

The non-catalytic cellulose-binding domain of a novel cellulase from *Pseudomonas fluorescens* subsp. *cellulosa* is important for the efficient hydrolysis of Avicel

Judith HALL,* Gary W. BLACK,* Luis M. A. FERREIRA,† Sarah J. MILLWARD-SADLER,* Bassam R. S. ALI,‡
Geoffrey P. HAZLEWOOD‡ and Harry J. GILBERT*§

*Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K., †Faculdade de Medicina Veterinária, Rua Gomes Freire, 1199 Lisboa Codex, Portugal, and ‡Department of Cellular Physiology, The Babraham Institute, Babraham, Cambridge CB2 4AT, U.K.

A genomic library of *Pseudomonas fluorescens* subsp. *cellulosa* DNA, constructed in λ ZAPII, was screened for carboxymethyl-cellulase activity. The pseudomonad insert from a recombinant phage which displayed elevated cellulase activity in comparison with other cellulase-positive clones present in the library, was excised into pBluescript SK⁻ to generate the plasmid pC48. The nucleotide sequence of the cellulase gene, designated *celE*, revealed a single open reading frame of 1710 bp that encoded a polypeptide, defined as endoglucanase E (CelE), of M_r 59 663. The deduced primary structure of CelE revealed an N-terminal signal peptide followed by a 300-amino-acid sequence that exhibited significant identity with the catalytic domains of cellulases belonging to glycosyl hydrolase Family 5. Adjacent to the catalytic domain was a 40-residue region that exhibited strong sequence identity to non-catalytic domains located in two other endoglucanases and a xylanase from *P. fluorescens*. The C-terminal 100 residues of CelE were similar to Type-I cellulose-

binding domains (CBDs). The three domains of the cellulase were joined by linker sequences rich in serine residues. Analysis of the biochemical properties of full-length and truncated derivatives of CelE confirmed that the enzyme comprised an N-terminal catalytic domain and a C-terminal CBD. Analysis of purified CelE revealed that the enzyme had an M_r of 56 000 and an experimentally determined N-terminal sequence identical to residues 40–54 of the deduced primary structure of full-length CelE. The enzyme exhibited an endo mode of action in hydrolysing a range of cellulosic substrates including Avicel and acid-swollen cellulose, but did not attack xylan or any other hemicelluloses. A truncated form of the enzyme, which lacked the C-terminal CBD, displayed the same activity as full-length CelE against soluble cellulose and acid-swollen cellulose, but exhibited substantially lower activity than the full-length cellulase against Avicel. The significance of these data in relation to the role of the CBD is discussed.

INTRODUCTION

It is widely recognized that many cellulases have a modular structure, comprising a catalytic domain linked to a non-catalytic cellulose-binding domain (CBD) by flexible linker sequences rich in hydroxyamino acids [1]. Based on sequence identity there are at least four families of CBDs in which all members of a given family are thought to have evolved from a common progenitor sequence [2]. The role of the CBD in enzyme function has been investigated in several enzymes. Although these domains do not appear to affect the activity of fungal cellobiohydrolases against soluble substrates, they significantly enhance the capacity of the enzymes to hydrolyse crystalline cellulose [3]. Knowles et al. [4] suggested that fungal CBDs, in binding to cellulose chains, disrupt the interchain hydrogen bonding and funnel the glucose polymers into the active site of the enzyme. Bacterial CBDs can either exhibit a high affinity for their ligand, such as the Type-I domains [5], or a low affinity for cellulose as demonstrated by the CBD of endoglucanase E (CelE) from *Clostridium thermocellum* [6]. It appears that this latter domain decreases the specific activity and increases the affinity of the enzyme for soluble cellulosic substrates such as β -glucan and carboxymethylcellulose (CMC). The enzyme exhibits no activity against insoluble cellulose. In contrast, it is clear that at least

some high-affinity Type-I CBDs enhance the capacity of cellulases to hydrolyse crystalline cellulose [7,8].

Studies in our laboratories have focused on the molecular architecture of plant cell wall hydrolases of *Pseudomonas fluorescens* subsp. *cellulosa*. This aerobic soil bacterium expresses multiple endoglucanases and xylanases in addition to enzymes that cleave xylan side-chain substituents [9,10]. The pseudomonad also expresses other plant cell wall hydrolases that attack mannan, galactan and arabinan ([11]; H. J. Gilbert, unpublished work). To date, the primary structures of two endoglucanases [12,13], two xylanases [14,15], an arabinofuranosidase [15], an acetyl xylan esterase [5] and a cellodextrinase [16] have been determined. Each enzyme has a modular structure comprising of a catalytic domain, a middle non-catalytic domain of unknown function and a terminal Type-I CBD. The location of CBDs in a range of plant cell wall hydrolases that do not attack cellulose led to the suggestion that in *P. fluorescens* subsp. *cellulosa*, CBDs could provide a general mechanism by which a consortium of hydrolases accumulate on the surface of the plant cell wall resulting in an enhancement, through their close proximity, of the synergism exerted between the enzymes [12,17]. To investigate this possibility it is necessary to analyse the relationship between the structure and function of a cellulase from *P. fluorescens* subsp. *cellulosa*. Although two cellulase genes from this bacterium

Abbreviations used: BMCC, bacterial microcrystalline cellulose; CBD, cellulose-binding domain; CelE, endoglucanase E; CelE', truncated CelE; CMC, carboxymethylcellulose; LB, Luria broth; ORF, open reading frame; PC, 50 mM potassium phosphate/12 mM citric acid buffer (pH 6.5).

§ To whom correspondence should be addressed.

The nucleotide sequence data for *celE* reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X86798.

have previously been cloned, they were poorly expressed in *Escherichia coli*, and we were unable to purify sufficient quantities of the encoded enzymes to evaluate the role of their CBDs in cellulose cleavage. In this report we describe the relationship between the structure and function of a third *Pseudomonas* cellulase that was successfully overexpressed in *E. coli*. The data revealed that the enzyme had a modular structure consisting of an N-terminal catalytic domain, a small non-catalytic domain also found in other pseudomonad cellulases, and a C-terminal Type-I CBD. The cellulase displayed a typical endo-mode of action and exhibited significant activity against insoluble and substantially crystalline cellulose. Removal of the CBD did not affect the activity of the enzyme against soluble or amorphous cellulose, but there was a reduction in the activity of the truncated cellulase against Avicel. It is apparent, therefore, that Type-I CBDs from *P. fluorescens* subsp. *cellulosa* cellulases can play a direct role in the hydrolysis of substantially crystalline substrates; the domain does not simply provide a mechanism by which a repertoire of hydrolases can bind to the plant cell wall.

MATERIALS AND METHODS

Bacterial strains, culture conditions and vectors

E. coli JM83 was cultured in Luria broth (LB) at 37 °C. Phage were propagated in *E. coli* grown on NZY medium. Media were supplemented with 100 µg/ml of ampicillin and 2 µg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside to select for *E. coli* transformants and recombinants, respectively. To identify recombinant *E. coli* strains that expressed cellulase activity, bacterial colonies, grown on LB-agar supplemented with medium-viscosity CMC, were stained with 1% Congo Red for 5 min and destained with 1 M NaCl. Cellulase-expressing *E. coli* colonies were surrounded by clear haloes against a red background [18]. The phage and plasmids employed in this work were λZAPII, pBluescriptSK⁻ (Stratagene) and the pMTL series of vectors [19].

Recombinant DNA methodology

Plasmid DNA was purified using Qiagen columns (Hybaid) following the instructions of the manufacturer. Transformation of *E. coli*, the general use of nucleic acid-modifying enzymes and Southern hybridization were as described by Sambrook et al. [20]. DNA probes were labelled with [α -³²P]dCTP by random priming. *Pseudomonas* genomic inserts in recombinants of λZAPII expressing cellulase activity were excised into pBluescript SK⁻ using VCS-M13 helper phage.

To sequence *celE* a series of nested deletions of the fragment carrying the gene was constructed using the exonuclease III/S1-nuclease kit supplied by Promega. The nested deletions were sequenced by the dideoxy chain termination method of Sanger et al. [21] both manually, using Sequenase (United States Biochemicals) and spin dialysis to purify denatured DNA [22], and by an automated method in which the nested deletions were sequenced using a PRISM ready reaction dye-deoxy terminator cycle, supplied by Applied Biosystems Inc., and the products analysed using an Applied Biosystems automated sequencer (model 373A). Sequences were compiled and ordered using the computer programs of Staden [23]. Both strands of *celE* were sequenced.

Assays

Avicel PH105 was obtained from FMC Corporation; bacterial microcrystalline cellulose (BMCC; [8]) was a generous gift from

Dr. Bernard Henrissat. Amorphous (acid-swollen) cellulose was prepared from crystalline cellulose as described by Wood [24].

Cellulase, xylanase and mannanase activities were assayed at 37 °C, unless otherwise stated, in 50 mM potassium phosphate/12 mM citric acid buffer (pH 6.5) (PC). The polysaccharides were incubated with enzyme at a final concentration of 0.2% (unless otherwise stated), and the level of reducing sugar released was determined [25]. One unit of enzyme activity was defined as the amount of enzyme that releases 1 µmol of reducing sugar per min. When using insoluble substrates, the residual polysaccharide was removed by centrifugation, before analysing the liquid phase for reducing sugar. Hydrolysis of red debranched arabinan and azurin-cross-linked galactan (both from Megazyme) was determined by culturing *E. coli* harbouring the *celE* gene on substrate-containing agar and looking for the appearance of clear halos against a red background or the solubilization of the blue AZCL dye, respectively. Cellulo-oligosaccharides (200 ng/ml) were incubated with CelE and aliquots were removed at regular intervals, boiled for 5 min and applied to a Dionex PA1 column. Separation of cellulo-oligosaccharides was achieved using a gradient elution of 0–75 mM sodium acetate (in 100 mM NaOH) from 0–16 min, and isocratic elution with 75 mM sodium acetate (in 100 mM NaOH) from 16–25 min. The sugars were detected with a pulsed amperometer. The α or β configuration of the anomeric carbon of celotriose, generated by the hydrolysis of cellopentaose by CelE, was determined as described previously [26], except that the two anomers of celotriose were separated on two Spherisorb HPLC columns (S5 ODS2; 25 cm × 4.6 mm) run in series. The oligosaccharides were detected as above after the addition of post-column reagent (300 mM NaOH). Protein was determined by the dye binding method of Sedmak and Grossberg [27] using BSA as the standard. The sizes of polypeptides were determined by SDS/PAGE [28]. Viscometric assays were performed as described by Zhou et al. [29]. The N-terminal sequence of CelE was determined by the method of Hunkapillar et al. [30].

Purification of CelE and CelE'

E. coli XL1-blue harbouring full-length *celE* in plasmid pC48 was cultured overnight in LB broth (3 l) containing ampicillin (100 µg/ml). A cell-free extract was prepared by sonicating the harvested cells in 20 mM Tris/HCl, pH 8.0, and CelE contained in the cell-free extract was precipitated between 40% and 60% ammonium sulphate saturation. Precipitated protein was redissolved in a minimum volume of 20 mM Tris/HCl, pH 7.2, and applied to a Sephacryl S-100 HR column (25 mm × 610 mm) which was eluted with 20 mM Tris/HCl, pH 7.2, containing 150 mM NaCl at a flow rate of 0.5 ml/min. Purification of full-length CelE was greatly facilitated by the apparent interaction of CelE with the Sephacryl matrix, which resulted in retention of the enzyme on the column during the passage of almost a full column volume of the elution buffer.

A truncated derivative of CelE (CelE'), having the same N-terminal amino acid sequence as the full-length enzyme, but lacking the C-terminal CBD, was generated by treating CelE with trypsin. Purified full-length CelE was incubated for 16 h at 37 °C in Puck's saline containing trypsin (0.125 mg/ml) and EDTA (50 µg/ml). SDS/PAGE revealed that CelE (M_r 56000) had been quantitatively converted into a smaller polypeptide (CelE') with an M_r of 33000. After dialysing against 20 mM Tris/HCl, pH 8.0, CelE' was purified by FPLC on Mono Q (Pharmacia) using a 1 ml analytical column and a 40 ml gradient containing from 0 to 1 M KCl in 20 mM Tris/HCl, pH 8.0, at a flow rate of 60 ml/h.

RESULTS

Isolation and characterization of *celE*

The primary objective of this study was to evaluate whether CBDs from *P. fluorescens* subsp. *cellulosa* cellulases enhance cellulolysis. This was not feasible with the two previously cloned *Pseudomonas* endoglucanase genes, as the gene products were synthesized at low levels in *E. coli*. To determine whether it was possible to isolate a *Pseudomonas* cellulase gene that can be highly expressed in *E. coli*, a genomic library of *P. fluorescens* subsp. *cellulosa* DNA, constructed in λ ZAPII, was screened for clones that exhibited carboxymethylcellulase activity. A recombinant phage, which appeared to produce high levels of cellulase activity, was selected for further study. The pseudomonad genomic insert was excised from the recombinant phage into pBluescriptSK⁻ to generate the plasmid pC48. Cross-hybridization experiments showed that the cellulase gene (designated *celE*) encoded by pC48 exhibited no detectable sequence identity with either of the two previously characterized *P. fluorescens* endoglucanase genes. A restriction map of the *Pseudomonas* DNA in pC48 is displayed in Figure 1. Deletion and subcloning experiments located the position of *celE* within the pseudomonad insert in pC48 between map positions 1.7 and 3.5.

Characterization of CelE

The *celE* gene product, designated CelE, was purified from *E. coli* cell-free extract and its biochemical properties evaluated. The enzyme had an M_r of 56000 (Figure 2) and a broad pH optimum between pH 4.5 and 8.0 (data not shown). The cellulase was also unusually thermostable for an enzyme derived from a mesophilic bacterium; significant thermal inactivation of CelE occurred at temperatures in excess of 60 °C. The N-terminal sequence of the cellulase was DVAPLSVQGNKILAN. Analysis of the substrate specificity of CelE revealed that the enzyme hydrolysed a range of cellulosic substrates, exhibiting maximum activity against barley β -glucan (Table 1). The major products generated from CMC hydrolysis were celotriose and cellobiose (data not shown). Analysis of the activity of the enzyme against cellulo-oligosaccharides showed that the cellulase had no activity against cellobiose or celotriose, although it hydrolysed longer-chain cellulo-oligosaccharides such as cello-

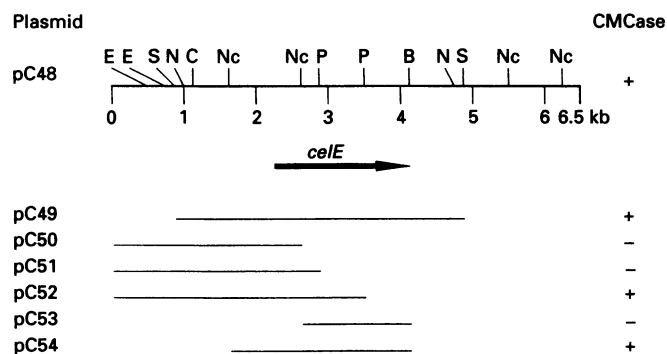


Figure 1 Restriction map of recombinant plasmids containing *celE*

Cleavage sites for restriction endonucleases are as follows: *Bcl*I (B), *Cla*I (C), *Eco*RV (E), *Nco*I (Nc), *Nru*I (N), *Pst*I (P) and *Sph*I (S). The location and orientation of *celE* is indicated by an arrow. In pC53 the 5'-truncated derivative of *celE* was in-frame with the *lacZ* translational start codon of the vector (pMTL22). The catalytic activity of CelE encoded by the recombinant plasmids is shown by + or -.

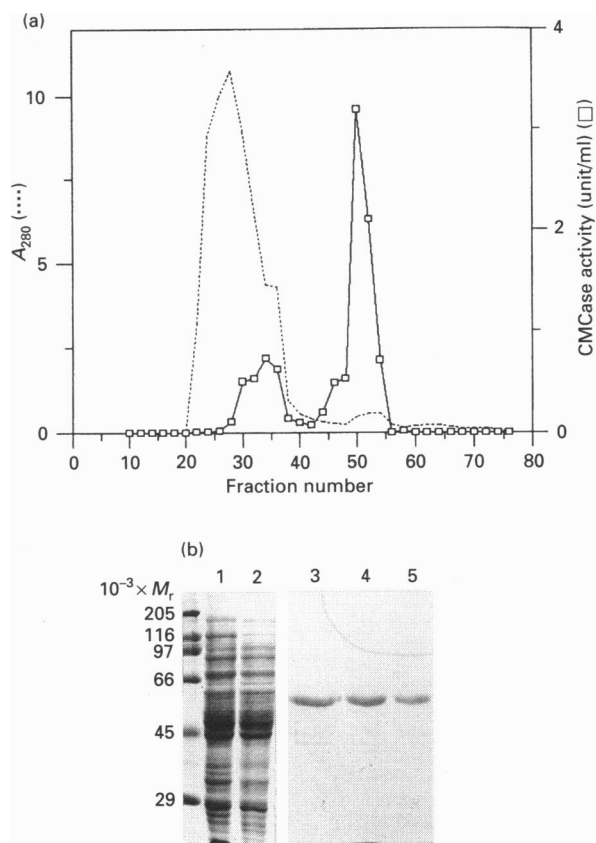


Figure 2 Purification of full-length CelE by gel filtration chromatography on Sephacryl S-100 HR

(a) Column chromatography was performed as described in the Materials and methods section. Endoglucanase activity (\square) and protein concentration (.....) were measured by quantifying reducing sugar release after fractions were incubated with CMC, and by monitoring A_{280} , respectively. (b) SDS/PAGE analysis using a 10% (w/v) polyacrylamide gel of cell-free extract from *E. coli* containing pC48 (lane 1), the ammonium sulphate fraction (lane 2) and Sephacryl S-100 HR column fractions 50 (lane 3), 51 (lane 4) and 53 (lane 5). Protein molecular mass markers were run on the extreme left.

Table 1 Substrate specificity of CelE

The specific activity of purified full-length CelE against barley β -glucan was 350 units/mg of protein. Abbreviation: NA, no activity detectable.

Substrate	Relative activity (%)
Barley β -glucan	100
Carboxymethylcellulose	45
Lichenan	120
Xylan	NA
Galactan	NA
Arabinan	NA
Galactomannan	NA
Laminarin	NA
Mannan	NA

tetraose and cellopentaose displaying 2.5 times higher activity against the larger oligosaccharide. The major products generated from cellopentaose were cellobiose and celotriose, while cello-tetraose was hydrolysed primarily to cellobiose (data not shown).

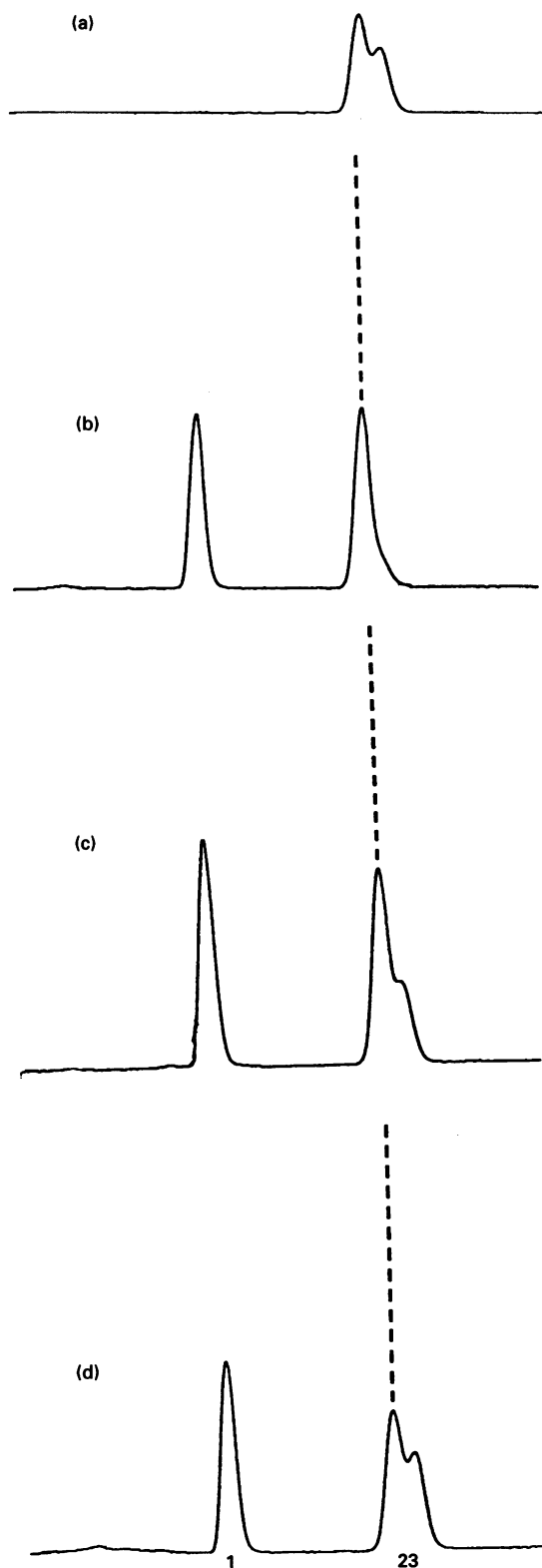


Figure 3 Anomeric configuration of cellotriose generated by CelE

Cellopentaose was incubated with CelE for 5 min (b) and 1 h (c) and the products generated were immediately analysed by HPLC as described in the Materials and methods section. In (d) the oligosaccharides generated in (b) were boiled for 5 min, to mutarotate the β anomer, prior to HPLC analysis. (a) Contains cellotriose. The retention times of cellobiose (1), the β anomer of cellotriose (2) and α anomer of cellotriose (3) were 6.94 ± 0.06 min ($n = 5$), 10.48 ± 0.03 min ($n = 6$) and 10.89 ± 0.03 min ($n = 4$), respectively.

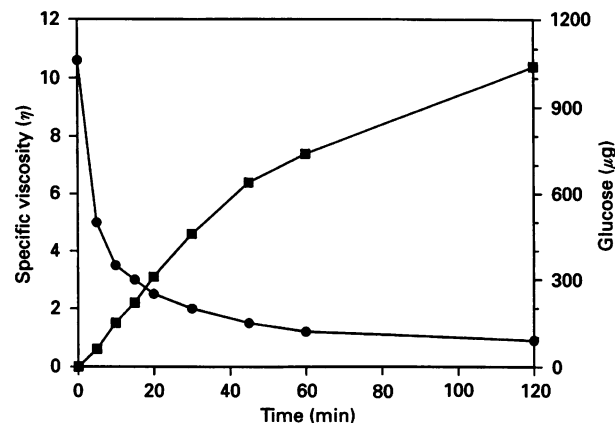


Figure 4 Effect of CelE on the viscosity of CMC

The substrate (2%, w/v) was incubated with CelE in PC buffer, pH 6.5, at 37 °C. Specific viscosity (●) and reducing sugar (■) were measured as described in the Materials and methods section.

The cellotriose released from cellopentaose by CelE (5 min enzyme reaction) co-chromatographed with the β anomer of cellotriose (Figure 3). After incubating these reaction products for a further 1 h at room temperature, cellotriose had mutarotated and significant quantities of the α anomer were observed, confirming that the separation method resolves the two anomeric forms of the trisaccharide. Thus, the cellulase retains the configuration of the anomeric carbon during bond cleavage. CelE therefore hydrolyses glycosidic bonds by a double displacement mechanism. To assess whether CelE was an endo-acting enzyme the reduction of viscosity and release of reducing sugar were determined during incubation with barley β -glucan. The data (Figure 4) showed that CelE exhibited a typical endo-mode of action, generating a large decrease in viscosity while releasing only small amounts of reducing sugar. The enzyme exhibited no detectable activity against a range of non-cellulosic plant structural polysaccharides including xylans, mannan, galactan and arabinan, or against β -1,3-linked glucose polymers such as laminarin (Table 1). Although these results indicate that CelE is a typical endo- β -1,4-glucanase, the enzyme displayed significant activity against both acid-swollen cellulose and crystalline cellulose (Avicel) (Table 2), substrates which are not normally hydrolysed to such a great extent by the majority of endo-glucanases.

Nucleotide sequence of *celE* and the primary structure of CelE

The nucleotide sequence of the region of pC48 containing *celE* was determined in both strands. The data revealed a single open reading frame (ORF) of 1710 bp encoding a protein of M_r 59663 (Figure 5). The codon usage of the ORF was very similar to other *P. fluorescens* plant cell wall hydrolases [5]. Upstream (9 bp) of the putative ATG translational start codon is the sequence AGAGA, which exhibits similarity with the ribosome binding sequences of genes from Gram-negative bacteria. Residues 40–54 of the deduced primary structure of the encoded protein were identical to the experimentally determined N-terminus of mature CelE. The sequence of CelE preceding the mature N-terminus conforms to a typical prokaryotic signal peptide; the hydrophilic N-terminus of the full-length enzyme contains five basic residues and is followed by 15 small hydrophobic residues capable of

Table 2 Relative activity of full-length and truncated derivatives of CelE against cellulose

Abbreviation: ASC, acid-swollen cellulose.

CelE structure	Size (M_r)	N-terminal sequence	Enzyme activity (katal/mol) ^a		
			CMC	ASC	Avicel
Full-length	56 000	DVAPLSVQGNKILAN	70	39	0.51
Truncated ^b	34 000	DVAPLSVQGNKILAN	85	36	0.13

^a katal/mol; molecules of product generated/s per molecule of enzyme.^b Full-length CelE was treated with trypsin and purified by FPLC as described in the Materials and methods section.

forming an α -helix. The M_r of recombinant CelE is similar to the deduced M_r of the mature cellulase (55 768).

The primary structure of CelE revealed domains observed in other microbial plant cell wall hydrolases. The 300 N-terminal residues of mature CelE exhibited significant sequence identity with the catalytic domains of endoglucanases belonging to glycosyl hydrolase Family 5 [31,32]. There was substantial sequence identity between residues 377–417 of CelE and the non-catalytic domains, of unknown function, from xylanase A, endoglucanases A and B and cellodextrinase C of *P. fluorescens* subsp. *cellulosa* (Figure 6; [12–14,16]). The C-terminal 100 amino acids of CelE displayed extensive sequence identity with prokaryotic Type-I CBDs (Figure 6) which are particularly prevalent in *P. fluorescens* subsp. *cellulosa* and *Cellulomonas fimi* plant cell wall hydrolases [2]. The major residues which are conserved in all Type-I CBDs, including four tryptophans and N- and C-terminal cysteines are retained in the CelE CBD homologue. In common with other *P. fluorescens* subsp. *cellulosa* plant cell wall hydrolases, the domains of CelE were linked by sequences rich in serine residues.

Relationship between the structure and function of CelE

To evaluate whether the putative N-terminal catalytic domain homologue constitutes a functional catalytic domain, truncated derivatives of *celE* were constructed and their capacity to direct the synthesis of a functional cellulase in *E. coli* was evaluated. The data presented in Figure 1 demonstrated that truncated CelE lacking 158 C-terminal residues (pC52) retained catalytic activity, but a derivative of the cellulase from which 100 N-terminal residues had been deleted (pC53) did not exhibit endoglucanase activity. These data confirm that the N-terminal 300 residues of mature CelE constitute the catalytic domain. To evaluate whether the C-terminal CBD homologue was biologically active, the capacity of full-length CelE to bind to cellulose was evaluated. The data presented in Figure 7 showed that a polypeptide of M_r 56 000, expressed by *E. coli* containing pC48 bound to Avicel. The polypeptide also bound to acid-swollen cellulose but not to xylan (data not shown). N-terminal sequence analysis of the polypeptide confirmed its identity as CelE. In contrast, a catalytically active but truncated derivative of CelE which had an M_r of 34 000 and the N-terminal sequence of DVAPLSVQGNKILAN (corresponding to residues 40–54 of full-length CelE), and was equivalent to the catalytic domain alone (predicted M_r 32 879) did not adhere to Avicel. These data suggest that the C-terminal 100-residue domain of CelE constitutes a functional CBD.

To evaluate the importance of the CBD in the capacity of CelE to attack various cellulosic substrates, the relative activities of full-length CelE and a truncated form of the enzyme, lacking the

CBD, against CMC, acid-swollen cellulose, Avicel and BMCC were determined. The data, presented in Table 2, clearly showed that although both forms of the cellulase exhibited similar activities against soluble and amorphous cellulose, the full-length enzyme displayed approximately four times higher activity against Avicel than the truncated derivative. Full-length and truncated CelE exhibited no detectable activity against BMCC.

DISCUSSION

The data presented in this report describe the cloning and characterization of a highly expressed endoglucanase, designated CelE, from *P. fluorescens* subsp. *cellulosa*. The enzyme comprises multiple domains with the catalytic domain derived from glycosyl hydrolase Family 5. In contrast, the catalytic domains of the other two *P. fluorescens* endoglucanases previously sequenced ([12,13]) are located in glycosyl hydrolase Families 9 and 45, respectively [31,32]. It is apparent, therefore, that the multiple endoglucanases of *P. fluorescens* resulted from the transfer of several discrete cellulase genes into the pseudomonad. The acquisition of multiple cellulases by lateral gene transfer is a common phenomenon among prokaryotes, while in lower eukaryotes isoforms of endo- and exo-glucanases often arise through gene duplication [32].

Within the classification established for glycosyl hydrolase enzymes, Family 5 is one of the largest groupings and contains enzymes of bacterial and fungal origin. Enzymes of this family are typically endoglucanases with relatively wide substrate ranges, which hydrolyse CMC or barley β -glucan preferentially and display variable but relatively low level activity against acid-swollen cellulose or Avicel. When compared with other endoglucanases belonging to Family 5, CelE from *P. fluorescens* subsp. *cellulosa* exhibited much higher activity against acid-swollen cellulose and Avicel. Another unusual feature of CelE is its relatively high thermostability for an enzyme that is derived from a mesophilic soil bacterium. It is possible that this property of the enzyme is a consequence of the same structural elements that confer resistance to proteolytic attack on the catalytic domain. Support for this view is provided by Fontes et al. [33] who showed that there was a clear relationship between the thermostability of plant cell wall hydrolases and their reduced susceptibility to proteolytic attack. It would be rational for CelE to evolve resistance to proteinases as its extracellular location would bring the cellulase into contact with protein-degrading enzymes secreted by other soil bacteria.

Previous studies in our laboratories have identified CBDs in a range of plant cell wall hydrolases expressed by *P. fluorescens* subsp. *cellulosa*, several of which do not attack cellulose directly. This led to the hypothesis that in *P. fluorescens* subsp. *cellulosa*

ACGGCCCTTTCTCGGCCCTTGAGATGAAAGATTTCTTCAAGTATGGCGGCACTTTCAGG
 -700
 GATCAGAGGCTGGCCAGAACACAGCGTGTGTATACGGAAGTGGTGGTCAAGCTTCTCCAA
 AGCAGGCTCAGAGCACTCTAAGTTCAGCAGAACATAGTAGCCTTGTGTACTTTTGTGAT
 -600
 GCTGTAGGCCAAGTGTAGCAGACGGCCCCAGTCTTCGAACGGGTGACCTGGCCACCGTCAAG
 -500
 TTTCACGGTAGAGGTATTAACCTCCACCATGGCCAGTACTTGTTCGCTTTGATCGGGGTG
 -400
 GATCAAACTACGACTTCGTATAGCGATATTAGCTTCTTACGGGTTGAAGTCTCGGGCC
 -300
 AGTGGCGTGTAGCAAGGAGAGACGGTTTTTATGCGCGGGCGGCAAGGCCACTTGCATAAC
 -200
 CCCCGGTCAAGTTTGGGTCGCAATTTCTACGGATTTGCGGTAAACACGCCAAGAAAGGT
 -100
 CTGCCAAGGGGGCTTCTCCCTTCGGAAGGTATCTTGGCTTATAGGGCTCAAAATGA
 CAGCGCTGCTTTTTTGTTCATTAGCAAAAATATGATTTTTTGGACATAATCCCGAACA
 -100
 GCTAATTTGCTCTAATTTTCTGCTTAAATTTTCAAGTCTTGGTGGTGTCCAGGAGCGGGGA
 -100
 TAGGGGGGAATTTTGGCAATTTATCGCTTTTTTTTATCTGACCGGGCCCTGAAATTTCCG
 +1
 GGGGATACCGGAATTTGCTGAAGGATCGAGGGCTTCGATAAACAATCTCTCAAAAATTTCCC
 TAGATGAGAGAACAGATATATGAAATAGAAATCTGACACTTTCGACCAAAACCAATCGG
 100
 CAAAGCAATCCGCAATCCGCTCCGATATCTGCTGGCCACTTGGACTCTCTCGCGG
 K A I R R A A I L L L G S F G L L G G

GTCTAGCCACAGCTGATGTGGGGCGCTGCTGCTAGAGGCCAACAAATCTCTGGCCAA
 V S A Q A D A D V A P L S V Q G M K I L A N
 200
 CGSGAGCGCCGCTCTCGGGCATGAGCTGTCTCGAGCAATAACGAGTGGGGCGG
 G Q P A S F S G H S L F W S M T E W G G
 CGAAAGTATTAACAGCCCAAGTGGTCAAGTGGCTTAAATCCGACTGGAAACGCCAAGCT
 E K Y Y N A A Q V V S W L K S D W N A K L
 300
 GGTGGCGCCCAATGGCGCTGGAGAGTGAAGCGGCTACTGACTGATCGGGCAATAA
 V R A A H G V E D E G G Y L I D P A N K
 400
 AGACAGGCTACCCAGTGGTGGATCGGCTATCGCCATGATATGATGATATCGA
 D R V T Q V V D A A I A N D M Y V I I D
 CTGGACTCGCAACAGCCCAACCAATACAGAGCCAGGCCAATTCCTTTTTCCAGGAAT
 W H S H N A H Q Y Q S Q A I A F Q G E H
 500
 GGCAAGATGTAAGCGCCCAACCACTGATCTATGAGATTTTACAGCAGACCGCTGCA
 A R K Y A N H V I Y E I Y N E P L Q
 GGTACTGGAGCAACCACTACAGCCCTAAGCCAGGGGGTGGATCGGGCTATTCGGC
 V S W S N T I K P Y A Q A V I A A I R A
 600
 CATAGCCCGCAACCTGATTTATGTCGCGCACCCCACTGCTGGTGGAGATGGATGT
 I D P D N L I I V G T P T W S Q D V D V
 700
 CGCCGCAATGACCTATACCGGCTACCGAATATGCTATACCTGCACTTTTATGC
 A A N D P I T T G Y Q M A I T Y T L H F Y A
 CGGCACCGGCAATACCTGCGCAGCAGCCCAAGCCCTTACCGGGTATAGC
 G I H G Q I L R D K A Q T A L N R G I A
 800
 CCTGTGTACCGAGTGGGTTCACTGCAATGCGCAATGGCGATGGTGTGTGGCTAACAG
 L F V T E W G S V N A M G D G A V A N S
 CGAGCAAAAGCCCTGGCTACGCTTATGAAACCAACCAATATCAAGCAACGCCAAGCTGGGC
 E T N A W V S P H K T N H I S N A N W A
 900
 GCTCATGATAGTCCAGGGCGCTTGGGGCTGGTTCGGGGAGTCCCAAGGGCG
 L N D K V E G G A S A L V F P G A S A N G G
 1000
 TTGGTCAACTCAACAATCAACGCTCGCGCCCTGGCCAAAGACTTATCAAGCGCTG
 W V M S Q L T A S G A L A K S I I S G W

GGCCAGCTACACAGCAGCAGCAGCTGGCGCTGCGAGCCAAACCCAGCTGTCAG
 P S Y N T S S S S A V S S Q T O V S S
 1100
 CTTCAACCCAGCCCGCGTAGTCCAGCTCCAGTCTCTCCACGGCCAGCTGGGTAGTGTCTC
 S S O A P V V S S S S T A S S V V S S
 CGCGGTATCCGGCCAGCAGTGAACACTGGTACGGCACCTCTACCCGCTGTGCAATGACAC
 A V S G Q Q C M W Y G T L Y P L C S T T
 1200
 CACCAATGGCTGGGCTGGGAGAACAAATGCCAGTGTGTATGGCCGGCCACTGACGGG
 I N G W G W E M N A S C I A R A T C S S G
 1300
 TCAGCCAGCCCTGGGGATTTGGTGGAGTACCAGCAGCCAGGCGAGTTTCCAGCGT
 Q P A P W G I V G G S T E S Q A S S S V
 CCGCTCAGCTCCTCCAGCGTGTTCAGCTCGGCTCCAGCAGCTCTTCCAGCGTTC
 R S S S S S L V S S S R S S S S S V Q
 AAGCTCAAGTGGCCAGCTCGTGGCCAGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
 S S A P S S S V A S S G S S S G Q C S
 CTATACCTGACCAACCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
 I T V I N Q W S N G F T A S I R I A N N
 1500
 CGGGACAGTCCAAATCAATGGCTGGAACTCAGCTGGAGTTACAGCGATGGCTCCCGCT
 G T S P I M G W M L S W S Y S D G S R V
 GACCAATAGTGGAAATGCCAATGTCTCGGGCAACAAATCCCTACACGGGTATCCAACTGGG
 T M S W N A M V S G M N P Y T V S N L G
 TTGGAACGGTATGATFCCAGCCCGGGCCAGGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGG
 W M G S I Q P G Q A V E F G F Q G T K N
 1700
 CAATAGCGGGCTATCCGAGCTAAGCGTAAATGTTGTAACTAATTAATGATCA
 N S A A A I P T L S G N V C N M
 TCCAGAGCGGGAGTATCTCGGGCGCTTCCATCACACAGATATTTTTCCCC

Figure 5 Nucleotide sequence of *celE* and the derived amino acid sequence of *CelE*

The experimentally determined N-terminal sequence of *CelE* is indicated by underlining. The proposed linker sequences are boxed.

(a)

CeLE	377	QCNWYGTLYPLCSTTTNGWGWENNASCARATCSGQPAPWIVGGS	417
XYLA	183	QCNWYGTLYPLCVTTTNGWGWEDQSRSCARSTCAQAPFVIGVSG	228
EGB	180	ACNWWYTLTLCNNTSNGWYEDGRSCVARTTCSAQPAPYGIIVSTS	225
EGA	669	NCNWYGTLYPLCVTTTQSGWGWENSSQCSIASSTCSAQPAPYIVGAA	714
CELC	183	QCNWYGTLYPLCVSTTSGWYENNRSCISPSSTCSAQPAPYIVGGS	228
		***** * ***** ** *****	

(b)

CeLE	471	QCSYTVTNQWSNGFTASIRIANNNGTSPINGWNLWSY	
XYLB/C/D	38	ACTYTIIDSEW TGFTANITLKNDTGAATNWNWNWQY	
XYLA	30	TCSYNIITNEWNTGYTGDIITINRGSAINGWVSNWQY	
EGB	31	ACEYRVITNEWGSGFTASIRIKNNGSSITINGWSVSNWY	
EGA	864	NCQYVVTNQWNGFTAVIRVRNNGSSAINRWSVSNWY	
EGC	39	GCEYVVTNSWGSFTAAIRITNSTSSVINGWVSNWY	
		* * * * * * * * * * * * * * * * *	

CeLE		SDGSRVTSNWNANVSGNPNFYVSNLWNGSIQPGQA	
XYLB/C/D		SS NRMTSGWNAFSGSNPYASNLWNGNIQPGQS	
XYLA		AT NRLSSSNANVSGSNPYASNLWNGNIQPGQS	
EGB		TD SRVTSNWNAGLGSANPYASNPVWNTSIPIGSS	
EGA		SD SRITNSWNAVVTGNPNYAASALGWNAIQPGQT	
EGC		NS NRVTNLWNPNSLGSNPYASNLWNGTIQPGQT	
		*** ** * * * * * * * * * * * * * *	

CeLE		VEFGFQGTKNNSAAAIPITLGSNVC	568
XYLB/C/D		ISFGLQGEKNGSTAERPVTVGAAC	133
XYLA		VSPFGQVKNKGGSAERPVSQVGGIC	126
EGB		VEFGVQGNNGSSRAQVPAVTGAIC	127
EGA		AEFGFQGTGAGSRQVPAVTGSVC	960
EGC		VEFGFQGVTSNGTVEPTVNGAAC	134
		* * * * * * * * * * * * * *	

Figure 6 Alignment of CelE with homologous prokaryotic sequences

The middle non-catalytic domain (a) and CBD (b) of CelE were aligned with plant cell wall hydrolases from *P. fluorescens* subsp. *cellulosa*, and those residues which showed sequence identity or similarity in all sequences compared are indicated by an asterisk. The enzymes compared were as follows: endoglucanase A (EGA; [13]), endoglucanase B (EGB; [12]), xylanase A (XYLA; [14]), xylanase B (XYLB; [15]), cellodextrinase (CELC; [16]), xylan arabinofuranosidase (XYLC; [15]) and acetyl xylan esterase (XYLD; [5]). The location of the N- and C-terminal residues, in the respective full-length sequences, are indicated.

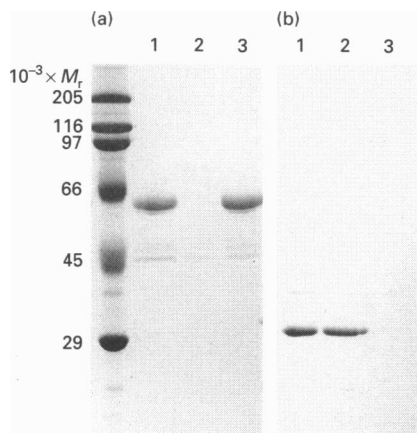


Figure 7 Affinity of CelE for cellulose

Purified full-length (a) and truncated (b) CelE (lane 1) were incubated with Avicel (5 mg) in 100 mM Tris/HCl, pH 7.6, at 0 °C for 30 min with gentle shaking. The unbound material was removed by centrifugation and the Avicel pellet was washed three times in 100 mM Tris/HCl, pH 7.6, containing 150 mM NaCl. Unbound (lane 2) and bound (lane 3) material were analysed by SDS/PAGE [using a 10% (w/v) polyacrylamide gel] and Coomassie Blue staining. Protein molecular mass markers were run on the extreme left.

the role of the CBDs is to bring an array of plant cell wall hydrolases into close contact with native substrates, increasing the synergistic interactions between these enzymes [12,17]. However, the data described in this report indicate that the CBD in

CeLE plays a direct role in facilitating the hydrolysis of substantially crystalline cellulosic substrates such as Avicel. It could be argued that the CBD is mediating an increase in the hydrolysis of cellulose by bringing the enzyme into intimate contact with its substrate. However, the observation that while the CBD binds to acid-swollen cellulose, it does not enhance the activity of the endoglucanase against this substrate, suggests that the CBD is not causing its effect by simply increasing the concentration of the enzyme at the surface of the polysaccharide. An alternative possibility is that the CBD disrupts the crystalline structure of Avicel facilitating its hydrolysis by CelE. However, it is apparent that while the cellulase cleaves Avicel it does not attack BMCC. BMCC is among the most highly crystalline forms of cellulose with a crystallinity index (CI) of approximately 80. In contrast, Avicel has a CI of 50 and thus contains amorphous and crystalline regions. Full-length CelE was four times more active against Avicel than the truncated form which lacked the C-terminal CBD. Since CelE has no detectable activity against crystalline cellulose, this suggests that the CBD does not facilitate the hydrolysis of Avicel by disrupting the crystalline structure of the polysaccharide. It would appear, therefore, that the CBD increases the activity of CelE against Avicel by exposing new regions of amorphous cellulose, although the precise mechanism by which the domain mediates increased substrate availability remains to be elucidated. In contrast to the action of the *Pseudomonas* CBD, Din et al. [34] demonstrated that the isolated CBD from *Cellulomonas fimi* endoglucanase A could open up the structure of ramie cellulose, making the substrate more accessible to enzyme attack. Similarly, Coutinho et al. [8] demonstrated that removal of the Type-I CBD from *C. fimi* endoglucanase A diminished the activity of the enzyme against highly crystalline (BMCC), but not amorphous cellulose. Although these data suggest that Type-I CBDs generally increase the capacity of *C. fimi* cellulases to hydrolyse celluloses which contain a crystalline component, not all Type-I CBDs have the same effect on enzyme activity or exhibit affinity for precisely the same ligands. For example, the C-terminal CBDs of *C. fimi* xylanase D and *Thermomonospora fusca* xylanase A exhibit equal affinity for both cellulose and xylan [35,36], while the internal Type-I CBD-homologue of the *C. fimi* enzyme binds exclusively to xylan [35]. In addition, while fusion of the CBD from *T. fusca* endoglucanase E to *Bacillus subtilis* endoglucanase A increased the activity of the enzyme against ball-milled cellulose [7], the cellulase was also more active against amorphous cellulose. Furthermore, the Type-I CBD from *P. fluorescens* subsp. *cellulosa* xylanase B did not elicit Avicel hydrolysis when fused to endoglucanases E and A from *Cl. thermocellum* and *Ruminococcus albus*, respectively [6]. However, it has been suggested that the lack of activity of the hybrid enzymes against insoluble cellulose could be a consequence of the inability of these catalytic domains to accommodate the insoluble substrate, irrespective of the properties of the CBD [2,7]. It should also be noted that the CBD of *C. fimi* endoglucanase A greatly increased the activity of the enzyme against BMCC, a substrate not attacked by CelE, even though the pseudomonad enzyme contains a Type-I CBD. This observation supports the view that although CBDs are important in facilitating cellulose hydrolysis, the nature of the active site of endo-acting cellulases also plays a pivotal role in defining the substrates accessible to a specific endoglucanase.

To conclude, it is apparent that the CBD of at least one cellulase from *P. fluorescens* subsp. *cellulosa* enhances cellulolysis of Avicel by increasing the exposure of amorphous cellulose to the enzyme, through a mechanism which does not involve the disruption of the crystalline regions of the substrate. This suggests that Type-I CBDs of *P. fluorescens* cellulases and hemicellulases

do not all play an equivalent role, and argues against the view that the role of the CBDs of the pseudomonad enzymes is to simply increase the synergistic interactions between these plant cell wall hydrolases. However, it remains to be elucidated whether CBDs from *P. fluorescens* subsp. *cellulosa* hemicellulases can play an identical role to the CBD from CelE.

We wish to thank the Biotechnology and Biological Sciences Research Council for supporting this work (LE13/138).

REFERENCES

- 1 Gilbert, H. J. and Hazlewood, G. P. (1993) *J. Gen. Microbiol.* **139**, 187–194
- 2 Béguin, P. and Aubert, J.-P. (1994) *FEMS Microbiol. Rev.* **13**, 25–58
- 3 Tomme, P., van Tilbeurgh, H., Pettersson, G., van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeysens, M. (1988) *Eur. J. Biochem.* **170**, 575–581
- 4 Knowles, J., Lehtovaara, P. and Teeri, T. T. (1987) *Trends Biotechnol.* **5**, 255–261
- 5 Ferreira, L. M. A., Wood, T. M., Williamson, G., Faulds, C., Hazlewood, G. P., Black, G. W. and Gilbert, H. J. (1993) *Biochem. J.* **294**, 349–355
- 6 Poole, D. M., Durrant, A. J., Hazlewood, G. P. and Gilbert, H. J. (1991) *Biochem. J.* **279**, 787–792
- 7 Maglione, G., Matsushita, O., Russell, J. B. and Wilson, D. B. (1992) *Appl. Environ. Microbiol.* **58**, 3593–3597
- 8 Coutinho, J. B., Gilkes, N. R., Kilburn, D. G., Warren, R. A. J. and Miller, R. C., Jr. (1993) *FEMS Microbiol. Lett.* **113**, 211–218
- 9 Gilbert, H. J., Sullivan, D. A., Jenkins, G., Kellett, L. E., Minton, N. P. and Hall, J. (1988) *J. Gen. Microbiol.* **134**, 3239–3247
- 10 Gilbert, H. J., Jenkins, G., Sullivan, D. A. and Hall, J. (1987) *Mol. Gen. Genet.* **210**, 551–556
- 11 Braithwaite, K. L., Black, G. W., Hazlewood, G. P., Ali, B. R. S. and Gilbert, H. J. (1995) *Biochem. J.* **305**, 1005–1010
- 12 Gilbert, H. J., Hall, J., Hazlewood, G. P. and Ferreira, L. M. A. (1990) *Mol. Microbiol.* **4**, 759–767
- 13 Hall, J. and Gilbert, H. J. (1988) *Mol. Gen. Genet.* **213**, 112–117
- 14 Hall, J., Hazlewood, G. P., Huskisson, N. S., Durrant, A. J. and Gilbert, H. J. (1989) *Mol. Microbiol.* **3**, 1211–1219
- 15 Kellett, L. E., Poole, D. M., Ferreira, L. M. A., Durrant, A. J., Hazlewood, G. P. and Gilbert, H. J. (1990) *Biochem. J.* **272**, 369–376
- 16 Ferreira, L. M. A., Hazlewood, G. P. and Gilbert, H. J. (1991) *Biochem. J.* **279**, 793–799
- 17 Millward-Sadler, S. J., Poole, D. M., Henrissat, B., Hazlewood, G. P., Clarke, J. H. and Gilbert, H. J. (1994) *Mol. Microbiol.* **11**, 375–382
- 18 Teather, R. M. and Wood, P. J. (1982) *Appl. Environ. Microbiol.* **43**, 777–780
- 19 Chambers, S. P., Prior, S. E., Barstow, D. A. and Minton, N. P. (1988) *Gene* **68**, 139–149
- 20 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor, New York
- 21 Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161–178
- 22 Murphy, G. and Kavanagh, T. (1988) *Nucleic Acids Res.* **16**, 5198
- 23 Staden, R. (1980) *Nucleic Acids Res.* **16**, 3673–3694
- 24 Wood, T. M. (1971) *Biochem. J.* **121**, 353–362
- 25 Miller, G. L. (1959) *Anal. Chem.* **31**, 426–428
- 26 Braun, C., Meinke, A., Ziser, L. and Withers, S. G. (1993) *Anal. Biochem.* **212**, 259–262
- 27 Sedmak, J. J. and Grossberg, S. E. (1977) *Anal. Biochem.* **69**, 544–552
- 28 Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- 29 Zhou, L., Xue, G.-P., Orpin, C. G., Black, G. W., Gilbert, H. J. and Hazlewood, G. P. (1994) *Biochem. J.* **297**, 359–364
- 30 Hunkapillar, M. W., Hewick, R. E., Dreyer, W. J. and Hood, L. E. (1983) *Methods Enzymol.* **91**, 399–413
- 31 Henrissat, B. (1991) *Biochem. J.* **280**, 309–316
- 32 Henrissat, B. and Bairoch, A. (1993) *Biochem. J.* **293**, 781–788
- 33 Fontes, C. M. G. A., Hall, J., Hirst, B. H., Hazlewood, G. P. and Gilbert, H. J. (1995) *Appl. Microbiol. Biotechnol.* **43**, 52–57
- 34 Din, N., Gilkes, N. R., Tekant, B., Miller, R. C., Jr., Warren, R. A. J. and Kilburn, D. G. (1991) *Bio/Technology* **9**, 1096–1099
- 35 Black, G. W., Hazlewood, G. P., Millward-Sadler, S. J., Laurie, J. I. and Gilbert, H. J. (1995) *Biochem. J.* **307**, 191–195
- 36 Irwin, D., Jung, E. D. and Wilson, D. B. (1994) *Appl. Environ. Microbiol.* **60**, 763–770