Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 349 (2005) 67–75
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*

Tetsushi Inoue[a,b,*], Shigeharu Moriya[a,c], Moriya Ohkuma[a,b], Toshiaki Kudo[a,c]

[a]*JST Bio-Recycle Project/Environmental Molecular Biology Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[b]*JST PRESTO, Kawaguchi, Saitama 332-0012, Japan*
[c]*Laboratory of Environmental Molecular Biology, Graduate School of Yokohama City University, Yokohama 230-0045, Japan*

Received 2 September 2004; received in revised form 15 November 2004; accepted 26 November 2004
Available online 2 March 2005
Received by A. Roger

Abstract

The endo-β-1,4-glucanase gene was cloned from a cDNA library constructed from the mixed population of symbiotic protists in the hindgut of the lower termite, *Coptotermes formosanus*, using the lambda ZAP II vector. The recombinant phage library was screened for cellulolytic activity by the Congo red staining procedure. The nucleotide sequence comprised 941 nucleotides including a polyA tail sequence and showed high sequence similarity with endoglucanase genes belonging to glycosyl hydrolase family 5. Determination of the 5′ end of the cellulase gene using the 5′RACE method showed that the full-length cDNA comprised a 921-bp ORF, encoding a putative 33,620 Da protein. The organismal source of this cellulase gene was identified using PCR with gene-specific primers and whole-cell in situ hybridization as the smallest symbiotic hypermastigote protist, *Spirotrichonympha leidyi*. The optimal pH and temperature of the cellulase heterologously expressed in *Escherichia coli* were 5.8–6.0 and 70 °C, respectively. The Km and Vmax values on carboxymethyl cellulose (CMC) substrate were 1.90 mg/ml and 148.2 units/mg protein, respectively.

Keywords: cDNA library; Cellulose degradation; Glycosyl hydrolase; *Spirotrichonympha leidy*

1. Introduction

The relationship between xylophagous termites and their intestinal protists is one of the most famous examples of symbiosis. Cellulose digestion in termites has fascinated biologists for a long time, particularly following the discovery of the cellulolytic properties of the flagellated protists in lower termites. However, cellulases of protist origin have not yet been analyzed extensively at the molecular level as pure protist cultures are very difficult to establish and only a few species have been axenically cultured (reviewed in Inoue et al., 2000).

Glycosyl hydrolases, including cellulases, show great multiplicity of resulting from the extensive variety of carbohydrate structures. In contrast to the conventional classification of enzymes based on the type of reactions catalyzed and substrate-specificity, a classification of glycosyl hydrolases into families based on amino acid sequence similarities was introduced to integrate both structural and mechanistic features of these enzymes (Henrissat, 1991) and is updated regularly (Henrissat and Bairoch, 1996). Over the years, the number of glycosyl hydrolase families (GHFs) has grown steadily and currently there are 85 families (Bourne and Henrissat, 2001). This classification scheme is considered to take into account evolutionary events such as divergence or convergence (Henrissat, 1991).

Abbreviations: CMC, carboxymethylcellulose; cDNA, DNA complementary to RNA; bp, base pair(s); PCR, polymerase chain reaction; HCA, hydrophobic cluster analysis; GHF, glycosyl hydrolase family; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; pfu, plaque-forming unit(s); IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria–Bertani (medium).

* Corresponding author. Environmental Molecular Biology Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. Tel.: +81 48 467 9648; fax: +81 48 462 4672.
E-mail address: i-tetsu@riken.jp (T. Inoue).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.11.048

Despite the critical role of symbiotic flagellates in cellulose digestion by lower termites, the first sequenced cellulase gene was of termite origin (Watanabe et al., 1998). The presence of two separate cellulolytic systems in the lower termite is now well established (Nakashima et al., 2002a,b), and this dual system seems to result in the high assimilation rate (greater than 90%) of wood glucan by termites (Breznak and Brune, 1994). The endogenous cellulases of termites were classified into GHF9 and their homologues were also found in cockroaches and decapods (Lo et al., 2003).

On the other hand, diverse genes encoding cellulase homologues belonging to GHF45 were identified from the symbiotic parabasalian protists of the termite, *Reticulitermes speratus* (Ohtoko et al., 2000) and *Mastotermes darwiniensis* (Li et al., 2003). Furthermore, cellulase components were isolated from the hindgut of the termites, *Coptotermes* spp., and the isolated sequences showing similarity to catalytic domains of GHF7 members were obtained from the symbiotic protists, *Pseudotrichonympha grassii* and *Holomastigotoides mirabile* (Nakashima et al., 2002a; Watanabe et al., 2002).

Symbiotic flagellates found in the hindgut of the lower termites belong to the orders Trichomonadida, Hypermastigida, and Oxymonadida (Yamin, 1979). Trichomonadida and Hypermastigida are classified into the class Parabasalea or into the phylum Parabasalia.

Functional differences in cellulose metabolism among protistan species have been demonstrated in the lower termites (reviewed in Inoue et al., 2000) and the diversity of cellulase genes of protist origin is a possible explanation of the functional differences. Therefore, the goal of this study is to elucidate the diversity of cellulase genes of symbiotic protist origin using a culture-independent approach. In this study, we constructed a cDNA library from the mixed population of protists in the hindgut of the subterranean termite, *Coptotermes formosanus*, which harbors three species of parabasalian flagellates in the hindgut, namely *Spirotrichonympha leidyi*, *H. mirabile*, and *P. grassii*. The previous phylogenetic analysis suggested that the *Pseudotrichonympha* group might be the most ancient lineage of parabasalids based on the sequences of small subunit ribosomal RNA genes (Ohkuma et al., 2000).

Here we describe the characterization of a cellulase gene, which belongs to GHF5, from symbiotic protist, *S. leidyi*. The isolated gene was heterologously expressed in *Escherichia coli* and the properties of the recombinant protein were determined.

2. Materials and methods

2.1. Construction and screening of a symbiotic protist cDNA library

Field colonies of *C. formosanus* were collected from Shirahama, Wakayama Prefecture, Japan. Gut contents (50 termites), including 3 species of protists, were suspended in 200 μl of Solution U (Trager, 1934) and washed twice with 1 ml of Solution U. Total RNA was extracted using ISOGEN (Nippon Gene, Japan) and poly(A)-mRNA was purified using Oligotex-dT30 (Super) (Nippon Roche, Japan) according to the manufacturer's instructions. mRNA (ca. 4 μg from 250 termites) and oligo (dT)20 primer (200 pmol) were heat denatured and first-strand cDNA was synthesized by reverse transcriptase (Superscript II, Gibco BRL, USA) in a final volume of 50 μl according to the manufacturer's instructions. After completion of the second-strand synthesis, the uneven termini of the double-stranded cDNAs were filled in with T4 DNA polymerase, and then *Eco*RI linkers (5′-GCTTGAATTCAAGC-3′) were ligated to the blunt ends. cDNA fragments were ligated into the *Eco*RI site of the lambda ZAP II vector (Stratagene, USA), and the ligation mixture was packaged into phages using the GigapackIII Gold packaging extract (Stratagene). The primary library, containing approximately 3×10^5 independent clones was amplified (titer: 3×10^9 pfu/ml).

The recombinant phage library was screened for cellulolytic activity by plating cells on 0.7% (w/v) 2xYT top agar containing 0.5% (w/v) carboxymethyl cellulose (CMC) (medium viscosity; Sigma, USA) along with 15 μl of 50 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Plaques having carboxymethyl cellulase activity were recognized by the formation of clear halos on a red background after staining with 0.1% Congo red and destaining with 1 M NaCl. Positive plaques displaying enzyme activity were purified at least three times. DNA from positive clones was transferred to the plasmid vector pBluescript SK(−) by in vivo excision using the ExAssist/SOLR system (Stratagene). Then, cDNA fragment *CFP-eg1* was isolated and the nucleotide sequences of both strands were sequenced.

2.2. Amplification of 5′ end cDNA

First-strand cDNA synthesis was carried out by the method of Moriya et al. (2001). In brief, after removal of the CAP structures from the 5′ end of the immobilized mRNA, capping oligoribonucleotide fragments (5′-GAGCACUGUUGGCCUACUGG-3′) were ligated to the 5′ ends. Then cDNA was amplified by PCR using the primer corresponding to the capped oligoribonucleotide sequence and the poly(dT)20 primer. The amplified cDNAs were purified and ligated into a pGEM-T plasmid (Promega, USA). The cDNA 5′ end was isolated from the purified plasmid using the gene-specific primer for *CFP-eg1* (5′-ACTCAGGCCACCAATTATGCCA-3′) combined with the primer derived from flanking sequences in the vector (T7 promoter primer). The isolated cDNA was cloned into pGEM-T plasmids (Promega) and sequenced.