CMLS, Cell. Mol. Life Sci. 58 (2001) 1167–1178
1420-682X/01/091167-12 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Review

Animal cellulases

H. Watanabe [a, *] and G. Tokuda [b]

[a] National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8634 (Japan), Fax + 81 298 38 6028, e-mail: hinabe@affrc.go.jp
[b] Department of Chemistry, Biology and Marine Science, Faculty of Science, University of the Ryukyus, 1 Senbaru, Nishihara-cho, Okinawa 903-0213 (Japan)

Received 19 December 2000; received after revision 5 March 2001; accepted 6 March 2001

Abstract. Previous dogma has maintained that cellulose, ingested by xylophagous or herbivorous animals, is digested by cellulolytic symbiotes. The first evidence in conflict with this contention involved the demonstration of cellulolytic activities in symbiote-free secreting organs (e.g., the salivary glands of termites) or defaunated guts. Following these demonstrations, possible endogenous cellulase components were purified from several cellulose-digesting invertebrates, but this research did little to change the general view concerning animal cellulose digestion. Thanks to recent developments in molecular biology, the existence of cellulases of animal origin has been firmly established. To date, cellulase genes have been reported from arthropods (insects and a crayfish) and nematodes. This paper describes and discusses the presence and nature of endogenous cellulases in higher animals.

Key words. Endo-β-1,4-glucanase; glycoside hydrolase; invertebrate; gene; purification; symbiosis; endogenous.

Introduction

Cellulose is the most abundant biomass on the earth [1]. It is a polymerized form of glucose molecules with β-1,4-linkages, consisting of composite forms of highly crystallized microfibrils among amorphous matrixes, thus refusing access to hydrolyzing enzymes [1]. Cellulolytic fungi and bacteria have developed complex forms of cellulase systems which actively hydrolyze cellulose fibrils [1]. In plants, cellulases hydrolyze their cell walls at various developmental stages (e.g., bean abscission, fruit ripening and abscission, and pedicel abscission) [2–11]. However, higher animals have not been generally recognized to produce endogenous cellulases [12].

The presence of symbiotic protozoa or bacteria has been used to explain cellulose digestion in invertebrates and herbivorous cattle. This theory was first proposed by Cleveland, in his work with termites, who demonstrated the disappearance of the viability of *Reticulitermes flavipes* after defaunation of protozoa from the hindgut [13], thus concluding that a hindgut protozoan fauna was responsible for the cellulose digestion of the hosts. Since his demonstration, symbiotic cellulase production became the most favored explanation for cellulose digestion in higher animals [14, 15].

Following Cleveland's experiments, much research was devoted to demonstrating symbiotic cellulase production by bacteria and fungi in addition to protozoa, which are harbored in invertebrate digestive systems. However, few reports were presented demonstrating symbiotic relationships between host invertebrates and intestinal microorganisms [16]. Termites belonging to the family Termitidae, which are called 'higher termites', generally do not contain protozoan fauna, but many of them ingest cellulosic material [17]. The cellulose-digesting ability of higher termites was interpreted as due to the acquisition of cellulolytic bacteria in the hindgut [15], but such enrollment of intestinal bacteria has not been confirmed

* Corresponding author.

[16, 18]. The intestinal bacteria were proposed as to produce acetate and fix nitrogen to digest cellulose consumed by the host [16, 18].

Later, the focus of research on the digestion of xylophagous insects shifted from inside to outside the digestive tract. Possible acquisitions of cellulase components by two species of fungus-growing termites (subfamily Macrotermitinae; Termitidae) from fungus combs grown in termite nests were independently reported in 1978 [19, 20]. In addition, cellulase, amylase, and xylanase activities in the gut of fungus-growing termites were also construed as acquired from symbiotic fungi in the nest [21]. In another fungus-growing termite, several cellulase components were purified to homogeneity from the gut and fungal nodules. A high similarity was demonstrated in enzymatic character between one component from the gut and another from the nodules [22]. However, whether such acquired cellulase activity in fungus-growing termites meets the physiological requirements of the hosts is controversial [18, 23, 24]. Acquisition of digestive enzymes has been explored in other xylophagous arthropods and interpreted as a general phenomenon in xylophagous arthropods including woodwasps, isopods, cerambycid beetles, and attine ants [25].

Apart from arthropods, several cellulase studies were undertaken in molluscs, including snails [26–30], a sea slug [31], a periwinkle [32], and some bivalves [33–39]. Some reported possible endogenous enzyme sources such as the hepatopancreas [27, 32], gastric teeth [31], and crystalline styles (needlelike structures made of crystalline proteins forming a motor organ in the stomach of bivalves [40]) [33, 39]. Later, a cellulolytic nitrogen-fixing bacterium was isolated from the Deshayes gland {a specialized bacteria-culturing organ found in the gills of shipworms (family Teredinidae) [40]} [34–37]. Possible bacterial cellulase production was also reported in another wood-boring bivalve species in the genus *Xylophaga* [38]. Such symbiotic cellulase production, however, was not reported as a general phenomenon among mollusca.

Yokoe and Yasumasu [33] investigated cellulase activity in 74 higher animals including several vertebrates and demonstrated that the activities of the possible endogenous cellulases in these animals were distributed according to phylogenetic relationships and not according to food habits. In fact, cellulase activities in insects are observed not only in strict xylophagous species like termites and wood-eating cockroaches, but also in many other species with different food habits which do not have symbiotic relationships with cellulolytic protozoa or microorganisms, such as domestic omnivorous cockroaches [41]. Recent advances in molecular biology will clarify whether cellulase genes in invertebrates are indeed distributed according to phylogenetic development, as supposed by Yokoe and Yasumasu [33].

In this review, we discuss the nature of cellulases from invertebrates, and describe their physiological and phylogenetic origins.

Challenges to the theory of completely symbiotic cellulose digestion

Discrepancies with theories based on symbiotic cellulose digestion arose from apparent contradictions between cellulolytic activity and the locations – or absence – of symbiotes as enzyme sources. In the snail, *Helix pomatia*, cellulase and chitinase activities were detected in sterile hepatopancreas [27]. In addition, both cellulase and chitinase activity in the digestive juices (luminal fluid of the gut) increased in proportion to body weight and the total protein content of the hepatopancreas, but not in relation to increases in digestive juice bacterial counts when, for example, the snails emerge from hibernation and start eating. The conclusion was drawn, therefore, that these enzymes were produced endogenously [27]. Similar arguments have also been put forward based on contradictions observed between the presence or absence of cellulolytic activities and gut microorganisms in earthworms [42–49]. To avoid repetition of the same arguments, our discussions here will mainly focus on termites and cockroaches for which research has reached the genetic level.

In termites, the first putative endogenous cellulase was demonstrated in the Japanese subterranean termite *Leucotermes (Reticulitermes) speratus* [50]. A drastic reduction in hindgut protozoa by heat exposure of these termites did not affect cellulase activity as measured by a decrease in the viscosity of carboxymethylcellulose (CMC) solution. Thus cellulase activity was attributed to the host [50]. Later, salivary gland cellulase activity against CMC and that in hindgut protozoan fauna against filter-paper and cotton were separated by Yamaoka and Nagatani [51], who proposed that these two cellulases synergistically break down native cellulose in an analogy to the fungal cellulase system [52]. Undoubtedly, ingested cellulose will come into contact with both the salivary and protozoan cellulases, but whether both cellulases act simultaneously in a synergistic action [52] has yet to be clarified.

A comparative study of the Australian arboreal higher termite *Nasutitermes walkeri* with the mound-building lower (= protozoa-inhabited) termite *Coptotermes lacteus* showed that the intestinal fauna and flora play no (*N. walkeri*) or a relatively small (*C. lacteus*) role in the cellulase production of these species [53]. Removal of intestinal bacteria and protozoa by tetracycline treatments (assumed to eliminate protozoa indirectly by the reduction of bacteria) did not affect cellulase activities in the guts of either species except for a 20% decrease seen in the hindgut of *C. lacteus* [53]. To explain the cellulose di-

gestion of higher termites (which make up three-quarters of known termite species [54]), the isolation of cellulolytic bacteria from the gut has been attempted, but only cellulolytic facultative heterotrophs and inherent non-cellulolytic spirochetes have been found [18, 23]. In *N. walkeri* and the mound-building higher species *N. exitiosus*, the majority of the cellulase activity was located in the epithelia and the luminal contents of the midgut [55].

Cockroaches are a group of insects including fossil species from the Paleozoic period. They are taxonomically cross-related to termites [56] and join a monophyletic clade which includes termites [57]. It is estimated that there are about 3500 cockroach species, of which less than 1% are domestic omnivorous pests. Most of the others live in the ground, or commensally in the nests of ants, wasps, or termites of tropical rainforests, while a few species bore into and inhabit wood [56]. Elucidation of cellulose digestion in such cockroaches may help to clarify its origin in termites.

In a series of experiments with the wood-eating cockroach *Cryptocercus punctulates*, Cleveland [58] concluded that this cockroach depends on hindgut protozoa for cellulose digestion, as does the lower termite *R. flavipes* [13]. On the other hand, Slaytor and his colleagues demonstrated that the distribution of cellulase in the gut of the Australian wood-eating cockroach *Panesthia cribrata* was limited to the fore- and midgut, and was not found in protozoa inhabiting the hindgut [59]. *P. cribrata* was maintained on filter paper with tetracycline (or crystalline cellulose) for over 12 weeks without reduction of cellulase activity [59]. During this experiment, the respiratory quotient value (produced CO_2/consumed O_2) stayed between 0.98 and 1.03 (0.7 for starved individuals) [59], showing that the *P. cribrata* cockroaches lived on the cellulose, not on stored energy, similar to the other Australian wood-eating cockroaches *Calolampra elegans* and *Geoscapheus dilatatus* [60]. A possible role for acquired fungal cellulases was also proposed in *P. cribrata* [15], but was not supported by quantification of the fungal cellulase in food material [61].

Purification and character of animal cellulases

The research mentioned above includes mostly quantitative analyses of cellulolytic activities in cellulose-digesting animals and thus there are 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 to identify the origins of the cellulases [22, 25].

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; randomly hydrolyzing the middle of cellulose chains), 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) [1, 62]. 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 [1, 63, 64]. For termites and cockroaches, Slaytor [23] proposed a unique system that lacks CBH components, while Martin [25] proposed the enrollment of CBH acquired from fungi in many xylophagous species. In this review, we mainly focus on the purification of endogenous components.

From the cockroach clade (cockroaches + termites [57]), 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* [65], another two (41 and 42 kDa) from the lower termite *R. speratus* [66], and one more (47 kDa) from the Japanese arboreal higher termite *N. takasagoensis* [67].

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* [68]. Aside from insects, two cellulolytic components (30 and 40 kDa; 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 [69].

Apart from arthropods, several cellulolytic components have been purified to homogeneity from molluscs. 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* [28] and *Achatina fulica* [30], respectively. A cellulase seeming to be an EG component (44 kDa) was purified from the gastric teeth of the sea slug *Dolabella auriclularia* [31]. The purification of various other EG components (51 kDa on SDS-PAGE, which separates into six different 31-kDa components on the DISC-PAGE) from this species has also been reported [70]. An EG component was also purified from the marine perivinkle *Littorina* sp. [32]. 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* [39].

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 some of the glucose residues, which prevents the enzyme from attacking from the chain ends to the

inside [71]. Many animal EG components have been identified according to this rule as EG, even though many cut out cellobiose units when they attack cellooligomers similarly to CBHs [28, 31, 65–67]. This quality of animal EG components may help them to digest 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 molluscs, which can break down food material into fine particles [29, 72]. In termites, ingested wood is ground to less than 50 μm in *Coptotermes formosanus* [73] or into two size ranges (10–30 and 100–300 mm) in *N. takasagoensis* [74]. 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 nearly a century has passed since the first cellulase research began with snails [75], 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 [76, 77].

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 [78, 79]. From amino acid sequences of uncharacterized proteins (svp39 from *G. rostochiensis* and svp49 from *H. glycines*), four different EG cDNAs (*HG-eng-1*, 1615 bp encoding a 49.8-kDa protein; *HG-eng-2,* 1191 bp encoding a 34.7 kDa protein; *GR-eng-1,* 1546 bp encoding a 49.7-kDa protein, and *GR-eng-2,* 1300 bp encoding a 42-kDa protein) were cloned from cDNA libraries of these species with the aid of degenerate and gene-specific rapid amplification of cDNA ends (RACE) amplifications [76]. 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 [76, 80]. 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 [76]. Recently, another EG-encoding cDNA, *MI-eng1* (1668 bp), which encoded a protein (506 amino acids) including a putative catalytic domain and a CBD, was isolated from the root-knot nematode *Meloidogyne incognita* [81], and its precise properties were reported [82]. The putative catalytic domain of MI-eng1 shared 52.5% identity with those of HG-eng1 and HG-eng2 [81].

From the phytophagous beetle *Phaedon cochleariae*, a putative EG cDNA was isolated from a gut cDNA library by random sequencing followed by RACE amplifications [83]. 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 [83].

From termites and cockroaches, 19 different EG cDNAs have been isolated to date [57, 77, 84]. 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 1466 bp cDNAs (*RsEG* and *RsEG2*) were identified [77, 84]. 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 [77, 84]. 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 [84]. A cDNA (1685 bp, *NtEG* cDNA) including a complete open reading frame was then obtained by additional gene-specific RACE amplifications [84]. The *NtEG* cDNA encoded 448 amino acids consisting of a signal peptide (16 amino acids) and a single catalytic domain, which showed 78.3% identity with the amino acid sequence of RsEG [77, 84]. 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 [57]. These showed 63–69% identity with amino acid sequences of EGs from the termites *R. speratus* and *N. takasagoensis* [57].

From the hepatopancreas of the Australian yabby *Cherax quadricarinatus* (a freshwater 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 [85]. The EG encoded by the cDNA from *C. quadricarinatus* consisted of 469 amino acids including a putative signal peptide of 24 amino acids [85]. From another four cockroach and three

termite species in addition to *R. sepratus*, *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. quardricarinatus* [57].

Structural character of amino acid sequences of animal EGs

As mentioned above, 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, conceals their phylogenetic relationships [86]. To compensate for this problem, Henrissat et al. [80] advocated the classification of cellulolytic enzymes by hydrophobic cluster analysis of amino acid sequences, which reflects structural similarities in the grouping [87]. They later broadened the application of their method to include all *O*-glycoside hydrolases (EC 3.2.-.-) by directly comparing amino acid sequences in addition to hydrophobic clustering, and then classified them into 78 families made up, to date, of ten clans [groups of families sharing a structural resemblance (http://afmb.cnrsmrs.fr/~pedro/CAZY/ghf.html)] [88, 89]. 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 [76, 82] or homology in amino-acid sequence [81]. GHF5 members share a general motif of an $(\alpha/\beta)_8$ barrel structure [88]. The three-dimensional structure of one member was investigated by X-ray crystallography of CelCCA from the gram-positive bacterium *Clostridium cellulolyticum* [90]. 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 [90]. 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, *Caenorhabditis elegans,* for which the entire genomic sequence is known [91], does not possess any cellulase genes [82, 92]. Since the nematode EGs showed the highest homology to bacterial GHF5 members, a possible ancient horizontal gene transfer from prokaryotes has been advocated [92].

EG coded by cDNA from the beetle *P. cochleariae* (*PcEG*) was placed in GHF45 because of its similarity in amino acid sequence [83]. 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 [93]. The nucleophile and proton donor of matured EG V are Asp10 and Asp121, respectively, and are thought to be critical for activity. However, only the nucleophile Asp residue was preserved as Asp26 in matured PcEG (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) [57, 77, 84]. 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 [94]. 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 the 18 amino acid residues located beside the cellohexaose bound by the catalytic center of E4-68 [94], 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) [94] are well preserved except for substitutions of Phe205 to Tyr in PcEG1, His376 to Ser in PcEG1 and PcEG2, and to Thr in RsEG, RsEG2, PcEG1, and PcEG2, 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 [94] are preserved in all animal GHF9 EGs listed in table 1 (fig. 3).

All known animal EGs, except for HG-eng1 and GR-eng1 from plant-parasitic nematodes, have consisted only of catalytic domains [57, 77, 83–85]. However, many fungal cellulase components active against crystalline cellulose consist not only of catalytic domains, but also CBDs and linkers of a Pro/Thr/Ser-rich residue, which hinge catalytic domains and CBDs [1, 95]. Although the large production of EGs generally seen in termites and cockroaches could perhaps compensate for the low activity of EGs on crystalline cellulose [23], an underlying mechanism for the efficient degradation of crystalline cellulose by insect cellulases that do not include CBDs has yet to be clarified.

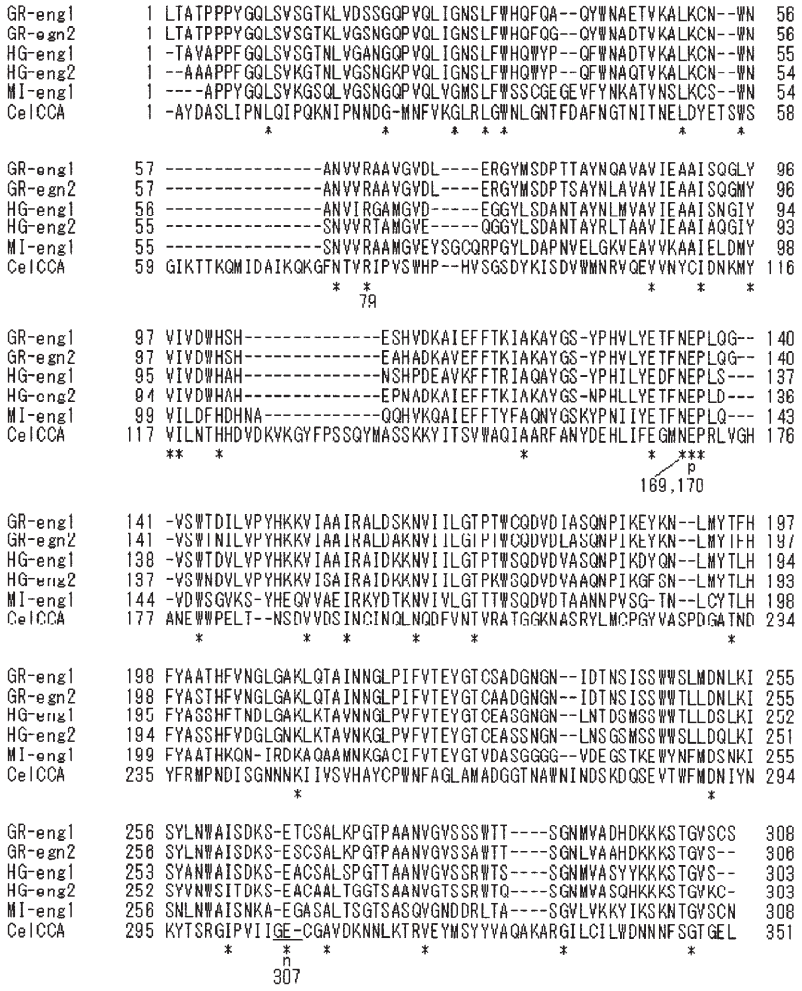

Figure 1. Multiple alignments of catalytic domains of endo-β-1,4-glucanase components in the glycoside hydrolase family 5. Alignment of endo-β-1,4-glucanase components from the bacterium *Clostridium cellulolyticum* (CelCCA [AAA23222] (numbers in brackets are the protein i. d. of GenBank/DDBJ or Swissprot) and the plant-parasitic nematodes *Globodera rastochiensis* (GR-eng1 [AAC48325.1] and GR-eng2 [AAC48341.1]), *Heterodera glycines* (HG-eng1 [AAC48327.1] and HG-eng2 [AAC48326.1]), and *Meloidogyne incognita* (MI-eng1 [AAD45868.1]). The alignments were achieved using the computer program CLUSTAL X (1.8) for Windows (National Center for Biotechnology Information, Bethesda, Md) 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-eng1 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 were realigned 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.

Animal cellulase genes

As already mentioned, 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.

Two EG genes from each of two plant-parasitic nematodes, *G. rostochiensis* and *H. glycines* (*GR-eng1*, 2492 bp and *GR-eng2*, 2388 bp; *HG-eng1*, 2151 bp and *HG-eng2*, 2324 bp, respectively) were isolated from recombinant phage genomic DNA libraries by probing the EG cDNAs [92]. 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. takasaogensis* by combining PCR amplification of genomic DNA with sets of primers designed from *NtEG* cDNA sequences. A complete 13,005-bp gene including ten exons separated by nine introns was then identified. In *R. speratus*, three introns of *RsEG* were identified interrupting exons at the corresponding

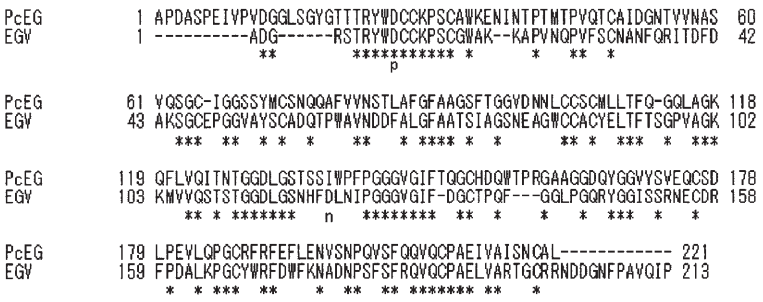

Figure 2. Multiple alignments of catalytic domains of endo-β-1,4-glucanase components in the glycoside hydrolase family 45. Alignment of endo-β-1,4-glucanase components from the fungus *Humicola insolens* (EGV [P43316]) and the beetle *Phaedon cochleariae* (PcEG [CAA76931.1]). The alignment was accomplished by the same method described in the legend to Figure 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.

positions and phases of *NtEG* [84]. 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 [84].

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, polyadenylation and cleavage signals (GATAAA) in appropriate sites, 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) [84, 92]. 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 [92], and a CCAAT promoter sequence in the upstream region of *NtEG* [84].

The structure of the EG genes from these nematodes and this termite demonstrate well 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* [91] imply that a possible ancestral EG gene among the four EG genes existed before the divergence of these plant parasitic nematodes [92]. Thus, an ancient horizontal transfer [96] is suggested as a possible explanation for the acquisition of EG genes by the cyst nematodes [92]. Highly homologous GHF9 EG precursors, which form a monophyletic clade among other GHF9 members, have been isolated from termites, cockroaches, and the crayfish *C. quadricarinatus* [57, 77, 84, 85], pointing to an inheritance and divergence of an ancestral EG gene through radiation from a common arthropod progenitor. However, the evidence is insufficient to establish these hypotheses as firm theories. Further investigations of EG genes from other animals and/or elucidation of other

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	accession																	
E4-68	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	BAA34050.1		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	AAD38027.1					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 beside Glc(+2), but show no particular hydrogen bonding [93]. Only substituted amino acids are displayed. Empty boxes indicate identical amino acids to E 4-68 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 (Bethesda, MD) with Gonnet 250 pairwise parameters.

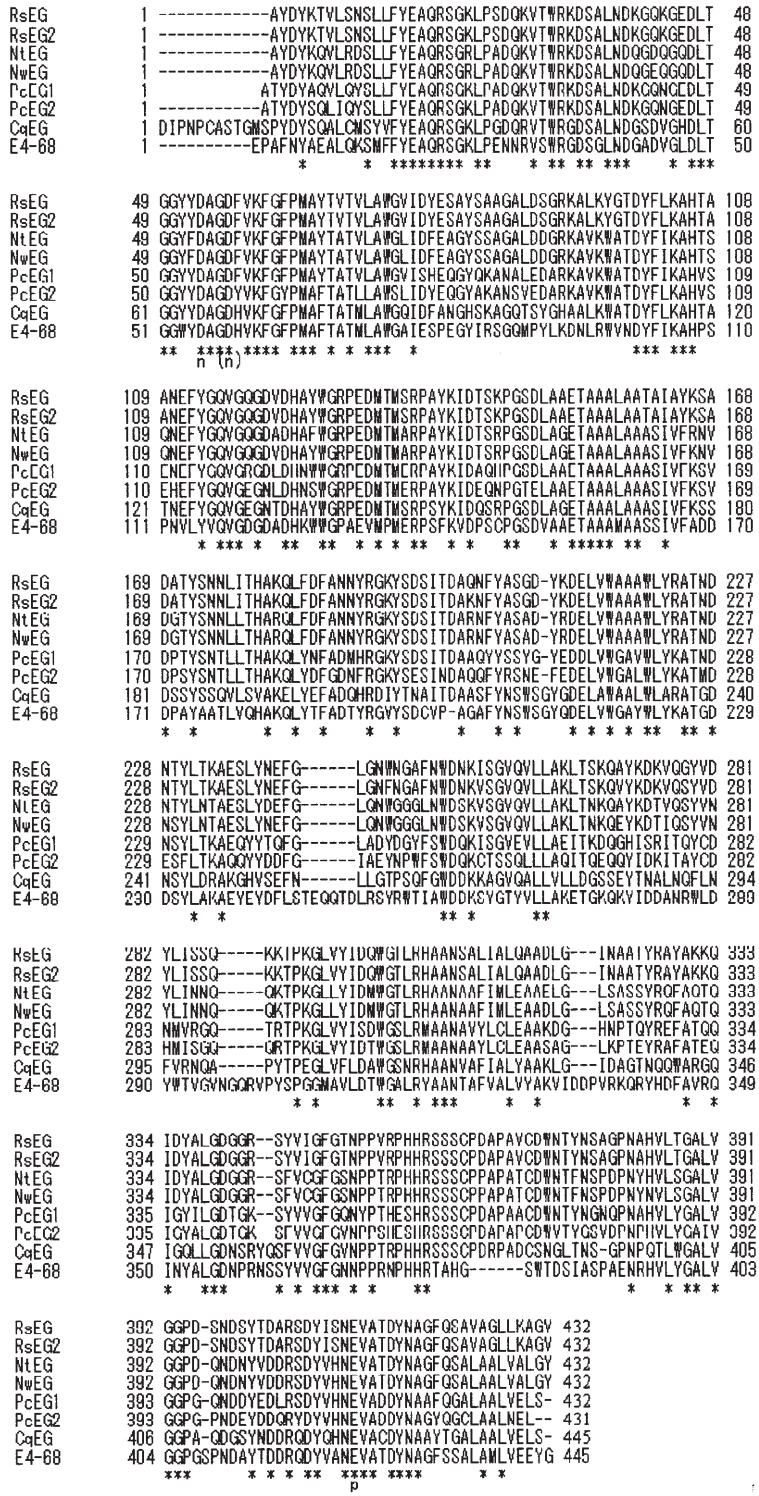

Figure 3. Multiple alignments of catalytic domains of endo-β-1,4-glucanase components in the glycoside hydrolase family 9. Alignment of endo-β-1,4-glucanase components from the bacterium *Thermomonospora fusca* (E4-68; catalytic domain of E4 [part of AAB42155.1]), the crayfish *Cherax quadricarinatus* (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 procedure described in the legend to figure 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.

genes in animals maintaining EG genes are required to compare the genomic structure and GC contents 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 identify 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 [66]. 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 in the host genome was confirmed by isolation of the gene or its fragment and/or by southern blotting, their expression has been localized primarily by RT-PCR and in situ hybridization [76, 77, 84, 85]. 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 [76]. Expression of a 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 [85]. For higher and lower termites, contrasting expression sites have been clarified. In the higher termite *N. takasagoensis*, 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 [84]. On the other hand, in the lower termite *R. speratus*, 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 [66, 77, 84]. The midgut and hepatopancreas are phylogenetically homologous organs of arthropods [40], and it is therefore understandable that the higher termite and the crayfish express similar EGs in those tissues. However, the expression of EGs in the salivary glands of lower termites (*R. speratus*, *Neotermes koshunensis*, *Hodotermopsis japonica* and *Mastotermes darwiniensis*) [57] implies a dynamic convergence of EG expression sites and the digestive mechanism in the course of termite divergence.

Conclusion

The historical debate concerning the presence of endogenous cellulase in higher animals has been settled by molecular evidence. There is no doubt that a cellulase is a protein that can be potentially produced in any life form providing there is a corresponding gene encoding it. However, this conclusion does not answer the queries, 'Why do some animals have endogenous cellulases and others not?' or 'Why isn't cellulase a common enzyme in animals like amylase or proteinase?'

We believe that further investigations into animal cellulose digestion can eventually answer these questions. To illustrate this, two recent findings of new members in GHF45 and GHF5 are noteworthy. Members of GHF45, one of which was found in the beetle *Phaedon cochleariae* [83], were recently also found in hindgut protozoa of the termite *R. speratus* [97]. These protozoa are considered among of the most ancient forms of eukaryotes [98, 99]. Members of GHF5, which had previously been isolated only in fungi, bacteria, and plant-parasitic nematodes [76, 81, 92], have also recently been isolated in the plant *Arabidopsis thaliana* (GenBank accession numbers: AL161572, AC006234, AC007357, AL391147, AB013389, AL391150, AB016878, AB024038, AB024038, AL162351, AC064879, AC011708 and AC011708 from the *Arabidopsis* genome project [100]). At this stage, these findings can only imply dynamic interchanges of ancestral genes among different organisms, but as more diverse cellulases come under molecular investigation, we will more precisely understand their origins and roles in animals.

Acknowledgements. This work was supported by the Promotion of Basic Research Activities for Innovative Biosciences fund from the Bio-oriented Technology Research Advancement Institution (BRAIN; Omiya, Saitama, 331-8537 Japan; www.brain.go.jp) and by the Pioneer Research Project Fund (No. PRPF-0022) from the MAFF (Ministry of Agriculture, Forestry and Fisheries of Japan).

1 Tomme P., Warren R. A. J. and Gilkes N. R. (1995) Cellulose hydrolysis by bacteria and fungi. Adv. Microb. Physiol. **37:** 1–81
2 Bennett A. B. and Christoffersen R. E. (1986) Synthesis and processing of cellulase from ripening avocado fruit. Plant Physiol. **81:** 830–835
3 Tucker M. L., Durbin M. L., Clegg M. T. and Lewis L. N. (1987) Avocado cellulase: nucleotide sequence of a putative full-length cDNA clone and evidence for a small gene family. Plant Mol. Biol. **9:** 197–203
4 Tucker M. L. and Milligan S. B. (1991) Sequence analysis and comparison of avocado fruit and bean abscission cellulases. Plant Physiol. **95:** 928–933
5 Lashbrook C. C., Gonzales-Bosch C. and Bennett A. B. (1994) Two divergent endo-β-1,4-glucanase genes exhibit overlapping expression in ripening fruit and abscising flowers. Plant Cell **6:** 1485-1493
6 Milligan S. B. and Gasser C. S. (1995) Nature and regulation of pistil-expressed genes in tomato. Plant Mol. Biol. **28:** 691–711

7 Campillo E. and Bennett A. B. (1996) Pedicel breakstrength and cellulase gene expression during tomato flower abscission. Plant Physiol. **111:** 813–820

8 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-β-D-glucanase gene. Plant J. **17:** 203–208

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

10 Trainotti L., Ferrarese L., Vecchia F. D., Rascio N. and Casadoro G. (1999) Two different endo-β-1,4-glucanases contribute to the softening of the strawberry fruits. J. Plant Physiol. **154**: 355–362

11 Brummell D. A., Lashbrook C. C. and Bennett A. B. (1994) Plant endo-1,4-β-D-glucanases. ACS Symp. Ser. **566:** 100–129

12 Foley B. (2001) Heterotrophic nutrition in animals. In: Biology 2nd ed., pp. 486–515, Knox B., Ladiges P., Evans B. and Saint R. (ed), McGraw-Hill, Australia, Roseville, Australia

13 Cleveland L. R. (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

14 Chapman R. F. (1998) The Insects, 4th edn, Cambridge University Press, Cambridge, UK

15 Martin M. M. (1991) The evolution of cellulose digestion in insects. Phil. Trans. R. Soc. Lond. B **333:** 281–288

16 Prins R. A. and Kreulen D. A. (1991) Comparative aspects of plant cell wall digestion in insects. Anim. Feed Sci. Technol. **32:** 101–118

17 Krishna K. (1969) Introduction. In: Biology of Termites, vol. 1, pp. 1–13, Krishna K. and Weesner F. M. (eds) Academic Press, New York

18 O'Brien R. W. and Slaytor M. (1982) Role of microorganisms in the metabolism of termites. Aust. J. Biol. Sci. **35:** 239– 262

19 Abo-Khatwa N. (1978) Cellulase of fungus-growing termites: a new hypothesis on its origin. Experientia **34:** 559–560

20 Martin M. M. and Martin J. S. (1978) Cellulose digestion in the midgut of the fungus-growing termite *Macrotermes natalensis*: the role of acquired digestive enzymes. Science **199:** 1453–1455

21 Rouland C., Lenoir F. and Lepage M. (1991) The role of the symbiotic fungus in the digestive metabolism of several species of fungus-growing termites. Comp. Biochem. Physiol. **99A:** 657–663

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

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

24 Veivers P. C., Mühlemann R., Slaytor M., Leuthold R. H. and Bignel D. E. (1991) Digestion, diet and polyethism in two fungus-growing termites: *Macrotermes subhyalinus* Rambur and *M. michaelseni* Sjøstedt. J. Insect Physiol. **37:** 675–682

25 Martin M. M. (1987) Invertebrate-Microbial Interactions, Ingested Fungal Enzymes in Arthropod Biology, Comstock, Ithaca, NY

26 Parnas I. (1961) The cellulolytic activity in the snail, *Levantina hierosolyma* Boiss. J. Cell. Comp. Physiol. **58:** 195–201

27 Strasdine G. A. and Whitaker D. R. (1963) On the origin of the cellulase and chitinase of *Helix pomatia*. Can. J. Biochem. Physiol. **41:** 1621–1626

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

29 Marshall J. J. and Grand R. J. A. (1976) Characterization of a β-1,4-glucan hydrolase from the snail, *Helix pomatia.* Comp. Biochem. Physiol. **53B:** 231–237

30 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

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

32 Elyakova L. A., Sova V. V. and Vaskovsky V. E. (1968) Cellulase of marine mollusc *Littorina* sp. Biochim. Biophys. Acta **167:** 462–464

33 Yokoe Y. and Yasumasu I. (1964) The distribution of cellulase in invertebrates. Comp. Biochem. Physiol. **13:** 323–338

34 Waterbury J. B. (1983) A cellulolytic nitrogen-fixing bacterium cultured from the gland of Deshayes in shipworms (Bivalvia: Teredinidae). Science **221:** 1401–1403

35 Griffin H. L., Freer, S. N. and Greene R. V. (1987) Extracellular endoglucanase activity by a novel bacterium isolated from marine shipworm. Biochem. Biophys. Res. Commun. **144:** 143–151

36 Greene R. V., Griffin H. L. and Freer S. N. (1988) Purification and characterization of an extracellular endoglucanase from the marine shipworm bacterium. Arch. Biochem. Biophys. **267:** 334–341

37 Iman S. H., Greene R. V. and Hockridge M. E. (1993) Zymographic analyses of carboxymethylcellulases secreted by the bacterium from wood-boring marine shipworms. Biotechnol. Tech.**7:** 579–584

38 Distel D. L. and Roberts S. J. (1997) Bacterial endosymbionts in the gills of the deep-sea wood-boring bivalves *Xylophaga atlantica* and *Xylophaga washingtona*. Biol. Bull. **192:** 253–261

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

40 Brusca R. C. and Brusca G. J. (1990) Invertebrates, Sinauer, Sunderland, Mass., USA

41 Martin M. M. (1983) Cellulose digestion in insects. Comp. Biochem. Physiol. **75A:** 313–324

42 Tracey M. V. (1951) Cellulase and chitinase of earthworms. Nature **167:** 776–777

43 Nielsen C. O. (1962) Carbohydrases in soil and litter invertebrates. Oikos **13:** 200–215

44 Parle J. N. (1963) Micro-organisms in the intestines of earthworms. J. Gen. Microbiol. **31:** 1–11

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

46 Mishra P. C. and Dash M. C. (1980) Digestive enzymes of some earthworms. Experientia **36:** 1156–1157

47 Urbasek F. (1990) Cellulase activity in the gut of some earthworms. Rev. Ecol. Biol. Sol. **27:** 21–28

48 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

49 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

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

51 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

52 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

53 O'Brien G. W., Veivers P. C., McEwen S. E., Slaytor M. and O'Brien R. W. (1979) The origin and distribution of cellulase in the termites, *Nasutitermes exitiosus* and *Coptotermes lacteus*. Insect Biochem. **9:** 619–625

54 Krishna K. (1970) Taxonomy, phylogeny, and distribution of termites. In: Biology of Termites, vol. 2, pp. 127–150, Krishna K. and Weesner F. M. (eds), Academic Press, New York

55 Hogan M., Veivers P. C., Slaytor M. and Czolij R. T. (1988) The site of cellulose breakdown in higher termites (*Nasutitermes walkeri* and *Nasutitermes exitiosus*). J. Insect Physiol. **34:** 891–899

56 Cornwell P. B. (1968) The Cockroach, vol. 1, Hutchinson, London

57 Lo N., Tokuda G., Watanabe H., Rose H., Slaytor M. Maekawa K. et al. (2000) Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. Curr. Biol. **10:** 801–804

58 Cleveland L. R. (1934) The wood-feeding roach *Cryptocercus,* its protozoa, and symbiosis between protozoa and roach. Mem. Am. Acad. Arts Sci. **17:** 185–342

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

60 Zhang J., Scrivener A. M., Slaytor M. and Rose H. A. (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

61 Scrivener A. M. and Slaytor M. (1994) Cellulose digetion in *Panesthia cribrata* Saussure: does fungal cellulase play a role? Comp. Biochem. Physiol. **107B:** 309–315

62 Radford A., Stone P. J. and Taleb F. (1996) Cellulase and amylase complexes. Mycota **3:** 269–294

63 Béguin P. and Lemaire M. (1996) The cellulosome: an exocellular, multiprotein complex specialized in cellulose degradation. Crit. Rev. Biochem. Mol. Biol. **31:** 201–236

64 Bayer E. A., Chanzy H., Lamed R. and Shoham Y. (1998) Cellulose, cellulases and cellulosomes. Curr. Opin. Struct. Biol. **8:** 548–557

65 Scrivener A. M. and Slaytor M. (1994) Properties of the endogenous cellulase from *Panesthia cribrata* Saussure and purification of major endo-β-1,4-glucanase components. Insect Biochem. Mol. Biol. **24:** 223–231

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

67 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-β-1,4-glucanase. Zool. Sci. **14**, 83–93

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

69 Xue X. M., Anderson A. J., Richardson N. A., Anderson A. J., Xue G. P. and Mather P. B. (1999) Characterisation of cellulase activity in the digestive system of the redclaw crayfish (*Cherax quadricarinatus*). Aquaculture **180:** 373–386

70 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. Agric. Biol. Chem. **52:** 633–640

71 Klesov A. A. (1991) Biochemistry and enzymology of cellulose hydrolysis. Biokhimiya **55:** 1295–1318

72 Snodgrass R. E. (1993) Principles of Insect Morphology, Cornell University Press, Ithaca, NY

73 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

74 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

75 Emert G. H., Gum E. K. Jr, Lang J. A., Liu T. H. and Brown R. D. Jr (1974) Cellulases. In: Food Related Enzymes (Advances in Chemistry Series, vol. 136), pp. 79–100, Whitaker J. R. (ed.), American Chemical Society, Washington, DC

76 Smant G., Stokkermans J. P. W. G., Yan Y., De Boer J. M., Baum T. J., Wang X. et al. (1998) Endogenous cellulases in animals: isolation of β-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes. Proc. Natl. Acad. Sci. USA **95:** 4096–4911

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

78 De Boer J. M., Smant G., Goverse A., Davis E. L., Overmars H. A. and Pomp H. et al. (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 Microb. Interact. **9:** 39–46

79 Smant G., Goverse, A. Stokkermans J. P. W. G. and De Boer J. M. (1997) Potato root diffusate-induced secretion of soluble, basic proteins originating from the subventral esophageal glands of potato cyst nematodes. Phytopathology **87:** 839–845

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

81 Rosso M. N., Favery B., Piotte C., Arthaud L., De Boer J. M. and Hussey R. S. et al. (1999) Isolation of a cDNA encoding a β-1,4-endoglucanase in the root-knot nematode *Meloidogyne incognita* and expression analysis during plant parasitism. Mol. Plant Microbe Interact. **12:** 585–591

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

83 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

84 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

85 Byrne K. A., Lehnert S. A., Johnson S. E. and Moore S. S. (1999) Isolation of a cDNA encoding a putative cellulase in the red claw crayfish *Cherax quadricarinatus*. Gene **239:** 317–324

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

87 Gaboriaud C., Bissery V., Benchetrit T. and Mornon J. P. (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett. **224:** 149–155

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

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

90 Ducros V., Czjzek M., Belaich A., Gaudin C., Fierobe H.- P., Belaich J.- P. et al. (1995) Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. Structure **3:** 939–949
91 Bird D. M., Opperman C. H., Jones S. J. M. and Baillie D. L. (1999) The *Caenorhabditis elegans* genome: a guide in the post genomics age. Annu. Rev. Phytopathol. **37**: 247–265
92 Yan Y., Smant G., Stokkermans J., Qin L., Helder J. and Baum T. et al. (1998) Genomic organization of four β-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications. Gene **220:** 61–70
93 Davies G. J., Dodson G. G., Hubbard R. E., Tolley S. P., Dauter Z., Wilson K. S. et al. (1993) Structure and function of endoglucanase V. Nature **365:** 362–364
94 Sakon J., Irwin D., Wilson D. B. and Karplus P. A. (1997) Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca.* Nat. Struct. Biol. **4:** 810–818
95 Béguin P. and Aubert J.-P. (1994) The biological degradation of cellulose. FEMS Microbiol. Rev. **13:** 25–58
96 Syvanen M. (1994) Horizontal gene transfer: evidence and possible consequences. Annu. Rev. Genet. **28:** 237–261
97 Ohtoko K., Ohkuma M., Moriya S., Inove 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 pseratus.* Extremophiles **4:** 343–349
98 Moriya S., Ohkuma M. and Kudo T. (1998) Phylogenetic position of symbiotic protist *Dinemypha exilis* in the hindgut of the termite *Reticulitermes speratus* inferred from the protein phylogeny of elongation factor 1a. Gene **210:** 221–227
99 Ohkuma M., Ohtoko K., Grunau C., Moriya S. and Kudo T. (1998) Phylogenetic identification of the symbiotic hypermastigote *Trichonympha agilis* in the hindgut of the termite *Reticulitermes speratus* based on small-subunit rRNA sequence. J. Euk. Microbiol. **45:** 439–444
100 The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815