Insect Biochemistry and Molecular Biology 40 (2010) 611–620

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

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

Functional and translational analyses of a beta-glucosidase gene (glycosyl hydrolase family 1) isolated from the gut of the lower termite *Reticulitermes flavipes*[☆]

Michael E. Scharf[a,*], Elena S. Kovaleva[b], Sanjay Jadhao[a], James H. Campbell[b], George W. Buchman[b], Drion G. Boucias[a]

[a] *Entomology and Nematology Department, University of Florida, PO Box 110620, Gainesville, FL 32611-0620, USA*
[b] *Chesapeake-PERL Inc., Savage, MD 20763, USA*

ARTICLE INFO

Article history:
Received 26 February 2010
Received in revised form
4 June 2010
Accepted 7 June 2010

Keywords:
GHF1
Beta-glucosidase
Laminaribose
Bioethanol
Termite

ABSTRACT

This research focused on digestive beta-glucosidases from glycosyl hydrolase family (GHF) 1 from the gut of the lower termite *Reticulitermes flavipes*. In preceding studies on *R. flavipes*, we characterized beta-glucosidase activity across the gut and its inhibition by carbohydrate-based inhibitors, and subsequently we identified two partial beta-glucosidase cDNA sequences from a host gut cDNA library. Here, we report on the full-length cDNA sequence for one of the *R. flavipes* beta-glucosidases (*RfBGluc-1*), the expression of its mRNA in the salivary gland and foregut, the production of recombinant protein using a baculovirus–insect expression system, optimal recombinant substrate specificity profiles and parameters, and significant inhibition by the established beta-glucosidase inhibitor cellobioimidazole. We also report the partial cDNA sequence for a second gut beta-glucosidase (*RfBGluc-2*), and show that like *RfBGluc-1* its mRNA is localized mainly in the salivary gland. Other results for *RfBGluc-1* showing activity against laminaribose, a component of microbial cell walls, suggest that *RfBGluc-1* may serve dual functions in cellulose digestion and immunity. These findings provide important information that will enable the testing of hypotheses related to collaborative host–symbiont lignocellulose digestion, and that contributes to the development of next-generation termiticides and novel biocatalyst cocktails for use in biomass-to-bioethanol applications.

1. Introduction

Most termites are well-adapted to survive on plant-derived lignocellulose and do so through a collaboration of endogenous (“host”) and symbiont-derived lignocellulase enzymes (Watanabe and Tokuda, 2001; Ohkuma, 2003; Scharf and Tartar, 2008). In termites, cellulose digestion is accomplished through the collaboration of endoglucanases (EC 3.2.1.4), exoglucanases and cellobiohydrolases (EC 3.2.1.91), and beta-glucosidases (EC 3.2.1.21) (Breznak and Brune, 1994). The latter enzymes acting in this process, beta-glucosidases, are typically from glycosyl hydrolase families (GHF) 1 and 3 (Henrissat and Bairoch, 1993; Davies and Henrissat, 1995; http://www.cazy.org/). Beta-glucosidases catalyze the hydrolysis of short beta-linked oligosaccharide metabolites released by endoglucanase and exoglucanase action, typically less than 5 units in length, to release free glucose (Ljungdahl and Erickson, 1985). However, there is also evidence in the insect literature that beta-glucosidases can hydrolyze sugar-aromatic conjugates to liberate kairomone and pheromone signaling molecules (Mattiacci et al., 1995; Cornette et al., 2003; Matsuura et al., 2009). Thus, because termites possess considerable lignocellulose digestion and chemically-mediated social communication capabilities (Wilson, 1971), termite beta-glucosidases likely serve dual functions in nutrient acquisition and pheromonal signaling (Korb et al., 2009).

To date, full-length beta-glucosidase genes have been sequenced and functionally investigated from the three lower termites *Neotermes koshunensis* (Tokuda et al., 2002; Ni et al., 2007), *Cryptotermes secundus* (Weil et al., 2007; Korb et al., 2009), and *Coptotermes formosanus* (Zhang et al., 2010), as well as the higher termite *Nasutitermes takasagoensis* (Tokuda et al., 2010). Another

---

[☆] Data Deposition: The nucleotide sequences reported in this paper have been deposited in the Genbank database with accession numbers HM152540, FL640173, FL639268, FL637619, FL637754, FL639498, FL636506, FL636125, FL635251, and FL635576.

[*] Corresponding author. Tel.: +1 352 273 3954; fax: +1 352 392 0190.
E-mail address: mescharf@ufl.edu (M.E. Scharf).

0965-1748/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2010.06.002

prior study suggests beta-glucosidases (sequence not reported) as a proteinaceous pheromone for egg recognition in the lower termite *Reticulitermes speratus* (Matsuura et al., 2009). The current work focused on digestive beta-glucosidases from the gut of *Reticulitermes flavipes*. In preceding work, beta-glucosidase activity across the *R. flavipes* gut and its inhibition by carbohydrate-based inhibitors was demonstrated (Wheeler et al., 2007; Zhou et al., 2008). More recently, two partial beta-glucosidase cDNA sequences from a *R. flavipes* host gut cDNA library were identified (Tartar et al., 2009).

The goals of this study were to obtain the full-length cDNA sequence for one of the *R. flavipes* beta-glucosidases (*RfBGluc-1*), characterize its expression profile in the gut, produce recombinant protein using a baculovirus–insect expression system, define optimal recombinant activity profiles and parameters, and investigate inhibition by known beta-glucosidase inhibitors. Additionally, we also report the partial cDNA sequence for a second gut beta-glucosidase (*RfBGluc-2*), as well as its expression profile across the *R. flavipes* gut. These findings provide novel basic information on *R. flavipes* beta-glucosidases that builds upon preceding biochemical data. These findings also provide essential background information for ongoing functional genomics research that is "investigating" termite immunity and host–symbiont collaboration in lignocellulose digestion, as well as developing novel biocatalyst cocktails for use in biomass-to-bioethanol applications and next-generation termiticides.

2. Materials and methods

2.1. Termites

R. flavipes colonies were collected in Gainesville, FL, USA and maintained in sealed plastic boxes (30 × 24 × 10 cm) in complete darkness (0:24 L:D), at 22 °C and 70 %RH. Colonies were maintained without soil for more than 3 months before use and provisioned with moist brown paper towels and pine wood shims. The identity of colonies as *R. flavipes* was verified by a combination of soldier morphology and 16S-mt-rDNA gene sequence. Only worker termites were used because of their significant lignocellulose digestion capability.

2.2. Sequencing and sequence analyses

Original sequence tags were obtained by Sanger sequencing from a host gut cDNA library as described in Tartar et al., (2009). *RfBGluc-1* was represented by eight library clones that aligned into a single contig ("Contig 771"; Accession Nos. FL640173, FL639268, FL637619, FL637754, FL639498, FL636506, FL636125, FL635251). *RfBGluc-2* was represented by one clone (Accession No. FL635576). Sequence alignments were performed using the ClustalW algorithm in Megalign™ (Lasergene software package; Madison, WI). Sequence features (signal peptide, polyadenylation site, poly-A tail) were identified based on expected features of eukaryotic genes (Lewin, 1997). N-glycosylation sites and signal peptides were predicted using online tools available at http://www.cbs.dtu.dk/services/. Active site motifs were identified as described previously (Tokuda et al., 2002).

2.3. Quantitative PCR

Quantitative real-time PCR (qRT-PCR) was conducted to determine beta-glucosidase expression profiles across the *R. flavipes* worker gut. RNA isolations were performed using the SV Total RNA kit (Promega; Madison, WI) following procedures detailed in Scharf and Tartar, (2008). Two replicate colonies were tested, which provided for two biological replicates with triplicate determinations. RNA was isolated from whole gut, whole carcass (*i.e.*, remaining body tissues after head and gut removal), and the four individual gut tissues foregut, salivary gland, midgut and hindgut. RNA quality and quantity were verified by agarose electrophoresis and spectrophotometry. cDNA was synthesized from total RNA using the iScript cDNA kit, which included a combination of oligo-dT and random hexamer primers (Bio-Rad; Hercules, CA). All PCR was performed in 20 μL volumes using SensiMix SYBR & Fluorescein one-step PCR reagent (Bioline; Taunton, MA), 1 μL cDNA template, and 0.5 μM primer concentrations. PCR primer sequences shown 5′–3′ (and accession numbers) were as follows: *RfBGluc-1* (FL637619) ("21C3 L" TGCTTCTTCATGGCTCAGAGT; "21C3 R" TGGTCTCCAGGTTGTGTATCC); *RfBGluc-2* (FL635576) ("309K21 L" GAGAGCTCCTCAACCAACGAT; "309K21 R" CCTGTAACAAGCACCGGAGTA); *beta-actin* (DQ206832) (Forward AGAGGGAAATCGTGCGTGAC; Reverse CAATAGTGATGACCTGGCCGT), and *NADH-dh* (BQ788168) (Forward GCTGGGGGTGTTATTCATTCCTA; Reverse GGCATACCACAAAGAGCAAAA). Nucleotide alignments and primer positions for the two beta-glucosidases are provided in Supplementary Fig. S1. Whole gut vs. carcass comparisons were made by viewing reaction products after 35 cycles of amplification on 2% agarose gels. Relative expression within gut tissues was determined by quantitative real-time PCR (qRT-PCR), using NADH-dh as the reference gene. Reaction efficiencies were determined using 2-fold serial dilutions of template as described in the MyIQ Real-Time PCR System user's manual (Bio-Rad). qRT-PCR data were analyzed by the $2^{-\Delta CT\ \Delta CT}$ method (Livak and Schmittgen, 2001), with normalization to the midgut as detailed by Scharf and Tartar (2008). Statistical analysis of qRT-PCR data was performed by analyzing ΔCT data for all tissues. Mean separations were made using the Tukey HSD method (SAS software; Cary, NC).

2.4. Recombinant protein production and purification

RfBGluc-1 was expressed as a recombinant protein for use in functional characterization studies. Three beta-glucosidase-encoding cDNA clones as described in Tartar et al., (2009) (TG_02_G7, TG_303_K20 and TG_21_C3) were sequenced using universal plasmid primers M13 forward and reverse. The entire *RfBGluc-1* ORF was present only in clone TG_21_C3. Thus, the full-length ORF sequence was amplified from clone TG_21_C3. Restriction sites and a C-terminal nucleotide sequence plus 6 histidine residues were introduced into the amplicon using the following primers: forward, 5′- GTCGAC**ATG**AGGTTACAGACGGTTTGC-3′ (*Sal*1 sites underlined, start codon shown in bold); and reverse, 5′-CTGCAG**TTA***GTGATGATGGTGATGATG*GTCTAGGAAGCGTTCTGGAA-3′ (*Pst*1 site underlined, stop codon shown in **bold** and 6x histidine-coding nucleotides *italicized*). This PCR amplicon, which encoded the full-length *RfBGluc-1* ORF sequence (amino acids 1 to 495) and 6x histidine tag at the C-terminus, was cloned into *Sal*1-*Pst*1 site of pFastBac1 transfer vector (Invitrogen; Carlsbad, CA). Recombinant baculovirus was generated using the Bac-to-Bac system (Invitrogen). The inserted *RfBGluc-1* gene was confirmed by sequencing and used for bacmid preparation and transfection into Sf9 cells, which were then incubated at 27 °C for 4 days (Passage 0). The cell pellet from Passage 0 was tested by Western blotting with anti-His antibody to confirm expression of His-tagged protein, and recombinant virus was harvested and injected to *T. ni* larvae as described previously (Liu et al., 2007; Kovaleva et al., 2009). Larvae were orally infected with active pre-occluded baculovirus (Liu et al., 2007; Kovaleva et al., 2009), harvested in large scale, and stored at −80 °C for later processing. Recombinant protein was recovered with greater than 98% purity from *T. ni* homogenates by (1) centrifugation, (2) 0.22 μM supernatant filtration, (3) tandem