Journal of Microbiology (2013) Vol. 51, No. 3, pp. 329–335
DOI 10.1007/s12275-013-2697-5

Cloning and Functional Characterization of Endo-β-1,4-Glucanase Gene from Metagenomic Library of Vermicompost[§]

Muhammad Yasir[1,2], Haji Khan[2], Syed Sikander Azam[3], Amar Telke[2], Seon Won Kim[2], and Young Ryun Chung[2]*

[1]*King Fahd Medical Research Center, King Abdulaziz University, Jeddah, 21589, Saudi Arabia*
[2]*Division of Applied Life Science (BK 21), Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju 660-701, Republic of Korea*
[3]*National Center of Bioinformatics, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad-45320, Pakistan*

(Received December 20, 2012 / Accepted February 19, 2013)

In the vermicomposting of paper mill sludge, the activity of earthworms is very dependent on dietetic polysaccharides including cellulose as energy sources. Most of these polymers are degraded by the host microbiota and considered potentially important source for cellulolytic enzymes. In the present study, a metagenomic library was constructed from vermicompost (VC) prepared with paper mill sludge and dairy sludge (fresh sludge, FS) and functionally screened for cellulolytic activities. Eighteen cellulase expressing clones were isolated from about 89,000 fosmid clones libraries. A short fragment library was constructed from the most active positive clone (cMGL504) and one open reading frame (ORF) of 1,092 bp encoding an endo-β-1,4-glucanase was indentified which showed 88% similarity with *Cellvibrio mixtus* cellulase A gene. The endo-β-1,4-glucanase *cmgl*504 gene was overexpressed in *Escherichia coli*. The purified recombinant cmgl504 cellulase displayed activities at a broad range of temperature (25–55°C) and pH (5.5–8.5). The enzyme degraded carboxymethyl cellulose (CMC) with 15.4 U, while having low activity against avicel. No detectable activity was found for xylan and laminarin. The enzyme activity was stimulated by potassium chloride. The deduced protein and three-dimensional structure of metagenome-derived cellulase cmgl504 possessed all features, including general architecture, signature motifs, and N-terminal signal peptide, followed by the catalytic domain of cellulase belonging to glycosyl hydrolase family 5 (GHF5). The cellulases cloned in this work may play important roles in the degradation of celluloses in vermicomposting process and could be exploited for industrial application in future.

Keywords: cellulose, metagenomic, vermicompost, glycosyl hydrolase family 5

*For correspondence. E-mail: yrchung@gnu.ac.kr; Tel.: +82-55-772-1326; Fax: +82-55-772-1329
[§]Supplemental material for this article may be found at http://www.springerlink.com/content/120956.

Introduction

There has been increasing interest in cellulolytic enzymes of microbial origin in recent years due to their capability of digesting plant cellulosic materials. Cellulosic biomass is one of the best sustainable sources available for biofuel production, by virtue of its relatively low cost, abundant supply, and cleanliness (Lynd *et al*., 2008). However, the extensive utilization of this important resource is hindered by the general deficiency of low-cost technology for hydrolysis of cellulosic biomass into soluble sugar. A potential strategy to overcome this difficulty involves the using of cellulolytic enzymes of microorganisms that can degrade cellulose to glucose through the interactive hydrolysis of three classes of cellulase: endo-β-1,4-glucanase (EC3.2.1.4), cellobiohydrolase (EC3.2.1.91), and β-glucosidase (EC3.2.1.21) (Lynd *et al*., 2002). Among these three enzymes, endo-β-1,4-glucanase (EC 3.2.1.4) is a key enzyme involved in cleavage of β-1,4-glycosidic bonds within cellulose chains, releasing smaller fragments of random length (Watanabe and Tokuda, 2001). Consequently, there has been a fleet growth in demand for endo-β-1,4-glucanase in various industrial applications including those related to waste management, biopolishing process and deinking (Kuhad *et al*., 2011). In addition, hydrolysis of the glycosidic bond of β-glucan by endoglucanase is important to decrease the viscosity of β-glucan solution, benefits beer brewing, improves the quality of animal feed and is also proved promising for biofinishing in the textile industry (Kuhad *et al*., 2011). Specifications of these applications proposed the investigation to identify a broad range of endoglucanases with varying optimal pH and temperature.

Cellulolytic microorganisms are widespread in the environment and are mostly found in areas where plant biomass tends to accumulate and deteriorate, such as vermicompost (VC) (Lednicka *et al*., 2000; Akasaka *et al*., 2003; Aira *et al*., 2007). Vermicomposting is widely used for the treatment of industrial organic wastes, such as paper mill sludge that is mainly composed of polymer carbohydrates, such as cellulose and hemicelluloses (Elvira *et al*., 1998; Singh *et al*., 2008). Although some species of earthworm can produce cellulases but in general decomposition and transformation of paper mill sludge during vermicomposting is attributed to microbiota inhabiting the earthworm gut (Aira *et al*., 2007; Aira and Dominguez, 2011). Especially, bacterial diversity is increased during vermicomposting (Yasir *et al*., 2009). The vast microbial diversity of vermicompost presents an enormous genetic and biological collection that can be used for the recovery of novel genes and biomolecules (Cowan, 2000). However, only 0.1% to 1.0% of bacteria are culturable

using typical laboratory cultivation techniques. Therefore, few microbial community genes have been studied using culture-based approaches (Rondon *et al*., 2000). So far, the majority of natural cellulases have been isolated from cultured microorganisms.

Characterization of cellulases from uncultured bacteria could reveal much about the cellulolytic diversity and the industrial applications of the bacteria (Feng *et al*., 2007; Kim *et al*., 2011). The recent development of metagenomic strategies has enabled the isolation of novel genes having broad substrates specific cellulase activities from various environmental samples including soil, compost and rabbit cecum (Feng *et al*., 2007; Kim *et al*., 2008, 2011). In present study, metagenomic libraries were constructed to explore the cellulolytic potential in the metagenome of the microbial consortia from VC of paper mill sludge. The use of both culture-dependent and culture-independent approaches to analyze the microbial consortia in VC could provide new biotechnological strategies for the conversion of cellulose, hemicellulose and lignocellulosic biomass into monosaccharides.

Materials and Methods

Preparation of VC samples

VC samples were prepared from paper mill sludge supplemented with dairy sludge (Fresh sludge, FS) to adjust the C:N ratio, and the fresh sludge samples were treated with earthworms (*Eisenia fetida*) in a 10:1 (w/w) ratio. The newly produced VC samples were collected after 15 days composting as described previously (Yasir *et al*., 2009).

Construction and functional screening of metagenomic libraries

Metagenomic DNA was extracted according to Zhou *et al*. (1996). The extracted high molecular weight DNA was separated by size in a 1% low melting point (LMP) agarose gel containing 2% polyvinyl-pyrrolidone (PVP) at 34 V for 24 h. The 30–48 kb metagenomic DNA was ligated into a pEpiFOS-5 fosmid vector (Epicenter Biotechnologies, USA) according to the manufacturer's instructions. The ligation mixture was packaged into lambda phages using MaxPlax Lambda Packaging Extracts (Epicenter Biotechnologies). The packaged library was transduced into *Escherichia coli* EPI-100 and *E. coli* transformants were selected on LB agar medium supplemented with chloramphenicol (20 μg/ml) according to the manufacture protocol (Epicenter Biotechnologies). The library clones were stored in 25% glycerol as clone pools, with approximately 400 clones per pool. The libraries were spread onto LB agar supplemented with 0.25% CMC and chloramphenicol (20 μg/ml). Metagenomic libraries were screened for cellulolytic activity by a Congo red staining method (Feng *et al*., 2007). Colonies with clear zone in the red background were selected for CMC hydrolysis. The active clones were reconfirmed by isolation of fosmids and retransformed into *E. coli* EPI-100. Cellulase (CMCase) activity of these was checked by replating them as described above. The fosmid DNAs from active clones were digested with *Bam*HI and their digestion pattern was analysed by agarose gel electrophoresis.

Identification of cellulase gene

Among the cellulolytic clones identified from metagenomic libraries, the most active cellulase positive fosmid clone cMGL504 was digested with the restriction enzyme *Eco*RI, subcloned in pUC19 (Invitrogen, USA) and screened on LB agar plates containing 0.25% CMC and 50 μg/ml ampicillin with Congo red staining as described previously (Feng *et al*., 2007). The smallest fragment showing cellulase activity was sequenced through primer walking. Possible open reading frames (ORFs) were identified with the ORF finder at the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov). Basic local alignment search tool (BLAST) search of the DNA fragment sequences were made with BlastN, BlastX, and BlastP at the NCBI website. Module structures of the enzyme were predicted by the simple modular architecture research tool (SMART; http://smart.embl-heidelberg.de). Amino acids sequences of the cloned cellulase cmgl504 and closely related matches of glycosyl hydrolase family GH5 retrieved from the BLAST search were used for the phylogenetic analysis.

Overexpression and purification of the recombinant endoglucanase

The newly identified cellulase gene *cmgl504* was amplified by polymerase chain reaction (PCR) with ORF specific primer pairs of cMGL504F (GATCGAATTCATGACTACAG CAAAAACTCTC) and cMGL504R (GATCAAGCTTCTT ATTGTCCTGCAACATCAC). The *Eco*RI and *Hin*dIII sites were added to the 5′ ends of forward and reverse primers, respectively. The PCR product was purified, digested with *Eco*RI and *Hin*dIII, and cloned into expression vector pET-28a(+) (Novagen, USA) for expression of *cmgl504* with 6xHis tag at the N-terminus. The recombinant plasmid was transformed into *E. coli* BL21 (DE3) pLysS (Novagen) and positive clones were selected by kanamycin resistance and Congo red staining. For expression of recombinant protein, the cells containing the recombinant plasmid were first grown to an optical density of 0.6 at 600 nm in LB medium containing 30 μg/ml of kanamycin at 37°C on a rotary shaker (200 rpm) and the expression of the *cmgl504* gene was induced by adding 0.5 mM isopropyl-β-D-thio-galactoside (IPTG) to the medium. Following a further incubation of cells at 25°C and 100 rpm overnight, the recombinant protein was extracted from the cytoplasmic fraction of the cell lysates. The extracted protein was purified by affinity chromatography through Ni-NTA columns (Qiagen, USA) according to the manufacture's protocol. The level of purity of cmgl504, as well as the molecular mass, was determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Homology modeling

The amino acid sequence of cmgl504 was submitted to NCBI PSI-BLAST (Altschul *et al*., 1997), a database similarity search tool, which searches the Brookhaven PDB database to detect distant evolutionary relationships by identifying the closest homologous structure. A template structure for a given target

sequence was recognized by taking into account the significance of the score, which indicates fitness of the target to the template; the E-value or the sequence identity are the two most common criteria. Computational studies were carried out on an Intel® Core™2 Duo 3.33GHz equipped with a Linux environment running under open SUSE 11.4. Multiple sequence alignment was done by ClustalW (Larkin *et al.*, 2007). The homology modeling was performed by Modeller version 9.9 (Sali and Blundell, 1993). Validation of protein structures was determined by Procheck (Laskowski *et al.*, 1993), Verify 3D (Bowie *et al.*, 1991), and Errat plots. Visualization was carried out by Chimera software (Pettersen *et al.*, 2004). Out of five predicted iterative models, the best model was selected after running validation methods.

Enzymatic assays

Cellulase activity was performed with 1% (w/v) carboxymethyl-cellulose in a total volume of 1.0 ml for a 30 min reaction time. The reaction mixture contained 20 µg of purified cmgl504 cellulase and 50 mM sodium acetate buffer (pH 5.5). The substrate specificity of cmgl504 cellulase was checked using the cellulosic substrates avicel (1%, w/v), laminarin (1%, w/v) and birch wood xylan (1%, w/v). The dinitrosalicylic acid (DNS) method was used to detect the product formation in these reactions (Miller *et al.*, 1960). One unit of enzyme activity was expressed as the amount of enzyme that released 1 µmol reducing sugar per min. The optimum temperature for the activity of the purified cmgl504 cellulase was determined by incubating the mixture of enzyme and 1% (w/v) CMC in 50 mM sodium phosphate buffer (pH 5.5) for 30 min at temperatures ranging from 25°C to 70°C. For determination of optimal pH, the reaction mixture of the enzyme and substrate was incubated in 50 mM of the appropriate buffer. Sodium acetate (pH 4–5.5), sodium phosphate (pH 6–8), Tris-HCl (pH 8–9) and glycine-NaOH (pH 9–10) buffers were used to maintain the respective pH. The reaction mixtures in different pH buffers were incubated for 30 min at 50°C and the enzyme activity was measured by the DNS method. Effect of different metal ions on the enzyme activity was investigated under the standard assay condition. One mM of each metal ions, KCl, $LiCl_2$, $CaCl_2$, $CuCl_2$, $FeCl_2$, $FeCl_3$, $ZnCl_2$, $CoCl_2$, $MgCl_2$, and $MnCl_2$ was applied in the reaction mixture. All enzyme assays carried out in triplicates and average value was calculated.

Phylogenetic analysis

Multiple alignments were preformed with ClustalX2.0 program (Larkin *et al.*, 2007). The phylogenetic trees were constructed by the distance method (neighbor-joining) using MEGA4 software (Tamura *et al.*, 2007). Bootstrap values were calculated on the basis of 1,000 replications. The related sequences in the phylogenetic tree were retrieved from the GenBank database (http://blast.ncbi.nlm.nih.gov). The *cmgl504* gene sequence has been deposited to GenBank under accession numbers JN012243.

Results and Discussion

Construction of the metagenomic libraries and screening of cellulase clones

To clone metabolically active genes from the VC microbial community, fosmid libraries were constructed with metagenomic DNA isolated directly from VC samples. The libraries yielded about 89,000 clones, and their qualities and insert sizes were analyzed by restriction digestion of purified fosmid from 18 randomly chosen clones. The inserted DNA fragments of these selected clones showed distinct restriction patterns, indicating that each clone represented a unique sequence (data not shown). The average insert size for the clones was estimated to be 36 kb and the full library size was about 3,204 Mb. Eighteen clones expressing cellulase activities were identified from the VC metagenomic libraries. Fosmid isolated from all active clones were re-transformed into *E. coli* EPI-100, and all transformants retained the cellulase activity that was encoded by the respective fosmid inserts. The hit obtained from these libraries indicated the enrichment of cellulolytic genes in the total VC bacterial community. This result was consistent with other metageno-

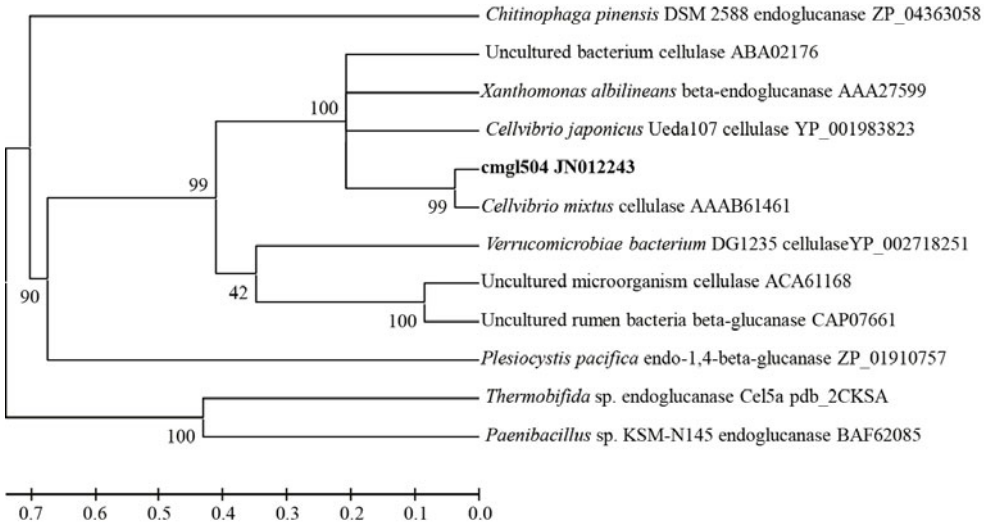

Fig. 1. Phylogenetic tree of catalytic domains of cmgl504 and GHF 5 cellulases was generated by the neighbor-joining method. The amino acid sequences of family 5 glycosyl hydrolases were retrieved from GenBank by homology searching against protein database with the cloned cellulase of cmgl504, and submitted to SMART for determining the catalytic domain of each enzyme. Bootstrap values are based on 1,000 replications shown at branch point.

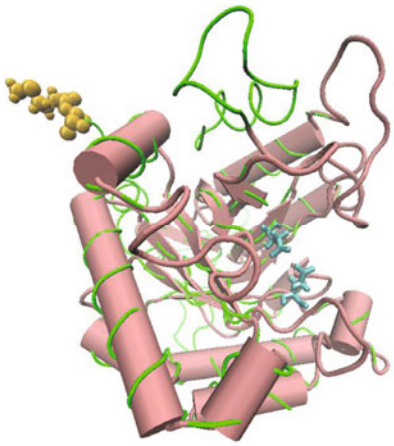

Fig. 2. Three dimensional structure of cmgl504 was constructed by utilizing MODELLER version 9.9. Visualization was carried out by Visual Molecular Dynamics (VMD). The catalytic domain from position 62–338 is represented in cartoon; 1–25 single peptide is represented in VDW; glutamate residue 193 (proton donor) and 289 (nucleophile) is represented in licorice.

mic libraries constructed from rumen samples and screened for relevant enzymes (Feng *et al.*, 2007; Duan *et al.*, 2009). Additionally, the functional analysis of the cultured isolates also indicates that the prevalence of cellulolytic bacteria in VC. These bacteria associated with vermicomposting of paper mill sludge probably play a significant role during composting of the fresh sludge (Aira *et al.*, 2007).

The most active cellulase-positive fosmid clone cMGL504 was initially selected for further studies. To locate the cellulase genes from the positive clone and facilitate sequencing, we further subcloned the cMGL504 with shorter insert sizes. A subclone of 4.2 kb that retained the corresponding cellulase activity was characterized by complete sequencing for cellulase genes and their neighboring ORF. The subclone harbored a complete ORF encoding for endo-β-1,4-glucanase. The predicted products of the cellulase gene of *cmgl504* consisted of 363 amino acids cellulase; the molecular mass was 41.6 kDa (Supplementary data Fig. 1S). The amino acid sequence of the cmgl504 cellulase showed that it contained a glycosyl hydrolase (GH) family 5 catalytic domain and a N-terminal signal peptide. The deduced amino acid sequence of the isolated cellulase gene was compared with the published sequences of other cellulases for phylogenetic analysis. The phylogenetic tree showed that the cellulase of cmgl504 was closely related to cellulase of *Cellovibrio mixtus* (Fig. 1). The large-scale metagenomic sequencing of hindgut bacteria of a wood feeding higher termite and cellulase clones from other environments also supported the view that the GH5 families predominate in nature (Voget *et al.*, 2006; Feng *et al.*, 2007; Warnecke *et al.*, 2007). Another clone from metagenomic libraries of VC belonged to the GH9 family (data not shown), which indicates that cellulase genes from other GH families also exist in VC. Hydrolysis of the polysaccharides requires the combined action of enzymes expressed from the genes of different GH families that are organized in clusters formed in the genomes of the microbiota, and their proteins function synergistically to degrade the cellulosic materials (DeBoy *et al.*, 2008). The existence of different GH families clones in VC is probably responsible for the degradation of different polysaccharides and cellulose fibers in the paper mill sludge engulfed by the earthworm.

Structure prediction

The amino acid sequence of cmgl504 based on a PSI-BLAST homology search revealed 14 hits out of which the crystal structure of the catalytic domain of *Thermobifida fusca* endoglucanase (2CKS) was considered as template. The 2CKS showed the similar occupancy pattern, 76% sequence similarity and an E-value of 9e-18 with cmgl504. Five models of cmgl504 cellulase were constructed by MODELLER. Model 3 revealed an excellent agreement between the predicted secondary structure and the experimentally determined three-dimensional (3D) structure from *T. fusca* endoglucanase (Fig. 2). The geometrical and structural consistencies of the modeled protein were evaluated by different approaches. The stereochemical quality and protein structure of the final models were validated by the Procheck program; Model 3 was the best, according to the comparative values in Table 1. According to the Ramachandran plot, about 84% of the residues fell in the most favorable region and 11.9% in the additional regions. Five residues (1.6% of the total) were present in the disallowed region. The phi and psi distributions of Ramachandran plots are summarized in Supplementary data Fig. 2S. In addition, two more protein evaluation programs were utilized to check the stereochemistry of the 3D models. Verification of the 3D scores established the com-

Table 1. Analysis of predicted protein structure of cmgl504 cellulase by Procheck

Regions	Model 1	Model 2	**Model 3**	Model 4	Model 5
Residues in most favoured regions [A,B,L]	258 81.1%	252 79.2%	267 84.0%	258 81.1%	259 81.4%
Residues in additional allowed regions [a,b,l,p]	46 14.5%	45 14.2%	38 11.9%	50 15.7%	42 13.2%
Residues in generously allowed regions [~a,~b,~l,~p]	09 2.8%	12 3.8%	08 2.5%	07 2.2%	14 4.4%
Residues in disallowed regions	05 1.6%	09 2.8%	05 1.6%	03 0.9%	03 0.9%
Number of non-glycine and non-proline residues	318	318	318	318	318
Number of end-residues (excl, Gly and Pro)	02	02	02	02	02
Number of glycine residues (shown as triangles)	23	23	23	23	23
Number of proline residues	17	17	17	17	17

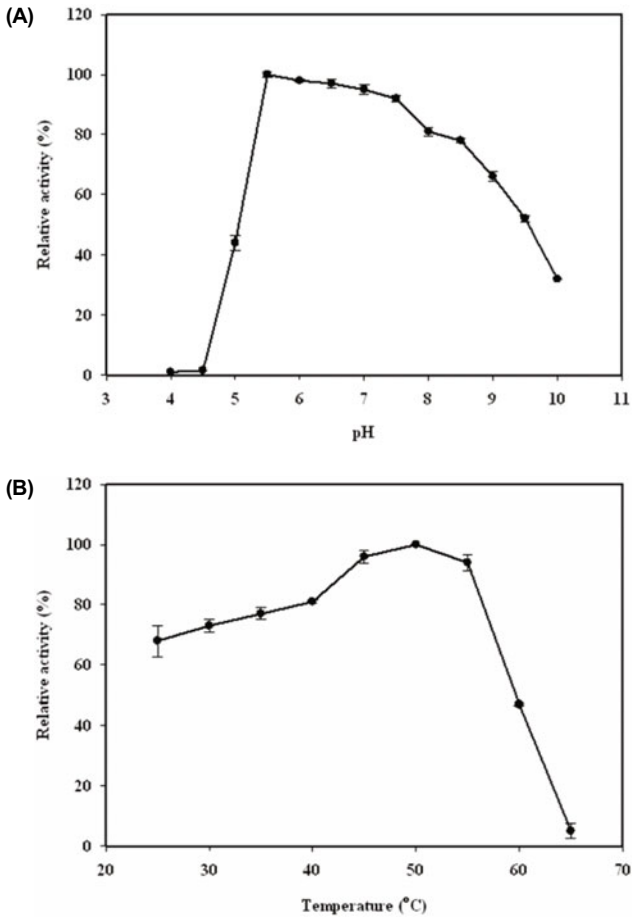

Fig. 3. Effects of pH and temperature on the recombinant cellulase activity of the cmgl504. (A) The activity was measured in acetate buffer (pH 4.0–6.0), phosphate buffer (pH 6-8.0), Tris-HCl buffer (pH 8.0–9.0) and glycine-NaOH buffer (pH 9.0–10.0). Enzyme assays were performed at indicated pH at 50°C for 30 min. (B) The activity was determined in a temperature range from 25°C to 70°C at pH 6.0 (phosphate buffer).

patibility between amino acid sequences and the environment of the amino acid side chains in the model. In the verified 3D plot, the majority of amino acid residues had scores lower than zero, indicating they were sufficiently suited according to their environment with exception of only three amino acid residues. The Errat program was utilized to assess the arrangement of different types of atoms with respect to one another in the protein models. It is a sensitive technique that is suitable for identifying incorrectly folded regions in preliminary protein models. The program yielded an overall quality factor of about 70% for cmgl504 cellulase. The cmgl504 shared a general α/β barrel architecture with other GHF5 cellulases (Fig. 2). Eight α-helices were observed along with three incomplete α helices showing only a single turn. These α-helices forming ring surrounding eight parallel β-sheets. Three anti-parallel β-sheets were also observed interacting with the loops. The whole structure was packed with loops and turns inter-connecting all the secondary structural elements. The cmgl504 was found to possess a 25 amino acid signal peptide at N-terminal preceding the catalytic domain. The cleavage site of the signal peptide was predicted between amino acid 25 and 26 with 0.99 probability by the SignalP program. Protein homologies analysis confirmed the presence in cmgl504 of a family GH5 catalytic domain located between amino acid 62 and 338 with no distinct non-catalytic cellulose-binding domain (CBD). On the basis of sequence homologies with other members of GH5 family, two conserved catalytic residues, Glu193 and Glu289 could be predicted as the putative proton donor and nucleophile, respectively. The comparative SMART and SignalP program analysis between cmgl504 and CelA revealed that signal peptide is not present in the closely related CelA cellulase. Moreover, very least sequence similarity which was observed between the two genes at the initial 78 nucleotides from the 5′ end that are crucial for the endoglucanase activity of CelA indicates the structural difference between the two endoglucanases.

Characterization of enzyme activities

The cmgl504 cellulase displayed activity over a broad pH range with an optimum at pH 5.5–8.5 as compared to the closely related cellulase from *C. mixtus* with the optimum pH between7.0–8.5 (Fontes *et al.*, 1997). The purified cmgl504 cellulase enzyme exhibited optimal activity (100%) at pH 5.5 (Fig. 3A). At pH 6.0–8.5, the residual enzyme activity ranged from 78–98% of the optimal activity. The activity still remained in 52–66% over pH 9, but it was reduced to 32% at pH 10. As cmgl504 cellulase is a metagenome-derived from the vermicompost sample that was slightly acidic and was not obvious source for extremophiles. The enzymes isolated from the mesophilic environment have adapted to cope with limited fluctuations of the environment and loss of their activity at more stringent conditions, which are probably the reason for the severe loss of cmgl504 cellulase activity below pH5 (Voget *et al.*, 2006). The predicted isoelectric point of cmgl504 was 7.9 calculated by using DNASTAR software. The activity of cmgl504 cellulase at alkaline conditions was probably a consequence of its high isoelectric point that allowed maintenance of proper structure in the alkaline environment. The activity of cmgl504 cellulase was measured at various temperatures ranging from 25°C to 65°C. As shown in Fig. 3B, the enzyme exhibited maximum activity (100%) at 50°C. The relative enzyme activity ranged from 68±5.2% to 81±0.5% between 25°C and 40°C. The enzyme activity at 45°C and 55°C was 96±2.1% and 94±8.0%, respectively. However, the activity was remarkably reduced to 47±0.4% of the optimal activity at 60°C and at 65°C the residual activity was only 5±2.5%. The compact structure of cmgl504 as shown in Fig. 3 supports the enzymatic activity

Table 2. Substrate specificity of the purified cmgl504 cellulase on polysaccharides

Substrate	Enzyme activity (U/mg of protein)	Relative activity (%)[b]
CMC	15.4	100.0 ± 6.0
Avicel	0.003	0.019 ± 0.005
Xylan	ND	ND
Laminarin	ND	ND

[a] ND, Not detected
[b] Relative enzyme activity was the ratio of the activity of each sample to the maximum activity of the sample. Relative activity was calculated from the specific activity of enzyme. The results are mean values from three experiments.

Table 3. Effect of metal ions on the cellulase activity of cmgl504

Metal ion[a]	Relative activity (%)
None	100
K^+ (KCl)	119 ± 2.8
Li^+ (LiCl)	94 ± 4.9
Mg^{2+} ($MgCl_2$)	95 ± 2.1
Mn^{2+} ($MnCl_2$)	25 ± 3
Co^{2+} ($CoCl_2$)	52 ± 0.1
Ca^{2+} ($CaCl_2$)	84 ± 3.2
Cu^{2+} ($CuCl_2$)	49 ± 3.3
Fe^{2+} ($FeCl_2$)	73 ± 3.3
Fe^{3+} ($FeCl_3$)	66 ± 2.6
Zn^{2+} ($ZnCl_2$)	68 ± 8.1

[a] 1 mM of each metal ion was added to the reaction mixture.

at high temperature due to strong amino acid interaction that helps to maintain the tertiary structure of the enzyme at high temperature (Fontes *et al*., 1997). The substrate specificity of the purified cmgl504 cellulase was investigated using different substrates. The enzyme degraded CMC with 15.4 U, while having low activity against avicel. No detectable activity was found for xylan and laminarin (Table 2) as compared to the closely related CelA cellulase from *C. mixtus* (Fontes *et al*., 1997). The relative activity for CMC was 100±6%. The effect of metal ions, KCl, $LiCl_2$, $CaCl_2$, $CuCl_2$, $FeCl_2$, $FeCl_3$, $ZnCl_2$, $CoCl_2$, $MgCl_2$, and $MnCl_2$ on the enzyme activity was investigated. The residual activity measured in the presence of each metal ion is shown in Table 3. With addition of 1 mM of KCl, the enzyme activity was increased by 119±2.8% of the non-added control. However, other ions such as $LiCl_2$, $MgCl_2$, and $CaCl_2$ slightly reduced the enzyme activity by 94±4.9%, 95±2.1%, and 84±3.2%, respectively. In the presence of 1 mM metal ions of $FeCl_2$, $ZnCl_2$, $FeCl_3$, $CoCl_2$, and $CuCl_2$ the cellulase activity was reduced to 73±3.3%, 68±8.1%, 66±2.6%, 52%±0.1%, and 49%±3.3%, respectively. The residual activity of the enzyme was only 25±3% of the original acitivity in the presence of 1 mM $MnCl_2$. In the presence of low concentration of $MnCl_2$, $CuCl_2$, $CoCl_2$, and $FeCl_3$, the cellulase activity was reduced to less than 70% of the control activity, which distinguished this CMGL504 cellulase from most of the other metagenomic derived cellulases (Voget *et al*., 2006; Feng *et al*., 2007). Several studies have demonstrated different activation and inhibition pattern of the cellulases (Clarke and Adams, 1987; Mawadza *et al*., 2000; Voget *et al*., 2006). Mawadza *et al*. (2000) observed that different metal ions including Na^+, K^+, Mg^{2+}, Cu^{2+}, Ni^{2+}, Ca^{2+}, Zn^{2+}, and Fe^{3+} at 1 mmol/L concentration had little effect on the activity of cellulase whereas the inhibitory effect of Co^{2+} was significant at the same concentration. In addition, other metal ions, particularly Ag^+, Mn^{2+}, and Hg^{2+} showed a tendency to inhibit cellulase activity (Clarke and Adams, 1987). Santose *et al.* (2012) proposed from the functional and crystallographic studies of GH5 cellulase that Mn^{2+} ions increase the stability and lower the enzymatic efficiency of BsCel5A endoglucanase probably due to a consequence of the higher overall molecular rigidity, which prevents the conformational modification necessary for optimal catalysis. This observation is supported by the other studies that previously described the fundamental role of ligand-induced conformational changes in the catalytic activity of family 5A cellulases (Varrot *et al*., 2002).

Microbial cellulases, particularly endoglucanase have shown their potential applications in various industries including textile, laundry, brewery and wine, animal feed, pulp and paper industries as well as in agriculture and for research purposes. For a variety of industrial applications, broad pH range and relatively high thermostability is an attractive and desirable characteristic of an enzyme (Fujinami and Fujisawa, 2010; Li *et al*., 2011). The purified cellulase cmgl504 displayed activity over a broad pH range with an optimum at pH 5.5–8.5 as compared to the closely related cellulases of CelA and Cel5A isolated from *C. mixtus* and uncultured microorganism, respectively (Fontes *et al*., 1997; Voget *et al*., 2006). In addition, cmgl504 showed activity at alkaline pH and in the presence of a range of divalent cations compared to the closely related Cel5A that showed different pattern of activity in the presence of 1 mM Fe^{2+}, Fe^{3+}, Cu^{2+}, Zn^{2+}, Ca^{2+}, and Mn^{2+}. Most of the commercial cellulases produced by *Bacilli* and other cellulolytic strains are optimum at neutral pH (Sukumaran *et al*., 2005). All the characteristics together with broad pH and temperature range support the further investigation of cmgl504 by mutagenesis and codon optimization that can probably enhance the efficiency, pH and temperature tolerance of cmgl504 cellulase. These will make this enzyme a potential candidate for the broad application in textile, paper, bio-polishing and detergent industries.

Acknowledgements

This work was supported by Brain Korea (BK) 21 project in 2007–2008, the Ministry of Education, Science and Technology, and partially supported by the R&D program of KEIT (#10034753) in 2009–2010, the Ministry of Knowledge Economy, Korea.

References

Aira, M. and Dominguez, J. 2011. Earthworm effects without earthworms: inoculation of raw organic matter with worm-worked substrates alters microbial community functioning. *PLoS One* **6**, e16354.

Aira, M., Monroy, F., and Dominguez, J. 2007. Earthworms strongly modify microbial biomass and activity triggering enzymatic activities during vermicomposting independently of the application rates of pig slurry. *Sci. Total Environ.* **385**, 252–261.

Akasaka, H., Izawa, T., Ueki, K., and Ueki, A. 2003. Phylogeny of numerically abundant culturable anaerobic bacteria associated with degradation of rice plant residue in Japanese paddy field soil. *FEMS Microbiol. Ecol.* **43**, 149–161.

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

Bowie, J.U., Lüthy, R., and Eisenberg, D. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, 164–170.

Clarke, A.J. and Adams, L.S. 1987. Irreversible inhibit-transition metal ions. *Biochim. Biophysica Acta (BBA)–Protein Structure and Molecular Enzymology* **916**, 213–219.

Cowan, D.A. 2000. Microbial genomes the untapped resource. *Trends Biotechnol.* **18**, 14–16.

DeBoy, R.T., Mongodin, E.F., Fouts, D.E., Tailford, L.E., Khouri, H., Emerson, J.B., Mohamoud, Y., Watkins, K., Henrissat, B., Gilbert, H.J., and *et al.* 2008. Insights into plant cell wall degradation from the genome sequence of the soil bacterium *Cellvibrio japonicus*. *J. Bacteriol.* **190**, 5455–5463.

Duan, C.J., Xian, L., Zhao, G.C., Feng, Y., Pang, H., Bai, X.L., Tang, J.L., Ma, Q.S., and Feng, J.X. 2009. Isolation and partial characterization of novel genes encoding acidic cellulases from metagenomes of buffalo rumens. *J. Appl. Microbiol.* **107**, 245–256.

Elvira, C., Sampedro, L., Benítez, E., and Nogales, R. 1998. Vermicomposting of sludges from paper mill and dairy industries with *Eisenia andrei*: A pilot-scale study. *Bioresour. Technol.* **63**, 205–211.

Feng, Y., Duan, C.J., Pang, H., Mo, X.C., Wu, C.F., Yu, Y., Hu, Y.L., Wei, J., Tang, J.L., and Feng, J.X. 2007. Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases. *Appl. Microbiol. Biotechnol.* **75**, 319–328.

Fontes, C.M., Clarke, J.H., Hazlewood, G.P., Fernandes, T.H., Gilbert, H.J., and Ferreira, L.M.A. 1997. Possible roles for a non-modular, thermostable and proteinase-resistant cellulase from the mesophilic aerobic soil bacterium *Cellvibrio mixtus*. *Appl. Microbiol. Biotechnol.* **48**, 473–479.

Fujinami, S. and Fujisawa, M. 2010. Industrial applications of alkaliphiles and their enzymes-past, present and future. *Environ. Technol.* **31**, 845–856.

Kim, D., Kim, S.N., Baik, K.S., Park, S.C., Lim, C.H., Kim, J.O., Shin, T.S., Oh, M.J., and Seong, C.N. 2011. Screening and characterization of a cellulase gene from the gut microflora of abalone using metagenomic library. *J. Microbiol.* **49**, 141–145.

Kim, S.J., Lee, C.M., Han, B.R., Kim, M.Y., Yeo, Y.S., Yoon, S.H., Koo, B.S., and Jun, H.K. 2008. Characterization of a gene encoding cellulase from uncultured soil bacteria. *FEMS Microbiol. Lett.* **282**, 44–51.

Kuhad, R.C., Gupta, R., and Singh, A. 2011. Microbial cellulases and their industrial applications. *Enzyme Res.* doi:10.4061/2011/280696.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., and *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: a program to check the stereo chemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.

Lednicka, D., Mergaert, J., Cnockaert, M.C., and Swings, J. 2000. Isolation and identification of cellulolytic bacteria involved in the degradation of natural cellulosic fibers. *Syst. Appl. Microbiol.* **23**, 292–299.

Li, D.C., Li, A.N., and Papageorgiou, A.C. 2011. Cellulases from thermophilic fungi: Recent insights and biotechnological potential. *Enzyme Res.* doi:10.4061/2011/308730.

Lynd, L.R., Laser, M.S., Bransby, D., Dale, B.E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J.D., Sheehan, J., and *et al.* 2008. How biotech can transform biofuels. *Nat. Biotechnol.* **26**, 169–172.

Lynd, L.R., Weimer, P.J., van Zyl, W.H., and Pretorius, I.S. 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**, 506–577.

Mawadza, C., Hatti-Kaul, R., Zvauya, R., and Mattiasson, B. 2000. Purification and characterization of cellulases produced by two *Bacillus* strains. *J. Biotechnol.* **83**, 177–187.

Miller, G.L., Blum, R., Glennon, W.E., and Burton, A.L. 1960. Measurement of carboxymethylcellulase activity. *Anal. Biochem.* **2**, 127–132.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. 2004. UCSF Chimera-A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., and *et al.* 2000. Cloning the soil metagenome: A strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* **66**, 2541–2547.

Sali, A. and Blundell, T.L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815.

Santos, C.R., Paiva, J.H., Sforca, M.L., Neves, J.L., Navarro, R.Z., Cota, J., Akao, P.K., Hoffmam, Z.B., Meza, A.N., Smetana, J.H., and *et al.* 2012. Dissecting structure-function-stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168. *Biochem. J.* **441**, 95–104.

Singh, R., Sharma, R.R., Kumar, S., Gupta, R.K., and Patil, R.T. 2008. Vermicompost substitution influences growth, physiological disorders, fruit yield and quality of strawberry (*Fragaria* x *ananassa* Duch.). *Bioresour. Technol.* **99**, 8507–8511.

Sukumaran, R.K., Singhania, R.R., and Pandey, A. 2005. Microbial cellulases - Production, applications and challenges. *J. Sci. Ind. Res.* **64**, 832–844.

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

Varrot, A., Frandsen, T.P., Driguez, H., and Davies, G.J. 2002. Structure of the *Humicola insolens* cellobiohydrolase Cel6A D416A mutant in complex with a non-hydrolysable substrate analogue, methyl cellobiosyl-4-thio-beta-cellobioside, at 1.9 a. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 2201–2204.

Voget, S., Steele, H.L., and Streit, W.R. 2006. Characterization of a metagenome-derived halotolerant cellulase. *J. Biotechnol.* **126**, 26–36.

Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., Cayouette, M., McHardy, A.C., Djordjevic, G., Aboushadi, N., and *et al.* 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**, 560–565.

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

Yasir, M., Aslam, Z., Kim, S.W., Lee, S.W., Jeon, C.O., and Chung, Y.R. 2009. Bacterial community composition and chitinase gene diversity of vermicompost with antifungal activity. *Bioresour. Technol.* **100**, 4396–4403.

Zhou, J., Bruns, M.A., and Tiedje, J.M. 1996. DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* **62**, 316–322.