Nucleotide sequence and characteristics of endoglucanase gene engB from Clostridium cellulovorans

FRANCES FOONG, TETSUO HAMAMOTO, ODED SHOSEYOV and ROY H. DOI*

Department of Biochemistry and Biophysics, University of California, Davis, CA 95616, USA

(Received 5 September 1990; revised 13 March 1991; accepted 16 April 1991)

An endoglucanase gene, engB, from Clostridium cellulovorans, previously cloned into pUC19, has been further characterized and its product investigated. The enzyme, EngB, encoded by the gene was secreted into the periplasmic space of Escherichia coli. The enzyme was active against carboxymethylcellulose, xylan and lichenan but not Avicel (crystalline cellulose). The sequenced gene showed an open reading frame of 1323 base pairs and coded for a protein with a molecular mass of 48.6 kDa. The mRNA contained a typical Gram-positive ribosome-binding site sequence GGAGG and a sequence coding for a putative signal peptide. There is high amino acid and base sequence homology between the N-terminal regions of EngB and another C. cellulovorans endoglucanase, EngD, but they differ significantly in their C-termini. Deletion analyses revealed that up to 32 amino acids of the N-terminus and 52 amino acids of the C-terminus were not required for catalytic activity. The conserved reiterated domains at the C-terminus of EngB were similar to those from endoglucanases from other cellulolytic bacteria. According to our deletion analyses, this region is not needed for catalytic activity.

Introduction

Clostridium cellulovorans is a spore-forming, mesophilic cellulolytic bacterium (Sleat et al., 1984) that can degrade crystalline cellulose. In order to elucidate the mechanism of cellulose degradation by this organism, we have initiated a study to investigate the properties of the subunits of this extracellular cellulase complex (Hamamoto et al., 1990; Shoseyov & Doi, 1990; Shoseyov et al., 1990) and to clone the genes for the subunits. By cloning the genes we plan to analyse their structure and organization, their regulation, their overproduced products and the mechanism of interaction of the subunits with each other and with their solid substrate.

The cellulase complex of clostridia contains many different proteins, including enzymic and non-enzymic subunits (Beguin, 1990; Cavedon et al., 1990; Lamed & Bayer, 1988; Wood & Scott, 1988; Wu & Demain, 1988). A large number of genes have been identified that code for various endoglucanases (Beguin et al., 1985; Grepinet & Beguin, 1986; Joliff et al., 1986; McGavin & Forsberg, 1988; Shoseyov et al., 1990; Wu & Demain, 1988). Some of these genes, in the same or different bacterial species, appear to be related, showing domains of conserved amino acid sequences (Faure et al., 1989).

When the cellulase complex is dissociated, the individual subunits may show endoglucanase or cellobiose-dase activity, but they are unable to degrade crystalline cellulose. Our working criterion for calling an enzyme a cellulase is that it must be able to degrade crystalline cellulose. We have recently shown that a large (170 kDa) non-enzymic subunit is part of the cellulase complex of C. cellulovorans and is essential for the cellulase activity of the complex (Shoseyov & Doi, 1990). To obtain a better understanding of the enzymic subunits of the C. cellulovorans cellulase complex, we are studying the properties of a number of endoglucanase genes and their enzyme products (Hamamoto et al., 1990; Shoseyov et al., 1990). Here we report the cloning of an endoglucanase gene, engB, from this organism, whose product EngB can hydrolyse a variety of cellulosic substrates, although not crystalline cellulose.

Methods

Bacterial strains and plasmids. The C. cellulovorans endo-1,4-β-glucanase gene, engB, was cloned in pUC19 as described previously (Shoseyov et al., 1990). Escherichia coli XL1-Blue (BRL) was used as the

Abbreviations: CMC, carboxymethylcellulose; CMCase, carboxymethylcellulase; PCR, polymerase chain reaction; PNPC, PNPG, PNPX, p-nitrophenyl cellobiose, glucoside, xyloside.

The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession number M37465.
host for the recombinant plasmids pC2 and pC2-engB, and for the engB deletion mutants N1, N2, N3, N4, N5, N6, N7, C1, C2, C3 and C4. The vectors pUC19 and pUC18 were used for construction of all recombinant plasmids. E. coli JM101 was used as the host for M13mp18 and M13mp19 vectors for gene sequencing.

**Gene sequencing.** The dideoxynucleotide method with M13 single strand template (Maniatis et al., 1982) was used for sequencing engB. The procedure supplied with the sequencing kit, Sequenase Version 2.0 (US Biochemical Corp.) was followed exactly and forward and reverse primer sequences were used in a few cases. The sequence of both strands was determined.

**Cell fractionation.** E. coli cells harbouring pC2-engB were grown to 150 Klett units and isopropyl β-D-thiogalactoside (IPTG) was added to a final concentration of 1 mM. Then the cells were grown to 250 Klett units. The cell suspension was centrifuged and the supernatant (extracellular fraction) was collected, concentrated and assayed for enzyme activity. Assay conditions used for the deletion clones were performed using total cell extract: the cells were sonicated in PC buffer (50 mM-potassium phosphate, 12 mM-sodium citrate, 1 mM-sodium azide), pH 6.0, and the supernatant obtained after centrifugation was used for the substrate assay. The osmotic shock method (Neu & Heppel, 1965) was used to obtain the periplasmic fraction; the osmotically shocked cells were sonicated in 100 mM-Tris/HCl buffer, pH 8.0, 1 mM-phenylmethylsulphonyl fluoride (PMSF) to obtain the cytoplasmic fraction. The fractions were diluted approximately 200-fold and the buffers exchanged with 100 mM-Tris/HCl, pH 8.0, 1 mM-PMSF and concentrated with a Centriprep-10 concentrator (Amicon).

**Substrate specificity and enzyme assays.** The E. coli periplasmic fraction was used for substrate specificity assays for the cloned enzyme. Initial testing of carboxymethylcellulase (CMCase), xylanase and lichenanase activity was done on 0.1% carboxymethylcellulose (CMC), xylan or lichenan agar (1.5%) plates. Staining with 0.1 M-red and destaining with 1 M-NaCl was used to visualize the haloes. To test for activity in a strain, the colonies were grown overnight at 37°C, exposed to chloroform to lyse the cells and incubated for another 6 h at 37°C. Then the colonies were washed off with distilled water and the agar plate was stained with 0.1% Congo Red and destained with 1 M-NaCl. CMCase and other substrate activities were quantified by incubating 150 μL of the periplasmic fraction with 140 μL (w/v) of the appropriate substrate in PC buffer, pH 6.0, at 37°C for 10 h. For the deletion clones, total cell extract was used instead. All the substrates, CMC (sodium salt, medium viscosity), microgranular cellulose, fibrous cellulose (medium), polygalacturonic acid, laminarin, xylan, lichenan, p-nitrophenyl β-D-β-glucoside (PNPG) and p-nitrophenyl β-D-xylanoside (PNPX) were obtained from Sigma. PNPG and PNPX (5 mm) were incubated with 150 μL of enzyme fraction in PC buffer, pH 6.0.

The amount of reducing sugars obtained at the end of incubation was measured by the Nelson-Somogyi test for reducing sugars (Wood & Scott, 1988). One unit of activity is defined as 1 μmol of reducing sugar liberated (mg protein)^-1 min^-1. E. coli XL1-Blue containing pUC19 or pUC18 was used as a control in all cases, and the final values were calculated by subtracting the control values. For the 0.1% CMC, 0.1% SDS 7.5% denaturing PAGE zymograms, the gels were run and prepared as described before (Shoseyov & Doi, 1990).

Alkaline phosphatase and malate dehydrogenase activities were measured by the methods of Nisman (1968) and Kitto (1969), respectively. Protein concentration was determined by the BCA Protein Assay System (Pierce).

**Construction of recombinant plasmid pC2-engB.** Primers were synthesized for the 5’ region of engB, just preceding the putative ribosome-binding site, and for the 3’ region just downstream from the putative stop codon. Sequences for NcoI and EcoRI for the 5’ primer (5’-CGGAATTCCGAGGATGAAACCATGGATAAAAGATTATC-ACGGGGGAAG-3’) and BamHI and Smal for 3’ primer (5’-TTCCCCGGGATCTCATTGGCTCCAGTATAAGATAACCA-3’) were introduced for the bases in the sequence underlined in Fig. 1. Polymerase chain reaction (PCR) amplification using pC2 template with the synthesized primers were done according to PCR protocols (Innis et al., 1990). The PCR fragment was restricted with EcoRI and BamHI and cloned into pUC18 to form the recombinant plasmid pC2-engB. The insert was checked by restriction with the same enzymes and by one other enzyme internal to the gene.

**Construction of deletion mutants.** The deletion mutants were constructed either by deleting the amino acid codons from the 5’ or 3’ terminus of engB with restriction enzymes or by synthesizing primers for the 5’ and 3’ termini with restriction sites and then performing PCR amplification using the pC2-engB template with the synthesized primers. The fragments were restricted with the appropriate restriction enzymes and cloned into either pUC18 or pUC19. The insert was checked by restriction with both enzymes to cleave out the complete fragment and with one other enzyme internal to the fragment to check for the correct fragment sizes.

**Results**

**Nucleotide sequence of engB.**

The *C. cellulorans* engB gene was isolated by shotgun cloning of DNA fragments into pUC19 (Shoseyov et al., 1990). An *E. coli* clone containing a 3.2 kb *C. cellulorans* chromosomal DNA fragment was found to possess CMCase activity by initial screening on 0.1% CMC agar plates stained with Congo Red. The plasmid bearing this fragment was named pC2. The fragment was subcloned into M13mp18 and M13mp19 for dideoxy sequencing. An open reading frame of 1323 bases was found on this fragment (Fig. 1), which coded for a protein with a molecular mass of 48.6 kDa. A putative ribosome-binding site, GGAGG, was found 8 bases upstream of the putative ATG methionine start codon. About 172 bases downstream of the translation stop codon, an imperfect inverted repeat, followed by a stretch of 9 thymidine and 3 adenosine residues, was present which could be the transcript termination signal for the gene since it resembles a rho-independent transcription terminator.

Fig. 1. Nucleotide sequence of the 2 kb DNA fragment which contains engB, and the deduced amino acid sequence of EngB. The putative Shine-Dalgarno sequence GGAGG is overlined. The putative signal peptide in *E. coli* is underlined. The translation termination codon is marked with an asterisk. Palindromic sequences at the 3’ end of the gene for the possible transcription termination signal are overlined. Primers synthesized for the 5’ (−13 to +27) upstream sequence and 3’ (+1393 to +1419) downstream sequence of engB are underlined.
engB gene of Clostridium cellulovorans

-399 ATCATACATTATTTAATATCAACTGACCTATAATCATGC

-360 ATTTAGGGAAAAGAAGAAGATACGTTGATATGCGGATATGAGGACAGATTTTAAGAGAATGATGAGGTC

-270 GAGTAGCTAGCTACAACTGACCTCTAATGAAATGAAAATGACGTTGAGGAAGAGATGATTGTTAAGATGGGTAAAATTA

-180 GAGGTAGTCAATAATAAGCTGATCATGTTATAGTGGTGAAGAAGAGCAGAAGATTTTAAGAGAATGGATTGTTAAGATGGGTAAAATTA

-90 CCATGTATAGTCTTCTTCTTTCTCATATTATATAATGTAATCGGTATGATACCATAAAATTGGACAAATAATATATTTT~

1 ATGAATAAAAGATTATCACGGGGAAAGATATCTCTTTTAGCATCAGTTTTCGTTACCACAACTTTTATGGGGGGAGTAAATGTTCTCGCA

~NKRLSRGKISLLASVFVTTTFMGVNVLA

91 TCTACAGCTAAGACAGGTATCGTGACATAACTTCTCAACAAGTTGTTAAGGAAATGAAGGTTGGTTGGAACTTAGGAAATACAATGGAT

S

181 GTACTAGTGACTAAAAATGCTAGTACTCAGTTTCTTTTATATTTAGTTTTTTTCACATTTTTTAGATAGAATATGGAAAATAT

w

271 CTTTCAATAACTTGGGATOGTCATATTGGAGCAGCACCAGATTATGCTATTGATGC~CATGGATGAATAGAGTCGAAG~TAGCAAAT

EngB121 YTWDGHIGAAPDYAIDATWAMNVEIAN

361 CGTCCATAGGGGAAAGATGAAATGGACACAAATTGCAAATCCCT

EngB151 KAATKAVQIANRFKDYGDYLIFETMNEP

451 AAGGCAAATAAACAAATGCGGAAAAACATCGCTTAAATAAAAATCTTTCAACAGCATATGATGATGCTTTATTTTGACAGCT

EngB181 RPVGAADEWSGGSYENRDVMNRYNLTAVNT

541 CTTCCCAAATCCTGGGATGGCTATGTGGAGCCAGTACCTATAGCTATGACACGCTATGATAGCTGAAATAATAGTGAAGCT

EngB211 IATGCAATATGACATTTGATGATAGCTGAAATAATAGTGAAGCT

EngB271 TAATGCAATATGACATTTGATGATAGCTGAAATAATAGTGAAGCT

EngB301 VIVEGMGTINKNLDSRSVFKHREYAVEATV

EngB331 AGACAACTCTCCAAATGACGGGAAACTTGCTGCTATGATGATGCTGAAATAATAGTGAAGCT

EngB361 ACTTTTGGTTGCTCCAAATGACGGGAAACTTGCTGCTATGATGATGCTGAAATAATAGTGAAGCT

EngB391 AGACAACTCTCCAAATGACGGGAAACTTGCTGCTATGATGATGCTGAAATAATAGTGAAGCT

EngB421 MKNKDGVNALDDLAVLKKMLLS

1351 CTAGTGGAATATAGAAGACAGCTTCTTTAGGTTATAGGTAATAGAGTCTCTCTTGTTAAGCTGACTCGAGCATATGATGAGCT

1441 TCTACGTGGAATATAGAAGACAGCTTCTTTAGGTTATAGGTAATAGAGTCTCTCTTGTTAAGCTGACTCGAGCATATGATGAGCT

1531 TAGGTACTTTTATTATACGACTACATTACAAACAGCAGCATATGATGAGCT
termination signal (Rosenberg & Court, 1979). The gene sequence was confirmed by subcloning this ORF (pC2-engB); the CMCase produced from the subclone had the same molecular mass as that from the original pC2 DNA fragment (Fig. 2). The stretch of 30 amino acids at the N-terminus, from methionine to alanine, with a molecular mass of 3.2 kDa (Fig. 1) conforms to the structure of the E. coli signal peptide. The codon usage of the gene has a strong bias (79.3%) towards the third nucleotide being either adenosine or thymidine.

**Cellular localization of EngB produced in E. coli**

E. coli cells harbouring pC2, pC2-engB or pUC19 as control were grown to 250 Klett units and the cells were harvested and fractionated into extracellular, periplasmic and cytoplasmic fractions for enzyme assays. Cells containing pC2 or pC2-engB grew very slowly. CMCase, and the marker enzymes malate dehydrogenase (cytoplasmic) and alkaline phosphatase (periplasmic) were measured in the intracellular, periplasmic and extracellular fractions (Table 1). Ninety-eight percent of the CMCase activity was found in the periplasmic fraction, and the remainder in the cytoplasmic fraction. No activity was detected in the extracellular fraction. These results indicated that the EngB protein was probably secreted into the periplasmic space, although the enzyme could also have leaked out into the periplasm because of lysis of the inner cell membrane.

**Enzymic properties of EngB**

The DNA fragment which contained the ORF (pC2-engB) was amplified by PCR. Initial testing for activity of pC2-engB and pC2 clones was done on CMC plates. The haloes formed after incubation at 37 °C were exactly the same size for both clones. The control plasmid pUC18 in E. coli XL1-Blue did not produce any halo. When the 1.4 kb pC2 fragment was cloned in the opposite direction (in pUC19) the halo formed was very much smaller. Since the 1.4 kb gene fragment did not include any promoter sequences, it is possible that there was readthrough transcription from the vector.

The periplasmic fractions of pC2 and pC2-engB clones were run on CMC-SDS-PAGE denaturing gels and the activity bands on the zymogram revealed that both samples produced proteins with a molecular mass around 42 kDa (Fig. 2). The EngB protein was active on CMC, lichenan, and xylan but not on fibrous or microcrystalline cellulose, Avicel, polygalacturonic acid, mannan, laminaran, PNPC, PNPG, or PNPX. This indicated that the enzyme was only active at an endo-glycolytic level and that it probably only cleaved β-1,4 glycosidic bonds. The specific activity of EngB was 1.2 units for CMCase, 8.1 units for lichenanase and 1.1 units for xylanase.

**Activity of the deletion mutants**

Deletions from the 5' and 3' termini of engB indicated that when 24 amino acids (counting from the ATG
methionine start site) were deleted from the N-terminus of the EngB protein, the activity for the substrates CMC, xylan and lichenan was unaffected. When 32 amino acids were deleted there was a slight drop in activity, and when 42 amino acids were deleted, there was only residual activity (Fig. 3). Deletion of 52 amino acids from the C-terminus, removing the conserved, reiterated region of the gene, resulted in no loss of activity. Deletion of up to 71 amino acids from the C-terminus resulted in a low level of residual activity.

Discussion

The nucleotide sequence of engB consists of an open reading frame of 1323 bases. By calculation, the gene can code for a protein of 48.6 kDa. The zymograms of the periplasmic fraction of E. coli clones harbouring pC2 or pC2-engB showed activity bands at approximately 42 kDa, indicating that pC2 and pC2-engB coded for the same protein. The molecular mass of the product expressed in E. coli was 5-6 kDa smaller than the estimated molecular mass of the protein deduced from the gene sequence. Since most of the CMCase activity was found in the periplasmic space, the protein was probably processed by the E. coli host and secreted into the periplasmic space.

The zymogram of the native cellulase isolated from the extracellular fluid of C. cellulovorans (Shoseyov & Doi, 1990) showed enzyme activity bands ranging from 40 to 100 kDa, so the cloned gene (engB) may code for one of the enzymes of the cellulase complex. Similarities at the amino acid level between EngB and endoglucanase genes from other species range from 30 to 50% similarity (GCG sequence software package analysis, Madison, Wisconsin; Bestfit for gap and homology alignment) with similarities of 43% with C. thermocellum celA, 50% with celB and 46% with celD. Most of the homology was found to be at the C-terminus (Fig. 4), where 50-60 residues were highly conserved (Beguin, 1990). The region contains two homologous segments linked by 8-20 amino acid residues. Similar structures were found in an endoglucanase sequence from C. cellulolyticum, egccA (Faure et al., 1989), and a xylanase gene from C. thermocellum, cloxynZ (Hall et al., 1988). The function of these conserved, reiterated regions of homology is unknown (Faure et al., 1989) but it is not essential for catalytic activity (Grepinet et al., 1988; Hall et al., 1988). The truncated C. thermocellum enzyme (Chavaux et al., 1990) without its conserved C-terminal region was also very similar to the enzyme, EgD, expressed from the intact gene in terms of activity and calcium-binding effects.

The deletion of 52 amino acids from the C-terminus of C. cellulovorans EngB, which eliminated the conserved, reiterated, region, resulted in no activity loss, thereby confirming reports that it is not needed for catalytic activity. When 71 amino acids were deleted, there was still residual activity, as detected by Congo Red agar plate assay, which was more sensitive than the reducing sugar assay. Perhaps at this point the drop in activity was due to instability of the enzyme rather than to loss of the
catalytic site. When 24 amino acids were deleted from the N-terminus, the activity for all three substrates remained the same whereas there was a slight drop in activity when 32 amino acids were deleted. This region can probably be deleted without much effect since this may be the signal peptide of the native enzyme. The N-terminus, the activity for all three substrates catalytic site. When 24 amino acids were deleted from the 5' end of the endoglucanase gene without affecting the catalytic secondary structure of the first 30 amino acids (Chou & Fasman, 1978) of the protein revealed a typical signal peptide (Neuwald & Stauder, 1989). Up to 157 bp could be deleted from the 5’ end of the C. thermocellum endoglucanase EGE gene without affecting the catalytic activity of EGE produced by E. coli (Hazlewood et al., 1990). However, the full-size enzyme is 75 kDa and the protein was still active when it was proteolytically processed to 40 kDa, indicating that a large part of the protein is also not required for catalytic activity. The manner and the ratios in which the activities for the substrates were lost indicated that the catalytic site for the three substrates of EngB is the same.

The engB gene and EngB product both had high homology (75%) at the 5’ end and N-terminus, respectively, with another C. cellulovorans endoglucanase gene, engD, and its product EngD (unpublished data). However, there were significant differences between the two gene products at their C-termini and EngB did not contain the ‘hinge’ region (O’Neill et al., 1986; Wong et al., 1986) present in EngD. The homology between engB and engD ends at this ‘hinge’ region. Since catalytic activity was not lost from engB when nearly all of the C-terminus was deleted, the catalytic site is probably at the N-terminus.

It has been suggested that the conserved, reiterated region may play a role in the interaction between macromolecules carrying reiterated determinants, e.g. the substrate or possibly a scaffolding protein of the cellulosome (Wu & Demain, 1988). A protein of this sort, of 170 kDa, has been purified from C. cellulovorans (Shoseyov & Doi, 1990). It has no enzymic activity but has very strong binding affinity to cellulose. This protein may act as a core protein to which the other catalytic subunits bind to form the cellulosome complex, thus aiding in the cooperative action of the enzymes required for digestion of crystalline cellulose. Perhaps the conserved region of EngB interacts with this 170 kDa protein.

The zymogram patterns of pC2 and pC2-engB clones both showed the presence of an endoglucanase of around 42 kDa. Purified EngB from pC2-engB-containing cells (unpublished data), after being boiled in SDS, run in SDS-PAGE and stained with Coomassie Blue, showed the same molecular mass as that of the enzymes in the zymogram. Since the 1-4 kb PCR product in pC2-engB contains only the nucleotide sequence of the engB gene together with its putative RBS, the transcription initiation must be from the lac promoter on the pUC vector.

<table>
<thead>
<tr>
<th>Clocela</th>
<th>170 kDa</th>
<th>386</th>
<th>416</th>
<th>GDVNQGDNVNSTSDLTMLKRYLVKGVSTINREDVNGAINSSDMTTLKRYLKL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clocelb</td>
<td>157 bp</td>
<td>501</td>
<td>386</td>
<td>GDVNQGDNVNSTSDLTMLKRYLVKGVSTINREDVNGAINSSDMTTLKRYLKL</td>
</tr>
<tr>
<td>Cloceld</td>
<td>157 bp</td>
<td>538</td>
<td>386</td>
<td>GDVNQGDNVNSTSDLTMLKRYLVKGVSTINREDVNGAINSSDMTTLKRYLKL</td>
</tr>
<tr>
<td>CloxynZ</td>
<td>157 bp</td>
<td>472</td>
<td>384</td>
<td>GDVNQGDNVNSTSDLTMLKRYLVKGVSTINREDVNGAINSSDMTTLKRYLKL</td>
</tr>
<tr>
<td>Egcca</td>
<td>427 bp</td>
<td>414</td>
<td>386</td>
<td>GDVNQGDNVNSTSDLTMLKRYLVKGVSTINREDVNGAINSSDMTTLKRYLKL</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison of homologous regions at the C-terminus of EngB with three endoglucanases from C. thermocellum, ClocelA, ClocelB, ClocelD; a xylanase from C. thermocellum, CloxynZ; and an endoglucanase from C. cellulovorans, Egcca. Identical amino acids are indicated with a line, similar amino acids with two dots, and less similar amino acids with one dot. The numbers on the left and right indicate the amino acid positions. These partial amino acid sequences were obtained from the corresponding gene sequences of ClocelA (Beguin et al., 1985), ClocelB (Grepinet & Beguin, 1986), ClocelD (Joliffet al., 1986), CloxynZ, and Egcca (Faure et al., 1989).
although transcription can probably be initiated from a Clostridium promoter sequence by the E. coli RNA polymerase (Zappe et al., 1986).

EngB was active against CMC, xylan and lichenan. Several other cloned gene products have been shown to possess activities on multiple substrates (Gilkes et al., 1988; Gilkes et al., 1984; Hamamoto et al., 1990; McGavin & Forsberg, 1988; Taylor et al., 1987). The enzymes that EngB is capable of using all have β,1-4 glycosyl linkages. The enzyme has very high activity on CMC, but has no detectable activity against lichenan (predominantly β,1-3-glycan), indicating that cleavage is predominantly at the β,1-4 linkages. It also has no activity against Avicel cellulose or PNP Council, which is a substrate for cellobiose activity, suggesting that this enzyme may need the synergistic action of exoglucanases in order to degrade cellulose effectively. The multiple activities of EngB are beneficial since it is possible for leaf complex to contain fewer enzyme species to degrade cellulose material such as cellulose, hemicellulose (xylans), and lichenan. C. cellulovorans was isolated from a methanogenic digester of poplar wood (Sleat et al., 1984). Perhaps cellulolytic organisms have evolved enzymes of wide specificity which make them capable of digesting several types of substrates found in woody material.

The codon usage of the engB gene was strongly biased, since 79-3% (342/431 codons) of the third nucleotides were either adenosine or thymidine. This reflects the very low G+C content (26-27 mol%) of C. cellulovorans (Sleat et al., 1984).

The engB gene cassette produced from the PCR reaction contains restriction sites, which facilitates transfer of the gene from one vector to another. This construct should be useful for studying gene regulation and expression, overproduction of the enzyme, and secretion.

This work was supported in part by grant GM 19673 from the National Institute of General Medical Sciences (R.H.D.) and Grant # DE-FG03-87ER13705 from the Department of Energy (R.H.D). O.S. was supported by fellowship SI 0057-87 from the United States-Israel Binational Agricultural Research and Development Fund.

References


