Short Communication

Analysis of DNA Sequence Coding for a Partial Cellobiohydrolase Gene from Aspergillus terreus SUK-1

Shaiful Adzni Sharifuddin, Sahidan Senafi and Nik Marzuki Sidik
School of Biosciences and Biotechnology, Faculty of Science and Technology, 43600 Universiti Kebangsaan Malaysia, Bangi, Malaysia.

Abstract

Cellulolytic enzymes are generally composed of multienzyme systems and are divided into three classes, endoglucanase, exoglucanase and β-glucosidase. They work synergistically to hydrolyze cellulose molecules to glucose. Fungal exoglucanase (CBH) is a unique enzyme, capable of degrading highly ordered crystalline cellulose. Here we report the isolation of both genomic and cDNA clone and coding for a partial exoglucanase gene from Aspergillus terreus SUK-1. Two oligonucleotide primers were synthesised based on conserved regions of CBH genes information obtained from database which act as specific primers for gene amplification. PCR amplification was done on both genomic DNA and cDNA. The reactions amplified approximately 538 bp nucleotides from genomic DNA and 486 bp from cDNA which were then sequenced. The sequences were analysed using various platforms such as BLASTX, BLASTN, PSI-BLAST and Pfam. Both genomic and cDNA sequences showed high similarity towards exoglucanase gene and protein from various type of microorganisms. When scanned against Pfam, a high match to glycosyl hydrolase family 7 was observed. Analysis of the genomic sequence showed that it contains a 52 bp intron.

Keywords: Aspergillus terreus SUK-1, cellobiohydrolase, cellulase, gene, sequence.

Introduction

Cellulose is the most abundant biological polymer available on earth. The chemical composition of cellulose consists of D-glucose residues linked by β-1,4-glucosidic bonds to form linear polymeric chains of over 10,000 glucose residues. Individual chains are adhered to each other along their lengths by hydrogen bonding and Van Der Waals forces. This gives the cellulose its insolubility, tensile strength and resistance to enzymes and chemical reagents. Where chains are less ordered or frayed, the celluloses are amorphous, more readily hydrated and more accessible to enzymes. Cellulolytic enzymes are generally composed of multienzyme systems. It can be divided into three classes, endoglucanase (EG) or carboxymethylcellulase (CMCase), exoglucanase or cellobiohydrolase (CBH) and β-1,4-glucosidase. All cellulolytic enzymes share the same specificity for β-1,4-glucosidic bonds and different modes of action of cellulolytic enzymes on the polymeric substrate are commonly described as endo and exo types of attack. Cellulase enzyme complexes work synergistically to hydrolyze cellulose molecules to glucose [1].

Fungal CBH is a unique enzyme, capable of degrading highly ordered crystalline cellulose. Potent cellulolytic fungi generally produces two different CBHs. In Trichoderma reesei, CBH1 and CBH2 make up 60% and 20% of the total cellulolytic protein respectively. Both enzymes were initially thought to act at non-reducing ends of cellulose chains. But, recent evidence indicates that CBH2 acts at the non-reducing end, while CBH1 seems to attack the reducing end of the chains [2]. Here we report the isolation of a genomic clone coding for a partial putative CBH gene from A. terreus SUK-1. A. terreus SUK-1 is a local fungal strain isolated from palm oil waste and its ability to produce high amounts of cellulase enzymes have been reported [3, 4].

Materials and methods

Organism and growth condition

Aspergillus terreus SUK-1 was obtained from Universiti Kebangsaan Malaysia, Malaysia and was cultured on Potato Dextrose Agar (PDA) at 30°C. After 2 weeks incubation it was inoculated into Mendel’s broth medium [5], and incubated in a shaking incubator at 150 rpm and 30°C for five days.

Isolation of genomic DNA and RNA

Cells were harvested by filtering through a piece of nylon gauze. The cells was then frozen and ground in the presence of liquid nitrogen. Genomic DNA was extracted using DNAeasy Plant Mini kit (Qiagen, Germany) and RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Germany). mRNA was isolated from total RNA using the Oligotex mRNA Mini Kit (Qiagen, Germany) and used to generate cDNA using the SuperScript™ cDNA Synthesis System (Invitrogen, USA).

Author for correspondence: Dr. Nik Marzuki Sidik, School of Biosciences and Biotechnology, Faculty of Science and Technology, 43600 Universiti Kebangsaan Malaysia, Bangi, Malaysia. E-mail: nms@ukm.my
PCR amplification and cloning

Two oligonucleotides were synthesised based on the conserved region of CBH gene information obtained from database. The forward primer, CBHf1 was 5'-GTCATTGACGCCAACTGGGCTTG-3' and the reverse primer, CBHr1 was 5'-ACGCTCCCAGCCCTCAGT-3'. PCR amplification was done on both genomic DNA and cDNA. The PCR cycles for amplification were: 94 °C for 3 min for one cycle, followed by 94 °C for 45 sec, 56.5 °C for 45 sec and 72 °C for 1 min for 30 cycles, then followed by a final cycle at 72 °C for 5 min. The pCR product was then separated on 1 % agarose gel and visualized with the presence of ethidium bromide under UV light. The amplified products were purified from the gel using QIAquick Gel Extraction Kit (Qiagen, Germany). The purified PCR products were then cloned into pCR®II-TOPO (Invitrogen, USA) and transformed into E. coli DH5α cells made competent as described by Sambrook et al. [6]. Recombinant plasmids were isolated from the bacterial culture using QIAprep Spin Miniprep Kit (Qiagen, Germany) and sent to First Base Technology, Malaysia for sequencing.

Sequence analysis

The DNA sequences were analysed against nucleotide and protein databases using BLASTX, PSI-BLAST and BLASTN for similarity search. The sequences were also scanned against Protein Families of Alignments (Pfam) and HMMs (http://www.sanger.ac.uk/software/pfam). Sequence search for predicted introns, exons and splicing sites was done using GENSCAN and WISE2, while pair wise alignment between genomic and cDNA sequences was done using Bioedit with Blosum62 matrix.

Results and discussion

After a series of optimisation, the designed primers have successfully amplified a specific DNA band from both templates. A DNA fragment of 538 bp was amplified from genomic DNA (Figure 1) while a 486 bp DNA fragment was amplified from cDNA (Figure 2). Both DNA fragments were cloned into competent E. coli DH5α cells and the inserts were known as SA1 and SA4 respectively.

Alignment between both sequences showed that the SA1 sequence has an intron of 52 bp. The intron splicing sites were identified as GT and AG and were located at 403 and 454 downstream of the SA1 sequence (Figure 3). Coding regions for both SA1 and SA4 sequences were 100% homology and they were coded for the same amino acid sequences. GENSCAN and WISE 2 software are widely used to locate the intron splicing sites within DNA sequences. We have successfully identified AG at the 3' end and GT at the 5' end as the intron splicing sites in the SA1 sequence. It is known that AG and GT are the 3' and 5' boundary or donors/receptors site for introns in most eukaryotes [7, 8, 9].

Homology search with the databases showed that the SA4 sequence shares high percentage of homology with CBH genes and proteins from various organisms. SA4 nucleotide sequences showed 81% homology toward CBH1 gene from Penicillium chrysogenum. High percentage homologies were also detected between SA4 sequence toward catalytic domains of various sources of CBH gene such as from Phanerochaete chrysosporium [10], A. niger [11], Talaromyces emersonii, Thermoaeropus aurantius and A. nidulans (Table 1). On the other hand, SA4 amino acid sequence showed 85%, 83% and 80% homology toward CBH protein from P. chrysogenum, T. aurantius [12] and A. aculeatus [13] respectively. 76% homology was also detected between SA4 amino acid sequence and CBH protein sequences from A. nidulans [14] and P. chrysosporium [10] (Table 2).

When scanned against Pfam, a match with Glycosyl hydrolase family 7 was observed. Glycosyl hydrolases are a widespread group of enzyme that hydrolyses the glycosidic bonds between two or more carbohydrates. Glycosyl hydrolase family 7 comprises enzymes with...
Figure 3. Nucleotide sequence of SA4 encoding for a partial cellobiohydrolase gene from A. terreus SUK-1. Deduced amino acid sequence is shown in capital letter. The 3' end and 5' end splicing sites of the intron are shown in bold. The GenBank accession number AY864863.

Table 1: Homology between SA4 nucleotide sequence and CBH gene sequences from various organisms.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Identity (%)</th>
<th>Organisms</th>
<th>GenBank accession number/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>cbh1</td>
<td>81</td>
<td>P. chrysogenum</td>
<td>AY790330.1</td>
</tr>
<tr>
<td>cbh1</td>
<td>86</td>
<td>P. chrysosoporum</td>
<td>M22220.1</td>
</tr>
<tr>
<td>cbh</td>
<td>83</td>
<td>A. niger</td>
<td>AF156169.1</td>
</tr>
<tr>
<td>cbhA</td>
<td>80</td>
<td>T. emersonii</td>
<td>AY081766.1</td>
</tr>
<tr>
<td>cbh</td>
<td>80</td>
<td>T. aurantiacus</td>
<td>AF421954.1</td>
</tr>
<tr>
<td>cbhA</td>
<td>80</td>
<td>A. nidulans</td>
<td>AF420019.1</td>
</tr>
</tbody>
</table>
several known activities such as EG and CBH. Therefore, based on high percentage homology of the SA4 nucleotide and amino acid sequences toward the CBH gene and protein as shown, we conclude that SA4 codes for a partial CBH gene from *A. terreus* SUK-1.

Acknowledgments

This work was supported by IRPA grant 09-02-02-0108-EA260, Ministry of Science, Technology and Innovation, Malaysia.

References