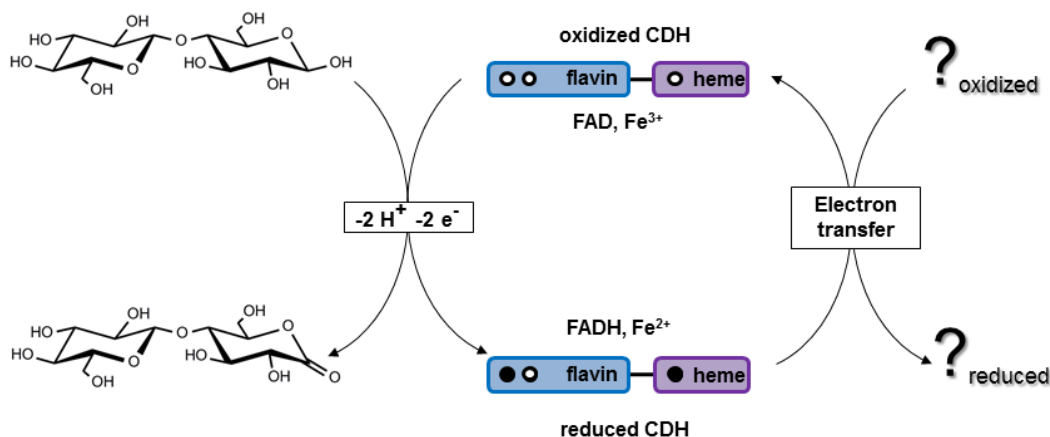

**Figure 1.** Schematic of CDH catalyzed reaction.

During turnover electrons are transferred from the cellodextrin to the CDH flavin domain and then a single electron is transferred to the heme domain (Figure 1). The biologically relevant terminal electron acceptor for CDH is unknown. Molecular oxygen is a poor substrate for CDH, however it reacts rapidly with quinones, metal ions, or organic dyes. A prominent hypothesis for CDH involves the generation of hydroxyl radicals through reduction of extracellular ferric complexes. Once reduced, the ferrous iron takes part in Fenton chemistry to generate hydroxyl radicals that randomly depolymerize plant cell wall polysaccharides.

The biological function of CDH is unknown; however, recent studies in *N. crassa* suggest that it plays a role in cellulose degradation. CDH shows no beneficial effect when added to a mixture of pure cellulases, but when added to a *N. crassa* secretome deficient in CDH activity, enhances cellulase activity by 1.6-2.0 folds. The enhancement of cellulose degradation is not due to Fenton chemistry, but does require other unknown factors, hypothesized to be metalloproteins. In this report we use a CDH from *M. thermophila* and a fractionation strategy to identify a novel family of copper dependent enzymes in *N. crassa* that are necessary for CDH dependent enhancement of cellulose degradation.

5.3 Results

5.3.1 Identification of GH61 proteins as part of the CDH dependent cellulose degradation pathway

As reported in chapter 4, addition of CDH to a mixture of purified cellulases showed no enhancement of activity, suggesting that additional factor(s) were required. The strong inhibition by EDTA noted in chapter 4, is consistent with the hypothesis that these factors were metalloproteins. Based on this information, an approach was developed to identify the other factors in the *N. crassa* secretome involved in CDH dependent enhancement of cellulose degradation. The secretome of the *Δcdh-1* strain grown on Avicel for 7 days was fractionated using a 10/100 GL MonoQ column. The flow through, and fractions eluting from the column were then added to a core mixture of *N. crassa* cellulases, including CBH-1 (NCU07340), GH6-2, a CBH-2; GH5-1, an endoglucanase; and GH3-4, a β-glucosidase. If the factor was present in the fraction, then addition of CDH would increase Avicel hydrolysis, and addition of EDTA would block the stimulation. A typical fractionation is summarized in Figure 2.
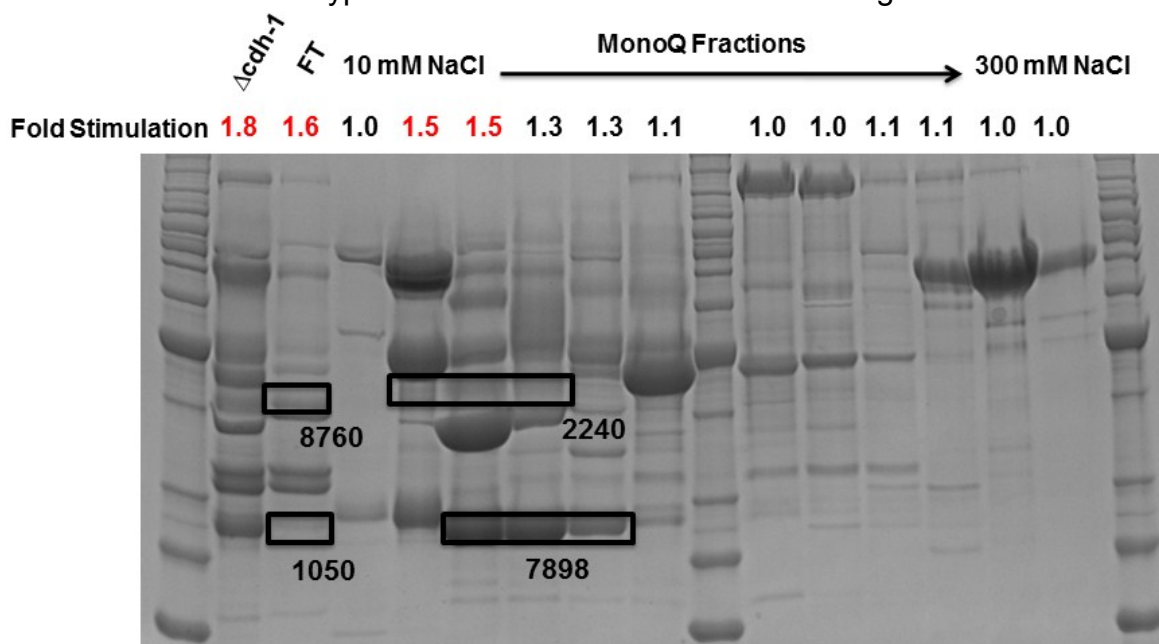
$$\text{Fold stimulation} = \frac{\text{Activity(Core + Fraction + CDH-1)}}{\text{Activity(Core + Fraction)}}$$

Figure 2. Fractionation of the *Δcdh-1* secretome to identify other factors involved in CDH dependent enhancement of cellulase activity. SDS-PAGE analysis of fractions from MonoQ separation of the *Δcdh-1* secretome. Above each lane is the fold stimulation of the fraction when added to a mixture of purified *N. crassa* cellulases in the presence or absence of CDH-1. Fractions highlighted in red were analyzed by mass spectrometry and showed the presence of four glycosyl hydrolase family 61 proteins (indicated by black boxes).

From the initial fractionation, it was clear that the flow-through fractions and fractions eluting from the column between 3-8 mS/cm contained the factors required for stimulation of hydrolysis by CDH. Since the stimulatory activity was present in multiple fractions, it was likely that multiple proteins could potentiate the CDH-dependent stimulation of cellulose hydrolysis. The proteins present in all of the fractions from the MonoQ column were identified by in-gel and solution tryptic digests coupled with LC-MS/MS. A clear trend that emerged from this analysis was that in fractions where there was CDH-dependent stimulation of cellulase activity, there were also GH61 proteins (Figure 2). The GH61 proteins identified in the fractions showing stimulatory activity were NCU01050, NCU02240, NCU07898, and NCU08760. Each of the fractions with stimulatory activity; however, contained at least two GH61 proteins, so further purification was required to confirm if a single GH61 protein could potentiate cellulose degradation.

5.3.2 Purification strategy to isolate high purity native GH61 proteins from *N. crassa*

After the initial MonoQ fractionation, the protein was less than 10% pure as judged by SDS-PAGE. Previous reports on the x-ray crystal structure of GH61 proteins had revealed an absolutely conserved divalent metal binding site on the surface of the protein. We hypothesized that the affinity of GH61 proteins for the MonoQ column, which binds proteins based on the amount of surface negative charge, would change if the protein was in apo or fully metal bound forms. To test this hypothesis, the MonoQ flow-through fraction, which contained NCU08760 and NCU01050, was treated with 1.0 mM EDTA for 12 hours and then reloaded onto the column. After EDTA treatment, the majority of the proteins still flowed through the column; however, on the application of a slight salt gradient two well defined peaks eluted between 1 and 3 mS/cm (Figure 3).
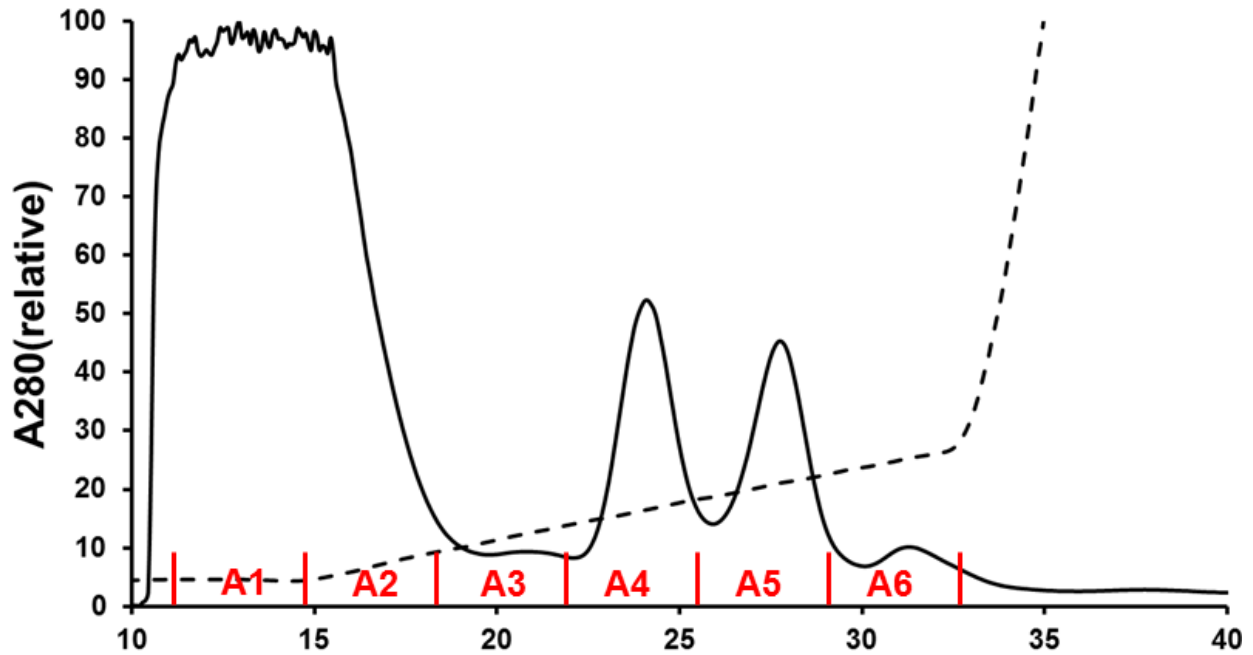
Figure 3. Fractionation of EDTA treated *N. crassa* MonoQ flow-through. Peaks indicated with arrows contain NCU08760 and NCU01050.

Figure 4. SDS-PAGE and lane guide for MonoQ fractionation of EDTA treated original MonoQ flow through. Comparing the load to the new flow through, it is clear that the bands between 25-30 kDa and between 40-50 kDa are dramatically decreased or absent in the new flow through and then elute with a shallow salt gradient. NCU01050 is predominantly in fraction A4 and NCU08760 is in fraction A5.

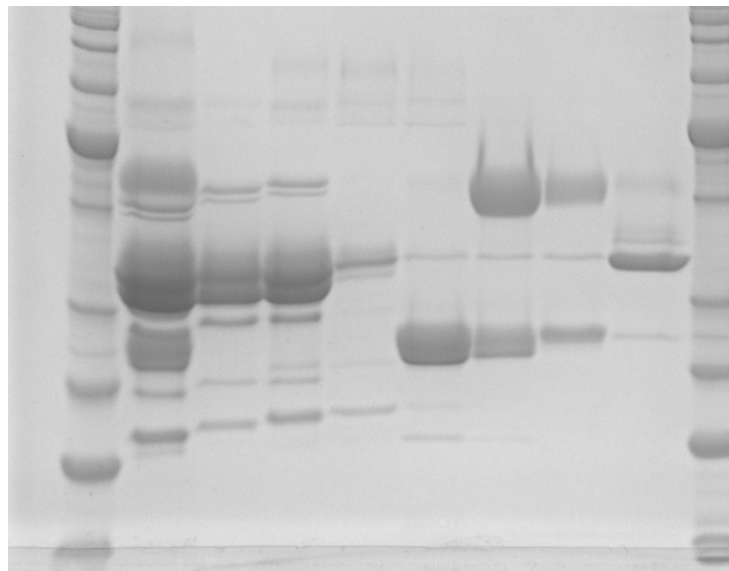SDS-PAGE analysis of the fractions is shown in Figure 4. After EDTA treatment, the MonoQ flow-through is clearly deficient in two proteins. One migrates between 40-50 kDa and the other between 25-30 kDa. The 25-30 kDa protein elutes from the column in the first sharp peak, and the 40-50 kDa protein in the second sharp peak. Tryptic digests and LC-MS/MS analysis of the proteins in these two fractions confirmed the presence of NCU01050 and NCU08760. These results are consistent with removal of positive charge from the surface of the protein causing them to bind the column more tightly. The same approach was used to separate NCU02240 and NCU07898 from other proteins in the secretome (see Materials and Methods). After that native metal is stripped from GH61 proteins, the metal site can be reloaded with exogenous divalent metal ions including zinc and cobalt. The reconstituted protein can again be fractionated on a MonoQ column to increase purity. After the GH61 proteins are separated from all the other proteins in the secretome, a gel filtration column can be used to separate GH61 proteins with cellulose binding modules from GH61 proteins without cellulose binding modules. The four *N. crassa* GH61 proteins are more than 98% pure based on SDS-PAGE analysis (Figure 5)
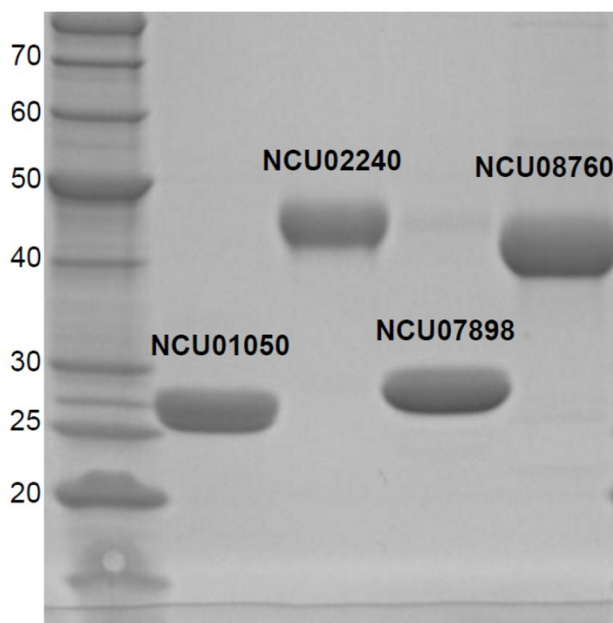
Figure 5. SDS-PAGE of natively purified GH61 proteins.

5.3.3 Metal analysis of the secretome and natively purified GH61 proteins

Previous studies have reported that GH61 proteins can bind to a variety of divalent metal ions. We wanted to determine the native metal present in GH61 proteins produced by *N. crassa*. After extensive dialysis, the metal content of the *Δcdh-1* secretome was analyzed using inductively coupled plasma atomic emission spectroscopy (ICP-AES). Only copper and zinc were present in the secretome, with copper 10-fold more abundant than zinc (Figure 6). However, no *N. crassa* proteins

secreted during growth on cellulose are predicted to bind copper. Metal analysis of the fractions from the 10/100 GL MonoQ column showed that fractions containing GH61 proteins were also enriched in copper (Figure 7). Three of the natively purified GH61 proteins were analyzed by ICP-AES and found to bind copper with a 1:1 stoichiometry (Figure 8). NCU02240 was not present in sufficient purity or yield for further characterization.

5.3.4 Stimulation of Avicel and pretreated corn stover hydrolysis by purified GH61 proteins.

Next, we studied the ability of the natively purified GH61 proteins to stimulate the hydrolysis of Avicel or pretreated corn stover by a defined mixture of *N. crassa* cellulases (Figure 9). In the absence of CDH, there was only a 1.1-1.2 fold enhancement of cellulase activity on Avicel by each of the GH61 proteins tested. When CDH was added to the core cellulases and NCU01050 or NCU02240 there was a substantial increase in hydrolytic activity. However, for NCU07898 and NCU08760, there was no stimulatory effect on cellulase activity.

Degradation of pretreated corn stover was enhanced by the GH61 proteins to a greater extent in the absence of CDH (1.3 to 1.5 fold) than Avicel hydrolysis was (Figure 9b). When CDH was added to the core cellulase plus GH61 mixture, a further enhancement in cellulase activity was observed. On pretreated corn stover, a stimulatory effect on hydrolysis was observed for all four GH61 proteins tested. NCU08760 was especially stimulatory, with an overall stimulation in the presence of CDH of more than 2 folds.
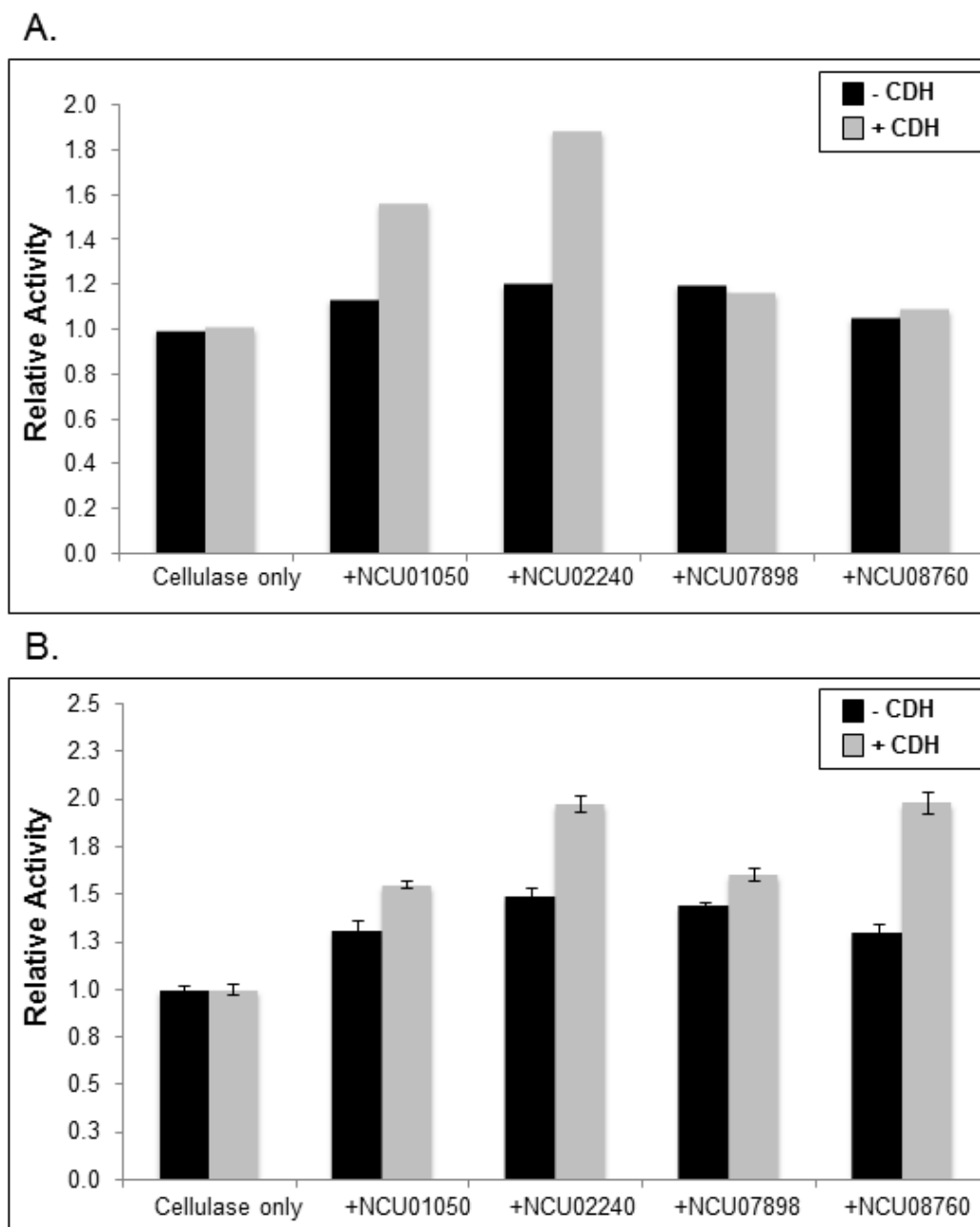
Figure 9. Stimulation of (A) Avicel or (B) pretreated corn stover hydrolysis by natively purified GH61 proteins. GH61 proteins (0.015 mg/mL) were added to *N. crassa* cellulases (0.05 mg/mL CBH-1, GH6-2, and the ß-glucosidase GH3-4 (0.005mg/mL) in the presence (black bars) or absence (gray bars) of MtCDH-1 (0.004 mg/mL).

5.4 Discussion

During growth on cellulose many different proteins are secreted by *N. crassa*. CDH-1 is the major known oxidoreductase secreted by *N. crassa* and constitutes approximately 2-3% of the secretome by weight. Despite only making up 2-3% of the secretome, deletion of CDH-1 reduced aggregate cellulase activity by approximately 50%. Here we show that CDH-1 dependent enhancement of cellulose degradation also requires the presence of GH61 proteins. GH61 proteins constitute 15-20% of the *N. crassa* secretome and their biological function is unknown.
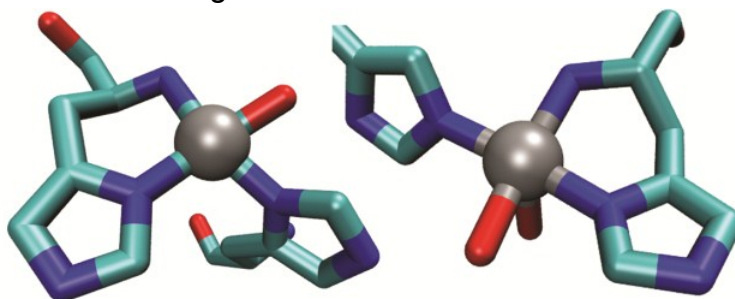

Figure 10. Comparison of the metal binding sites in CopC (left) and GH61E from *Thielavia terrestris* (right). Copper is bound in the CopC structure and zinc is bound in GH61E.

GH61 proteins have been studied by x-ray crystallography and contain a highly conserved metal binding site that includes two histidine residues. One of the histidines is the N-terminal residue and functions as a bidentate ligand involving the amine and ring N2. Crystal structures of GH61 proteins with nickel (*161*), magnesium, or zinc (*162*), have been reported. Redox chemistry involving O_2 generally requires a transition metal. Copper was found to be the metal natively bound to GH61 proteins from *N. crassa* and is found in many enzymes that activate molecular oxygen. These results suggest that copper-bound GH61 proteins may accept electrons from the reduced CDH-heme domain. Hence, the copper in GH61 proteins could function as redox couple with CDH. Two other proteins that contain bidentate N-terminal histidine ligands were found in the protein databank. One of them, CopC, is part of the cop operon involved in copper resistance in bacteria and has been shown to bind Cu(II) with picomolar affinity (Figure 10) (*163*). The other, particulate methane monooxygenase (pMMO), uses a trinuclear copper site to oxidize methane (*164*). Combining our experimental results with the similar ligation in CopC and pMMO, we conclude that copper is the native metal in GH61 proteins.

	[836]	[1050]	[2240]	[2344]	[2916]	[3328]	[5969]	[7760]	[7898]	[8760]
[836]	100	30	42	38	41	39	26	26	31	38
[1050]	30	100	57	39	46	34	25	28	34	40
[2240]	42	57	100	37	48	33	26	25	30	41
[2344]	38	39	37	100	36	50	21	27	33	30
[2916]	41	46	48	36	100	27	32	35	34	40
[3328]	39	34	33	50	27	100	26	25	28	36
[5969]	26	25	26	21	32	26	100	64	38	25
[7760]	26	28	25	27	35	25	64	100	38	29
[7898]	31	34	30	33	34	28	38	38	100	34
[8760]	38	40	41	30	40	36	25	29	34	100

Figure 11. Identity matrix for the ten GH61 proteins upregulated two folds or more during growth on cellulose or lignocellulose.

The *N. crassa* genome contains fourteen genes encoding proteins predicted to contain GH61 domains. Of these fourteen genes, ten of them were upregulated 2 fold or more during growth on cellulose or lignocellulose. Six of these predicted GH61 proteins also contain C-terminal fungal cellulose binding domains (CBM1), suggesting that they directly interact with crystalline cellulose. There is a high amount of sequence diversity in the *N. crassa* GH61 proteins, with most sharing approximately 30% pairwise sequence identity. When incubated with CDH and a mixture of core cellulases, Avicel hydrolysis was enhanced by GH61 proteins NCU01050 and NCU02240, while NCU07898 and NCU08760 showed little or no beneficial effects. Both NCU07898 and NCU08760 are probably functional, because they did show a stimulatory effect on pretreated corn stover hydrolysis. Differences in the activities of GH61 proteins on different cellulosic materials might be due to sequence variation on the surface of the protein.

5.5 Materials and Methods

Growth of *N. crassa*. The Δ*cdh-1* strain of *N. crassa* was inoculated onto slants of Vogel's minimal media and grown for 3 days at 30 °C in the dark followed by 7 days at room temperature with ambient lighting. A conidial suspension was then inoculated into 100 mL of Vogel's salts supplemented with 2% Avicel® PH101 (Sigma) in a 250 mL Erlenmeyer flask. After 7 days of growth on Avicel®, cultures were filtered over 0.2 μm polyethersulfone (PES) filters.

Purification of GH61 proteins. *N. crassa* GH61 proteins were partially purified from the *N. crassa* Δ*cdh-1* strain. Culture filtrates from multiple flasks were pooled and concentrated 100-fold using a tangential flow filtration system with a 5,000 MWCO PES membrane (Millipore, Billerica, MA). The concentrated culture filtrate was then buffer

exchanged into 10 mM Tris pH 8.5 using a HiPrep 26/10 desalting column. The concentrated and buffer exchanged protein was then fractionated using an AKTAexplorer FPLC system (GE Healthcare) and a 10/100 GL MonoQ column. The mobile phases for the anion-exchange fractionation were buffer A: 10 mM Tris pH 8.5 and buffer B: 10 mM Tris pH 8.5 with 1.0 M sodium chloride. For each run on the MonoQ column approximately 100 mg of total secretome protein was loaded and eluted from the column with a linear gradient from 0 to 50% buffer B over six column volumes. NCU08760 and NCU01050 do not bind the column under these conditions and are present in the flow through. NCU02240 and NCU07898 elute from the column between 4-9 mS/cm. Fractions containing the target proteins were pooled and treated with 1.0 mM EDTA overnight to strip bound metals. The EDTA treated samples were then concentrated using 3,000 MWCO PES spin concentrators and desalted into 10 mM Tris pH 8.5 using a 26/10 desalting column. Removal of the bound metal ion increases the affinity of the GH61 proteins for the anion-exchange resin and causes them to elute at higher salt concentrations. For NCU08760 and NCU01050, removal of bound metal ion causes the proteins to bind to the MonoQ column. The apo NCU01050 and NCU08760 are then eluted from the column using the same buffers as above with a linear gradient from 0 to 3.5% buffer B over 3 column volumes. NCU01050 elutes at ~1.6 mS/cm and NCU08760 at ~2.2 mS/cm conductivity. The apo forms of NCU02240 and NCU07898 elute from the MonoQ column between 7-9 mS/cm. The purity of the PMOs can be further improved by reconstituting the apo GH61 proteins with Zn(II) and repeating the anion-exchange fractionation. Other proteins in the secretome do not change their affinity for the resin in the presence or absence of metal ions and can be effectively removed from GH61 proteins by cycles of stripping the metal, fractionating, reconstituting, and fractionating again. NCU01050 and NCU08760 can then be separated from one another by size exclusion chromatography using a Sephacryl S100 column with a mobile phase of 10 mM Tris pH 8.5 with 150 mM sodium chloride. NCU02240 and NCU07898 are separated from one another using the same method as used for NCU01050 and NCU08760. The four natively purified PMOs from *N. crassa* are stable for several months at 4 °C.

Analysis of copper and zinc in the secretome by ICP-AES. The *N. crassa Δcdh-1* strain was grown on 2% w/v Avicel® as previously described with an additional 5 µM copper sulfate, and 30 µM zinc sulfate supplemented in the minimal media. The culture filtrate was concentrated using tangential flow filtration and buffer exchanged into 10 mM Tris pH 8.5. The concentrated and buffer exchanged culture filtrate was loaded onto a 10/100 GL MonoQ column and separated into 5 fractions with a linear salt gradient. Each fraction was then analyzed for the presence of copper or zinc. Metal analysis was performed using a Perkin Elmer inductively coupled plasma atomic emission spectrometer (ICP-AES). and nm were The wavelengths used for

quantification for copper were 327.393 or 324.752 nm and for zinc, 213.857 or 206.200 nm. Add a statement here saying what other metals we looked for and that they were present at less than 1 uM in our mixture of concentrated proteins.

Copper stoichiometry of apo-PMOs.   Apo-PMO stocks of NCU01050, NCU07898, and NCU08760 were diluted to a final concentration of 1.0 mg/mL in 10 mM Tris pH 8.5 buffer and incubated with 200 μM copper sulfate at room temperature for 16 hours. After reconstitution, the protein was diluted 5-fold into 10 mM Tris pH 8.5 and desalted using a 26/10 desalting column to remove unbound copper. The desalted protein was concentrated to a final volume of 2.5 mL using 3,000 MWCO PES spin concentrators. The absorption at 280 nm was recorded for each sample and used to determine the concentration of the protein. The concentration of copper in the sample was measured using a Perkin Elmer 7000 series ICP-AES. The wavelengths used for copper quantification were 327.393 and 324.752 nm.

5.6 Acknowledgements

*Footnote to chapter 5

This chapter, presented here with modifications from its original format, represents part of a peer-reviewed paper published in *ACS Chemical Biology* (2011), In press. The authors are Christopher M. Phillips, William T. Beeson, Jamie H.D. Cate and Michael A. Marletta.

Chapter 6: Oxidative cleavage of cellulose by polysaccharide monooxygenases

6.1 Abstract

Glycosyl hydrolase family 61 (GH61) proteins were identified as essential to the cellobiose dehydrogenase (CDH) dependent stimulation of cellulase activity. Here, the mechanism of action of these proteins is elucidated. GH61 proteins were found to oxidatively cleave cellulose in the presence of oxygen and CDH or chemical reductants. The products of the reaction included oligosaccharides oxidized at the C1 or C4 position, depending on the GH61 examined. Isotope labeling experiments showed that an oxygen atom from molecular oxygen was inserted into the oxidized oligosaccharides. Based on these results a potential chemical mechanism for the enzymes is proposed. Our data supports that GH61 proteins should be renamed polysaccharide monoyxgenases.

6.2 Introduction

In chapter 5, GH61 proteins were found to be copper metalloproteins that were necessary for CDH dependent enhancement of cellulase activity. The mechanism by which the GH61 protein was enhancing cellulose degradation is unknown. Many copper metalloproteins are involved in the binding or activation of molecular oxygen. The CDH dependent enhancement of cellulase activity was found in chapter 3 to be oxygen dependent. Taken together, it is likely that the GH61 protein is binding to molecular oxygen and using it to enhance cellulose degradation.

Recently, CBP21, a protein with structural homology to GH61 proteins (less than 10% sequence identity), was identified as an oxidative enzyme important for chitin degradation in the bacterium *Serratia marcescens*. When CBP21 was incubated with crystalline chitin and a reductant, a series of oxidized chito-oligosaccharides were released from the chitin. These oligosaccharides were found to be oxidized at the C1 position, forming aldonic acids. A metal and oxygen dependence was also reported for the formation of oxidized products. All of these properties are similar to what we had observed for the CDH dependent enhancement of cellulose degradation in *N. crassa*. In the *N. crassa* system, CDH could be acting as the reductant, substituting for ascorbic acid in the bacterial chitin degrading system. If GH61 proteins were making cellodextrins oxidized at the C1 position, they may have been missed because CDH generates the same product.

Here we use a combination of analytical techniques to better understand the mechanism by which GH61 proteins and CDH enhance fungal cellulose degradation. In contrast to the widely accepted Fenton model, which involves random attack of hydroxyl radicals on the surface of cellulose, we find that CDH and GH61 proteins work together to directly oxidize internal regions of the cellulose chains. This direct oxidative cleavage probably enhances cellulose degradation by making new sites for exoglucanases to act.

6.3 Results

6.3.1 Product analysis of GH61 proteins.

The phylogenetic diversity of 10 *N. crassa* GH61s whose transcripts are upregulated during growth on cellulose (*43*) suggests that these enzymes may target a wide array of substrates in lignocellulose, or generate different products (Fig. 2). To investigate the reaction products of the purified GH61s, assays were performed on phosphoric acid swollen cellulose (PASC). When PASC was treated with GH61 and CDH, a series of aldonic acids two to nine glucose residues in length (A2-A9) were identified by high performance anion exchange chromatography (HPAEC). In addition to aldonic acids, the combination of CDH and GH61s NCU01050 or NCU07898 produced peaks at a later retention time (Fig. 3A). Product analysis by liquid chromatography-mass spectrometry confirmed the presence of aldonic acids (Gx+15

a.m.u.), as well as masses of Gx + 13 a.m.u. and Gx + 31 a.m.u (Fig. 3B). The Gx +13 mass is consistent with a doubly oxidized cellodextrin. Cellulose cleavage by these GH61s likely results in oxidation at the non-reducing end followed by oxidation at the reducing end by CDH. Given the necessity to cleave a 1,4-glycosidic bond, these products are likely oligosaccharides with a 4-keto sugar at the non-reducing end. The Gx + 31 mass is consistent with the hydrate of this product, a ketal. Ketoaldoses are unstable in aqueous solution and are known to decompose spontaneously into many different species (*165*). The third purified GH61, NCU08760, did not form the late eluting peak on the HPAEC, or Gx + 13 and Gx + 31 species (Fig. 3A and 3B), consistent with oxidation exclusively at the reducing end on C1 to form aldonic acids. Incubation of PASC with GH61 alone led to the formation of low amounts of hydrolytic products. The formation of hydrolytic products could be due to low levels of cellulase contamination.
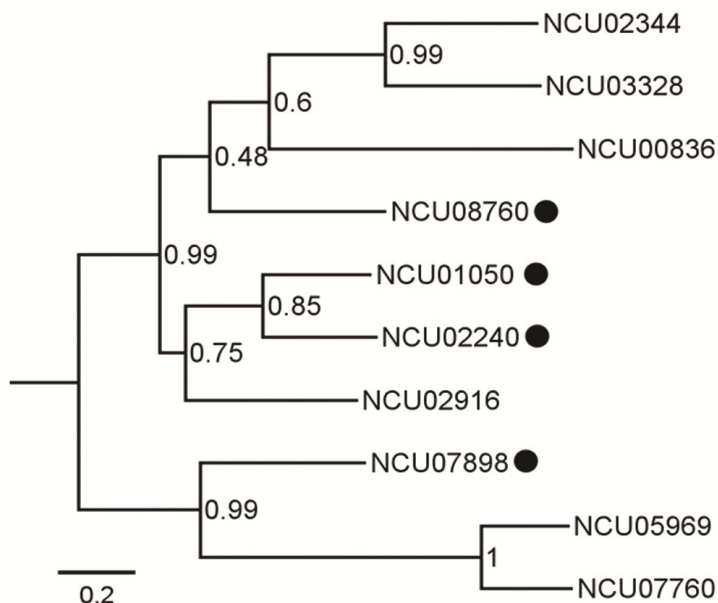

Figure 2. Phylogeny of GH61 proteins in *N. crassa.* Shown are the upregulated GH61 proteins. Of the 14 GH61 proteins in the *N. crassa* genome, 10 are upregulated >2-fold in response to growth on cellulose relative to sucrose (*43*). Proteins identified via proteomics and purified here are marked (●).

Since CDH is known to oxidize the C1 position of cellodextrins, a reaction with NCU08760 was carried out with ascorbic acid substituted for CDH. In the presence of ascorbic acid and copper, NCU08760 produced a ladder of aldonic acids (Fig. 4). Under identical conditions, NCU01050 produced a ladder of products with a later retention time when analyzed by HPAEC and 100-fold less aldonic acid than NCU08760 (Fig. 4b). The formation of oxidized sugars by the GH61s was also oxygen dependent, suggesting that GH61s are oxidases (data not shown).
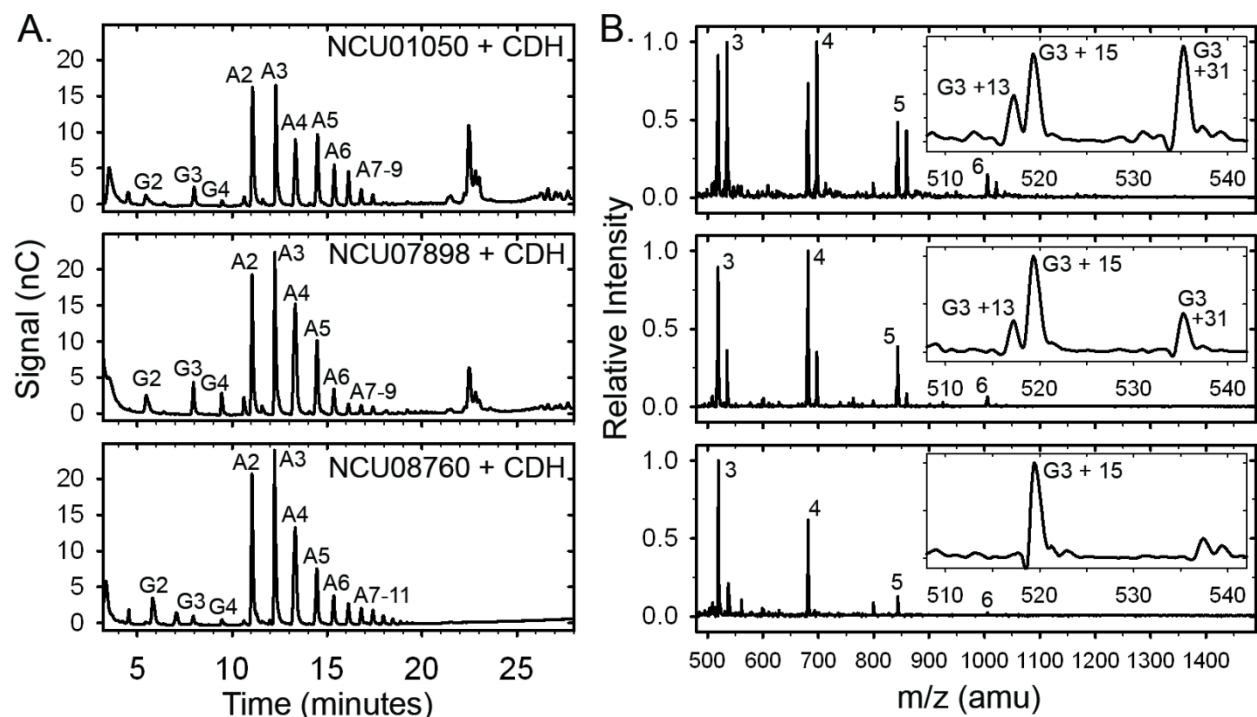
Figure 3. Reaction products of cellulose cleavage by CDH and GH61. GH61 (5.0 μM) and 0.5 μM CDH-2 were incubated with 5 mg/mL phosphoric acid swollen cellulose for 2 hours in 10 mM ammonium acetate pH 5.0 at 40°C. Products of CDH-2 in combination with each GH61 protein (NCU01050-top, NCU07898-middle), or NCU08760-bottom)) were analyzed by HPAEC (A) and LC-MS (B). HPAEC standards were used to determine retention times of cellodextrins (G1-G6) and the respective aldonic acid derivatives (A2-A9). LC-MS of reaction mixtures in negative ion mode shows reaction mixtures comprised of masses consistent with a ladder of aldonic acids (Gx + 15 a.m.u.) separated by an anhydroglucose unit. Inset is a zoom around the G3 series. NCU01050 and NCU07898 show additional masses consistent with a keto-aldonic acid (Gx + 13 a.m.u.) or the hydrate of that product, a ketal-aldonic acid (Gx + 31 a.m.u.).
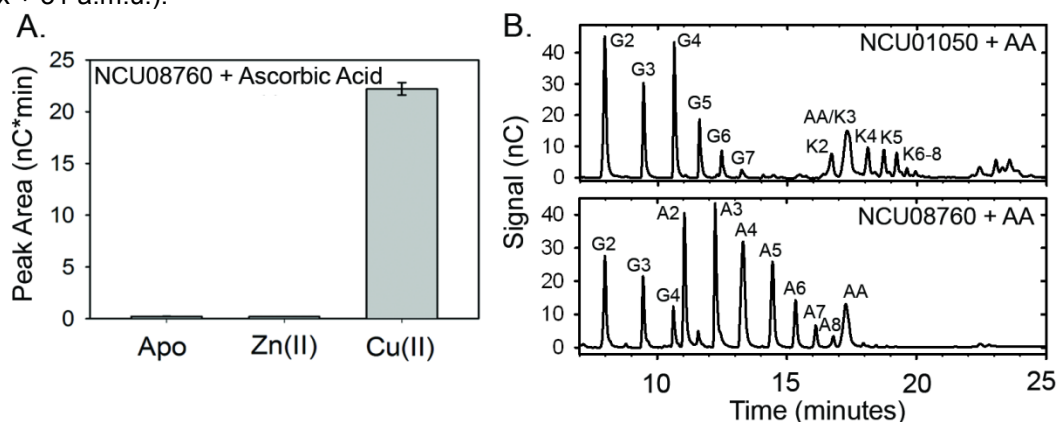

Figure 4. (A) Incubation of 5.0 μM apo-, Zn-bound, or Cu-bound NCU08760 with 2 mM ascorbic acid confirms that Cu-bound GH61 is required for generation of oxidative products. Error bars represent standard deviation of assays performed in triplicate (B) Ascorbic acid (2 mM) and 5.0 μM GH61 were assayed on PASC and analyzed by HPAEC. Products of NCU01050 include a ladder of cello-oligosaccharides (G2-G7) and a ladder of later eluting products that are likely oxidized at the non-reducing end to generate 4-keto sugars (K2-8). Products of NCU08760 include a ladder of cello-oligosaccharides (G3-G6) and aldonic acids (A2-A7). Trace amounts of aldonic acids were produced by NCU01050 but these are 100-fold less abundant than those produced by NCU08760 suggesting different regiospecificities for the 2 proteins. AA designates the peak due to ascorbic acid.

86

Next experiments were designed to determine if oxygen from molecular oxygen was incorporated into the oxidized cellodextrins. To show oxygen insertion from O_2, reactions were prepared anaerobically that contained NCU08760, ascorbic acid, and 5.0 mg/mL phosphoric acid swollen cellulose (PASC). The solutions were sealed in 1.0 mL vials and then the headspace was replaced with $^{18}O_2$ on a Schlenk line or opened to atmospheric oxygen. Cellulase and our previous GH61 reactions were typically performed at pH 5.0. A more basic pH was chosen here in order to limit the exchange of bulk water into the products. Figure 5 shows the mass spectra confirming incorporation of 1 atom of ^{18}O into the aldonic acid products, providing evidence that the enzyme functions as a monooxygenase. Similar experiments were performed where the ascorbic acid was substituted with CDH (Figure 6). ^{18}O incorporation into the products was also observed at both pH 5.0 and pH 7.8. NCU01050 did not produce products labeled with ^{18}O; however, the predicted 4-keto-aldose products are present as hydrates in aqueous solutions and would rapidly exchange with bulk water.

The identity of the non-reducing end product generated by the NCU01050 was then investigated. Direct chemical hydrolysis, using trifluoro-acetic acid (TFA), of the NCU01050 and CDH reaction products generated three peaks on a Dionex PA-200 column (Figure 7). As expected, glucose and gluconic acid are formed from the internal carbohydrates and the oxidation of the reducing end by CDH. The third peak, eluting at 6.4 minutes, is absent in control reactions with NCU08760 and CDH and probably is the GH61-oxidized non-reducing end carbohydrate. If a ketone functional group was introduced in the ring, reduction with sodium borohydride would convert it back to a racemic mixture of alcohols at that position (Figure 8a). Galactose is the C4 epimer of glucose and can be readily separated from glucose using a Dionex PA-20 column. Initial reaction products from the NCU01050 reaction were treated with sodium borohydride. Analysis of the monosaccharide products on a Dionex HPAEC confirmed the formation of galactose (Figure 8b). In control experiments where NCU08760 was substituted for NCU01050, 10-fold less galactose was formed (data not shown).
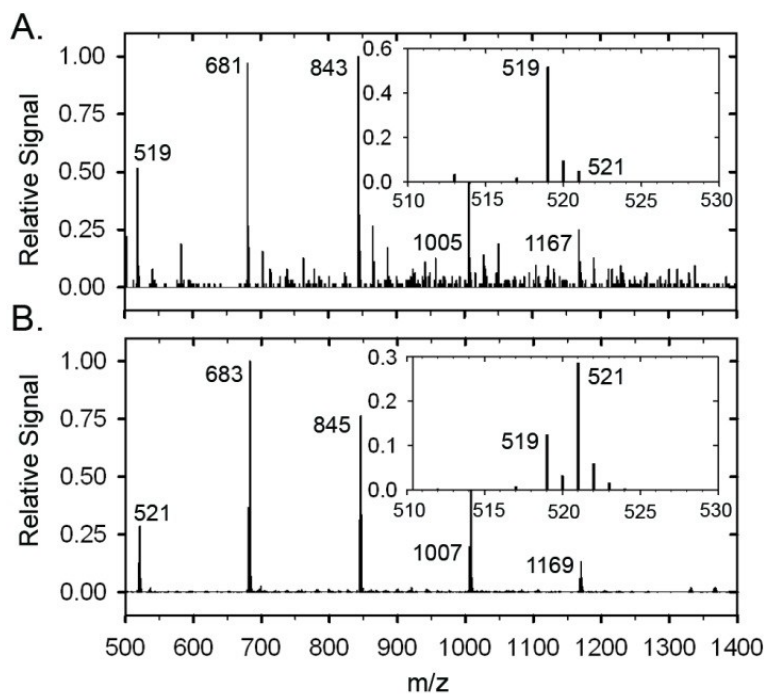
87

Figure 5. Mass spectra illustrating the incorporation of ^{18}O from $^{18}O_2$ into the aldonic acid products of NCU08760 with ascorbic acid as the reductant. Assays were performed with $^{16}O_2$ (A) or $^{18}O_2$ (B) and the products were analyzed by LC-MS. Shown is a ladder of aldonic acid products (DP 3-7). The inset is an expanded region around the mass of the cellotrionic acid ion.

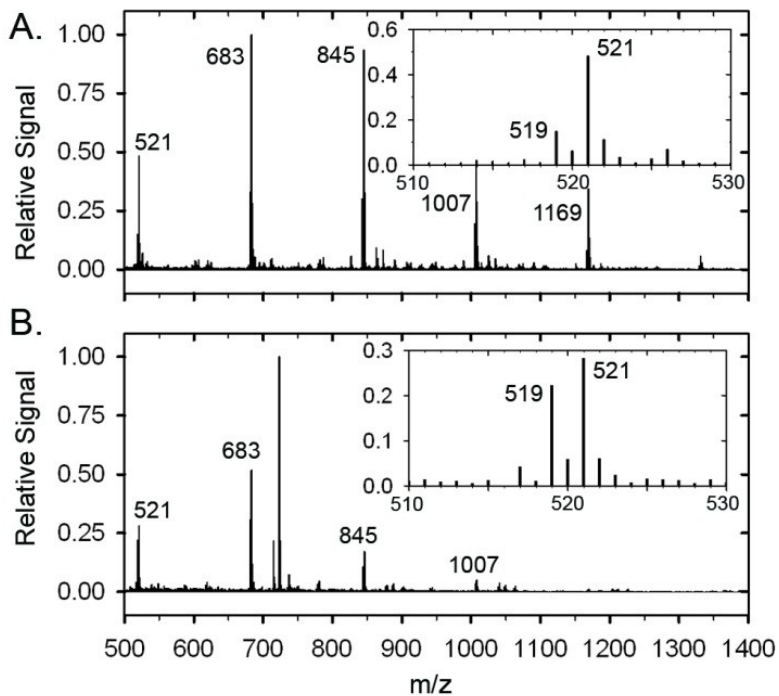**Figure 6.** Mass spectra illustrating the incorporation of ^{18}O from $^{18}O_2$ into the aldonic acid products of NCU08760 with CDH as the reductant. Assays were performed with NCU08760 and CDH in $^{18}O_2$ at pH 7.8 (A) or pH 5.0 (B) and the products were analyzed by LC-MS.
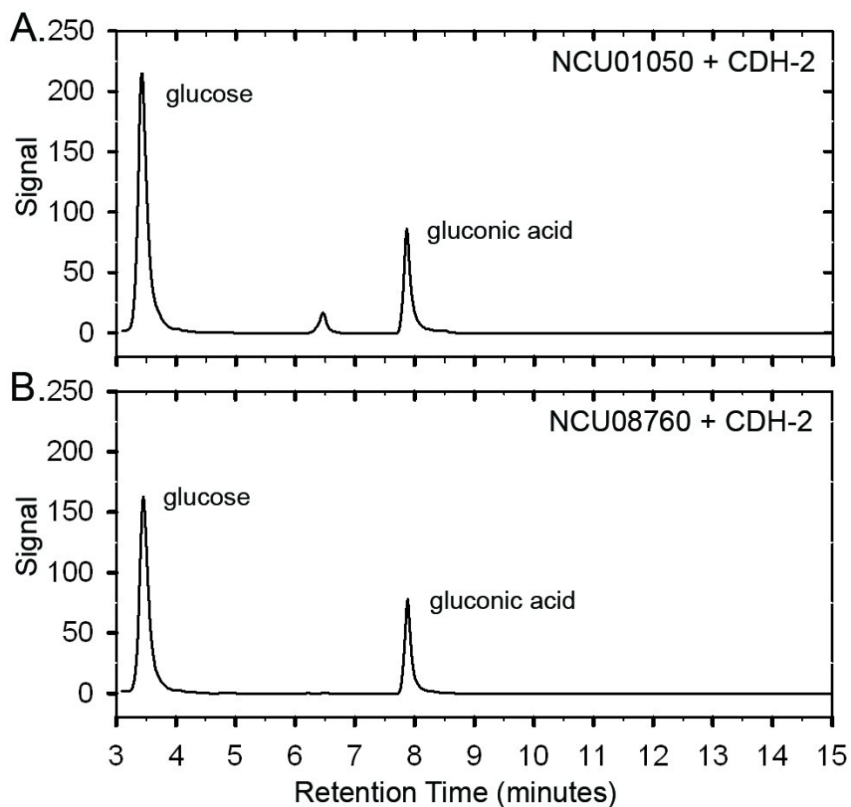
Figure 7. TFA hydrolysis of PMO reaction products. The products of PASC degradation by MtCDH-2 and NCU08760 (A) or NCU01050 (B) were hydrolyzed by TFA for 1 hour and analyzed by HPLC (Dionex PA-200). Both reactions produced a mixture of glucose and gluconic acid. Gluconic acid is formed from the reaction of CDH on the reducing ends of cellodextrins and is also a product of NCU08760. The NCU01050 reaction contained an additional peak with a retention time of 6.4 minutes.

Figure 8. Products generated by NCU01050 after sodium borohydride reduction and TFA hydrolysis. (A) Chemical structures of predicted products. (B) Dionex HPAEC of products showing production of galactose.

89

6.4 Discussion

Evidence that GH61s are copper enzymes requiring electron transfer from CDH to cleave cellulose in an oxygen dependent manner provides the basis to propose a chemical mechanism for a new group of enzymes acting as polysaccharide monooxygenases (PMOs)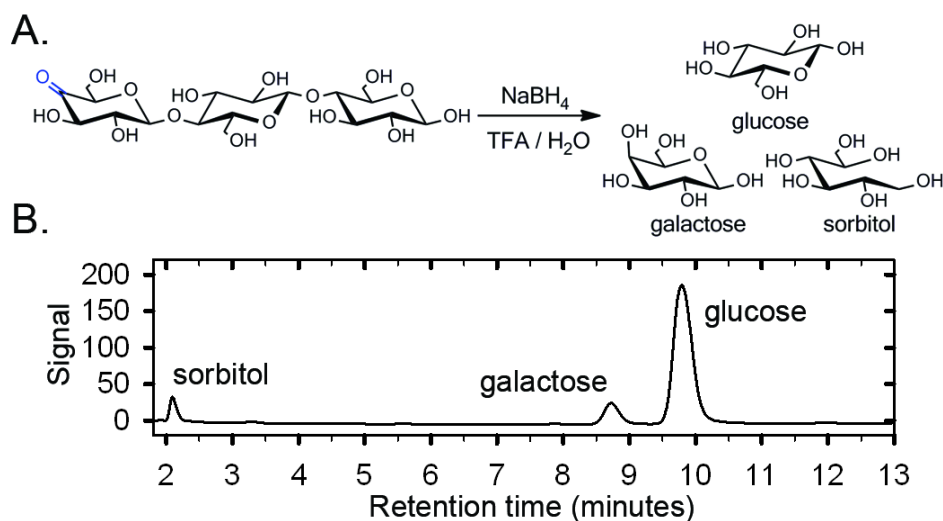 (Fig. 9). Precedent is drawn from the well-studied copper monooxygenases (*166, 167*). The work reported here supports one electron reduction of PMO-Cu(II) to PMO-Cu(I) by the CDH heme domain followed by oxygen binding and internal electron transfer to form a copper superoxo intermediate. Hydrogen atom abstraction by the copper superoxo at the 1-position (by NCU08760) or the 4-position (by NCU01050) of an internal carbohydrate then takes place, generating a copper hydroperoxo intermediate and a substrate radical. The 2nd electron from CDH then facilitates O-O bond cleavage releasing water and generating a copper oxo radical that couples with the substrate radical, thereby hydroxylating the polysaccharide at the 1- or 4-position. The additional oxygen atom destabilizes the glycosidic bond leading to elimination of the adjacent glucan and formation of a sugar lactone or ketoaldose. This elimination would be facilitated by a general acid, possibly a third absolutely conserved histidine that is located on the surface of all fungal PMO proteins near the metal binding site (*161*). It is possible that a 2-electron reduction of oxygen to a Cu-OOH intermediate could abstract the hydrogen. However, peroxide is not able to shunt the reaction in the absence of CDH, and catalase showed no inhibitory effect (data not shown).
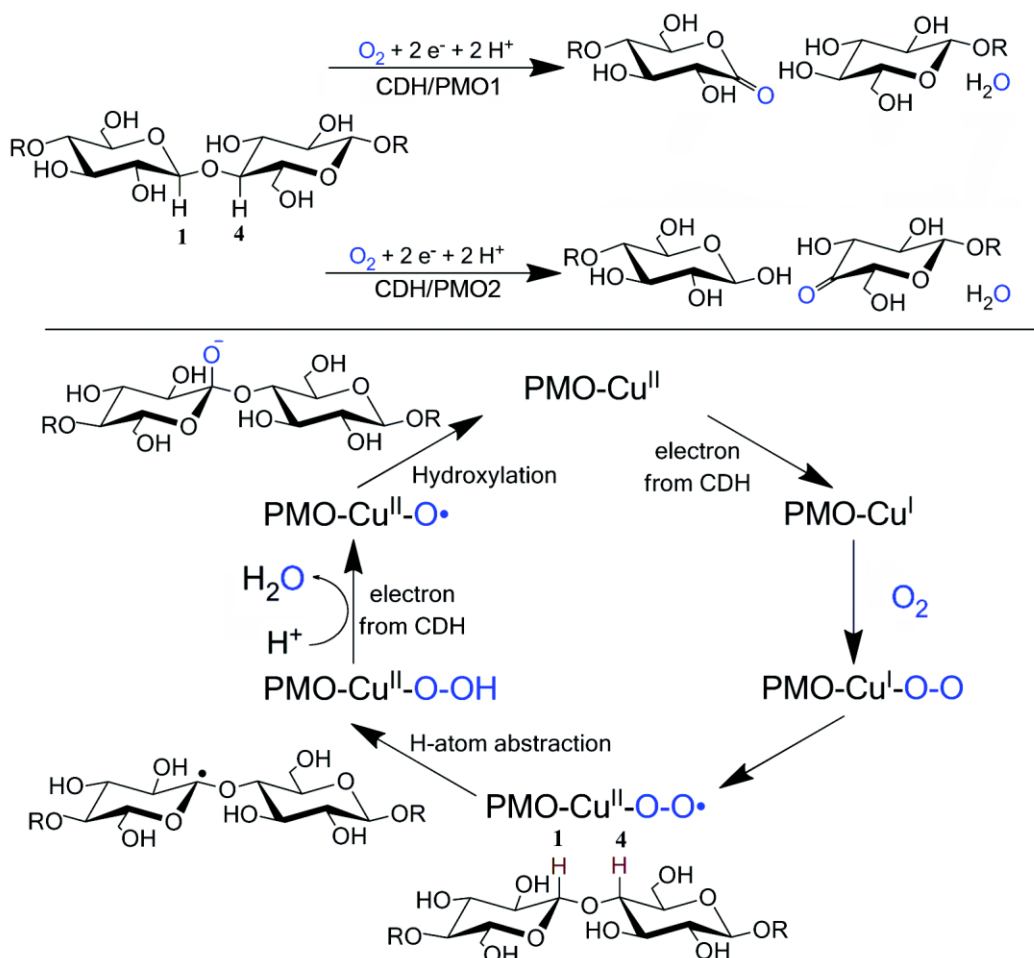
Figure 9. PMO reactions and proposed mechanism. (Top) Type 1 PMOs abstract a hydrogen atom from the 4 position leading to formation of 4-ketoaldose products. Type 2 PMOs catalyze hydrogen atom abstraction from the 1 position leading to formation of aldonic acids. (Bottom) PMO mechanism: An electron from the heme domain of CDH reduces the PMO Cu(II) to Cu(I) and then O_2 binds. Internal electron transfer takes place to form a copper superoxo intermediate, which then abstracts a H• from the 1 or 4 position on the carbohydrate. A second electron from CDH leads to homolytic cleavage of the Cu-bound hydroperoxide. The copper oxo species (Cu-O•) then couples with the substrate radical, hydroxylating the substrate. Addition of the oxygen atom destabilizes the glycosidic bond and leads to elimination of the adjacent glucan.

6.5 Materials and methods

Assays on PASC Phosphoric acid swollen cellulose (PASC) was prepared by addition of 10g Avicel® to 500 mL 85% phosphoric acid and blended for 30 minutes. The cellulose was then precipitated by the addition of 4L of ice cold water and washed with water multiple times until the pH was >4. The concentration of PASC was determined by the phenol-sulfuric acid assay. Assays contained 5.0 µM PMO, 0.5 µM CDH-2, and 9 mg/mL PASC in 10 mM ammonium acetate pH 5.0 and were performed at 40 °C unless otherwise noted. In some assays, 2 mM ascorbic acid was used in place of CDH-2.

Product analysis by HPAEC. Cellulase assays were mixed with 1 volume of 0.1 M NaOH to quench the reaction and then centrifuged to remove the supernatant. Samples were analyzed on a Dionex ICS-3000 HPAEC. Products were separated on a PA-200 HPAEC column using 0.1 M NaOH in the mobile phase with the concentration of sodium acetate increasing from 0 to 140 mM (14 min), 140-300 mM (8 min), 300 to 400 mM (4 min) and then held constant at 500 mM (3 min) before re-equilibration in 0.1 M NaOH (4 min). The flow rate was set to 0.4 mL/min, the column was maintained at a temperature of 30 °C, and samples were detected on an electrochemical detector. Authentic standards of glucose, cellodextrins, glucono-δ-lactone and cellobiono-δ-lactone were used to determine retention times and for quantification. Cellobiono-δ-lactone was synthesized as previously described (*168*).

Product analysis by LC-MS. Samples were analyzed by an Agilent HPLC (1200 series) connected to an electrospray ionization emitter in a linear ion trap mass spectrometer (LTQ XL, Thermo Scientific). Carbohydrates were separated using a SeQuant ZIC®-HILIC column (150 x 2.1 mm, 3.5 µM 100Å) with a SeQuant ZIC®-HILIC guard column (20 x 2.1 mm, 5 µm). Solvent A was 5 mM ammonium acetate pH 7.2 and solvent B was 90% acetonitrile and 10 mM ammonium acetate pH 6.5. Samples were prepared by centrifugation of the assay mixture followed by the addition of 1 volume of 100% acetonitrile and 1% formic acid to the supernatant. Sample injection was set to 5 µL. The elution program consisted of a linear gradient from 80% B to 20% B over 14 minutes followed by 5 minutes at 20% B then re-equilibration for 2 minutes at 80% B. The column temperature was maintained at 25 °C and the flow rate was 0.2 mL/minute. Mass spectra were acquired in negative ion mode over the range m/z = 310-2000. Data processing was performed using Xcalibur software (version 2.2, Thermo Scientific).

Metal dependence of PMO activity. Apo-PMO was prepared by treatment of as purified PMO with 10 mM EDTA for 24 hours. Protein was then concentrated in a 3 kDa

spin concentrator and loaded onto a Sephacryl S100 column with 10 mM Tris pH 8.0 and 100 µM EDTA in the mobile phase. Following elution, Sigma TraceSELECT grade buffers, metals, and water were used for all assays and only extensively washed and rinsed plastics were used due to problems with copper contamination. The protein was buffer exchanged > 100-fold into 10 mM sodium acetate (Sigma Cat # 59929 and #07692) in water (Sigma Cat # 14211) to a final concentration of >40 µM PMO. Cu- or Zn-bound PMO was then produced by reconstitution with a 2-fold molar excess of $CuSO_4$ (Sigma Cat #: 203165) or $ZnSO_4$ (Sigma Cat #: 204986).

PASC Assays to determine the metal dependence of NCU08760 activity were performed as described above except TraceSelect grade buffer and water were used with equimolar amounts of apo-, Zn-, or Cu-bound PMO. Before the assay, PASC was mixed in a large volumetric excess of 100 uM EDTA for 48 hours. The PASC was then washed multiple times with Traceselect water until the concentration of residual EDTA was < 1nM.

Reactions in ^{18}O or ^{16}O gas. ^{18}O labeled molecular oxygen was purchased from Sigma (99 atom % ^{18}O). For experiments with $^{18}O_2$, all reagents were made anaerobic by incubation in an anaerobic glove bag for at least 48 hours. Reactions were setup in the glove bag (Coy) in oxygen impermeable vials with double faced PTFE/silicone septa. The sample was then attached to a Schlenk line, then vacuum applied followed by $^{18}O_2$ delivery to the headspace of the sample.

Reduction with sodium borohydride. PMO reaction supernatant (100 µL) was mixed with freshly prepared 20 mg/mL sodium borohydride in 1.0 M ammonium hydroxide (100 µL). The reaction was allowed to proceed for 2 hours at room temperature with occasional mixing. Glacial acetic acid was added until bubbling ceased (25 µL). Next, 200 µL of a 9:1 methanol:acetic acid mixture was added to the reaction and the samples were then dried under a stream of nitrogen at 40 °C. The resulting sample was hydrolyzed by TFA as described below.

TFA hydrolysis. Samples were incubated with 2.0 M TFA at 121 °C on a heating block for 1 hour. The samples were then removed from the heat block, cooled on ice, centrifuged briefly, and dried under a stream of nitrogen at 40 °C. After drying the sample was washed with 300 µL of isopropanol and dried under a stream of nitrogen twice. Dried samples were then dissolved in 1.0 mL of water for analysis on a PA-20 column (Dionex) or 1.0 mL 0.1 M sodium hydroxide for analysis on a PA-200 column (Dionex).

Monosaccharide analysis on Dionex HPAEC. The products of sodium borohydride reduction and TFA hydrolysis were separated isocratically on a PA-20 column.

Products were eluted with 10 mM potassium hydroxide for 20 minutes followed by a 5 minute wash with 100 mM potassium hydroxide and a 5 minute equilibration at 10 mM potassium hydroxide. Standards for glucose, galactose, sorbitol, gluconic acid, and glucuronic acid were purchased from Sigma.

Isotope labeled oxygen incorporation analysis by LC-MS. Samples were analyzed on an Agilent HPLC (1200 series) connected to an electrospray ionization emitter in a linear ion trap mass spectrometer (LTQ XL, Thermo Scientific) as described previously(*169*) with the following modifications. To limit the exchange of bulk water into the oxidized products, the mobile phases were changed to contain 10 mM ammonium bicarbonate pH 8.0 in both solvent A and solvent B. Following centrifugation of the assays and removal of the supernatant, 1 volume of 100% acetonitrile was added to the samples. Samples were immediately injected into the LC-MS and the injection was set to 5 μL. The elution program consisted of a linear gradient from 80% B to 20% B over 7 minutes followed by 5 minutes at 20% B then re-equilibration for 2 minutes at 80% B.

6.6 Acknowledgements

7. Conclusions and Outlook.

In closing, the data reported in this thesis shows that oxidative enzymes are key components of the enzyme cocktails secreted by fungi for plant cell wall degradation. Based on the results from numerous biochemical experiments we have proposed a chemical mechanism for the action of a new family of metalloenzymes, the polysaccharide monooxygenases. In contrast to the widely discussed Fenton model, our data supports a direct enzymatic oxidation of cellulose leading to glycosidic bond cleavage. The genetic and biochemical experiments reported here with CDH show that in *N. crassa,* CDH-1 functions as the sole extracellular reductase and that the heme domain of CDH is required for enhancement of cellulose degradation. In many highly cellulolytic ascomycetes, CDH contains a C-terminal cellulose-binding module which targets the enzyme to the cellulose surface, probably to facilitate electron transfer to PMOs. Basidiomycete CDHs do not contain CBMs; however, many basidiomycetes secrete high levels of a protein with a CDH heme domain fused to a CBM (*170*). These "free" heme domains would require electrons from another, unidentified reductase, to potentiate the action of PMOs. Reductases functionally similar to CDH presumably exist in bacteria, but to our knowledge none have been identified.

Bacterial proteins with structural homology to fungal PMOs have recently been biochemically characterized and shown to oxidize cellulose (*171*) and chitin (*153*). These bacterial proteins have the same conserved metal ligands as fungal PMOs and are likely copper metalloenzymes employing a similar mechanism. Further work is needed to confirm that the activity of these proteins is dependent on bound copper. PMOs are conserved in every cellulolytic fungus studied to date, and are generally expressed at very high levels during growth on cellulose. In the white rot fungus, *P. chrysosporium,* and the brown rot fungus, *S. lacrymans* expression profiling experiments revealed that a PMO is the most highly upregulated protein in response to growth on biomass (*156, 172*).

N. crassa has a potent genome defense mechanism, RIP, which prevents nearly all gene duplications (*61*). Surprisingly, in the *N. crassa* genome*,* there are 14 genes encoding predicted PMOs. Previous expression profiling studies showed that expression of at least 10 of the PMOs was induced during growth on cellulose. The average pairwise sequence identity between PMOs in *N. crassa* is only 33%, suggesting that these proteins may have diverse functions. The biochemical characterization of three members of the PMO superfamily reported here showed that different PMOs catalyze reactions with different regiospecificity. At this point, phylogenetic inference in uncharacterized clades does not allow for the prediction of regioselectivity. Further work exploring the activity of the PMOs encoded by NCU00836, NCU03328, and NCU07760 may reveal new reactions or phylogenetic trends that allow for functional prediction.

In addition to their prevalence in nature, supplementation of cellulase cocktails with PMOs can significantly reduce the enzyme loading required for saccharification of lignocellulose. Because several different PMOs are produced by fungi, and PMOs work through a mechanism completely orthogonal to that of cellulases, it is likely that addition of multiple PMOs to cellulase cocktails will further reduce enzyme loadings. Additional mechanistic insights into this large family of enzymes may facilitate their development for commercial applications.

8 References

1. Reese, E., Mandels, M., and Weiss, A. (1972) Cellulose as a novel energy source Advances in Biochemical Engineering, Volume 2, pp 181-200, Springer Berlin / Heidelberg.
2. Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., and Foust, T. D. (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production, *Science 315*, 804-807.
3. Perlack, R. D., Wright, L. L., Turhollow, A. F., Graham, R. L., Stokes, B. J., and Erbach, D. C. (2005) Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply., In *DOE/GO-102005-2135*, Oak Ridge National Laboratory Oak Ridge, TN.
4. Somerville, C., Youngs, H., Taylor, C., Davis, S. C., and Long, S. P. (2010) Feedstocks for lignocellulosic biofuels, *Science 329*, 790-792.
5. Aden, A. (2008) Biochemical production of ethanol from corn stover: 2007 state of technology model., In *NREL/TP-510-43205*, National Renewable Energy Laboratory.
6. Wang, D., Portis, A. R., Jr., Moose, S. P., and Long, S. P. (2008) Cool C4 photosynthesis: pyruvate Pi dikinase expression and activity corresponds to the exceptional cold tolerance of carbon assimilation in Miscanthus x giganteus, *Plant Physiol 148*, 557-567.
7. Dohleman, F. G., Heaton, E. A., Leakey, A. D. B., and Long, S. P. (2009) Does greater leaf-level photosynthesis explain the larger solar energy conversion efficiency of Miscanthus relative to switchgrass?, *Plant Cell Environ 32*, 1525-1537.
8. Beale, C. V., and Long, S. P. (1995) Can perennial C4 grasses attain high efficiencies of radiant energy conversion in cool climates?, *Plant, Cell & Environment 18*, 641-650.
9. Perlack, R. D., and Stokes, B. J. (2011) U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry., (Energy, U. S. D. o., Ed.), p 227, Oak Ridge, TN.
10. Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S., and Youngs, H. (2004) Toward a systems approach to understanding plant-cell walls, *Science 306*, 2206-2211.
11. http://upload.wikimedia.org/wikipedia/commons/5/53/Plant_cell_wall_diagram.svg.
12. OSullivan, A. C. (1997) Cellulose: the structure slowly unravels, *Cellulose 4*, 173-207.
13. Fernandes, A. N., Thomas, L. H., Altaner, C. M., Callow, P., Forsyth, V. T., Apperley, D. C., Kennedy, C. J., and Jarvis, M. C. (2011) Nanostructure of cellulose microfibrils in spruce wood, *P Natl Acad Sci USA 108*, E1195-1203.
14. Somerville, C. (2006) Cellulose synthesis in higher plants, *Annu Rev Cell Dev Bi 22*, 53-78.
15. Himmel, M. E. (2008) *Biomass recalcitrance : deconstructing the plant cell wall for bioenergy*, Blackwell Pub., Oxford.
16. Balakshin, M. Y., Capanema, E. A., and Chang, H. M. (2007) MWL fraction with a high concentration of lignin-carbohydrate linkages: Isolation and 2D NMR spectroscopic analysis, *Holzforschung 61*, 1-7.
17. Lloyd, T. A., and Wyman, C. E. (2005) Combined sugar yields for dilute sulfuric acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids, *Bioresour Technol 96*, 1967-1977.

18. Freudenberg, K., and Neish, A. C. (1968) *Constitution and biosynthesis of lignin*, Springer, Berlin, Heidelberg, New York,.

19. Domalski, E. S., Jobe, T. L., Milne, T. A., and American Society of Mechanical Engineers. Research Committee on Industrial and Municipal Wastes. (1987) *Thermodynamic data for biomass materials and waste components*, American Society of Mechanical Engineers, New York, N.Y. (United Engineering Ctr., 345 E. 47th St., New York 10017).

20. Yoshida, M., Liu, Y., Uchida, S., Kawarada, K., Ukagami, Y., Ichinose, H., Kaneko, S., and Fukuda, K. (2008) Effects of cellulose crystallinity, hemicellulose, and lignin on the enzymatic hydrolysis of Miscanthus sinensis to monosaccharides, *Biosci Biotechnol Biochem 72*, 805-810.

21. Malhi, Y. (2002) Carbon in the atmosphere and terrestrial biosphere in the 21st century, *Philos T Roy Soc A 360*, 2925-2945.

22. Hess, M., Sczyrba, A., Egan, R., Kim, T. W., Chokhawala, H., Schroth, G., Luo, S., Clark, D. S., Chen, F., Zhang, T., Mackie, R. I., Pennacchio, L. A., Tringe, S. G., Visel, A., Woyke, T., Wang, Z., and Rubin, E. M. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen, *Science 331*, 463-467.

23. Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., Cayouette, M., McHardy, A. C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S. G., Podar, M., Martin, H. G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., Szeto, E., Salamov, A., Barry, K., Mikhailova, N., Kyrpides, N. C., Matson, E. G., Ottesen, E. A., Zhang, X., Hernandez, M., Murillo, C., Acosta, L. G., Rigoutsos, I., Tamayo, G., Green, B. D., Chang, C., Rubin, E. M., Mathur, E. J., Robertson, D. E., Hugenholtz, P., and Leadbetter, J. R. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite, *Nature 450*, 560-565.

24. Falkowski, P., Scholes, R. J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hogberg, P., Linder, S., Mackenzie, F. T., Moore, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., and Steffen, W. (2000) The global carbon cycle: A test of our knowledge of earth as a system, *Science 290*, 291-296.

25. Wei, H., Xu, Q., Taylor, L. E., Baker, J. O., Tucker, M. P., and Ding, S. Y. (2009) Natural paradigms of plant cell wall degradation, *Curr Opin Biotechnol 20*, 330-338.

26. Stajich, J. E., Berbee, M. L., Blackwell, M., Hibbett, D. S., James, T. Y., Spatafora, J. W., and Taylor, J. W. (2009) The Fungi, *Curr Biol 19*, R840-R845.

27. Adrio, J. L., and Demain, A. L. (2003) Fungal biotechnology, *Int Microbiol 6*, 191-199.

28. Gusakov, A. V. (2011) Alternatives to Trichoderma reesei in biofuel production, *Trends Biotechnol 29*, 419-425.

29. Bhat, M. K. (2000) Cellulases and related enzymes in biotechnology, *Biotechnol Adv 18*, 355-383.

30. Dadi, A. P., Varanasi, S., and Schall, C. A. (2006) Enhancement of cellulose saccharification kinetics using an ionic liquid pretreatment step, *Biotechnology and Bioengineering 95*, 904-910.

31. Garg, H. G., Cowman, M. K., and Hales, C. A. (2008) *Carbohydrate chemistry, biology and medical applications*, 1st ed., Elsevier, Amsterdam ; Boston.

32. Belitz, H. D., Grosch, W., and Schieberle, P. (2004) *Food chemistry*, 3rd rev. ed., Springer, Berlin ; New York.

33.	Xiang, Q., Lee, Y. Y., Pettersson, P. O., and Torget, R. (2003) Heterogeneous aspects of acid hydrolysis of alpha-cellulose, *Appl Biochem Biotechnol 105*, 505-514.

34.	Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, *Nucleic Acids Res 37*, D233-D238.

35.	Henrissat, B., Driguez, H., Viet, C., and Schulein, M. (1985) Synergism of Cellulases from Trichoderma-Reesei in the Degradation of Cellulose, *Bio-Technol 3*, 722-726.

36.	Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K., Teeri, T. T., and Jones, T. A. (1994) The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from Trichoderma reesei, *Science 265*, 524-528.

37.	MacKenzie, L. F., Sulzenbacher, G., Divne, C., Jones, T. A., Woldike, H. F., Schulein, M., Withers, S. G., and Davies, G. J. (1998) Crystal structure of the family 7 endoglucanase I (Cel7B) from Humicola insolens at 2.2 A resolution and identification of the catalytic nucleophile by trapping of the covalent glycosyl-enzyme intermediate, *Biochem J 335 (Pt 2)*, 409-416.

38.	Kleywegt, G. J., Zou, J. Y., Divne, C., Davies, G. J., Sinning, I., Stahlberg, J., Reinikainen, T., Srisodsuk, M., Teeri, T. T., and Jones, T. A. (1997) The crystal structure of the catalytic core domain of endoglucanase I from Trichoderma reesei at 3.6 A resolution, and a comparison with related enzymes, *Journal of molecular biology 272*, 383-397.

39.	Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., Penttila, M., Ando, T., and Samejima, M. (2011) Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface, *Science 333*, 1279-1282.

40.	Igarashi, K., Koivula, A., Wada, M., Kimura, S., Penttila, M., and Samejima, M. (2009) High speed atomic force microscopy visualizes processive movement of Trichoderma reesei cellobiohydrolase I on crystalline cellulose, *The Journal of biological chemistry 284*, 36186-36190.

41.	Lehtio, J., Sugiyama, J., Gustavsson, M., Fransson, L., Linder, M., and Teeri, T. T. (2003) The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules, *P Natl Acad Sci USA 100*, 484-489.

42.	Shoseyov, O., Shani, Z., and Levy, I. (2006) Carbohydrate binding modules: biochemical properties and novel applications, *Microbiology and molecular biology reviews : MMBR 70*, 283-295.

43.	Tian, C., Beeson, W. T., Iavarone, A. T., Sun, J., Marletta, M. A., Cate, J. H., and Glass, N. L. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa, *Proc Natl Acad Sci U S A 106*, 22157-22162.

44.	Reese, E. T., Siu, R. G., and Levinson, H. S. (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis, *J Bacteriol 59*, 485-497.

45.	Medve, J., Karlsson, J., Lee, D., and Tjerneld, F. (1998) Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from Trichoderma reesei: adsorption, sugar production pattern, and synergism of the enzymes, *Biotechnology and bioengineering 59*, 621-634.

46.	Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology, *Microbiol Mol Biol Rev 66* 506-577.

47. Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C. P., Ferreira, P., Ruiz-Duenas, F. J., Martinez, A. T., Kersten, P., Hammel, K. E., Vanden Wymelenberg, A., Gaskell, J., Lindquist, E., Sabat, G., Bondurant, S. S., Larrondo, L. F., Canessa, P., Vicuna, R., Yadav, J., Doddapaneni, H., Subramanian, V., Pisabarro, A. G., Lavin, J. L., Oguiza, J. A., Master, E., Henrissat, B., Coutinho, P. M., Harris, P., Magnuson, J. K., Baker, S. E., Bruno, K., Kenealy, W., Hoegger, P. J., Kues, U., Ramaiya, P., Lucas, S., Salamov, A., Shapiro, H., Tu, H., Chee, C. L., Misra, M., Xie, G., Teter, S., Yaver, D., James, T., Mokrejs, M., Pospisek, M., Grigoriev, I. V., Brettin, T., Rokhsar, D., Berka, R., and Cullen, D. (2009) Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion, *Proc Natl Acad Sci U S A 106 (6)*, 1954-1959.

48. Mason, M. G., Nicholls, P., and Wilson, M. T. (2003) Rotting by radicals--the role of cellobiose oxidoreductase?, *Biochem Soc Trans 31*, 1335-1336.

49. Fenton, H. J. H. (1894) LXXIII.-Oxidation of tartaric acid in presence of iron, *Journal of the Chemical Society, Transactions 65*, 899-910.

50. Haber, F., and Weiss, J. (1934) The Catalytic Decomposition of Hydrogen Peroxide by Iron Salts, *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences 147*, 332-351.

51. Kirk, T. K., Ibach, R., Mozuch, M. D., Conner, A. H., and Highley, T. L. (1991) Characteristics of Cotton Cellulose Depolymerized by a Brown-Rot Fungus, by Acid, or by Chemical Oxidants, *Holzforschung 45*, 239-244.

52. Kersten, P. J., and Kirk, T. K. (1987) Involvement of a New Enzyme, Glyoxal Oxidase, in Extracellular H2o2 Production by Phanerochaete-Chrysosporium, *J Bacteriol 169*, 2195-2201.

53. Baldrian, P., and Valaskova, V. (2008) Degradation of cellulose by basidiomycetous fungi, *Fems Microbiol Rev 32*, 501-521.

54. Ek, M., Gierer, J., and Jansbo, K. (1989) Study on the Selectivity of Bleaching with Oxygen-Containing Species, *Holzforschung 43*, 391-396.

55. Henriksson, G., Johansson, G., and Pettersson, G. (2000) A critical review of cellobiose dehydrogenases, *J Biotechnol 78*, 93-113.

56. Kremer, S. M., and Wood, P. M. (1992) Production of Fenton's reagent by cellobiose oxidase from cellulolytic cultures of Phanerochaete chrysosporium, *European journal of biochemistry / FEBS 208*, 807-814.

57. Zamocky, M., Ludwig, R., Peterbauer, C., Hallberg, B. M., Divne, C., Nicholls, P., and Haltrich, D. (2006) Cellobiose dehydrogenase--a flavocytochrome from wood-degrading, phytopathogenic and saprotropic fungi, *Curr Protein Pept Sci 7*, 255-280.

58. Hallberg, B. M., Bergfors, T., Backbro, K., Pettersson, G., Henriksson, G., and Divne, C. (2000) A new scaffold for binding haem in the cytochrome domain of the extracellular flavocytochrome cellobiose dehydrogenase, *Structure 8*, 79-88.

59. Tejirian, A., and Xu, F. (2010) Inhibition of cellulase-catalyzed lignocellulosic hydrolysis by iron and oxidative metal ions and complexes, *Applied and environmental microbiology 76*, 7673-7682.

60. Bao, W., and Renganathan, V. (1992) Cellobiose oxidase of Phanerochaete chrysosporium enhances crystalline cellulose degradation by cellulases, *FEBS Lett 302*, 77-80.

61. Galagan, J. E., Calvo, S. E., Borkovich, K. A., Selker, E. U., Read, N. D., Jaffe, D., FitzHugh, W., Ma, L. J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C. B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M. A., Werner-Washburne, M., Selitrennikoff, C. P., Kinsey, J. A., Braun, E. L., Zelter, A., Schulte, U., Kothe, G. O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R. L., Perkins, D. D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R. J., Osmani, S. A., DeSouza, C. P., Glass, L., Orbach, M. J., Berglund, J. A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D. O., Alex, L. A., Mannhaupt, G., Ebbole, D. J., Freitag, M., Paulsen, I., Sachs, M. S., Lander, E. S., Nusbaum, C., and Birren, B. (2003) The genome sequence of the filamentous fungus Neurospora crassa, *Nature 422*, 859-868.

62. Davis, R. H., and Perkins, D. D. (2002) Neurospora: a model of model microbes [Review], *Nature Reviews Genetics 3*, 397-403.

63. Perkins, D. D., Turner, B. C., and Barry, E. G. (1976) Strains of Neurospora collected from nature, *Evolution 30*, 281-313.

64. Pandit, A., and Maheshwari, R. (1996) Life-history of Neurospora intermedia in a sugar cane field, *Journal of Biosciences (Bangalore) 21*, 57-79.

65. Smith, M. L., Micali, O. C., Hubbard, S. P., Mir-Rashed, N., Jacobson, D. J., and Glass, N. L. (2000) Vegetative incompatibility in the *het-6* region of *Neurospora crassa* is mediated by two linked genes, *Genetics 155*, 1095-1104.

66. Eberhart, B. M., Beck, R. S., and Goolsby, K. M. (1977) Cellulase of Neurospora crassa, *J. Bacteriol. 130*, 181-186.

67. Romero, M. D., Aguado, J., L., G., and Ladero, M. (1999) Cellulase production by *Neurospora crassa* on wheat straw, *Enzyme Microbial Technol 25*, 244-250.

68. Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., Chapman, J., Chertkov, O., Coutinho, P. M., Cullen, D., Danchin, E. G., Grigoriev, I. V., Harris, P., Jackson, M., Kubicek, C. P., Han, C. S., Ho, I., Larrondo, L. F., de Leon, A. L., Magnuson, J. K., Merino, S., Misra, M., Nelson, B., Putnam, N., Robbertse, B., Salamov, A. A., Schmoll, M., Terry, A., Thayer, N., Westerholm-Parvinen, A., Schoch, C. L., Yao, J., Barabote, R., Nelson, M. A., Detter, C., Bruce, D., Kuske, C. R., Xie, G., Richardson, P., Rokhsar, D. S., Lucas, S. M., Rubin, E. M., Dunn-Coleman, N., Ward, M., and Brettin, T. S. (2008) Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina), *Nat Biotechnol 26*, 553-560.

69. Dunlap, J. C., Borkovich, K. A., Henn, M. R., Turner, G. E., Sachs, M. S., Glass, N. L., McCluskey, K., Plamann, M., Galagan, J. E., Birren, B. W., Weiss, R. L., Townsend, J. P., Loros, J. J., Nelson, M. A., Lambreghts, R., Colot, H. V., Park, G., Collopy, P., Ringelberg, C., Crew, C., Litvinkova, L., DeCaprio, D., Hood, H. M., Curilla, S., Shi, M., Crawford, M., Koerhsen, M., Montgomery, P., Larson, L., Pearson, M., Kasuga, T., Tian, C., Basturkmen, M., Altamirano, L., Xu, J., and Jay, C. D. (2007) Enabling a Community to Dissect an Organism: Overview of the Neurospora Functional Genomics Project, In *Advances in Genetics*, pp 49-96, Academic Press.

70. Kasuga, T., and Glass, N. L. (2008) Dissecting colony development of Neurospora crassa using mRNA profiling and comparative genomics approaches, *Eukaryot Cell 7*, 1549-1564.

71. Tian, C., Kasuga, T., Sachs, M. S., and Glass, N. L. (2007) Transcriptional profiling of cross pathway control in *Neurospora crassa* and comparative analysis of the Gcn4 and CPC1 regulons, *Eukaryotic Cell 6*, 1018-1029.

72. Kasuga, T., Townsend, J. P., Tian, C., Gilbert, L. B., Mannhaupt, G., Taylor, J. W., and Glass, N. L. (2005) Long-oligomer microarray profiling in Neurospora crassa reveals the transcriptional program underlying biochemical and physiological events of conidial germination, *Nucleic Acids Res 33*, 6469-6485.

73. Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M., and Mewes, H. W. (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes., *Nucleic Acids Res 32*, 5539-5545.

74. Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology, *Microbiol Mol Biol Rev 66*, 506-577, table of contents.

75. Linder, M., and Teeri, T. T. (1996) The cellulose-binding domain of the major cellobiohydrolase of Trichoderma reesei exhibits true reversibility and a high exchange rate on crystalline cellulose, *Proc Natl Acad Sci U S A 93*, 12251-12255.

76. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics, *Nucleic Acids Res 37*, D233-238.

77. Bouws, H., Wattenberg, A., and Zorn, H. (2008) Fungal secretomes--nature's toolbox for white biotechnology, *Appl Microbiol Biotechnol 80*, 381-388.

78. Espagne, E., Lespinet, O., Malagnac, F., Da Silva, C., Jaillon, O., Porcel, B. M., Couloux, A., Aury, J. M., Segurens, B., Poulain, J., Anthouard, V., Grossetete, S., Khalili, H., Coppin, E., Dequard-Chablat, M., Picard, M., Contamine, V., Arnaise, S., Bourdais, A., Berteaux-Lecellier, V., Gautheret, D., de Vries, R. P., Battaglia, E., Coutinho, P. M., Danchin, E. G., Henrissat, B., Khoury, R. E., Sainsard-Chanet, A., Boivin, A., Pinan-Lucarre, B., Sellem, C. H., Debuchy, R., Wincker, P., Weissenbach, J., and Silar, P. (2008) The genome sequence of the model ascomycete fungus Podospora anserina, *Genome Biol 9*, R77.

79. Foreman, P. K., Brown, D., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N. S., Goedegebuur, F., Houfek, T. D., England, G. J., Kelley, A. S., Meerman, H. J., Mitchell, T., Mitchinson, C., Olivares, H. A., Teunissen, P. J., Yao, J., and Ward, M. (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei, J Biol Chem 278*, 31988-31997.

80. Vanden Wymelenberg, A., Gaskell, J., Mozuch, M., Kersten, P., Sabat, G., Martinez, D., and Cullen, D. (2009) Transcriptome and secretome analyses of Phanerochaete chrysosporium reveal complex patterns of gene expression, *Appl Environ Microbiol 75*, 4058-4068.

81. Seiboth, B., Messner, R., Gruber, F., and Kubicek, C. P. (1992) Disruption of the *Trichoderma reesei cbh2* gene coding for cellobiohydrolase II leads to a delay in the triggering of cellulase formation by cellulose, *J Gen Microbiol. 138*, 1259-1264.

82. Suominen, P. L., Mantyla, A. L., Karhunen, T., Hakola, S., and Nevalainen, H. (1993) High frequency one-step gene replacement in Trichoderma reesei. II. Effects of deletions of individual cellulase genes, *Mol Gen Genet 241*, 523-530.

83. Seiboth, B., Hakola, S., Mach, R. L., Suominen, P. L., and Kubicek, C. P. (1997) Role of four major cellulases in triggering of cellulase gene expression by cellulose in Trichoderma reesei, *J Bacteriol 179*, 5318-5320.

84. Stricker, A. R., Mach, R. L., and de Graaff, L. H. (2008) Regulation of transcription of cellulases- and hemicellulases-encoding genes in Aspergillus niger and Hypocrea jecorina (Trichoderma reesei), *Appl Microbiol Biotechnol 78*, 211-220.

85. Vogel, H. J. (1956) A convenient growth medium for *Neurospora, Microbiol. Genet. Bull. 13*, 42-46.

86. Canevascini, G. (1988) Cellobiose Dehydrogenase from Sporotrichum-Thermophile, *Method Enzymol 160*, 443-448.

87. Townsend, J. P. (2004) Resolution of large and small differences in gene expression using models for the Bayesian analysis of gene expression levels and spotted DNA microarrays, *BMC Bioinformatics 5*, 54.

88. Yang, Y. H., and Speed, T. (2002) Design issues for cDNA microarray experiments, *Nat Rev Genet 3*, 579-588.

89. Vinciotti, V., Khanin, R., D'Alimonte, D., Liu, X., Cattini, N., Hotchkiss, G., Bucca, G., de Jesus, O., Rasaiyaah, J., Smith, C. P., Kellam, P., and Wit, E. (2005) An experimental evaluation of a loop versus a reference design for two-channel microarrays, *Bioinformatics 21*, 492-501.

90. Meiklejohn, C. D., and Townsend, J. P. (2005) A Bayesian method for analysing spotted microarray data, *Brief Bioinform 6*, 318-330.

91. Seo, J., Bakay, M., Chen, Y. W., Hilmer, S., Shneiderman, B., and Hoffman, E. P. (2004) Interactively optimizing signal-to-noise ratios in expression profiling: project-specific algorithm selection and detection p-value weighting in Affymetrix microarrays, *Bioinformatics 20*, 2534-2544.

92. Dementhon, K., Iyer, G., and Glass, N. L. (2006) VIB-1 is required for expression of genes necessary for programmed cell death in Neurospora crassa, *Eukaryot Cell 5*, 2161-2173.

93. Margolin, B. S., Freitag, M., and Selker, E. U. (1997) Improved plasmids for gene targeting at the his-3 locus of Neurospora crassa by electroporation, *Fungal Genet Newslett 44*, 34-36.

94. Bhat, K. M., and Maheshwari, R. (1987) *Sporotrichum thermophile* growth, cellulose degradation, and cellulase activity, *Appl. Environ. Microbiol. 53*, 2175-2182.

95. Katapodis, P., Vrsanska, M., Kekos, D., Nerinckx, W., Biely, P., Claeyssens, M., Macris, B. J., and Christakopoulos, P. (2003) Biochemical and catalytic properties of an endoxylanase purified from the culture filtrate of *Sporotrichum thermophile*, *Carbohydr. Res. 338*, 1881-1890.

96. Canevascini, G., Borer, P., and Dreyer, J. L. (1991) Cellobiose dehydrogenases of *Sporotrichum-(chrysosporium) thermophile Eur. J. Biochem. 198*, 43-52.

97. Zamocky, M., Ludwig, R., Peterbauer, C., Hallberg, B. M., Divne, C., Nicholls, P., and Haltrich, D. (2006) Cellobiose dehydrogenase - A flavocytochrome from wood-degrading, phytopathogenic and saprotropic fungi, *Curr. Protein Peptide Sci. 7*, 255-280.

98. Conchie, J., and Levvy, G. A. (1957) Inhibition of glycosidases by aldonolactones of corresponding configuration, *Biochem. J. 65*, 389-395.

99. Parry, N. J., Beever, D. E., Owen, E., Vandenberghe, I., Van Beeumen, J., and Bhat, M. K. (2001) Biochemical characterization and mechanism of action of a thermostable beta-glucosidase purified from *Thermoascus aurantiacus*, *Biochem. J. 353*, 117-127.

100. Iyayi, C. B., Bruchmann, E. E., and Kubicek, C. P. (1989) Induction of cellulase formation in *Trichoderma-reesei* by cellobiono-1,5-lactone, *Arch. Microbiol. 151*, 326-330.

101. Sawyer, D. T., and Bagger, J. B. (1959) The lactone acid salt equilibria for D-glucono-delta-lactone and the hydrolysis kinetics for this lactone, *J. Am. Chem. Soc. 81*, 5302-5306.

102. Collard, F., Collet, J. F., Gerin, I., Veiga-da-Cunha, M., and Van Schaftingen, E. (1999) Identification of the cDNA en coding human 6-phosphogluconolactonase, the enzyme catalyzing the second step of the pentose phosphate pathway, *FEBS Lett. 459*, 223-226.

103. Hager, P. W., Calfee, M. W., and Phibbs, P. V. (2000) The *Pseudomonas aeruginosa* devB/SOL homolog, pgl, is a member of the hex regulon and encodes 6-phosphogluconolactonase, *J. Bacteriol. 182*, 3934-3941.

104. Zimenkov, D., Gulevich, A., Skorokhodova, A., Biriukova, I., Kozlov, Y., and Mashko, S. (2005) *Escherichia coli* ORF ybhE is pgl gene encoding 6-phosphogluconolactonase (EC 3.1.1.31) that has no homology with known 6PGLs from other organisms, *FEMS Microbiol. Lett. 244*, 275-280.

105. Thomason, L. C., Court, D. L., Datta, A. R., Khanna, R., and Rosner, J. L. (2004) Identification of the *Escherichia coli* K-12 ybhE gene as pgl, encoding 6-phosphogluconolactonase, *J. Bacteriol. 186*, 8248-8253.

106. Kupor, S. R., and Fraenkel, D. G. (1969) 6-Phosphogluconolactonase mutants of *Escherichia coli* and a maltose blue gene, *J. Bacteriol. 100*, 1296-1301.

107. Bruchmann, E. E., Schach, H., and Graf, H. (1987) Role and properties of lactonase in a cellulase system, *Biotechnol. Appl. Biochem. 9*, 146-159.

108. Westermark, U., and Eriksson, K. E. (1974) Cellobiose-quinone oxidoreductase, a new wood-degrading enzyme from white-rot fungi, *Acta Chemica Scandinavica Series B-Organic Chemistry and Biochemistry B 28*, 209-214.

109. Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol. 340*, 783-795.

110. Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., and Bateman, A. (2010) The Pfam protein families database, *Nucleic Acids Res. 38*, D211-D222.

111. Kajander, T., Merckel, M. C., Thompson, A., Deacon, A. M., Mazur, P., Kozarich, J. W., and Goldman, A. (2002) The structure of *Neurospora crassa* 3-carboxy-cis,cis-muconate lactonizing enzyme, a beta propeller cycloisomerase, *Structure 10*, 483-492.

112. Mazur, P., Henzel, W. J., Mattoo, S., and Kozarich, J. W. (1994) 3-carboxy-cis,cis-muconate lactonizing enzyme from *Neurospora crassa* - An alternate cycloisomerase motif, *J. Bacteriol. 176*, 1718-1728.

113. Tarighi, S., Wei, Q., Camara, M., Williams, P., Fletcher, M. P., Kajander, T., and Cornelis, P. (2008) The PA4204 gene encodes a periplasmic gluconolactonase (PpgL)

which is important for fitness of *Pseudomonas aeruginosa*, *Microbiology-Sgm 154*, 2979-2990.

114. Tian, C. G., Beeson, W. T., Iavarone, A. T., Sun, J. P., Marletta, M. A., Cate, J. H. D., and Glass, N. L. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*, *Proceedings of the National Academy of Sciences of the United States of America 106*, 22157-22162.

115. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq, *Nat. Methods 5*, 621-628.

116. Claeyssens, M., Vantilbeurgh, H., Tomme, P., Wood, T. M., and McRae, S. I. (1989) Fungal cellulase systems - Comparison of the specificities of the cellobiohydrolases isolated from *Penicillium pinophilum* and *Trichoderma reesei*, *Biochem. J. 261*, 819-825.

117. Martinez, D., Larrondo, L. F., Putnam, N., Gelpke, M. D., Huang, K., Chapman, J., Helfenbein, K. G., Ramaiya, P., Detter, J. C., Larimer, F., Coutinho, P. M., Henrissat, B., Berka, R., Cullen, D., and Rokhsar, D. (2004) Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78, *Nat Biotechnol 22*, 695-700.

118. Frush HL, I. H. (1963) Aldonic acids, lactonization of aldonic acids., In *Methods in Carbohydrate Chemistry* (Whistler, R. L., and Wolfrom, M. L., Eds.), pp 13-14, Academic Press, New York.

119. Vogel, H. (1956) A convenient growth medium for *Neurospora*, *Microbiology Genetics Bulletin 13*, 42-43.

120. Hestrin, S. (1949) The reaction of acetylcholine and other carboxylic acid derivatives with hydroxylamine, and its analytical application, *J. Biol. Chem. 180*, 249-261.

121. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389-3402.

122. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment, *Journal of Molecular Biology 302*, 205-217.

123. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., and Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist, *Nucleic Acids Res. 36*, W465-W469.

124. Tian, C. G., Kasuga, T., Sachs, M. S., and Glass, N. L. (2007) Transcriptional profiling of cross pathway control in *Neurospora crassa* and comparative analysis of the Gcn4 and CPC1 regulons, *Eukaryot. Cell 6*, 1018-1029.

125. Li, H., Ruan, J., and Durbin, R. (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores, *Genome Res. 18*, 1851-1858.

126. Wei, H., Xu, Q., Taylor, L. E., 2nd, Baker, J. O., Tucker, M. P., and Ding, S. Y. (2009) Natural paradigms of plant cell wall degradation, *Curr Opin Biotechnol 20*, 330-338.

127. Vanden Wymelenberg, A., Gaskell, J., Mozuch, M., Sabat, G., Ralph, J., Skyba, O., Mansfield, S. D., Blanchette, R. A., Martinez, D., Grigoriev, I., Kersten, P. J., and Cullen, D. (2010) Comparative transcriptome and secretome analysis of wood decay fungi Postia placenta and Phanerochaete chrysosporium, *Appl Environ Microbiol 76*, 3599-3610.

128. Tian, C., Beeson, W. T., Iavarone, A. T., Sun, J., Marletta, M. A., Cate, J. H., and Glass, N. L. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*, *Proc Natl Acad Sci U S A 106*, 22157-22162.

129. Bouws, H., Wattenberg, A., and Zorn, H. (2008) Fungal secretomes--nature's toolbox for white biotechnology, *Appl Microbiol Biotechnol 80*, 381-388.
130. Adrio, J. L., and Demain, A. L. (2003) Fungal biotechnology, *Int Microbiol 6*, 191-199.
131. Aden, A., and Foust, T. (2009) Technoeconomic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol, *Cellulose 16*, 535-545.
132. Dunlap, J. C., Borkovich, K. A., Henn, M. R., Turner, G. E., Sachs, M. S., Glass, N. L., McCluskey, K., Plamann, M., Galagan, J. E., Birren, B. W., Weiss, R. L., Townsend, J. P., Loros, J. J., Nelson, M. A., Lambreghts, R., Colot, H. V., Park, G., Collopy, P., Ringelberg, C., Crew, C., Litvinkova, L., DeCaprio, D., Hood, H. M., Curilla, S., Shi, M., Crawford, M., Koerhsen, M., Montgomery, P., Larson, L., Pearson, M., Kasuga, T., Tian, C., Basturkmen, M., Altamirano, L., and Xu, J. (2007) Enabling a community to dissect an organism: overview of the Neurospora functional genomics project, *Adv Genet 57*, 49-96.
133. Zhang, R., Fan, Z., and Kasuga, T. (2011) Expression of cellobiose dehydrogenase from *Neurospora crassa* in *Pichia pastoris* and its purification and characterization, *Protein Expr Purif 75*, 63-69.
134. Harreither, W., Sygmund, C., Augustin, M., Narciso, M., Rabinovich, M. L., Gorton, L., Haltrich, D., and Ludwig, R. (2011) Cellobiose dehydrogenases from ascomycetes: Catalytic properties and classification, *Appl Environ Microbiol*.
135. Fang, J., Qu, Y. B., and Gao, P. J. (1997) Wide distribution of cellobiose-oxidizing enzymes in wood-rot fungus indicates a physiological importance in lignocellulosics degradation, *Biotechnol Tech 11*, 195-197.
136. Hallberg, B. M., Henriksson, G., Pettersson, G., and Divne, C. (2002) Crystal structure of the flavoprotein domain of the extracellular flavocytochrome cellobiose dehydrogenase, *J Mol Biol 315*, 421-434.
137. Zamocky, M., Ludwig, R., Peterbauer, C., Hallberg, B. M., Divne, C., Nicholls, P., and Haltrich, D. (2006) Cellobiose dehydrogenase -- A flavocytochrome from wood-degrading, phytopathogenic and saprotropic fungi, *Curr. Protein Pept. Sci. 7*, 255-280.
138. Cavener, D. R. (1992) GMC oxidoreductases. A newly defined family of homologous proteins with diverse catalytic activities, *J Mol Biol 223*, 811-814.
139. Zamocky, M., Hallberg, M., Ludwig, R., Divne, C., and Haltrich, D. (2004) Ancestral gene fusion in cellobiose dehydrogenases reflects a specific evolution of GMC oxidoreductases in fungi, *Gene 338*, 1-14.
140. Canevascini, G., Borer, P., and Dreyer, J. L. (1991) Cellobiose dehydrogenases of *Sporotrichum (Chrysosporium) thermophile*, *Eur J Biochem 198*, 43-52.
141. Kremer, S. M., and Wood, P. M. (1992) Production of Fenton reagent by cellobiose oxidase from cellulolytic cultures of *Phanerochaete Chrysosporium*, *Eur. J. Biochem. 208*, 807-814.
142. Mason, M. G., Nicholls, P., and Wilson, M. T. (2003) Rotting by radicals--the role of cellobiose oxidoreductase?, *Biochem Soc Trans 31*, 1335-1336.
143. Ayers, A. R., Ayers, S. B., and Eriksson, K. E. (1978) Cellobiose oxidase, purification and partial characterization of a hemoprotein from *Sporotrichum Pulverulentum*, *Eur. J. Biochem. 90*, 171-181.
144. Roy, B. P., Paice, M. G., Archibald, F. S., Misra, S. K., and Misiak, L. E. (1994) Creation of metal-complexing agents, reduction of manganese-dioxide, and promotion of

manganese peroxidase-mediated Mn(III) production by cellobiose-quinone oxidoreductase from *Trametes Versicolor*, *J. Biol. Chem. 269*, 19745-19750.

145. Colot, H. V., Park, G., Turner, G. E., Ringelberg, C., Crew, C. M., Litvinkova, L., Weiss, R. L., Borkovich, K. A., and Dunlap, J. C. (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors, *Proc Natl Acad Sci U S A 103*, 10352-10357.

146. Ninomiya, Y., Suzuki, K., Ishii, C., and Inoue, H. (2004) Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining, *Proc Natl Acad Sci U S A 101*, 12248-12253.

147. Bao, W., Usha, S. N., and Renganathan, V. (1993) Purification and characterization of cellobiose dehydrogenase, a novel extracellular hemoflavoenzyme from the white-rot fungus *Phanerochaete chrysosporium*, *Arch Biochem Biophys 300*, 705-713.

148. Yoshida, M., Igarashi, K., Wada, M., Kaneko, S., Suzuki, N., Matsumura, H., Nakamura, N., Ohno, H., and Samejima, M. (2005) Characterization of carbohydrate-binding cytochrome b(562) from the white-rot fungus *Phanerochaete chrysosporium*, *Appl. Environ. Microbiol. 71*, 4548-4555.

149. Tullius, T. D. (1988) DNA footprinting with hydroxyl radical, *Nature 332*, 663-664.

150. Harris, P. V., Welner, D., McFarland, K. C., Re, E., Poulsen, J. C. N., Brown, K., Salbo, R., Ding, H. S., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S., and Lo Leggio, L. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family, *Biochemistry 49*, 3305-3316.

151. Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M. F., Synstad, B., and Eijsink, V. G. H. (2005) The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation, *J. Biol. Chem. 280*, 28492-28497.

152. Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H. K., Eijsink, V. G. H., and van Aalten, D. M. F. (2005) Crystal structure and binding properties of the *Serratia marcescens* chitin-binding protein CBP21, *J. Biol. Chem. 280*, 11313-11319.

153. Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sorlie, M., and Eijsink, V. G. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides, *Science 330*, 219-222.

154. Sun, J., Phillips, C. M., Anderson, C. T., Beeson, W. T., Marletta, M. A., and Glass, N. L. (2011) Expression and characterization of the Neurospora crassa endoglucanase GH5-1, *Protein Expr Purif 75*, 147-154.

155. Phillips, C. M., Iavarone, A. T., and Marletta, M. A. (2011) Quantitative Proteomic Approach for Cellulose Degradation by Neurospora crassa, *J Proteome Res*.

156. Eastwood, D. C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., Asiegbu, F. O., Baker, S. E., Barry, K., Bendiksby, M., Blumentritt, M., Coutinho, P. M., Cullen, D., de Vries, R. P., Gathman, A., Goodell, B., Henrissat, B., Ihrmark, K., Kauserud, H., Kohler, A., LaButti, K., Lapidus, A., Lavin, J. L., Lee, Y. H., Lindquist, E., Lilly, W., Lucas, S., Morin, E., Murat, C., Oguiza, J. A., Park, J., Pisabarro, A. G., Riley, R., Rosling, A., Salamov, A., Schmidt, O., Schmutz, J., Skrede, I., Stenlid, J., Wiebenga, A., Xie, X., Kues, U., Hibbett, D. S., Hoffmeister, D., Hogberg, N., Martin, F., Grigoriev, I. V., and Watkinson, S. C. (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi, *Science 333*, 762-765.

157. MacDonald, J., Doering, M., Canam, T., Gong, Y., Guttman, D. S., Campbell, M. M., and Master, E. R. (2011) Transcriptomic responses of the softwood-degrading white-rot fungus Phanerochaete carnosa during growth on coniferous and deciduous wood, *Appl Environ Microbiol 77*, 3211-3218.

158. Phillips, C. M., Iavarone, A. T., and Marletta, M. A. (2011) Quantitative Proteomic Approach for Cellulose Degradation by Neurospora crassa, *J Proteome Res*, In press.

159. Harreither, W., Sygmund, C., Augustin, M., Narciso, M., Rabinovich, M. L., Gorton, L., Haltrich, D., and Ludwig, R. (2011) Catalytic properties and classification of cellobiose dehydrogenases from ascomycetes, *Appl Environ Microbiol 77*, 1804-1815.

160. Westermark, U., and Eriksson, K. E. (1975) Purification and properties of cellobiose: quinone oxidoreductase from Sporotrichum pulverulentum, *Acta Chem Scand B 29*, 419-424.

161. Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C., and Sandgren, M. (2008) The first structure of a glycoside hydrolase family 61 member, Cel61B from Hypocrea jecorina, at 1.6 A resolution, *J Mol Biol 383*, 144-154.

162. Harris, P. V., Welner, D., McFarland, K. C., Re, E., Navarro Poulsen, J. C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S., and Lo Leggio, L. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family, *Biochemistry 49*, 3305-3316.

163. Zhang, L., Koay, M., Maher, M. J., Xiao, Z., and Wedd, A. G. (2006) Intermolecular transfer of copper ions from the CopC protein of Pseudomonas syringae. Crystal structures of fully loaded Cu(I)Cu(II) forms, *J Am Chem Soc 128*, 5834-5850.

164. Himes, R. A., and Karlin, K. D. (2009) Copper-dioxygen complex mediated C-H bond oxygenation: relevance for particulate methane monooxygenase (pMMO), *Curr Opin Chem Biol 13*, 119-131.

165. Freimund, S., and Kopper, S. (2004) The composition of 2-keto aldoses in organic solvents as determined by NMR spectroscopy, *Carbohydr Res 339*, 217-220.

166. Klinman, J. P. (2006) The copper-enzyme family of dopamine beta-monooxygenase and peptidylglycine alpha-hydroxylating monooxygenase: resolving the chemical pathway for substrate hydroxylation, *J Biol Chem 281*, 3013-3016.

167. Solomon, E. I., Ginsbach, J. W., Heppner, D. E., Kieber-Emmons, M. T., Kjaergaard, C. H., Smeets, P. J., Tian, L., and Woertink, J. S. (2011) Copper dioxygen (bio)inorganic chemistry, *Faraday Discuss 148*, 11-39.

168. Beeson, W. T., Iavarone, A. T., Hausmann, C. D., Cate, J. H., and Marletta, M. A. (2011) Extracellular Aldonolactonase from *Myceliophthora thermophila*, *Appl Environ Microbiol 77*, 650-656.

169. Phillips, C. M., Beeson, W. T., Cate, J. H., and Marletta, M. A. (2011) Cellobiose Dehydrogenase and a Copper-Dependent Polysaccharide Monooxygenase Potentiate Cellulose Degradation by Neurospora crassa, *ACS Chem Biol*.

170. Yoshida, M., Igarashi, K., Wada, M., Kaneko, S., Suzuki, N., Matsumura, H., Nakamura, N., Ohno, H., and Samejima, M. (2005) Characterization of carbohydrate-binding cytochrome b562 from the white-rot fungus Phanerochaete chrysosporium, *Applied and environmental microbiology 71*, 4548-4555.

171. Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunaes, A. C., Stenstrom, Y., Mackenzie, A., Sorlie, M., Horn, S. J., and Eijsink, V. G. (2011) Cleavage of cellulose by a CBM33 protein, *Protein Sci 20*, 1479-1483.

172. Vanden Wymelenberg, A., Gaskell, J., Mozuch, M., Sabat, G., Ralph, J., Skyba, O., Mansfield, S. D., Blanchette, R. A., Martinez, D., Grigoriev, I., Kersten, P. J., and Cullen, D. (2010) Comparative transcriptome and secretome analysis of wood decay fungi Postia placenta and Phanerochaete chrysosporium, *Applied and environmental microbiology 76*, 3599-3610.