Insect science

Insect Science (2010) 17, 253–264, DOI 10.1111/j.1744-7917.2010.01323.x

ORIGINAL ARTICLE

Identification of proteins involved in lignocellulose degradation using in gel zymogram analysis combined with mass spectroscopy-based peptide analysis of gut proteins from larval Asian longhorned beetles, *Anoplophora glabripennis*

Scott M. Geib[1], Ming Tien[1] and Kelli Hoover[2]

Departments of [1]Biochemistry and Molecular Biology and [2]Entomology, The Pennsylvania State University, University Park, Pennsylvania, USA

Abstract Enzyme activities toward lignocellulose substrates were analyzed in the gut of larval Asian longhorned beetle (*Anoplophora glabripennis*). Total protein was extracted from gut contents of wild collected larvae from an invasive population in Worchester, MA, USA. From these protein extracts, lignocellulolytic activities were measured (β-1,4-endoglucanase, β-1,4-glucosidase and birch wood xylanase). β-1,4-glucosidase activity was 0.075 μmol glucose/mg protein per min, endoglucanase activity was measured at 0.41 μmol glucose/mg protein per min and xylanase activity was 0.058 μmol xylose/mg protein per min. To identify specific enzymes that may provide these activities, zymogram analysis was performed to detect enzymes active toward carboxymethyl cellulose (CMC), 4-methylumbelliferyl-β-D-glucopyranoside and birch wood xylan. Three protein bands were found to be active toward CMC, three displayed β-1,4-glucosidase, and one displayed xylanase activity. Proteins from active bands from these zymograms were then identified by in-gel trypsin digestions followed by peptide identification by matrix-assisted laser desorption ionization – time of flight – time of flight mass spectrometry (MS). A custom *A. glabripennis* transcriptome database was used for peptide identification, giving highly significant matches in all MS analyses. These matches were then searched against the National Center for Biotechnology Information database to provide annotation to the transcripts and provide possible classification. From these analyses, we were able to detect enzymes active toward cellulose and xylan, and proteins putatively involved in lignocellulose degradation in the gut of this wood-feeding insect. Future research will be focused on characterizing these enzymes through cloning and expression experiments and understanding how the lignocellulose degradation system functions in the gut of this insect.

Key words beta-glucosidase, cellulase, Cerambycidae, endoglucanase, proteomics, xylanase

Introduction

The 30 × 30 Biofuels Initiative set by the US Department of Energy is to replace 30% of the current gasoline consumption with biofuels by 2030. It is expected that a major contributor to biofuel stock will be cellulosic ethanol. While cellulose represents the largest source of

Correspondence: Scott M. Geib, 408 Althouse Laboratory, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA. Tel: 814 865 0658; fax: 814 863 7024; email: smg283@psu.edu

renewable carbon on earth, accessibility to the sugar component required to produce ethanol is limited by the high costs associated with extracting (depolymerizing) sugars from a complex lignocellulose substrate that is heavily cross-linked and crystalline. This study aims to identify enzymes produced in the gut of a xylophagous insect, the Asian longhorned beetle (*Anoplophora glabripennis*) that are involved in lignocellulose degradation. Mining this environment for such enzymes can potentially yield new enzymes for processing lignocellulolytic material into cellulosic ethanol. Little is known about how wood-feeding beetles, specifically longhorned beetles (Family: Cerambycidae), degrade lignocellulose. Thus, the potential for discovery of novel enzymes from these species is considerable. Larval *A. glabripennis* grow and develop on the inner-wood of living hardwood tree species. This insect is unusual because it feeds in a wide diversity of species of both healthy and weakened (dead/dying) trees (Hanks, 1999). The beetles' main food source contains fully polymerized, cross-linked lignocellulose. Lignocellulose degradation was hypothesized to be carried out, at least in part, by gut symbionts. We found that the *A. glabripennis* gut harbors a wide diversity of microbes, many of which may produce enzymes capable of degrading lignocellulose (Geib *et al.*, 2008, 2009).

Wood is composed of three polymeric materials: cellulose, hemicellulose and lignin. Cellulose is a linear polymer of glucose linked by β-1,4 glycosidic bonds, accounting for approximately 45% of wood by weight. Its linear structure and extensive hydrogen bonding increases crystallinity of this macromolecule and decreases the accessibility of hydrolytic enzymes. Hemicellulose accounts for approximately 25% of wood by weight and is also linked by β-1,4 linkages. Unlike cellulose, hemicellulose has much greater structural heterogeneity, containing sugars (the most predominate being xylose), sugar acids and acetyl esters as side groups from the linear polymer. These side groups prevent efficient packing of the hemicellulose fibrils and render hemicelluloses non-crystalline. Lignin imparts structural rigidity in woody biomass with phenylpropanoid units being the precursors of lignin (van Rensburg *et al.*, 2000). Oxidation of these phenols yields free radicals, which undergo radical coupling to form a polymer linked by over 12 types of chemical bonds. It is the random nature of lignin cross-linking and its condensed and insoluble properties that makes lignin resistant to most forms of microbial attack.

Cellulose digestion occurs widely in different taxa of insects (Martin, 1983) and the proportion of ingested cellulose digested can be extremely high (up to 99%) (Prins & Kreulen, 1991). Cellulolytic enzymes may originate from gut symbionts, ingestion of enzymes produced by wood decay fungi, the insect itself, or some combination of these (Martin, 1983; Breznak & Brune, 1994; Watanabe & Tokuda, 2001; Brune, 2003; Suh *et al.*, 2005). Overall, very little is known about how cellulose is digested in insects, with the exception of termites or other insects living in decayed wood. Some longhorned beetles make their own endoglucanases and β-glucosidases, but to-date no insects have been found to produce exoglucanases (Sugimura *et al.*, 2003; Lee *et al.*, 2004).

It is well known that the digestibility of cellulose decreases as the lignin content of plant biomass increases (Hatfield *et al.*, 1999). Some wood-eating insects are able to overcome the lignin barrier, and obtain help with cellulose digestion by feeding on wood that was previously degraded by environmental fungi (Kukor *et al.*, 1988). In fact, in a group of higher termites (Termitidae, Macrotermitidae), the colony actually cultivates a basidiomycete fungus within the nest (Hyodo *et al.*, 2000; Taprab *et al.*, 2005). Basidiomycetes are known to cause white-rot decay in wood and can produce lignin and Mn peroxidases, as well as laccases, which degrade and depolymerize the lignin macromolecule (Taprab *et al.*, 2005). This preconditioning permits greater access to cellulose. Other insects feed on dead, rotting wood colonized with white or brown rot fungi. It was previously thought that substantial lignin degradation did not occur within the gut of wood-feeding insects (Ohkuma, 2003). However, recently we reported that lignin is modified during passage of wood through the *A. glabripennis* gut, as well as the gut of the Pacific dampwood termite (Geib *et al.*, 2008). While the biochemical process of lignin degradation was partially characterized, we do not know what enzymes are involved in overcoming the lignin barrier in *A. glabripennis*. Herein, we examine the lignocellulolytic ability of larval *A. glabripennis*. Enzyme assays were performed on total gut enzyme extracts. In addition, zymogram analyses toward hydrolysis of different lignocellulose components were performed along with in-gel trypsin digestion/peptide sequencing to identify specific enzymes involved in wood degradation in this insect gut.

Materials and methods

Insect collection

Anoplophora glabripennis were field collected in conjunction with eradication efforts by the US Department of Agriculture Animal and Plant Health Inspection Service Plant Protection and Quarantine (USDA-APHIS-PPQ). Infested trees were located by the presence of exit holes and dieback of the trees. Several trees were

cut for this study in January 2009, from a single small stand of mature silver maple (*Acer saccharinum*) on a property in urban Worcester, MA, US (42°18′18″N, 71°48′00″W). Trees were cut into short (0.5–1.0 meter) bolts and transported to the Pennsylvania State University Quarantine Research Lab following our USDA permit guidelines. Due to the time of year that the wood was collected, the larvae were not actively feeding so the bolts were stored in an aluminum screen cage within our quarantine facility and maintained at room temperature (≈22°C) for 2 months to allow the insects to resume feeding when collected for dissection, as evidenced by frass (feces) production from the bolts. At this point the bolts were split and late instar larva were removed for dissection. Larvae were immediately put on ice after removal from trees and dissected.

Anoplophora glabripennis *larval gut dissection and protein extraction*

Larvae removed from trees were immediately chilled and dissected within 1 h in a laminar flow hood to maintain sterility using sterile dissection tools. Before dissection, larvae were surface-sterilized in 70% ethanol for 1 min and rinsed in sterile water. To remove the entire, intact gut, the cuticle was cut laterally and the gut was ligated at either end to isolate an intact gut section between the anterior midgut and posterior hindgut. This portion was then carefully transferred to a new sterile dissection dish. The gut contents were then removed by carefully cutting the gut open and transferring the contents into a microfuge tube containing 0.05 mol/L sodium citrate buffer, pH 5.5. Gut contents from 15 larvae were combined to obtain sufficient protein. The sample was then homogenized with a micropestel, centrifuged at 12 000 *g* for 5 min, and the supernatant was transferred to a new tube. This supernatant was used as gut enzyme extract. Three replicate samples were created in this manner.

Cellulase and xylanase assays

Activities of β-1,4-glucosidase, β-1,4-endoglucanase and β-1,4-xylanase were measured from *A. glabripennis* gut content enzyme extracts incubated with cellulose or xylan substrates based on the release of reducing sugars measured by the dinitrosalicylic acid (DNS) assay (Bernfeld, 1955; Miller, 1959). Each assay was performed in triplicate. Protein concentration was measured using the Bradford method (Bradford, 1976; Bollag *et al.*, 1996) using bovine serum albumin as a protein standard (0–20 μg) and samples were diluted to a working concentration of approximately 40 μg/mL in sodium citrate buffer (50 mmol/L, pH 5.5). For β-glucosidase activity, 750 μL of a 2% salicin solution (in 50 mmol/L sodium citrate buffer, pH 5.5) was combined with 30 μg (750 μL of 40 μg/mL dilution) of gut enzyme extract. For β-endoglucanase activity 750 μL of a 2% carboxymethyl cellulose (CMC) solution (Kukor *et al.*, 1988) (in 50 mmol/L sodium citrate buffer, pH 5.5) was combined with 750 μL of gut content enzyme extract. For xylanase activity, 750 μL of a 1% xylan solution (in 50 mmol/L sodium citrate buffer, pH 5.5, xylan from birch wood, Sigma-Aldrich, St Louis, MO, US) was combined with 30 μg (750 μL of 40 μg/mL dilution) of enzyme extract. For all assays, background sugars were subtracted from the end product by removing 100 μL from the reaction mixture and measuring sugar content at time 0. Reaction mixtures were incubated at 37°C with 100 μL aliquots removed after 30, 60, 90, 120 and 240 min. For each aliquot, 100 μL DNS reagent were added to halt enzyme activity (Miller, 1959). Samples were then incubated in a boiling water bath for 8 min and absorbance of a 150 μL aliquot was read at 540 nm on a SpectraMax(tm) 190 microplate reader (Molecular Devices Corp., Sunndale, CA, US) along with glucose standards (standard curve, 20–1000 μg).

Zymogram analysis

Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) or native PAGE gels (Laemmli, 1970) were performed with some alterations to detect activity of cellulases, β-glucosidase or xylanases through zymogram techniques (Schwarz *et al.*, 1987; Painbeni *et al.*, 1992; Her *et al.*, 1999; Chavez *et al.*, 2002). For the CMC zymogram, a 12% SDS-PAGE separation gel was poured containing 0.1% carboxymethyl cellulose. *A. glabripennis* gut enzyme extracts were loaded onto half of the gel at three different amounts of protein (7, 20 and 40 μg protein/lane). This loading pattern was repeated on the other half of the gel so that the gel could be cut vertically in half after electrophoresis to produce two identical acrylamide gels. The first half of the gel was stained with colloidal blue stain to visualize proteins. The second half was used for zymogram analysis. For zymogram analysis, the gel was rinsed in sodium citrate buffer (50 mmol/L, pH 5.5) containing 1% Triton X-100 for 1 h at room temperature to remove SDS (Her *et al.*, 1999). This was followed by incubation for 1.5 h in sodium citrate buffer (50 mmol/L, pH 5.5) to allow for enzyme activity against the substrate. Then the gel was stained with 0.1% Congo red for 30 min and destained in 1 mol/L NaCl to reveal

zones of clearing. The gel was imaged under ultraviolet light and aligned with colloidal blue stained gels.

For β-glucosidase and xylanase zymograms, we used native PAGE to maintain protein activity during electrophoresis. Twelve percent native PAGE gels (Laemmli, 1970) were created lacking SDS, and contained either 2 mmol/L 4-methylumbelliferyl-β-D-glucopyranoside, or 0.1% birch wood xylan for the β-glucosidase or xylanase zymograms, respectively (Schwarz *et al.*, 1987; Painbeni *et al.*, 1992; Her *et al.*, 1999; Chavez *et al.*, 2002). *A. glabripennis* gut enzyme extracts were loaded onto half of the gels at three different amounts of protein (7, 20 and 40 μg protein/lane) as described above with half for protein staining and the other half for zymograms. Electrophoresis was carried out in standard Tris-glycine buffer (25 mmol/L Tris, 192 mmol/L glycine, pH 8.8) lacking SDS and the gel was maintained at 4°C during electrophoresis. Zymogram analysis was performed by incubating the gels for 1.5 h in sodium citrate buffer (50 mmol/L, pH 5.5) to allow for enzyme activity against the substrates. At this point, the β-glucosidase zymogram gel was visualized under ultraviolet light to detect fluorescence of 4-methylumbelliferone from the substrate where β-glucosidase activity was present. Xylanase activity was visualized by first staining the gel with 0.1% Congo red for 30 min and destaining in 1 mol/L NaCl to reveal zones of clearing.

Trypsin digestion of selected protein bands and MALDI-TOF-TOF analysis

Bands correlating to activity based on zymogram analysis were excised from reference gels and prepared for matrix-assisted laser desorption ionization – time of flight – time of flight (MALDI-TOF-TOF) analysis by trypsin digestion. Bands of interest (labeled in Figs. 2–4) were excised using a clean razor blade and cut into 1 mm^2 pieces. Gel pieces were destained with 200 μL of 100 mmol/L ammonium bicarbonate (pH 8.0) in 50% acetonitrile, with two repeated 45 min washes at 37°C. After destaining, gel pieces were dried in a SpeedVac (Thermo Scientific, Waltham, MA, US), reduced in 10 mmol/L dithiothreitol in 25 mmol/L ammonium bicarbonate (pH 8.0) for 15 min at 37°C, and alkylated in 20 mmol/L iodoacetamide in 25 mmol/L ammonium bicarbonate (pH 8.0) for 45 min at 37°C in the dark. Three washes with 25 mmol/L ammonium bicarbonate were then performed to remove iodacetamide. The gel pieces were dried again and then rehydrated in 0.02 μg/μL of Trypsin Gold, MS grade (Promega Corporation, Madison, WI, US) in 10% acetonitrile, 40 mmol/L ammonium bicarbonate (pH 8.0) and 0.1% w/v n-octylglucoside. Digestion was performed at 37°C for 18 h. After digestion, peptides were extracted in 50% acetonitrile, 0.1% w/v n-octylglucoside and dried in a SpeedVac. The dried sample was then rinsed three times with 100 μL of water, allowing the samples to dry completely between each wash. After the final wash, the peptides were resuspended in water. Samples were then concentrated and cleaned for mass spectroscopy using a ZipTip SCX pipette tip (Millipore Corporation, Billerica, MA, US).

Samples were analyzed at the Penn State Hershey Medical Center Mass Spectrometry Core Research Facility on an Applied Biosystems 4800 Proteomic Analyzer (MALDI-TOF-TOF). For each sample, the molecular mass of each peptide fragment was measured, and then the amino acid sequence of the top 10 most abundant peptides was determined using ProteinPilot 3.0 software coupled with the MALDI-TOF-TOF analyzer (Applied Biosystems, Foster City, CA, US). Within ProteinPilot, proteins were identified by searching a custom *A. glabripennis* gut transcriptome database consisting of 5 959 assembled contigs (average contig size ~900 bp) from a cDNA library sequenced using a 454 FLX pyrosequencer (Roche Technologies, Branford, CT, US). The transcriptome library was translated into all possible amino acid sequences across all six frames, and this resulting amino acid database was used in the ProteinPilot software. In addition, the *Tribolium castaneum* predicted protein database (obtained from http://beetlebase.org), as well as the termite gut metagenome, *Fusarium solani*, and *Phanerochaete chrysosporium* predicted protein databases (obtained from http://genome.jgi-psf.org) were used, but either no significant matches were obtained from these databases or a higher significance score was obtained when using our custom ALB (Asian longhorned beetle) database. ProteinPilot identifies proteins from double mass spectroscopy (MS/MS) spectra using the Paragon Algorithm, with an "unused" score higher than 1.3 representing a significant protein match (at 95% confidence). For each sample, significant protein matches to the transcriptome database were identified using this method, and the identified contigs from the transcriptome database were searched against the National Center for Biotechnology Information (NCBI) nr database and annotated.

Results

Enzyme activity of A. glabripennis *gut contents*

Activities of several lignocellulolytic enzymes were assayed from gut extracts of larval *A. glabripennis*.

β-1,4-glucosidase activity was 0.075 μmol glucose/mg protein per min based on the release of β-1,4-linked glucose from salicin (Fig. 1A). Endoglucanase activity, determined by reducing sugar release from carboxymethyl cellulose, was 0.41 μmol glucose/mg protein per min (Fig. 1B) and xylanase activity, determined by the release of reducing sugars from birch wood xylan was 0.058 μmol xylose/mg protein per min (Fig. 1C).

Zymogram analyses

Zymogram analysis for cellulase activity as measured by degradation of carboxymethyl cellulose showed that at least three enzymes are responsible for the measured cellulase activity. Three major zones of clearing were observed on the zymogram gel, with sizes of approximately 28, 30 and 40 kDa (Fig. 2, lane C). Examination of the colloidal blue-stained SDS-PAGE gel showed protein bands corresponding with these zones of clearing, with a major band at 28 kDa, a less abundant band at 40 kDa, and a faint band at 30 kDa (Fig. 2, lane B). β-glucosidase zymogram analysis revealed two active bands at approximately 38 and 60 kDa. The birch wood xylan zymogram analysis revealed a single band of activity at approximately 34 kDa (Figs. 3 and 4).

MALDI-TOF-TOF identification of proteins from active bands

Bands corresponding to activity on the endoglucanase, β-glucosidase and xylanase zymograms were excised for peptide analysis (denoted with arrows in Figs. 2–4). MALDI-TOF-TOF analysis run on the Applied Biosystems 4800 Analyzer identifies proteins in a sample by comparing the peptide sequence obtained from the analysis to a database. The significance of a match is calculated by the ProteinPilot software with significant scores having an "unused" value greater than 1.3. Table 1 presents the significant matches for each gel band analyzed, with the percentage coverage of the peptides identified against the matching protein in the database, and the number of peptides that matched the protein at 95% confidence. In addition, the contig length from the transcriptome database, as well as the number of 454 FLX reads that were assembled to create that contig, are listed along with a protein description based on the top NCBI nr BLAST match and various BLAST statistics (Table 1). Also, the corresponding amino acid sequence of each identified protein from the database is listed in FASTA format (supplementary information).

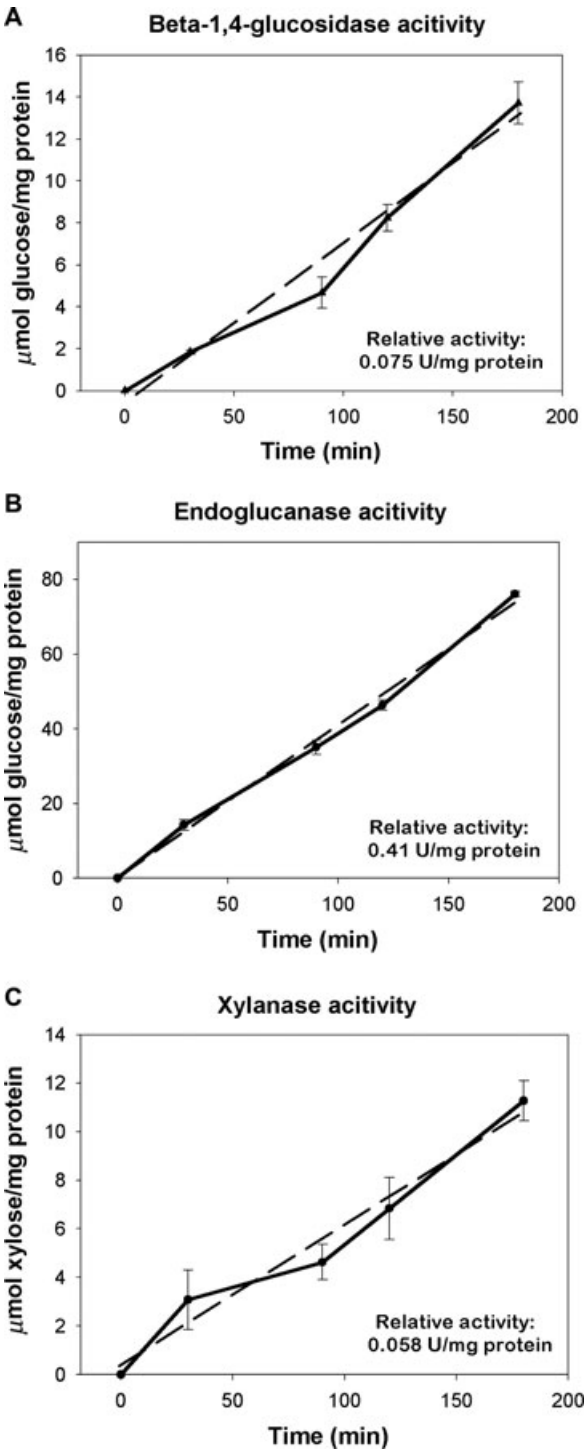

Fig. 1 Cellulase and xylanase activity of *A. glabripennis* gut enzyme extract. Enzyme activity of (A) beta-glucosidase, (B) endoglucanase, and (C) xylanase, measured by release of reducing sugar from each substrate by 30 μg of total gut protein over 180 min. One unit (U) of activity is the amount of enzyme to hydrolyze 1 μmol of reducing sugar per minute.

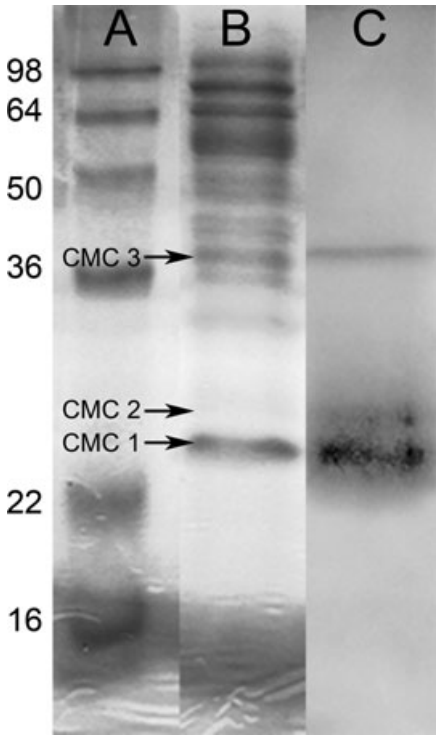

Fig. 2 Carboxymethyl cellulose (CMC) zymogram of *A. glabripennis* larval gut contents. Electrophoresis of 20 μg of *A. glabripennis* gut contents. Protein standard is on the left (lane A), with band sizes labeled (kDa). Lane B is a colloidal blue-stained sample and lane C is the corresponding CMC zymogram stained with Congo red. Dark bands represent zones of clearing (under ultraviolet light). Three major zones of clearing are present at approximately 28, 30 and 40 kDa and matched to protein bands on the colloidal blue-stained lane. Arrows depict corresponding protein bands that were excised for matrix-assisted laser desorption ionization – time of flight – time of flight analysis; band name with protein identification is listed in Table 1.

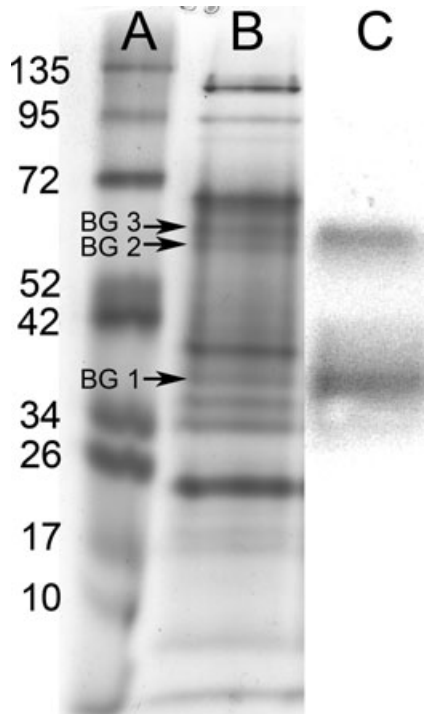

Fig. 3 Beta-glucosidase zymogram of *A. glabripennis* gut contents protein extract using native polyacrylamide gel elecrophoresis. Electrophoresis of 30 μg of *A. glabripennis* gut contents. Protein standard is on the left (lane A), with band sizes labeled (kDa). Lane B is a colloidal blue-stained sample and lane C is the corresponding beta-glucosidase zymogram visualized under ultraviolet (UV) light. Dark bands represent zones of clearing (under UV light). Two major zones of clearing are present at approximately 38 and 60 kDa and matched to protein bands on the colloidal blue-stained lane. Arrows depict corresponding protein bands that were excised for matrix-assisted laser desorption ionization – time of flight – time of flight analysis; band name with protein identification is listed in Table 1.

For the CMC zymogram, none of the three bands on the gel gave protein matches that could be clearly identified as cellulases. The first band was identified as a "thaumatin-like" protein, the second as a "serine proteinase," and the third consisted of three proteins that either had no significant match in the NCBI database or matched to an unknown protein with no annotation.

For the β-glucosidase gel bands, both produced matches to glycoside hydrolases, including β-glucosidases along with other proteins. The band excised from the xylanase zymogram analysis also produced a match to a glycoside hydrolase, as well as a sucrose-6-phosphate hydrolase.

Discussion

In this study, we characterized the enzymes involved in lignocellulose degradation found in the larval stages of *A. glabripennis*. Figure 1 presents the enzyme activity of the Asian longhorned beetle larval gut contents toward β-1,4-endoglucanase, β-1,4-glucosidase and xylanase activity. The enzyme activities measured here in the Asian longhorned beetle are similar to those seen in termites and other cerambycid beetle species (Kukor *et al.*, 1988; Tokuda *et al.*, 2005; Smith *et al.*, 2009).

Zymograms were performed to assess the number of enzymes responsible for the observed activities. The

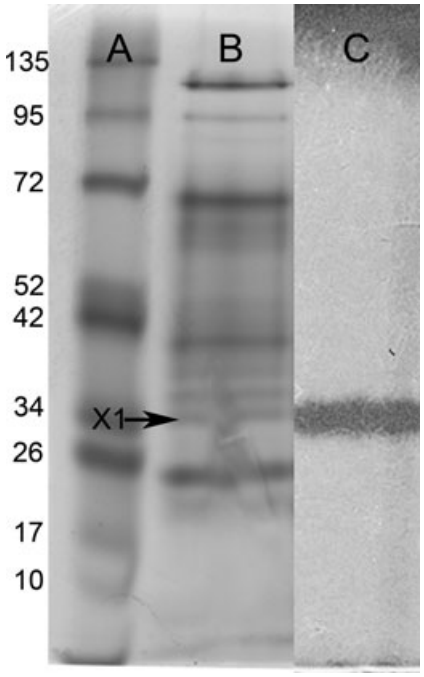

Fig. 4 Birch wood xylan zymogram of *A. glabripennis* gut contents protein extract using native polyacrylamide gel elecrophoresis. Electrophoresis of 30 μg of *A. glabripennis* gut contents. Protein standard is on the left (lane A), with band sizes labeled (kDa). Lane B is a colloidal blue-stained sample and lane C is the corresponding birch wood xylan zymogram stained with Congo red. Dark bands represent zones of clearing (under ultraviolet light). One major zone of clearing is present at approximately 34 kDa and can be matched to a protein band on the colloidal blue-stained lane. An arrow depicts the corresponding protein band that was excised for matrix-assisted laser desorption ionization – time of flight – time of flight analysis; band name with protein identification is listed in Table 1.

zymograms using carboxymethyl cellulose as a substrate showed several active bands, two at approximately 28–30 kDa, and one slightly larger than 36 kDa. This is consistent with several enzymes present in this gut system that are able to act upon non-crystalline cellulose and liberate reducing sugars. The zymograms were of gels separating the proteins only in one dimension. The Coomassie-blue staining of similar gels showed that while the activity bands did not reflect pure proteins, there was still a possibility that they could be identified by MALDI-TOF-TOF. Bands CMC-1, CMC-2 and CMC-3 did not yield any peptides that could be clearly identified to correspond to a cellulase gene. For CMC-1, a single protein type was identified as thaumatin-like proteins, which we would not expect to exhibit hydrolase activity. From band CMC-2 a single proteinase was identified, which is again not likely to be involved in cellulose degradation. CMC-3 identified three unknown proteins. One, contigE00782, had a significant match to a *T. castaneum* protein, NCBI accession number XP_975796, but the function of this protein is not known. The other two proteins matched to contigs E00961 and E04148, which had no significant matches to any sequence in the NCBI database, suggesting that these are previously unknown proteins. While it is possible that one or all of these proteins are involved in cellulose hydrolysis, the more likely explanation is that the proteins of interest in the CMC bands are not the most abundant proteins from the fragments excised. Further separation of these proteins, for example through two-dimensional gel electrophoresis, may yield more pure protein spots, permitting identification of specific cellulase enzymes. Our transcriptome database has several contigs that match to putative cellulase genes. Ideally, we would like to demonstrate their cellulolytic activity through zymogram analysis to confirm the annotation. In other cerambycid species, several families of glycoside hydrolases have been described that have endoglucanase activity. In the mulberry longicorn beetle, *Apriona germari*, three insect endoglucanase genes have been identified, two classified in glycoside hydrolase family 45 and one in family 5 (Lee *et al.*, 2004, 2005). In the yellow-spotted longicorn beetle, *Psacothea hilaris*, an endogenous family 5 endoglucanase has been characterized (Sugimura *et al.*, 2003).

Activity was also observed with zymograms for β-glucosidase and birch wood xylanase. Unlike the zymograms containing SDS described above, the molecular weight obtained from the native gels used here may not accurately reflect the actual mass of the proteins since proteins do not migrate solely based on mass under native conditions (Bollag *et al.*, 1996). From the β-glucosidase zymogram, three fragments were excised from the protein gel for MALDI-MS-MS, which corresponded with the two active zones on the zymogram. Peptide analysis of the first band, BG-1, gave rise to six peptides that matched to the transcriptome dataset giving a significant score ("unused" > 1.3), matching to four different contigs within this dataset. Interestingly, all four of these contigs had the highest BLAST match to the same gene from *Tribolium castaneum*, NCBI accession number XP_972437. This gene is a putative glycoside hydrolase family 1 β-glucosidase based on conserved domains. However, all four contigs from the ALB transcriptome were distinct, suggesting that there may be multiple hydrolase genes present that perform similar functions.

It was not clear which protein band matched the second active zone due to a wide band of activity (Fig. 3), so two bands were excised, BG-2 and BG-3. Analysis of the second band, BG-2, gave rise to 10

Table 1 Proteins identified from in gel digestion/MALDI-TOF-TOF analysis of active bands of zymogram gels.

Gel band	Protein number from band	Unused score	% coverage	% coverage (95% confidence)	# peptides matched (95% confidence)	Transcriptome contig matched	Contig length	# reads in contig	Sequence description	Organism	Hit (NCBI accession #)	E-value	Similarity	Score	Alignment length	Positives
CMC-1	1	2.56	36.6	14.1	1	contigE04500	213	18	Thaumatin-like protein	*Tribolium castaneum*	XP_975175	3.942E-15	64	83.96	65	42
	2	2.28	17.5	10.9	1	contigE00172	434	49	Thaumatin-like protein	*Tribolium castaneum*	XP_975175	1.273E-29	92	132.11	67	62
	3	2	32.9	12.5	1	contigE04351	266	33	Thaumatin-like protein	*Tribolium castaneum*	XP_968724	2.423E-28	80	127.87	87	70
CMC-2	1	5.05	7.8	5.3	2	contigE00651	846	24	Chymotrypsin-like serine proteinase	*Costelytra zealandica*	ABZ04021	2.574E-43	60	179.49	264	159
CMC-3	1	6	22.5	22.5	5	contigE00782	740	13	Unknown protein	*Tribolium castaneum*	XP_975796	8.391E-10	51	67.78	209	107
	2	3.05	8.8	4.3	1	contigE00961	1147	77	No BLAST match							
	3	2	4.5	4.5	1	contigE04148	1153	96	No BLAST match							
BG-1	1	2.85	7.1	4.0	2	contigP00192	768	10	Beta-glucosidase	*Tribolium castaneum*	XP_972437	8.127E-80	74	252.68	129	96
	2	2.01	4.1	2.3	2	contigE00211	1308	71	Glycoside hydrolases	*Tribolium castaneum*	XP_972437	1.47E-143	73	513.46	429	315
	3	2	2.7	2.7	1	contigE04507	1096	118	Beta-glucosidase	*Tribolium castaneum*	XP_972437	5.71E-111	72	404.83	353	255
	4	1.4	3.4	1.5	1	contigE03980	1600	186	Glycoside hydrolases	*Tribolium castaneum*	XP_972437	3.71E-171	74	605.52	499	370
BG-2	1	8	5.9	5.9	4	contigE00216	2492	176	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967103	0	63	677.17	735	464
	2	4	6.2	4.0	2	contigP00027	1505	101	Carboxylesterase	*Tribolium castaneum*	XP_972251	1.73E-90	59	337.42	495	296
	3	2.06	7.1	4.5	1	contigE04148	1153	96	No BLAST match							
	4	2	5.3	5.3	1	contigE00651	846	24	Chymotrypsin-like serine proteinase	*Costelytra zealandica*	ABZ04021	2.574E-43	60	179.49	264	159
	5	1.4	2.7	2.7	2	contigE03986	2692	93	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967022	0	66	463.00	408	270
BG-3	1	8	19.1	19.1	4	contigE00961	1147	77	No BLAST match							
	2	3.3	10.0	10.0	2	contigE00651	846	24	Chymotrypsin-like serine proteinase	*Costelytra zealandica*	ABZ04021	2.574E-43	60	179.49	264	159
	3	2	4.4	1.5	1	contigE04156	1556	203	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967103	5.15E-170	70	601.67	521	368

Continued

Table 1 Continued

Gel band	Protein number from band	Unused score	% coverage	% coverage (95% confidence)	# peptides matched (95% confidence)	Transcriptome contig matched	Contig length	# reads in contig	Sequence description	Organism	Hit (NCBI accession #)	E-value	Similarity	Score	Alignment length	Positives
	4	2	5.3	2.0	1	contigE04111	1198	261	Carboxylesterase	*Tribolium castaneum*	XP_970253	3.263E-70	60	269.63	383	231
	5	2	1.9	1.9	1	contigP00021	1585	49	Carboxylesterase	*Tribolium castaneum*	XP_970253	1.67E-123	64	447.20	506	326
X-1	1	4.19	9.2	6.1	2	contigP00237	1261	25	Sucrose-6-phosphate hydrolase	*Streptococcus pneumoniae* SP3-BS71	ZP_01817526	1.34E-100	70	321.63	168	119
	2	2	6.3	6.3	1	contigP01396	475	20	Glycoside hydrolases	*Tribolium castaneum*	XP_972437	5.323E-52	79	206.45	159	127
H-1	1	5.7	10.7	8.8	3	contigP01221	1125	24	Gram negative bacteria binding protein 1		XP_970010	4.95E-142	76	508.06	368	281
	2	4	14.3	14.3	3	contigE00782	740	13	Unknown protein	*Tribolium castaneum*	XP_975796	8.391E-10	51	67.78	209	107
	3	3.87	5.0	2.7	2	contigE00144	2635	85	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967022	0	68	704.90	427	294
	4	2	12.5	12.5	1	contigP00184	241	2	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967103	1.51E-14	68	82.03	80	55
H-2	1	3.55	6.4	4.2	2	contigE00151	1523	303	Plasma alpha-l-fucosidase	*Tribolium castaneum*	XP_971511	0	80	639.80	432	348
H-3	1	5.07	5.9	4.0	2	contigE00144	2635	85	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967022	0	68	704.90	427	294
	2	2	8.6	8.6	1	contigP01299	492	17	Chitinase 3	*Monochamus alternatus*	BAF49605	1.403E-47	75	191.82	137	103
	3	2	2.1	2.1	1	contigP01221	1125	24	Gram-negative bacteria binding protein 1		XP_970010	4.95E-142	76	508.06	368	281
	4	2	12.5	12.5	1	contigP00184	241	2	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967103	1.51E-14	68	82.03	80	55
	5	1.7	8.4	8.4	1	contigP01217	467	4	No BLAST match							
	6	1.3	4.4	2.9	1	contigE04156	1556	203	Neutral alpha-glucosidase ab	*Tribolium castaneum*	XP_967103	5.15E-170	70	601.67	521	368

BLAST, basic local alignment search tool.

peptides that matched to five contigs in the dataset. Of these, contigE00216 had the highest score (8.0) with four of the peptides matching to this contig. Searching this contig against NCBI gave a match to a glycoside hydrolase family 31 α-glucosidase from *T. castaneum*, NCBI accession number XP_967103. In addition, a second contig, E03986 matched to another family 31 α-glucosidase, NCBI accession number XP_967022. Analysis of protein band BG-3 identified contigE04156 matching to the same *T. castaneum* α-glucosidase as the first protein identified in BG-2. These findings suggest that there are at least three β-glucosidases in the *A. glabripennis* gut that appear to have activity toward β-1,4 linkages. Known β-glucosidases fall into three glycoside hydrolase families, GH 1, 3 and 9, but animal-derived β-glucosidases have only been identified in families 1 and 9. In other coleopteran species, the number of β-glucosidases identified to date have been relatively large. For example, two β-glucosidases were identified from *Ergates faber* with sizes of 57 and 70 kDa (Chararas *et al.*, 1983). In *Tenebrio molitor*, two β-glucosidase genes were identified, βGly1 and βGly2, which fall into GH family 1 (Ferreira *et al.*, 2001).

Only one single insect-derived xylanase gene has been cloned. It encodes a 75 kDa protein in GH family 11 derived from *Phaedon cochleariae* (Girard & Jouanin, 1999). Several insect-associated microbes have been shown to produce xylanases used by insects, which include but are not limited to protists in termite gut systems (Arakawa *et al.*, 2009), bacteria associated with beetles (Zhou *et al.*, 2009), and fungal species associated with siricids (Kukor & Martin, 1983). Analysis of the protein band that matched to hydrolysis of birch wood xylan identified two contigs. Of these, contig P01396 matched to a glycoside hydrolase in *T. castaneum*, NCBI accession number XP_972437. This is the same gene that matched to the proteins in band BG-1, but the match in this case had a lower score. While the database suggests that this is a β-glucosidase, our zymogram analysis suggests that the activity is directed toward β-1,4-xylan.

We demonstrated here that larval Asian longhorned beetles have enzymes present in their gut that aid in the degradation of both xylan and cellulose. Zymogram analysis was used to visualized specific activities of these enzymes; coupled with tandem mass spectrometry analysis of peptide digestions we were able to identify specific novel proteins involved in wood degradation in this system. These results will lead to future studies, where these targeted proteins that have lignocellulose activity will be cloned, expressed and characterized as done previously in other insect and invertebrate systems (Smant, 1998; Sugimura *et al.*, 2003; Suzuki *et al.*, 2003; Lee *et al.*, 2004, 2005). Expression and characterization of specific enzymes will allow for comparison of enzyme activity and substrate specificity compared to other known insect cellulases/xylanases, as well as microbial and fungal-derived enzymes. In addition, analysis of combinations of enzymes, for example cellulases and beta-glucosidases together, to identify the complexes that are most efficient in conversion of lignocellulose are of interest. Together, these future studies will lead toward a better understanding of how insect-derived enzymes can be utilized in industrial lignocellulose processing and if there are advantages over microbial systems.

Acknowledgments

We thank A. Sawyer's group at USDA APHIS and K. Gooch at the Massachusetts Department of Conservation and Recreation for assistance collecting beetles from Worcester, MA and A. and B. Stanley at the Penn State Hershey Medical Center Mass Spectrometry Core Research Facility for mass spectrometry analysis. Funding for this project was provided by USDA-NRI-CRSEES 2008-35504-04464, USDA-NRI-CREES 2009-35302-05286, and the Alphawood Foundation, Chicago, IL.

References

Arakawa, G., Watanabe, H., Yamasaki, H., Maekawa, H. and Tokuda, G. (2009) Purification and molecular cloning of xylanases from the wood-feeding termite, *Coptotermes formosanus* Shiraki. *Bioscience, Biotechnology, and Biochemistry*, 73, 710–718.

Bernfeld, P. (1955) Amylases alpha and beta. *Methods in Enzymology* (eds. S.P. Colowick & N.O. Kaplans), pp. 149–150. Academic Press, New York.

Bollag, D.M., Rozycki, M.D. and Edelstein, S.J. (1996) *Protein Methods*, Wiley-Liss, New York. 415 pp.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72, 248–254.

Breznak, J.A. and Brune, A. (1994) Role of microorganisms in the digestion of lignocellulose by termites. *Annual Review of Entomology*, 39, 453–487.

Brune, A. (2003) Symbionts aiding digestion. *Encyclopedia of Insects* (eds. V. Resh & R. Cardes), pp. 1102–1107. Academic Press, New York.

Chararas, C., Eberhard, R., Courtois, J.E. and Petek, F. (1983) Purification of three cellulases from the zylophageous

larvae of *Ergates faber* (Coleoptera: Cerambycidae). *Insect Biochemistry*, 13, 213–218.

Chavez, R., Schachter, K., Navarro, C., Peirano, A., Aguirre, C., Bull, P. and Eyzaguirre, J. (2002) Differences in expression of two endoxylanase genes (*xynA* and *xynB*) from *Penicillium purpurogenum*. *Gene*, 293, 161–168.

Ferreira, A.H.P., Marana, S.R., Terra, W.R. and Clélia, F. (2001) Purification, molecular cloning, and properties of a beta-glycosidase isolated from midgut lumen of *Tenebrio molitor* (Coleoptera) larvae. *Insect Biochemistry and Molecular Biology*, 31, 1065–1076.

Geib, S.M., Filley, T.R., Hatcher, P.G., Hoover, K., Carlson, J.E., Jimenez-Gasco, M.D.M., Nakagawa-Izumi, A., Sleighter, R.L. and Tien, M. (2008) Lignin degradation in wood-feeding insects. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 12932–12937.

Geib, S.M., Jimenez-Gasco, M.D.M., Carlson, J.E., Tien, M. and Hoover, K. (2009) Effect of host tree species on cellulase activity and bacterial community composition in the gut of larval Asian longhorned beetle. *Environmental Entomology*, 38, 686–699.

Girard, C. and Jouanin, L. (1999) Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. *Insect Biochemistry and Molecular Biology*, 29, 1129–1142.

Hanks, L.M. (1999) Influence of the larval host plant on reproductive strategies of cerambycid beetles. *Annual Review of Entomology*, 44, 483–505.

Hatfield, R.D., Ralph, J. and Grabber, J.H. (1999) Cell wall structural foundations: Molecular basis for improving forage digestibilities. *Crop Science*, 39, 27–37.

Her, S., Lee, H.-S., Choi, S.-J., Choi, S.-W., Choi, H.-J., Yoon, S.-S. and Oh, D.-H. (1999) Cloning and sequencing of β-1,4-endoglucanase gene (*celA*) from *Pseudomonas* sp. YD-15. *Letters in Applied Microbiology*, 29, 389–395.

Hyodo, F., Inoue, T., Azuma, J.I., Tayasu, I. and Abe, T. (2000) Role of the mutualistic fungus in lignin degradation in the fungus-growing termite *Macrotermes gilvus* (Isoptera; Macrotermitinae). *Soil Biology & Biochemistry*, 32, 653–658.

Kukor, J.J., Cowan, D.P. and Martin, M.M. (1988) The role of ingested fungal enzymes in cellulose digestion in the larvae of cerambycid beetles. *Physiological Zoology*, 61, 364–371.

Kukor, J.J. and Martin, M.M. (1983) Acquisition of digestive enzymes by siricid woodwasps from their fungal symbiont. *Science*, 220, 1161–1163.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680–685.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Sohn, H.D. and Jin, B.R. (2004) cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comparative Biochemistry and Physiology, B: Comparative Biochemistry*, 139, 107–116.

Lee, S.J., Lee, K.S., Kim, S.R., Gui, Z.Z., Kim, Y.S., Yoon, H.J., Kim, I., Kang, P.D., Sohn, H.D. and Jin, B.R. (2005) A novel cellulase gene from the mulberry longicorn beetle, *Apriona germari*: Gene structure, expression, and enzymatic activity. *Comparative Biochemistry and Physiology, B: Comparative Biochemistry*, 140, 551–560.

Martin, M. (1983) Cellulose digestion in insects. *Comparative Biochemistry and Physiology*, 75A, 313–324.

Miller, G.L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry*, 31, 426–428.

Ohkuma, M. (2003) Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Applied Microbiology and Biotechnology*, 61, 1–9.

Painbeni, E., Valles, S., Polaina, J. and Flors, A. (1992) Purification and characterization of a *Bacillus polymyxa* beta-glucosidase expressed in *Escherichia coli*. *Journal of Bacteriology*, 174, 3087–3091.

Prins, R.A. and Kreulen, D.A. (1991) Comparative aspects of plant-cell wall digestion in insects. *Animal Feed Science and Technology*, 32, 101–118.

Schwarz, W.H., Bronnenmeier, K., Grabnitz, F. and Staudenbauer, W.L. (1987) Activity staining of cellulases in polyacrylamide gels containing mixed linkage β-glucans. *Analytical Biochemistry*, 164, 72–77.

Smant, G. (1998) Endogenous cellulases in animals: Isolation of β-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 4906–4911.

Smith, J.A., Scharf, M.E., Pereira, R.M. and Koehler, P.G. (2009) Changes in *Reticulitermes flavipes* (Kollar) gut cellulolytic activities in response to hardwood, softwood and cellulose diets. *Journal of Entomological Science*, 44, 256–263.

Sugimura, M., Watanabe, H., Lo, N. and Saito, H. (2003) Purification, characterization, cDNA cloning and nucleotide sequencing of a cellulase from the yellow-spotted longicorn beetle, *Psacothea hilaris*. *European Journal of Biochemistry*, 270, 3455–3460.

Suh, S.O., Mchugh, J.V., Pollock, D.D. and Blackwell, M. (2005) The beetle gut: a hyperdiverse source of novel yeasts. *Mycological Research*, 109, 261–265.

Suzuki, K., Ojima, T. and Nishita, K. (2003) Purification and cDNA cloning of a cellulase from abalone *Haliotis discus hannai*. *European Journal of Biochemistry*, 270, 771–778.

Taprab, Y., Johjima, T., Maeda, Y., Moriya, S., Trakulnaleamsai, S., Noparatnaraporn, N., Ohkuma, M. and Kudo, T. (2005)

Symbiotic fungi produce laccases potentially involved in phenol degradation in fungus combs of fungus-growing termites in Thailand. *Applied and Environmental Microbiology*, 71, 7696–7704.

Tokuda, G., Lo, N. and Watanabe, H. (2005) Marked variations in patterns of cellulase activity against crystalline- vs. carboxymethyl-cellulose in the digestive systems of diverse, wood-feeding termites. *Physiological Entomology*, 30, 372–380.

Van Rensburg, H., Anterola, A.M., Levine, L.H., Davin, L.B. and Lewis, N.G. (2000) Monolignol compositional determinants in loblolly pine: Aromatic amino acid metabolism and associated rate-limiting steps. *Lignin: Historical, Biological, and Materials Perspectives* (eds. W.G. Glasser, R.A. Northey, & T.P. Schultz), pp. 118–144. American Chemical Society, Washington, DC.

Watanabe, H. and Tokuda, G. (2001) Animal cellulases. *Cellular and Molecular Life Sciences*, 58, 1167–1178.

Zhou, J., Huang, H., Meng, K., Shi, P., Wang, Y., Luo, H., Yang, P., Bai, Y., Zhou, Z. and Yao, B. (2009) Molecular and biochemical characterization of a novel xylanase from the symbiotic *Sphingobacterium* sp. TN19. *Applied Microbiology and Biotechnology*, 85, 323–333.

Accepted December 15, 2009