Intronless *celB* from the anaerobic fungus *Neocallimastix patriciarum* encodes a modular family A endoglucanase

Liqing ZHOU,* Gang-ping XUE,† Colin G. ORPIN,† Gary W. BLACK,‡ Harry J. GILBERT‡ and Geoffrey P. HAZLEWOOD*§

*Department of Cellular Physiology, AFRC Babraham Institute, Babraham, Cambridge CB2 4AT, U.K., †CSIRO Division of Tropical Crops and Pastures, 306 Carmody Road, St. Lucia, Queensland 3067, Australia and ‡Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K.

The cDNA designated *celB* from the anaerobic rumen fungus *Neocallimastix patriciarum* contained a single open reading frame of 1422 bp coding for a protein (CelB) of M_r 53070. CelB expressed by *Escherichia coli* harbouring the full-length gene hydrolysed carboxymethylcellulose in the manner of an endoglucanase, but was most active against barley β-glucan. It also released reducing sugar from xylan and lichenan, but was inactive against crystalline cellulose, laminarin, mannan, galactan and arabinan. The rate of hydrolysis of cellulo-oligosaccharides by CelB increased with increasing chain length from cellotriose to cellopentaose. The predicted structure of CelB contained features indicative of modular structure. The first 360 residues of CelB constituted a fully functional catalytic domain that was homologous with bacterial endoglucanases belonging to cellulase family A, including five which originate from three different species of anaerobic rumen bacteria. Downstream from this domain, and linked to it by a serine/threonine-rich hinge, was a non-catalytic domain containing short tandem repeats, homologous to the C-terminal repeats contained in xylanase A from the same anaerobic fungus. Unlike previous fungal cellulases, genomic *celB* was devoid of introns. This lack of introns and the homology of its encoded product with rumen bacterial endoglucanases suggest that acquisition of *celB* by the fungus may at some stage have involved horizontal gene transfer from a prokaryote to *N. patriciarum*.

INTRODUCTION

Microbial cellulases are typically composed of several different enzymes that cleave β-1,4-glycosidic linkages and act cooperatively in converting cellulose into glucose. Each enzyme may occur as a number of isoenzymes, usually encoded by multiple genes; this is particularly true of the endoglucanase (EC 3.2.1.3) and cellobiohydrolase (EC 3.2.1.91) components (Knowles et al., 1987; Hazlewood et al., 1988). Recent molecular studies have established that many individual endoglucanases and cellobiohydrolases are modular in structure, being composed of one or more catalytic domains, linked by sequences rich in hydroxy amino acids, to a non-catalytic cellulose-binding domain (see Gilkes et al., 1991). Sequence comparisons, based primarily on the pattern of clustering of hydrophobic residues, suggest that the catalytic domains of these enzymes can be grouped into well-defined families, indicating that their genes evolved divergently from relatively few ancestral sequences (Henrissat et al., 1989; Gilkes et al., 1991). The fact that genes belonging to a single family have been isolated from different genera of cellulolytic bacteria, as well as from cellulolytic fungi, further suggests that the evolution of cellulase systems has involved extensive transfer of genes between unrelated species.

The structure and composition of microbial cellulase systems are subject to considerable variability. For example, in aerobic fungi such as *Trichoderma reesei*, cellulase components, including essential isoenzymic cellobiohydrolases, do not aggregate to form an identifiable complex, but nevertheless act in concert to hydrolyse crystalline cellulose (Wood, 1992). By comparison, the cellulase systems of anaerobic bacteria, of which the cellulosome of *Clostridium thermocellum* is a good example, typically consist of multiple endoglucanases which associate with a non-catalytic scaffolding protein via a highly conserved recognition sequence

to form an ordered multiprotein complex (Lamed and Bayer, 1988; Tokatlidis et al., 1991; Fujino et al., 1992).

Anaerobic fungi belonging to the order *Chytridiomycetes* and including the species *Neocallimastix patriciarum* have evolved side by side with anaerobic cellulolytic bacteria and are among the most potent producers of cellulases in the herbivore gut (Orpin and Joblin, 1988; Teunissen et al., 1992). Biochemically, their cellulase system resembles that of anaerobic bacteria in consisting of a large multiprotein complex (Wilson and Wood, 1992); unlike the cellulases of aerobic fungi, it is devoid of a true cellobiohydrolase. To establish whether similarity with the bacterial cellulases is evident at the molecular level, and to correlate the structure of these enzymes with their function, we have initiated studies to determine the molecular architecture of the cellulases produced by *N. patriciarum*. In a previous report, Xue et al. (1992) isolated several cDNAs (*celA*, *celB* and *celC*) encoding fungal enzymes that hydrolyse carboxymethylcellulose (CMC). In this paper, we present the first nucleotide sequence of an endoglucanase gene (*celB*) from the anaerobic fungus *N. patriciarum*, and describe in detail the catalytic activity of the encoded enzyme. Unlike previously described fungal cellulases, the gene is devoid of introns and encodes a modular protein; a small non-catalytic domain at the C-terminus of the encoded endoglucanase contains repeated elements, conserved in xylanase A from the same fungus. The catalytic domain exhibits significant homology with family A endoglucanases.

MATERIALS AND METHODS

Microbial strains, vectors and culture conditions

N. patriciarum was isolated from the rumen of a sheep by Orpin and Munn (1986) and was cultured in anaerobic medium containing Avicel as sole carbohydrate source (Kemp et al.,

---

Figure 1 Restriction map of recombinant plasmids containing celB

Enzyme sites are as follows: R, EcoRI; RV, EcoRV; D, DraI; Pf, PflMI; Pv, PvuII; S, SpeI; X, XhoI. Vector sequences are shown by a single line. The large arrow denotes the extent and orientation of the celB open reading frame. The phenotype of E. coli XL1-blue strains harbouring derivatives of celB is shown as CMCase⁺ or CMCase⁻. Areas shaded ▨ and ■ signify the C-terminal tandem repeats and serine/threonine linker sequence respectively.

1984). The *Escherichia coli* host strain XL1-blue was obtained from Stratagene and was grown in Luria broth (LB) containing ampicillin (100 μg/ml) to select transformants and, where necessary, isopropyl β-D-thiogalactopyranoside (IPTG; 0.5 mM) to induce *lacZ'p* activity.

General recombinant DNA procedures

DNA isolation, agarose-gel electrophoresis and the transformation of *E. coli* were carried out as described by Sambrook et al. (1989). Restriction endonucleases and other DNA-modifying enzymes were used in accordance with the manufacturer's instructions. For Southern hybridization, DNA restriction fragments were blotted on to Hybond N membrane (Amersham). Probe DNA was labelled with [α-³²P]dCTP by random priming (Sambrook et al., 1989). Prehybridization, hybridization and washing sequences were as described in the Amersham protocol.

Construction and screening of the genomic library

Genomic DNA was isolated as described by Brownlee (1988) from *N. patriciarum* mycelium that had been cultured in anaerobic medium containing cellobiose (0.5%, w/v) as carbon source. After partial digestion with Sau3A, fragments in the size range 4–12 kb were ligated into λZAPII which had been digested with XhoI and half filled. A library consisting of 3 × 10⁴ recombinant phage was plated on *E. coli* XL1-blue and plaques were lifted on to Hybond N and screened for hybridization to a radiolabelled probe consisting of the cDNA insert from pL9 (Figure 1) using the Amersham protocol.

Nucleotide sequencing

Double-stranded plasmid DNA was sequenced by the dideoxy-chain-termination method of Sanger et al. (1977) using a protocol recommended for the Sequenase DNA Sequencing Kit (United States Biochemical). After denaturation by alkali, plasmid DNA was neutralized by spin dialysis (Murphy and Kavanagh, 1988). To generate overlapping sequences, appropriate restriction fragments were cloned into pBluescript (SK⁻). Where required, nested deletions were generated by digesting plasmid DNA with

Figure 2 Nucleotide sequence of full-length celB cDNA and the derived amino acid sequence of endoglucanase B (CelB)

The proposed translational start codon, the serine/threonine linker sequence and the C-terminal tandem repeats are boxed. The 3′ ends of the truncated fragments cloned in pL9Exo1 and pL9Exo2 are denoted by black triangles. Primers used for extending the incomplete celB cDNA by 5′-RACE were the reverse complement of the sequence indicated.

Exonuclease III (ExoIII) according to the protocol recommended for use with the Erase-a-Base system (Promega). The complete sequence of the cDNA contained in pL9 was determined in both

strands, and was compiled and ordered using software written by the Genetics Computer Group at the University of Wisconsin.

Extension of *celB* cDNA by rapid amplification of cDNA ends (5′-RACE)

With *N. patriciarum* mRNA as the template, single-stranded cDNA was synthesized by reverse transcription, using primer A (Figure 2), and was appended with a poly(A) tail. Two rounds of PCR amplification were carried out using primer B (Figure 2) and an *Spe*I adaptor–(dT)$_{17}$ primer [CATCGATATCTAGA-TACTAG(T)$_{17}$]. A third and final round of amplification was completed with primer C (Figure 2) and products were cleaved with *Spe*I and *Pfl*M1 and the fragment generated, which contained the 5′ extension of the *celB* cDNA, was ligated into pL9EEV from which the *Spe*I/*Pfl*M1 fragment had been removed. (pL9EEV was generated previously by removing the *Eco*RV/*Xho*I fragment from pL9.)

In vitro transcription/translation

Full-length *celB* cloned in-phase with the *lacZ′* translation initiation codon of pBluescript, was subjected to *in vitro* transcription/translation using the Amersham prokaryotic DNA-directed translation kit. Polypeptides, labelled with [^{35}S]methionine, were separated by PAGE (10% acrylamide) and were visualized by fluorography.

Assays

Hydrolysis of the polymeric substrates CMC, laminarin, lichenan, oat-spelt xylan (all from Sigma), and barley β-glucan (MegaZyme) by *E. coli* XL1-blue recombinants containing *celB*, or its truncated derivatives, was investigated by overlaying LB plates on which the strains had been cultured, with agar (1.5%, w/v) containing the appropriate substrate (0.25% or 0.5%, w/v), incubating at 37 °C and subsequently staining with Congo Red (Hazlewood et al., 1990). Dyed arabinan, galactan or mannan (MegaZyme) were similarly incorporated into agar overlays, and the plates examined for release of dye during incubation at 37 °C. Quantitative assessment of enzyme activities was carried out at 37 °C in PC buffer (50 mM K$_2$HPO$_4$, 12 mM citric acid, pH 6.5), with substrate concentrations in the range 2.5–10 mg/ml. Reducing sugar was measured using the dinitro-salicylic acid reagent (Miller, 1959). Hydrolysis of cellulo-oligosaccharides was monitored by h.p.l.c. as described by Rixon et al. (1992). Protein was measured by dye binding (Sedmak and Grossberg, 1977) or by the Lowry method with BSA as standard. Viscometric assays were performed as described previously (Romaniec et al., 1987).

RESULTS AND DISCUSSION

Sequence of *celB*

A number of *celB* cDNAs coding for the putative endoglucanase CelB were isolated by screening a cDNA library of 4 × 10^5 clones prepared from *N. patriciarum* grown with Avicel as sole carbon source (Xue et al., 1992). The largest was estimated by restriction mapping to be 1.7 kb in length and, although adjudged by Northern blotting to be incomplete, coded for a functional endoglucanase (pL9; Figure 1). To obtain the complete sequence of the protein-coding region, the incomplete cDNA cloned in pL9 was extended by 5′-RACE.

Full-length *celB* cDNA contained an open reading frame (ORF) of 1422 nucleotides, coding for a polypeptide of 473 amino acids with M_r 53070. The nucleotide sequence of *celB* and the deduced primary structure of the encoded protein are shown in Figure 2. Identification of the ORF was supported by the observation that sequences 5′ and 3′ of the proposed protein-coding region were extremely rich in A and T residues, a feature which is highly characteristic of the non-coding regions of the *Neocallimastix* genome (Gilbert et al., 1992). The presence of translational stop codons in all three reading frames upstream of the proposed start codon, and the absence of further ATG codons upstream of the ORF are evidence for the validity of the putative translational start. *In vitro* transcription/translation of recombinant pBluescript containing full-length *celB* in-phase with the *lacZ′* translation initiation codon revealed a major 30000-M_r polypeptide, corresponding to β-lactamase, and a number of higher-molecular-mass polypeptides, the largest of which had an M_r of 52000 (results not shown). This is consistent with the predicted size of CelB based on the nucleotide sequence of *celB*.

In order to investigate the occurrence of introns within the protein-coding region of *celB*, we isolated and sequenced a 3 kbp genomic fragment that spanned the *celB* gene. The nucleotide sequences of genomic *celB* and the full-length cellulase cDNA were essentially identical, indicating that there were no introns in the genomic sequence. Without exception, previously described cellulases from aerobic filamentous fungi have contained introns (Knowles et al., 1989). The finding that *celB* is devoid of introns is unusual and could be regarded as the first line of evidence that *celB* is prokaryotic in origin.

Structure of CelB

Inspection of the deduced primary sequence of CelB revealed a number of interesting and informative features, the first of these being a region between residues 366 and 387 with the sequence SIKTRTTIRTTTTTTSQSQPT. In other microbial cellulases and hemicellulases, sequences such as this, rich in the hydroxy amino acids serine and threonine, serve as hinges or linkers between different functional domains, and are therefore strongly indicative of modular architecture (Gilkes et al., 1991). Down-stream of the linker sequence, a repeated sequence of some 30 residues comprised the C-terminus of CelB. Each element of the repeat sequence had significant identity to a similar-sized repeated sequence occurring in an identical position of xylanase A from the same organism (Gilbert et al., 1992). The degree of identity between the repeated regions of the two enzymes was 63% at the protein level. Short tandem repeats such as these are not uncommon in microbial polysaccharide hydrolases, and some progress has been made in understanding their function. For example, in the anaerobic, thermophilic bacterium *Clostridium thermocellum* a 23-amino-acid reiterated domain, found in several endoglucanases and in a xylanase that are all subunits of the multiprotein cellulosome complex, has been shown to mediate attachment to a large non-catalytic scaffolding protein during cellulosome assembly (Tokatlidis et al., 1991; Fujino et al., 1992). In the light of biochemical evidence supporting the view that the cellulase system of *Neocallimastix* spp. constitutes a large multiprotein complex (Wilson and Wood, 1992), it is tempting to speculate that the terminal repeats in CelB and XylA fulfil an analogous function.

Deletion of 231 nucleotides from the 3′ end of *celB* in pL9 (Figure 1) produced pL9Exo1 (Figure 1) and resulted in a truncated form of CelB, which, although lacking the C-terminal domain, attacked the same range of substrates and displayed the

Table 1   Percentage sequence identity between CelB from the anaerobic fungus *N. patriciarum* and other family A cellulases

All of the enzymes are endo-β-1,4-glucanases.

Organism	Enzyme	Identity (%)	Size of overlap	Reference
Bacteroides ruminicola	EgI	35.8	307	Matsushita et al. (1990)
Butyrivibrio fibrisolvens	End1	38.2	340	Berger et al. (1989)
Clostridium cellulolyticum	CelA	42.6	263	Faure et al. (1990)
Clostridium thermocellum	CelE	44.6	350	Hall et al. (1988)
Ruminococcus albus F-40	EgI1	41.4	345	Ohmiya et al. (1989)
Ruminococcus albus SY3	CelA	43.5	345	Poole et al. (1990)
Ruminococcus albus SY3	CelB	41.5	347	Poole et al. (1990)

Figure 3   Comparison of the amino acid sequence (between residues 51 and 365) of CelB from *N. patriciarum* with the sequences of the catalytic domains of *Butyrivibrio fibrisolvens* End1 (Berger et al., 1989), *Clostridium cellulolyticum* CelA (Faure et al., 1990), *Clostridium thermocellum* CelE (Hall et al., 1988), *Ruminococcus albus* EgI1 (Ohmiya et al., 1989) and *R. albus* CelA and CelB (Poole et al., 1990)

Residues identical in all sequences are boxed.

same ratio of activities against barley β-glucan, CMC and xylan as the full-length form. Further truncation of the 3' region of the gene generated derivatives of CelB that were inactive (Figure 1). Thus the region of CelB extending from the N-terminus to immediately upstream of the putative linker constitutes a catalytic domain of some 365 residues. When compared with the known sequences of other cellulases and xylanases, the N-terminal domain between residues 51 and 365 of CelB exhibited the highest percentage sequence identity (36% to 45%) with the catalytic domains of seven different bacterial cellulases belonging to family A as defined by Henrissat et al. (1989) and reviewed by Gilkes et al. (1991) (Table 1 and Figure 3). Interestingly, five of the endoglucanases were from anaerobic cellulolytic bacteria that, like *N. patriciarum*, normally inhabit the rumen. Percentage sequence identity between CelB and the remaining family A endoglucanases, including endoglucanase II from the aerobic fungus *T. reesei*, was 26% or less. The evolution of rumen cellulases and xylanases is an interesting question as these

enzymes have evolved within an essentially closed ecosystem. Although the rumen is open to inoculation by micro-organisms associated with the diet, the proliferation of exogenous organisms in the rumen is, in general, prevented by intense competition for nutrients and the prevailing anaerobic conditions. The fact that previously reported cellulases of rumen bacteria belong almost exclusively to family A (Gilkes et al., 1991) suggests that the rumen cellulases have evolved from a much smaller pool of progenitor sequences than the cellulases of organisms occupying other natural ecosystems, although it could be argued that recovery of cellulase genes from rumen bacteria has been biased in favour of a few readily isolated genes. Furthermore, the observed homology between the catalytic domain of CelB from *N. patriciarum*, and similar domains from the cellulases of three different rumen bacteria suggests that horizontal transfer of genes between rumen bacteria and fungi may have occurred during their occupancy of the rumen. The lack of similarity between the bacterial enzymes and the homologous C-terminal
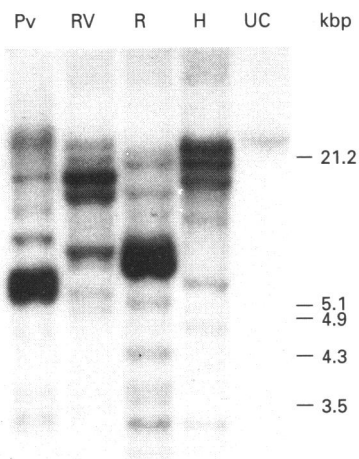
Figure 4   Southern hybridization of celB cDNA excised from pL9 (Figure 1), with PvuII (Pv), EcoRV (RV), EcoRI (R) and HindIII (H) digests, and with undigested (UC) N. patriciarum genomic DNA
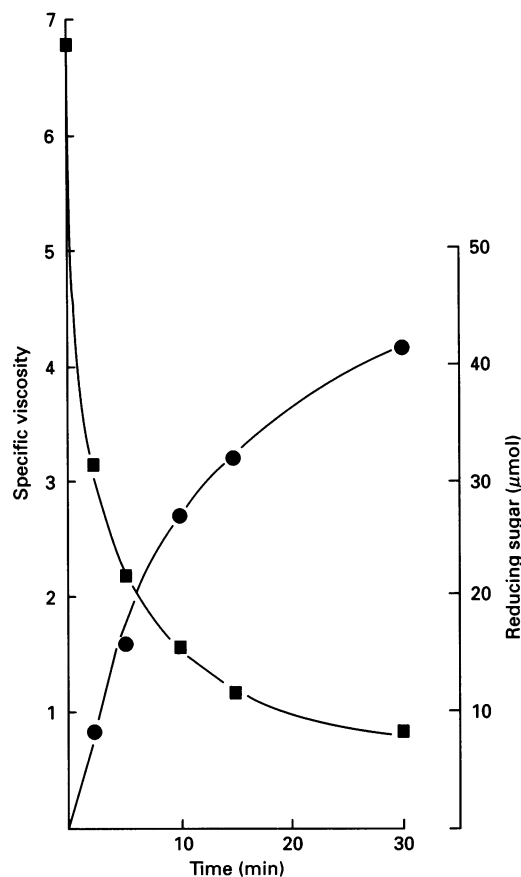

Figure 5   Effect of CelB on the viscosity of CMC (0.75%, w/v) in PC buffer, pH 6.5, at 37 °C

Specific viscosity (■) and reducing sugar (●) were measured as described in the Materials and methods section.

repeats of CelB and XylA is evidence that this domain has been acquired or has evolved subsequently to serve a function that, although not essential for catalytic activity, is specific to N. patriciarum.

Although ineffective as a secretory signal in E. coli, the N-terminal residues of CelB had the appearance of a typical signal peptide, with a positively charged N-terminus preceding a region containing more than 12 hydrophobic residues.

Reiteration of celB

A labelled probe consisting of the celB cDNA excised from pL9 was hybridized under high-stringency conditions with N. patriciarum genomic DNA that had been restricted with PvuI, EcoRV, EcoRI and HindIII. In each case, the probe hybridized to multiple bands (Figure 4) suggesting the following possibilities; (i) celB is present in the N. patriciarum genome as multiple copies; (ii) the celB cDNA isolated in pL9 and sequenced here is one of a number of non-identical but closely related celB homologues occurring within the N. patriciarum genome; (iii) part of celB codes for a protein domain that is conserved between cellulases and hemicellulases of N. patriciarum, and is therefore carried by a number of different cellulase and hemicellulase genes which are borne on distinct genomic fragments.

Although it is not possible at present to discount any of these explanations, either of the latter two possibilities would seem most likely, as the differing intensities resulting from hybridization of the probe to genomic fragments (Figure 4) would suggest that it recognizes similar as well as identical sequences.

Characteristics of CelB

CelB produced by E. coli XL1-blue harbouring the full-length celB gene cultured on LB agar, hydrolysed CMC, xylan, barley

β-glucan and lichenan that had been incorporated into an overlay at 0.5% (w/v); laminarin, arabinan, galactan and mannan were not hydrolysed. CelB from an E. coli cell-free extract had a maximal specific activity against CMC of 1.2 units/mg of protein, and hydrolysed it in a manner typical of many endoglucanases by releasing reducing sugar and promoting a rapid decline in viscosity (Figure 5). The major products generated during prolonged hydrolysis of CMC were cellobiose and cellotriose. The substrate range for CelB was similar to that of the endo-glucanases CelI and CelF from C. thermocellum (Hazlewood et al., 1993); highest activity was observed against the mixed-linkage β-glucan from barley endosperm cell walls, while a much lower level of activity was seen against lichenan and soluble xylan (Table 2). Activities against CMC and xylan were 20.6% and 2.2% respectively of the activity against barley β-glucan, and this ratio of activities against the three substrates was maintained for full-length (pL9; Figure 1) and truncated (pL9Exo1; Figure 1) CelB, indicating that the repeated regions contained in the C-terminal domain are not involved in catalytic activity. The observed substrate specificity was reflected in the apparent K_m value for barley β-glucan (0.24 mg/ml), which was substantially less than the K_m for either CMC (3 mg/ml) or soluble xylan (8 mg/ml), indicating a much higher affinity of the enzyme for this substrate than for either CMC or xylan (Table 2). More recalcitrant forms of cellulose such as acid-swollen cellulose, filter paper and Avicel were not hydrolysed, even after

Table 2   Substrate specificity of CelB

Abbreviations: nd, not determined; NA, no activity.

Substrate	Relative activity	K_m (mg/ml)
Barley β-glucan	100%	0.24
CMC	20.6%	3.0
Soluble xylan	2.2%	8.0
Lichenan	0.3%	nd
Laminarin	NA	
Mannan	NA	
Galactan	NA	
Arabinan	NA	
Acid-swollen cellulose	NA	
Filter paper	NA	
Avicel	NA	

Table 3   Hydrolysis of cellulo-oligosaccharides by CelB

Abbreviations: G1, glucose; G2, cellobiose; tr, trace.

Substrate	Relative rate of hydrolysis (%)	Products after 5 h (mol %)			
		G1	G2	G3	G4
Cellotriose (G3)	2	tr	tr		
Cellotetraose (G4)	42	6	69	25	–
Cellopentaose (G5)	100	1	47	46	6

prolonged incubation with high concentrations of CelB. The rate of hydrolysis, by CelB, of water-soluble cellulo-oligosaccharides increased with increasing chain length from cellotriose to cellopentaose; no activity was observed against cellobiose. The endoglycolytic mode of action of CelB was confirmed by analysing the products obtained during prolonged incubation of the enzyme with cellulo-oligosaccharides at 2 mM concentration (Table 3). Cellotetraose was converted into mainly cellobiose, and cellopentaose yielded equal quantities of cellotriose and cellobiose. When the initial concentration of cellopentaose was increased to 6 mM, cellulo-oligosaccharides with retention volumes greater than cellopentaose were observed in low yield (results not shown). On addition of glucose or cellobiose at 20 mM concentration, the accumulation of higher oligomers was diminished. These results suggest that at high substrate concentration CelB may act as a transglycosylase, transferring the product of cellopentaose cleavage to a second substrate molecule, rather than to a water molecule. The activity of CelB against CMC was pH-dependent and showed a maximum around pH 6.5.

Conclusions

In conjunction with an earlier study (Gilbert et al., 1992), this report suggests that cellulases and hemicellulases of the rumen anaerobic fungi have a modular structure incorporating a highly conserved non-catalytic domain. The precise role of this domain is the focus of current work in our laboratories. Biochemical analysis of the enzyme (CelB) encoded by the *celB* gene from *N. patriciarum* indicates that it is an endo-acting β-glucanase, principally active against barley β-glucan, but also having significant activity against CMC, soluble xylan and cellulo-oligosaccharides. The absence of introns from the *celB* gene, and the observed homology between the catalytic domain of CelB and a number of family A endoglucanases from anaerobic rumen bacteria, though not definitive, could be regarded as the first line of evidence that the evolution of cellulases within the herbivore digestive tract has involved gene transfer between prokaryotes and eukaryotes.

REFERENCES

Berger, E., Jones, W. A., Jones, D. T. and Woods, D. R. (1989) Mol. Gen. Genet. **219**, 193–198

Brownlee, A. G. (1988) Fungal Gen. News. **35**, 8–9

Faure, E., Belaich, A., Bagnara, C., Gaudin, C. and Belaich, J.-P. (1990) Gene **65**, 51–58

Fujino, T., Béguin, P. and Aubert, J.-P. (1992) FEMS Microbiol. Lett. **94**, 165–170

Gilbert, H. J., Hazlewood, G. P., Laurie, J. I., Orpin, C. G. and Xue, G.-P. (1992) Mol. Microbiol. **6**, 2065–2072

Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. and Warren, R. A. J. (1991) Microbiol. Rev. **55**, 303–315

Hall, J., Hazlewood, G. P., Barker, P. J. and Gilbert, H. J. (1988) Gene **69**, 29–38

Hazlewood, G. P., Romaniec, M. P. M., Davidson, K., Grépinet, O., Béguin, P., Millet, J., Raynaud, O. and Aubert, J.-P. (1988) FEMS Microbiol. Lett. **51**, 231–236

Hazlewood, G. P., Davidson, K., Laurie, J. I., Romaniec, M. P. M. and Gilbert, H. J. (1990) J. Gen. Microbiol. **136**, 2089–2097

Hazlewood, G. P., Davidson, K., Laurie, J. I., Huskisson, N. S. and Gilbert, H. J. (1993) J. Gen. Microbiol. **139**, 307–316

Henrissat, B., Claeyssens, M., Tomme, P., Lemesle, L. and Mornon, J.-P. (1989) Gene **81**, 83–95

Kemp, P., Lander, D. J. and Orpin, C. G. (1984) J. Gen. Microbiol. **130**, 27–37

Knowles, J. K. C., Lehtovaara, P. and Teeri, T. T. (1987) Trends Biotechnol. **5**, 255–261

Knowles, J. K. C., Teeri, T. T., Pentilla, M., Saloheimo, M., Harkki, A. and Lehtovaara, P. (1989) in Enzyme Systems for Lignocellulose Degradation (Coughlan, M. P., ed.), pp. 51–56, Elsevier Applied Science, London and New York

Lamed, R. and Bayer, E. A. (1988) Adv. Appl. Microbiol. **33**, 1–46

Matsushita, O., Russell, J. B. and Wilson, D. B. (1990) J. Bacteriol. **172**, 3620–3630

Miller, G. L. (1959) Anal. Chem. **31**, 426–428

Murphy, G. and Kavanagh, T. (1988) Nucleic Acids Res. **16**, 5198

Ohmiya, K., Kajino, T., Kato, A. and Shimizu, S. (1989) J. Bacteriol. **171**, 6771–6775

Orpin, C. G. and Joblin, K. N. (1988) in The Rumen Microbial Ecosystem (Hobson, P. N., ed.), pp. 129–150, Elsevier Applied Science, London and New York

Orpin, C. G. and Munn, E. A. (1986) Trans. Br. Mycol. Soc. **86**, 178–180

Poole, D. M., Hazlewood, G. P., Laurie, J. I., Barker, P. J. and Gilbert, H. J. (1990) Mol. Gen. Genet. **223**, 217–223

Rixon, J. P., Ferreira, L. M. A., Durrant, A. J., Laurie, J. I., Hazlewood, G. P. and Gilbert, H. J. (1992) Biochem. J. **285**, 947–955

Romaniec, M. P. M., Clarke, N. G. and Hazlewood, G. P. (1987) J. Gen. Microbiol. **133**, 1297–1307

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) in Molecular Cloning: A Laboratory Manual, 2nd Edn., Cold Spring Harbor, New York

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467

Sedmak, J. J. and Grossberg, S. E. (1977) Anal. Biochem. **69**, 544–552

Teunissen, M. J., de Kort, G. V. M., Op den Camp, H. J. M. and Huis in't Veld, J. H. J. (1992) J. Gen. Microbiol. **138**, 1657–1664

Tokatlidis, K., Salamitou, S., Béguin, P., Dhurjati, P. and Aubert, J.-P. (1991) FEBS Lett. **291**, 185–188

Wilson, C. A. and Wood, T. M. (1992) Enzyme Microb. Technol. **14**, 258–264

Wood, T. M. (1992) Biochem. Soc. Trans. **20**, 46–53

Xue, G.-P., Orpin, C. G., Gobius, K. S., Aylward, J. H. and Simpson, G. D. (1992) J. Gen. Microbiol. **138**, 1413–1420