Journal of Insect Physiology 57 (2011) 300–306

Contents lists available at ScienceDirect

Journal of Insect Physiology

journal homepage: www.elsevier.com/locate/jinsphys

Identification, cloning, and expression of a GHF9 cellulase from *Tribolium castaneum* (Coleoptera: Tenebrionidae)

Jonathan D. Willis [a], Brenda Oppert [b], Cris Oppert [a], William E. Klingeman [c], Juan L. Jurat-Fuentes [a,*]

[a] Department of Entomology and Plant Pathology, University of Tennessee, 2431 Joe Johnson Drive, 205 Ellington Plant Sciences Building, Knoxville, TN 37996-4560, USA
[b] USDA-ARS Center for Grain and Animal Health Research, Manhattan, KS 66502, USA
[c] Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996-4560, USA

ARTICLE INFO

Article history:
Received 28 September 2010
Received in revised form 23 November 2010
Accepted 23 November 2010
Available online 30 November 2010

Keywords:
Cellulase
Tribolium castaneum
GHF9
Cellulolytic activity
Endoglucanase
Biofuels

ABSTRACT

The availability of sequenced insect genomes has allowed for discovery and functional characterization of novel genes and proteins. We report use of the *Tribolium castaneum* (Herbst) (red flour beetle) genome to identify, clone, express, and characterize a novel endo-β-1,4-glucanase we named TcEG1 (*T. castaneum* endoglucanase 1). Sequence analysis of a full-length TcEG1 cDNA clone (1356 bp) revealed sequence homology to enzymes in glycosyl hydrolase family 9 (GHF9), and verified presence of a change (Gly for Ser) in the conserved catalytic domain for GHF9 cellulases. This TcEG1 cDNA clone was predicted to encode a 49.5 kDa protein with a calculated pI of 5.39. Heterologous expression of TcEG1 in *Drosophila* S2 cell cultures resulted in secretion of a 51-kDa protein, as determined by Western blotting. The expressed protein was used to characterize TcEG1 enzymatic activity against two cellulose substrates to determine its specificity and stability. Our data support that TcEG1 as a novel endo-β-1,4-glucanase, the first functional characterization of a cellulase enzyme derived from an insect genome with potential applications in the biofuel industry due to its high relative activity at alkaline pH.

1. Introduction

Advances in DNA sequencing in the past decade have resulted in the completion of several insect genome and transcriptome sequencing projects. The availability of insect genomes has greatly facilitated large-scale genomic analyses and discovery and functional characterization of new families of proteins. The *Tribolium castaneum* (Herbst) genome (Richards et al., 2008) has been used as a tool to perform functional genomic studies (Siebert et al., 2008; Trauner et al., 2009; Wang et al., 2008). The *T. castaneum* genome has also allowed identification of genes and proteins with relevant roles in diverse physiological processes (Morris et al., 2009).

One of the physiological processes poorly characterized that can greatly benefit from availability of completed insect genomes is insect carbohydrate digestion (Kunieda et al., 2006). Genomic studies on cellulase gene expression in termites suggest that they are not only important for digestion of lignocellulosic biomass, but are directly involved in development and caste differentiation (Scharf et al., 2003, 2005; Weil et al., 2007). With growing need for novel cellulolytic enzymes to serve as biocatalyst for efficient conversion of cellulose to bioethanol, interest has increased in insect systems involved in cellulose digestion (Watanabe and Tokuda, 2010). The complete degradation of plant-derived cellulose into glucose, followed by fermentation to ethanol, requires the synergistic actions of endo-β-1,4-glucanases (EG; EC. 3.2.1.4), exo-β-1,4-cellobiohydrolases (CBH; EC. 3.2.1.91), and β-glucosidases (EC. 3.2.1.21) (Clarke, 1997). Although insect cellulose digestion has traditionally been hypothesized to occur through enzymes produced by symbiotic gut microbes in insects (Martin et al., 2001; Martin, 1983), current evidence also supports a crucial role for endogenous cellulolytic enzymes in some insects (Nakashima et al., 2002; Tokuda et al., 2007; Zhou et al., 2008). While discovery of insect cellulolytic enzymes is usually challenged by complex purification protocols, several endogenous insect cellulases have been identified using genomic data, including cDNA libraries and EST datasets (Girard and Jouanin, 1999; Kim et al., 2008; Lee et al., 2004, 2005; Wei et al., 2006b; Yuki et al., 2008). Data gleaned from sequenced genomes have also allowed identification and characterization of novel cellulolytic enzyme genes in *Apis mellifera* (Kunieda et al., 2006), *Acyrthosiphon*

Abbreviations: CMC, carboxymethyl cellulose; MCC, microcrystalline cellulose; EG, endo-β-1,4-glucanase; CBH, exo-β-1,4-cellobiohydrolases; DNSA, 3,5-dinitrosalicylic acid.

* Corresponding author. Tel.: +1 865 974 5931; fax: +1 865 974 4744.
E-mail address: jurat@utk.edu (J.L. Jurat-Fuentes).

0022-1910/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jinsphys.2010.11.019

pisum (Sabater-Munoz et al., 2006), and *T. castaneum* (Morris et al., 2009). We recently identified a putative cellulase through sequence homology searches using the *T. castaneum* genome (Morris et al., 2009). Adults and larvae of this insect are devastating pests of stored grain and feed on cereals and grains that have coatings consisting of chaff and bran, which may help explain the need for cellulolytic enzymes in their digestive system. In fact, examination of the *T. castaneum* genome identifies many glycosyl hydrolase homologs, some of which are predicted to be expressed in the gut (Morris et al., 2009). The putative *T. castaneum* cellulase identified by genome surveying displayed similarity to cellulases found in the *A. mellifera* genome that are proposed to be involved in digestion of the inner wall of pollen grains (Kunieda et al., 2006).

In this report, we describe the cloning and expression of the *T. castaneum* cellulase gene and further characterization of the activity of this enzyme. Since this gene represents the first predicted endo-β-1,4-glucanase reported for *T. castaneum*, we named it as *TcEG1* (for *T. castaneum* endoglucanase 1). We cloned and expressed the full-length *TcEG1* cDNA in cultured *Drosophila* S2 insect cells as a secreted protein. Recombinant TcEG1 demonstrated activity on carboxymethyl cellulose (CMC), but not on microcrystalline cellulose (MCC), providing evidence for typical endoglucanase activity. Our data represent the first functional characterization of a predicted insect cellulase derived using a genome sequence survey, thereby demonstrating the utility of genomic datasets for functional identification of novel enzymes that may have utility for degradation of lignocellulosic plant biomass during bioethanol synthesis.

2. Materials and methods

2.1. Insects

Adults and larvae of *T. castaneum* were obtained from the USDA-ARS Center for Grain and Animal Health Research in Manhattan, Kansas. Insects were fed wheat flour and oats and were maintained in an incubator at 26 °C in darkness with 70% relative humidity. Under these conditions larvae required 2–3 weeks to reach their 6th stadium used for dissections to purify RNA.

2.2. Cloning and construction of TcEG1 expression cassette

Sixth-instar *T. castaneum* larvae were collected and anesthetized on ice for 10 min before dissection. Head capsules, which contain salivary and labial glands, were isolated and placed into individual micro-centrifuge tubes (10 per tube) with 100 μl of RNAlater (Ambion). Tubes containing head tissues were incubated at 4 °C overnight to allow RNAlater to infiltrate and preserve the RNA. Tubes were quick-spun and RNAlater removed, followed by flash freezing and mortar pestle grinding. RNA was extracted using the RNeasy kit (Qiagen) following manufacturer's instructions, and the product was evaluated on a 1% agarose gel and stored at −80 °C until used. Concentrations of RNA were estimated using the Quant-iT RNA assay kit (Invitrogen) following manufacturer's instructions.

Superscript III reverse transcriptase (Invitrogen) was used to synthesize cDNA from RNA collected from head. Primers for TcEG-1 were designed for amplification of the full-length cDNA sequence and to assist with proper insertion into the pIZT/V5-His (Invitrogen) expression vector so that a 6×His tag was added to the C-terminus of the expressed protein. Cloning was facilitated by EcoRI and NotI restriction enzyme sites engineered in primers used for amplification (underlined): TcpIZTFwd 5′-GGAATTCGATGTTC-TACTCATTGTGGGTGCTACTATTT-3′ and TcpIZTRev 5′- ATAGTT-TAGCGGCCGCCAATTTATTCTCATTTTCAATATAAAT-3′. Accuprime Pfx supermix (Invitrogen) was used for PCR reactions containing 5% DMSO, 20 mmol of each primer, and 1 μl of cDNA. The PCR protocol consisted of an initial 5 min at 95 °C, followed by 30 cycles of 30 s at 95 °C, 30 s at 60 °C, 90 s at 68 °C, and a final 10 min extension at 68 °C. The PCR products were cloned into TOPO TA (Invitrogen) and constructs transformed into *E. coli* TOP10 competent cells (Invitrogen). Positive clones were determined by sequencing transformants in forward and reverse directions (University of Tennessee Sequencing Facility). Both the native and His-tagged TcEG1 inserts in the TOPO plasmid were excised using EcoRI and NotI digestion and gel-purified (QIAquick Gel Extraction Kit, Qiagen). Purified inserts were ligated using T4 DNA ligase (Invitrogen) into pIZT/V5-His to generate the pIZT/V5/TcEG1-His expression cassette. Constructs were transformed into *E. coli* TOP10 cells (Invitrogen) and positive transformants as well as reading frames were confirmed by sequencing in forward and reverse directions (University of Tennessee Sequencing Facility). Plasmid for insect cell transfection was produced and purified using a HiPure Plasmid Maxiprep kit (Invitrogen).

2.3. Sequence alignment and phylogenetic analysis

The TcEG1 sequence (Glean 15370) was annotated from the *T. castaneum* genome (Morris et al., 2009) and was translated by ExPASy translate and analyzed with the compute pI/MW tool to predict isoelectric point (pI) and molecular weight (MW) of the predicted protein (http://expasy.org/tools). DictyOGlyc 1.1 (http://www.cbs.dtu.dk/services/DictyOGlyc) and NetNGlyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc) software packages were used to predict O- and N-glycosylation sites, respectively. Signal peptide prediction was performed using SignalP 3.0 (Bendtsen et al., 2004) (http://www.cbs.dtu.dk).

For phylogenetic comparisons, we performed a BLASTP search (Altschul et al., 1997) of the NCBInr database (accessed September 15, 2010) limiting by organism group to the Insecta Class, using the TcEG1 protein sequence as query, to find insect homologs of this protein. Sequences with *e* values $<e^{-100}$ (corresponding to about 54% sequence identity) were selected for alignment and tree construction. Predicted endogenous cellulase sequences from coleopteran species in the NCBInr database that were not returned by the BLASTp search (*Psacothea hilaris*, *Apriona germari* and *Phaedon cochlearia*) were also included in sequence alignments for phylogenetic and molecular evolutionary analyses. Retrieved sequences were aligned with the predicted TcEG1 protein sequence for construction of a maximum parsimony phylogenetic tree using MEGA version 4 (Tamura et al., 2007).

For prediction of the three dimensional (3D) structure of TcEG1, we search structural databases using the Phyre server (http://www.sbg.bio.ic.ac.uk/~phyre/index.cgi) to identify accurate models (Kelley and Sternberg, 2009). The cellulase from *Nasutitermes takasagoensis* (Accession number BAA33708.1) was identified as optimum model and used to determine alterations in protein structure due to a specific Gly to Ser change observed in the TcEG1 sequence compared to other GHF9 members. Prediction of phosphorylation of the ^{70}Ser residue in TcEG1 was performed using Phos 2.0 (www.cbs.dtu.dk/services/NetPhos/).

2.4. Transient expression of TcEG1 in Drosophila S2 cell cultures

Cultured *Drosophila* S2 cells were maintained in serum-free insect cell medium (HyQ SFX-Insect, Hyclone) and transfected as previously described (Jurat-Fuentes and Adang, 2006). Briefly, approximately 2×10^6 S2 cells from a confluent culture were suspended in 2 ml fresh media and allowed to adhere overnight to surface-treated 60 × 15 mm polystyrene dishes (Falcon). Plasmid transfection mixtures were prepared by mixing either 4 μg of pIZT, or 5 μg of pIZT/V5/TcEG1-His-plasmids with 1 ml of serum-free

insect medium and 20 μl of Cellfectin reagent (Invitrogen). Transfection mixtures were incubated with cells at 20 °C for 4 h, after which fresh media (3 ml final volume) was added to the plates. Cell cultures were incubated at 26 °C for three days before collection and centrifugation at room temperature (6000 × *g* for 3 min). Cell pellets and supernatant media were retained to test for TcEG1 expression. Supernatant media was concentrated approximately 10-fold using a SpeedVac (Thermo Scientific). The S2 cell pellets were lysed by resuspension in native purification buffer (25 mM NaH_2PO_4 pH 8.0, 250 mM NaCl) and two freeze-thaw cycles. After lysis, samples were cleared by centrifugation as above and supernatants collected and stored at −80 °C until used.

2.5. Detection of TcEG1-His expression

Expression of His-tagged TcEG1 expressed in S2 cell cultures was detected by Western blotting using antisera against the 6×His tag. Protein concentrations in S2 cell lysate and media supernatant were estimated using the Coomassie Protein Assay Reagent (Pierce) following manufacturer's instructions using bovine serum albumin (BSA) as standard. Proteins (1 μg) in S2 cell lysate or media supernatant were solubilized in 2× sample buffer (Laemmli, 1970) and heat denatured for 20 min. Proteins were separated by electrophoresis on SDS–10%PAGE gels, then electro-transferred to polyvinylidene difluoride (PVDF) filters overnight at 4 °C under constant voltage (20 V). Filters were blocked with 3% BSA in PBST (135 mM NaCl, 2 mM KCl, 10 mM Na_2HPO_4, 1.7 mM KH_2PO_4, pH 7.5, 0.1% Tween-20) for 1 h at room temperature. After blocking, filters were probed with a 1:10,000 dilution of antisera (Novus Biologicals) against 6×His tag conjugated to horseradish peroxidase (HRP) for 1 h in blocking buffer. After several washes with PBST and 0.1% BSA, blots were developed with the Supersignal West Pico chemiluminescence substrate (Pierce).

2.6. TcEG1 enzymatic activity: thermostability and pH optima

Gut and head digestive fluid samples were obtained from dissected tissues of *T. castaneum* larvae by homogenization in distilled water and centrifugation as previously described (Willis et al., 2010). Supernatants were collected as digestive fluid samples and protein concentrations estimated as for the S2 cell pellets above. Cellulase activity in media and cell pellets from S2 cell cultures expressing TcEG1 and mock-transfected cell cultures was determined using a modified dinitrosalicylic acid (DNSA) assay (Miller 1959) with carboxymethyl cellulose sodium salt (CMC, Sigma–Aldrich) or microcrystalline cellulose (MCC, Acros Organics) as substrates. Protein samples (50 μg and 150 μg for CMC and MCC assays, respectively) or glucose standards (ranging from 0 to 313 μg) were added to substrate solutions (2%, w/v in 50 mM sodium citrate buffer, pH 6.0) and incubated for 1 h (CMC) or 2 h (MCC) at 50 °C. A modified DNSA reagent containing Rochelle salt (Miller, 1959) was added to samples to halt enzymatic activity, after which color was developed at 100 °C for 15 min. Samples were centrifuged at 2000 × *g* for 2 min to precipitate any remaining substrate. Supernatants were transferred to polystyrene microplates and spectral absorbance at 595 nm read on a Synergy HT microplate reader (BioTek) using the KC4 software (v. 3.1). Background amounts of reducing sugars were corrected by subtracting final from initial values of the calculated reducing sugars in the sample. One unit of cellulolytic activity is defined as the amount of enzyme that produces 1 μmol of reducing sugar (glucose equivalents) min^{-1} at 50 °C and pH 6.0. As previously described, we used a commercial preparation of endocellulase from *Aspergillus niger* as positive internal control for cellulolytic activity (Oppert et al., 2010). Specific activities were reported as units per g of protein and represent averages from experiments using three independent recombinant enzyme batches. Statistical significance was tested by the *t*-test using the SigmaPlot v11.0 software (Systat Software IN, IL).

For thermal stability assays of recombinant TcEG1 enzymes secreted by S2 cells in culture supernatants (25 μg total protein) were incubated at 30 °C, 40 °C, 50 °C, 60 °C, or 70 °C for 10 min before testing in DNSA assays as described above using CMC as substrate. The effect of pH on TcEG1 activity was measured by incubating recombinant proteins from S2 cell cultures (10 μg) in 2% (w/v) CMC dissolved in buffer at pH 2.0 (50 mM maleic acid), pH 5.0 (50 mM citrate buffer), pH 7.0 (50 mM phosphate buffer), or pH 8.5 (50 mM glycine buffer), and CMC degradation at 50 °C for 1 h detected and measured as above. All absorbance readings represent averages from at least three experiments with material from independent recombinant enzyme or larval batches (biological replicates) performed in triplicate (technical replicates). Statistical significance was tested through one-way analysis of variance (ANOVA) and pairwise multiple comparison procedures (Holm-Sidak method, overall $\alpha = 0.05$) using the SigmaPlot v11.0 software (Systat Software IN, IL).

3. Results and discussion

T. castaneum is one of the most economically significant storage cereal grain pests worldwide. Extensive genomic and proteomic resources, including a complete genome sequence, have been generated for this insect. In turn, these resources have helped advance *T. castaneum* as an insect model for developmental studies (Richards et al., 2008). Using these resources, we previously described the expression of a putative cellulase, now named TcEG1, in the head and gut tissue of *T. castaneum* larvae (Morris et al., 2009).

The full-length cDNA (1,356 bp) from the *TcEG1* gene encodes a predicted 50 kDa protein (TcEG1) with a pI of 5.39. Sequence analysis revealed that TcEG1 contains a signal peptide at the N-terminus that targets the protein for secretion (Fig. 1A). Three N-(^{148}NMTR^{151}, ^{184}NQSY^{187}, and ^{335}NISQ^{338}) and one O-(^{304}Thr) glycosylation sites were predicted for TcEG1 (Fig. 1A). Protein alignment followed by phylogenetic and molecular evolutionary analyses revealed no TcEG1 homologues among described predicted endogenous cellulases in Coleoptera. Proteins displaying highest homology (between 60% and 65%) were predicted endoglucanases from Hemiptera (*Acyrthosiphon pisum* XP_001944774.1), Hymenoptera (*A. mellifera* XP_396791.2 and *Nasonia vitripennis* XP_001606454.1), and Phthiraptera (*Pediculus humanus corporis* XP_002426465). Interestingly, all these predicted EGs were found by genome annotation (Kunieda et al., 2006; Sabater-Munoz et al., 2006). Lower sequence identity was found for cellulases from diverse species of Isoptera and Dictyoptera (data not shown). The uniqueness of the TcEG1 cellulase was reflected in sequence alignments and corresponding phylogenetic trees (Fig. 1B). In these analyses, bootstrap values supported that TcEG1 is not related to any of the described coleopteran cellulases, and was more closely related to cellulases from alternative taxonomic insect orders.

Insect cellulases displaying high similarity to TcEG1 belong to the GHF9 family. This glycosyl hydrolase family consists of EG, CBH, and EC enzymes from prokaryotic and eukaryotic sources, which present a typical 3D structure containing 6 alpha helices (Khademi et al., 2002). The majority of identified and cloned insect EGs belong to this family, which is considered as evidence for vertical transfer of these genes from an ancient insect species (Davison and Blaxter, 2005; Watanabe and Tokuda, 2010). The fact that TcEG1 represents the first GHF9 enzyme reported for a coleopteran species supports the existence of GHF9 cellulases in a common insect ancestor. The conserved catalytic domain for GHF9

A

```
(1)   MFYSLWVLLFITFGSAYDYRHALKLSLLFYEAQRSGPLPASNRITWRSDS   (50)
(51)  ALDDHGLNNEDLTGGYYDASDYVKFSFTMAFTTTILTWGMLSFRHAYNEA   (100)
(101) GQYQHGLDAIKWATDYFIKCHASKYEFYGQIGDFSLDHAFWGRPEEMNMT   (150)
(151) RPAYKIDIEHPGSDLAGEAVAALAAASLLFENFNQSYSDELLRHATELYD   (200)
(201) FATMYRGLYHDSIPGAKVYYESTGYGDELTWAAIWLYKKTKDVKYIEQAE   (250)
(251) TFYTKFRIKDRPNEFFYNKKVAGIQLLLAELTLRPEYISTIKNFCDYTIK   (300)
(301) EQIRTPKGLIFIDKSGTLSHAANVAFICLHAGLTLNISQAAYVSFAKEQI   (350)
(351) NYMLGSTGQSFVVGYGQNYPKQPHHSASSCPNLPEPCGWKQFTWKGPNPQ   (400)
(401) ILYGALVSGPDQNDHYEDVREEFLYNEVTLDYNAGFQSTLAGLIYIENENKL (452)
                                  *
```

B

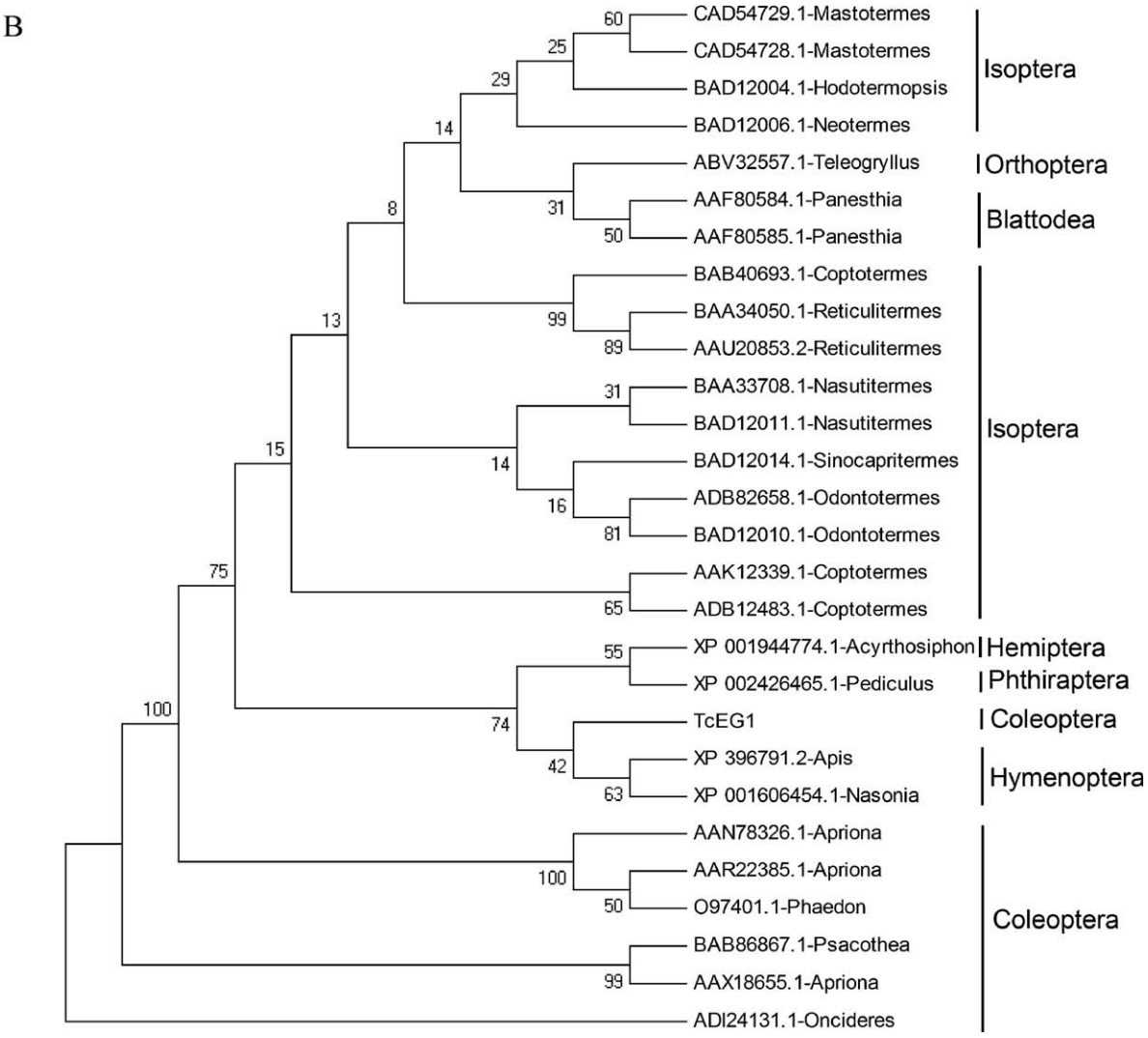

Fig. 1. Predicted features of the TcEG1 protein sequence (A) and phylogenetic tree comparing TcEG1 to alternative insect GHF9 cellulases (B). (A) TcEG1 sequence with predicted signal peptide for secretion highlighted in grey. Predicted N-glycosylation and O-glycosylation sites are either underlined or highlighted in black, respectively. Square region denotes proton–donor catalytic region. Proton acceptor is Glu with an asterisk below. (B) Maximum parsimony tree for amino acid sequences of TcEG1 and reported insect endogenous cellulases. NCBI GenBank database sequence accession number, genera, and taxonomic order are presented for each sequence. Numbers on the branches represent bootstrap values for 1000 replicates.

members is present in TcEG1 (^{68}Asp and ^{71}Asp and the ^{427}Gln proton acceptor at the C-terminus). However, a conserved serine residue at position 70 among GHF9 members is substituted by a glycine in TcEG1 (Fig. 2). Glycine residues are known to contribute to conformational flexibility of loops associated with enzymatic catalysis, thus a Gly to Ser change in the TcEG1 catalytic site may introduce conformational restrictions, which sometimes affect enzymatic function (Okoniewska et al., 2000). Analysis of the TcEG1 sequence with protein structure prediction software (Kelley and Sternberg, 2009) returned a significant structural match (100% estimated precision, *e* value = 0) to an endo-β-1,4-glucanase from *Nasutitermes takasagoensis* (structural classification of proteins code d1ks8a). Structural alignment of TcEG1 and d1ks8a revealed that the ^{68}DASD^{71} catalytic site is predicted to be part of a long coil connecting two beta strands, and that this secondary structure is not affected by the Ser to Gly change found in TcEG1. Phosphorylation of serine residues can contribute to alterations in the protein structure, however analysis of ^{70}Ser in TcEG1 revealed that this residue is predicted to not to be phosphorylated. Future research using site-specific mutations would be necessary

```
                                                    **  *        *          **
Acyrthosiphon pisum PREDICTED XP 001944774.1    (91) DAGD (327) KTP (448) NEV
                  Apis mellifera XP 396791.2   (113) DAGD (349) RTP (472) TEV
       Coptotermes acinaciformis AAK12339.1    (69) DAGD (305) KTP (426) NEV
          Coptotermes formosanus ACI45756.1    (69) DAGD (305) KTP (426) NEV
          Nasonia vitripennis XP 001606454.1   (92) DAGD (328) RTR (451) TEV
       Nasutitermes takasagoensis BAA33708.1   (69) DAGD (305) KTP (426) NEV
    Nasutitermes takasagoensis NtEG BAA76619.1 (69) DAGD (305) KTP (426) NEV
       Nasutitermes walkeri NwEG BAA33709.1    (69) DAGD (305) KTP (426) NEV
             Panesthia cribata AAF80584.1      (72) DAGD (308) RTP (429) NEV
             Panesthia cribata AAF80585.1      (70) DAGD (306) RTP (427) NEV
Reticulitermes flavipes endogenous AAU20853.2  (69) DAGD (305) KTP (426) NEV
      Reticulitermes speratus EG2 BAA34050.1   (69) DAGD (305) KTP (426) NEV
  Reticulitermes speratus salivary BAA31326.1  (69) DAGD (305) KTP (426) NEV
          Teleogryllus emma ABV32557.1         (74) DAGD (310) KTP (431) NEV
          Teleogryllus emma ACA04897.1         (74) DAGD (310) KTP (431) NEV
            Tribolium castaneum TcEG-1         (68) DA[S]D (304) RTP (426) NEV
```

Fig. 2. Sequence alignment of GHF 9 enzymes from insect systems. Black arrow denotes proton–donor catalytic region, (*) denotes sequence conservation. The catalytic domain mutation detected for TcEG1 (Gly) is denoted by a square. Conserved O-glycosylation (Thr) is denoted by a black circle; proton acceptor (Glu) is denoted by a grey arrow.

to examine the implications of this substitution with regard to catalytic activity of TcEG1.

Previously TcEG1 expression was shown to be higher in head compared to gut tissue in gene expression analysis using microarrays and with quantitative RT-PCR (Morris et al., 2009). Based on this information, we designed specific PCR primers based on the genomic *TcEG1* sequence and used the primers to amplify the full length cDNA encoding TcEG1 from larval *T. castaneum* head. The full length *TcEG1* cDNA was expressed in *Drosophila* S2 insect cell cultures (Fig. 3A). We chose this heterologous system for expression based on the reported dependence of insect cellulase activity on glycosylation (Wei et al., 2006a) and glycosylation ability in S2 cells (Benting et al., 2000). In agreement with the predicted molecular weight for TcEG1 containing a secretory signal peptide, antisera to the 6×His tag in expressed TcEG1 detected a protein of about 50-kDa in media from S2 cell cultures transfected with the pIZT/V5/TcEG1-His construct (Fig. 3A). In contrast, no TcEG1 expression was detected in the S2 cell lysates or in samples from mock transfected cells.

Cellulolytic activity of heterologously-expressed TcEG1 was compared to samples from mock-transfected cell cultures and gut fluids from *T. castaneum* larvae. Two different cellulase substrates, CMC and MCC, were used as proxies for detection of either endoglucanase (Xiao et al., 2005) or exoglucanase activity (Chundawat et al., 2008), respectively. Activity against MCC was not detectable for any of the tested recombinant samples (Fig. 3B). In contrast, media from S2 cell cultures expressing TcEG1 displayed activity against CMC (Fig. 3B), validating TcEG1 as an endoglucanase. Fluids from *T. castaneum* larval gut and head tissue were previously found to display detectable EG activity (Oppert et al., 2010), which suggests expression of TcEG1 and possibly other cellulolytic enzymes in this beetle. Activity levels of TcEG1 activity against CMC substrate (12.9 U/mg) were much lower than reported for *Apriona germari* (992 U/mg) and *T. emma* (3118.4 U/mg) endoglucanases expressed in Sf9 cells (Kim et al., 2008; Lee et al., 2004; Wei et al., 2005). Since similar methods were used for these activity assays, differences probably reflect differences in protein production or purification methods. Alternatively, it is also possible that these differences may represent lack of optimized glycosylation or reflect diverse efficiencies in degrading cellulosic material between these insect species.

Once we obtained TcEG1 secreted from S2 cells, we characterized cellulase activity against CMC using thermostability and pH optima assays. In the thermostability assays, activity against CMC in TcEG1 was highest at 50 °C, then sharply decreased as temperatures were increased (Fig. 4A). Similar thermostability data have been reported for other cellulases from coleopteran (Lee et al., 2004; Lee et al., 2005; Sami and Shakoori, 2008; Wei et al., 2006b) and orthopteran (Kim et al., 2008) species, while lower thermostability has been reported for some termite cellulases (Zhang et al., 2009). Although TcEG1 activity against CMC was almost negligible at pH 2, the activity increased from pH 5 to 8.5,

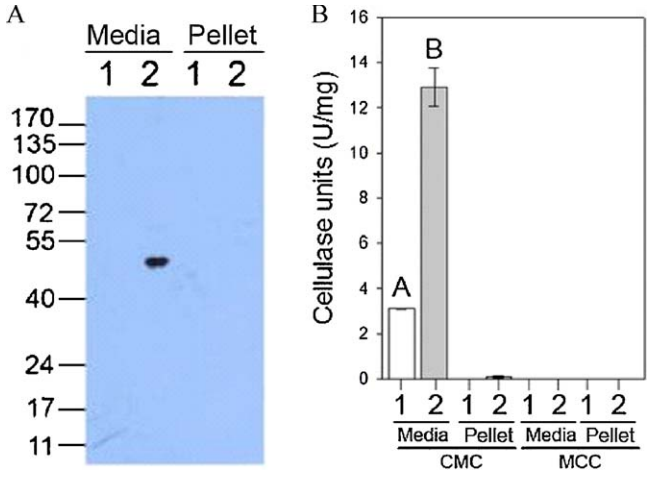

Fig. 3. Expression of TcEG1 in *Drosophila* S2 cell cultures and detection of cellulolytic activity. (A) Expression of TcEG1 in S2 cell cultures. Proteins in media supernatant (Media) or cell pellets (Cells) from S2 cell cultures transfected with pIZT/V5-His (lanes 1), or pIZT/V5/TcEG1-His (lanes 2) were separated by electrophoresis and blotted on PVDF filters. Expressed TcEG1 was detected by probing with antisera against the C-terminal 6×His tag on TcEG1. Molecular weight markers (in kDa) are presented on the side of the image for TcEG1 mass estimation. (B) Comparison of activity against CMC or MCC cellulose substrates in TcEG1 protein expressed in S2 cell cultures. Media from S2 cells transfected with pIZT/V5-His (columns labeled 1) or pIZT/V5/TcEG1-His (columns labeled 2). Data for each experiment represents the mean and standard deviation from three independent determinations of CMC degradation from three independent biological samples. Activity for commercial cellulase from *A. niger* (internal control for cellulase activity) was 292 U/mg against CMC and 0.636 U/mg against MCC substrates (data not shown in graph). Also not shown in the graph are levels of activity against CMC in digestive fluids from head and guts of *T. castaneum* larvae, which were 0.024 U/mg and 0.016 U/mg, respectively. Bars display mean values and standard errors calculated from three biological replicates. Different letters on top of columns represent significant differences (*t*-test, $P < 0.05$).

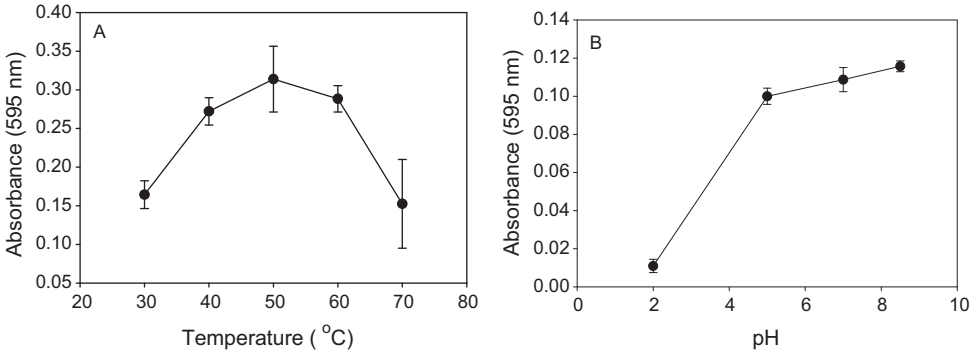

Fig. 4. Determination of thermostability and pH optima for TcEG1. (A) Thermostability of activity against CMC substrate in crude extracts of native or His-tagged TcEG1 (TcEG1 and TcEG1-His, respectively). (B) Dependence of activity of TcEG1 toward CMC on pH of the solution. Activity was tested using buffers of different pH, as indicated in the figure. Data on the *Y*-axis represents absorbance at 595 nm corrected for background signal of cell culture media at each pH tested.

the highest pH tested (Fig. 4B). By contrast, insect cellulases characterized in other reports have displayed highest activities at more acidic pH levels, often between pH 4 and 6 (Kim et al., 2008; Lee et al., 2004; Lee et al., 2005; Wei et al., 2006b). An endoglucanase from *Aulacophora foveicollis* (Lucas) had a slightly alkaline (pH 7.8) pH optimum (Sami and Shakoori, 2008), and a cellulase from *C. formosanus* retained 70% activity at pH 9 (Zhang et al., 2009). To our knowledge, the high relative activity at pH 8.5 found for TcEG1 is unique among described insect cellulases, which may indicate greater molecular stability (Boer and Koivula, 2003). This feature of TcEG1 may have potential utility for development of industrial enzymes that would efficiently hydrolyze lignocellulosic biomass at alkaline pH (Qin et al., 2008; Wang et al., 2005). For example, ionic liquids with high pH conditions have been suggested for accelerated lignocellulose degradation (Zhao et al., 2009), but currently available cellulolytic enzymes are not stable under these conditions. Further testing is needed to determine TcEG1 functionality under different biorefinery conditions, as well as to identify specific regions of this enzyme amenable to engineering for optimal performance in biorefineries.

Our data represent the first functional evidence for an EG from *T. castaneum*. Future research on TcEG1 will focus on localizing expression of the protein in the *T. castaneum* digestive tract, and evaluation of conditions affecting change in expression, such as different diets. If TcEG1 activity is critical for larval survival, it could be exploited as a potential insecticide target using RNAi, as previously demonstrated for the *Cell-1* cellulase gene from *Reticulitermes flavipes* (Kollar) (Zhou et al., 2008), or inhibitors, as reported for *A. foveicollis* cellulase (Pant and Ramana, 1989). Optimization of TcEG1 via genetic engineering could enhance activity and thermostability leading to incorporation of TcEG1 in enzyme mixtures for bioreactors used to degrade lignocellulosic material for production of ethanol. Alternatively, specific regions of TcEG1 could be incorporated in other cellulases for enhanced performance at alkaline pH conditions.

Acknowledgements

This research was supported by a grant from the Southeastern Sun Grant Center to JLJF and BO. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

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

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. Journal of Molecular Biology 340, 783–795.

Benting, J., Lecat, S., Zacchetti, D., Simons, K., 2000. Protein expression in *Drosophila* Schneider cells. Analytical Biochemistry 278, 59–68.

Boer, H., Koivula, A., 2003. The relationship between thermal stability and pH optimum studied with wild-type and mutant *Trichoderma reesei* cellobiohydrolase Cel7A. European Journal of Biochemistry 270, 841–848.

Chundawat, S.P., Balan, V., Dale, B.E., 2008. High-throughput microplate technique for enzymatic hydrolysis of lignocellulosic biomass. Biotechnology and Bioengineering 99, 1281–1294.

Clarke, A.J., 1997. Biodegradation of Cellulose: Enzymology and Biotechnology. Technomic Pub. Co., Lancaster, PA.

Davison, A., Blaxter, M., 2005. Ancient origin of glycosyl hydrolase family 9 cellulase genes. Molecular Biology and Evolution 22, 1273–1284.

Girard, C., 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.

Jurat-Fuentes, J.L., Adang, M.J., 2006. The *Heliothis virescens* cadherin protein expressed in *Drosophila* S2 cells functions as a receptor for *Bacillus thuringiensis* Cry1A but not Cry1Fa toxins. Biochemistry 45, 9688–9695.

Kelley, L.A., Sternberg, M.J., 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols 4, 363–371.

Khademi, S., Guarino, L.A., Watanabe, H., Tokuda, G., Meyer, E.F., 2002. Structure of an endoglucanase from termite, *Nasutitermes takasagoensis*. Acta Crystallographica Section D Biological Crystallography 58, 653–659.

Kim, N., Choo, Y.M., Lee, K.S., Hong, S.J., Seol, K.Y., Je, Y.H., Sohn, H.D., Jin, B.R., 2008. Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket *Teleogryllus emma*. Comparative Biochemistry and Physiology 150B, 368–376.

Kunieda, T., Fujiyuki, T., Kucharski, R., Foret, S., Ament, S.A., Toth, A.L., Ohashi, K., Takeuchi, H., Kamikouchi, A., Kage, E., Morioka, M., Beye, M., Kubo, T., Robinson, G.E., Maleszka, R., 2006. Carbohydrate metabolism genes and pathways in insects: insights from the honey bee genome. Insect Molecular Biology 15, 563–576.

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., Jin, B.R., 2004. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. Comparative Biochemistry and Physiology 139B, 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., 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 V 140B 551–560.

Martin, B.S., Ruiz-Gomez, M., Landgraf, M., Bate, M., 2001. A distinct set of founders and fusion-competent myoblasts make visceral muscles in the *Drosophila* embryo. Development 128, 3331–3338.

Martin, M.M., 1983. Cellulose digestion in insects. Comparative Biochemistry and Physiology 75, 426–428.

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

Morris, K., Lorenzen, M.D., Hiromasa, Y., Tomich, J.M., Oppert, C., Elpidina, E.N., Vinokurov, K., Jurat-Fuentes, J.L., Fabrick, J., Oppert, B., 2009. The *Tribolium castaneum* larval gut transcriptome and proteome: a resource for the study of the coleopteran gut. Journal of Proteome Research 8, 3889–3898.

Nakashima, K., Watanabe, H., Saitoh, H., Tokuda, G., Azuma, J.I., 2002. Dual cellulose-digesting system of the wood-feeding termite, *Coptotermes formosanus* Shiraki. Insect Biochemistry and Molecular Biology 32, 777–784.

Okoniewska, M., Tanaka, T., Yada, R.Y., 2000. The pepsin residue glycine-76 contributes to active-site loop flexibility and participates in catalysis. Biochemical Journal 349, 169–177.

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B., Jurat-Fuentes, J.L., 2010. Prospecting for cellulolytic activity in insect digestive fluids. Comparative Biochemistry and Physiology 155B, 145–154.

Pant, R., Ramana, D., 1989. Cellulolytic activity in a phytophagous lepidopteran insect *Philosamia ricini*: the origin of the enzymes. Insect Biochemistry 19, 269–276.

Qin, Y., Wei, X., Song, X., Qu, Y., 2008. Engineering endoglucanase II from *Trichoderma reesei* to improve the catalytic efficiency at a higher pH optimum. Journal of Biotechnology 135, 190–195.

Richards, S., Gibbs, R.A., Weinstock, G.M., Brown, S.J., Denell, R., Beeman, R.W., Gibbs, R., Bucher, G., Friedrich, M., Grimmelikhuijzen, C.J., Klingler, M., Lorenzen, M., Roth, S., Schroder, R., Tautz, D., Zdobnov, E.M., Muzny, D., Attaway, T., Bell, S., Buhay, C.J., Chandrabose, M.N., Chavez, D., Clerk-Blankenburg, K.P., Cree, A., Dao, M., Davis, C., Chacko, J., Dinh, H., Dugan-Rocha, S., Fowler, G., Garner, T.T., Garnes, J., Gnirke, A., Hawes, A., Hernandez, J., Hines, S., Holder, M., Hume, J., Jhangiani, S.N., Joshi, V., Khan, Z.M., Jackson, L., Kovar, C., Kowis, A., Lee, S., Lewis, L.R., Margolis, J., Morgan, M., Nazareth, L.V., Nguyen, N., Okwuonu, G., Parker, D., Ruiz, S.J., Santibanez, J., Savard, J., Scherer, S.E., Schneider, B., Sodergren, E., Vattahil, S., Villasana, D., White, C.S., Wright, R., Park, Y., Lord, J., Oppert, B., Brown, S., Wang, L., Weinstock, G., Liu, Y., Worley, K., Elsik, C.G., Reese, J.T., Elhaik, E., Landan, G., Graur, D., Arensburger, P., Atkinson, P., Beidler, J., Demuth, J.P., Drury, D.W., Du, Y.Z., Fujiwara, H., Maselli, V., Osanai, M., Robertson, H.M., Tu, Z., Wang, J.J., Wang, S., Song, H., Zhang, L., Werner, D., Stanke, M., Morgenstern, B., Solovyev, V., Kosarev, P., Brown, G., Chen, H.C., Ermolaeva, O., Hlavina, W., Kapustin, Y., Kiryutin, B., Kitts, P., Maglott, D., Pruitt, K., Sapojnikov, V., Souvorov, A., Mackey, A.J., Waterhouse, R.M., Wyder, S., Kriventseva, E.V., Kadowaki, T., Bork, P., Aranda, M., Bao, R., Beermann, A., Berns, N., Bolognesi, R., Bonneton, F., Bopp, D., Butts, T., Chaumot, A., Denell, R.E., Ferrier, D.E., Gordon, C.M., Jindra, M., Lan, Q., Lattorff, H.M., Laudet, V., von Levetsow, C., Liu, Z., Lutz, R., Lynch, J.A., da Fonseca, R.N., Posnien, N., Reuter, R., Schinko, J.B., Schmitt, C., Schoppmeier, M., Shippy, T.D., Simonnet, F., Marques-Souza, H., Tomoyasu, Y., Trauner, J., Van der Zee, M., Vervoort, M., Wittkopp, N., Wimmer, E.A., Yang, X., Jones, A.K., Sattelle, D.B., Ebert, P.R., Nelson, D., Scott, J.G., Muthukrishnan, S., Kramer, K.J., Arakane, Y., Zhu, Q., Hogenkamp, D., Dixit, R., Jiang, H., Zou, Z., Marshall, J., Elpidina, E., Vinokurov, K., Oppert, C., Evans, J., Lu, Z., Zhao, P., Sumathipala, N., Altincicek, B., Vilcinskas, A., Williams, M., Hultmark, D., Hetru, C., Hauser, F., Cazzamali, G., Williamson, M., Li, B., Tanaka, Y., Predel, R., Neupert, S., Schachtner, J., Verleyen, P., Raible, F., Walden, K.K., Angeli, S., Foret, S., Schuetz, S., Maleszka, R., Miller, S.C., Grossmann, D., 2008. The genome of the model beetle and pest *Tribolium castaneum*. Nature 452, 949–955.

Sabater-Munoz, B., Legeai, F., Rispe, C., Bonhomme, J., Dearden, P., Dossat, C., Duclert, A., Gauthier, J.P., Ducray, D.G., Hunter, W., Dang, P., Kambhampati, S., Martinez-Torres, D., Cortes, T., Moya, A., Nakabachi, A., Philippe, C., Prunier-Leterme, N., Rahbe, Y., Simon, J.C., Stern, D.L., Wincker, P., Tagu, D., 2006. Large-scale gene discovery in the pea aphid *Acyrthosiphon pisum* (Hemiptera). Genome Biology 7, R21.

Sami, A.J., Shakoori, A.R., 2008. Biochemical characterization of endo-1,4-β-D-glucanase activity of a green insect pest *Aulacophora foveicollis* (Lucas). Life Science Journal 5, 30–36.

Scharf, M.E., Wu-Scharf, D., Pittendrigh, B.R., Bennett, G.W., 2003. Caste- and development-associated gene expression in a lower termite. Genome Biology 4, R62.

Scharf, M.E., Wu-Scharf, D., Zhou, X., Pittendrigh, B.R., Bennett, G.W., 2005. Gene expression profiles among immature and adult reproductive castes of the termite *Reticulitermes flavipes*. Insect Molecular Biology 14, 31–44.

Siebert, K.S., Lorenzen, M.D., Brown, S.J., Park, Y., Beeman, R.W., 2008. Tubulin superfamily genes in *Tribolium castaneum* and the use of a tubulin promoter to drive transgene expression. Insect Biochemistry and Molecular Biology 38, 749–755.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.

Tokuda, G., Watanabe, H., Lo, N., 2007. Does correlation of cellulase gene expression and cellulolytic activity in the gut of termite suggest synergistic collaboration of cellulases? Gene 401, 131–134.

Trauner, J., Schinko, J., Lorenzen, M.D., Shippy, T.D., Wimmer, E.A., Beeman, R.W., Klingler, M., Bucher, G., Brown, S.J., 2009. Large-scale insertional mutagenesis of a coleopteran stored grain pest, the red flour beetle *Tribolium castaneum*, identifies embryonic lethal mutations and enhancer traps. BMC Biology 7, 73.

Wang, S., Lorenzen, M.D., Beeman, R.W., Brown, S.J., 2008. Analysis of repetitive DNA distribution patterns in the *Tribolium castaneum* genome. Genome Biology 9, R61.

Wang, T., Liu, X., Yu, Q., Zhang, X., Qu, Y., Gao, P., 2005. Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. Biomolecular Engineering 22, 89–94.

Watanabe, H., Tokuda, G., 2010. Cellulolytic systems in insects. Annual Review of Entomology 55, 609–632.

Wei, Y.D., Lee, K.S., Gui, Z.Z., Yoon, H.J., Kim, I., Je, Y.H., Lee, S.M., Zhang, G.Z., Guo, X., Sohn, H.D., Jin, B.R., 2006a. N-linked glycosylation of a beetle (*Apriona germari*) cellulase Ag-EGase II is necessary for enzymatic activity. Insect Biochemistry and Molecular Biology 36, 435–441.

Wei, Y.D., Lee, K.S., Gui, Z.Z., Yoon, H.J., Kim, I., Zhang, G.Z., Guo, X., Sohn, H.D., Jin, B.R., 2006b. Molecular cloning, expression, and enzymatic activity of a novel endogenous cellulase from the mulberry longicorn beetle, *Apriona germari*. Comparative Biochemistry and Physiology 145B, 220–229.

Wei, Y.D., Lee, S.J., Lee, K.S., Gui, Z.Z., Yoon, H.J., Kim, I., Je, Y.H., Guo, X., Sohn, H.D., Jin, B.R., 2005. N-glycosylation is necessary for enzymatic activity of a beetle (*Apriona germari*) cellulase. Biochemical and Biophysical Research Communications 329, 331–336.

Weil, T., Rehli, M., Korb, J., 2007. Molecular basis for the reproductive division of labour in a lower termite. BMC Genomics 8, 198.

Willis, J.D., Klingeman, DW.E., Oppert, C., Oppert, B., Jurat-Fuentes, J.L., 2010. Characterization of cellulolytic activity from digestive fluids of *Dissosteira carolina* (Orthoptera: Acrididae). Comparative Biochemistry and Physiology 157B, 267–272.

Xiao, Z., Storms, R., Tsang, A., 2005. Microplate-based carboxymethylcellulose assay for endoglucanase activity. Analytical Biochemistry 342, 176–178.

Yuki, M., Moriya, S., Inoue, T., Kudo, T., 2008. Transcriptome analysis of the digestive organs of *Hodotermopsis sjostedti*, a lower termite that hosts mutualistic microorganisms in its hindgut. Zoological Science 25, 401–406.

Zhang, D., Lax, A.R., Raina, A.K., Bland, J.M., 2009. Differential cellulolytic activity of native-form and C-terminal tagged-form cellulase derived from *Coptotermes formosanus* and expressed in *E. coli*. Insect Biochemistry and Molecular Biology 39, 516–522.

Zhao, H., Jones, C.L., Baker, G.A., Xia, S., Olubajo, O., Person, V.N., 2009. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. Journal of Biotechnology 139, 47–54.

Zhou, X., Wheeler, M.M., Oi, F.M., Scharf, M.E., 2008. RNA interference in the termite *Reticulitermes flavipes* through ingestion of double-stranded RNA. Insect Biochemistry and Molecular Biology 38, 805–815.