Journal of Bioscience and Bioengineering
VOL. 119 No. 4, 384–391, 2015

ELSEVIER

www.elsevier.com/locate/jbiosc

Identification of glycosyl hydrolases from a metagenomic library of microflora in sugarcane bagasse collection site and their cooperative action on cellulose degradation

Pattanop Kanokratana,[1,2] Lily Eurwilaichitr,[1] Kusol Pootanakit,[2] and Verawat Champreda[1,*]

Enzyme Technology Laboratory, Bioresources Technology Unit, National Center for Genetic Engineering and Biotechnology (BIOTEC), 113 Thailand Science Park, Phahonyothin Road, Klong Luang, PathumThani 12120, Thailand[1] and Institute of Molecular Biosciences, Mahidol University, Salaya, Nakornpathom 73170, Thailand[2]

Received 19 May 2014; accepted 13 September 2014
Available online 30 October 2014

Lignocellulose decomposition is a natural process involving the cooperative action of various glycosyl hydrolases (GH) on plant cell wall components. In this study, a metagenomic library was constructed to capture the genetic diversity of microbes inhabiting an industrial bagasse collection site. A variety of putative genes encoding GH families 2, 3, 5, 9, 11, and 16 were identified using activity-based screening, which showed low to moderate homology to various cellulases and hemicellulases. The recombinant GH9 endoglucanase (Cel9) and GH11 endo-xylanase (Xyn11) were thermophilic with optimal activity between 75°C and 80°C and the maximal activity at slightly acidic to neutral pH range. The enzymes exhibited cooperative activity with *Trichoderma reesei* cellulase on the degradation of lignocellulosic substrates. Mixture design showed positive interactions among the enzyme components. The optimal combination was determined to be 41.4% Celluclast, 18.0% Cel9, and 40.6% Xyn11 with the predicted relative reducing sugar of 658% when compared to Celluclast alone on hydrolysis of alkaline-pretreated bagasse. The work demonstrates the potential of lignocellulolytic enzymes from a novel uncultured microbial resource for enhancing efficiency of biomass-degrading enzyme systems for bio-industries.

[**Key words:** Glycosyl hydrolases; Lignocellulose; Metagenome; Mixture design; Sugarcane bagasse]

In the global biogeochemical organic carbon cycle, decomposition of lignocellulosic plant biomass by microbes is an essential process. The complex physical structure and chemical composition of lignocellulose confers high physicochemical stability; hence, lignocellulose decomposition involves cooperative actions of a diverse range of microbes producing various lignocellulolytic enzymes targeting cellulosic and non-cellulosic plant cell wall components (1). Lignocellulose has received increasing attention as a promising alternative renewable feedstock to depleting fossil resources, and it provides the basis of the biorefinery industry in which it is converted to fuels and value-added chemicals. Enzymatic hydrolysis of plant biomass to sugar is the key step in the sugar platform biorefinery. Exploration of microbial lignocellulose degradation mechanisms in nature and discovery of efficient lignocellulolytic enzyme systems are thus a key issue in current biotechnology research with potential for application in biomass industry.

Synergistic or cooperative actions among lignocellulolytic enzymes have been reported for many cellulolytic and hemicellulolytic enzyme systems for both aerobic and anaerobic bacteria (2–5). The synergy of enzyme catalysis on hydrolysis of lignocelluloses is based on different mechanistic models of the enzymes on the substrates; for example, differences in their substrate specificities and their mode of actions among a variety of glycosyl hydrolases and auxiliary enzymes (6), proximity effects of enzyme components in cellulosomal systems (2), and physical alteration of the substrates by non-catalytic proteins, e.g., expansins (7). These enzyme cooperative interactions form the basis of efficient lignocellulose degradation in nature and provide a basis for development of efficient lignocellulolytic enzyme systems for biotechnological application.

Culture-independent metagenomic technology has been used to explore novel genes and metabolic pathways from uncultured microbes, which can make up to 99% of total microbial diversity in environments. Different techniques, including functional-based and sequence-based gene identification and direct high-throughput pyrosequencing have been used to explore environmental metagenomes and reveal complex structures of microbial assemblages in various ecosystems related to plant biomass decomposition (8,9). These approaches have been used to identify a number of genes encoding various novel lignocellulose degrading enzymes from termite gut (10,11), cow rumen (12), compost (13), and symbiotic microbial consortia (14). A number of them show desirable properties for industrial applications such as alkaliphilic xylanases for pulp bio-bleaching (10).

Industrial bagasse collection sites at sugar mills represent interesting habitats to explore lignocellulose decomposition because of their relatively high temperature, low nitrogen availability, and indigenous microbes. Recently, microbial communities occupying different regions of a bagasse pile have been explored for

* Corresponding author. Tel.: +66 2564 6700x3473; fax: +66 2564 6707.
E-mail address: verawat@biotec.or.th (V. Champreda).

1389-1723/$ – see front matter
http://dx.doi.org/10.1016/j.jbiosc.2014.09.010

their metabolic potential using high-throughput pyrosequencing approaches, which revealed complex microbial communities with different niches varying according to microenvironmental conditions (15). In particular, substantial numbers of lignocellulolytic enzymes are found in the inner pile based regions suggesting that this unique environment provides a rich potential resource of biomass degrading microbes and enzymes.

In this study, a fosmid metagenomic library was constructed to represent the diversity of microbes residing in a bagasse pile. Activity-based screening was used to identify genes encoding various glycosyl hydrolases attacking lignocellulosic components. A recombinant thermophilic endo-glucanase and a xylanase from the metagenome were characterized and their cooperative action on enhancing efficiency of *Trichoderma reesei* cellulase was demonstrated using a mixture design approach. The work shows potential of enzymes from unexplored microbial resources and provides the basis for development of efficient biomass-degrading enzyme systems for biotechnological application.

MATERIALS AND METHODS

Sample collection Sugarcane bagasse samples were collected from an industrial bagasse collection site at Phu Khieo Bio-Energy, Chaiyapoom province, Thailand (N 16°28′54″, W 102°07′05″). Bagasse was collected from a large open-air pile, approximately 10 m in height covering the area of several acres. The bagasse had been left in the field for approximately 6 months at the time of collection. A bagasse sample was taken at a depth of 1 m from the base of bagasse pile in June 2009. The temperatures of the samples inside the pile at the collection time ranged from 49°C to 52°C. The sample was rapidly frozen in liquid nitrogen and kept at −80°C for subsequent experiments.

Lignocellulosic biomass preparation Intact bagasse sample collected from the same site was physically processed using a SM2000 cutting mill (Retsch, Haan, Germany) and sieved to <0.2 mm^2. The biomass was pretreated using 5% (w/v) NaOH (with a liquid/solid ratio of 3:1) and at 80°C for 90 min. The pretreated solid was washed with distilled water until its pH decreased to 7 and it was then dried at 60°C before use as a substrate for subsequent enzymatic hydrolysis experiments.

Metagenomic DNA extraction and purification Metagenomic DNA was extracted directly from 5 g of sample by the SDS-based DNA extraction procedure (16), with slight modifications (17). High molecular weight DNA with size ranging between 30 and 50 kb was selected and purified by pulsed-field gel electrophoresis and electroelution. Pulsed-field gel electrophoresis was performed using a CHEF DRIII system (Bio-Rad, Hercules, CA, USA) in 0.5× TBE at 14°C, using a 0.1–14.0 s switch time at 6 V/cm for 12 h. DNA was electroeluted in a dialysis bag (Spectrapor 4, Spectrum Laboratories, Rancho Dominguez, CA, USA) using a field strength of 70 V at 4°C for 2 h. The purified DNA solution was collected and subsequently concentrated using an Amicon Ultra filter unit (Millipore, Billerica, MA, USA).

Fosmid library construction A metagenomic fosmid library from bagasse sample was constructed using a CopyControl fosmid library production kit (Epicentre Biotechnologies, Madison, WI, USA) according to the manufacturer's instruction, with slight modifications. The purified DNA was end-repaired to generate blunt 5′-phosphorylated ends and then ligated to pCC1FOS vector at 25°C for 3 h. The ligated DNA was packaged using the lambda packaging extract supplied with the kit and subsequently transformed into *Escherichia coli* EPI300-T1R. The transformants were selected on LB agar plates supplemented with 12.5 μg/ml of chloramphenicol. The library was stored at −80°C in 15% glycerol as pooled colonies.

Fosmid library screening The fosmid library was screened for genes encoding lignocellulolytic enzymes by the activity-based method using plates overlaid with insoluble chromogenic AZCL-linked substrates (Megazyme, Wicklow, Ireland). The fosmid library was diluted in LB broth before spreading on LB plates supplemented with 12.5 μg/ml chloramphenicol and 0.001% (w/v) arabinose. After incubation overnight at 37°C, colonies on the plates were overlaid with 0.7% agar containing 0.05% of AZCL-xylan, AZCL-HE-cellulose, AZCL-β-glucan, or AZCL-xyloglucan in ENZhance cell permeabilizing reagent pH 7.0 (Biotec, Thailand) and incubated at 37°C from 1 to 4 h. Positive colonies producing enzymes that can degrade the substrates were detected based on blue zone resulting from degradation of the substrates. Insert fragments from positive clones were analyzed by either end-sequencing using forward sequencing primer (FP: 5′- GGATGTGCTGCAAGGCGATTAAGTTGG -3′) and reverse sequencing primer (RP: 5′- CTCGTATGTTGTGTGGAATTGTGAGC -3′) flanking the fosmid cloning site, or restriction fragment length polymorphism (RFLP) analysis using *Not*I (Fermentas, Vilnius, Lithuania) digestion.

Gene identification Recombinant fosmids from positive colonies were extracted with a GeneJET Plasmid Miniprep kit (Fermentas). The fosmid DNAs were then partially digested with *Bsp*143I (Fermentas) and fragments larger than 1 kb were purified using a Wizard DNA Clean-Up System (Promega, Madison, WI, USA). The purified DNA fragments were ligated into pZErO-2 vector that had been digested with *Bam*HI at 16°C for 1 h. The ligated DNA was then transformed into *E. coli* TOP-10 (Invitrogen, Carlsbad, CA, USA) by electroporation. The fosmid sub-cloned plasmid libraries were cultured overnight on LB plates containing 50 μg/ml kanamycin and colonies carrying the target gene were screened again using the overlaying technique with the respective AZCL-substrates as described above. Plasmids were extracted from the positive colonies using a GeneJET Plasmid Miniprep kit (Fermentas) and inserts were sequenced at Macrogen (Seoul, Korea). Homology search of the obtained sequences was performed using BLASTX against the NCBI database. Full-length gene sequences were obtained by gene walking with internal primers designed from the partial gene sequences and recombinant fosmid as a template. The gene sequences are deposited in Genbank with accession numbers AGN70389–AGN70395.

Recombinant enzyme expression and purification Mature genes excluding sequences likely encoding signal peptide according to cleavage site prediction by the Signal P server (18) were amplified from the fosmid vectors using *Pfu* DNA polymerase (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's protocol using *cel*9F (5′- *ATCA*CCATGGGCGCCCCCGGTAACTACAAC -3′) and *cel*9R (5′- *ATCA*CTCGAGGACGATCGTACAGGCAG -3′) for the *cel*9 gene and xyn11F (5′- *ATCA*CCATGGCACAGCAGTGCATCACCTCG -3′) and xyn11R (5′- *ATCA*CTCGAGCGGCGTAT TACCGTAGCC -3′) for the *xyn*11 gene (Restriction sites for cloning are underlined). The PCR conditions were: pre-denaturation at 94°C for 1 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 7 min, and a prolonged final extension at 72°C for 10 min. The amplicons were purified from agarose gel using a GeneJet Gel Purification kit (Thermo Scientific). The amplified genes were digested with *Nco*I and *Xho*I. The purified gene fragments were ligated with pET28a(+) or pET32a(+) (Millipore) which was pre-digested with the same enzymes for expression of the encoded proteins in C-terminal His_6-tagged fusion or N-terminal Trx/C-terminal His_6 tagged forms, respectively. Ligation reactions were transformed into *E. coli* DH5α. The recombinant plasmids were then isolated and transformed into *E. coli* Rosetta and *E. coli* Rosetta-gami (Millipore) for protein expression.

A starter culture of recombinant *E. coli* harboring the target gene was grown in LB supplemented with appropriate antibiotics (34 μg/ml of chloramphenicol and 50 μg/ml kanamycin for pET28-*cel*9, and 34 μg/ml of chloramphenicol and 50 μg/ml ampicillin for pET32-*xyn*11) at 37°C for overnight. The culture was then inoculated into the fresh LB containing appropriate antibiotics (1% (v/v) inoculum) and induced with 1 mM isopropyl-thio-β-D-galactopyranoside (IPTG) at 30°C or 37°C for 4 h. The induced cells were then harvested by centrifugation. The cell pellet was resuspended in 100 mM sodium phosphate buffer, pH 7.4 and disrupted by sonication. The cell lysate was clarified by centrifugation at 12,000 ×*g* and the supernatant was recovered. The soluble protein was purified using a HisTrap HP column (GE Healthcare, Buckinghamshire, UK). The eluted protein fractions were analyzed by 12% SDS-PAGE electrophoresis and fractions with the purified protein were buffer-exchanged into 20 mM sodium phosphate buffer pH 7.4 and concentrated using a Vivaspin ultrafiltration spin column MWCO 10 kDa (Millipore).

Enzyme activity assays Polysaccharide-degrading activities were determined using the 3,5-dinitrosalicylic acid (DNS) method by measuring the amount of reducing sugars liberated (19). Reactions of 0.32 ml were performed in 0.05 M phosphate buffers, pH 6.0 containing barley β-glucan (0.5% w/v), carboxymethylcellulose (CMC, 1% w/v), Avicel (1% w/v), or beechwood xylan (1% w/v) as the substrates with an appropriate enzyme concentration, for determining β-glucanase, endoxylanase, Avicelase, and endo-xylanase activities, respectively. The reactions were incubated for 10 min at 75°C and 80°C for Cel9 and Xyn11, respectively, before the addition of 0.68 ml of DNS solution. The reactions were stopped by boiling for 10 min. The amount of reducing sugars was determined by measuring the absorbance at 540 nm using a VICTOR^3 V microplate reader (Perkin Elmer, Waltham, MA, USA) and interpolation from a standard curve prepared using the corresponding sugar standards. One unit of enzyme activity is defined as the amount of enzyme required to release 1 μmol of reducing sugars from a substrate in 1 min under the assay condition. Effects of pH were studied in McIlvaine's buffers (20) for pH 3–8 and Tris–HCl for pH 8–10.

Cooperative enzyme interaction analysis Interaction between the recombinant enzymes and the *T. reesei* cellulase (Celluclast 1.5 L, Novozymes, Bagsvaerd, Denmark) commercial enzyme was tested using an experimental mixture design approach. A {3,3}-augmented simplex lattice design was examined with the Minitab 16.0 software (Minitab Inc., State College, PA, USA) to define an optimal enzyme formulation. The design consisted of 13 experimental points with three components and a lattice degree of 3. The three independent variables in the mixture design were Celluclast (X1), Cel9 (X2) and Xyn11 (X3). The sum of all the components in the mixture was 100% based on volume. The relative reducing sugar yield (Y1) was applied as a dependent variable for the analysis and simulation of the respondent model. The percent relative reducing sugar yield (Y1) is the reducing sugar produced by each combination divided by the reducing sugar produced by 100% Celluclast. After regression analysis, the full cubic model was used to simulate the optimized ratio of the mixture components. The assay reactions (120 μl) were performed based on the microplate assay reaction (21). The reactions contained 2.5 mg of alkaline