Heterologous Expression and Enzymatic Properties of Exocellulase CelA

Ming Sui¹, Jiachao Rong², Yang Zhang², Peng Qu³, Yusong Zhu³

¹Department of Wine and Food Engineering, Sichuan Technology & Business College, Dujiangyan, Sichuan Province, China
²Department of economics and management, Tongren Polytechnic College, Gui Zhou, Tongren, 554300, China
³Xinjiang Oilfield Company

Abstract. The gene of glycosyl hydrolase CelA is derived from Bacillus licheniformis and the GenBank accession number is AAU40776.1. The sequence analysis of this gene was carried out by online software, and the celA gene was cloned from the genome of the strain by homologous amplification. The gene was ligated into the pMAL-c2x vector, transformed into E. coli Transetta for heterologous expression, and the catalytic properties of the recombinant enzyme CelA were studied.

1. Introduction
In recent years, the intensification of the resource crisis has led to a sharp increase in the demand for bioenergy. Lignocellulose is the most abundant biomass resource, and its biotransformation research has attracted much attention. Cellulase, especially progressive exocellulase, is considered to be an important component of the biotransformase system and has an important influence on the rate of transformation. In this study, the GH48 family exocellulase CelA derived from Bacillus licheniformis was cloned and heterologously expressed, and the enzymatic properties and related functional residues of the enzyme were studied. The CBM domain was also explored. Effect on the catalytic properties of CelA.

2. Medium and buffer
(a) SOXTI}E buffer (50 times agarose gel running buffer): Weigh 242 g of Tris dissolved in 800 mL of deionized water, add 57.1 mL of acetic acid, stir and mix, and dilute to 1 L with deionized water. Adjust the pH of the buffer to 8.5 and store at room temperature.
(b) Agarose gel: Weigh agarose in a concentration of 0.8% to 1.5% (W/V), add it to 1*TAE buffer, dissolve it in a microwave oven, cool it to a certain temperature at room temperature, add the nucleic acid dye, mix and quickly pour Into the glue tank, solidified at room temperature.
(c) LB medium: 10 g of trypsin, Sg yeast extract and 10 g of NaCl were dissolved in 1 L of deionized water to adjust the pH to 7.0. 1.8% agar was added to the solid medium, and all liquid and solid media were sterilized in a sterilizer at 115 °C for 30 min.
(d) Ampicillin antibiotics: Acetomycin was weighed at a concentration of 100 mg/mL and dissolved in sterile water. Dissolved in 1.5 mL sterile EP tubes and stored at -20 °C.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd
(e). Isopropylthio-galactose (IPTG): IPTG was weighed in sterile water at a concentration of 1 mol/L, filtered and sterilized, and dispensed in 1.5 mL of sterile EP tube and stored at -20 °C.

Table 1. Formulation of citric acid-sodium citrate buffer

| Buffer pH | Citric acid (mL) | Sodium citrate (mL) |
|-----------|------------------|---------------------|
| 3.0       | 18.6             | 1.4                 |
| 3.0       | 13.1             | 6.9                 |
| 5.0       | 8.2              | 11.8                |
| 6.0       | 3.8              | 16.2                |

Table 2. Formulation of Tris-HCL Buffer

| Buffer pH | 0.1 mol/L Tris(mL) | 0.1 mol/L HCL(mL) | H2O (mL) |
|-----------|--------------------|-------------------|----------|
| 8.0       | 50                 | 29.2              | 21.8     |
| 9.0       | 50                 | 7.0               | 43.0     |

3. Sequence analysis and homology modeling

The gene sequence of exo-glucanase CeIA was obtained from NCBI database, and the amino acid sequence was analyzed by NCBI BLAST program. The domain of the enzyme and its family were predicted. The online software ProtParam was used to predict CeIA. The physicochemical properties and amino acid composition of the enzyme were analyzed; the signal peptide sequence of the enzyme was predicted using the online software SignalP-5.0 Server; the transmembrane region of the recombinant protein was predicted using the online software TMHMM Server v.2.0. Homology modeling was performed using SWISS-MODEL online protein structure prediction software.

The Bacillus lichen rmisATCC14580 strain was inoculated into Sml LB liquid medium, shaken overnight at 37 °C, the cells were collected by centrifugation, and the whole genome DNA of the strain was extracted using the bacterial whole genome DNA extraction kit, and stored in a refrigerator at -80 °C.

PCR amplification was carried out using the strain Bacillus lichen such as rmisATCC14580 whole genome DNA as template. The PCR reaction conditions are shown in Table 3:

Table 3. Reaction conditions

| T°C  | Reaction time | Reaction times |
|------|---------------|----------------|
| 94°C | 5 min         | 1 time         |
| 94°C | 30 s          | 30 times       |
| 56°C | 30 s          | 30 times       |
| 72°C | 2 min 30 s    | 30 times       |
| 72°C | 10 min        | 1 time         |

The PCR product was detected by agarose gel electrophoresis, and the target band was recovered using a gel recovery kit.

4. Expression and purification of recombinant enzyme

The sequence-verified recombinant plasmid was transformed into the expression host E. cohTransetta (DE3) for heterologous expression according to the method of 2.3.2.4.

(a). Randomly pick a single colony to prepare the seed solution, inoculate the seed solution in a 1 / 40 ratio into an Erlenmeyer flask containing 200 ml of LB liquid medium, shake culture at 37 °C, 190 rpm, and add IPTG at a final concentration of 1 mM. The culture temperature was adjusted to 30 °C, the shaking speed was adjusted to about 165 rpm, and the culture was continued for about 15 hours.
(b). The fermentation broth obtained in the previous step was centrifuged at 4 °C in a 10,000 rpm sub-supersonic cryogenic centrifuge for 20 min. The supernatant was decanted, and the precipitate was washed by adding Tris-HCL buffer, and then centrifuged to collect a precipitate.

(c). 8 ml of pMAL column buffer was added to the precipitate, suspended and mixed, placed in an ice bath, and the cells were disrupted by an ultrasonic cell pulverizer.

(d). The liquid to be crushed was cleared, and it was dispensed into a 2 mL EP tube, and centrifuged at 4 °C, 10000 rpm high-speed low-temperature centrifuge for 20 min, and the supernatant and the precipitate were separated and placed in a refrigerator at 4 °C for use.

Purification by amylose resin affinity chromatography: 2 mL of Amylose Resin injection column stored in 20% ethanol, and washed with sterile water to add pMAL Column Buffe: balance column material; Add the crushed supernatant, make the enzyme solution fully contact with Amylose Resin, mix at low temperature for 2h, and gently flip it during the period; fix the column, let Amylose Resin settle to the bottom naturally, install the sieve plate and add the elution buffer sequentially. Elute and collect. Each part of the eluate was stored at -20 °C for use.

5. Determination of protein concentration
The BCA working kit was used to determine the protein concentration. The bovine serum albumin (BSA) was used as a standard. The 20 ul system was added with 200 ul of BCA working solution, placed at 37 °C for 35 min, and the absorbance at 570 nm was measured by a microplate reader to map the protein. standard curve line. After the sample to be tested is diluted to a suitable concentration, the same operation is performed, and the protein concentration of the sample is calculated from the standard curve and the dilution.

Take RAC, methyl cellulose, cotton wool, microcrystalline cellulose, corn stover, cellulose powder CF11 as substrates, add appropriate amount of enzyme, react at 50 °C for 30 min, use DNS method to measure the yield of reducing sugar, calculate reorganization The enzymatic activity of the enzyme on different substrates to characterize the substrate specificity of the enzyme.

The pure enzyme components were separately added to each pH buffer and allowed to stand at 4 °C for 24 h. The enzymatic solution and the RAC substrate were thoroughly mixed under optimal conditions for enzymatic reaction, and the cellulase activity was determined according to the DNS method. The highest enzyme activity was 100, and the relationship between the relative enzyme activity and the incubation pH was plotted.

Add methanol, ethanol, isopropanol, glycerol, Triton X 100 at a final concentration of 1% and 10%, respectively, to the enzyme reaction system, and add EDTA, PMSF, and urea at a final concentration of 1 mM and 10 mM, respectively. After the reaction under the reaction conditions, the cellulase activity was measured according to the DNS method, and the relative enzyme activity of the enzyme in the reaction system was calculated by using the enzyme activity of 100% without any reagent.

The maltose-removing recombinase is mixed with 5 mg/mL of cellobiase, cellotetraose and cellobiose in a certain ratio, and reacted under the optimum reaction conditions for a period of time, and the boiling water bath is used for 10 min to terminate the reaction and centrifuge. The supernatant was taken and the product analysis was carried out by the TLC method.

6. Sequence analysis and homology modeling
According to the NCBI database, the sequence of the CeIA gene derived from Bacillus lichen rnis ATCC14580 is 2115 bp long and encodes a protein sequence of 704 amino acids. The protein has only one functional domain belonging to the GH48 family. The amino acid sequence of CeIA was analyzed by the online software ProParam, and the molecular weight of the protein was predicted to be 79463.69, and the theoretical pI was 4.92. This indicates that the protein is structurally stable. The half-life in E. coli is estimated to be >10 hours and greater than 20 hours in yeast. The total number of negatively charged residues of the amino acid of the enzyme was 82, and the total number of positively charged residues was 51. The aliphatic index was 57.19 and the average hydrophilicity was -0.670. The enzyme signal surface was predicted by the online software SignalP 4.1 Server. The
The predicted result is shown in Figure 1. CeIA has no obvious signal peptide sequence and is a non-secreted protein.

![Signal peptide prediction of CeIA](image1)

**Figure 1.** Signal peptide prediction of CeIA

The transmembrane region of the enzyme was analyzed using the online software TMHMMv2.0. The predicted results are shown in Figure 2. CeIA is an extramembrane protein and there is no transmembrane region.

![CeIA transmembrane channel protein prediction results](image2)

**Figure 2.** CeIA transmembrane channel protein prediction results

The structure determines the function, and the three-dimensional structure of the protein helps to understand the function of the protein. With the development of bioinformatics technology and the analysis of more and more protein structures, one can use the 3D model of protein homology with high homology as a template to model the structure of unknown enzyme molecules to understand its structure feature. This paper uses the online software SWISS-MODEL to model the homology of CeIA. The template is the three-dimensional structure of GH48 derived from Bacillus pumilus. The sequence similarity is 73.43%, the GMQE score is 0.94, and the QMEAN score is 0.23. The structure of CeIA is shown in Figure 3. It has the same barrel folding as other enzymes of the GH48 family, which is a typical structural feature of the GH48 family of enzymes. At the same time, the molded structure also showed that there is a long closed substrate at the N-terminus of the enzyme molecule in combination with the tunnel, and the tunnel is followed by an open crack, which may be the outlet of the hydrolyzate.

![Structure of CeIA](image3)
7. Conclusion
In this paper, the exo-cellulase derived from Bacillus licheniformis ATCC14580 was cloned and heterologously expressed, and the properties of the recombinant enzyme were studied. The results obtained were: CelA enzyme gene was 2115 bp, encoding one containing 704 amino acids. A sequence of proteins that have only one functional domain that is part of the GH48 family. Homology modeling shows that CelA has barrel-like folding like other enzymes of the GH48 family, which is a typical structural feature of the GH48 family of enzymes. The celA gene was cloned by homologous amplification and ligated into pMAL-c2X vector, heterologously expressed in the host, and the expressed product was isolated and purified to obtain a recombinant protein with exocellulase activity.

References
[1] Das KC, Haque N, Baruah KK, et al. Comparative nutrient utilization, growth, and rumen enzyme profile of mithun (Bos frontalis) and Tho-tho cattle (Bos indicus) fed on tree-leaves-based ration [J]. Tropical animal health and production, 2016, 43 (1): 209-214.
[2] Ma G, Chang H, Li S, et al. Phylogenetic relationships and status quo of colonies for gayal based on analysis of cytochrome 6 gene partial sequences[J]. Journal of genetics and genomics, 2007, 34 (5): 413-419.
[3] Ye Daiyong, Huang Hong, Fu Heqing. Progress in cellulose chemistry research [J]. Journal of Chemical Industry and Engineering (China),2016, (8):1782-1791.
[4] Voon L K, Pang S C, Chin S F. Regeneration of cello-oligomers via oelective oepolymerization of cellulose fibers derived from printed paper pastes [J]. Carbohydrate polymers, 2016, 142:31-37.
[5] Ye D, Farriol X. Improving accessibility and reactivity of cellulosates of annual plants for the synthesis of methylcellulose [J]. Cellulose, 2015, 12 (5): 507-515.
[6] Zugenmaier P. Conformation and packing of various crystalline cellulose fibers [J]. Progress in polymer science, 2011, 26 (9): 1341-1417.