CHAPTER 3

New endo-β-1,4-glucanases from the parabasalian symbionts, *Pseudotrichonympha grassii* and *Holomastigotoides mirabile* of *Coptotermes* termites

Abstract

An endo-β-1,4-glucanase (EG) was purified from the hindgut of an Australian mound-building termite, *Coptotermes lacteus*. The hindgut extract formed a peak separate from those for extracts obtained from the salivary glands and the midgut in Sephacryl S-200 gel chromatography, and also demonstrated an origin different from the endogenous EGs of the termite itself. The recovery was further purified by SDS-PAGE, and its N-terminal amino-acid sequence was analyzed. The obtained N-terminal sequence showed high homology to EGs from glycoside-hydrolase family (GHF) 7. PCR-based cloning methods were applied to the hindgut content of *C. lacteus* and the protozoan symbionts (*Pseudotrichonympha grassii* and *Holomastigotoides mirabile*) of *C. formosanus*. As a result, cDNAs encoding putative EGs homologous to GHF7 members were identified. The functionality of one of the putative proteins was confirmed by its expression in *E. coli*.

Introduction

Although termites are well known as xylophagous animals, until recent reports of endogenous and protozoan cellulose genes (Ohtoko et al., 2000; Tokuda et al., 1999; Watanabe et al., 1998), little had been known about cellulolytic enzymes apart from chromatographic separations of cellulolytic activities (Slaytor, 2000; Watanabe and Tokuda, 2001). Termites (order Isoptera), which consist of 7 families,

The content of this chapter has been published in; H. Watanabe *et al.* (2002) New endo-β-1,4-glucanases from the parabasalian symbionts, *Pseudotrichonympha grassii* and *Holomastigotoides mirabile* of *Coptotermes* termites. Cell. Molec. Life Sci. **59**: 1983-1993.

are divided into two categories; one made up of protozoa-independent species (family Termitidae), and the other of protozoa-dependent species (families Mastotermitidae, Kalotermitidae, Termopsidae, Hodotermitidae, Rhinotermitidae and Serritermitidae)(Krishna, 1970). The first group does not harbor cellulolytic protozoa in its gut, thus to account for the deviating distribution of cellulase activity detected in this group's midgut, it has been presumed by termite researchers that cellulolytic enzymes must be produced in the midgut tissue where no cellulolytic symbionts have been identified (Breznak and Brune, 1994; O'Brien et al., 1979; Slaytor, 1992, 2000). Still, the endogenous production of cellulolytic enzymes had not been generally accepted until the recent finding of a gene encoding an endo-β-1,4-glucanase (EG; EC 3.2.1.4; a type of cellulolytic enzyme which often represents cellulase) of glycoside-hydrolase family (GHF) 9 from the protozoan-independent termite, *Nasutitermes takasagoensis* (Tokuda et al., 1999; Watanabe and Tokuda, 2001).

In the second group, composed of protozoa-dependent species, the original proposal for the protozoan cellulose digestion was made by Cleveland (Cleveland, 1924), and the combined cellulose digestion of endogenous and protozoan cellulases in termites were proposed because of the simultaneous distribution of cellulolytic activities both in secreting organs (the salivary glands and midgut, which do not have cellulolytic symbionts), and non secreting organ (the hindgut) harboring protozoan fauna (O'Brien et al., 1979; Yamaoka and Nagatani, 1975). The endogenous and protozoan enzymes have been considered to share responsibility for cellulose digestion (Breznak and Brune, 1994; Hogan et al., 1988a; O'Brien and Slaytor, 1982; Yamaoka and Nagatani, 1975). In spite of these historical background, only the role of hindgut protozoan fauna has been emphasized, but gene level demonstration had been done only for the endogenous EGs (GHF9) (Watanabe et al., 1998) until recent finding of a GHF 45 cellulase from hindgut hypermastigote protozoa of

Reticulitermes speratus (Ohtoko et al., 2000). Nakashima and Azuma purified a cellulase to be placed GHF7 from the termite, *Coptotermes formosanus*, and attributed the salivary glands as the producer of this enzyme (Nakashima and Azuma, 2000), but it was renounced and re-attributed to the hindgut protozoan fauna based on an identification and relocation of each cellulases among different digestive organs (the salivary glands, fore-gut, midgut, and the hindgut) (Nakashima et al., 2002a). Yamaoka and Nagatani (1975) demonstrated presence of two different enzymes in the gut of the protozoa inhabited termite, *Reticulitermes speratus*; one from the salivary glands was active against carboxymethylcellulose (Cx-cellulase; endo-β-1,4-glucanase), and the other from the hindgut protozoan fauna was active against filter paper (C1-cellulase; cellobiohydrolase), which were assumed to digest crystalline cellulose synergistically (Yamaoka and Nagatani, 1975). Despite the identification of cellulolytic activities in the termite digestive systems, the exact contribution of each species of protozoa in cellulose digestion largely remains to be clarified. The report of cellulase genes from the hindgut protozoa of *R. speratus* has partially addressed to elucidate the cellulolytic enzymes produced by the symbiotic protozoa (Ohtoko et al., 2000), however it does not provide a general explanation for composition of protozoan enzymes in other protozoa-inhabited termite species.

In the present study, a possible protozoan EG was purified from the hindgut extract of the Australian mound-building termite, *Coptotermes lacteus*, analyzed amino-acid sequences and elucidated the major protozoan cellulolytic component. Following the isolation of EG, expression of EG coding genes in the hindgut contents (contain protozoan fauna) of *C. lacteus* and, in addition, that in the parabasalian symbionts of the worldwide pest species, *C. formosanus,* which shares common hindgut protozoan composition of genera *Pseudotrichonympha, Holomastigotoides* and *Spirotrichonympha* with other *Coptotermes* species (Kitade and Matsumoto,

1998a; Koizumi, 1921; Yamin, 1979) were demonstrated by employing PCR-based cloning techniques.

Materials and methods

Termites

Coptotermes lacteus workers were collected from a termite mound in Wyong, N.S.W., Australia and used for protein purifications. Part of *C. lacteus* workers were preserved in absolute acetone for following cDNA cloning (Fukatsu, 1999) because of its limited supply and access for the collection site. *C. formosanus* workers were collected on Iriomote Island in Okinawa Prefecture, Japan and used for cDNA cloning in fresh form. Termites for fresh use were kept with their nest materials and tissue paper or sawdust for food for several days until used for the experiments.

Detection of EG activity

Enzymatic samples (crude extracts or fractions of gel filtration, 10 μl) were incubated 200 μl of 2% (w/v) sodium carboxymethylcellulose (CMC; standard molecular weight: 250,000, degree of carboxymethyl substitution: 0.7 (w/v); Aldrich) in sodium acetate buffer (0.1 M, pH 5.5) at 37 °C for 5 min. Reducing sugars were detected with tetrazolium blue regent (Sigma) as glucose equivalents (Jue and Lipke, 1985).

Gel filtration and SDS-PAGE

The hindgut, salivary glands and the midgut including the esophagus and proventriculus from 200 workers of *C. lacteus* were homogenized separately in 500 μl of ammonium acetate buffer (0.3 M, pH 5.0). Each homogenate was centrifuged for 5 min at 9,000 *g* and the supernatant was applied to a gel filtration column

(HIPREP 16/60 Sephacryl S-200 High Resolution, Amersham Biosciences). Samples were eluted with ammonium acetate buffer at a flow rate 0.5 ml/min, and the peak fractions and its shoulders were separately stored. Shoulder fractions were applied to the same column and eluted under the same conditions. The recovered peak fraction was added to the first recovery. The stored recovery was denatured with SDS-PAGE sample buffer and applied for SDS-PAGE (0.75%, 1 mm in thick and 100 mm running length) according to the conventional method (Laemmli, 1970).

N-terminal amino-acid analysis

Proteins on the SDS-PAGE gels were transferred to PVDF membrane using a semi-dry blotting cell (Bio-Rad) according to the method of Hirano and Watanabe (1990). Protein bands on the membrane were cut out and applied to an N-terminal amino-acid sequencer (Beckman).

cDNA cloning

mRNA was extracted from the hindgut contents using the QuickPrep micro mRNA extraction kit (Amersham Biosciences). First-strand cDNA was constructed with oligo-dT primer with (for 3'-RACE) or without an anchor sequence and Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's recommended conditions. PCR amplification of the first-strand cDNA was performed with a set of degenerate primers designed from the N-terminal sequence of the purified protein with recombinant Taq polymerase (Takara, Japan) according to the manufacturer's recommended buffer and substrate conditions. For identification of 5'- and 3'-flanking regions, RACE (rapid amplification of cDNA ends) was employed with gene specific primers. For cloning cDNA of EG from the hindgut of *C. lacteus*, a degenerate primer was redesigned from its N-terminal amino-acid sequence,

and 3'-RACE method was directly applied to amplify the cDNA (Tokuda et al., 2002). All RACE amplifications were performed using the SMART RACE cDNA amplification kit (Clontech). Nested PCR was performed for RACE of the 5'-flanking regions. The cloning strategy employed is summarized in Fig. 1.

Identification of protozoan species expressing EG genes

To identify EG expressing protozoan species, RT-PCR based on individual protozoan cells (single cell RT-PCR) was employed. The hindgut contents of the workers were suspended in a 90% concentration of Solution U (Trager, 1934), and each cell of protozoa was separated using a pipette with 200 μl tips under a microscope. Subsequently each protozoan cell was ejected into 100 μl of acetone in a 0.5 ml PCR tube to be dehydrated. The acetone was dried by heating the tube on a heating block (100 °C) for 1 min. Ten μl of the first-strand mixture (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 5 mM $MgCl_2$, 1 μM of anti-sense side gene-specific primer, each 0.25 μM of dNTPs, 0.5 μl of reverse transcriptase [Superscript II, Invitrogen] and 0.25 μl of human placenta ribonuclease inhibitor [Toyobo, Japan]) was added to the dried specimen and incubated for 1 hour at 42 °C. The reaction was stopped by heating at 70 °C for 5 min, following which the tube was placed on ice. Supplemental regent solution was added to make 50 μl of PCR reaction mixture (10 mM Tris-HCl [pH 8.3], 50 mM KCl, 15 mM $MgCl_2$ and finally, 2.5 units of recombinant Taq polymerase [Takara]). The PCR reaction was performed using the hot start technique. This technique was used for identification of origin of particular mRNA but inadequate for cloning for full length cDNA because of the small recovery of cDNA recovered from the reverse-transcription step.

DNA sequencing

For all PCR amplifications, the products were cloned into a plasmid cloning vector using the pMOSBlue blunt end cloning kit (Amersham Biosciences), or pGEM-T plasmid cloning vector (Promega) with JM109 bacterial host. The constructed plasmids were extracted from the host bacterial cells using conventional alkaline mini-prep and sequenced with an ABI377 or ABI3700 automated DNA sequencer with BigDye Terminator with appropriate primers (designed for vector sequences or the same primers as used for the PCR amplifications). All gene handling procedures were performed according to textbook methods (Ausubel et al., 1989) unless otherwise noted.

Expression of protozoan EG cDNA in E. coli

The region coding from N-terminal to the stop codon (i.e., that without putative signal peptides) of one protozoan EG cDNA was amplified by PCR using a sense side primer (HmEx, Table 1) and an anchor primer (joined at the 3'-end) following the method for 3'-RACE, and then ligated in a *Lac Z* coding frame at the EcoRV site of the pMOSBlue plasmid vector. The construct was transformed by pMOSBlue host cells (Amersham Biosciences). Expression and activity detection were performed using the method described previously (Tokuda et al., 1999).

Sequence analysis

Homology searches of the determined sequences were performed using the BLAST program (Altschul et al., 1997) through the NCBI server at http://www.ncbi.nlm.nih.gov/BLAST/. Alignment of the searched homologues was performed using CLUSTAL X (Ver. 1.8) for Windows (National Center for Biotechnology Information) with Gonnet 250 pairwise parameters with matured-form amino-acid sequences (Altschul et al., 1990). Phylogenic trees were constructed

by TREE-PUZZLE tree reconstruction software, with quartet-sampling and neighbor-joining of 1,000 puzzling steps using the WAG substitution model (Strimmer and Von Haeseler, 1996), based on alignments of amino-acid sequences recorded on the GenBank, EMBL and DDBJ databases by CLUSTAL X.

Results

Purification of the protozoan EG and N-terminal sequencing

Hindgut, midgut, and salivary gland extracts of *C. lacteus* were separately applied to the gel filtration column. EG activity of the hindgut formed a peak at 69.6 ml, while the others formed peaks at 74.4 ml (salivary glands), or 72.0 and 74.4 ml (midgut), as seen in Fig. 2. The hindgut peak fraction from each chromatography was stored and applied for SDS-PAGE. Proteins on the SDS-PAGE gel were transblotted on PVDF membrane where one band of EG was separated from contaminations. N-terminal amino-acid sequencing was conducted for the EG band and the amino-acid sequence (EKHPKFTWQEVTK©G©TNDN [©: blank positions assumed to be Cys residues, which are undetectable using conventional amino-acid sequencers]) showed homology to other EGs in GHF7 by the BLAST X search.

Cloning of protozoan EG cDNAs

Degenerate primers (DS for sense side and DA for anti-sense side) (Table 1) were designed from the N-terminal amino-acid sequence of the purified hindgut EG of *C. lacteus*. First-strand cDNAs were synthesized with oligo-dT primer with or without an anchor sequence from mRNA extracted from the hindgut contents from *C. lacteus*, which had been dehydrated and preserved in absolute acetone, and from that of freshly collected *C. formosanus*. The first-strand cDNAs were applied for PCR

amplifications to isolate target cDNA. The overall PCR strategies for cloning are summarized in Fig. 1. From the hindgut contents of *C. lacteus*, two cDNAs (*CLhgEG1* [GenBank accession: AB089800] and *CLhgEG2* [AB089801], 1,031 bp and 1,085 bp, respectively) were obtained (Fig. 1a). From *C. formosanus*, although the N-terminal coding region of the intended cDNA was not amplified by the designed degenerate primers, a partial sequence of protozoan EG cDNA (corresponding to the 274th-892nd bases of *PgEG1a* with the anti-sense side primer sequence [DA, 20 bp, Table 1] at both ends) was obtained (Fig. 1b). From this partial sequence, a sense side primer (PCEL-F1) and an anti-sense primer (PCEL-R1) were designed (Table 1). Using these primers, two different groups of partial cDNAs were identified from the protozoa, *Pseudotrichonympha grassii* and *Holomastigotoides mirabile* (*PgEGs* and *HmEGs*, respectively) by the single cell RT-PCR (Fig. 1a and b). The amplification of *HmEGs* from *H. mirabile* was considered to be due to the unexpected similarity of PCEL-F1 and PCEL-R1 in corresponding sequences (Table 1). From these partial sequences, anti-sense primers for 5'-RACE (Pg5'-1 and Hm5'-1) and for nested amplifications (Pg5'-2 and Hm5'-2) were designed (Table 1). By the 5'-RACEs and the nested PCR amplifications which followed, 5'-regions of *PgEGs* and *HmEGs* were obtained (Fig. 1b and c). The obtained DNA sequences of the 5'-regions encoded a putative N-terminal amino-acid sequence homologous to that of the purified EG from the hindgut of *C. lacteus*. Sense-side primers (*PgEx* and *HmEx*) were designed from the DNA sequences of the putative N-terminal coding regions (Table 1) and conducted 3' RACEs with these. Ten different cDNAs were identified for *PgEGs* and three for *HmEGs*. These sequences were aligned with the 5'-RACE sequences. As a result, *PgEG1a* (GenBank accession: AB071001), *HmEG1* (AB071011), *HmEG2* (AB071012) and *HmEG3* (AB071013) were identified with their region spanning from the putative start codons to downstream of the stop codons,

while for the other cDNAs (*PgEG1b-PgEG1h*, AB071002-AB071010), the 5' RACE did not reach either to the putative start codons or to the following signal peptide coding regions.

Origin of hindgut EGs

The single cell RT-PCR method elucidated the origin of the cloned EG cDNAs as well as clarifying their nucleotide sequences. Each protozoan cell from the hindgut of *C. formosanus* was carefully separated with a pipette while keeping record of the species with a digital camera. RT-PCR was performed for each single protozoan cell, priming the first-strand cDNA with PCEL-F1, and amplifying with PCEL-F1 and PCEL-R1 (Table 1, Fig. 1b and c). The PCR products were then cloned and sequenced. All *PgEG* and *HmEG* homologues were amplified from *Pseudotrichonympha grassii* and *Holomastigotoides mirabile*, respectively. No cross-contamination or exceptions occurred, as verified by the recordings of the preparation of the protozoa. No amplification was obtained using background Solution U as a control sample. Fresh samples of *C. lacteus* were not available to perform the single cell protozoan RT-PCR.

Sequence comparison of protozoan EG cDNAs

Cloned cDNAs from *P. grassii* and *H. mirabile* encoded *PgEG1* (328 amino-acids on *PgEG1a*), *HmEG1*, *HmEG2* and *HmEG3* (326 amino-acids on each), and part of *PgEG1* (from the putative N-terminal to the putative C-terminal, 318 amino-acids on *PgEG1b-PgEG1h*). The putative amino-acid sequences encoded on the isolated cDNAs showed over 99% identity in nucleotides within species except CLhgEG1 and CLhgEG2, and 61-62% between protozoan species. The putative proteins encoded by each cDNA had one amino-acid substitution between each other,

except for *HmEG3*, which had 13 substitutions against *HmEG1* and 14 against *HmEG2*. The cDNAs obtained from the hindgut content of *C. lacteus* (*CLhgEG1* and *CLhgEG2*) encoded 307 and 303 amino-acids, respectively, corresponding from the downstream of the N-terminal region used to design the degenerate primer (CL5', Table 1, Fig. 1a) to the down stream of putative C-terminal and stop codon. CLhgEG1 showed 76.9%, 88.6-89.9%, and 69.1-69.9% identities in amino-acids to CLhgEG2, HmEGs, and PgEGs, respectively, while CLhgEG2 showed 68.9% and 77.8-79.8% identities in amino-acids to PgEGs and HmEGs, respectively. By BLAST X homology searches, the *PgEGs*, *HmEGs* and *CLhgEGs* showed high homology to EGs and endo-β-1,4-glucanasecellobiohydrases (CBH; E.C. 3.2.1.91) both from GHF7 (Fig. 3). The *PgEGs*, *HmEGs* and *CLhgEGs* encoded only putative catalytic domains. No other domains homologous to linkers or cellulose-binding domains of multi-domain cellulases were encoded. Specifically, the PgEGs, HmEGs and CL-hindgut EGs showed 32-39% identity in amino-acid sequence to TrEG1 (=Cel7B), and 37-43% identity to TrCBH1 (=Cel7A) from *Trichoderma reesei*, whose crystal structures have been firmly elucidated (Divne et al., 1994; Divne et al., 1998; Kleywegt et al., 1997). The protozoan and hindgut EGs lacked all of the loop forming regions and many of the shift loop forming regions covering over the catalytic clefts of TrCBH1 and TrEG1, respectively (Divne et al., 1994; Divne et al., 1998; Kleywegt et al., 1997) (Fig. 3). Phylogenetic analysis of protozoan EGs and CLhgEGs with other members of GHF7 showed that protozoan (and *C. lacteus* hindgut) EGs, fungal EGs, and CBHs form distinctive clades (Fig 4).

Expression of the protozoan EG

To confirm the functionality of the protozoan EGs, one of the obtained EG cDNAs (*HmEG3*) was sub-cloned to pMOSBlue cloning vector and the expression of

a functional EG was confirmed by the formation of halos on an activity-stained CMC-LB plate (Fig. 5). The control, transformed pMOSBlue without insert, did not create any recognizable halos (Fig. 5).

Discussion

There have been several descriptions on *Holomastigotoides* sp. in *Coptotermes* (Grassi, 1917; Koizumi, 1921; Yamin, 1979). In *C. formosanus*, "*Holomastigotoides hartmanni*" was originally identified as solo *Holomastigotoides* species by Koizumi (1921) (Koizumi, 1921). Recently, phylogenic analyses of the hypermastigotes in *C. formosanus* were performed by Ohkuma et al. (2000), in which, *Holomastigotoides* sp. was described as *H. mirabile* (Ohkuma et al., 2000). Although the name *H. hartmanni* is still supported by many researchers, the present study has followed Ohkuma's description in order to avoid possible confusion on DNA databases.

The gel filtration using Superdex 75 distinguished EGs from the salivary glands, the midgut and the hindgut of *C. lacteus*. The simultaneous presence of endogenous and protozoan cellulase activities was first proposed by Yamaoka and Nagatani (1975) based on their investigations of the salivary glands and hindgut of *Reticulitermes speratus* from the family, Rhinotermitidae, of which *Coptotermes* species are members. Following this report, endogenous and protozoan activities were also identified in the digestive system of *C. lacteus* and chromatographically separated using a Superose 6B molecular sieving column (O'Brien et al., 1979). The present results coincide with these earlier studies.

In addition to the peak at the same fraction as that from the salivary gland extract, the *C. lacteus* midgut extract formed a second active peak at higher molecular weight, which was assumed an aggregate form of the protein consisting in the first

peak (Fig. 2). In termites, all known endogenous EG genes encode proteins consisting of 448 amino-acids and their expression in the salivary glands and/or midgut have been established (Nakashima et al., 2002a; Tokuda et al., 1999; Watanabe et al., 1998; Watanabe and Tokuda, 2001). Homologues to the termite's endogenous EG genes have also been found in cockroaches, including omnivorous species, and a crayfish, which are evolutionarily cross-related to termites and to insects, respectively (Lo et al., 2000; Watanabe and Tokuda, 2001).

The esophagus and proventriculus (foregut) included in the midgut extract are unlikely to be origins of EG activity, since no secretion or stable symbionts have ever been reported in these organs in insects as far as the author knows, and the foregut is generally considered a non-secreting organ (Chapman, 1998). The salivary glands are connected to the esophagus through the salivary canals, both enzymes would be produced by the salivary glands and midgut would contribute to the activity from the extract of foregut and midgut (Nakashima et al., 2002a).

The purified hindgut EG and CLhgEGs from *C. lacteus* showed high homology to PgEGs and HmEGs, consisting of 315 and 318 amino-acids, respectively. Therefore, the hindgut EG of *C. lacteus* is expected to have similar molecular weights, and thus, they would be smaller the termite endogenous EGs with 432 amino-acids without signal peptides in molecular weights (Nakashima et al., 2002a; Tokuda et al., 1999; Watanabe et al., 1998), although the hindgut EG of *C. lacteus* was eluted earlier than those from its salivary glands and midgut from Superdex 75 (Fig. 2). This might again be explained by molecular clustering due to the experimental conditions.

The result demonstrated that *H. mirabile* and *P. grassii* (both from Hypermastigida) of *C. formosanus* have genes encoding GHF7 members, that is expected to be the case for the hindgut fauna of *C. lacteus* sharing *Holomastigotoides*

and *Pseudotrichonympha* as one of *Coptotermes* species (Kitade and Matsumoto, 1998a; Yamin, 1979). Possible prokaryote origins (externally attached spirochaeta, secondary bacterial or archeal intracellular-symbionts *et al.*) of the hindgut cellulases would be ruled out since our experimental methodology recovered poly-A tailed mRNA specific to eukaryotes. The negative result for the RT-PCR amplification of the background solution U, which possibly contained those micro organisms except protozoa, also supports this conclusion. In another protozoon, *Trichonympha sphaerica* (Hypermastigida) from the hindgut of *Zootermopsis* sp., it was formally concluded that the cellulose digestion in the protozoon is independent of its endosymbiotic bacteria (Yamin, 1980). The fact that no GHF7 members have ever been reported outside of Fungi would disclaim possible prokaryote origins, too. The phylogenetic analysis of GHF7 members also supports the protozoan origin of PgEGs, HmEGs and CLhgEGs by forming a clade unique for them (Fig. 4). GHF45 members or their cDNA could not be found in the hindguts both of *C. lacteus* and *C. formosanus*. This might be explained by the difference in protozoan fauna between *Coptotermes* and *Reticulitermes* termites. Many of isolated GHF45 cDNAs were unidentified of their origin, but part of them were attributed to *Teranympha mirabilis* and *Trichonympha agili* of *R. speratus* (Ohtoko et al., 2000), which were absent in the hindgut of *Coptotermes* (Kitade and Matsumoto, 1998a; Yamin, 1979).

The glycoside-hydrolase family 7 consists of both EGs and cellobiohydrases (CBHs) (Henrissat, 1991) (http:afmb.cnrsmrs.fr/~pedro/CAZY/ghf.html). The three-dimensional structure of GHF7 has been elucidated on cellobiohydrase (CBH1) and endo-β-1,4-glucanase (EG1) from *T. reesei* (Divne et al., 1994; Divne et al., 1998; Kleywegt et al., 1997). CBH1 from *T. reesei* has a β sandwich structure consisting of two large anti-parallel β sheets opposite to each other (Divne et al.,

1994; Divne et al., 1998; Kleywegt et al., 1997). The comparison of amino-acid sequences suggests that this overall structure of CBH1 seems to be conserved among GHF7 members including the protozoan EGs (Fig. 3). CBH1 and EG1 from *T. reesei* are retention-type cellulases which have two glutamate residues, one acting as a nucleophile and the other as a Bronsted acid/base catalyst, and one aspartic acid residue, which acts as a nucleophile supporter at the catalytic center (corresponding to Glu 212, Glu 217 and Asp 214 of TrCBH1, respectively, Fig. 3) (Divne et al., 1994; Divne et al., 1998; Kleywegt et al., 1997). These three residues are conserved in all protozoan (hindgut) EGs cloned in the present study and in the other GHF7 members compared with them (Fig. 3). These structural similarities and homologies of the protozoan EGs to other GHF7 members reported previously suggest that these putative proteins are functional, and thereafter the functionality of one protozoan EG (*HmEG3*) was experimentally confirmed (Fig. 5). The high homology between the N-terminals of the putative proteins coded on the protozoan EG cDNAs from *C. formosanus* and the purified protozoan EG from *C. lacteus* also supports this conclusion.

The catalytic domain of CBH1 from *T. reesei* forms a tunnel-like structure comprised of two anti-parallel β-sheets with loop-forming residues stitching the sheets together at both sides, thus access to substrate cellulose chains is limited to only from the terminals (Divne et al., 1994). However, in EG1 from *T. reesei*, less than half of the loop-forming amino-acid residues are conserved in the sequence (Kleywegt et al., 1997). The stitching function of the missing residues is substituted by shift loop-forming regions which form a more open structure at the catalytic center; thus allowing this enzyme to randomly access the middle of cellulose chains (Kleywegt et al., 1997; Rouvinen et al., 1990). This open vs. closed structural difference in the catalytic centers determines the enzyme's affinity to either

crystalline or amorphous cellulose, respectively, since in crystalline cellulose the only parts accessible to the enzymes are at the ends of the chains, while amorphous cellulose, like CMC, allows access to the middle of the chains if a cellulase have open-form catalytic center (Teeri, 1997). In the termite protozoan EGs examined here, many of amino-acid residues corresponding to both the loop-forming and shift loop-forming regions of CBH1/EG1 are absent or substituted (Fig. 3), although there should be some residues which hold two β-sheets together to form a catalytic cleft. Therefore, these protozoan enzymes would be randomly accessible to amorphous-form cellulose chains like EGs usually are. This is also supported by the activities, against CMC as a specific substrate for EGs (Klesov, 1991), of purified hindgut cellulase of *C. lacteus* and HmEG3 (a homologue of HmEG1, HmEG2, PgEGs, and CLhgEGs) expressed in *E. coli*. It was previously demonstrated that the detectable amount of endogenous EGs stored in the midgut of *C. formosanus* does not flow into the hindgut and the crude extracts of the midgut and the hindgut hydrolyze crystalline cellulose independently (Nakashima and Azuma, 2000). Here, it was concluded that the hindgut protozoa of *Coptotermes* termites produce EGs ("Cx-cellulase" (Yamaoka and Nagatani, 1975)) by themselves.

Table 1 Primers used for the experiments

Name	Design	Corresponding sequence (position), side
DS	5′ CAY CCN AAR TTY ACN TGG CAR 3′	"HPKFTWQ" (3-9), sense side
DA	5′ RTT RTC RTT NGT RCA NCC RC 3′	"CGCTNDN" (14-20), antisense side
PCEL-F1	5′ TGG ATA TGG TTA CTG TGA TGC GAA CTG CGT 3′ (C) (C) (C) (G)	 *PgEG1a* (438-467), sense side *HmEG1* (442-471), sense side
PCEL-R1	5′ GAT CGT CTT CCC ACC CTG CAC ATA CAA ACG 3′	*PgEG1a* (735-706), antisense side *HmEG1* (736-707), antisense side
Hm5′-1	5′ CCG TTA GCG AMC CTC CTG ACC CCA C 3′ (C)	*HmEG1* (698-674), antisense side
Hm5′-2	5′ CGA MCC TCC TGA CCC CAC AAA STG 3′ (C) (C)	*HmEG1* (691-667), antisense side
HhEx	5′ CGA GAA GCA TCC GAA GTT TGT G 3′	*HmEG1* (43-64), sense side
Pg5′-1	5′ CCC ATT CAA GTT ACC AGA AGC ATC C 3′ (A) (G)	*PgEG1a* (700-675), antisense side
Pg5′-2	5′ CAA GTT ACC AGA AGC ATC CGC ACG G 3′ (G)	*PgEG1a* (693-669), antisense side
PgEx	5′ CGA GAA GCA TCC GAA CTT TAC 3′	*PgEG1a* (39-59), sense side
CL5′	5′ GAR AAR CAY CCN AAR TTY ACN TGG CA 3′	"EKHPKFTWQ" (1-9), sense side

The degenerate primers (DS and DA) were designed from the N-terminal amino-acid sequence of protozoan endo-β-1,4-glucanase purified from the hindgut of *Coptotermes lacteus*. The characters in parentheses indicate actual bases finally confirmed in the corresponding sequences. Abbreviations for positions with degeneracies are: Y for C+T, R for A+G, M for A+C, S for C+G, and N for A+T+G+C. DDBJ/GenBank Accession Numbers for *PgEG1a* and *HmEG1* are AB071001 and AB071011, respectively.

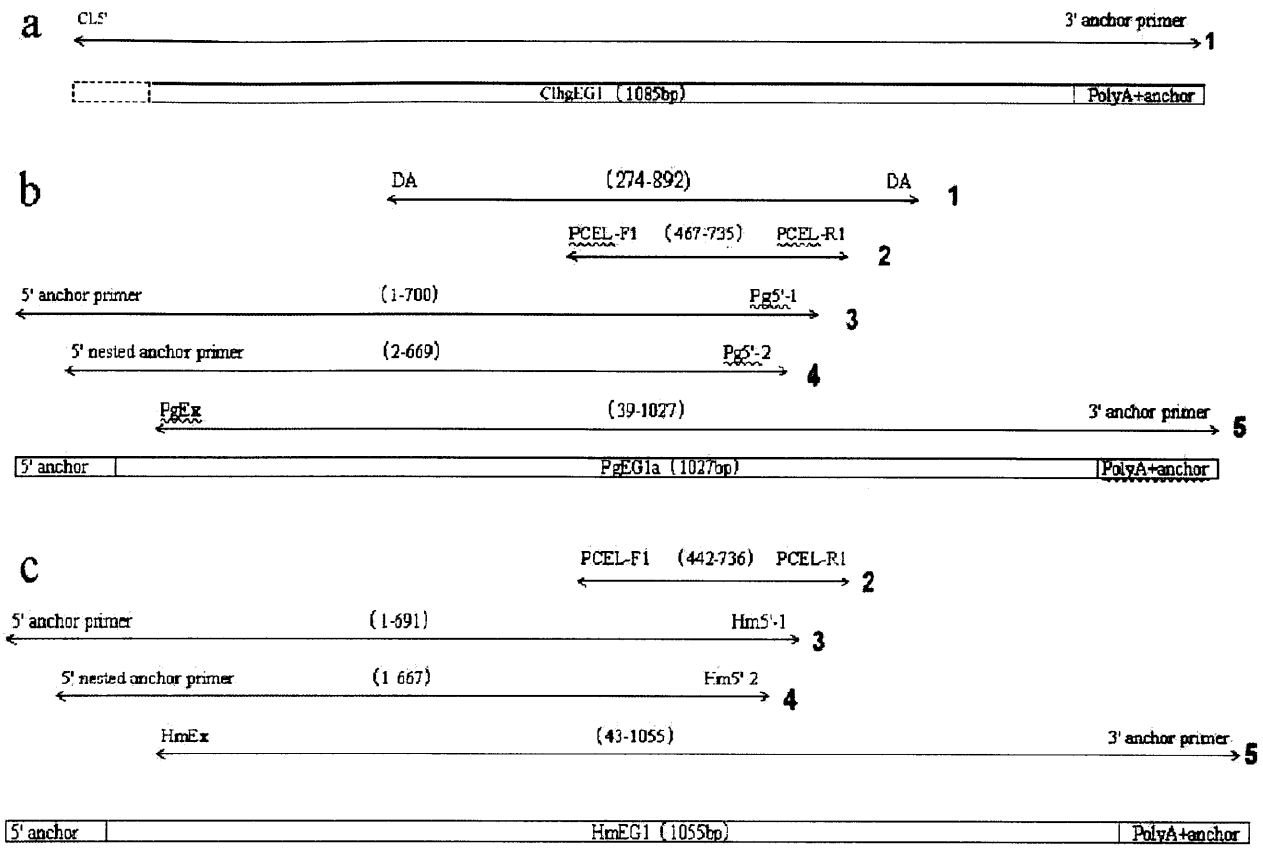

Figure 1. Cloning strategies for endo-β-1,4-glucanase cDNAs from the hindgut of the termites, *Coptotermes lacteus* (*CLhgEG1* [a]) and *C. formosanus* (*PgEG1a* [b] and *HmEG1* [c]). The ranges amplified by PCR are designated by the arrow-headed lines and the nucleotide numbers (5' and 3' terminals in parentheses) including the primer sequences except 5' end of *CLhgEG1* [a]. Numbers in bold characters indicate orders in PCR amplifications of each strategy (a, b, and c). The amplification [**2**] was performed by the single cell RT-PCR, while the other amplifications [**1**, **3**, and **4**] were performed on mRNA extracted from the hindgut contents by RT-PCR. Other homologues (*CLhgEG2*, *PgEG1b-PgEG1h*, *PgEG2*, *PgEG3*, *HmEG2* and *HmEG3*) were also cloned by the same strategy for each host species. Details of the experimental conditions are described in the text.

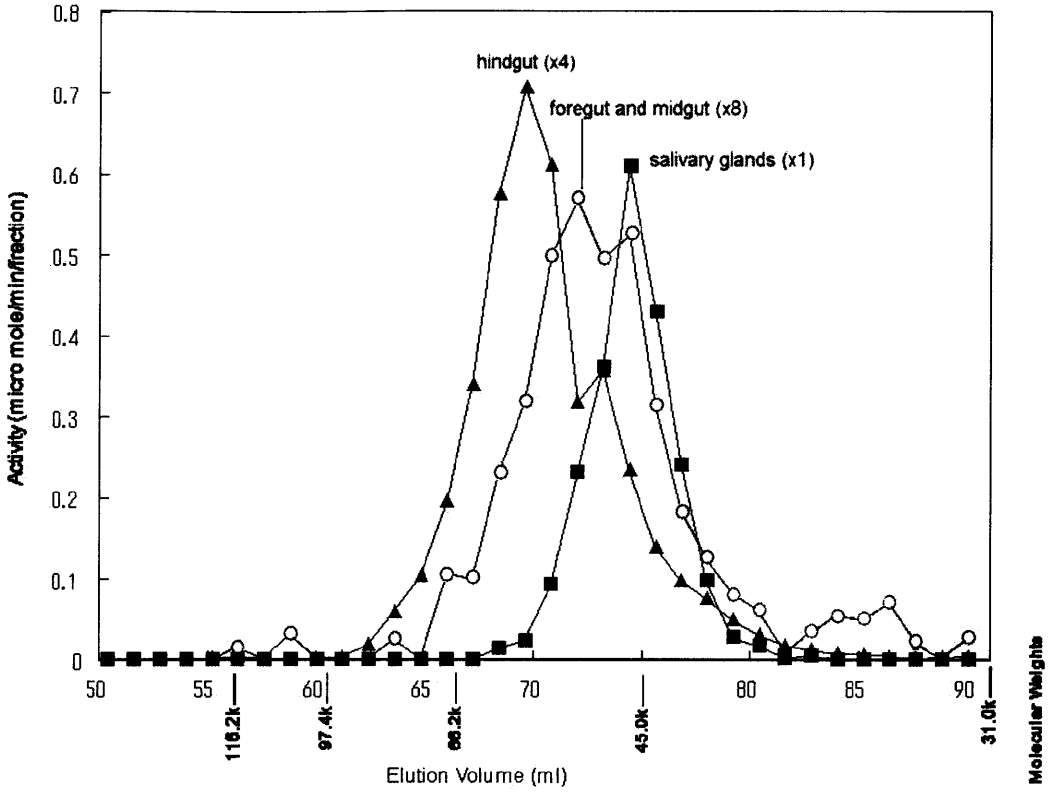

Figure 2. Elution profiles of endo-β-1,4-glucanase activities from the salivary gland (solid squares), the foregut and midgut (open circles), and the hindgut (solid triangles) extracts of the termite, *Coptotermes lacteus* on Superdex 75. Enzymatic activity of the hindgut extract and that of the foregut and midgut extract were magnified to four and eight times the original values, respectively. Details of the experimental conditions are described in the text.

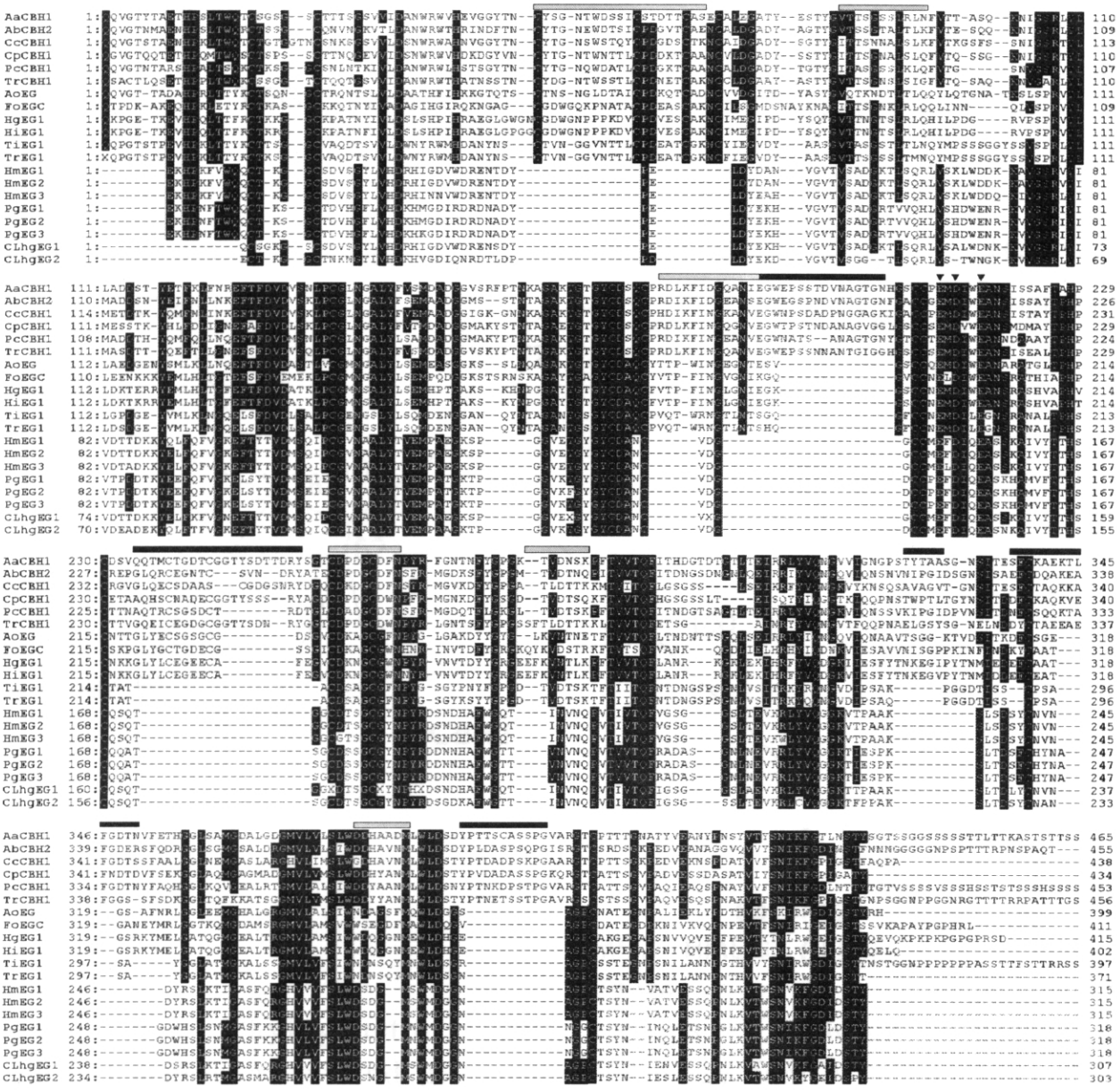

Figure 3. Multiple alignments of amino-acid sequences of protozoan endo-β-1,4-glucanases and fungal cellulases from glycoside-hydrolase family 7 (GHF7). Shaded characters indicate the coincidence of the majority of the amino-acids in each column. The alignment methods are described in the text. CBH and EG in the abbreviated names refer to cellobiohydrases (EC 3.2.1.91) and endo-β-1,4-glucanases (EC 3.2.1.4), respectively. Closed wedges indicate important catalytic residues for GHF7 members corresponding to Glu212, Asp214 and Glu217 of TrCBH1. The GenBank accession numbers are O59843 (AaCBH1 from *Aspergillus aculeatus*), Q92400 (AbCBH2 from *Agaricus bisporus*), Q00328 (CcCBH1 from *Cochliobolus carbonum*), Q00548 (CpCBH1 from *Cryphonectria parasitica*), P13860 (PcCBH1 from *Phanerochaete chrysosporium*), P00725 (TrCBH1 from *Trichoderma reesei* [=*Hypocrea jecorina*]), BAA22589 (AoEG from *Aspergillus oryzae*), P46237 (FoEGC from *Fusarium oxysporum*), Q12622 (HgEG1 from *Humicola grisea*), P56680 (HiEG1 from *H. insolens*), Q12714 (TlEG1 from *T. longibrachiatum*), P07981 (TrEG1 from *Trichoderma reesei*, BAB64553, BAB64561, BAB64562 (PgEG1, PgEG2, PgEG3, respectively, from the hindgut protozoa, *Pseudotrichonympha grassii* of the termite, *C. formosanus*), BAB64563, BAB64564, BAB64565 (HmEG1, HmEG2, HmEG3, respectively, from the hindgut protozoa, *Holomastigotoides mirabile* of *C. formosanus*), BAB089800, BAB089801 (CLhgEG1, CLhgEG2, respectively from the hindgut contents of the termite,

Coptotermes lacteus).

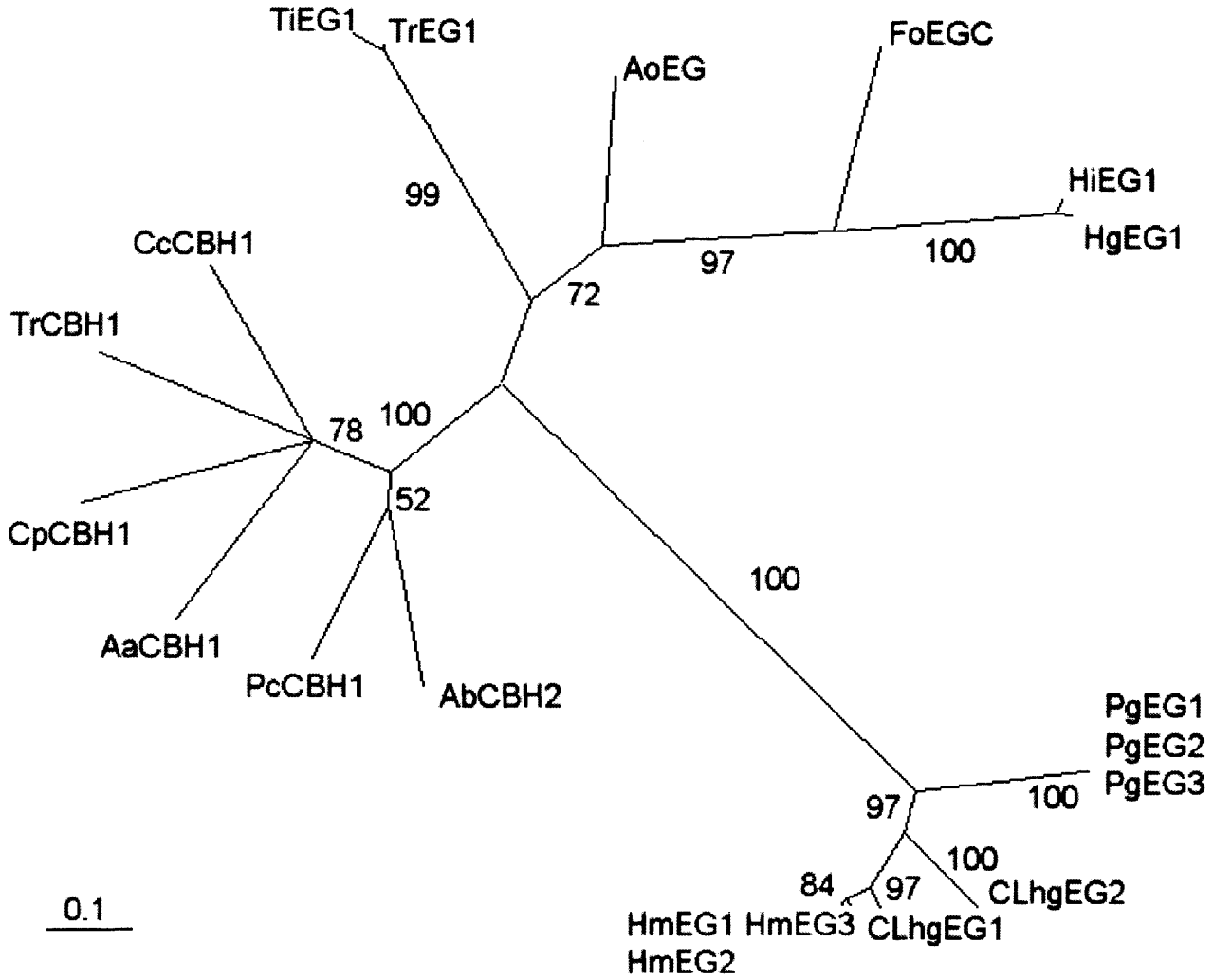

Figure 4. Phylogenetic relationships between protozoan endo-β-1,4-glucanases and glycoside-hydrolase family 7 members. The tree was constructed by the maximum-likelihood method with 282 aligned amino-acid positions, corresponding to the putative catalytic domains, of 12 fungal genes (6 endo-β-1,4-glucanase and 6 cellobiohydrolase genes) in GenBank and putative protozoan endo-β-1,4-glucanases. Numbers at the internal branches indicate the percent occurrence of a group in 1,000 quartet puzzling steps. The scale bar indicates the number of substitutions per position as given by TREE-PUZZLE (Strimmer and Von Haeseler, 1996). The GenBank (protein) accession numbers were described in the footnote of Fig. 3.

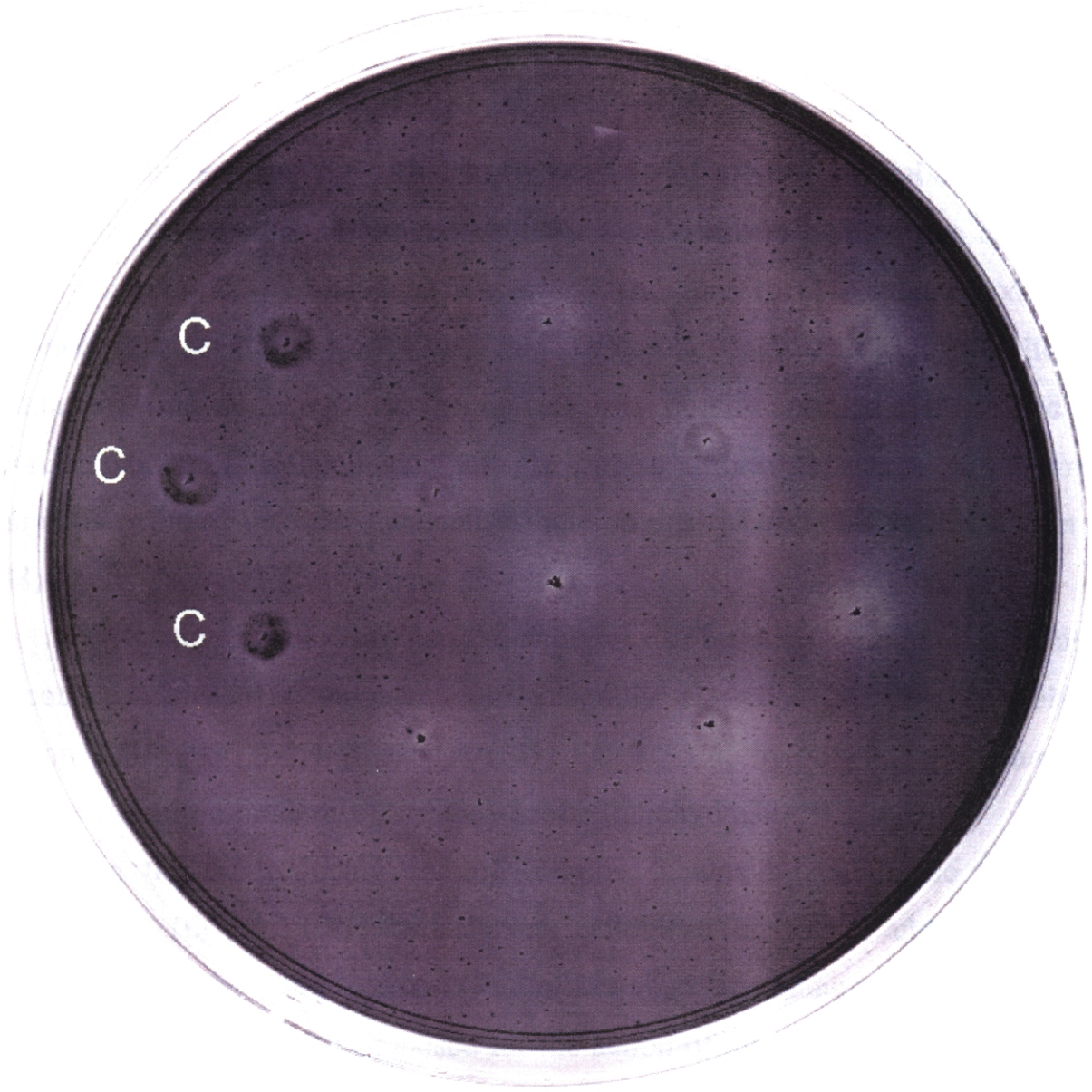

Figure 5. Expression of endo-β-1,4-glucanase in *Escherichia coli*. *PgEG1* cDNA from the hindgut protozoa, *Pseudotrichonympha grassii* of the termite, *C. formosanus* was ligated into pMOSBlue and transformed into *E. coli*. The control (C) was transformed non-inserted pMOSBlue. The halo seen in the image on the right was made by the enzymatic degradation of CMC in agarose. The experimental conditions are described in the text.

GENERAL DISCUSSION

Cellulases of animal origins

Purification and characterization of animal cellulases

The early studies on animal cellulases were mostly quantitative analyses of cellulolytic activities in the digestive organs and thus there were methodological limitations to separating or identifying symbiotic and endogenous cellulases. Qualitative studies using electrophoresis or liquid-chromatography, and the characterization of isolated components appear more persuasive than simple quantitative studies in order to identify their origins (Martin, 1987; Rouland et al., 1988).

Purification has revealed functional differences among cellulolytic components. Isolated components are generally classified by their hydrolyzing functions into endo-β-1,4-glucanase (EG; EC 3.2.1.4; hydrolyzing the middle of cellulose chains randomly), exo-β-1,4-cellobiohydrase (CBH, EC 3.2.1.91; hydrolyzing the non-reducing end of cellulose chains by cellobiose units), and β-glucosidase (EC 3.2.1.21; hydrolyzing cellobiose or longer chains from the non-reducing end by glucose units) (Radford et al., 1996; Tomme et al., 1995). Two different systems have been proposed for the action of these components. One is the synergistic action of free components (EG, CBH and β-glucosidase), while the other involves cellulosomes binding different components into a large structure (Bayer et al., 1998; Beguin and Lemaire, 1996; Tomme et al., 1995). In termites and cockroaches, Slaytor proposed a unique system that lacks CBH components (Slaytor, 1992), while Martin proposed the enrollment of CBH acquired from fungi in many xylophagous species (Martin, 1987). In this chapter, the main focus was given on the purification of endogenous components.

From the cockroach clade (cockroaches + termites, (Lo et al., 2000)), five EG components have been purified to homogeneity. Two components (53.6 and 48.8 kDa) were isolated from the wood-eating cockroach *P. cribrata* (Scrivener and

Slaytor, 1994b), another two (41 and 42 kDa) from the lower termite *R. speratus* (Watanabe et al., 1997), and one more (47 kDa) from the Japanese arboreal higher termite *Nasutitermes takasagoensis* (Tokuda et al., 1997).

In another insect taxon, Coleoptera, which is less related to termites and cockroaches, two β-glucosidase components (57 and 70 kDa) and a CBH component (25 kDa) were purified from larvae of the xylophagous beetle *Ergates faber* (Chararas et al., 1983). Aside from insects, two cellulolytic components (30 and 40 kD; possibly EG or CBH) were purified from crayfish. These cleave β-1,3 glycosidic bonds in addition to β-1,4 glycosidic bonds of cellulose (Xue et al., 1999).

Apart from arthropods, several cellulolytic components have been purified to homogeneity from mollusca. In snails, which are one of the earliest animals reported as being cellulolytic, an EG and a CBH (23 kDa) component were purified from the digestive juice of *H. pomatia* (Marshall, 1973) and *Achatina fulica* (Maeda et al., 1996), respectively. A cellulase seeming to be an EG component (44 kDa) was purified from the gastric teeth of the sea slug *Dolabella auriclularia* (Anzai et al., 1984). The purification of different other EG components (51 kDa on the SDS-PAGE, which is separated into six different 31 kDa components on the DISC-PAGE) from this species has also been reported (Anzai et al., 1988). An EG component was also purified from the marine periwinkle *Littorina* sp (Elyakova et al., 1968). Recently, the purification of a low molecular mass EG (19.702 kDa; predicted from amino acid sequence) was reported in the blue mussel *Mytilus edulis* (Xu et al., 2000).

EG components sometimes act differently from their nomenclature. Cellulolytic enzymes which can hydrolyze CMC are referred to as "endoglucanases," since the number of open-ended cellulose terminals is considered limited on CMC by the frequent substitution of carboxymethyl groups on part of the glucose residues, which prevents the enzyme from attacking from the chain ends to the inside (Klesov, 1991). Many animal EG components have been identified according to this rule as EG, even though many of these cut out cellobiose units when

they attack cellooligomers similarly to CBHs (Anzai et al., 1984; Marshall, 1973; Scrivener and Slaytor, 1994b; Tokuda et al., 1997; Watanabe et al., 1997). This quality of animal EG components may help in the digestion of native-form cellulose in food whose surface area is limited compared to that of CMC. In addition, many xylophagous invertebrates possess grinding or crunching organs, such as the mandibles of insects, the crystalline styles and sacs of bivalves, and the radula teeth of herbivorous mollusca, which can break down food material into fine particles (Marshall and Grand, 1976; Snodgrass, 1993). In termites, ingested wood is ground to less than 50 μm in *C. formosanus* (Itakura et al., 1995) or into two size ranges (10-30 and 100-300 μm) in *N. takasagoensis* (Yoshimura et al., 1996). These grinding and crunching organs may also aid digestion by widening the surface areas of the food debris into substrates, which the cellulase components can then easily attack.

Cloning of animal cellulase cDNAs

Although it has been nearly a century since the first cellulase research began with snails (Emert et al., 1974), the structures of animal cellulases remained unknown until the earliest reports in 1998, when endogenous animal cellulase cDNAs were finally cloned from two nematodes and a termite (Smant et al., 1998; Watanabe et al., 1998).

In the two plant-parasitic cyst nematodes *Globodera rostochiensis* and *Heterodera glycines*, proteins produced in the subventral esophageal glands were surveyed in the course of a phytopathological study of these species (de Boer et al., 1996; Smant et al., 1997). From amino acid sequences of uncharacterized proteins (*svp39* from *G. rostochiensis* and *svp49* from *H. glycines*), four different EG cDNAs (*HG-eng-1*, 1,615 bp encoding a 49.8 kD protein; *HG-eng-2*, 1,191 bp encoding a 34.7 kDa protein; *GR-eng-1*, 1,546 bp encoding a 49.7 kDa protein; and *GR-eng-2*, 1,300 bp encoding a 42 kDa protein) were cloned from cDNA libraries of these species with the aid of degenerate and gene-specific RACE (rapid amplification of cDNA end) amplifications (Smant et al., 1998). Accompanied by a putative signal

peptide (16-27 amino acids), one EG catalytic domain (300-307 amino acids) was encoded for each cDNA from the nematodes (Henrissat et al., 1989; Smant et al., 1998). Following the catalytic domains, *HG-eng-1* and *GR-eng-1* encoded cellulose-binding domains (CBDs) (98 and 96 amino acids, respectively) which were joined to the catalytic domains by linkers (51 and 50 amino acids, respectively) of Pro/Thr/Ser-rich amino acid sequences. *HG-eng-2* encoded only a putative signal peptide and a catalytic domain, while *GR-eng-2* had an additional sequence following after the catalytic domain which corresponded to the linker of HG-eng-1, but had no connecting CBD (Smant et al., 1998). Recently, another EG encoding cDNA, *MI-engl* (1668bp), which encoded a protein (506 amino acids) including a putative catalytic domain and a CBD, was isolated from the root-knot nematode *Meloidogyne incognita* (Rosso et al., 1999), and its precise properties were reported (Bera-Maillet et al., 2000). The putative catalytic domain of MI-engl shared 52.5% identity with those of HG-engl and HG-eng2 (Rosso et al., 1999).

From the phytophagous beetle *Phaedon cochleariae*, a putative EG cDNA was isolated from a gut cDNA library by random sequencing followed by RACE amplifications (Girard and Jouanin, 1999). This putative EG cDNA encoded 242 amino acids including a putative leader peptide consisting of 21 amino acids, and contained about 10 % of the cDNAs identified from the larval-gut cDNA library of this species (Girard and Jouanin, 1999).

From termites and cockroaches, 19 different EG cDNAs have been isolated to date (Lo et al., 2000; Tokuda et al., 1999; Watanabe et al., 1998). From the termite *R. speratus,* fragments of EG cDNA were isolated from a recombinant phage cDNA library using a mouse antiserum raised against one of its EG components, then, following RACE amplifications, two of 1466 cDNAs (*RsEG* and *RsEG2*) were identified (Tokuda et al., 1999; Watanabe et al., 1998). Each cDNA encoded an EG consisting of 448 amino acids including a putative leader peptide (16 amino acid residues), and encoded a single catalytic domain (Tokuda et al., 1999; Watanabe et al., 1998). From another termite, *N. takasagoensis*, a fragment (95 bp) of EG cDNA was

amplified by a degenerate PCR based on the N-terminal amino acid sequence (33 peptides) of the purified EG component (Tokuda et al., 1999). A cDNA (1685 bp, *NtEG* cDNA) including a complete ORF was then obtained by additional gene specific RACE amplifications (Tokuda et al., 1999). The *NtEG* cDNA encoded 448 amino acids consisting of a single peptide (16 amino acids) and a single catalytic domain, which showed 78.3 % identity with the amino acid sequence of RsEG (Tokuda et al., 1999; Watanabe et al., 1998). From *N. walkeri*, another EG cDNA (*NwEG* cDNA, 1585 bp, 98.0% and 97.8% identical to NtEG at nucleotide and amino acid levels, respectively) was cloned using the same gene specific primers designed for *NtEG*. From the wood-eating cockroach *P. cribrata*, an EG cDNA fragment (96 bp) was amplified by a degenerate PCR based on the N-terminal amino acid sequence of a purified EG component, following which, two full length cDNAs (*PcEG1* cDNA, 1415 bp and *PcEG2* cDNA, 1437 bp) were cloned by RACE amplifications (Lo et al., 2000). These showed 63-69% identity with amino acid sequences of EGs from the termites *R. speratus* and *N. takasagoensis* (Lo et al., 2000). From the hepatopancreas of the Australian yabby *Cherax quadricarinatus* (a fresh water crayfish), a putative EG cDNA was identified by degenerate PCR based on the conserved amino acid sequences among EGs from the termite *R. speratus*, the slime mold *Dictyostelium discoideum*, and the kidney bean *Phaseolus vulgaris*, followed by additional RACE amplifications (Byrne et al., 1999). The EG encoded by the cDNA from *C. quadricarinatus* is consisted of 469 amino acids including a putative signal peptide of 24 amino acids (Byrne et al., 1999). From another four cockroach and three termite species in addition to *R. speratus*, *N. takasagoensis*, *N. walkeri* and *P. cribrata*, 14 different fragments (~1000 bp) were isolated by degenerate RT-PCR based on conserved amino acid sequences among EGs from *R. speratus*, *N. takasagoensis*, *N. walkeri*, and *C. quadricarinatus* (Lo et al., 2000).

Structural characters of amino acid sequences of animal EGs

As aforementioned, cellulolytic enzymes are usually classified into EG, CBH,

and β-glucosidase based on IUB-MB enzyme nomenclature (http://www.expasy.ch/enzyme/) which does not reflect the structural features of the enzyme molecules and, thus, neither their phylogenetic relationships (Henrissat, 1991). To compensate for this problem, Henrissat et al. advocated the classification of cellulolytic enzymes (Henrissat et al., 1989) by hydrophobic cluster analysis of amino acid sequences, which reflects structural similarities in the grouping (Gaboriaud et al., 1987). They later broadened the application of their method to include all *O*-glycoside hydrolases (EC 3.2.X.X) by using the direct comparison of amino acid sequences in addition to hydrophobic clustering, and then classified them into 78 families made up of 10 clans (groups of families showing a structural resemblance with each other) to date (http://afmb.cnrs-mrs.fr/CAZY/index.html) (Henrissat and Bairoch, 1996; Henrissat and Davies, 1997). Therefore, the cloning and sequencing of cellulolytic enzyme precursors from animals can be used not only to clarify their primary structure and possible catalytic functions, but also to help elucidate their phylogenetic origins.

The catalytic domains of five EGs from the cyst nematodes *H. glycines* and *G. rostochiensis* and the root-knot nematode *M. incognita* were all placed in the glycoside hydrolase family (GHF) 5 by hydrophobic cluster analysis (Bera-Maillet et al., 2000; Smant et al., 1998) or homology in amino-acid sequence (Rosso et al., 1999). GHF5 members share a general motif of $(\alpha/\beta)_8$ barrel structure (Henrissat and Bairoch, 1996). The three dimensional structure of one member was investigated by X-ray crystallography of CelCCA from the gram-positive bacterium *Clostridium cellulolyticum* (Ducros et al., 1995). GHF5 members generally conserve seven residues which correspond to Arg79, His122, Asn169, Glu170 (assumed as a proton donor of the catalytic center, which pulls away a proton from a water molecule and gives it to the non-reducing side of the β-1,4 glycosidic bond being catalyzed), His254, Tyr256, and Glu307 (assumed as a nucleophile of the catalytic center, which attracts an OH residue from the water molecule and gives it to the anomeric carbon at the reducing side of the β-1,4 glycosidic bond being catalyzed) of CelCCA (Ducros et

al., 1995). Among these residues, the catalytic domains of five nematode EGs preserved those corresponding to Arg79, His122, Asn169, Glu170 and Glu307 of CelCCA (Fig. 1). In contrast, *C. elegans,* for which the entire genomic sequence is known (Bird et al., 1999), does not possess any cellulase genes (Bera-Maillet et al., 2000; Yan et al., 1998). Since the nematode EGs showed the highest homology to bacterial GHF5 members, a possible ancient horizontal gene transfer from prokaryote has been advocated (Yan et al., 1998).

EG coded by cDNA from the beetle *P. cochleariae* (*PcEG*) was placed in GHF45 because of its similarity in amino acid sequence (Girard and Jouanin, 1999). The three dimensional structure of GHF45 EG was investigated in EG V from the fungus *Humicola insolens*, which is composed of six β-barrels, three α-helixes, and inter-domain regions connecting these structural units (Davies et al., 1993). The nucleophile and proton donor of matured EG V are Asp10 and Asp121, respectively, and are thought to be critical for activity. However, the region including the nucleophile Asp10 of EG V (Thr6-Cys16) was preserved in matured PcEG (Thr22-Cys32), whilst the region including nucleophile Asp121 in EG V (Asn118-Ile124) was not preserved (Fig. 2).

EG cDNAs isolated from termites, cockroaches and crayfish were all placed in GHF9 by amino acid homology with other GHF9 members (Fig. 3) (Lo et al., 2000; Tokuda et al., 1999; Watanabe et al., 1998). The three dimensional structure of GHF9 was investigated in the catalytic domain (E4-68) of endo/exocellulase E4 from the fungus *Thermomonospora fusca* by X-ray crystallography (Sakon et al., 1997). The catalytic center of the E4-68 bound a cellopentaose residue, and cleaved it into cellotetraose plus glucose or cellotriose plus cellobiose, with an inversion of the configuration at the anomeric carbon. The preservation and substitution of amino acid residues among the catalytic centers of E4-68 and animal EGs in GHF9, for which complete protein sequences are known, are listed in Table 1. Among the18 amino acid residues located beside the cellohexaose bound by the catalytic center of E4-68 (Sakon et al., 1997), all residues positioned next to Glc (glucose) (-4) and Glc (-3)

except for Asp261 are substituted in animal GHF9 EGs (Table 1). The other amino acid residues positioned beside Glc (-2) to Glc (+2) (Sakon et al., 1997) are well preserved except for substitutions of Phe205 to Tyr in PcEG1; His376 to Ser in PcEG1 and PcEG2; and Ile388 to Thr in RsEG, RsEG2, PcEG1, and PcEG2, Ile388 to Phe in NtEG and NwEG, or Ile388 to Thr in CqEG (Table 1). The nucleophile Asp55, co-nucleophile Asp58, and the proton donor Glu424 (Sakon et al., 1997) are conserved in all animal GHF9 EGs listed in Table 1.

All known animal EGs, except for HG-eng1 and GR-eng1 from plant parasitic nematodes, have consisted only of catalytic domains (Byrne et al., 1999; Girard and Jouanin, 1999; Lo et al., 2000; Tokuda et al., 1999; Watanabe et al., 1998). However, many fungal cellulase components active against crystalline cellulose consist not only of catalytic domains, but also cellulose binding domains (CBD) and linkers of a Pro/Thr/Ser-rich residue, which hinge catalytic domains and CBDs (Beguin and Aubert, 1994; Tomme et al., 1995). Although it has been proposed that the large production of EGs generally seen in termites and cockroaches would compensate for the low activity of EGs on crystalline cellulose (Slaytor, 1992), an underlying mechanism for the efficient degradation of crystalline cellulose by insect cellulases that do not include CBDs has yet to be clarified.

Animal cellulase genes

As aforementioned, there have been several reports on the cDNA cloning of cellulases from higher animals, but few on the genomic structure and sequences of their cellulase genes. Only reports on plant parasitic nematodes and the higher termite *N. takasagoensis* have covered the entire genomic structure of their cellulase genes.

In two plant parasitic nematodes, *G. rostochiensis* and *H. glycines* , two EG genes from each (*GR-eng1* , 2492 bp and *GR-eng2*, 2388 bp; and *HG-eng1*, 2151 bp and *HG-eng2*, 2324 bp, respectively) were isolated from recombinant phage genomic DNA libraries by probing the EG cDNAs (Yan et al., 1998). These four genes had seven introns at exactly corresponding positions in the catalytic domain coding

regions. Additionally, *GR-* and *HG-eng2* had one additional intron at corresponding positions in the putative CBD coding regions. An entire *NtEG* gene was isolated from the higher termite *N. takasagoensis* by combining PCR amplification of genomic DNA with sets of primers designed from NtEG cDNA sequences. A complete 13,005 bp gene including 10 exons separated by nine introns was then identified. In *R. speratus*, three introns of *RsEG* were identified interrupting exons at the corresponding positions and phases of *NtEG* (Tokuda et al., 1999). Other corresponding introns of *RsEG* and *RsEG2* have also been identified from the genomic DNA of *R. speratus* (data unpublished), however, no positional correlations were identified between the introns from these termite EG genes and those from other known EG genes in GHF9 (Tokuda et al., 1999).

In addition to the presence of introns, these animal cellulase genes maintain other qualities shared with other known eukaryotic genes. They all have TATA boxes in their upstream regions, and polyadenylation and cleavage signals (GATAAA) in appropriate regions; and their exon/intron borders follow the eukaryote "GT-AG" rule except for four minor cases in the nematode EG genes (GC instead of GU at the 5'end) (Tokuda et al., 1999; Yan et al., 1998). Other *cis* regulatory elements generally found in eukaryotic genes have also been found in these EG genes, such as a consensus core binding site sequence and a repressor/enhancer element sequence from nematode EG genes (Yan et al., 1998), and a CCAAT promoter sequence in the upstream region of *NtEG* (Tokuda et al., 1999).

The structure of the EG genes from these nematodes and this termite well demonstrate their eukaryotic and endogenous origins. The presence of corresponding introns among all four nematode EG genes and the complete absence of EG genes in other nematodes including *C. elegans* (Bird et al., 1999) imply that a possible ancestral EG gene among the four EG genes existed before the divergence of these plant parasitic nematodes (Yan et al., 1998). Thus, an ancient horizontal transfer (Syvanen, 1994) is suggested as a possible explanation for the acquisition of EG genes by the cyst nematodes (Yan et al., 1998). Highly homologous GHF9 EG

precursors, which form a monophyletic clade among other GHF9 members, have been isolated from termites, cockroaches and the crayfish *C. quaadricarinatus* (Byrne et al., 1999; Lo et al., 2000; Tokuda et al., 1999; Watanabe et al., 1998), connoting an inheritance and a divergence of an ancestral EG gene through radiation from a common arthropod progenitor. However, there is not enough evidence to establish these hypotheses as firm theories. Further investigations of EG genes from other animals and/or elucidation of other genes in animals maintaining EG genes are required to compare the genomic structure and GC contents among them and to complete the molecular phylogenetic analysis of EG genes.

Expression sites of animal EG genes

The organs producing endogenous cellulase components have been surmised on the basis of the distribution of cellulolytic activities in the digestive systems. This method is useful as an approach to locate the conceivable endogenous production of cellulolytic components, but is insufficient to conclude their origins. Simple measurements of cellulolytic activities in particular excreting organs cannot discriminate authentic endogenous activities from production by intracellular endosymbionts. Immunohistochemical observation using an antiserum raised against purified cellulolytic components can provide reasonable accuracy (Watanabe et al., 1997). However, ideally, a combination of molecular methods would be most desirable to meet today's requirements.

Whereas in the above-mentioned research, the presence of each EG gene on the host genome was confirmed by isolation of the gene or its fragment and/or by southern blotting, their expression was localized primarily by RT-PCR and *in situ* hybridization (Byrne et al., 1999; Smant et al., 1998; Tokuda et al., 1999; Watanabe et al., 1998). In plant parasitic nematodes, in addition to the cloning of EG cDNAs by degenerate PCR from N-terminal amino acid sequences, expression in the subventral glands was confirmed by *in situ* hybridization (Smant et al., 1998). The expression of putative EG gene (*CqEG*) in the crayfish *C. quadricarinatus* was located in the

hepatopancreas by site specific cDNA cloning using degenerate RT-PCR and *in situ* hybridization with gene specific probes (Byrne et al., 1999). For higher and lower termites, contrasting expression sites have been clarified. In the higher termite *N. takasagoensis*, the expression of the EG gene (*NtEG*) was located in midgut columnar cells by *in situ* hybridization and additional RT-PCR using tissue specific templates and gene specific primers (Tokuda et al., 1999). On the other hand, in the lower termite *R. speratus*, the localization of antigenic EG proteins was confirmed in the salivary glands by immunohistochemistry, and the EG gene expression was confirmed by tissue specific RT-PCR with gene specific primers (Tokuda et al., 1999; Watanabe et al., 1997; Watanabe et al., 1998). The midgut and hepatopancreas are phylogenetically homologous organs of arthropods (Brusca and Brusca, 1990), and therefore it is understandable that the higher termite and the crayfish expressed similar EGs in those tissues. However, the expression of EGs in the salivary glands of lower termites (*R. speratus*, *Neotermes koshunensis*, *Hodotermopsis japonica* and *M. darwiniensis*) (Lo et al., 2000) implies a dynamic convergence of EG expression sites and the digestive mechanism in the course of termite divergence.

Phylogenetic origins of animal cellulase genes and future insight for animal cellulase study

The historical debate concerning the presence of endogenous cellulase in higher animals has been settled by molecular evidence. There is no doubt that cellulases are proteins which can potentially be produced in any life form holding cellulase genes. Indeed, cellulase genes have been reported from a variety of metazoans as well as other life-forms (Fig. 4). In contrast to other digestive enzymes (i.e. β-amylase), animal cellulases have been found in three non-related families (GHF 5, 9, 45). GHF5 cellulase genes were reported from plant-parasitic nematodes and a longicorn beetle, which are both placed in GHF5 subfamily 2 (Sugimura et al., 2003; Yan et al., 1998) and formed a monophyletic group, but have not reported other nematode or beetle species (Lo et al., 2003). GHF9 cellulase genes were reported

from termites, cockroaches, an abalone, and a sea-squirt (Dehal et al., 2002; Lo et al., 2000; Suzuki et al., 2003; Watanabe and Tokuda, 2001), and they shared several intron positions each other (Lo et al., 2003). GHF45 cellulase genes were reported from a bivalve (Xu et al., 2001) and a leaf beetle (Girard and Jouanin, 1999). In all cases, still current available data do not give enough resolutions to form solid phylogenetic trees expect that GHF5 members from longicorn beetles and plant-parasitic nematodes formed a monophyletic clade (Lo et al., 2003). Although this situation, all of reported animal cellulase implies a possible common ancestral cellulase genes of animals in the latest ancestor of bilaterian animals (Fig. 3) (Lo et al., 2003).

However, this conclusion does not answer the query, "Why do some animals have endogenous cellulases and others do not?" or "Why isn't cellulase a common enzyme in animals like amylase or proteinase?" Further investigations into animal cellulose digestion will eventually answer these questions. Although recent rushing findings of animal cellulase genes imply dynamic horizontal interchanges of ancestral genes among different organisms, the acquisition of cellulase gene by animals would not so simple and late event in phylogenetic development of animals. As more diverse cellulases come under molecular investigation, their origins and roles in animals will be understood more precisely.

Interactions of cellulolytic symbionts and host animals

Protozoan cellulases and their synergism with termite cellulases

Current study showed that lower termites have two sources of cellulolytic enzymes, which produce enzymes different qualitatively. The salivary glands of *R. speratus* produced endo-β-1,4-glucanase components belonging to the glycosyl-hydrolase family 9 (GHF9), whereas the hindgut protozoan fauna of *C. formosanus* had another type of endo-β-1,4-glucanase components, which could be placed in GHF7. The presence of *RsEG* homologues in *C. formosanus* and its expression in the salivary glands were confirmed by RT-PCR (Nakashima et al.,

2002a) The presence of protozoan GHF7 cellulase genes (encoding both of endo-β-1,4-glucanase-like and the exocellobiohydrase-like amino-acid sequences) in the hindgut of *R. speratus* were also confirmed (Moriya et al., 2002). Thus, cellulose digesting systems of both species of lower termites consist of endogenous and protozoan cellulolytic enzymes, which have phylogenetically different origins from each other, and the prediction for the presence of endogenous termite cellulases by Yokoe (1964) and the theory of dual cellulose digesting system in termites by Yamaoka and Nagatani (1975) were confirmed, although the mode of action of both enzymes have not been cleared. Yamaoka and Nagatani (1975) proposed a synergism between endogenous and protozoan cellulases, which is that the endogenous enzyme from the salivary glands was assumed a Cx-cellulase (endo-β-1,4-glucanase; EG), while the protozoan cellulase was assumed a C1-cellulase (exocellobiohydrase; CBH) hydrolyzing crystalline cellulose actively (later it was found that the salivary cellulase is active against crystalline cellulose alone in current thesis (Watanabe et al., 1997)). They also assumed that the Cx-cellulase was absorbed in the protozoan cells with cellulose particles and act for crystalline cellulose digestion with the protozoan C1-cellulase. The original C1-Cx synergism for the fugal cellulase system by (Reese et al., 1950) was assumed that for C1- and Cx-cellulases act simultaneously against crystalline cellulose, and Yamaoka and Nagatani (1975) seemed to follow this theory. In *R. speratus*, the hindgut extract did not react against the anti-serum against YEG2 (Watanabe et al., 1997), and also in *C. formosanus*, the result does not support for the simultaneous presence of endogenous and protozoan cellulases in the hindgut (Nakashima et al., 2002a). Because of the presence of the pylorus valve narrowing the connection of the midgut and the hindgut, the digestive fluid including salivary (and midgut for *C. formosanus*) enzyme and digested soluble sugars would flow backward thought the lumen outside of the peritrophic membrane in the midgut. Thus, the amount of the endogenous cellulases entering into the hindgut would be limited. My colleagues and I found the expression of EG by the hindgut protozoa *P. grassii* and *H. mirabile* from *C. formosanus* in the current study (Watanabe et al., 2002), and

also CBH homologues only from *P. grassii* (Nakashima et al., 2002b). These results suggest that *P. grassii* has a combination of cellulolytic enzymes acting simultaneously by them selves, while a simultaneous synergism between the termite and protozoan cellulases is not seemed to be essential for those termites. Demonstration of microcrystalline-cellulose digesting activity in the extract of the termite symbiotic protozoa *Trichomitopsis termopsidis* cultured in a formulated medium also supports independency of protozoan cellulose digestion from the termite endogenous cellulase (Odelson and Breznak, 1985). Still, ingested cellulose would be partially attacked by the salivary enzymes before the protozoan cellulases affect it, because that the salivary (termite-own) canal connected to the esophagus located far upstream of the hindgut. In this meaning, endogenous enzymes might act synergistically with protozoan enzymes (the protozoan enzymes might show higher activity against crystalline cellulose which was pre-affected by the salivary enzymes than that against intact crystalline cellulose) in *R. speratus* and *C. formosanus*. While recently, a simultaneous presence of GHF9 EG and protozoan GHF45 cellulases (at least active against CMC) in cells of a termite-symbiotic protozoon *Koruga bonita* of the termite *Mastotermes darwiniensis* were reported. Because the GHF9 EG has an identical N-terminal sequence to the endogenous EG expressed in the salivary glands of *M. darwiniensis* and only GHF45 cDNA were expressed in *K. bonita*, it was concluded that GHF9 EG was acquired by the protozoa and help cellulose digestion in the protozoan cells (Li et al., 2003). Although, the function and of the GHF9 EG and the GHF45 cellulases have not been cleared, presence or absence of synergistic actions between termite-endogenous EGs and protozoan cellulases in each lower termite would not be over-generalized.

Protozoan cellulases and microbe community in host termites

The expression both of EG and CBH in the protozoa *P. grassii* (but not in *H. mirabile* which expressed only EG) (Nakashima et al., 2002b; Watanabe et al., 2002) agree with the observation by Kanai and Azuma (1982). They fed a variety of

saccharides to the workers of *C. formosanus* and found that *P. grassii* survived after 4 weeks only when the host termites were served celluloses (Avicel (crystalline cellulose) or wood), but could not survive for one week when the hosts were starved or served with saccharides or acetic acid other than celluloses and wood. While, *H. hartmanni* (= *H. mirabile*) and *S. leidyi* survived over a week in the host termites fed on acetic acid or all tested saccharides including polysaccharides with α -1,6-linkages, which tested workers could not survive on beyond eight weeks. These two protozoa even survived in more than quarter of tested termites that was starved for a week. It was concluded that *P. grassii* is primarily responsible for crystalline cellulose digestion (Kanai and Azuma, 1982). These results would also be interpreted that *P. grassii* is completely dependent on cellulose as its substrate while *H. hartmanni* and *S. leidyi* depends on the metabolite of the other hindgut microbes of the host termite. Moriya et al. also found cDNAs encoding CBH homologues from the protozoan fauna in the hindgut of *R. speratus* in addition to those of EG homologues (Kudo et al., 2003; Moriya et al., 2003; Ohkuma et al., 2002).

Itakura et al (2003) estimated that *C. formosanus* produces 40% of total pyruvate generated in the whole body, which are calculated by the pyruvate-dehydrogenase activity and carbon-dioxide emission from the worker termites (Itakura et al., 2003). Since pyruvate is direct product from glucose by glycolysis (one glucose molecule is converted into two molecules of pyruvate and ATP), the report suggest that the hindgut protozoa of *C. formosanus* contribute 60% of energy production, and this value is similar to the estimation by Inoue et al. (1997) that attributed 70% of total cellulase activity to the hindgut protozoan fauna in *R. speratus* (Inoue et al., 1997).

Because cellulose digestion of the hindgut protozoa is occurred inside the protozoan cells (Yamaoka, 1979), the symbiotic protozoa do not release free glucose to the host but supply energy obtained from cellulose in acetate form to the host termite (Breznak, 1982), but release acetate, molecular-form hydrogen, and carbon-dioxide like acetogenic anaerobes since Axostylata protozoa including

symbionts from the lower termites (Oxymonadida, Trichomonadida and Hypermastigida) have hydrogenosome but lack mitochondria (Muller, 1993).

Cellulose → Glucose

$$\text{Glucose } (C_6H_{12}O_6) + 2\ H_20 \rightarrow 2\ CH_3COOH + 2\ CO_2 + 4\ H_2$$

The generated carbon-dioxide and molecular-form hydrogen was consumed by CO2-reducing acetogenic bacteria and turned into another molecule of acetate and water (Brauman et al., 1992),

$$4\ H_2 + 2\ CO_2 \rightarrow CH_3COOH + 2\ H_2O$$

or, converted into methane and water by methanogenic Archaea *Methanobrevibacter* sp (Breznak and Brune, 1994) .

$$4\ H_2 + CO_2 \rightarrow CH_4 + 2\ H_2O$$

The protozoan habitat of the lower termites harbors the CO_2-reducing acetogenic bacteria and the methanogenic Archaea *Methanobrevibacter* sp. (Leadbetter and Breznak, 1996; Leadbetter et al., 1998; Ohkuma et al., 1999). *Methanobrevibacter* sp was also found inside of Trichomonadida protozoa (Lee et al., 1987) and appeared to be endosymbionts harbored within protozoan cells (Lee et al., 1987; Odelson and Breznak, 1985). No cellulase genes have been found from *Methanobrevibacter* sp or other possible endosymbionts in the protozoan cells to date. It is estimated that acetate generated by the CO_2-reducing acetogenic bacteria fulfills 1/3 of physiological energy demands of the host termite in *R. flavipes* (Odelson and Breznak, 1983) and with the amount produced by protozoa, it would be 60-70% of the total energy consumption calculated from cellulolytic activity and pyruvate production in *R. speratus* and *C. formosanus* (Inoue et al., 1997; Itakura et al., 2003), as

aforementioned. Since there are no reports for major cellulolytic microbes from the termite hindgut microbes other than Axostylata protozoa to the author's knowledge (only facultative cellulolytic bacteria was reported in the hindgut of *R. hesperus* (Thayer, 1976)), it would be concluded that cellulolytic protozoa, especially like the species possessing both of EG and CBH genes like *P. grassii* of *C. formosanus*, is one of the main engines for the lower termites as much as the endogenous cellulase is. In another word, because of the presence of protozoan cellulase genes, the microbe communities of the lower termites were formulated.

Table 1. Preservation and substitution of amino acid residues at the catalytic sites of endo/exocellulase E4 from *Thermomonospora fusca* among endo-β-1,4-glucanases from animals.

Glucose residue		Glc(-4)		Glc(-3)			Glc(-2)			Glc(-1)			Glc(+1)					Glc(+2)	
(protein) E4-68	(accession) AAB42155.1	D261	D262	W209	W256	Y318	F205	W260	R317	D55	D58	W313	H125	H376	R378	Y420	E424	W128	I388
RsEG	BAA31326.1		N	G	G	H													Y
RsEG2	BAA34650.1		N	G	G	H													Y
NtEG	BAA33708.1		S	A	G	H													F
NwEG	BAA33709.1		S	A	G	H													F
PcEG1	AAF80584.1		Q	Y	P	M	Y							S					Y
PcEG2	AAF80585.1		Q	N	S	M								S					Y
CqEG	AAD38827.1					H													T

The amino acid residues from E4-68 (catalytic domain of endo/exocellulase E4) show hydrogen bonding to the glucose residues designated, except for W128 and I38 which lie aside Glc(+2), but show no particular hydrogen bonding (Davies et al., 1993). Only substituted amino acids are displayed. Empty boxes mean the identical amino acids are located at corresponding positions in each protein. Amino acids are designated in single character notations. Accessions are the GenBank/DDBJ protein i.d. number of each protein. Protein sequences were aligned using Clustal X 1.8 for Windows (National Center for Biotechnology Information (Bldg 38A, NIH 8600 Rockville Pike, Bethesda, MD 20894, USA) with Gonnet 250 pairwise parameters.

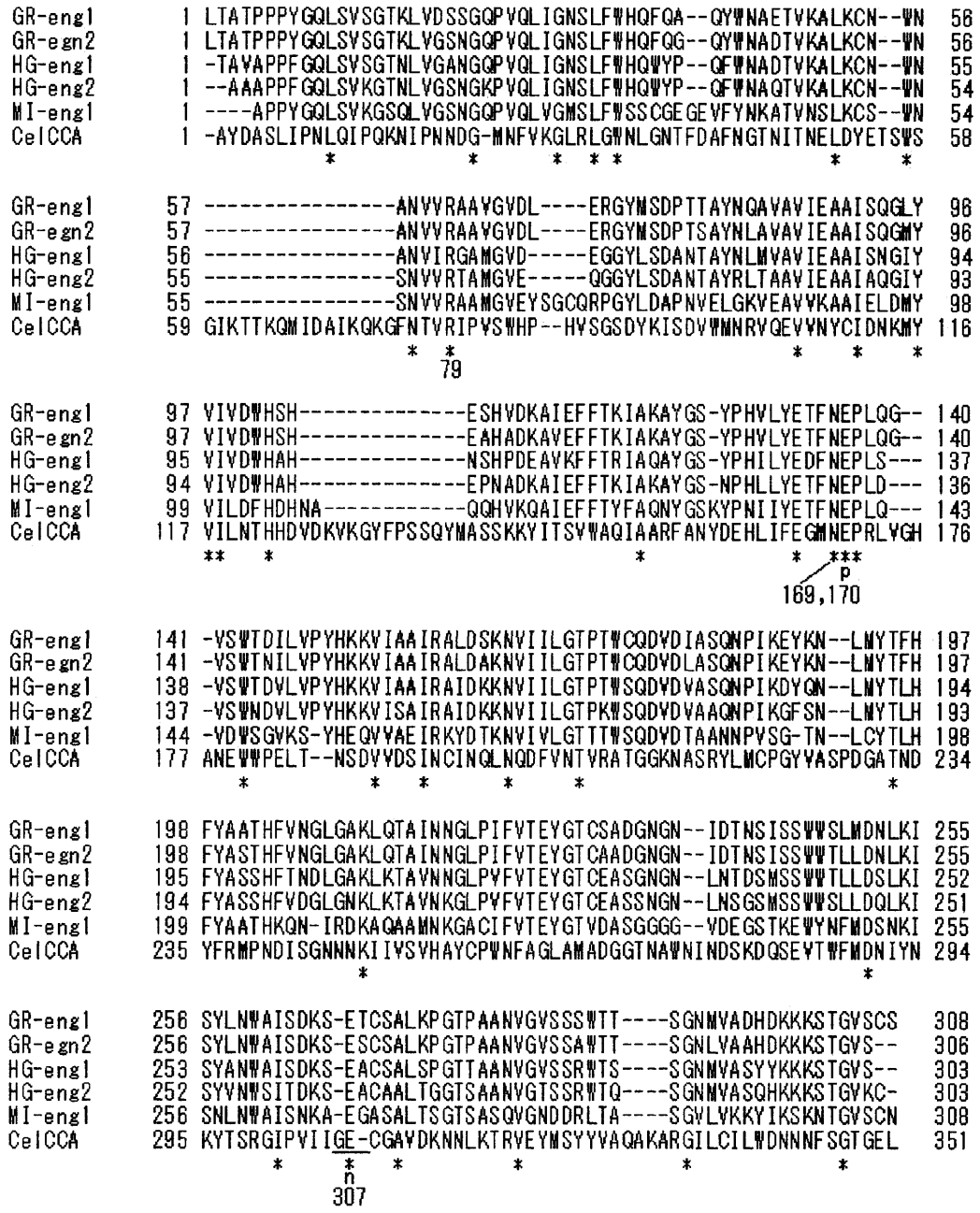

Figure 1. Multiple alignments of catalytic domains of endo-β-1,4-glucanase components in the glycoside hydrolase family 5.

Alignment of endo-β-1,4-glucanase components from the bacterium *Clostridium cellulolyticum* (CelCCA [AAA23222] (the numbers in [ ] are the protein i.d. of GenBank/DDBJ or Swissprot) and the plant-parasitic nematodes *Globodera rastochiensis* (GR-engl [AAC48325.1] and GR-eng2 [AAC48341.1]), *Heterodera glycines* (HG-engl [AAC48327.1] and HG-eng2 [AAC48326.1]), and *Meloidogyne incognita* (MI-engl [AAD45868.1]). The alignments were achieved using the computer program CLUSTAL X (1.8) for Windows (National Center for Biotechnology Information, Bldg 38A, NIH 8600 Rockville Pike, Bethesda, MD 20894, USA) with Gonnet 250 pairwise parameters in matured-form amino-acid sequences. "p"and "n" under the columns indicate proton donor and nucleophile residues of the catalytic center, respectively. The N-terminal of matured MI-engl was predicted using the SignalP V2.0.b2 server (http://www.cbs.dtu.dk/services/SignalP-2.0/ , Center for Biological Sequence Analysis Department of Biotechnology, The Technical University of Denmark). The underlined columns in were re-aligned manually. The numbers under the columns indicate the amino-acid number in CelCCA. Asterisks under columns indicate complete coincidence of the amino acid residues in the columns.

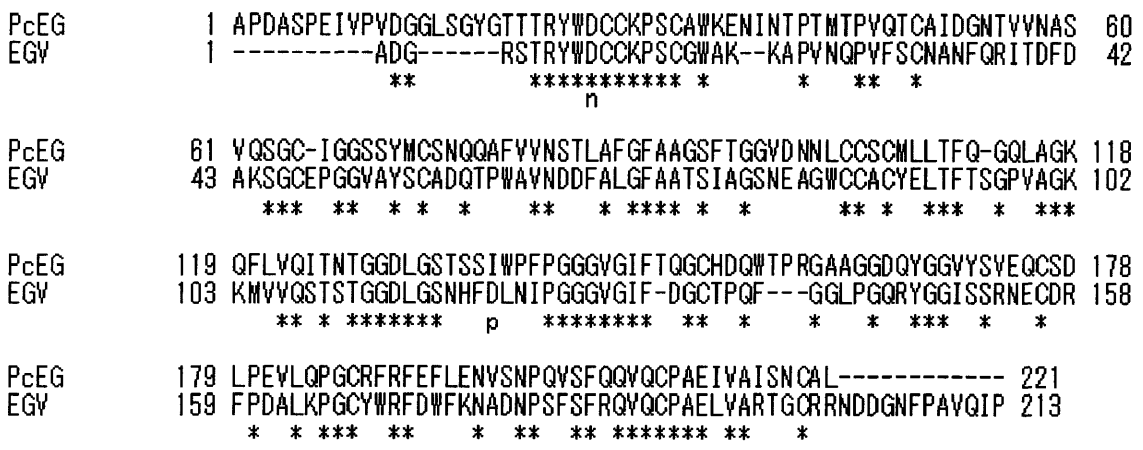

Figure 2. **Multiple alignments of catalytic domains of endo-β-1,4-glucanase components in the glycoside hydrolase family 45.**
Alignment of endo-β-1,4-glucanase components from the fungus *Humicola insolens* (EGV [P43316]) and the beetle *Phaedon cochleariae* (PcEG [CAA76931.1]). The alignment was accomplished by the same method for Fig. 1. "p" and "n" under the columns indicate proton donor and nucleophile residues of the catalytic center, respectively. Asterisks under columns indicate complete coincidence of the amino acid residues in the columns.

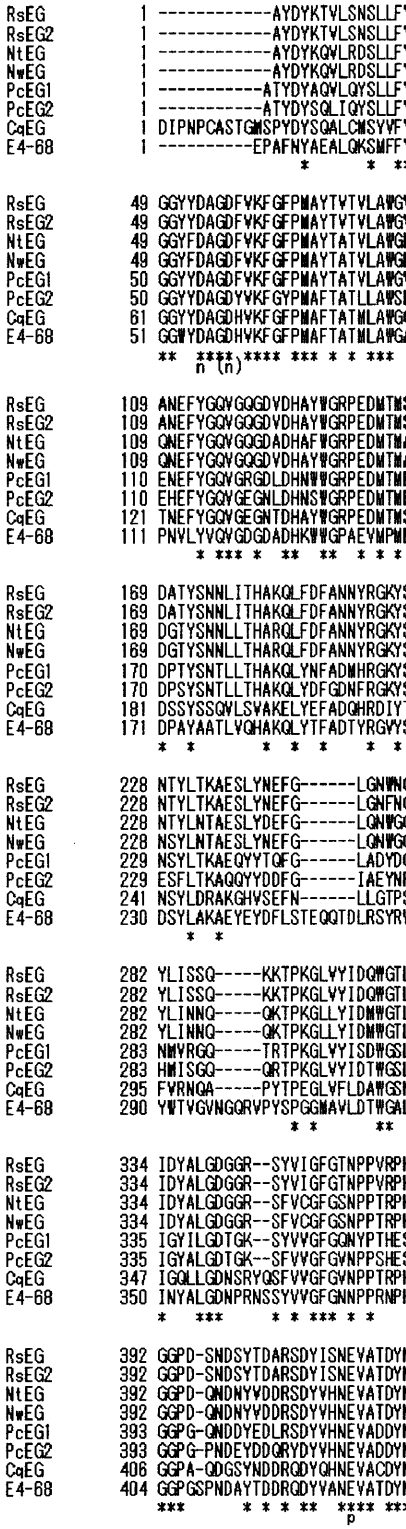

Figure 3. Multiple alignments of catalytic domains of endo-β-1,4-glucanase components in the glycoside hydrolase family 9.

Alignment of endo-β-1,4-glucanase components from the bacterium *Thermomonospora fusca* (E4-68 (catalytic domain of E4 [part of AAB42155.1]), the crayfish *Cherax quadricarinatus* (CqEG [AAD38027.1]), and the termites *Reticulitermes speratus* (RsEG [BAA31326.1] and RsEG2 [BAA34050.1]), *Nasutitermes takasagoensis* (NtEG [BAA33708.1]), *Nasutitermes walkeri* (NwEG [BAA33709.1]), *Panesthia cribrata* (PcEG1 [AAF80584.1] and PcEG2 [AAF80585.1]). The alignment was accomplished by the same method for Fig. 1. "p", "n" and "(n)" under the columns indicate proton donor, nucleophile and co-nucleophile residues of the catalytic center, respectively. Asterisks under columns indicate complete coincidence of the amino acid residues in the columns.

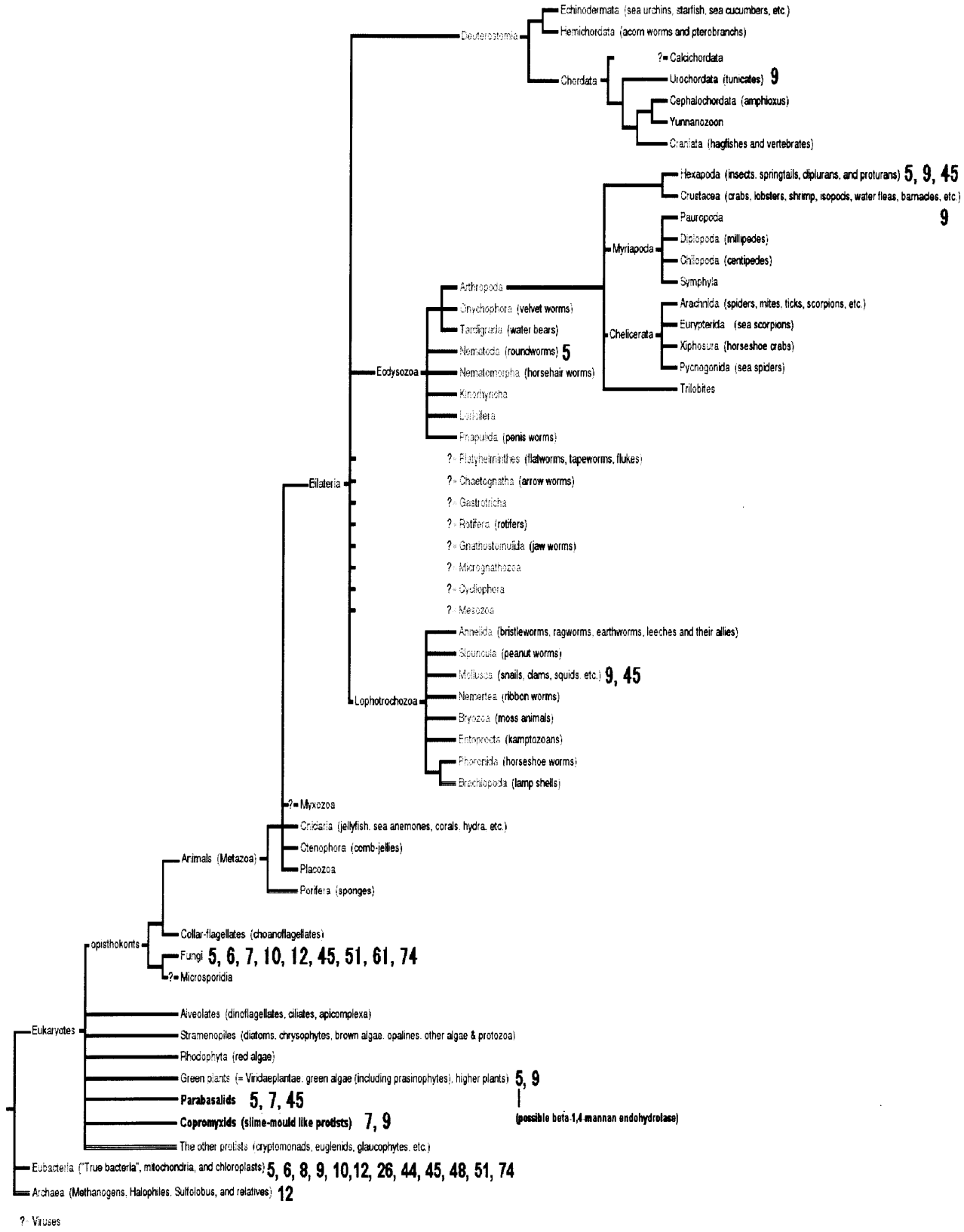

Fig. 4. Distribution of cellulase genes among life-forms.
Numbers by the each taxon shows distribution of cellulase genes and their gylcosil-hydrorase families. The schematic evolutionally tree were drawn based on Tree of Life Project (http://tolweb.org/tree/).

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ, 1990, Basic local alignment search tool. *J Mol Biol,* 215: 403-410.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ, 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res,* 25: 3389-3402.

Anzai H, Nisizawa K, and Matsuda K, 1984, Purification and characterization of a cellulase from *Dolabella auricularia. J Biochem (Tokyo),* 96: 1381-1390.

Anzai H, Uchida N, Nishide E, Nisizawa K, and Matsuda K, 1988, Isolation of a complex of endo-cellulases from the gastric teeth of *Dolabella auricularia* by affinity chromatography. *Agrc Biol Chem,* 52: 633-640.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, and Kevin S, 1989, Current protocols in molecular biology.

Azuma J, and Nishimoto K, 1984, Studies on digestive system of termites. II. Properties of carbohydrolases of termite *Coptotermes formosanus* Shiraki. *Wood Res,* 70: 1-16.

Bandi C, Sironi M, Damiani G, Magrassi L, Nalepa CA, Laudani U, and Sacchi L, 1995, The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. *Proc R Soc Lond B Biol Sci,* 259: 293-299.

Bayer EA, Chanzy H, Lamed R, and Shoham Y, 1998, Cellulose, cellulases and cellulosomes. *Curr Opin Struct Biol,* 8: 548-557.

Beguin P, and Aubert J-P, 1994, The biological degradation of cellulose. *FEMS Microbiol Rev,* 13: 25-58.

Beguin P, and Lemaire M, 1996, The cellulosome: an exocellular, multiprotein complex specialized in cellulose degradation. *Crit Rev Biochem Mol Biol,* 31: 201-236.

Bennett AB, and Christoffersen RE, 1986, Synthesis and processing of cellulase from ripening avocado fruit. *Plant Physiol,* 81: 830-835.

Bera-Maillet C, Arthaud L, Abad P, and Rosso MN, 2000, Biochemical characterization of MI-ENG1, a family 5 endoglucanase secreted by the root-knot nematode Meloidogyne incognita. *Eur J Biochem,* 267: 3255-3263.

Bird DM, Opperman CH, Jones SJ, and Baillie DL, 1999, The *Caenorhabditis elegans* genome: A Guide in The Post Genomics Age. *Annu Rev Phytopathol,* 37: 247-265.

Blin N, and Stafford DW, 1976, A general method for isolation of high molecluar weight DNA from eukaryotes. *Nucleic Acids Res,* 3: 2303-2308.

Brauman A, Kane MD, Labat M, and Breznak JA, 1992, Genesis of acetate and methane by gut bacteria of nutritionally diverse termites. *Science,* 257: 1384-1387.

Breznak JA, 1982, Intestinal microbiota of termites and other xylophagous insects. *Annu Rev Microbiol,* 36: 323-343.

Breznak JA, and Brune A, 1994, Role of microorganisms in the digestion of lignocellulose by termites. *Annu Rev Entomol,* 39: 453-487.

Brummell DA, Lashbrook CC, and Bennett AB, 1994, Plant endo-1,4-beta-D-glucanases. *ACS symposium series (USA),* 566: 100-129.

Brusca RC, and Brusca GJ, 1990, Invertebrates, Sinauer Associates, Inc. Publishers, Sunderland, MA.

Byrne KA, Lehnert SA, Johnson SE, and Moore SS, 1999, Isolation of a cDNA encoding a putative cellulase in the red claw crayfish *Cherax quadricarinatus. Gene,* 239: 317-324.

Campillo E, and Bennett AB, 1996, Pedicel breakstrength and cellulase gene expression during tomato flower abscission. *Plant Physiol,* 111: 813-820.

Chapman RF, 1998, The Insects 4th ed., Cambridge University Press, Cambridge.

Chararas C, Eberhard R, Courtois JE, and Petek F, 1983, Purification of three cellulases from the xylophagous larvae of *Ergates faber* (Coleoptera: Cerambycidae). *Insect Biochem,* 13: 213-218.

Cleveland LR, 1924, The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to *Reticulitermes flavipes* Kollar. *Biol Bull mar biol Lab,* 46: 117-227.

Cleveland LR, 1934, The wood-feeding roach *Cryptocercus*, its protozoa, and symbiosis between protozoa and roach. *Mem Am Acad Arts Sci,* 17: 185-342.

Cornwell PB, 1968, The Cockroach Vol. 1, Hutchinson & Co. Ltd., London.

Damonte A, Lombard A, Tourn ML, and Cassone MC, 1971, A modified solvent system and multiple detection technique for the separation and identification of mono and oligosaccharides on cellulose thin layers. *J Chromatogr,* 60: 203-211.

Davies GJ, Dodson GG, Hubbard RE, Tolley SP, Dauter Z, Wilson KS, Hjort C, Mikkelsen JM, Rasmussen G, and Schulein M, 1993, Structure and function of endoglucanase V. *Nature,* 365: 362-364.

de Boer JM, Smant G, Goverse A, Davis EL, Overmars HA, Pomp H, van Gent-Pelzer M, Zilverentant JF, Stokkermans JP, Hussey RS, Gommers FJ, Bakker J, and Schots A, 1996, Secretory granule proteins from the subventral esophageal glands of the potato cyst nematode identified by monoclonal antibodies to a protein fraction from second-stage juveniles. *Mol Plant Microbe Interact,* 9: 39-46.

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, et al., 2002, The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. *Science,* 298: 2157-2167.

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

Divne C, Stahlberg J, Teeri TT, and Jones TA, 1998, High-resolution crystal structures reveal how a cellulose chain is bound in the 50 A long tunnel of cellobiohydrolase I from *Trichoderma reesei. J Mol Biol,* 275: 309-325.

Ducros V, Czjzek M, Belaich A, Gaudin C, Fierobe HP, Belaich JP, Davies GJ, and Haser R, 1995, Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. *Structure,* 3: 939-949.

Elyakova LA, Sova VV, and Vaskovsky VE, 1968, Cellulase of marine mollusc *Littorina* sp. *Biochim*

Biophys Acta, 167: 462-464.

Emert GH, Gum EKJ, Lang JA, Liu TH, and Brown RDJ, 1974, Cellulases. in: *Food related enzymes* (Whitaker JR, ed.), American Chemical Society, Washington, D. C., pp. 79-100.

Foley B, 2001, Heterotrophic nutrition in animals. in: *Biology 2nd ed* (Knowx B, Ladiges P, Evans B, and Saint R, eds.), McGraw-Hill, Australia, Roseille, Australia, pp. 486-515.

Fukatsu T, 1999, Acetone preservation: a practical technique for molecular analysis. *Mol Ecol,* 8: 1935-1945.

Gaboriaud C, Bissery V, Benchetrit T, and Mornon JP, 1987, Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett,* 224: 149-155.

Girard C, and Jouanin L, 1999, Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae. Insect Biochem Mol Biol,* 29: 1129-1142.

Grassi GB, 1917, Flagelati viventi nei termiti. *Att Nazil Lincei Mem,* 12: 331-394.

Henrissat B, 1991, A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J,* 280 (Pt 2): 309-316.

Henrissat B, and Bairoch A, 1993, New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J,* 293 (Pt 3): 781-788.

Henrissat B, and Bairoch A, 1996, Updating the sequence-based classification of glycosyl hydrolases. *Biochem J,* 316 (Pt 2): 695-696.

Henrissat B, Claeyssens M, Tomme P, Lemesle L, and Mornon J-P, 1989, Cellulase families revealed by hydrophobic cluster analysis. *Gene,* 81: 83-95.

Henrissat B, and Davies G, 1997, Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol,* 7: 637-644.

Hirano H, and Watanabe T, 1990, Microsequencing of proteins electrotransferred onto immobilizing matrices from polyacrylamide gel electrophoresis: application to an insoluble protein. *Electrophoresis,* 11: 573-580.

Hogan ME, Schulz MW, Slaytor M, Czolij RT, and O'Brien RW, 1988a, Components of termite and protozoal cellulases from the lower termite, *Coptotermes lacteus* Frogatt. *Insect Biochem,* 18: 45-51.

Hogan ME, Veivers PC, Slaytor M, and Czolij RT, 1988b, The site of cellulose breakdown in higher termites (*Nasutitermes walkeri* and *Nasutitermes exitiosus*). *J Insect Physiol,* 34: 891-899.

Inoue T, Murashima K, Azuma J, Sugimoto A, and Slaytor M, 1997, Cellulose and xylan utilization in the lower termite *Reticulitermes speratus. J Insect Physiol,* 43: 235-242.

Itakura S, Tanaka H, Enoki A, Chappell DJ, and Slaytor M, 2003, Pyruvate and acetate metabolism in termite mitochondria. *J Insect Physiol,* 49: 917-926.

Itakura S, Ueshima K, Tanaka H, and Enoki A, 1995, Degradation of wood components by subterranean termite, *Coptotermes formosanus* Shiraki. *Mokuzai Gakkaishi,* 41: 580-586.

Jue CK, and Lipke PN, 1985, Determination of reducing sugars in the nanomole range with tetrazolium blue. *J Biochem Biophys Methods,* 11: 109-115.

Kanai K, and Azuma J, 1982, Studies on digestive system of termites. I. Digestion of carbohydrates by termite *Coptotermes formosanus* Shiraki. *Wood Res,* 68: 47-57.

Kim DW, Jeong YK, Jang YH, and Lee JK, 1994, Purification and characterization of endoglucanase and exoglucanase components from *Trichoderma viride*. *J Ferment Bioeng,* 77: 363-369.

Kitade O, and Matsumoto T, 1998a, Characteristics of the symbiotic flagellate composition within the termite family rhinotermitidae (Isoptera). *Symbiosis, Philadelphia, Pa (USA),* 25(1/3): 271-278.

Kitade O, and Matsumoto T, 1998b, Characteristics of the symbiotic flagellate composition within the termite family rhinotermitidae (Isoptera). *Symbiosis, Philadelphia, Pa (USA),* 25: 271-278.

Klesov AA, 1991, Biochemistry and enzymology of cellulose hydrolysis. *Biokhimiya,* 55: 1295-1318.

Kleywegt GJ, Zou JY, Divne C, Davies GJ, Sinning I, Stahlberg J, Reinikainen T, Srisodsuk M, Teeri TT, and Jones TA, 1997, The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 A resolution, and a comparison with related enzymes. *J Mol Biol,* 272: 383-397.

Koizumi M, 1921, Studies on the intestinal protozoa found in the termites of Japan. *Parasitology,* 13: 235-310.

Krishna K, 1970, Taxonomy, phylogeny, and distribution of termites. in: *Biology of termites* (Krishna K, and Weesner MF, eds.), Academic Press, New York, pp. 127-150.

Kudo T, Moriya S, Saita K, Inoue T, and Ohkuma M, 2003, Cellulose degradation and protozoa-prokaryote interaction in lower termite. in: *Mie Bioforum 2003, Biotechnology of lignocellulose degradation and biomass utilization, program and abstracts, November, 10-14, 2003, YAMAHA Resort Nemunosato, Ise-Shima, Japan,* pp. 51.

Kundu RK, Dube S, and Dube DK, 1988, Extracellular cellulolytic enzyme system of *Aspergillus japonicus*: 3. Isolation, purification and characterization of multiple forms of endoglucanase. *Enzyme microb Technol,* 10: 100-109.

Laemmli UK, 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227: 680-685.

Lashbrook CC, Gonzales-Bosch C, and Bennett AB, 1994, Two divergent endo-beta-1,4-glucanase genes exhibit overlapping expression in ripening fruit and abscising flowers. *The Plant Cell,* 6: 1485-1493.

Layne E, 1957, Spectrophotometric and turbidimetric methods for measuring proteins. *Method Enzymol,* 4: 447-454.

Leadbetter JR, and Breznak JA, 1996, Physiological ecology of Methanobrevibacter cuticularis sp. nov. and Methanobrevibacter curvatus sp. nov., isolated from the hindgut of the termite Reticulitermes flavipes. *Appl Environ Microbiol,* 62: 3620-3631.

Leadbetter JR, Crosby LD, and Breznak JA, 1998, Methanobrevibacter filiformis sp. nov., A filamentous methanogen from termite hindguts. *Arch Microbiol,* 169: 287-292.

Lee MJ, Schreurs PJ, Messer AC, and Zindr SH, 1987, Association of methanogenic bacteria with flagellated protozoa from a termite hindgut. *Curr Biol,* 15: 337-341.

Li L, Frohlich J, Pfeiffer P, and Konig H, 2003, Termite gut symbiotic archaezoa are becoming living metabolic fossils. *Eukaryot Cell,* 2: 1091-1098.

Liu X, Gao P, and Cheng Y, 1992, Purification and characterization of 1,4-beta-D-glucan glucanohydrolase from *Paecilomyces* sp S 28. *Biotechnol Appl Biochem,* 16: 134-143.

Lo N, Tokuda G, Watanabe H, Rose H, Slaytor M, Maekawa K, Bandi C, and Noda H, 2000, Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. *Curr Biol,* 10: 801-804.

Lo N, Watanabe H, and Sugimura M, 2003, Evidence for the presence of a cellulase gene in the last common ancestor of bilaterian animals. *Proc R Soc Lond B Biol Sci,* **270 Suppl 1**: S69-72.

Maeda I, Shimohigashi Y, Kihara H, and Ohno M, 1996, Purification and characterization of a cellulase from the giant snail *Achatina fulica. Biosci Biotech Biochem,* 60: 122-124.

Marialigeti K, 1979, On the community-structure of the gut-microbiota of *Eisenia lucens* (Annelida, Oligochaeta). *Pedobiologia,* 19: 213-220.

Marshall JJ, 1973, Purification of a -1,4-glucan hydrolase (cellulase) from the snail, *Helix pomatia. Comp Biochem Physiol,* 44B: 981-988.

Marshall JJ, and Grand RJA, 1976, Characterization of a beta-1,4-gllucan hydrolase from the snail, *Helix pomatia. Comp Biochem Physiol,* 53B: 231-237.

Martin MM, 1987, Invertebrate-Microbial Interactions, Ingested Fungal Enzymes in Arthropod Biology., Comstock, Ithaca, NY.

Martin MM, 1991, The evolution of cellulose digestion in insects. *Phil Trans R Soc Lond,* B, 333: 281-288.

Meinkoth J, and Wahl G, 1984, Hybridization of nucleic acids immobilized on solid supports. *Anal Biochem,* 138: 267-284.

Milligan SB, and Gasser CS, 1995, Nature and regulation of pistil-expressed genes in tomato. *Plant Molec Biol,* 28: 691-711.

Mishra PC, and Dash MC, 1980, Digestive enzymes of some earthworms. *Experientia,* 36: 1156-1157.

Mori Y, 1992, Purification and characterization of an endoglucanase from the cellulosomes (multi component cellulase complexes) of *Clostridium thermocellum. Biosci Biotech Biochem,* 56: 1198-1203.

Moriya S, Saita K, Ohkuma M, and Kudo T, 2002, Existence of diverse cellulase genes in the ecosystem of symbiotic protists in the termite gut. in: *Japan Science and Technology Corporation International Symposium, Bio-recycle Research on Termites and Their Symbiotic Microorganisms, 26th-27th February 2002 Suzuki Umetaro Hall, Bioscience Building, RIKEN, Saitama, Japan,* Kasetsart University and RIKEN, pp. 6P4.

Moriya S, Todaka N, Saita K, Ohkuma M, and Kubo T, 2003, Phylogeny of cellulase system of symbiotic protists of lower termite. in: *Mie Bioforum 2003, Biotechnology of lignocellulose degradation and biomass utilization, program and abstracts, November, 10-14, 2003, YAMAHA Resort Nemunosato, Ise-Shima, Japan,* pp. 95.

Muller M, 1993, The hydrogenosome. *J Gen Microbiol,* 139 (Pt 12): 2879-2889.

Nakashima K, and Azuma J, 2000, Distribution and properties of endo-beta-1,4-glucanase from a lower termite, *Coptotermes formosanus* (Shiraki). *Biosci Biotechnol Biochem,* 64: 1500-1506.

Nakashima K, Watanabe H, Saitoh H, Tokuda G, and Azuma JI, 2002a, Dual cellulose-digesting system of the wood-feeding termite, *Coptotermes formosanus* Shiraki. *Insect Biochem Mol Biol,* 32: 777-784.

Nakashima K, Watanabe H, and Azuma JI, 2002b, Cellulase genes from the parabasalian symbiont

Pseudotrichonympha grassii in the hindgut of the wood-feeding termite *Coptotermes formosanus*. *Cell Mol Life Sci,* 59: 1554-1560.

Nielsen CO, 1962, Carbohydrases in soil and litter invertebrates. *Oikos,* 13: 200-215.

Nishizawa S, 1974, Cellulase, in: *Kagakuno ryo-iki sensho*, Kounando, Tokyo.

O'Brien GW, Veivers PC, McEwen SE, Slaytor M, and O'Brien RW, 1979, The origin and distribution of cellulase in the termites, *Nasutitermes exitiosus* and *Coptotermes lacteus*. *Insect Biochem,* 9: 619-625.

O'Brien RW, and Slaytor M, 1982, Role of microorganisms in the metabolism of termites. *Aust J Biol Sci,* 35: 239-262.

Odelson DA, and Breznak JA, 1983, Volatile fatty acid production by the hindgut microbiota of xylophagous termites. *Appl Environ Microbiol,* 35: 239-262.

Odelson DA, and Breznak JA, 1985, Cellulase and other polymer-hydrolyzing activities of *Trichomitopsis termopsidis*, a symbiotic protozoan from termites. *Appl Environ Microbiol,* 49: 622-626.

Ohkuma M, Noda S, and Kudo T, 1999, Phylogenetic diversity of nitrogen fixation genes in the symbiotic microbial community in the gut of diverse termites. *Appl Environ Microbiol,* 65: 4926-4934.

Ohkuma M, Ohtoko K, Iida T, Tokura M, Moriya S, Usami R, Horikoshi K, and Kudo T, 2000, Phylogenetic identification of hypermastigotes, *Pseudotrichonympha*, *Spirotrichonympha*, *Holomastigotoides*, and parabasalian symbionts in the hindgut of termites. *J Eukaryot Microbiol,* 47: 249-259.

Ohkuma M, Ohtoko K, Takagi A, Moriya S, Inoue T, and Kubo T, 2002, Diversegenes for GHF45 and GHF7 cellulases in the gut symbiotic protists of termite. in: *Japan Science and Technology Corporation International Symposium, Bio-recycle Research on Termites and Their Symbiotic Microorganisms, 26th-27th February 2002 Suzuki Umetaro Hall, Bioscience Building, RIKEN, Saitama, Japan*, pp. 6P2.

Ohtoko K, Ohkuma M, Moriya S, Inoue T, Usami R, and Kudo T, 2000, Diverse genes of cellulase homologues of glycosyl hydrolase family 45 from the symbiotic protists in the hindgut of the termite *Reticulitermes speratus*. *Extremophiles,* 4: 343-349.

Okada G, 1975, Enzymatic studies on a cellulase system of *Trichoderma viride*. II. Purification and properties of two cellulases. *J Biochem,* 77: 33-42.

Okada G, 1976, Enzymatic studies on a cellulase system of *Trichoderma viride*. IV. Purification and properties of a less-random type cellulase. *J Biochem,* 80: 913-922.

Okada G, 1985, Purification and properties of a cellulase from *Aspergillus niger*. *Agric Biol Chem,* 49: 1257-1265.

Parle JN, 1963, Micro-organisms in the intestines of earthworms. *J Gen Microbiol,* 31: 1-11.

Prins RA, and Kreulen DA, 1991, Comparative aspects of plant cell wall digestion in insects. *Anim Feed Sci Technol,* 32: 101-118.

Radford A, Stone PJ, and Taleb F, 1996, Cellulase and amylase complexes. *Mycota,* 3: 269-294.

Reese E, Siu R, and Levinson H, 1950, The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J Bacteriol,* 59: 485.

Rosso MN, Favery B, Piotte C, Arthaud L, De Boer JM, Hussey RS, Bakker J, Baum TJ, and Abad P,

1999, Isolation of a cDNA encoding a beta-1,4-endoglucanase in the root-knot nematode *Meloidogyne incognita* and expression analysis during plant parasitism. *Mol Plant Microbe Interact,* 12: 585-591.

Rouland C, Civas A, Renoux J, and Petek F, 1988, Purification and properties of cellulases from the termite *Macrotermes mulleri* (Termitidae, Macrotermitinae) and its symbiotic fungus *Termitomyces* sp. *Comp Biochem Physiol,* 91B: 449-458.

Rouvinen J, Bergfors T, Teeri T, Knowles JK, and Jones TA, 1990, Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei. Science,* 249: 380-386.

Sakon J, Irwin D, Wilson DB, and Karplus PA, 1997, Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca. Nat Struct Biol,* 4: 810-818.

Schulz MW, Slaytor M, Hogan M, and O'Brien RW, 1986, Components of cellulase from the higher termite, *Nasutitermes walkeri. Insect Biochem,* 16: 929-932.

Scrivener AM, and Slaytor M, 1994a, Cellulose digetion in *Panesthia cribrata* Saussure: does fungal cellulase play a role? *Comp Biochem Physiol,* 107B: 309-315.

Scrivener AM, and Slaytor M, 1994b, Properties of the endogenous cellulase from *Panesthia cribrata* Saussure and purification of major endo-beta-1,4-glucanase components. *Insect Biochem Molec Biol,* 24: 223-231.

Scrivener AM, Slaytor M, and Rose HA, 1989, Symbiont-independent digestion of cellulose and starch in *Panesthia cribrata* Saussure, an Australian wood-eating cockroach. *J Insect Physiol,* 35: 935-941.

Scrivener AM, Zhao L, and Slaytor M, 1997, Biochemical and immunological relationships between endo-beta-1,4-glucanases from cockroaches. *Comp Biochem Physiol,* B,118: 837-843.

Slaytor M, 1992, Cellulose digestion in termites and cockroaches: what role do symbionts play? *Comp Biochem Physiol,* 103B: 775-784.

Slaytor M, 2000, Energy metabolism in the termite and its gut microbiota. in: *Termites: Evolution, Society, Symbioses, Ecology* (Higashi M, ed.), Kluwer Academic Publishers, Dordrecht, pp. 307-332.

Smant G, Goverse A, Stokkermans JPWG, and de Boer JM, 1997, Potato root diffusate-induced secretion of soluble, basic proteins originating from the subventral esophageal glands of potato cyst nematodes. *Phytopathology,* 87: 839-845.

Smant G, Stokkermans JP, Yan Y, de Boer JM, Baum TJ, Wang X, Hussey RS, Gommers FJ, Henrissat B, Davis EL, Helder J, Schots A, and Bakker J, 1998, Endogenous cellulases in animals: isolation of beta-1, 4-endoglucanase genes from two species of plant-parasitic cyst nematodes. *Proc Natl Acad Sci U S A,* 95: 4906-4911.

Snodgrass RE, 1993, Principles of insect morphology, Cornell Univ. Press, Ithaca, pp. 667.

Strasdine GA, and Whitaker DR, 1963, On the origin of the cellulase and chitinase of *Helix pomatia. Can J Biochem Physiol,* 41: 1621-1626.

Strimmer K, and Von Haeseler A, 1996, Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. *Mol Biol Evol,* 13: 964-969.

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*. *Eur J Biochem,* 270: 3455-3460.

Suzuki K, Ojima T, and Nishita K, 2003, Purification and cDNA cloning of a cellulase from abalone Haliotis discus hannai. *Eur J Biochem,* 270: 771-778.

Syvanen M, 1994, Horizontal gene transfer: evidence and possible consequences. *Annu Rev Genet,* 28: 237-261.

Teeri TT, 1997, Crystalline cellulose degradation: New insight into the function of cellobiohydrolases. *Trends Biotech,* 15: 160-167.

Thayer DW, 1976, Facultative wood-digesting bacteria from the hind-gut of the termite *Reticulitermes hesperus*. *J Gen Microbiol,* 95: 287-296.

Tokuda G, Lo N, Watanabe H, Slaytor M, Matsumoto T, and Noda H, 1999, Metazoan cellulase genes from termites: intron/exon structures and sites of expression. *Biochim Biophys Acta,* 1447: 146-159.

Tokuda G, Saito H, and Watanabe H, 2002, A digestive β-glucosidase from the salivary glands of the termite, *Neotermes koshunensis* (Shiraki): distribution, characterization and isolation of its precursor cDNA by 5'- and 3'-RACE amplifications with degenerate primers. *Insect Biochem Molec Biol,* **(in press)**.

Tokuda G, Watanabe H, Matsumoto T, and Noda H, 1997, Cellulose digestion in the wood-eating higher termite, *Nasutitermes takasagoensis* (Shiraki): distribution of cellulases and properties of endo-beta-1,4-glucanase. *Zoolog Sci,* 14: 83-93.

Tomme P, Kwan E, Gilkes NR, Kilburn DG, and Warren RA, 1996, Characterization of CenC, an enzyme from Cellulomonas fimi with both endo- and exoglucanase activities. *J Bacteriol,* 178: 4216-4223.

Tomme P, Warren RAJ, and Gilkes NR, 1995, Cellulose hydrolysis by bacteria and fungi. *Adv Microb Physiol,* 37: 1-81.

Tracey MV, 1951, Cellulase and chitinase of earthworms. *Nature,* 167: 776-777.

Trager W, 1934, The cultivation of a cellulose-digesting flagellate *Trichomonas termopsidis* and of certain other termite protozoa. *Biol Bull,* 66: 182-190.

Trainotti L, Ferrarese L, Vecchia FD, Rascio N, and Casadoro G, 1999a, Two different endo-beta -1,4-glucanases contribute to the softening of the strawberry fruits. *J Plant Physiol,* 154: 355-362.

Trainotti L, Spolaore S, Pavanello A, Baldan B, and Casadoro G, 1999b, A novel E-type endo-beta -1,4-glucanase with a putative cellulose-binding domain is highly expressed in ripening strawberry fruits. *Plant Mol Biol,* 40: 323-332.

Tucker ML, Durbin ML, Clegg MT, and Lewis LN, 1986, Avocado cellulase: nucleotide sequence of a putative full-length cDNA clone and evidence for a small gene family. *Plant Mol Biol,* 9: 197-203.

Tucker ML, and Milligan SB, 1991, Sequence analysis and comparison of avocado fruit and bean absiccion cellulases. *Plant Physiol,* 95: 928-933.

Urbasek F, 1990, Cellulase activity in the gut of some earthworms. *Rev Ecol Biol Sol,* 27: 21-28.

Urbasek F, and Pizl V, 1991, Activity of digestive enzymes in the gut of five earthworm species (Oligochaeta; Lumbricidae). *Rev Ecol Biol Sol,* 28: 461-468.

Veivers PC, Musca AM, O'Brien RW, and Slaytor M, 1982, Digestive enzymes of the salivary glands and gut of *Mastotermes darwiniensis*. *Insect Biochem,* 12: 35-40.

Vinceslas-Akpa M, and Loquet M, 1996, Activity and origin of cellulasic enzymes in gut of the earthworm *Eisenia fetida* andrei. *C R Acad Sci III,* 319: 1113-1117.

Watanabe H, Nakamura M, Tokuda G, Yamaoka I, Scrivener AM, and Noda H, 1997, Site of secretion and properties of endogenous endo-beta-1,4-glucanase components from *Reticulitermes speratus* (Kolbe), a Japanese subterranean termite. *Insect Biochem Mol Biol,* 27: 305-313.

Watanabe H, Nakashima K, Saito H, and Slaytor M, 2002, New endo-beta-1,4-glucanases from the parabasalian symbionts, *Pseudotrichonympha grassii* and *Holomastigotoides mirabile* of *Coptotermes* termites. *Cell Mol Life Sci,* 59: 1983-1992.

Watanabe H, Noda H, Tokuda G, and Lo N, 1998, A cellulase gene of termite origin. *Nature,* 394: 330-331.

Watanabe H, and Tokuda G, 2001, Animal cellulases. *Cell Mol Life Sci,* 58: 1167-1178.

Xu B, Hellman U, Ersson B, and Janson JC, 2000, Purification, characterization and amino-acid sequence analysis of a thermostable, low molecular mass endo-beta-1,4-glucanase from blue mussel, *Mytilus edulis*. *Eur J Biochem,* 267: 4970-4977.

Xu B, Janson JC, and Sellos D, 2001, Cloning and sequencing of a molluscan endo-beta-1,4-glucanase gene from the blue mussel, Mytilus edulis. *Eur J Biochem,* 268: 3718-3727.

Xue XM, Anderson AJ, Richardson NA, Anderson AJ, Xue GP, and Mather PB, 1999, Characterisation of cellulase activity in the digestive system of the redclaw crayfish (*Cherax quadricarinatus*). *Aquaculture,* 180: 373-386.

Yamaoka I, 1979, Selective ingestion of food by the termite protozoa, Trichonympha agilis. *Zool Mag,* 88: 174-179.

Yamaoka I, and Nagatani Y, 1975, Cellulose digestion system in the termite, *Reticulitermes speratus* (Kolbe). I. producing sites and physiological significance of two kinds of cellulase in the worker. *Zool Mag,* 84: 23-29.

Yamin MA, 1979, Flagellates of the orders Trichomonadida Kirby, *Oxymonadida Grasse* and *Hypermastigida Grassi* & *Foa* reported from lower termites (Isoptera families Mastotermitidae, Kalotermitidae, Hodotermitidae, Termopsidae, Rhinotermitidae, and Serritermitidae) and from the wood-feeding roach *Cryptocercus* (Dictyoptera: Cryptocercidae). *Sociobiol,* 4: 3-119.

Yamin MA, 1980, Cellulose metabolism by the flagellate *Trichonympha* from a termite is independent of endosymbiotic bacteria. *Science,* 211: 58-59.

Yan Y, Smant G, Stokkermans J, Qin L, Helder J, Baum T, Schots A, and Davis E, 1998, Genomic organization of four beta-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. *Gene,* 220: 61-70.

Yokoe Y, 1964, Cellulase activity in the termite, *Leucotermes speratus*, with new evidence in support of a cellulase produced by the termite itself. *Scient Pap Coll Gen Educ Tokyo,* 14: 115-120.

Yoshigi N, Taniguchi H, and Sasaki T, 1988, Purification and properties of a new endo-cellulase from *Robillarda* sp. Y 20. *Agric Biol Chem,* 52: 1389-1396.

Yoshimura T, Tsunoda K, and Takahashi M, 1996, Degradation of wood in the digestive tract of a higher

termite, *Nasutitermes takasagoensis* (Shiraki) (Isoptera: Termitidae). *Mokuzai Gakkaishi,* **42**: 1250-1257.

Yung M-H, Schaffer R, and Putterill J, 1999, Identification of genes expressed during early Arabidopsis carpel development by mRNA differential display: characterization of ATCEL2, a novel endo-1,4--D-glucanase gene. *Plant J,* **17**: 203-208.

Zhang J, Scrivener AM, Slaytor M, and Rose HA, 1993, Diet and carbohydrase activities in three cockroaches, *Calolampra elegans* Roth and Princis, *Geoscapheus dilatatus* Saussure and *Panesthia cribrata* Saussure. *Comp Biochem Physiol,* **104A**: 155-161.