Isolation and Expression of the *xynB* Gene and Its Product, XynB, a Consistent Component of the *Clostridium cellulovorans* Cellulosome

Sung Ok Han,[1] Hideaki Yukawa,[2] Masayuki Inui,[2] and Roy H. Doi[1]*

*Section of Molecular and Cellular Biology, University of California, Davis, California,[1] and
Research Institute of Innovative Technology for the Earth, Kyoto, Japan[2]*

The nucleotide sequence of the *Clostridium cellulovorans xynB* gene, which encodes the XynB xylanase, consists of 1,821 bp and encodes a protein of 607 amino acids with a molecular weight of 65,976. XynB contains a typical N-terminal signal peptide of 29 amino acid residues, followed by a 147-amino-acid sequence that is homologous to the family 4-9 (subfamily 9 in family 4) carbohydrate-binding domain. Downstream of this domain is a family 10 catalytic domain of glycosyl hydrolase. The C terminus separated from the catalytic domain by a short linker sequence contains a dockerin domain responsible for cellulosome assembly. The XynB sequence from mass spectrometry and N-terminal amino acid sequence analyses agreed with that deduced from the nucleotide sequence. XynB was highly active toward xylan, but not active toward carboxymethyl cellulose. The enzyme was optimally active at 40°C and pH 5.0. Northern hybridizations revealed that *xynB* is transcribed as a monocistronic 1.9-kb mRNA. RNA ligase-mediated rapid amplification of 5′ cDNA ends by PCR (RLM-5′RACE PCR) analysis of *C. cellulovorans* RNA identified a single transcriptional start site of *xynB* located 47 bp upstream from the first nucleotide of the translation initiation codon. Alignment of the *xynB* promoter region provided evidence for highly conserved sequences that exhibited strong similarity to the σ^A consensus promoter sequences of gram-positive bacteria. Expression of *xynB* mRNA increased from early to middle exponential phase and decreased during the early stationary phase when the cells were grown on cellobiose. No alternative promoter was observed by RLM-5′RACE PCR and reverse transcriptase PCR analyses during expression. The analysis of the products from xylan hydrolysis by thin-layer chromatography indicated its endoxylanase activity. The results suggest that XynB is a consistent and major cellulosomal enzyme during growth on cellulose or xylan.

Hemicellulose is the second most abundant renewable polysaccharide in nature after cellulose (3). β-1,4-Xylan is a major component of hemicellulose and has a backbone of β-1,4-linked D-xylopyranoside residues replaced by acetyl, arabinosyl, and uronyl side chains (50). The complete breakdown of xylan requires the action of several hydrolytic enzymes (3). The most important hydrolytic enzyme is endoxylanase (1,4-β-D-xylan xylanohydrolase) (EC 3.2.1.8). Xylanases have been widely detected in bacteria and fungi, and their properties have been characterized (3, 52). On the basis of the amino acid sequences of catalytic domains, xylanases have been classified into two groups: family 10 and family 11 of glycosyl hydrolases (25).

Clostridium cellulovorans ATCC 35296 (44) produces a highly active polysaccharolytic complex termed the cellulosome, in which several cellulases are tightly bound to a scaffolding protein called CbpA (1). Doi et al. (11) and Tamaru et al. (11, 49) have characterized the genes necessary for degradation of the plant cell wall of this bacterium. However, until now, only one xylanase gene, *xynA*, encoding a family 11 glycosyl hydrolase, had been cloned and expressed in *Escherichia coli* in agreement with the presence of xylanase activity in the cellulosome (27, 28).

In this paper, we describe the nucleotide sequence, expression, and characterization of the *xynB* gene, a family 10 glycosyl hydrolase, encoding one of the consistent components of the cellulosome and corresponding to component P6 of a subpopulation of the *C. cellulovorans* cellulosome (35). In addition, we performed a detailed transcriptional analysis of the *xynB* promoter and its regulation by growth phase in order to gain some understanding of the pattern of expression of this gene in *C. cellulovorans*. We also characterized the XynB enzyme purified from a recombinant *E. coli* strain.

MATERIALS AND METHODS

Bacterial strains and fosmid. *C. cellulovorans* ATCC 35296, used for isolation of the cellulosome fraction, was described previously (42). The *E. coli* strain used was EPI300 (Epicentre). A fosmid (Epicentre) containing 50-kb DNA fragments was used for determination of the DNA sequence of *xynB* and the production of the recombinant enzyme.

Nucleic acid isolation. Chromosomal DNA of *C. cellulovorans* was isolated by using the genomic DNA purification kit (Promega) according to the manufacturer's instructions. Total RNA was extracted from *C. cellulovorans* broth cultures using the QIAGEN RNeasy kit with the additional step of treatment with RNA*later* solution (Ambion) and RNase-free DNase (Promega) according to the manufacturers' instructions.

Fosmid library construction and screening. The *C. cellulovorans* fosmid library was constructed as follows. Briefly, genomic DNA was randomly sheared by passing it through a syringe needle. The ends of the sheared DNA were then repaired, and the DNA fragments were separated on 0.8% agarose in 1× TAE (1× TAE is 40 mM Tris, 20 mM acetic acid, and 1 mM EDTA) overnight at 30 V. Then 50-kb fragment pools were gel purified and cloned into the vector EpiFOS (Epicentre) according to the manufacturer's instructions. Ligated DNA was packaged with the Epicentre MaxPlax lambda packaging extract and used to transfect *E. coli* EPI300 cells (Epicentre). Transfected cells were selected on Luria-Bertani (LB) agar containing 12.5 μg of chloramphenicol per ml. The procedure yielded 2,000 recombinant clones. The LB agar plates, on which recombinant *E. coli* bacteria were grown, were overlaid with soft agar containing 0.3% xylan (birchwood; Sigma) and 0.7% agar in 25 mM sodium acetate buffer (pH 6.0). After incubation at 37°C for 16 h, the plates were stained with 0.3%

* Corresponding author. Mailing address: Section of Molecular and Cellular Biology, University of California, Davis, CA 95616. Phone: (530) 752-3191. Fax: (530) 752-3085. E-mail: rhdoi@ucdavis.edu.

TABLE 1. Oligonucleotide primers used in this study

Primer	Sequence (5′→3′)	Gene	Nucleotide position[a]	Use[b]
xynB-F	GTATTATATTTCTATAAAAGACC	*xynB*	269–291	NB, RT
xynB-R	TGCTTTCAATACGTCTGTTC	*xynB*	1027–1051	NB, RT
xynB-F1	GTTCTTAAGTTCACAAATACTCGAG	*xynB*	46–70	RT
xynB-F2	GATTTTCATATTATTATACTTAATTTTA	*xynB*	198–228	RT
5′-Inner	CGAAGGTGCTACTTATTAATAGAG	*xynB*	410–433	RACE
5′-Outer	TGCTTTCAATACGTCTGTTC	*xynB*	563–582	RACE
xynA-F	TGTTAGCCTCTTCTGC	*xynA*	6389–6404	NB
xynA-R	GATTCCAAGTGCCATAGC	*xynA*	6722–6739	NB
cbpA-exgS-F	ATGCAAAAAAAGAAATCGCTG	*cbpA-exgS*	883–903	NB
cbpA-exgS-R	GGTTGATGTTGGGCTTGCTGTT	*cbpA-exgS*	1215–1236	NB
arfA-F	ATGGAGGATTTTGGGTTGGG	*arfA*	553–572	NB
arfA-R	TCGGTGACTCTCCATC	*arfA*	809–906	NB

[a] The nucleotide position refers to the relevant position within the *xynB* sequence (GenBank accession no. AY604045), *xynA* sequence (28) (GenBank accession no. AF435978), *cbpA-exgS* sequence (43) (GenBank accession no. M73817), or *arfA* sequence (26) (GenBank accession no. AY128945).

[b] NB, Northern blot analysis; RT, RT-PCR; RACE, RLM-5′ RACE.

Congo red and destained with 1 M NaCl. The clones that formed halos were selected as xylanase-positive colonies. The xylanase-positive clones were picked and restreaked to confirm the formation of clearing zones around the colonies on LB agar overlaid with soft agar containing xylan.

Construction of the mutant library and DNA sequencing. The EZ::TN <*ori*V/KAN-2> insertion system (Epicentre), based on in vitro Tn5 transposition (18), was used to generate insertion knockout mutants of *xynB*. The xylanase-positive clone was incubated with transposase and an EZ::TN <*ori*V/KAN-2> transposon with a kanamycin resistance marker under the conditions suggested by the manufacturer (Epicentre). After the reaction, the reaction mixture was used to transform TransforMax EPI300 electrocompetent *E. coli*, and the knockout (xylanase-negative) transformants were selected on LB plates with 0.3% xylan, 50 μg of kanamycin per ml, and 12.5 μg of chloramphenicol per ml. The transposon insertion sites of mutant clones were mapped by nucleotide sequencing with the transposon-specific flanking primers (Epicentre).

RLM-RACE PCR. RNA ligase-mediated rapid amplification of cDNA ends by PCR (RLM-RACE PCR) was performed with total RNA extracted from a *C. cellulovorans* culture grown on cellobiose and was used to determine single or multiple transcription start points. Rapid amplification of 5′ cDNA ends (5′RACE) was performed using the FirstChoice RLM-RACE kit (Ambion) according to the manufacturer's instructions with the following exceptions. The nested PCR conditions for 5′ outer PCR were as recommended, with 10 pmol of gene-specific outer primer *xynB*-5′-Outer (5′-GAACAGACGTATTGAAAG CA), 1.25 U of SuperTaq polymerase (Ambion), 10 pmol of 5′RACE outer primer (5′-GCTGATGGCGATGAATGAACACTG) (Ambion), 1× SuperTaq PCR buffer (Ambion), 100 μM deoxynucleoside triphosphates, 1 ng of first-strand cDNA reaction mixture per μl, and H₂O to a volume of 50 μl. PCR was performed as follows: (i) 4 min at 94°C; (ii) 35 cycles, with 1 cycle consisting of 30 s at 94°C, 30 s at 60°C, and 40 s at 72°C; and (iii) 7 min at 72°C. 5′ inner PCR was performed, as recommended, with 10 pmol of gene-specific inner primer *xynB*-5′-Inner (5′-CTCTATTAATAAGTAGCACCTTCG) and 10 pmol of 5′RACE inner primer (5′-CGCGGATCCGAACACTGCGTTTGCTGGCTTT GATG) (Ambion) using the same conditions as for the 5′ outer PCR. PCR products were run on 2% agarose gels and sequenced.

RT-PCR analysis. Reverse transcriptase (RT) reactions were performed in a final volume of 20 μl, which contained 5 mM MgCl₂, RT buffer (10 mM Tris-HCl [pH 9.0], 50 mM KCl, 0.1% Triton X-100), 1 mM each of the four deoxynucleoside triphosphates, 1 U of recombinant RNasin RNase inhibitor (Promega), 15 U of avian myeloblastosis virus RT (Promega), 0.25 μM oligonucleotide primer, and 10 μg of substrate RNA. Reaction mixtures were incubated at 42°C for 60 min, and reactions were terminated by heating the mixtures at 95°C for 5 min, followed by incubation on ice for 5 min. The cDNA products were then amplified in 50-μl PCR mixtures using 1 μl of the RT reaction mixture as the template.

Northern blot analysis. RNA samples (5 μg) were denatured in RNA sample buffer at 65°C for 10 min. RNA sample buffer consists of 250 μl of formamide, 83 μl of 37% (wt/vol) formaldehyde, 83 μl of 6× loading dye (Promega), 50 μl of 10× MOPS (morpholinepropanesulfonic acid) buffer (10× MOPS buffer is 20 mM MOPS, 5 mM sodium acetate, and 1 mM EDTA [pH 7.0]), and 34 μl of distilled water. RNA samples were separated through 1% agarose gels in MOPS buffer with 2% (vol/vol) formaldehyde. DNA probes were synthesized by PCR using specific oligonucleotides derived from the *C. cellulovorans* sequence as a template (Table 1). The probes were nonradioactively labeled by random prim-

ing using the digoxigenin (DIG) High Prime kit (Roche). To add the correct amount of probe for hybridization, a serial dilution of each probe (0.05 to 10 pg) was spotted on a nylon membrane, and the labeling sensitivity (amount of labeled DNA per spot) was determined. RNA was transferred overnight to a positively charged nylon membrane (Roche) by capillary transfer using 20× SSC (20× SSC is 0.3 M NaCl plus 0.03 M sodium citrate [pH 7]). Hybridization was performed for 16 h at 50°C in DIG Eazy Hyb buffer solution (Roche). The membrane was washed and specific transcripts on the blots were detected by using the DIG luminescence detection kit (Roche) according to the protocol recommended for the kit.

Preparation of the recombinant enzyme. The recombinant enzyme was isolated from culture supernatants of a xylanase-positive clone or a knockout clone. The culture supernatants were obtained by centrifugation of cultures grown overnight (1,000 ml) in LB medium containing chloramphenicol (12.5 μg/ml). The supernatants were concentrated by 80% ammonium sulfate saturation and dialyzed. The dialyzed enzyme was concentrated with Ultrafree Biomax (10-kDa cutoff; Millipore). The concentration of fractionated protein was measured by the method of Bradford (5) with a protein assay kit from Bio-Rad, using bovine serum albumin as the standard.

SDS-PAGE, zymogram, and mass spectrometry analyses. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed with a 10% polyacrylamide gel by the method of Laemmli et al. (29). Proteins were fixed in the gels by soaking in a solution containing 40% (vol/vol) methanol and 10% (vol/vol) acetic acid for approximately 1 h and subsequently visualized by Coomassie blue staining (Genomic Solutions). The zymograms for xylanase and carboxymethyl cellulase were performed by using a 0.1% (wt/vol) concentration of each substrate incorporated into the polyacrylamide. After SDS-PAGE, the gels were renatured in renaturation buffer (100 mM succinic acid, 10 mM CaCl₂, 1 mM dithiothreitol [pH 6.3]) for 2 h at 25°C with gentle shaking. The renatured gel was then incubated in fresh renaturation buffer for 1 h at 37°C with gentle shaking. The clearing zones corresponding to enzyme activities were visualized with 0.3% (wt/vol) Congo red (stained for 10 min and destained with 1 M NaCl solution) (4). Mass spectrometric analysis was performed to identify the XynB protein on cellulosomes, and recombinant proteins were separated by SDS-PAGE as described previously (19).

Analysis of reaction products. Xylan degradation products were determined by thin-layer chromatography (TLC) on precoated TLC sheets (silica gel; Whatman) with acetone-ethyl acetate-acetic acid (2:1:1, vol/vol/vol) (34). The plates were visualized by spraying with a 1:1 (vol/vol) mixture of 0.2% methanolic orcinol and 20% sulfuric acid.

Nucleotide sequence accession number. The *xynB* gene sequence was submitted to GenBank and has been assigned accession number AY604045.

RESULTS AND DISCUSSION

Analysis of DNA sequence and transcriptional start site of *C. cellulovorans* xynB. Figure 1 shows the *C. cellulovorans* *xynB* coding sequence and flanking regions. There is an open reading frame composed of 1,821 nucleotides encoding a protein of 607 amino acids with a predicted molecular weight of 65,976.
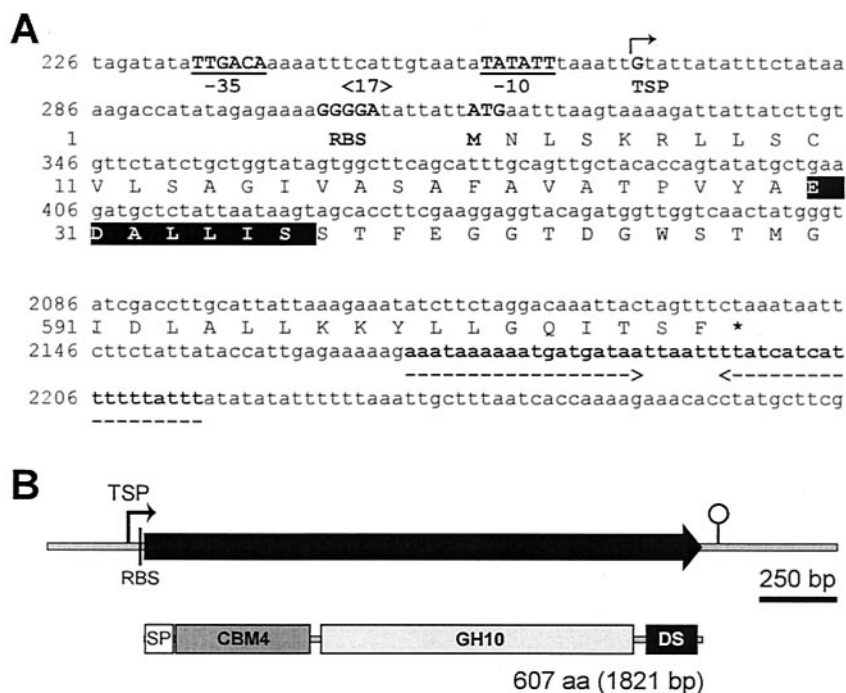
FIG. 1. Nucleotide and deduced amino acid sequences (A) and modular structure (B) of *xynB*. (A) Untranslated regions (lowercase letters) and putative promoter regions (consensus −10 and −35 regions, gap 17 bp) (underlined) are indicated. The transcription start point (+1) (TSP) is indicated by the bent arrow. The putative ribosome binding site (RBS) and the first codon (M) are indicated. The stop codon (asterisk) and a palindrome (broken arrows facing each other below the sequence) are indicated. The deduced amino acid sequence corresponding to the experimentally determined N-terminal sequence (35) is shown by white letters on a black background. Additional sequences upstream and downstream of *xynB* can be accessed at GenBank accession no. AY604045. (B) The putative signal peptide (SP) identified using SignalP prediction software (37) is boxed. A family 4–9 carbohydrate-binding domain (CBM4) is shaded and followed by a lightly shadowed family 10 catalytic domain of glycosyl hydrolases (GH10). The DS (or dockerin) domain is shown. aa, amino acids.

The assigned ATG initiation codon at nucleotide position 316 is preceded by a conserved putative ribosome binding site (GGGGA). The reading frame is terminated by the stop codon TAA at position 2136.

In order to localize the promoter, the transcription start site of *xynB* was determined by RLM-5′RACE PCR analysis. An 165-bp PCR fragment (5′-Inner/5′RACE inner primer) was generated for DNA sequencing of the region upstream of the transcription start site of *xynB*. One clear ending sequence was identified 47 bp upstream from the A of the proposed ATG start codon, and it was obtained by DNA sequencing of the RLM-5′RACE PCR product (Fig. 2A). In addition, this transcription start point was confirmed by RT-PCR analysis using multiple primer combinations (xynB-F1/xynB-R, xynB-F2/xynB-R, and xynB-F/xynB-R) (Fig. 2B and Table 1). The *xynB* mRNA start point suggested a putative promoter sequence, TTGACA and TATATT, with a 17-bp spacing between them (Fig. 2C). These sequences show high homologies to the consensus promoter sequences for σ^{70} factor found in *E. coli*, i.e., TTGACA and TATAAT with a 17-bp spacing (39). The consensus sequence of *xynB* and previously determined *C. cellulovorans* promoters (21) at −10 (5′-T-GAAA-3′) and at −35 (5′-TATAAT-3) are homologous to elements recognized by *Bacillus subtilis* (σ^A) and *E. coli* (σ^{70}) RNA polymerases (7, 56) (Fig. 2C), and the consensus sequence contained only one mismatch to the *E. coli* σ^{70} −10 consensus sequence (23). The spacing between the −35 and −10 regions was 17 bp, which is the optimal spacing observed for *B. subtilis* and *E. coli* pro-

moter sequences (7, 22). The close similarity of the promoter to the consensus σ^A sequence suggests that the *xynB* promoter, if not subjected to any regulatory constraints, would act as a strong promoter in vivo (10, 23). In addition, the *xynB* promoter was very similar to previously determined *C. cellulovorans* promoters (1, 21).

Northern hybridizations of the RNA with a *xynB* probe showed a single transcript of 1.9 kb (data not shown). The size of this mRNA agreed with the size of the *xynB* gene (1,821 bp) and indicated that *xynB* was a monocistronic gene. On the basis of the size of *xynB* mRNA and the location of the transcription start site, a putative transcription terminator that consists of a 44-bp palindromic sequence, corresponding to an mRNA hairpin loop with a ΔG of −67.6 kcal/mol (6, 45), followed by a long poly(U) tail (AUAUAUAUUUUUU) was found downstream of the TGA termination codon at nucleotide position 2171 (Fig. 1A). This structure is similar to the rho factor-independent terminator of *E. coli* (39).

Transcriptional analysis of *xynB*. To find alternative promoters and determine whether the *xynB* gene along with *C. cellulovorans* genes are regulated coordinately, changes in the level of mRNA expressed were monitored during cultivation of *C. cellulovorans* on cellobiose as the sole carbon source. RNA was prepared from the culture at different stages of growth. During the culture for 30 h, the pH was not changed and 0.5% (wt/vol) cellobiose was not consumed totally by *C. cellulovorans* (data not shown). The RNA was subjected to Northern blot analyses using primers that were specific to the *xynB*, *xynA*,
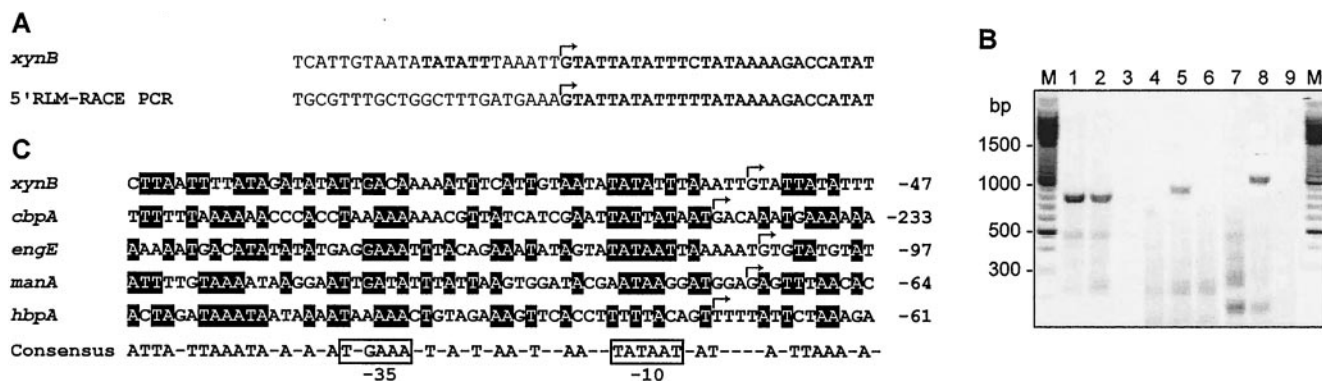
A

```
xynB            TCATTGTAATATATATTTAAATTGTATTATATTTCTATAAAAGACCATAT
5'RLM-RACE PCR  TGCGTTTGCTGGCTTTGATGAAAGTATTATATTTTTATAAAAGACCATAT
```

B

C

```
xynB      CTTAATTTTATAGATATATTGACAAAAATTTCATTGTAATATATATTTAAATTGTATTATATTT   -47
cbpA      TTTTTTAAAAAACCCACCTAAAAAAAAACGTTATCATCGAATTATTTATAATGACAAATGAAAAA  -233
engE      AAAAATGACATATATATGAGCAAATTTACAGAAATATAGTATATAATTAAAAATGTGTATGTAT   -97
manA      ATTTTGTAAAATAAGGAATGCATATTTATTAAGTGGATACGAATAAGGATGGAGAGTTTAACAC   -64
hbpA      ACTAGATAAAATAATATAAAATAAAAAACTGTAGAAAGTTCACCTTTTTTACAGTTTTTATTCTAAAGA -61
Consensus ATTA-TTAAATA-A-A-AT-GAAA-T-A-T-AA-T--AA--TATAAT-AT----A-TTAAA-A-
                                  -35                      -10
```

FIG. 2. Determination and alignment of putative *xynB* promoter located −10 and −35 bases upstream of the start point of transcription. (A) The transcription start point was determined by sequencing the RLM-5′RACE PCR product and indicated by the bent arrows. (B) Agarose gel electrophoresis of RT-PCR products. The oligonucleotide primers used in RT-PCR analysis of promoter region were xynB-F1 and xynB-R (lanes 1 to 3), xynB-F2 and xynB-R (lanes 4 to 6), and xynB-F and xynB-R (lanes 7 to 9) (more details in Table 1). RT-PCR analysis was performed using total RNA (lanes 1, 4, and 7) that was isolated from *C. cellulovorans* in the medium containing cellobiose. In controls, the reactions were performed with DNA template (lanes 2, 5, and 8) or in the absence of RT (lanes 3, 6, and 9). The positions of molecular size markers (M) (in base pairs) are shown to the left of the gel. (C) Transcription start points are indicated by the bent arrows. The nucleotide numbering begins from the first codon shown on the right. The consensus sequence derived from this alignment is given at the bottom. It consists of nucleotides that are present in any given position in more than 50% of the sequences. Promoter sequence nucleotides that match those of the consensus sequence are shown as white letters on a black background. The coordinates are those in the published nucleotide sequences for *cbpA* (43) (GenBank accession no. M73817), *engE* (48) (GenBank accession no. AF105331), and *manA* and *hbpA* (47, 49) (GenBank accession no. AF132735).

cbpA-exgS, and *arfA* genes. These genes represent the major subunits (*cbpA-exgS*) (17, 32, 43), xylanase gene (*xynA*) (28), and a noncellulosomal α-L-arabinofuranosidase gene, *arfA* (26). A semiquantitative measure of the level of *xynB* mRNA, using DIG-labeled probes and RNA isolated at different times during cell growth, was obtained by Northern blot analysis (Fig. 3A and B). The intensities of the bands were close approximations of their relative abundance. The levels of *xynB* mRNA increased simultaneously from early to middle exponential phase and dramatically decreased during the early stationary phase when the cells were grown on cellobiose (Fig. 3A, points 1 to 9, and B, lanes 1 to 9). Like *xynB* gene expression, cells contained high levels of *cbpA*, *exgS*, and *xynA* mRNAs during most of the exponential growth phase with the level peaking at the middle of the exponential phase (Fig. 3A and B). In contrast, the *arfA* transcripts increased during the middle of the stationary phase (Fig. 3A and B, lane 10) and peaked when the cells reached the stationary growth phase (Fig. 3A and B, lanes 9 and 10).

RLM-5′RACE PCR analysis was performed using primers 5′-Inner and 5′-Outer (Fig. 1A), which were specific for identifying the putative promoter. One clear band (165 bp) of the nested mRNA product from the RLM-5′RACE PCR was found during the early to late exponential phase (Fig. 3C) and disappeared during the stationary phase (Fig. 3A and C). The fact that nested PCR did not produce several bands indicates no alternative transcriptional start site. A negative control using the RNA that was not reverse transcribed was included to ensure that there was no contaminating DNA present in the RNA preparation (data not shown).

The mRNA of *xynB* and the mRNAs of *cbpA-exgS* and *xynA* are coordinately expressed, whereas the time course of expression of these genes (*xynB*, *cbpA-exgS*, and *xynA*) was greatly different from that of the noncellulosomal α-L-arabinofuranosidase gene (*arfA*). The transcriptional data also showed that whereas *xynB* and *cbpA-exgS* from cells grown on cellobiose were efficiently expressed, relatively low levels of *xynA* and *arfA* transcripts were present. In addition, the *xynB* gene and major cellulase genes, such as *cbpA-exgS* and *engE*, were constitutively expressed (19–21). On the basis of a general model for induction of xylanase expression in *Trichoderma reesei*, a sensor enzyme may be constitutively expressed; this enzyme hydrolyzes xylan into oligosaccharides that can enter the bacterium and activates the expression of all the xylanase genes (33, 57). It is possible that one or more of the constitutively expressed *C. cellulovorans* xylanases functions as the sensor enzyme. Cellulose and xylan are closely associated in nature, and it appears that *C. cellulovorans* has mechanisms to ensure efficient utilization of both types of polymers. Previous studies also suggest that a common regulatory mechanism may exist at the transcriptional level for cellulase and xylanase induction by cellobiose in this organism (19, 20). Cellulose metabolites, such as cellobiose, or a derivative of cellobiose may act as an inducer and may bind to a receptor protein in a signal transduction pathway, and this pathway would then lead to cellulase induction.

Amino acid sequences and domains of XynB. A previous study of *C. cellulovorans* showed that the bacterium produces high levels of xylanase activity when grown on xylan and concluded that at least four xylanase genes were located on the *C. cellulovorans* chromosome (27, 35). The xylanase subunit, XynA, was found recently, and its xylanase activity is ascribed to a family 11 catalytic domain (28). In this work, we have located *xynB* from the fosmid clone which contained a 50-kb insert DNA without *xynA*. High sequence identity to XynB is observed with XynC of *C. thermocellum* (44%) (24), XynB of *Polyplastron multivesiculatum* (38%) (9), and XynA of *Thermotoga neapolitana* (35%) (59).

Figure 1B shows schematically the molecular architecture of XynB. The deduced N-terminal sequence of 29 amino acids contains a sequence similar to the signal peptide sequences found in prokaryotic secretory proteins, which have a short
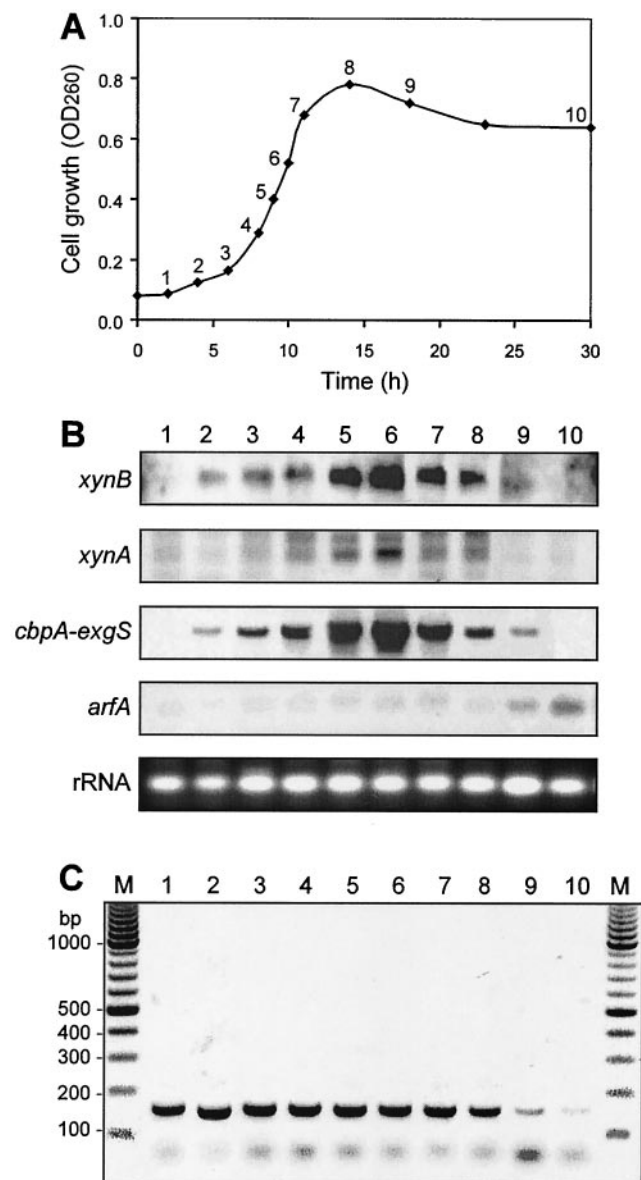
FIG. 3. Relative levels (B) and size (C) of transcripts of *xynB* at different growth phases (A) on cellobiose culture. (A) *C. cellulovorans* growth curve. OD260, optical density at 260 nm. (B) Northern blot analyses were conducted with the same concentration of RNA (5 μg) isolated from *C. cellulovorans* cultured grown on 0.5% cellobiose as the sole carbon source. Ethidium bromide staining of rRNA is shown as a loading control. (C) RLM-5′RACE PCR analysis for putative promoter location. Primers specific for *xynB* were used to amplify fragments by PCR. In negative controls, the reactions were performed in the absence of RT or RNA templates (data not shown). The numbers of the lanes in panels B and C are the numbers of the growth curve points in panel A. See Materials and Methods for details. The positions of molecular size markers (M) (in base pairs) are shown to the left of the gel.

region rich in positively charged amino acids (lysine-5 and arginine-6), followed by a sequence of predominantly hydrophobic residues (leucine-7, -8, and -12; valine-11, -17, and -23; alanine-14, -18, -20, -22, and –24; and isoleucine-16), a residue breaking the secondary structure (proline-26), and a cleavage site ending with alanine-29 (37).

Comparison of the amino acid sequence of XynB with those registered in SWISS PROT and GenBank databases clearly revealed that the mature XynB protein consisted of three distinct functional domains. The N-terminal domain of the mature form of XynB, 147 amino acid residues downstream of the signal peptide, is homologous with the family 4-9 (subfamily 9 in family 4) carbohydrate-binding domain of other xylanases (Fig. 4A), i.e., 37% sequence identity with XynC of *Clostridium thermocellum* (24), 33% identity with XynD of *Ruminococcus flavefaciens* (14), and 32% identity with XynA of *Thermotoga maritima* (55).

The family 10 catalytic domain of XynB, extending from positions 193 to 532, exhibited extensive sequence homology with the catalytic domains of the other xylanases in family 10 glycosyl hydrolase domains (Fig. 4B), i.e., 52% sequence identity with XynA of *R. flavefaciens* (58), 45% identity with XynY of *C. thermocellum* (15), and 38% identity with XynB of *Butyrivibrio fibrisolvens* (31). The catalytic domains of these enzymes are highly conserved in eight regions and presumably serve to form the overall structure of these enzymes (Fig. 4B), suggesting that these xylanases have diverged from a common evolutionary ancestor. In addition, two glutamic acid residues (Fig. 4B), which have been confirmed as the catalytic amino acids by site-directed mutagenesis and crystallography in *Cellulomonas fimi*, *C. thermocellum*, and *Streptomyces lividans* (8, 51, 53, 54), were conserved in two identical regions as the nucleophile and the proton donor, and it is assumed that the highly conserved histidine-427 of XynB presumably serves to stabilize the transition state with an oxocarbonium ion character (41).

The third domain in XynB is a dockerin domain located in the C terminus of the peptide. Dockerins that consist of a pair of well-conserved 22-residue repeats are highly conserved in cellulases and xylanases from *C. cellulovorans* (Fig. 4C) and play a role in cellulosome assembly by docking the various catalytic subunits to the scaffolding protein CbpA (11, 12, 38). Murashima et al. (35) fractionated the cellulosome of *C. cellulovorans* into subpopulations by ion-exchange chromatography. One subpopulation with high cell wall degrading activity contained an enzyme subunit called P6 with strong xylanase activity, and the N-terminal amino acid sequence of P6 was identical to that of XynB reported in this study. We also found XynB on cellulosomal enzymes isolated from different culture conditions, for example, not only in cultures grown on xylan as the sole carbon source but also in cultures grown on cellulose and cellobiose (data not shown). These findings along with the presence of a dockerin domain suggest that XynB is a consistent and major catalytic subunit of the cellulosome of *C. cellulovorans*.

Identification and characterization of XynB in recombinant protein. The gene product of *xynB* was recovered from the supernatant of *E. coli* EPI300 grown on LB by ammonium sulfate precipitation. The cellulosomal proteins were purified from *C. cellulovorans* grown on xylan-containing medium by a cellulose affinity column, HiLoad 26/60 Superdex 200 prep grade column and RESOURCE Q (35, 42). We analyzed the recombinant proteins by zymogram analysis and compared them with the cellulosomal protein. The recombinant proteins gave one major band in the zymogram with xylan, and the molecular weight of the enzyme was estimated to be 65,000

A

```
C.v-XynB  ( 34) LISSTEGGGTDGWSTM---GTTTISPSTLS-HSGNGSIYVSGRAASWAG-PCNIRTDVLKAGNTYNLSSYVMYNDASAADTQPISFMTKYTDSTNK-AF
C.t-XynC  ( 35) LIYDDFETG-LNGWGPR---GPETVELTTEEAYSGRYSLVSGRTSTWNG-PMVDKTDVLTIGESVKLGVYVKFVGDSYSNEQRFSLQLQYNDGAG--DV
T.s-XynA  (199) -ANETFENGNTSGWIGT---GSSVVKAVYGVAHSGDYSILTTGRTANWNG-PSYDLTGKIVPGQQYNVD--FWVKFVNGDTEQTKATVKATSDKDN---
T.m-XynA  (206) IYETSFENG-VGDWQPR---GDVNIEASSEVAHSGKKSIFISNRQKGWQG-AQINLKGILKTCKTYAFEA---WYQNSGQDQTHIMTQRKYSSDASTQ
C.s-XynB  (357) ---CNFESG-TAEDFVARGSGVTVTVVDNVYYHSETKALYVSGRADTWQG-AQIDMTNLLEKGKDQOES---IWYQNSGSDQKITLTYQRKNADDS-TN
R.a-XynA  (270) LYSATFENG-TNYWSGR---GSASWASSSSKAYECSKSIYVSGRTDNWNGGEVEVNTSTFKPCNSWSFS-----AMVNPAESTTVQLSMQY-DQNGT-TK
C.t-XynY  (568) YYHDTFEGS-VGQWTAR---GPAEVLLSGRTAYKCSESILVRNRIAAWNGAQRALNPRTFVPGNTYCFSV--VASFIEGASSTTFCMKLQYVDGSCT-QR
R.f-XynD  (261) YLKENFESG-EGNWSGR---GSAKVKSSSG--YDCTKGIFVSGREDTWNGASINLDELTFKAGETYSLGT--A-VMQDFESSVDFKLTLQYTDADCK-EN
                          **   **              *               *         *   **    * **    *       *
```

```
C.v-XynB  ( 94) YVPIATVTVN-KGEWTKIENTAYKIPAFAT--SAMIYWETTG-------TMINYYVDDVTAYGP
C.t-XynC  (128) YQNIKTATVY-KGTWTLEG-QLTVFSHAK--DVKIYVETEFKNSPSPQDLMDFYIDDFTATPA
T.s-XynA  (289) MIQVNDFANVNKGEWTEIKG-SFTLEVADYS-GISIYVESQN-------PTLEFYIDDFSVIGE
T.m-XynA  (298) MEWIKSATVP-SGQWVQLSG-IYTIPAGVTVEDLTLYFESQN-------PTLEFYVDDVKIVDT
C.s-XynB  (449) YDTIKWQQTVPSGVWTEVSG-SYTVFQTAT--QLIFYVESPN-------ATLDFYLDDFTVIDK
R.a-XynA  (359) YTQTAEGSCT-AGKWTKIENTSFTIPSGAS--NVKMYVEAPD-------SLCDIYVDRAIIADK
C.t-XynY  (661) YDTTDMKTVG-PNQWVHLYNPQYRIPSDAT--DMYVYVETAD-------DTINFYIDEAIGAVA
R.f-XynD  (351) YDEVKTVTVAA-KGQWVDLSNSSYTIPSGAT--GLVLYVEVPE-------SKTDFYMDGAYAGVK
                          *         *   *       *   *                   **   *
```

B

```
C.v-XynB  (193) LKNVFGKYFDIGCAATPSEVSLQV-AKDIVKTHYNNLTIGNELKPDYVLDKAACQASGNNV--NPQVKLDSARSLLKYCAENNIEVRGHVLVWHSQTPSW
R.f-XynA  (630) LKGAPSKYFKICTSVSPHELNS---GADFLKKHYNSITPENELKPESIDDQGACQQKGNNV--NTQISLSRRAQTLKFCEQNGIALRGHTFVWYSQTPQW
C.t-XynY  (195) LKDMYANYFRVGSVLNSGTVNNSS-IKALILREFNSITGENEMKPDATLVQSGS----TNT--NIRVSLNRAASILNFCAQNNIAVRGHTLVWHSQTPEW
B.f-XynB  (  3) LKTAYEPYFKICAAISRWNLHTPAH-TKFLAEQYNSITPENELKPMYYWDREANKKDPEKYNLSPALTDENAIPYLEFAKDNKIAMRGHTLVWHNQTPKW
C.r-XynA  (356) ICEKYKNYFKICVAVPYRALTNPVD-VEVIKRHFNSITFENEMKPE-SLOPYEGG-----------ESFSIADEYVDFCKKDNISIRGHTLVWHQOTPSW
B.h-XynA  ( 57) LSERYQEQDICAAVEEYQLEGRQ---AQILKKHYNSLVAENAMKPE-SLQPREGE-----------WNWEGADKIVEFARKHNMELRFHTLVWHSGVPEW
C.s-XynB  ( 47) LAEAFRDYFPICAAIEPGYTTGQI--AEIYKKHVNMIVAENAMKEA-SLOPTEGN-----------FQWADADIKVDVFAKENGVELRFHTLVWHNQTPTG
T.s-XynA  (358) LYSVFKDYFPICVAVDPSRLNDADPHAQLTAKHFNMLVAENAMKPE-SLQPTEGN-----------FTDNADKIVDYIAHNMKMRGHTLIWHNCVPDW
C.t-XynX  (210) LSSVFKDYFPICVAVDPSRLNDTDPHAQLTAKHFNMLVAENAMKPE-SLQPTEGN-----------FTDNADRIVDYAIAHNMKMRGHTLIWHNCVPDW
G.s-XynA  (  8) LREVGANDFRICAAVNEVTLEAQQS---ILIRHVNSITEMKKFB-HLOPEEGR----------FTDDIAIKSSSTSPFS-SHGVRGHTLVWHNQTPSW
T.m-XynA  (370) LKEVLKDYFKVGVALPSKVFLNPKD-IEITTKHFNSITAENEMKPDSLLAGIENGK--------LKPRDETADKYIOFVEENGYVIRGHTLVWHNQTPEW
T.n-XynA  (366) LKDYVLKDYFRVGVALPSKVFINQKD-IALISKHSNSSTAENEMKPDSLLAGIENGK--------LKPRDETADKYIEFAQQNGYVVRGHTLVWHNQTPEW
                *   **                  *   *                                           *    * **   *  *  ***
```

```
C.v-XynB  (290) FFKENFS--------DTGATVSKDVMNQRLENYIKNLFAAINAEFPT-LKIYAWDVVNECYLDGGNLRTAGF---PETACKEASANNLVYGDDSYIDNAF
R.f-XynA  (725) FFRENFS--------QNGAYVSKDIMNQRLESMIKNTFAALKSQYPN-LDVYSYDVCNELFLNNG----------GGMRCADNSNAVKIYGDDSFVINAF
C.t-XynY  (288) FFKDNFQ--------DNGNWVSQSVMDQRLESYIKNMFAEIQRQYPS-LNLYAYDVVNEAVSDDANRTRYYGGAREPGYCNGRSPVVQIYGDNKFIEKAF
B.f-XynB  (102) FFCERYN--------ENFPMADRETILARLESYIHGVLDFVQTNYPG--IIYAWDVVNBIVD---------------EGAFRKSIWTETVGED-FFIKAF
C.r-XynA  (443) FFTNPET--------GEKLTNSEKDRELVELLDRIKTHIQTVVGTYPG--WDVVNEAIDENQ-----------PDCYFRSDVYNILCPE-YIEKAR
B.h-XynA  (143) FEIDEDGNRMVDETDPDKREANKQLLLENHIKIVVERYKDDVTS------WDVVNEVID--D-----------GGGLRFESEWYQITGTD-YIKVAF
C.s-XynB  (133) FSLDKEGKPMVEETDPQKREENRKLLQRLENYIRAVVLRYKDDIKS------WDVVNEVIEPND-----------PGCMRNSPWYQITGTE-YIEVAF
T.s-XynA  (446) FFQODP--------SDPSK-SASRDLLLQRLKTHITTVLDHFKTKYGSQNPIIGWDVVNEVLD--D-----------NGNLRNSKWLQILGPD-YIEKAF
C.t-XynX  (298) FFQODP--------SDPTK-PASRDLLLQRLKTHITTVLDHFKTKYGAQNPIIGWDVVNEVLD--D-----------NGSLRNSKWLQILGPD-YIEKAF
G.s-XynA  ( 92) VFQODS-----------QGHFVGRDVLLERMKSHITVVQRYKGKVYC------WDVINBAVADEG-----------SEWLRSSTWRQIIGDI-FIQQAF
T.m-XynA  (461) FFKDEN------GNLLS-----KEAMTELRIKEYIHTVVGHFKGKVYA------WDVVNEAVDPNQ-----------PDGLRRSTWYQIMCFD-YIELAF
T.n-XynA  (457) FFKDEN------GNLLS-----KEEMTERLREYIHITVVGHFKGKVYA------WDVVNEAVDPNQ-----------PDGLRRSTWYQIMCFD-YIELAF
                *                                                *   ** **     *                * *     *      **
```

```
C.v-XynB  (378) TYARKYAPAGVKLFYNDFNEYIQ-SKRNAIYTMAMRLKSKG--IIDGIGMQSFLDMGFPDTN---TYKAALTKFGSTGLEVQVTELDITTSDTS-------
R.f-XynA  (806) KYARQYAPAGCKTLYNDYNEYIF-ARTNDIYNMAMRLKQLG-YIDGIGMQSHLATNYPDAN---TYETALKKFLSTGLEVQITELDITCTNS-------
C.t-XynY  (379) TYARKYAEANCKLYYNDYNEYWD-HKRDCIASICANLYNKG--LLDGVGMQSHLLMDHPDISEYRTALEMYGSTG---LQIHITELDMHNADPS-----E-
B.f-XynB  (176) EFARKYAAPEVSLFYNDFTAQG-WKRDFILEKVLGPLIDKKLIDGMGMQSHLLMDHPDISEYRTALEMYGSTG---LQIHITELDMHNADPS-----E-
C.r-XynA  (519) IWAHEADEK-AKLFYNDYSTEDP-YKREFIYKIANLLEKGVHGVHCVGLOCHISLDWPDVSEIEETVKLFSRIPG--LEIHFTEIDISIAKN------MTD
B.h-XynA  (222) ETARKYGGEEAKLYINDYNTEVE-SKRDDLYNLVKDLLEQGVPIDGVCHQTHINIDIYNEPVERHIESIKRFAGLG---LDNQVTELDMSLYGWPPTGAYTS
C.s-XynB  (214) RATREAGGSDIKLYINDYNTDDE-VKRDILVELVKNLLEGVPIDGVGHOTHIDIYNEPVERHIESIKRFAGLG---LDNITELDMSIYSWNDRSDYG-
T.s-XynA  (522) EYAHEADEKS-MKLFINDYNIENNGVKTQAMYDLVKKLKSEQVIDGIGMQMHININSN-IDNIKASIERLASLG---VBIQVTELDMNMN----GNISN-
C.t-XynX  (374) EYAHEADES-MKLFINDYNIENNGVKTQAMYDLVKKLKSEQVPISGIGMQMHININSN-IDNIKASIEKLASLG---VBIQVTELDMNMN----GNVSN-
G.s-XynA  (162) LYAHEADPE-AILFYNDYNECFF-EKREKIYTMVKSLRDKGIPIHGIGMQAFWSLNRETLDEIRAAIERYASLG---VILHTTELDISMFEFDDHRKDLA
T.m-XynA  (531) KBAREADPD-AKLFYNDYNTFEE-RKRDIIYNLVLDDLKPKG-IIDGIGMQCHISLAT-DIKQEEAIKKFSTIPG--IBIHITELDMSVYRDS----SSN
T.n-XynA  (527) KBAREADPN-AKLFYNDYNTFEP-KKRDIIYNLVKSLKPKG-IIDGIGMQCHISLAT-DIRQEEAIKKFSTIPG--IBIHITELDISVYRDS----TSN
                  *   **          *            *      *         * ***               *                 * **
```

```
C.v-XynB  (464) -ATGLANQATKYAAIMQSILDAKKAGT-ANITNVFWGITDGTSWRATR----------LPLLFDGNYYPKPAFDSVVK
R.f-XynA  (893) -----AEQAD----LMEKIFKLAMQNS-AQIPAVIIWGTODTVSWRSSQ----------NPLLFSAGYQPKPAYDRVMA
C.t-XynY  (470) --FSLOQOADKYKAVFQAAVDINRTSSVKLDINKTNLLDENSWLGSQN----------IPLIFNANNQPKPAYNAVAS
B.f-XynB  (266) --ESMHALATRYQEFFQTYLDAKKSGK-ANITSVTEWNLLDENSWLSGFR-----R--ETSYPIVRKGKCEAKEAYYAVLK
C.r-XynA  (610) DDAYNRYLIVQQAQKLKAIEDVLKKYR-NVVISVTFWGLKDDYSWLRG-----------DMPLLSDKDYQPKPAFWSLID
B.h-XynA  (318) YDDIPAELIQAQADRYDQLFELYEELA-ADIISSVTFWGIADNHTWLDGRARE-YNNGVGIDAAFVEDHNYRVKPAYWIID
C.s-XynB  (309) -DSIPDYILTLQAKRYQELFDALKENK-DIVSAVVFWGISDKYSWINGFPVK-RTN-----APLLFDRNFMPKPAFWAIKD
T.s-XynA  (612) -----EAILR-KQARLYKQLFDLFKAEK-QYITAVVFWGVSDDVTWLS------KPN-----APLLFDSKLQKPAGWAVVD
C.t-XynX  (464) -----EAILR-KQARLYKQLFDLFKAEK-QYITAVVFWGVSDDVTWLS------KPN-----APLLFDSKLQAKPAYWAIAD
G.s-XynA  (257) AP--TNEMVERQAERYEQIFSLFKEYR-DVIQNVTFWGIADDHTWIDHFPVQGRKN-----WPLLFDEQHNPKPAFWRVVN
T.m-XynA  (621) YPEAPRTALIEQAHKMMQLFEIFKKYS-NVITNVTFWGLKDDYSWRATR----R-N----DWPLIFDKDHQARIAYWAIVA
T.n-XynA  (617) YSEAPRTALIEQAHKMAQLFKIFKKYS-NVITNVTFWGLKDDYSWRATR----R-N----DWPLIFDKDYQARIAYWAIVA
                          * *    *      *     *  * *                        *          *  *  *
```

C

```
XynB  (547) GDVNNDGKINALDIALVKKGLLSEFTDPAAKV-AADVNKDGNTNAIDLALLKKYLL
EngB  (386) GDVNKDGKVNAIDYAVLKSILLGTNTNVDLS--VSDVNKDGKVNALDIAVLKKMLL
EngE  (978) GDVGNDVVNSLDFEILLKKYVLNNDTMINK-A-SADLNKDGKINIIDLAFLKKAI-
EngH  (495) GDVNNNGIVNSMDLAMLKKYILGYEVEMNK-E-ASDLNKDGKINAIDFALLKKLIL
EngK  (834) GDVNGDTKINAIDIAMLKKYILDNSTVIKT-A-NADMNNDGKINAIDLALLKKKLL
EngL  (469) GDVNKDGKVNAIDIAILKKNLLNGITGDLT---VSDMNKDGKVNAIDLALLKKKLL
EngM  (822) GDINNDAKINAIDLAILKKYILNNDIPIVA-I-NGDINDGIINAIDLILLLKKQLL
EngY  (678) GDLNNDGKINAIDMALMKKGILGGFTDSATQL-AGDVNKDGKTNAIDLALIKKYIL
ExgS  (673) GDVNNDGKINAIDIAILKKYILDSTTKINT-A-NSDMNGDGKVNAIDLALLKKALI
ManA  ( 30) GDVNVDGKVNSIDMAIVKQYILGTKIFSDAELKAADVDANGKVNAIDLALYKQYIL
PelA  (859) GDLSKDGKVNAIDLAFLKKILISGDTSGADMT-VADMNKDGKINAIDLALLKKQLL
XynA  (249) GDVNNGTVNALDIAILKKFLLGMIPSLPATG---DVDQNGRMNAIDFAMIKQHLL
             **           *            *        *   *  *   *   *
```
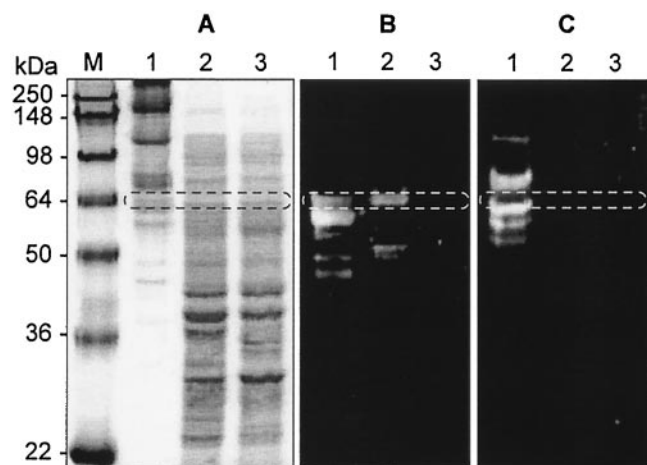
FIG. 5. Expression of XynB in *C. cellulovorans* and *E. coli*. Gels were stained with Coomassie blue (A) or stained for xylanase (B) or carboxymethyl cellulase (C) activity. Lanes: M, protein mass standards; 1, cellulosomal proteins of *C. cellulovorans*; 2, supernatant of recombinant *E. coli*; 3, supernatant of knockout XynB recombinant *E. coli*. The putative XynB (65,976 Da) is circled.

FIG. 6. TLC of hydrolysis products from xylooligosaccharides. Each xylooligosaccharide (xylobiose to xylopentaose) (2 mg) was incubated with either knockout recombinant XynB (1 μg) (lanes 1, 3, 5, and 7) or recombinant XynB (1 μg) (lanes 2, 4, 6, and 8) for 16 h, and the hydrolysates were analyzed by TLC. S, authentic oligosaccharides; X1, xylose; X2, xylobiose; X3, xylotriose; X4, xylotetraose; X5, xylopentaose.

(Fig. 5B, lane 2), which appeared to be the same size as the mature XynB deduced from the nucleotide sequence (65,976 Da) (Fig. 1). A band with an apparent molecular weight of 65,000 was detected in the cellulosomal proteins purified from *C. cellulovorans* (Fig. 5A, lane 1). The size of the protein was in good agreement with that of the full-length XynB produced by recombinant *E. coli* and the size calculated from the deduced amino acid sequence.

The mass spectrometry technique was also used for the identification of proteins separated by SDS-PAGE. The mass spectra of XynB showed the amino acid sequences of tryptic peptides LPLLFDGNYYPKPAFDSVVK, LVPTSDYYTPGYAL GDVNNDGK, IYAWDVVNECYLDGGNLR, YFDIGCAA TPSEVSLQVAK, ENFSDTGATVSK, and EASAWNLVY GDDSYIDNAFTYAR, which perfectly matched the entire sequence of XynB. The N-terminal amino acid sequence of this protein (65.3 kb) was previously identified as Glu-Asp-Ala-Leu-Leu-Ile-Ser (35), which was found in the deduced amino acid sequence of XynB at amino acid positions 31 to 36 (Fig. 1A). However, one additional smaller band was found at a molecular size of 52,000 (Fig. 5B, lane 2), and it is likely that the xylanase obtained here arose from the parental protein by partial proteolysis. The XynB protein showed specific activity

toward xylan and no activity toward carboxymethyl cellulose (Fig. 5C). The profiles based on SDS-PAGE, mass spectrometry, and zymogram analyses suggest that XynB is one of the major components of the cellulosome.

The optimum temperature for activity was found to be 40°C at pH 6.0 (data not shown). The optimum pH for activity was found to be pH 5 when the enzyme activity was assayed at 37°C in sodium acetic buffer solutions at various pHs (data not shown). The action of the enzyme on xylooligosaccharides was qualitatively analyzed. As shown in Fig. 6, XynB hydrolyzed xylotetraose and xylopentaose to yield mainly xylobiose and xylotriose, along with very small amounts of xylose. When xylan was treated with the enzyme, xylobiose and xylotriose were produced as end products accompanied by small amounts of xylose (data not shown). In contrast, this enzyme was less active toward xylotriose and not active at all toward xylobiose (Fig. 6).

The general characteristics of XynB, including temperature and pH for maximum activity, are very similar to those of xylanases (3). XynB hydrolyzed xylan substrates, releasing xylobiose and xylotriose as the major products. In addition, TLC analysis with xylooligosaccharides showed that all the xylooligosaccharides (xylotriose, xylotetraose, and xylopentaose)

FIG. 4. Alignment of domains of Xyn proteins from different species. (A) Alignment of family 4–9 carbohydrate-binding domains of *C. cellulovorans* (C.v) XynB (C.v-XynB) (GenBank accession no. AY604045), *Clostridium thermocellum* (C.t) XynC (BAA21516) (24), *Thermoanaerobacterium saccharolyticum* (T.s) XynA (P36917) (30), *Thermotoga maritima* (T.m) XynA (Q60037) (55), *Clostridium stercorarium* (C.s) XynB (P40942) (16), *Ruminococcus albus* (R.a) XynA (BAB39493) (36), *C. thermocellum* (C.t) XynY (P51584) (15), and *Ruminococcus flavefaciens* (R.f) XynD (Q53317) (14). (B) Alignment of family 10 catalytic domains of *C. cellulovorans* (C.v) XynB (C.v-XynB) (GenBank accession no. AY604045), *R. flavefaciens* (R.f) XynA (P29126) (58), *C. thermocellum* (C.t) XynY (P51584) (15), *Butyrivibrio fibrisolvens* (B.f) XynB (P26223) (31), *Caldicellulosiruptor* sp. strain Rt8B.4 (C.r) XynA (P40944) (13), *Bacillus halodurans* (B.h) XynA (P07528) (46), *Clostridium stercorarium* (C.s) XynB (P40942) (16), *Thermoanaerobacterium saccharolyticum* (T.s) XynA (P36917) (30), *C. thermocellum* (C.t) XynX (P38535), *Geobacillus stearothermophilus* (G.s) XynA (P45703) (2), *Thermotoga maritima* (T.m) XynA (Q60037) (55), and *Thermotoga neapolitana* (T.n) XynA (Q60042) (59). The two catalytic Glu residues are marked by vertical arrows. (C) Alignment of dockerin domains of XynB, EngB, EngE, EngH, EngK, EngL, EngM, EngY, ExgS, ManA, PelA, and XynA of *C. cellulovorans*. Gaps introduced to maximize alignment are indicated by dashes. Amino acids that are conserved in more than 50% of the sequences are indicated by white letters on a black background. Residues that are identical in all enzymes are indicated by an asterisk. Numbers at the start of the respective lines refer to amino acid residues; all sequences are numbered from Met-1 of the peptide.

were hydrolyzed except for xylobiose, suggesting that XynB hydrolyzes xylan in a manner typical of many endo-ββ-1,4-xylanases (3, 40). XynB as a xylanase in the cellulosome should contribute to the degradation of the xylan present in plant cell walls, allowing the cellulosome access to cellulose chains that are buried in xylan and are not readily accessible unless xylan is hydrolyzed and removed.

REFERENCES

1. **Attwood, G. T., H. P. Blaschek, and B. A. White.** 1994. Transcriptional analysis of the *Clostridium cellulovorans* endoglucanase gene, *engB*. FEMS Microbiol. Lett. **124:**277–284.
2. **Baba, T., R. Shinke, and T. Nanmori.** 1994. Identification and characterization of clustered genes for thermostable xylan-degrading enzymes, β-xylosidase and xylanase, of *Bacillus stearothermophilus* 21. Appl. Environ. Microbiol. **60:**2252–2258.
3. **Beg, Q. K., M. Kapoor, L. Mahajan, and G. S. Hoondal.** 2001. Microbial xylanases and their industrial applications: a review. Appl. Microbiol. Biotechnol. **56:**326–338.
4. **Bequin, P.** 1983. Detection of cellulase activity in polyacrylamide gels using Congo red-stained agar replicas. Anal. Biochem. **131:**333–336.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
6. **Breslauer, K. J., R. Frank, H. Blocker, and L. A. Marky.** 1986. Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. USA **83:** 3746–3750.
7. **Chang, B. Y., Y. T. Shyu, and R. H. Doi.** 1992. The interaction between *Bacillus subtilis* sigma-A (σ^A) factor and RNA polymerase with promoters. Biochimie **74:**601–612.
8. **Derewenda, U., L. Swenson, R. Green, Y. Wei, R. Morosoli, F. Shareck, D. Kluepfel, and Z. S. Derewenda.** 1994. Crystal structure, at 2.6-A resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of β-1,4-D-glycanases. J. Biol. Chem. **269:**20811–20814.
9. **Devillard, E., C. Bera-Maillet, H. J. Flint, K. P. Scott, C. J. Newbold, R. J. Wallace, J. P. Jouany, and E. Forano.** 2003. Characterization of XYN10B, a modular xylanase from the ruminal protozoan *Polyplastron multivesiculatum*, with a family 22 carbohydrate-binding module that binds to cellulose. Biochem. J. **373:**495–503.
10. **Doi, R. H.** 1991. Regulation of gene expression, p. 15–39. *In* U. N. Streips and R. E. Yasbin (ed.), Modern microbial genetics. Wiley-Liss, New York, N.Y.
11. **Doi, R. H., A. Kosugi, K. Murashima, Y. Tamaru, and S. O. Han.** 2003. Cellulosomes from mesophilic bacteria. J. Bacteriol. **185:**5907–5914.
12. **Doi, R. H., and Y. Tamaru.** 2001. The *Clostridium cellulovorans* cellulosome: an enzyme complex with plant cell wall degrading activity. Chem. Rec. **1:** 24–32.
13. **Dwivedi, P. P., M. D. Gibbs, D. J. Saul, and P. L. Bergquist.** 1996. Cloning, sequencing and overexpression in *Escherichia coli* of a xylanase gene, *xynA* from the thermophilic bacterium Rt8B.4 genus *Caldicellulosiruptor*. Appl. Microbiol. Biotechnol. **45:**86–93.
14. **Flint, H. J., J. Martin, C. A. McPherson, A. S. Daniel, and J. X. Zhang.** 1993. A bifunctional enzyme, with separate xylanase and β(1,3–1,4)-glucanase domains, encoded by the *xynD* gene of *Ruminococcus flavefaciens*. J. Bacteriol. **175:**2943–2951.
15. **Fontes, C. M. G. A., G. P. Hazlewood, E. Morag, J. Hall, B. H. Hirst, and H. J. Gilbert.** 1995. Evidence for a general role for non-catalytic thermostabilizing domains in xylanases from thermophilic bacteria. Biochem. J. **307:** 151–158.
16. **Fukumura, M., K. Sakka, K. Shimada, and K. Ohmiya.** 1995. Nucleotide sequence of the *Clostridium stercorarium xynB* gene encoding an extremely thermostable xylanase, and characterization of the translated product. Biosci. Biotechnol. Biochem. **59:**40–46.
17. **Goldstein, M. A., M. Takagi, S. Hashida, O. Shoseyov, R. H. Doi, and I. H. Segel.** 1993. Characterization of the cellulose-binding domain of the *Clostridium cellulovorans* cellulose-binding protein A. J. Bacteriol. **175:**5762–5768.
18. **Goryshin, I. Y., and W. S. Reznikoff.** 1998. Tn5 in vitro transposition. J. Biol. Chem. **273:**7367–7374.
19. **Han, S. O., H. Y. Cho, H. Yukawa, M. Inui, and R. H. Doi.** 2004. Regulation of expression of cellulosomes and noncellulosomal (hemi)cellulolytic enzymes in *Clostridium cellulovorans* during growth on different carbon sources. J. Bacteriol. **186:**4218–4227.
20. **Han, S. O., H. Yukawa, M. Inui, and R. H. Doi.** 2003. Regulation of expression of cellulosomal cellulase and hemicellulase genes in *Clostridium cellulovorans*. J. Bacteriol. **185:**6067–6075.
21. **Han, S. O., H. Yukawa, M. Inui, and R. H. Doi.** 2003. Transcription of *Clostridium cellulovorans* cellulosomal cellulase and hemicellulase genes. J. Bacteriol. **185:**2520–2527.
22. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15:**2343–2361.
23. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
24. **Hayashi, H., K.-I. Takagi, M. Fukumura, T. Kimura, S. Karita, K. Sakka, and K. Ohmiya.** 1997. Sequence of *xynC* and properties of XynC, a major component of the *Clostridium thermocellum* cellulosome. J. Bacteriol. **179:** 4246–4253.
25. **Henrissat, B., and A. Bairoch.** 1996. Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. **316:**695–696.
26. **Kosugi, A., K. Murashima, and R. H. Doi.** 2002. Characterization of two noncellulosomal subunits, ArfA and BgaA, from *Clostridium cellulovorans* that cooperate with the cellulosome in plant cell wall degradation. J. Bacteriol. **184:**6859–6865.
27. **Kosugi, A., K. Murashima, and R. H. Doi.** 2001. Characterization of xylanolytic enzymes in *Clostridium cellulovorans*: expression of xylanase activity dependent on growth substrates. J. Bacteriol. **183:**7037–7043.
28. **Kosugi, A., K. Murashima, and R. H. Doi.** 2002. Xylanase and acetyl xylan esterase activities of XynA, a key subunit of the *Clostridium cellulovorans* cellulosome for xylan degradation. Appl. Environ. Microbiol. **68:**6399–6402.
29. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
30. **Lee, Y. E., S. E. Lowe, and J. G. Zeikus.** 1993. Gene cloning, sequencing, and biochemical characterization of endoxylanase from *Thermoanaerobacterium saccharolyticum* B6A-RI. Appl. Environ. Microbiol. **59:**3134–3137.
31. **Lin, L. L., and J. A. Thomson.** 1991. Cloning, sequencing and expression of a gene encoding a 73 kDa xylanase enzyme from the rumen anaerobe *Butyrivibrio fibrisolvens* H17c. Mol. Gen. Genet. **228:**55–61.
32. **Liu, C. C., and R. H. Doi.** 1998. Properties of *exgS*, a gene for a major subunit of the *Clostridium cellulovorans* cellulosome. Gene **211:**39–47.
33. **Mach, R. L., and S. Zeilinger.** 2003. Regulation of gene expression in industrial fungi: *Trichoderma*. Appl. Microbiol. Biotechnol. **60:**515–522.
34. **Morag, E., E. A. Bayer, and R. Lamed.** 1990. Relationship of cellulosomal and noncellulosomal xylanases of *Clostridium thermocellum* to cellulose-degrading enzymes. J. Bacteriol. **172:**6098–6105.
35. **Murashima, K., A. Kosugi, and R. H. Doi.** 2002. Determination of subunit composition of *Clostridium cellulovorans* cellulosomes that degrade plant cell walls. Appl. Environ. Microbiol. **68:**1610–1615.
36. **Nakamura, M., T. Nagamine, A. Takenaka, R. I. Aminov, K. Ogata, K. Tajima, H. Matsui, Y. Benno, and H. Itabashi.** 2002. Molecular cloning, nucleotide sequence and characteristics of a xylanase gene (*xynA*) from *Ruminococcus albus* 7. Anim. Sci. J. **73:**347–352.
37. **Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne.** 1997. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int. J. Neural Syst. **8:**581–599.
38. **Park, J. S., Y. Matano, and R. H. Doi.** 2001. Cohesin-dockerin interactions of cellulosomal subunits of *Clostridium cellulovorans*. J. Bacteriol. **183:**5431–5435.
39. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:** 319–353.
40. **Saha, B. C.** 2003. Hemicellulose bioconversion. J. Ind. Microbiol. Biotechnol. **30:**279–291.
41. **Sakka, K., Y. Kojima, T. Kondo, S. Karita, K. Ohmiya, and K. Shimada.** 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.
42. **Shoseyov, O., and R. H. Doi.** 1990. Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA **87:**2192–2195.
43. **Shoseyov, O., M. Takagi, M. A. Goldstein, and R. H. Doi.** 1992. Primary sequence analysis of *Clostridium cellulovorans* cellulose binding protein A. Proc. Natl. Acad. Sci. USA **89:**3483–3487.
44. **Sleat, R., R. A. Mah, and R. Robinson.** 1984. Isolation and characterization of an anaerobic, cellulolytic bacterium, *Clostridium cellulovorans* sp. nov. Appl. Environ. Microbiol. **48:**88–93.
45. **Sugimoto, N., S. Nakano, M. Katoh, A. Matsumura, H. Nakamuta, T. Ohmichi, M. Yoneyama, and M. Sasaki.** 1995. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry **34:**11211–11216.
46. **Takami, H., K. Nakasone, Y. Takaki, G. Maeno, R. Sasaki, N. Masui, F. Fuji, C. Hirama, Y. Nakamura, N. Ogasawara, S. Kuhara, and K. Horikoshi.** 2000. Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*. Nucleic Acids Res. **28:**4317–4331.

47. **Tamaru, Y., and R. H. Doi.** 2000. The *engL* gene cluster of *Clostridium cellulovorans* contains a gene for cellulosomal *manA*. J. Bacteriol. **182:**244–247.

48. **Tamaru, Y., and R. H. Doi.** 1999. Three surface layer homology domains at the N terminus of the *Clostridium cellulovorans* major cellulosomal subunit EngE. J. Bacteriol. **181:**3270–3276.

49. **Tamaru, Y., S. Karita, A. Ibrahim, H. Chan, and R. H. Doi.** 2000. A large gene cluster for the *Clostridium cellulovorans* cellulosome. J. Bacteriol. **182:**5906–5910.

50. **Thomson, J. A.** 1993. Molecular biology of xylan degradation. FEMS Microbiol. Lett. **104:**65–82.

51. **Tull, D., S. G. Withers, N. R. Gilkes, D. G. Kilburn, R. A. Warren, and R. Aebersold.** 1991. Glutamic acid 274 is the nucleophile in the active site of a "retaining" exoglucanase from *Cellulomonas fimi*. J. Biol. Chem. **266:**15621–15625.

52. **Uffen, R. L.** 1997. Xylan degradation: a glimpse at microbial diversity. J. Ind. Microbiol. Biotechnol. **19:**1–6.

53. **Wang, Q., D. Tull, A. Meinke, N. R. Gilkes, R. A. Warren, R. Aebersold, and S. G. Withers.** 1993. Glu280 is the nucleophile in the active site of *Clostridium thermocellum* CelC, a family A endo-β-1,4-glucanase. J. Biol. Chem. **268:**14096–14102.

54. **White, A., S. G. Withers, N. R. Gilkes, and D. R. Rose.** 1994. Crystal structure of the catalytic domain of the β-1,4-glycanase cex from *Cellulomonas fimi*. Biochemistry **33:**12546–12552.

55. **Winterhalter, C., P. Heinrich, A. Candussio, G. Wich, and W. Liebl.** 1995. Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium *Thermotoga maritima*. Mol. Microbiol. **15:**431–444.

56. **Young, M., N. P. Minton, and W. L. Staudenbauer.** 1989. Recent advances in the genetics of the clostridia. FEMS Microbiol. Rev. **5:**301–325.

57. **Zeilinger, S., R. L. Mach, M. Schindler, P. Herzog, and C. P. Kubicek.** 1996. Different inducibility of expression of the two xylanase genes *xyn1* and *xyn2* in *Trichoderma reesei*. J. Biol. Chem. **271:**25624–25629.

58. **Zhang, J. X., and H. J. Flint.** 1992. A bifunctional xylanase encoded by the *xynA* gene of the rumen cellulolytic bacterium *Ruminococcus flavefaciens* 17 comprises two dissimilar domains linked by an asparagine/glutamine-rich sequence. Mol. Microbiol. **6:**1013–1023.

59. **Zverlov, V. V., K. Piotukh, O. Dakhova, G. Velikodvorskaya, and R. Borriss.** 1996. The multidomain xylanase A of the hyperthermophilic bacterium *Thermotoga neapolitana* is extremely thermoresistant. Appl. Microbiol. Biotechnol. **45:**245–247.