Journal of Bioscience and Bioengineering
VOL. 109 No. 4, 331 – 336, 2010

ELSEVIER

www.elsevier.com/locate/jbiosc

Enhancement of the thermostability and activity of mesophilic *Clostridium cellulovorans* EngD by *in vitro* DNA recombination with *Clostridium thermocellum* CelE

Chae Yoeng Lee,[1] Kyung Ok Yu,[1] Seung Wook Kim,[2] and Sung Ok Han[1,*]

School of Life Science and Biotechnology, Korea University, Seoul 136-713, Republic of Korea[1] *and Department of Chemical Biological Engineering, Korea University, Seoul, Korea*[2]

Received 27 July 2009; accepted 15 October 2009
Available online 6 November 2009

The thermal stability and catalytic activity of endoglucanase (EngD) from mesophilic *Clostridium cellulovorans* were improved by evolutionary molecular engineering. Thermostable mutants were isolated after staggered extension process (StEP) with *celE* from thermophilic *Clostridium thermocellum* performed to conduct family shuffling and overlay screening of the resultant mutant library. The relative activity of the best-evolved clone has been improved of about 2 times higher at 50 °C and showed a higher k_{cat}/K_m value than its *engD* parental clone. We determined that these variants had two amino acid substitutions (L157N, Q158E) and confirmed their effects by substituting these amino acids in the parental gene by site-directed mutagenesis. These substitutions resulted in an increase in hydrophilic or charged residues. Our results demonstrate that *in vitro* recombination is an effective approach to improve the thermostability and enzymatic activity of a mesophilic enzyme.

[**Key words:** *Clostridium cellulovorans*; *Clostridium thermocellum*; Cellulase; Thermostability; *in vitro* DNA recombination]

Over the last few decades, there has been increasing interest in the bioconversion of cellulosic biomass into bioethanol as a renewable energy source (1). Since plant polysaccharides are the most abundant renewable biomass, cellulolytic microorganisms play a major role in carbon turnover in nature (2). Cellulases are broadly classified into three groups based on amino acid sequence and substrate specificity: endoglucanase, exoglucanase, and β-glucosidase (3). Endo-β-1,4-glucanases randomly cleave internal β-1,4-linkages in cellulose polymers and completely convert cellulose to glucose by synergistic action with the other types of cellulases, cellobiohydrolase and β-glucosidase (4).

Enzymes are effective and specific biocatalysts widely used in industry, but their application in industrial processes often requires desirable functions not found in natural source enzymes. In order to obtain the desired enzymes, scientists have developed rational and non-rational design methods to enhance enzyme properties. Many cellulases, including stable enzymes, have been found and used for cellulose degradation. Degradation of cellulose at increased temperatures has many benefits, such as elevated cellulase activity, less energy cost for cooling, and diminished risk of contamination (5). In other words, thermostable enzymes are of considerable biotechnological interest since their improved stability could greatly reduce enzyme replacement costs or permit processes to be carried out at high temperature (6).

Directed evolution has been employed to improve the properties of enzymes and has shown its capacity for protein engineering. *In vitro* recombination methods generally offer higher recombination capability and greater experimental adaptability than *in vivo* approaches (7). For the past two decades, many methodologies describing protein thermostability improvement through the introduction of mutations have been used, including the comparison of naturally occurring homologous proteins, rational design, and directed evolution (8, 9, 10, 11).

The StEP (staggered extension process) method uses homologous genes as templates for the synthesis of chimeric genes. It consists of priming denatured templates followed by repeated cycles of denaturation and short annealing/extension steps. Unique events occur when the randomly extended primers anneal to recombinogenic templates via template switching events based on sequence complementarity and further extension (12).

In our study, we selected two cellulases as parental genes. One was *engD* from the mesophilic anaerobic bacteria *Clostridium cellulovorans* (13), which is a noncellulosomal cellulase that does not possess a dockerin domain (14). The other parental gene was *celE* from the thermophilic anaerobic bacteria *Clostridium thermocellum* (15). This gene encodes a cellulosomal cellulase that contains a dockerin domain (16). Due to their high sequence homology, these genes can participate in DNA recombination, which we conducted. Because of

* Corresponding author. Tel.: +82 2 3290 3151; fax: +82 2 927 9028.
E-mail address: samhan@korea.ac.kr (S.O. Han).

1389-1723/$ - see front matter © 2009, The Society for Biotechnology, Japan. All rights reserved.
doi:10.1016/j.jbiosc.2009.10.014

the thermal stability of CelE, we hoped to use family shuffling to enhance the enzyme stability of EngD.

In this study, we report the use of directed evolution to improve the thermostability of an endoglucanase from *C. cellulovorans* by *in vitro* recombination with an endoglucanase from *C. thermocellum*.

MATERIALS AND METHODS

Construction The *engD* (GenBank: M37434.1, ATCC 35296) gene was amplified by PCR using the forward primer (5′-GGA TCC GTC TAC TGC TTT TAC AGG TGT ACG T-3′) and reverse primer (5′-CTC GAG TTT TAC TGT GCA TTC AGT ACC AT-3′). The *celE* (GenBank: M22759.1, ATCC 27405D) gene was amplified by PCR using the forward primer (5′-GGA TCC GTC GGG AAC AAA GCT TTT G-3′) and the reverse primer (5′-CTC GAG TAT TGA GCG CAG TAG ATA TTT TTT-3′). The PCR program was run for 30 cycles under the following condition: 30 s at 94 °C, 30 s at 68 °C and 1 min at 72 °C with the exception of the first denaturation step and final extension step were carried out 5 min at 94 °C and 7 min at 72 °C. Each 1.4 kb PCR product was purified from an agarose gel (QIAGEN, Valencia, CA) using a gel extraction kit and then digested with *Bam*H and *Xho* (Takara, Shiga, Japan). Digested DNA fragments were inserted into the same restriction sites of pET25b (Novagen, Madison, WI). The resulting plasmids were designated pET25b-*engD* and pET25b-*celE*, respectively. These were used as the starting DNA sequences for the random recombination of the two genes.

DNA recombination and overlap PCR The *in vitro* recombination between the 1 kb PCR product of the *engD* and *celE* catalytic sites was done via the StEP. Aliquots of 125 ng of amplified *engD* and *celE* genes were mixed in 50 μL of PCR mixture containing 5 μL of 10× Taq buffer, 2 μL of dNTP mix, 10 pmol of each primer (5′-GGA TCC GTC GGG AAC AAA GCT TTT G-3′, 5′-CAA AAG TCT CTG CTT TTC CAG CAA CA-3′), distilled H_2O, and 2.5 U Taq DNA polymerase (Takara, Shiga, Japan) in a total volume of 50 μL (PCR program: 5 min at 95 °C, 80 cycles of 30 s at 94 °C, and 5 s at 55 °C) (17).

To prepare the fusion gene (mutant-CBD) encoding the catalytic domain sequence of the mutant and cellulose-binding domain (CBD) sequences of *engD*, we used the technique of splicing by overlap extension PCR (18). The StEP product that corresponded to the linker region of CBD was amplified. The CBD fragment was amplified using primers (5′-TGG AAA AGC AGA GAC TTT TGG TAT AT-3′, 5′-CTC GAG TTT TAC TGT GCA TTC AGT ACC AT-3′) that covered the C-terminal sequence of the StEP product. The PCR products were separated by electrophoresis on 0.8% agarose gel and stained with ethidium bromide. Fragments were purified using a gel extraction kit (QIAGEN, Valencia, CA) and fused together in a second PCR step (PCR program: 3 min at 94 °C, 35 cycles of 1 min at 94 °C, 2 min at 50 °C, 1 min at 72 °C, and finally 5 min at 72 °C). An equal molar concentration of each fragment was added to the PCR mixture (19). Overlap PCR product was purified from an agarose gel using a gel extraction kit (QIAGEN, Valencia, CA) and then digested with *Bam*H and *Xho* . About 1.4 kb digested DNA fragments were inserted into the restriction sites of pET25b (20).

Screening *E. coli* JM109 (DE3) (Promega, Madison, WI) cells containing recombination genes were plated on LB agar (Difco Laboratories, Detroit, MI) plates, and the cells harboring wild-type *engD* and *celE* were used as controls. After 15-h incubation at 37 °C, cells were transferred on fresh LB agar plates containing ampicillin by either nylon membranes or sterilized picker. The plates were then covered with 5 mL of top medium consisting of 0.5% carboxymethyl cellulose (CMC) (Sigma, St Louis, MO), 100 μg mL^{-1} ampicillin and 0.8% agar and incubated at 65 °C for 6 h. Then 0.1% Congo Red (Sigma, St Louis, MO) was added and plates were incubated for 20 min. At the end of the staining, plates were washed with 1 M NaCl. Positive colonies with visualized hydrolysis zones under thermal conditions were picked and applied to the thermostability test (21).

Site-directed mutagenesis Oligonucleotide-mediated mutagenesis of the *engD* gene was performed by a PCR-based procedure using pET25b-*engD* as the template. The mutagenic primers (5′-GAG CCA AGA CCA GTA GGC GCA AGC *AAT* CAG TGG ACA GGC GGT TCT TAT G-3′ and 5′-CAT AAG AAC CGC CTG TCC ACT G*AT T*GC TTG CGC CTA CTG GTC TTG GCT C-3′), containing the leucine (L) to asparagine (N) mutation (italics), were used in the amplification of the NQ (substitution L157N) mutant gene. The LE (substitution Q158E) mutant containing the glutamine (Q) to glutamic acid (E) mutation was introduced into the *engD* gene by means of the oligonucleotide primers (5′-GAG CCA AGA CCA GTA GGC AAG CTT A*GA G*TG GAC AGG CGG TTC TTA TG-3′ and 5′-CAT AAG AAC CGC CTG TCC A*CT C*TA AGC TTG CGC C-3′). The resulting clones were confirmed by DNA sequencing.

Enzyme production and purification For purification of native *engD* and evolved thermostable enzymes (mutant NE, site-directed mutation product NQ and LE), a 6xHis affinity tag part of the vector was fused to the C terminus of each enzyme. Each *E. coli* JM109 (DE3) strain harboring recombinant wild-type and thermostable mutants was grown at 37 °C in LB-amp-IPTG medium (100 μg mL^{-1} ampicillin, IPTG of final concentration 0.2 mM). Grown cells were treated by centrifugation and sonication. Crude cell extract in a solution containing 50 mM sodium phosphate (pH 8.0), 300 mM NaCl, and 10 mM imidazole was purified with nickel-nitrilotriacetic acid agarose resin (Qiagen, Hilden, Germany) and imidazole-containing buffer. After extensive washing, the bound protein was eluted with 250 mM imidazole in phosphate buffer. Fractions containing enzymatic activity were pooled and then concentrated by applying them to a centrifugal filtration device (Centricon Concentrators, Amicon Millipore). Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was used to determine the purity of the protein (6). To confirm protein activity, we conducted Zymogram (0.3% CMC sodium dodecyl sulfate–polyacrylamide gel electrophoresis and Congo Red staining methods).

Enzyme assay and determination of kinetic parameters Endoglucanase activity was estimated by measuring the reducing sugars released from CMC. The purified enzyme assay consisted of 50 μL 0.5% CMC in 50 mM sodium acetate buffer (pH 6.0) and 50 μL diluted enzyme solution. After incubation at 37 °C for 30 min, dinitrosalicylic acid reagent (100 μL) was added, and the mixture was heated in a boiling water bath for 5 min and 800 μL distilled H_2O was added (22). The absorbance was measured at 540 nm. One unit of enzyme activity was defined as the quantity of enzyme capable of releasing 1 μmol of glucose equivalent per min (with glucose as a standard). Kinetic parameters for enzyme activity were determined using substrate concentrations ranging from 0.125% to 2.0% CMC (pH 6.0) at 37 °C. For the half-life measurement of the enzymes, the enzyme solutions were incubated with 50 μg mL^{-1}, 50 mM sodium acetate buffer, pH 6.0, at 55 °C for various time periods, and the residual activity was determined with 0.5% substrate. And the half-life was calculated from the residual activity versus incubation times. Protein was measured by the Bradford method using bovine serum albumin as a standard (18).

Thermostability assay and ph profile The effect of pH and temperature on enzyme activity was evaluated by measuring activity at diverse temperatures (30–80 °C) and at different pH values using 0.2 M Na_2HPO_4–citric acid buffer (pH 4.0–5.0), 0.2 M Na_2HPO_4–NaH_2PO_4 buffer (pH 6.0–8.0), and 0.05 M glycine–NaOH buffer (pH 9.0–12.0). Thermal stability studies were carried out by preincubating the enzyme for 30 min in 50 mM sodium acetate buffer (pH 6.0) at various temperatures (30–80 °C). The residual enzyme activity was determined. For pH stability, relative activity was determined after the enzyme had been preincubated for 1 h with diverse pH buffers (1, 23).

3D structure modeling Structural models of parental and mutated enzymes were determined as high homologous template. The three-dimensional structure of EngD was modeled with the program PyMOL. The structure of CelCCA (PDB ID: 1EDG) was obtained from a *Clostridium cellulolyticum* template.

RESULTS

Screening the mutant library Approximately 8000 clones were screened for increased thermostability. At the initial screening using the overlay method, mutants showing a larger halo size than the wild-type clones were selected. Subsequently, the thermostability of selected mutants was assayed. By considering the enzyme activity and thermal stability, the best mutant was finally chosen for further purification and characterization. Sequence analysis of the selected mutant showed two amino acid substitutions.

Sequence analysis of the mutant The primary structure of several mutants was determined by DNA sequence analysis. The nucleotide sequence of the mutant endoglucanase was submitted to GenBank (accession no. GU060420). Sequenced hybrids containing mainly the *engD* amino acid sequence had been substituted with the *celE* amino acid sequence in the catalytic domain. The substitution pattern from EngD to the mutant and the amino acid composition of the mutant were determined. The results are shown in Fig. 1. There are two amino acid substitutions in the mutant proteins (eight base

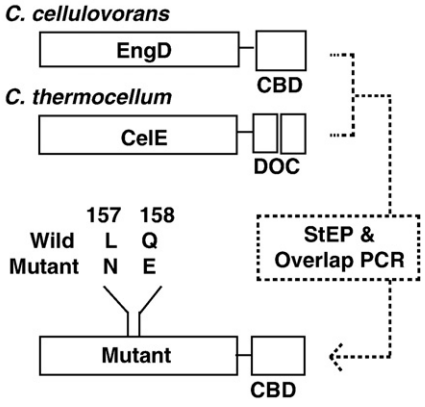

FIG. 1. Schematic image of library construction. CBD and DOC represent the cellulose-binding domain of EngD and the dockerin domain of CelE, respectively. The image shows two amino acid substitutions L157N (from leucine to asparagine) and Q158E (from glutamine to glutamic acid).