Microbiology (1997), **143**, 891–898 Printed in Great Britain

Structure of the *Clostridium stercorarium* gene *celY* encoding the exo-1,4-β-glucanase Avicelase II

Karin Bronnenmeier, Kerstin Kundt, Kathrin Riedel, Wolfgang H. Schwarz and Walter L. Staudenbauer

Author for correspondence: Karin Bronnenmeier. Tel: +49 89 2892 2637. Fax: +49 89 2892 2360.

Institute for Microbiology, Technical University Munich, Arcisstraße 21, D-80290 München, Federal Republic of Germany

The nucleotide sequence of the *celY* gene coding for the thermostable exo-1,4-β-glucanase Avicelase II of *Clostridium stercorarium* was determined. The gene consists of an ORF of 2742 bp which encodes a preprotein of 914 amino acids with a molecular mass of 103 kDa. The signal-peptide cleavage site was identified by comparison with the N-terminal amino acid sequence of Avicelase II purified from *C. stercorarium*. The *celY* gene is located in close vicinity to the *celZ* gene coding for the endo-1,4-β-glucanase Avicelase I. The CelY-encoding sequence was isolated from genomic DNA of *C. stercorarium* with the PCR technique. The recombinant enzyme produced in *Escherichia coli* as a LacZ′–CelY fusion protein could be purified using a simple two-step procedure. The properties of CelY proved to be consistent with those of Avicelase II purified from *C. stercorarium*. Sequence comparison revealed that CelY consists of an N-terminal catalytic domain flanked by a domain of 95 amino acids with unknown function joined to a type III cellulose-binding domain. The catalytic domain belongs to the recently proposed family L of cellulases (family 48 of glycosyl hydrolases).

Keywords: cellulase, *Clostridium*, Avicelase, exo-1,4-β-glucanase, cellobiohydrolase

INTRODUCTION

The existence of an exo-1,4-β-glucanase in cellulolytic bacteria was questioned for a long time, although Creuzet *et al*. (1983) presented evidence for the presence of a cellobiohydrolase in the thermoanaerobic bacterium *Clostridium stercorarium*. Recently, CelS, the most abundant catalytic subunit in cellulosomes of the cellulolytic thermophile *Clostridium thermocellum*, was characterized as an exo-1,4-β-glucanase (Morag *et al*., 1991; Wang *et al*., 1993; Kruus *et al*., 1995). Furthermore, CbhA and CbhB of *Cellulomonas fimi* (Meinke *et al*., 1994; Shen *et al*., 1995) and presumably CelF of *Clostridium cellulolyticum* (Reverbel-Leroy *et al*., 1996) were shown to represent this type of cellulolytic enzyme in mesophilic bacteria.

Avicelase II purified from *C. stercorarium* has previously been identified as a novel type of exoglucanase, acting as a cellodextrinohydrolase rather than a cellobiohydrolase (Bronnenmeier *et al*., 1991, 1996). Avicelase II is an essential component of the 'low-complexity' cellulase system of *C. stercorarium*. Acting together, Avicelase II and the endo-1,4-β-glucanase Avicelase I effect the efficient hydrolysis of microcrystalline cellulose (Bronnenmeier & Staudenbauer, 1988, 1990). Avicelase I, encoded by the gene *celZ*, was shown to be a member of cellulase family E (Jauris *et al*., 1990). In this paper we report the complete nucleotide sequence of the *celY* gene and describe the properties of the encoded enzyme, Avicelase II, purified from a recombinant *Escherichia coli* strain. Analysis of the deduced amino acid sequence shows that this enzyme belongs to a growing new family of cellulases, presently called family L (Shen *et al*., 1994), thus far comprising only bacterial enzymes.

METHODS

Bacterial strains and plasmids. *C. stercorarium* NCIB 11754 was obtained from the National Collection of Industrial and Marine Bacteria, Aberdeen, UK. The cloning vector/host strain combinations used for sequencing and expression studies were pUC19/*E. coli* JM83 (Hanahan, 1983; Yanish-Perron *et al*., 1985) and pBluescript II KS(−)/*E. coli* XL1-Blue

The EMBL accession number for the sequence reported in this paper is Z69359.

(Stratagene). The pIC20R-derived plasmid pCMC-130E and the pHC79-derived plasmids pCMC-107 and pCMC-130 have been described previously (Schwarz *et al*., 1989). *E. coli* strains were routinely grown in Luria–Bertani medium (LB; 1% peptone, 0·5% yeast extract, 0·5% sodium chloride, pH 7·2) supplemented with ampicillin at a concentration of 100 μg ml^{-1}.

Recombinant DNA techniques and sequence analysis. Preparation of chromosomal and plasmid DNA, endonuclease digestion, ligation and transformation were carried out using standard procedures (Sambrook *et al*., 1989). Enzymes for DNA modification were purchased from Boehringer Mannheim, Gibco/BRL and Stratagene. Nucleotide sequence determination was done via double-strand plasmid sequencing (Sanger *et al*., 1977) on both strands employing the DIG Taq DNA sequencing kit (Boehringer Mannheim). Sequence data were analysed and compared using the DNASIS/PROSIS (Hitachi America) software package.

Construction of pBS*celY*. Two synthetic oligonucleotide primers with the nucleotide sequences 5′ GCGAATCGAT-GGCTTATGCTTCTTCGG 3′ and 5′ GCATTCTAGA-ACCTCCTCCCTTCAACGG 3′, designated pCelY1 and pCelY2, respectively, were used in PCR to amplify the *celY*-coding region. The reaction mixture contained 500 ng chromosomal *C. stercorarium* DNA as template, 50 pmol each primer, deoxynucleotide triphosphates (200 μM each), 2 U *Pyrococcus furiosus* (Pfu) DNA polymerase and 10 μl 10× Pfu reaction buffer (Stratagene) in a total volume of 0·1 ml. The template strands were separated at 95 °C (3 min) before starting 35 reaction cycles (94 °C, 1 min; 62 °C, 45 s; 72 °C, 3 min). The PCR primers pCelY1 and pCelY2 contained recognition sites for *Cla*I and *Xba*I, respectively, which were used for in-frame fusion of the *celY* gene to the *lacZ′* gene sequence of pBluescript II KS(−) treated with the same restriction enzymes. The correct insertion of the PCR product into pBluescript II KS(−) was verified by nucleotide sequence analysis. The resulting plasmid, designated pBS*celY*, was used to transform *E. coli* XL1-Blue.

Preparation of crude extract and purification of recombinant CelY from *E. coli* XL1-Blue(pBS*celY*). Eight 2·5 l cultures of *E. coli* XL1-Blue(pBS*celY*) were grown in 2× LB medium supplemented with 100 μg ampicillin ml^{-1}. IPTG (1 mM) was added at OD_{600} 0·5. After 16 h, cells were harvested, washed and resuspended in 20 mM Tris/HCl pH 7·0. The cells were lysed by passage of the suspension through a French pressure cell (American Instrument Co.). Remaining whole cells and debris were removed by centrifugation. The cleared lysate was incubated at 65 °C for 30 min and centrifuged again to pellet the heat-labile host cell proteins. The supernatant was applied to a Q Sepharose fast-flow anion-exchange column (Pharmacia), equilibrated with 20 mM Tris/HCl pH 7·0. Elution was achieved with a 10-bed-volume gradient of 0–0·5 M sodium chloride in 20 mM Tris/HCl pH 7·0. The N-terminal amino acid sequence of the purified protein was determined by Edman degradation using a gas-phase amino acid Sequenator, model 477A (Applied Biosystems).

Assay of enzyme activity. Enzyme assays were done at 70 °C in 0·5 or 0·75 ml of reaction mixture containing 50 mM MES buffer pH 6·0, 10 mM $CaCl_2$ and 0·5% of one soluble polysaccharide substrate or 1% of one non-soluble substrate or 4 mM of *p*-nitrophenyl β-D-cellobioside. The enzymic release of reducing groups was determined using the dinitrosalicylic acid (DNSA) reagent (Wood & Bhat, 1988). One unit of enzyme activity was defined as the amount of enzyme needed to release 1 μmol glucose-equivalent reducing groups per minute. Thermostability was measured by determination of residual activity with the standard assay after preincubation of enzyme samples in 50 mM MES buffer pH 6·0 and 10 mM $CaCl_2$, without substrates.

Carboxymethylcellulose (CMC; low viscosity) and *p*-nitrophenyl β-D-cellobioside were purchased from Sigma. Avicel (microcrystalline cellulose 0·02 mm) and Cellulose MN300 were from Serva. Oat spelts xylan was from Fluka. Barley β-glucan was from Megazyme. Acid-swollen Avicel was prepared by incubating Avicel in concentrated hydrochloric acid with continuous agitation for 2 h (Sakamoto *et al*., 1984).

Analytical methods. Protein concentrations were measured by the method of Sedmak & Grossberg (1977) with bovine serum albumin as a standard. SDS-PAGE was carried out in 10% polyacrylamide gels in the presence of 0·1% SDS as described by Laemmli (1970). Cellodextrins were analysed by HPLC on an Aminex HPX 42A column (Bio-Rad).

RESULTS

Nucleotide sequence of the *celY* gene

The physical map of the region of the *C. stercorarium* chromosome containing the *celY* gene is shown in Fig. 1. This gene is located upstream of the *celZ* gene, which codes for the endo-1,4-β-glucanase Avicelase I (Jauris *et al*., 1990). The *celZ* gene had been identified in a genomic library of *C. stercorarium* DNA by screening the recombinant cosmid clones for enzymic activity (Schwarz *et al*., 1989). Due to lack of a sensitive assay procedure (Bronnenmeier *et al*., 1991) the *celY* gene escaped detection. In search of the transcription start point, the sequence upstream of *celZ* revealed another ORF. This ORF extended beyond the 5′ end of the genomic DNA insert of pCMC-130E (Jauris *et al*., 1990) and showed significant homology to the 3′ end of the *celZ* gene. Because of the proximity of the putative *cel* gene to *celZ*, it was predicted that it would encode the exo-1,4-β-glucanase Avicelase II already purified from *C. stercorarium* (Bronnenmeier *et al*., 1991). Restriction enzyme mapping and hybridization studies led to the identification of two cosmid clones (CMC 107 and CMC 130) containing inserts of sufficient size to completely encompass the *celY* gene. Subcloning of a *Sph*I/*Kpn*I fragment of pCMC130 in pUC19 and a *Sca*I fragment of pCMC107 in pBluescript II KS(−) yielded the plasmids pWS1404 and pWS1405, shown in Fig. 1. The complete nucleotide sequence of *celY* was obtained by sequencing the overlapping inserts of both subclones by the dideoxynucleotide chain-termination method (Sanger *et al*., 1977).

The nucleotide sequence of the *celY* gene and its flanking regions is presented in Fig. 2. There are two ORFs in this strand and its complement. The CelY-encoding sequence (ORF1, nucleotides 1–2742) begins with the ATG start codon at position +1 preceded, with a spacing of 6 bp, by a potential ribosome-binding site (AGGGGA) with a calculated free energy of Shine–Dalgarno pairing of −69·4 kJ mol^{-1}. At the 3′ end the coding region is flanked by two consecutive stop codons

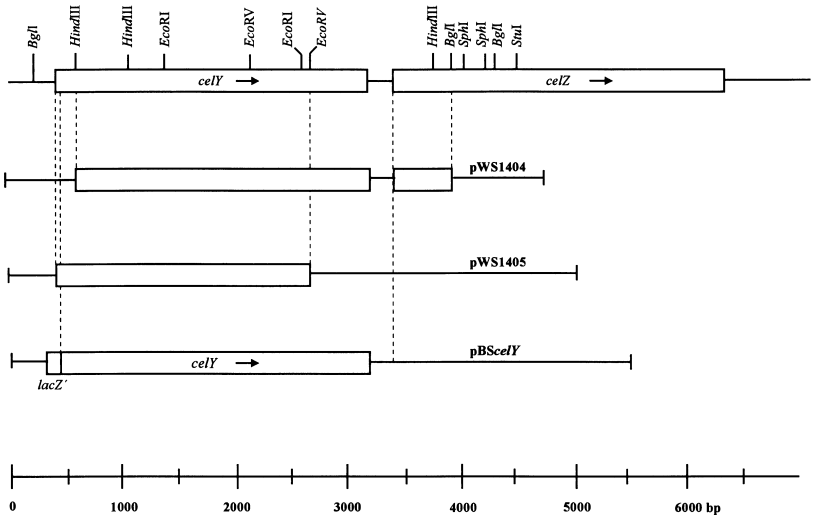

Fig. 1. Physical map of the *celY–celZ* region of the *C. stercorarium* genome. Plasmids pWS1404 and pWS1405 carrying the *C. stercorarium* DNA inserts used for sequence determination, and plasmid pBS*celY* carrying the complete *celY* gene are also shown. The boxed regions represent the *celY* and *celZ* coding sequence. The arrows indicate the direction of transcription.

(TAA and TAG). Thus, the primary translation product of the *celY* gene is 914 amino acids long and has a predicted molecular mass of 103 kDa. The codon utilization of the *celY* gene is very similar to that reported for *celZ* (Jauris *et al*., 1990) and *xynA* (Sakka *et al*., 1993) of *C. stercorarium*. ORF2, starting at position 2887, codes for the *celZ* gene, which is succeeded by a transcription terminator stem–loop (Jauris *et al*., 1990). The coding region of the *celY* gene has an average GC content of 46·3 mol% which is close to the 45 mol% determined for the *celZ* gene. As suggested by Gräbnitz *et al*. (1989), the GC content of the non-coding regions surrounding the two structural genes is approximately 10 mol% lower. The 138 bp nucleotide sequence that lies between the two stop codons of the *celY* gene and the initiation codon of the *celZ* gene contains an 18 bp palindrome. Although this sequence lacks the characteristic run of T residues found at the 3′ end of factor-independent transcription terminators, it could form a stable RNA hairpin ($\Delta G = -107$ kJ mol^{-1}) and might play a role in regulating the rate of expression of *celZ* relative to the *celY* gene. *celY* and *celZ* might be expressed as a single transcription unit, although the transcription start point of the putative *celYZ* operon has not been identified. Sequence inspection did not reveal any typical promoter sequence in the 5′ region preceding the *celY* start codon.

Protein structure

The domain organization of CelY is similar to that of CelZ from *C. stercorarium* (Jauris *et al*., 1990). The deduced amino acid sequence can be divided into several discrete segments: an N-terminal signal peptide, a catalytic region and a C-terminal non-catalytic region consisting of two domains (Fig. 3).

The identity of the N-terminal amino acid sequence of Avicelase II purified from *C. stercorarium* (Bronnenmeier *et al*., 1991) with residues 35–69 of the deduced amino acid sequence of CelY indicates that the protein encoded by the *celY* gene corresponds to the exo-1,4-β-glucanase present in the culture supernatant of *C. stercorarium*. The predicted sequence of residues 1–34 contains the characteristic arrangement of basic and hydrophobic regions typical of the signal peptides of bacterial exoproteins. Cleavage of the signal peptide was inferred to occur after Ala-34. This yields a mature protein of 99·2 kDa, which is in agreement with the apparent molecular mass of 94 kDa determined by SDS-PAGE for Avicelase II isolated from *C. stercorarium*.

On the basis of amino acid sequence similarity a catalytic domain comprising at least 620 residues was postulated. The C-terminus of the deduced CelY sequence comprises one copy of each of the two non-catalytic domains, B and C, which are present in duplicate in the cellulose-binding fragment of CelZ. Domain C is a type III cellulose-binding domain (Béguin & Aubert, 1994). Domain B is homologous to a hydrophilic repeat present four times in the cellulose-binding protein CbpA of *Clostridium cellulovorans* (Shoseyov *et al*., 1992), at least twice in the cellulose-binding protein CipC of *Clostridium cellulolyticum* (Pagès *et al*., 1996; Reverbel-Leroy *et al*., 1996) and twice in the C-terminal region of the endoglucanase CelB of *Bacillus lautus* (Jorgensen & Hansen, 1990). The function of domain B remains unknown. The amino acid sequence does not contain any obvious linker sequences between the domains.

Classification of CelY (Avicelase II)

Recently, a new family of 1,4-β-glycanases, designated family L (family 48 in the general classification of glycosyl hydrolases) was proposed (Shen *et al*., 1994, 1995). The catalytic domain of the CelY enzyme is highly homologous with the members of this family (Table 1). The degree of sequence identity within family L is remarkable: it ranges between 54% and 93%. The sequence of CelY is most closely related to the C-terminal catalytic domain of the multidomain cellulase CelA of '*Caldocellum saccharolyticum*' (Te'o *et al*.,

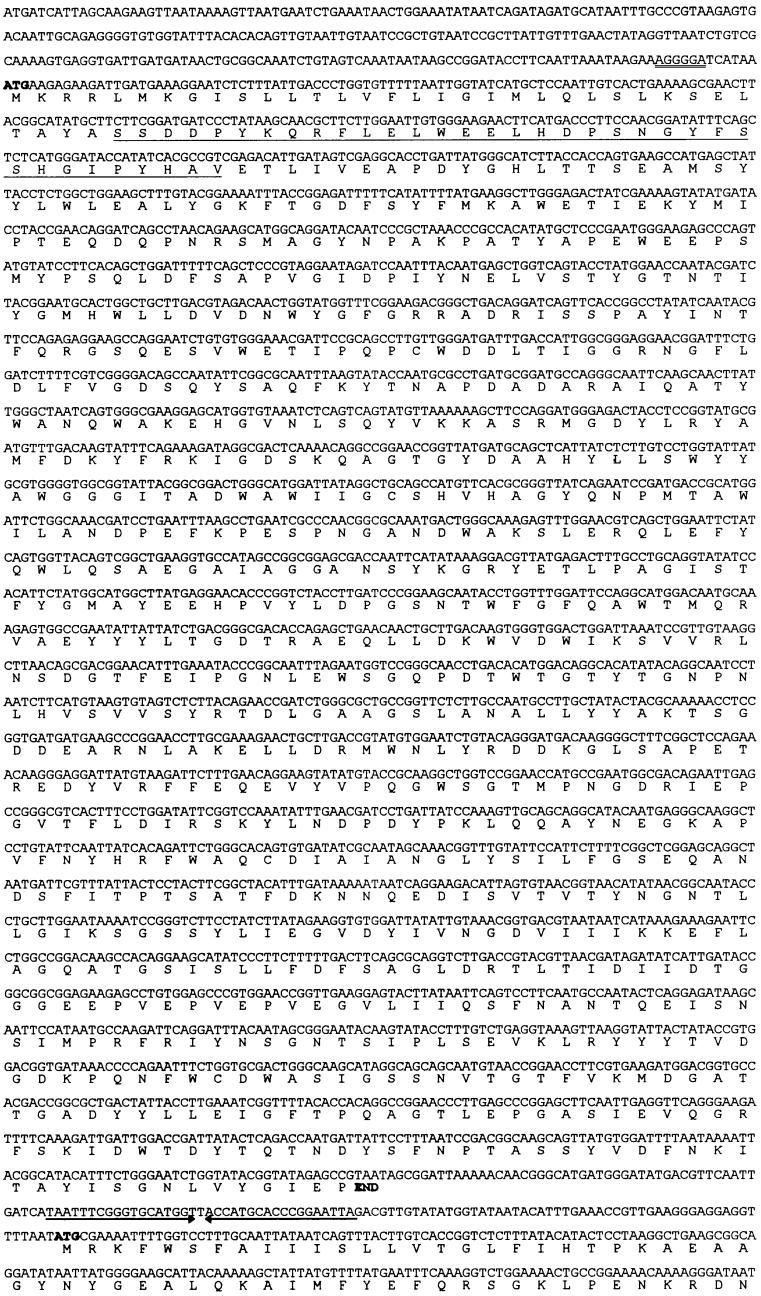

Fig. 2. Nucleotide sequence of the *C. stercorarium celY* gene and deduced amino acid sequence of Avicelase II. The presumptive Shine–Dalgarno ribosome-binding site preceding the start codon of *celY* is doubly underlined. The underlined amino acids were determined by liquid-phase sequencing. The palindromic sequence between *celY* and *celZ* is underlined by facing arrows. The numbers refer to the nucleotide (or amino acid) position, with numbering starting at the first residue of the coding sequences.

1995) which has recently been renamed *Caldicellulosiruptor saccharolyticus* (Rainey *et al*., 1994).

Expression in *E. coli* and purification of the recombinant enzyme

The entire sequence encoding the mature CelY polypeptide (nucleotides 88–2887 in Fig. 2) was isolated from *C. stercorarium* chromosomal DNA via PCR using appropriate oligonucleotide primers hybridizing with sequences at the 5′ and 3′ ends of the coding region (see Methods). It should be pointed out that the sequence encoding the CelY signal peptide was omitted. Insertion of the 2·8 kb PCR product downstream of the P_{lac} promoter and in frame with the *lacZ′* gene sequence into pBluescript II KS(−) yielded a plasmid, pBS*celY*, conferring IPTG-inducible expression of a LacZ′–CelY fusion protein upon *E. coli* transformant cells (Fig. 1).

The recombinant CelY produced in *E. coli* was purified using a simple two-step procedure (Table 2). The majority of host proteins were removed by heat treatment of the crude cell-extract preparation for 30 min at 65 °C. Extensive purification was achieved by FPLC anion-exchange chromatography on a Q Sepharose fast-flow column. A 14-fold purification was sufficient to yield an almost homogeneous protein preparation. This implies that the recombinant protein constitutes approximately 7% of the total soluble protein of the host. Analysis of the purified enzyme by SDS-PAGE

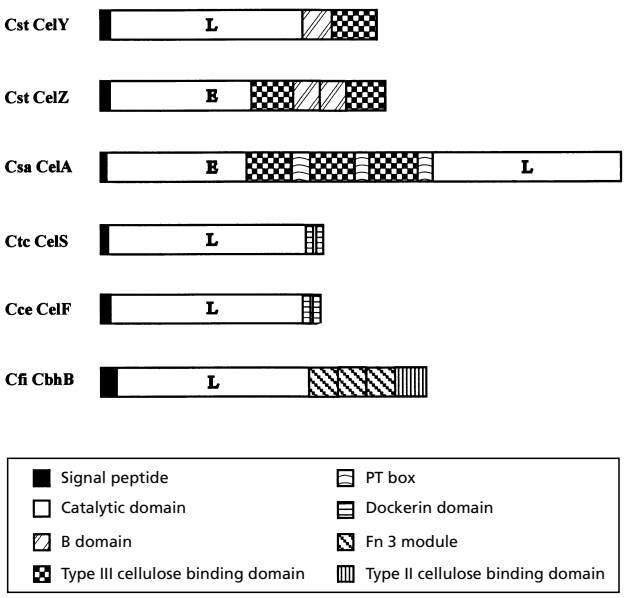

Fig. 3. Domain organization of CelY (Avicelase II) and CelZ (Avicelase I) of *C. stercorarium* as compared to the other members of cellulase family L. Domains showing significant similarity are indicated by the same pattern. Catalytic domain families are indicated by letters. Abbreviations: Cst, *Clostridium stercorarium*; Csa, *Caldicellulosiruptor saccharolyticus*; Ctc, *Clostridium thermocellum*; Cce, *Clostridium cellulolyticum*; Cfi, *Cellulomonas fimi*.

Table 1. Homology of catalytic domains within cellulase family L

Identity is defined as the number of identical amino acid residues in each pairwise comparison, expressed as a percentage of the total number of amino acids compared. Abbreviations: CstY, *C. stercorarium* CelY (EMBL/GenBank accession number Z69359); CsaA, *Caldicellulosiruptor saccharolyticus* C-terminal domain of CelA (Te'o *et al.*, 1995; L32742); CtcS, *C. thermocellum* CelS (Wang *et al.*, 1993; L06942); CfiB, *Cellulomonas fimi* CbhB (Shen *et al.*, 1995; L38827); CceF, *C. cellulolyticum* CelF (Reverbel-Leroy *et al.*, 1996; U30321); Cjo, *C. josui* deduced protein of ORF1 (Fujino *et al.*, 1993; D16670).

Core enzyme	Identity (%)					
	CstY	CsaA	CtcS	CfiB	CceF	Cjo
CstY	100	70	61	56	63	65
CsaA		100	65	57	68	68
CtcS			100	54	59	60
CfiB				100	55	55
CceF					100	93
Cjo						100

(Fig. 4) revealed a predominant protein band with a molecular mass of 97 kDa. This is consistent with the size predicted from the sequence data for the recombinant fusion protein (102 kDa).

Optimal conditions for enzyme action

The effects of pH and temperature on the activities of the recombinant CelY and the native Avicelase II enzymes were determined under exactly the same assay conditions. Both enzymes were maximally active around pH 5·0 and retained 40 % activity at pH values ranging from 5 to 8. When the enzymes were assayed for 120 min with acid-swollen Avicel, the temperature profiles of both enzymes coincided completely. The highest activity was observed at 75 °C, and 85 % of the maximal activity was still detectable at 80 °C. The pronounced thermostability reported earlier for Avicelase II (Bronnenmeier *et al.*, 1991) was also found for the recombinant enzyme. Microcrystalline Avicel was degraded at a constant rate for at least 4 d at 65 °C. A prerequisite for thermostability in the absence of substrate was the presence of Ca^{2+}. Under these conditions no loss of activity was observed during 4 d incubations at 60 °C and 70 °C. However, at 75 °C the enzyme exhibited a significantly lower half-life of 1–2 h.

Substrate specificity and degradation products

The highest level of CelY activity was observed with amorphous cellulose (Table 3). Microcrystalline cellulose preparations (Avicel, Cellulose MN300) were hydrolysed at a significantly lower rate. This agrees with the substrate specificity previously reported for Avicelase II (Bronnenmeier *et al.*, 1991). A kinetic analysis of the reaction products formed during hydrolysis of cellulosic substrates and cellodextrins indicated that CelY acts via the same exoglucanolytic mechanism as Avicelase II (data not shown). The recombinant enzyme showed significant activity towards the mixed-linkage substrate barley β-glucan and very low activity towards the substituted cellulose derivative CMC. Variable amounts of these activities had previously been detected in Avicelase II preparations of *C. stercorarium* and attributed to trace amounts of contaminating CelZ enzyme. It should be noted that the activity of the *C. stercorarium* endoglucanase CelZ towards β-glucan (481230 mU mg^{-1}; Bronnenmeier & Staudenbauer, 1990) is higher by several orders of magnitude than that of CelY (76 mU mg^{-1}). Compared to typical xylanases, the activities of CelY and Avicelase II towards xylan both appear insignificant. The xylanase activity found associated with the native enzyme purified from *C. stercorarium* was higher than that determined for the recombinant CelY. This difference is most likely due to a contamination of the Avicelase II preparation with one of the xylanases present in culture supernatants of *C. stercorarium* (Bronnenmeier *et al.*, 1996). Furthermore, the lack of aryl-glycosidase activity described as a characteristic property of Avicelase II could be confirmed with the recombinant enzyme.

DISCUSSION

Avicelase II encoded by the *celY* gene of *Clostridium stercorarium* has previously been characterized as a thermoactive exo-1,4-β-glucanase capable of releasing

Table 2. Purification of CelY

Activity was determined under standard assay conditions with acid-swollen Avicel as the substrate.

Purification step	Volume (ml)	Protein (mg)	Activity (U)	Specific activity (mU mg^{-1})	Purification (-fold)	Yield (%)
Cell extract	150	3225	40·0	12	1·0	100
Heat-treated extract	142	1207	38·0	31	2·6	95
Q Sepharose FF	135	101	17·2	170	14·2	43

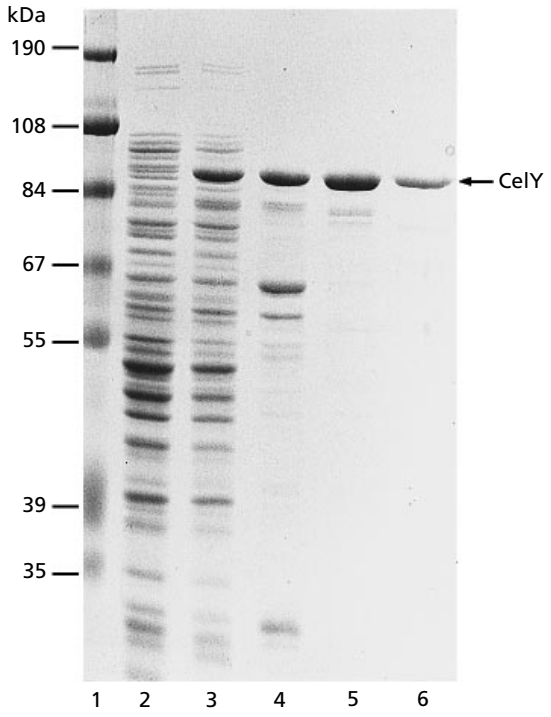

Fig. 4. SDS-PAGE of the fractions obtained during purification of recombinant CelY from *E. coli* cell extract. Lane 1, molecular mass markers; lane 2, crude extract of *E. coli* XL1-blue; lane 3, crude extract of *E. coli* XL1-Blue(pBS*celY*); lane 4, heat-treated extract; lane 5, pooled fractions from the Q Sepharose column; lane 6, Avicelase II purified from *C. stercorarium*.

Table 3. Substrate specificity of recombinant CelY purified from *E. coli*

Standard reaction mixtures as described in Methods were incubated for 48 h at 70 °C. The assay mixtures contained 75 μg of purified CelY. The individual values are means of triplicate assays.

Substrate	Specific activity (mU mg^{-1})
Avicel acid-swollen	170
Avicel	1·2
Cellulose MN300	0·4
Barley β-glucan	76
Xylan	1·6
Carboxymethylcellulose	0·8
p-nitrophenyl β-D-cellobioside	0
p-nitrophenyl β-D-xylopyranoside	0

cellobiose, cellotriose and cellotetraose from the non-reducing end of cellulose chains (Bronnenmeier *et al*., 1991). Sequence comparison revealed that CelY belongs to the recently proposed cellulase family L, as already supposed by virtue of the N-terminal amino acid sequence of Avicelase II (Shen *et al*., 1994). Four complete sequences and one partial sequence all showing a high degree of identity to the catalytic domain of CelY have been reported (Fig. 3). Three of the encoded enzymes, CelS from *Clostridium thermocellum* (Kruus *et al*., 1995), CbhB from *Cellulomonas fimi* (Shen *et al*., 1995) and, recently, CelF from *Clostridium cellulolyticum* (Reverbel-Leroy *et al*., 1996) have been characterized further. The reported enzymic properties are very similar to those determined for CelY and Avicelase II, respectively. In particular, the exoglucanolytic mode of action first described for Avicelase II was also found for these members of cellulase family L.

The *celY* gene is located in close vicinity to the endo-1,4-β-glucanase gene *celZ* (Jauris *et al*., 1990). Clustering of the two *cel* genes of *C. stercorarium* might be of functional significance since the encoded enzymes exhibit a synergistic cooperation in the degradation of microcrystalline cellulose (Creuzet *et al*., 1983; Bronnenmeier & Staudenbauer, 1988; K. Bronnenmeier, K. Riedel, J. Ritter & W. L. Staudenbauer, unpublished results). As shown in Fig. 3, the exoglucanase CelY and the family E endoglucanase CelZ of *C. stercorarium* both have a homologous counterpart in *Caldicellulosiruptor saccharolyticus*. However, in the latter organism both proteins are located on one polypeptide chain (Te'o *et al*., 1995). The N-terminal part of the multidomain cellulase CelA, designated CD1, was shown to exhibit endoglucanase activity. It consists of a catalytic domain belonging to cellulase family E flanked by a cellulose-binding domain with similarity to the domain C′ joining the catalytic domain of CelZ. The C-terminal part of CelA, designated CD2, contains a catalytic domain belonging to cellulase family L for which no activity could be detected. The two parts of the protein are separated by two cellulose-binding domains

with similarity to domain C found in the non-catalytic regions of the *C. stercorarium cel* genes. Thus, CelA of the extremely thermophilic *C. saccharolyticus* represents a naturally occurring fusion protein of the unique endo- and exoenzymes produced by the thermophilic bacterium *C. stercorarium*. By applying recombinant DNA technology, work is in progress to transplant the enzymes encoded by *celY* and *celZ* into one multienzyme. It will be of particular interest to evaluate the hydrolytic capability of the artificial construct in comparison with a synergistic admixture of CelY and CelZ.

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (grant SFB 145 to W. L. Staudenbauer and W. H. Schwarz; grant Br 1472/2 to K. Bronnenmeier).

REFERENCES

Béguin, P. & Aubert, J. P. (1994). The biological degradation of cellulose. *FEMS Microbiol Rev* **13**, 25–58.

Bronnenmeier, K. & Staudenbauer, W. L. (1988). Resolution of *Clostridium stercorarium* cellulase by fast protein liquid chromatography (FPLC). *Appl Microbiol Biotechnol* **27**, 432–436.

Bronnenmeier, K. & Staudenbauer, W. L. (1990). Cellulose hydrolysis by a highly thermostable endo-1,4-β-glucanase (Avicelase I) from *Clostridium stercorarium*. *Enzyme Microb Technol* **12**, 431–436.

Bronnenmeier, K., Rücknagel, K. P. & Staudenbauer, W. L. (1991). Purification and properties of a novel type of exo-1,4-β-glucanase (Avicelase II) from the cellulolytic thermophile *Clostridium stercorarium*. *Eur J Biochem* **200**, 379–385.

Bronnenmeier, K., Adelsberger, H., Lottspeich, F. & Staudenbauer, W. L. (1996). Affinity purification of cellulose-binding enzymes of *Clostridium stercorarium*. *Bioseparation* **6**, 41–45.

Creuzet, N., Berenger, J.-F. & Frixon, C. (1983). Characterization of exoglucanase and synergistic hydrolysis of cellulose in *Clostridium stercorarium*. *FEMS Microbiol Lett* **20**, 347–350.

Fujino, T., Karita, S. & Ohmiya, K. (1993). Nucleotide sequences of the *celB* gene encoding endo-1,4-β-glucanase-2, ORF1 and ORF2 forming a putative cellulase gene cluster of *Clostridium josui*. *J Ferment Bioeng* **76**, 243–250.

Gräbnitz, F., Rücknagel, K. P., Seiß, M. & Staudenbauer, W. L. (1989). Nucleotide sequence of the *Clostridium thermocellum bglB* gene encoding thermostable β-glucosidase B: homology to fungal β-glucosidases. *Mol Gen Genet* **217**, 70–76.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **16**, 557–580.

Jauris, S., Rücknagel, K. P., Schwarz, W. H., Kratzsch, P., Bronnenmeier, K. & Staudenbauer, W. L. (1990). Sequence analysis of the *Clostridium stercorarium celZ* gene encoding a thermostable cellulase (Avicelase I): identification of catalytic and cellulose-binding domains. *Mol Gen Genet* **223**, 258–267.

Jorgensen, P. L. & Hansen, C. K. (1990). Multiple endo-β-1,4-glucanase-encoding genes from *Bacillus lautus* PL236 and characterization of the *celB* gene. *Gene* **93**, 55–60.

Kruus, K., Wang, W. K., Ching, J. & Wu, J. H. D. (1995). Exoglucanase activities of the recombinant *Clostridium thermocellum* CelS, a major cellulosome component. *J Bacteriol* **177**, 1641–1644.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **277**, 680–685.

Meinke, A., Gilkes, N. R., Kwan, E., Kilburn, D. G., Warren, R. A. J. & Miller, R. C., Jr (1994). Cellobiohydrolase A (CbhA) from the cellulolytic bacterium *Cellulomonas fimi* is a β-1,4-exocellobiohydrolase analogous to *Trichoderma reesei* CBH II. *Mol Microbiol* **12**, 413–422.

Morag, E., Halevy, I., Bayer, E. A. & Lamed, R. (1991). Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. *J Bacteriol* **173**, 4155–4162.

Pagès, S., Belaich, A., Tardif, C., Reverbel-Leroy, C., Gaudin, C. & Belaich, J.-P. (1996). Interaction between the endoglucanase CelA and the scaffolding protein CipC of the *Clostridium cellulolyticum* cellulosome. *J Bacteriol* **178**, 2279–2286.

Rainey, F. A., Donnison, A. M., Janssen, P. H., Saul, D., Rodrigo, A., Berquist, P. L., Daniel, R. M., Stackebrandt, E. & Morgan, H. W. (1994). Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: an obligatory anaerobic, extremely thermophilic, cellulolytic bacterium. *FEMS Microbiol Lett* **120**, 263–266.

Reverbel-Leroy, C., Belaich, A., Bernadac, A., Gaudin, C., Belaich, J.-P. & Tardif, C. (1996). Molecular study and overexpression of the *Clostridium cellulolyticum celF* cellulase gene in *Escherichia coli*. *Microbiology* **142**, 1013–1023.

Sakamoto, R., Arai, M. & Murao, S. (1984). Enzymatic properties of hydrocellulase from *Aspergillus aculeatus*. *J Ferment Technol* **62**, 561–567.

Sakka, K., Kojima, Y., Kondo, T., Karita, S., Ohmiya, K. & Shimada, K. (1993). Nucleotide sequence of the *Clostridium stercorarium xynA* gene encoding xylanase A: identification of catalytic and cellulose binding domains. *Biosci Biotechnol Biochem* **57**, 273–277.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.

Schwarz, W. H., Jauris, S., Kouba, M., Bronnenmeier, K. & Staudenbauer, W. L. (1989). Cloning and expression of *Clostridium stercorarium* cellulase genes in *Escherichia coli*. *Biotechnol Lett* **11**, 461–466.

Sedmak, J. J. & Grossberg, S. E. (1977). A rapid, sensitive assay for protein using Coomassie brilliant blue G250. *Anal Biochem* **79**, 544–552.

Shen, H., Tomme, P., Meinke, A., Gilkes, N. R., Kilburn, D. G., Warren, R. A. J. & Miller, R. C., Jr (1994). Stereochemical course of hydrolysis catalysed by *Cellulomonas fimi* CenE, a member of a new family of β-1,4-glucanases. *Biochem Biophys Res Commun* **199**, 1223–1228.

Shen, H., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr & Warren, R. A. J. (1995). Cellobiohydrolase B, a second exo-cellobiohydrolase from the cellulolytic bacterium *Cellulomonas fimi*. *Biochem J* **311**, 67–74.

Shoseyov, O., Takagi, M., Goldstein, M. A. & Doi, R. H. (1992). Primary sequence analysis of *Clostridium cellulovorans* cellulose binding protein A. *Proc Natl Acad Sci USA* **89**, 3483–3487.

Te'o, V. S. J., Saul, D. J. & Bergquist, P. L. (1995). *celA*, another gene coding for a multidomain cellulase from the extreme thermophile *Caldocellum saccharolyticum*. *Appl Microbiol Biotechnol* **43**, 291–296.

Wang, W. K., Kruus, K. & Wu, J. H. D. (1993). Cloning and DNA sequence of the gene coding for *Clostridium thermocellum*

cellulase S_s (CelS), a major cellulosomal component. *J Bacteriol* **175**, 1293–1302.

Wood, T. M. & Bhat, K. M. (1988). Methods for measuring cellulase activities. *Methods Enzymol* **160**, 87–112.

Yanish-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18pUC vectors. *Gene* **33**, 103–119.

Received 12 August 1996; revised 10 October 1996; accepted 14 October 1996.