Isolation and analysis of two cellulase cDNAs from Orpinomyces joyonii

Xiao Qiu *,1, Brent Selinger 2, L.-J. Yanke, K.-J. Cheng

Lethbridge Research Station, Agriculture and Agri-Food Canada, PO Box 3000, Lethbridge, Alberta, Canada T1J 4B1

Received 30 August 1999; received in revised form 22 December 1999; accepted 13 January 2000

Received by V. Larionov

Abstract

Two cellulase cDNAs, celB29 and celB2, were isolated from a cDNA library derived from mRNA extracted from the anaerobic fungus, Orpinomyces joyonii strain SG4. The nucleotide sequences of celB2 and celB29 and the primary structures of the proteins encoded by these cDNAs were determined. The larger celB29 cDNA was 1966 bp long and encoded a 477 amino acid polypeptide with a molecular weight of 54 kDa. Analysis of the 1451 bp celB2 cDNA revealed an 1164 bp open reading frame coding for a 44 kDa protein consisting of 388 amino acids. Both deduced proteins had a high sequence similarity in central regions containing putative catalytic domains. Primary structure analysis revealed that CelB29 contained a Thr/Pro-rich sequence that separated the N-terminal catalytic domain from a C-terminal reiterated region of unknown function. Homology analysis showed that both enzymes belong to glycosyl hydrolase family 5 and were most closely related to endoglucanases from the anaerobic fungi Neocallimastis multiparametricum, Neocallimastis frontalis and Orpinomyces sp. The classification of CelB29 and CelB2 as endoglucanases was supported by enzyme assays. The cloned enzymes had high activities towards barley β-glucan, lichenan and carboxymethylcellulose (CMC), but not Avicel, laminarin, pachyman, xylan and pullulan. In addition, CelB29 and CelB2 showed activity against p-nitrophenyl-β-d-cellotrioside, p-nitrophenyl-β-d-cellobioside, p-nitrophenyl-β-d-cellopentaoside, p-nitrophenyl-β-d-glucopyranoside, and p-nitrophenyl-β-d-glucopentaoside with preferential activity against p-nitrophenyl-β-d-glucopyranoside (pNP-G5) but not p-nitrophenyl-β-d-glucopyranoside (pNP-G4). Based on these results, we proposed that CelB29 and CelB2 are endoglucanases with broad substrate specificities for short- and long-chain β-1,4-glucans. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: cDNA sequence; Cellulose degradation; Cloning; Rumen fungus

1. Introduction

Cellulose is the most abundant source of carbon found in nature. This linear β-glucan chain, consisting of β-glucose monomers linked by β-1,4-glycosidic bonds, is the principal structural component of plant cell walls. Individual cellulose molecules are formed into insoluble cellulose microfibrils through numerous intra- and intermolecular hydrogen bonds, thereby creating the very complex and recalcitrant physical structure of cellulose found in plant cell walls (Delmer and Amor, 1995; Carpita, 1996).

The value of cellulose as a renewable source of energy and carbon has made cellulose hydrolysis the subject of scientific research and industrial interest for many years. Over the past decades, a large amount of information has been accumulated on the hydrolysis of cellulose by cellulases. Many cellulosytic systems from a wide array of organisms, including higher flowering plants, protozoa, fungi and bacteria, have been studied. Conversion of cellulose to glucose involves at least three types of cellulases: (1) endoglucanases (EC 3.2.1.4), which randomly cleave internal β-1,4-glycosidic linkages, (2) cel-
lobiobiodratalas (EC 3.2.1.91), which hydrolyse \( \beta \)-1,4-glycosidic linkages of cellulose from the ends liberating cellobiose, and (3) \( \beta \)-glycosidases (EC 3.2.1.21), which release two glucose residues from one cellobiose molecule (Beguin, 1990). Despite the information available on these enzyme systems and the structure of plant cell walls (Gilbert and Hazelwood, 1993; Beguin and Lemaire, 1996), our application of this knowledge to cellulose degradation has met with limited success. The lack of success may be attributed to at least two factors: the inherent complexity and heterogeneity of native cellulose, and our lack of a thorough understanding of the basic processes of the hydrolysis (Lechini, 1995). Therefore, an understanding of the molecular mechanisms underlying cellulose degradation in combination with new and superior enzymes will certainly facilitate increased usage of this valuable renewable resource.

2. Materials and methods

2.1. Microbial strains and culture media

*Orpinomyces joyonii* strain SG4, a sheep ruminal isolate, was obtained from the Lethbridge Research Centre Culture Collection and grown in modified semi-defined medium (Lowe et al., 1985) with either Whatman No. 1 filter paper or straw as a carbon source. The fungus was cultivated using the anaerobic technique of Hungate (1950) as modified by Bryant and Burkey (1953). *Escherichia coli* XL1-Blue MRF’ (Stratagene Cloning Systems, La Jolla, CA), the host for cDNA library construction and vector propagation, was grown in LB medium (Sambrook et al., 1989). Ampicillin (100 \( \mu \)g/ml) was incorporated into media to culture plasmid-bearing *E. coli* strains.

2.2. RNA isolation

Fungal mycelia were harvested from culture media and lyophilized. The dried mycelia were ground under liquid nitrogen into powder with a mortar and pestle. Total RNA was isolated from powdered mycelia according to Qia and Erickson (1994) with some modifications. Briefly, following homogenization of the mycelia in extraction buffer (100 mM Tris–Cl, pH 8.0, 50 mM EDTA, 500 mM NaCl, 2% SDS, 1% \( \beta \)-mercaptoethanol), an equal volume of phenol:chloroform (1:1) was added. The suspension was agitated for 1 min with a vortex mixer. The mixture was centrifuged, and the resulting aqueous phase was removed to a clean centrifuge tube, extracted with phenol:chloroform and precipitated with one-third volume of 8 M lithium chloride. The RNA pellet was washed with 2 M LiCl, followed by water and ethanol, and then resuspended in diethyl pyrocarbonate-treated water. Poly(A) \(^{+} \) RNA was isolated from total RNA using oligo(dT)-cellulose chromatography following standard protocols (Sambrook et al., 1989).

2.3. Construction and screening of an *O. joyonii* cDNA library

An *O. joyonii* cDNA library was constructed with the aid of a Stratagene cDNA synthesis kit, Lambda...
2. Results and discussion

3.1. Isolation of cellulase cDNA clones

In order to identify the genes encoding cellulolytic enzymes, a cDNA library was constructed using poly(A)+ RNA from O. juelonii grown on a medium with either straw or filter paper as the sole carbon source. The resulting library contained a total of \(5 \times 10^8\) original plaque-forming units. Screening with the mixture of CMC-barley \(\beta\)-glucan overlays detected 69 plaques with hydrolytic activity.

To avoid analysing duplicate clones, restriction mapping and substrate plate assays were used to classify the positive clones into several preliminary groups. Two cDNAs from two separate groups, celB29 and celB2, were further analysed.

3.2. Nucleotide sequence analysis

CelB29 is 1451 bp in length and contains a coding region of 1431 bp with a start codon at position 167 and stop codon at position 1598 (Fig. 1). The 5′ untranslated leading sequence is rich in A/T residues (94%). The 3′ untranslated region is also A/T-rich (91%), and contains six potential eukaryotic polycadenylation signals (AATAAA) and a 19 bp poly(A) tail. Similar AT rich non-coding flanking regions have been reported for other sequences cloned from ruminal fungi (Gilbert et al., 1992; Zhou et al., 1994). The translational start site was established as a result of the following observations. There are no start codons and several stop codons present upstream of the proposed translation start site, and the first 20 amino acids of the deduced CelB29 polypeptide constitute a typical signal peptide (i.e. it is generally located at the N-terminus of the proprotein; von Heijne, 1986). The deduced CelB29 polypeptide contains 477 amino acids and has a molecular weight of 54 kDa and a predicted isoelectric point of 6.3. A series of 21 amino acids, rich in Thr and Pro, were also found in other microbial cellulases. They are hypothesized to function as hinges or linkers between two functional domains (Gilks et al., 1992; Zhou et al., 1994). The translational start site was identified according to Kawai et al. (1987). One unit of enzyme activity was defined as the amount of enzyme catalysing the formation of 1 \(\mu\)mol of product per minute. Cellulase binding assays were performed according to Ali et al. (1995).

For enzymatic assays of \(\beta\)-glucosidase, cellobiohydrolase and \(\beta\)-cellodextrinase, 0.1 ml of enzyme solution was mixed with 0.1 ml of 10 mM arylglycoside substrates in 20 mM phosphate buffer \((pH\ 6.8)\) at 37°C for 0.5 h. Released reducing sugars were quantified as described by Kawai et al. (1987). One unit of enzyme activity was defined as the amount of enzyme catalysing the formation of 1 \(\mu\)mol of product per minute. Cellulase binding assays were performed according to Ali et al. (1995).

For enzymatic assays of \(\beta\)-glucosidase, cellobiohydrolase and \(\beta\)-cellodextrinase, 0.1 ml of enzyme solution was mixed with 0.1 ml of 10 mM arylglycoside substrates in 20 mM phosphate buffer \((pH\ 6.8)\). The mixture was incubated at 37°C for 15 min. The reaction was stopped with the addition of 0.1 ml 1 M Na2CO3. The resulting mixture was measured spectrophotometrically at 405 nm. One unit of enzymatic activity was defined as the amount of enzyme catalysing the formation of 1 \(\mu\)mol of p-nitrophenol per minute.

Temperature and pH optima were determined according to the procedure of Denman et al. (1996).

3. Results and discussion

3.1. Isolation of cellulase cDNA clones

In order to identify the genes encoding cellulolytic enzymes, a cDNA library was constructed using poly(A)+ RNA from O. juelonii grown on a medium with either straw or filter paper as the sole carbon source. The resulting library contained a total of \(5 \times 10^8\) original plaque-forming units. Screening with the mixture of CMC-barley \(\beta\)-glucan overlays detected 69 plaques with hydrolytic activity.

To avoid analysing duplicate clones, restriction mapping and substrate plate assays were used to classify the positive clones into several preliminary groups. Two cDNAs from two separate groups, celB29 and celB2, were further analysed.
Fig. 1. Nucleotide and deduced protein sequences of CelB29. The putative catalytic region is indicated in bold italics. The signal peptide, a linker sequence and putative polyadenylation signals are shown in bold. Amino acid direct repeats are underlined.

166 X. Qiu et al. / Gene 245 (2000) 119–126

a predicted isoelectric point of 4.9. Like CelB29, there is a short amino acid direct repeat (NEGK/NW) in the C-terminus.

3.3. Homology analyses of the polypeptide sequences

A comparison of CelB29 and CelB2 primary sequences revealed that CelB2 is about 80 amino acids shorter than CelB29 due to about 40 residues missing from the N- and C-terminal regions. The overall identity of the two sequences is 57%. However, the highest degree of homology (66%) was found in the putative catalytic region (Fig. 2).

Database searches revealed that CelB2 and CelB29 exhibited a high degree of amino acid sequence similarity to CelA from N. freatalis (Accession No. AAC30949, Fujino et al., 1998), CelB from N. patriciarum (Accession No. Q12647, Zhou, et al., 1994) and CelB from Orpinomyces sp. (Accession No. AAD04193, Li et al., 1997). All of these enzymes belong to the cellulase family (Henrissat et al., 1989) or glycosyl hydrolase family 5 (Henrissat, 1991), which is characterized by an amino acid signature L–I–F–E–G–x–N–E–P–R with a conserved glutamic acid residue potentially involved in catalytic degradation of cellulose. CelB29 contains 16 tryptophan residues that are perfectly conserved in the homologous cellulases.
contain hydroxyl amino acid linker sequences in their central region. While CelB29 is more closely related to CelB from Orpinomyces sp. and CelB from N. patri- 

ciarum, they all have two distinct domains separated by a Thr/Pro-rich linker sequence.

3.4. Enzymatic properties of cloned enzymes

Native cellulosomes are complex and heterogeneous in nature. Therefore, it could not be used in characterizing the cellulases. Instead, we utilized a range of different model substrates (e.g. CMC, Avicel, barley $\beta$-glucan, lichenan, laminarin, pachyman, and xylan) to characterize the substrate specificity. A preliminary qualitative assessment of enzymatic activities was performed by inoculating E. coli strains bearing the cellulase cDNAs onto agar media containing oat spelt xylan, CMC, lichenan, barley $\beta$-glucan, Avicel, laminarin, pachyman or pullulan. The results indicated that both enzymes have very similar enzymatic properties, although there are distinct differences in the primary structure of the two deduced protein sequences. Except for strong activities on CMC, lichenan and barley $\beta$-glucan, neither of the cloned enzymes showed any activity to the rest of the substrates, suggesting that both enzymes are $\beta$-1,4-endoglucanases. CMC is a water-soluble long chain cellulose with carboxymethyl substitu-
tions that is commonly used as a model substrate for detecting $\beta$-1,4-endoglucanases (Teeri, 1997). Cleavage between the substituent groups can result in a rapid decrease in the degree of polymerization and solution viscosity. Digestion of lichenan and barley $\beta$-glucan (mixtures of $\beta$-1,3-1,4 linkages) might be mainly due to random cleavage of $\beta$-1,4-glycosidic linkage in the substrates since both enzymes were unable to digest laminarin and pachy-
man, and their major components are $\beta$-1,3-glucans.

The activities of CelB2 and CelB29 on different substrates were also measured quantitatively. These results (Table 1) coincided with the plate assays. When incubated at 37°C for 30 min, both enzymes could produce detectable reducing sugars on CMC, barley $\beta$-glucan and lichenan, but not Avicel, laminarin, pachy-
man and pullulan. CelB29 and CelB2 exhibited the highest activity towards barley $\beta$-glucan, then lichenan, and followed by CMC. Unlike the plate assays, CelB2 had some activities towards oat spelt xylan. The differ-
ence between the plate and quantitative assays regarding xylanase activity may be a reflection of the sensitivity of the latter method.

In order to reveal more about the catalytic mode of the cloned enzymes, quantitative assays using arylglyco-
sides as substrates were performed (Table 2). These cross-linked compounds are convenient substrates for detecting cellohexaose activity (Huang and Forchberg, 1988). Cellohexaoses are a class of short-chain soluble oligosaccharides, such as cellotriose, cellotetraose, cello-

Fig. 2. Comparison of deduced protein sequences of CelB29 and CelB2. Identical amino acids are indicated as ‘|’.
Fig. 3. Alignment of CelB29 and CelB2 with homologous cellulase CelA from Neocallimastix frontalis (Accession No. AAC63094, Fujino et al., 1998), CelB from Neocallimastix patriciarum (Accession No. Q12647, Zhou et al., 1994) and CelB from Orpinomyces sp. (Accession No. AAD04193, Li et al., 1997). ’*’ perfectly conserved; ’.’ well conserved. The conserved signature sequence of glycosyl hydrolase family 5 and the threonine-rich linker sequence are framed in a box. Conserved glutamic acids and tryptophan residues are highlighted.

pentose and cellohexaose. Artificial substrates such as methyllumelliteryl cellobioside (MUC) and p-nitrophenyl-\(\beta\)-D-glucoside (pNP-G\(\beta\)_1) may also be used to measure the cellobiohydrolase activity. The results showed that while both enzymes lacked \(\beta\)-glucosidase activity (i.e. they did not cleave p-nitrophenyl-\(\beta\)-D-glucose), they did show a broad range of substrate specificity for cellodextrins (i.e. p-nitrophenyl-\(\beta\)-D-cellobioside to p-nitrophenyl-\(\beta\)-D-cellopentaoside). The highest activities were towards pNP-G\(\beta\)_3, then pNP-G\(\beta\)_2,
followed by pNP-G₄, indicating that CelB29 and CelB2 possess β-cellobiohydrolase and β-cellobiohydrodolase. Cleavage of pNP-G₃ might be due to β-cellobiohydrolase (cellobiohydrolase), but not the combined action of cellobiohydrodolase and β-glucosidase since both cloned enzymes lack the activity of β-glucosidase. Based on the same rationale, cleavage of pNP-G₂ should be the result of either cellobiohydrolase or the combination of cellobiohydrodolase and cellobiohydrolase. Compared to CelB29, CelB2 had a higher activity towards all arylglycoside substrates, indicating that a C-terminal shortage of a number of amino acids in CelB2 did not affect the enzymatic capacity. Based on these results, we propose that CelB29 and CelB2 are endoglucanases with broad substrate specificities for short- and long-chain β-1,4-glucans, and these enzymes showed catalytic activities including endoglucanase, cellobiohydrolase and cellobiohydrodolase. However, it is worth noting that due to the overlapping specificity on several substrates supplied, the clear-cut classification of the cloned enzymes into a specific group is difficult. A similar phenomenon has also been found in celD, a polysaccharide hydrolase cDNA from N. patriciarum encoding three multi-functional catalytic domains with high endoglucanase, cellobiohydrolase and xylanase activities (Xue et al., 1992b).

In view of the findings that many characterized cellulases consist of two distinct domains, a cellulose-binding domain (CBD) and a catalytic domain (Gilkes et al., 1991), a search for CBD sequences was conducted. No distinct CBDs were found in either CelB29 or CelB2. This result was confirmed with a cellulose-binding assay that showed that neither CelB29 nor CelB2 was able to bind Avicel, a microcrystalline cellulose (data not shown). The absence of CBDs and the presence of reiterated scaffold binding sequences in CelB29 and CelB2 suggested that these enzymes may be immobilized to the cellulose, a cellulose hydrolytic complex. More than 80% of extracellular cellulase activities of Orpinomyces are associated with a cellulose-binding complex (Ali et al., 1995). As such, the enzymes may function more effectively.

The effect of pH and temperature of both cloned cellulases on CMC as a substrate was measured in 20 mM sodium phosphate bufrer over a pH range of 4.2–7.8 with an interval of 0.4 and a temperature range of 20–100°C with 5°C intervals. Results revealed that both cloned enzymes were active over a broad pH range between 5.4 and 7.4. CelB29 had the highest activity at pH 5.8, while CelB2 was most active at pH 6.6. Temperature optima for CelB2 and CelB29 were 45 and 50°C, respectively. Both enzymes retained at least 65% of their activity over the temperature range of 50–60°C.

In conclusion, we have isolated two cDNA celB29 and celB2 from O. joyonii that encode endoglucanases. CelB29 and CelB2 share a high amino acid identity (37%) throughout the sequence and belong to the same group of glycansases: cellulase A or glycosyl hydrolase family 5. Neither protein contains a cellulose-binding domain. Tandem repeats in the C-terminal region suggested that both enzymes were immobilized to the cellulose. Although there are differences in the primary structure of the two deduced polypeptide sequences, both enzymes appeared to have very similar enzymatic properties. Both could hydrolyse barley β-glucan, lichenan, CMC, but not Avicel, laminarin, pachyman and pullulan, indicating their β-1,4-endoglucanase activity. In addition, both enzymes were able to cleave pNP-G₁, pNP-G₂, pNP-G₄ and pNP-G₅, but not pNP-G₃, indicating their cellobiohydrolase and cellobiohydrodolase but not β-glucosidase activities.

Table 1

<table>
<thead>
<tr>
<th>Substrates</th>
<th>Major components</th>
<th>CelB29 (units/mg)</th>
<th>CelB29 (units/ml)</th>
<th>CelB2 (units/mg)</th>
<th>CelB2 (units/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMC</td>
<td>β-1,4-glucan</td>
<td>0.245</td>
<td>2.302</td>
<td>0.350</td>
<td>2.338</td>
</tr>
<tr>
<td>Barley β-glucan</td>
<td>β-1,3-1,4-glucan</td>
<td>0.403</td>
<td>3.257</td>
<td>0.505</td>
<td>3.374</td>
</tr>
<tr>
<td>Oat spelt xylan</td>
<td>β-1,4-sylan</td>
<td>0</td>
<td>0.008</td>
<td>0.008</td>
<td>0.051</td>
</tr>
<tr>
<td>Lichenan</td>
<td>β-1,3,1-4-glucan</td>
<td>0.540</td>
<td>2.752</td>
<td>0.428</td>
<td>2.656</td>
</tr>
<tr>
<td>Avicel</td>
<td>Crystalline cellulose</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lammarin</td>
<td>β-1,3-glucan</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pachyman</td>
<td>β-1,3-glucan</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pullulan</td>
<td>n-1,6-glucan</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Crude cell lysate was used for the assay of the enzyme activities. One unit of enzyme activity was defined as the amount of enzyme catalysing the formation of 1 nmol of p-nitrophenol per minute.

Table 2

<table>
<thead>
<tr>
<th>Trivial name</th>
<th>Full name</th>
<th>CelB29</th>
<th>CelB2</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNP-G₁</td>
<td>p-nitrophenol-β-D-glucopyranoside</td>
<td>0°</td>
<td>0</td>
</tr>
<tr>
<td>pNP-G₂</td>
<td>p-nitrophenol-β-D-cellobioside</td>
<td>16</td>
<td>51</td>
</tr>
<tr>
<td>pNP-G₃</td>
<td>p-nitrophenol-β-D-cellotrioside</td>
<td>73</td>
<td>221</td>
</tr>
<tr>
<td>pNP-G₄</td>
<td>p-nitrophenol-β-D-cellotetraoside</td>
<td>31</td>
<td>56</td>
</tr>
<tr>
<td>pNP-G₅</td>
<td>p-nitrophenol-β-D-cellpentoside</td>
<td>19</td>
<td>22</td>
</tr>
</tbody>
</table>

* Crude cell lysate was used for the assay of the enzyme activities.
Acknowledgements

The authors would like to thank Adele Popp, Yan Ge and Therese Despins for the technical assistance. Dr Akio Takenaka for the helpful discussion during the course of the experiment and Dr Sam Mackenzie for critical reading of the manuscript. This research was partially funded by the Canadian Government National Biotechnology Strategy fund and Alberta Agriculture Research Institute Matching Grant no. 96M005.

This paper is LRC contribution no. 3879703

References


