

Sequence of xynC and properties of XynC, a major component of the *Clostridium thermocellum* cellulosome.

H Hayashi, K I Takagi, M Fukumura, T Kimura, S Karita, K Sakka and K Ohmiya
J. Bacteriol. 1997, 179(13):4246.

Updated information and services can be found at:
<http://jb.asm.org/content/179/13/4246>

CONTENT ALERTS

These include:

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), [more»](#)

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

Sequence of *xynC* and Properties of XynC, a Major Component of the *Clostridium thermocellum* Cellulosome

HIDENORI HAYASHI,¹ KEN-ICHIRO TAKAGI,¹ MASAYUKI FUKUMURA,¹ TETSUYA KIMURA,¹ SHUICHI KARITA,² KAZUO SAKKA,^{1*} AND KUNIO OHMIYA¹

Faculty of Bioresources¹ and Center for Molecular Biology and Genetics,² Mie University, Tsu 514, Japan

Received 24 December 1996/Accepted 1 May 1997

The nucleotide sequence of the *Clostridium thermocellum* F1 *xynC* gene, which encodes the xylanase XynC, consists of 1,857 bp and encodes a protein of 619 amino acids with a molecular weight of 69,517. XynC contains a typical N-terminal signal peptide of 32 amino acid residues, followed by a 165-amino-acid sequence which is homologous to the thermostabilizing domain. Downstream of this domain was 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 N-terminal amino acid sequence of XynC-II, the enzyme purified from a recombinant *Escherichia coli* strain, was in agreement with that deduced from the nucleotide sequence although XynC-II suffered from proteolytic truncation by a host protease(s) at the C-terminal region. Immunological and N-terminal amino acid sequence analyses disclosed that the full-length XynC is one of the major components of the *C. thermocellum* cellulosome. XynC-II was highly active toward xylan and slightly active toward *p*-nitrophenyl- β -D-xylopyranoside, *p*-nitrophenyl- β -D-cellobioside, *p*-nitrophenyl- β -D-glucopyranoside, and carboxymethyl cellulose. The K_m and V_{max} values for xylan were 3.9 mg/ml and 611 μ mol/min/mg of protein, respectively. This enzyme was optimally active at 80°C and was stable up to 70°C at neutral pHs and over the pH range of 4 to 11 at 25°C.

β -1,4-Xylan, which is a major component of the hemicellulose fraction of plant cell walls, has a backbone of β -1,4-linked xylopyranosyl residues. The β -1,4-xylopyranosyl chain is replaced by mainly acetyl, arabinofuranosyl, and glucuronosyl residues (7).

Two kinds of enzymes are generally involved in microbial hydrolysis of the main chain, i.e., endo-1,4- β -xylanase (EC 3.2.1.8) and β -xylosidase (EC 3.2.1.37) (7). Many xylanase and xylosidase genes along with their translated products from fungi and bacteria have been isolated and characterized (48). On the basis of amino acid sequence homology, xylanases can be divided into two substantial groups: family 10 and family 11 catalytic domains of glycosylhydrolase (24). These two domains are quite different from each other in their structures. The family 10 enzymes form closely related eight-stranded α/β barrel structures (11, 12, 23, 53). On the other hand, the family 11 enzymes comprise a single domain of two or three β -sheets and one helix (26, 49).

Strong consumers of cellulosic materials such as *Trichoderma reesei* (49) and *Cellulomonas fimi* (10, 54) produce xylanase(s) in addition to a series of cellulases for efficient degradation of plant cell walls, since xylan exists in the plant cell walls as a major component and associates with other components (7).

Clostridium thermocellum is a spore-forming anaerobic thermophilic bacterium which secretes a highly active cellulolytic complex, termed the cellulosome (4, 6, 13). The cellulosome is a complex aggregate of at least 14 subunits, detectable by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), whose molecular weights range from 20,000 to 250,000, and it has a total molecular mass of more than 2 MDa (29). However, the cellulosome is too stable to be disrupted by urea, guanidine hydrochloride, and various detergents. Al-

though treatment of the cellulosome with SDS in the presence of EDTA and thiols is the most effective way to dissociate it (37, 57), this treatment may change the native structure of the components somewhat. Recently, more moderate treatment, i.e., incubation of the cellulosome in 50 mM Tris(hydroxymethyl)aminomethane buffer containing 0.1 M NaCl and 5 mM EDTA was shown to disintegrate the cellulosome into polypeptides, but this was followed by the formation of truncated polypeptides (9). So far, only CelA (37) and CelS (57), which are known to be major components of the *C. thermocellum* cellulosome, have been successfully purified from the cellulosome in the presence of SDS; however, there are no reports concerning the purification of xylanase from the cellulosome. Difficulty in dissociating the cellulosome prevents us from isolating its subunits and studying the function of each subunit.

Under these circumstances, many genes encoding catalytic subunits have been cloned, sequenced, and expressed in *Escherichia coli* in addition to several genes encoding noncatalytic structural proteins such as CipA (18), e.g., endoglucanase genes *celA* (5), *celB* (20), *celD* (25), *celE* (22), *celF* (38), *celG* (31), *celH* (58), and *celJ* (1); an exoglucanase gene, *celS* (51); and a lichenase gene, *licB* (46). In agreement with the presence of xylanase activity in the cellulosome (27, 35), the xylanase genes *xynY* (15) and *xynZ* (21) were cloned and expressed in *E. coli*. Immunological analysis suggested that XynY and XynZ existed in the cellulosome, although these enzymes were not equivalent to the major components specified as S1 to S14 of the *C. thermocellum* YS cellulosome (29).

We constructed a gene library of *C. thermocellum* F1, which was isolated from a compost heap, and cloned eight endoglucanase genes, two xylanase genes, and a β -glucosidase gene (42). On comparing the restriction maps of the plasmids constructed from *C. thermocellum* F1 genomic DNA and those of *C. thermocellum* NCIB 10682 genes, we found that four of eight endoglucanase genes cloned from strain F1 are homologous with the genes from the type strain, i.e., *celC*, *celF*, *celH*,

* Corresponding author. Phone: (81) 592-31-9621. Fax: (81) 592-31-9684. E-mail: sakka@bio.mie-u.ac.jp.

and an uncharacterized endoglucanase gene (42). Comparison of the nucleotide sequences of the homologous *celC* genes showed only six substitutions in the coding region, resulting in three amino acid changes (44). Recently, we have reported the nucleotide sequence of *celJ*, which encodes the largest catalytic component of the cellulosome (1) and appears to correspond to the component S2 identified in the cellulosome of *C. thermocellum* YS (29) and Jm160 in *C. thermocellum* NCIB 10682 (3). These observations suggested that the organization of the cellulosome of *C. thermocellum* F1 resembles that of *C. thermocellum* NCIB 10682.

In this paper, we describe the nucleotide sequence of the *xynC* gene encoding one of the major components of the cellulosome, which corresponds to component J4 of a subpopulation of the *C. thermocellum* NCIB 10682 cellulosome and probably to component S9 or S10 of the *C. thermocellum* YS cellulosome. We also deal with characterization of the enzyme purified from a recombinant *E. coli* strain.

MATERIALS AND METHODS

Bacterial strains and plasmids. *C. thermocellum* F1, used for isolation of the cellulosome fraction, was described previously (43). The *E. coli* strains used were XL1-Blue {*supE44 hsdR17 recA1 endA1 gyrA46 thi-1 relA1 lac* [*F'*proAB⁺ laqI^q lacZΔM15 Tn10 (Tet^r)]} and XL2-Blue MRF' {Δ(*mcra*)183 Δ(*mcrcB*-*hsdSMR*-*mrr*)173 *supE44 recA endA1 gyrA96 λ⁻ thi-1 relA1 lac* [*F'*proAB⁺ laqI^q lacZΔM15 Tn10 (Tet^r) Amy Cam^r]}. Plasmids pBluescript II KS(+) and KS(-) (Stratagene) were used for determination of the DNA sequence of *xynC* and production of the recombinant enzyme. Plasmid pKS103 is a derivative of pBR322 in which a 5.2-kbp DNA fragment containing the *xynC* gene of the *C. thermocellum* F1 was cloned (42).

Subcloning and DNA sequencing. A 3.3-kbp *EcoRI-Eco81I* fragment of pKS103 containing the coding sequence was inserted at the *EcoRI-SalI* sites of pBluescript II KS(+) and KS(-) to yield pKS103-1 and pKS103-2, respectively. A series of nested deletion mutants from pKS103-1 and pKS103-2 was constructed by using the exonuclease III-mung bean nuclease digestion protocol from Toyobo Co., Ltd. (Osaka, Japan). The dideoxy chain termination reaction was done with a single-stranded DNA template, a dye-labeled custom primer (T3 or T7 primer), and *Taq* DNA polymerase, using a Dye Primer Cycle Sequencing Kit (Applied Biosystems), and products were analyzed on a model 373A automated DNA sequencer system (Applied Biosystems). Nucleotide and amino acid sequences were analyzed with GENETYX-MAC computer software (version 7.3; Software Development Co., Ltd., Tokyo, Japan). Standard techniques described by Sambrook et al. (45) were used for other DNA manipulations.

Purification of the recombinant enzyme. All purification procedures were performed at 4°C. Cells of *E. coli* XL1-Blue(pKS103-1) were harvested from an overnight culture (5 liters) in L medium containing ampicillin (50 μg/ml; Nacal Tesque Co. Ltd., Kyoto, Japan) by centrifugation at 5,000 × *g* for 10 min and osmotically shocked according to the method of Neu and Heppel (39). After removal of the cells by centrifugation at 5,000 × *g* for 10 min, the periplasmic fraction was precipitated by adding solid ammonium sulfate to 60% saturation. The precipitate was dissolved in 20 ml of buffer A (20 mM Tris-HCl, pH 7.5) and dialyzed against 2 liters of the same buffer. The dialyzed sample was loaded onto a column of DEAE-Toyopearl 650M (2.0 by 10.5 cm; Tosoh Co., Tokyo, Japan) equilibrated with buffer A. The column was washed with buffer A and eluted at 4 ml/min with a linear gradient of NaCl ranging from 0 to 0.5 M in buffer A. Fractions with xylanase activity were combined, dialyzed against buffer A, and then put on a MonoQ HR5/5 column (0.5 by 5 cm; Pharmacia Biotech) equilibrated with buffer A. The column was washed with buffer A and eluted at 0.5 ml/min with a linear gradient of 27.5 ml of NaCl ranging from 0 to 0.5 M in the same buffer. Pooled eluates containing xylanase activity were loaded onto a HiLoad 16/60 Superdex 200-pg column (1.6 by 60 cm; Pharmacia Biotech) equilibrated with buffer A containing 0.2 M NaCl and eluted at 1 ml/min with the same buffer solution. The enzyme thus obtained was used for characterization of enzymatic properties.

Enzyme assays. Xylanase activity was measured in a 10-min incubation at 60°C in 50 mM sodium succinate buffer (pH 5.5) or Britton and Robinson's universal buffer (50 mM phosphoric acid–50 mM boric acid–50 mM acetic acid; the pH was adjusted to 2 to 12 with 1 N NaOH) in the presence of 0.75% oat-spelt xylan (Fluka AG, Buchs, Switzerland). Reducing sugars released from the substrate were measured with the 3,5-dinitrosalicylic acid reagent as described by Miller (34). One unit of xylanase activity was defined as the amount of enzyme releasing 1 μmol of xylose equivalent per min from xylan. β-Xylosidase, β-cellobiosidase, and β-galactosidase activities were assayed at 60°C with *p*-nitrophenyl-β-D-xylopyranoside (PNPX; Sigma), *p*-nitrophenyl-β-D-cellobioside (PNPC; Sigma), and *p*-nitrophenyl-β-D-glucopyranoside (PNPG; Sigma), respectively. One unit of enzyme activity toward PNP derivatives was defined as the amount of enzyme liberating 1 μmol of *p*-nitrophenol per min. Enzyme activity on carboxymethyl

cellulose (CMC) was assayed as described previously (42). Protein concentrations were determined by the method of Lowry et al. (33) with bovine serum albumin (Sigma) as a standard.

Isolation of the *C. thermocellum* cellulosome. A cellulosome fraction of *C. thermocellum* F1 was prepared from the culture supernatant by the affinity digestion procedure described by Morag et al. (36). In brief, cellulase and xylanase complex having affinity for insoluble cellulose in the culture supernatant was selectively bound to acid-swollen cellulose at 4°C and recovered after digestion of the cellulose at 50°C.

SDS-PAGE and zymogram analysis. SDS-PAGE was done by the method of Laemmli (29). Zymogram analysis was performed as described by Ali et al. (3), with an SDS–10% polyacrylamide gel containing 0.1% oat-spelt xylan.

Preparation of antiserum and immunoblotting. The purified enzyme (100 μg) was mixed with an equal volume of Freund's complete adjuvant and injected subcutaneously into a BALB/c mouse. The second injection was administered at an interval of 2 weeks, with the same amount of the protein with Freund's incomplete adjuvant. The serum was collected 2 weeks after the second injection. Protein samples were fractionated by SDS-PAGE (28) and transferred onto nitrocellulose membranes by using an electroblotting apparatus, Sartoblot II (Sartorius, Göttingen, Germany) (49). Immunoreactive proteins were detected on Western blots by enzyme immunoassay using peroxidase-conjugated goat anti-mouse immunoglobulins (Tago Inc., Burlingame, Calif.) and 3,3'-diaminobenzidine tetrahydrochloride.

Determination of N-terminal amino acid sequences. The xylanase purified from a recombinant *E. coli* strain and the cellulosomal proteins from *C. thermocellum* were fractionated by SDS-PAGE (28) and transferred onto Immobilon P transfer membranes (Millipore Corp., Bedford, Mass.) by electroblotting. The blotted proteins were cut out from the blots and were subjected to automated amino acid sequencing on an Applied Biosystems model 476A protein sequencer.

Analysis of hydrolysis products. Xylooligosaccharides (xylobiose to xylooctaose, each 5 mg) were incubated with 0.1 U of the purified enzyme in 1 ml of 50 mM sodium succinate buffer (pH 5.5) at 60°C. Thin-layer chromatography of the hydrolysis products was performed on a DC-Fertigplatten SIL G-25 plate (Macherey-Nagel, Dorne, Germany) developed with a solvent of 1-propanol-water (85:15, vol/vol), and xylooligosaccharides were visualized by spraying the plate with an aniline-diphenylamine reagent (17).

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been submitted to the DDBJ, EMBL, and GenBank nucleotide sequence databases under the accession no. D84188.

RESULTS

Nucleotide sequence of the *xynC* gene. Figure 1 shows the *xynC* structural gene along with its flanking regions. There is an open reading frame composed of 1,857 nucleotides encoding a protein of 619 amino acids with a predicted molecular weight of 69,517. The assigned ATG initiation codon at nucleotide position 599 is preceded by a putative Shine-Dalgarno sequence, GGAGG, a typical ribosome binding site in *C. thermocellum* (52). The reading frame is ended by the stop codon TGA at position 2,458. A possible promoter sequence, TTGACA for the –35 region and TATGAA for the –10 region, with a 19-bp spacing between them, was observed. These sequences show high homologies to the consensus promoter sequences for σ⁷⁰ factor found in *E. coli*, i.e., TTGACA and TATAAT with a 17-bp spacing (41). A possible transcription terminator that consists of a 35-bp palindromic sequence, corresponding to an mRNA hairpin loop with a Δ*G* of –27 kcal/mol (ca. –106 kJ/mol) (8), followed by 3 T's was found downstream of the TGA termination codon. This structure is similar to the rho factor-independent terminator of *E. coli* (41).

Molecular architecture of XynC. The deduced N-terminal sequence of 32 amino acids contains a sequence similar to the signal peptide sequences found in prokaryotic secretory proteins, which all share general characteristics, such as a short region rich in positively charged amino acid species, followed by a sequence of predominantly hydrophobic residues, a residue breaking the secondary structure (glycine or proline), and a cleavage site ending with alanine, glycine, or serine (53).

Comparison of the amino acid sequence of XynC with those registered in protein databases such as SWISS PROT and PIR clearly revealed that the mature XynC consists of three distinct functional domains, i.e., an N-terminal domain which is ho-

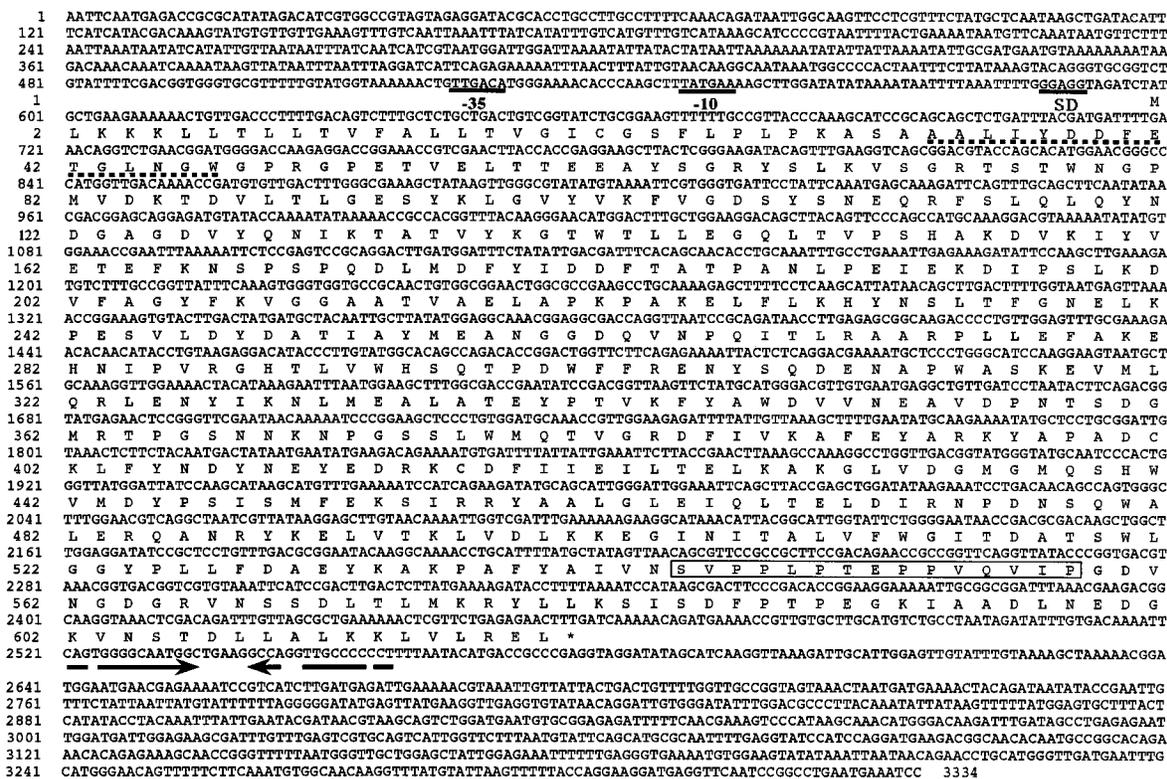


FIG. 1. Nucleotide and deduced amino acid sequences of *xynC*. The putative promoter and Shine-Dalgarno (SD) sequences are underlined. The amino acids underlined with a dotted line were identified for XynC and XynC-II by automated sequencing. A palindrome is indicated by arrows facing each other. The linker sequence rich in Pro between a family 10 catalytic domain and a dockerin domain is boxed.

mologous with the stretches found in several glycanases, a family 10 catalytic domain of glycosyl hydrolases, and a dockerin domain (listed in order from the N terminus). Figure 2 shows schematically the molecular architecture of XynC along with the related enzymes. The family 10 domain of XynC, extending from position 198 to 541, exhibited extensive sequence homology with the catalytic domains of the other xylanases in family 10 (Fig. 3), e.g., 44.2% sequence identity with XynB of *Butyrivibrio fibrisolvens* (32), 43.4% identity with XynA of *Ruminococcus flavefaciens* (59), 43.4% identity with XynY of *C. thermocellum* YS (15), 40.5% identity with XynA of *Thermotoga maritima* MSB8 (56), 38.2% identity with XynX of *C. thermocellum* ATCC 27405 (GenBank accession no. M67438), 38.1% identity with XynA of *Thermoanaerobacterium saccharolyticum* B6A-RI (30), 34.8% identity with XynB of *Clostridium stercoararium* F-9 (16), 34.5% identity with XynZ of *C. thermocellum* NCIB 10682 (21), and 29.0% identity with XynC of *C. fimi* (10). As shown in Fig. 4, the N-terminal domain of the mature form of XynC, about 160 amino acid residues downstream of the signal peptide, exhibited 30.2 and 32.1% sequence identities with residues 43 to 199 and residues 200 to 356, respectively, of XynA from *T. saccharolyticum* B6A-RI (31); 31.1% identity with residues 254 to 412 of XynC from *C. fimi* (10), 31.8% identity with residues 42 to 202 of XynX from *C. thermocellum* ATCC 27405 (GenBank accession no. M67438); 34.7% identity with residues 570 to 725 of XynY from *C. thermocellum* YS (16); 26.8 and 29.9% identities with residues 54 to 207 and residues 208 to 368, respectively, of XynA from *T. maritima* MSB8 (56); 24.0 and 30.0% identities with residues 45 to 197 and residues 198 to 354, respectively, of XynA from

thermophilic bacterium strain Rt8.B4 (GenBank accession no. L18965); and 35.2% identity with residues 263 to 420 of XynD from *R. flavefaciens* (14). These sequences have been recently referred to as the thermostabilizing domain by Fontes et al. (15) based on the findings that removal of this domain from *C. thermocellum* XynY (15) and *T. saccharolyticum* XynA (30) decreased their optimum temperatures and thermal stabilities. The third domain in XynC, which is separated from the catalytic domain by a short linker sequence rich in Pro, is a dockerin domain located in the C terminus of the peptide. Dockerins that consist of a pair of well-conserved 25-residue repeats are highly conserved in cellulases and xylanases from *C. thermocellum* and other cellulosome-forming clostridia (Fig. 5) and play a role in cellulosome assembly by docking the various catalytic subunits to a noncatalytic scaffolding protein, CipA (4, 6).

Purification of the xylanase encoded by *xynC* from a recombinant *E. coli* strain. The gene product of *xynC* was purified 163-fold from the periplasmic fraction of *E. coli* XL1-Blue(pKS103-1), with a recovery of 7% by ammonium sulfate precipitation and DEAE-Toyopearl 650M, MonoQ HR5/5, and HiLoad 16/60 Superdex 200-pg column chromatographies. The final preparation gave a single band in SDS-PAGE, and the molecular weight of the enzyme was estimated to be around 64,000 (Fig. 6A). The N-terminal amino acid sequence of this protein was identified as Ala-Ala-Leu-Ile-Tyr-Asp-Asp-Phe-Glu-Thr-Gly-Leu-Asn-Gly-Trp, which was found in the deduced amino acid sequence of XynC at amino acid positions 33 to 47 (Fig. 1), indicating that the N-terminal sequence of 32 amino acids mediates secretion of the protein to the periplas-

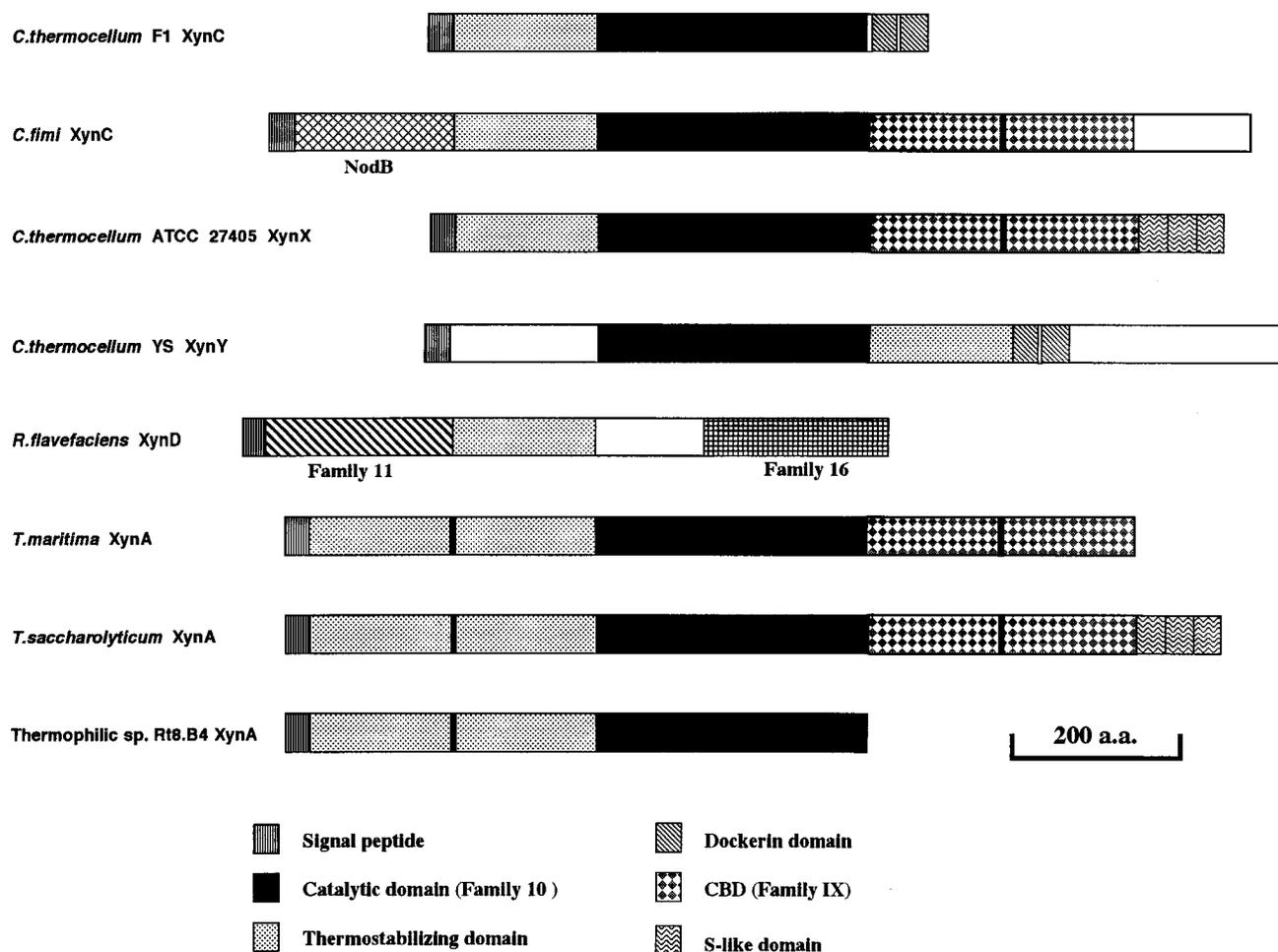


FIG. 2. Modular structure of enzymes containing thermostabilizing domains. Structures of *C. thermocellum* XynC, *C. fimi* XynC, *C. thermocellum* XynX, *C. thermocellum* XynY, *R. flavefaciens* XynD, *T. maritima* XynA, *T. saccharolyticum* XynA and thermophilic bacterial strain Rt8.B4 XynA are shown. Boxes representing domains with significant amino acid sequence similarity are filled in with the same pattern. NodB of *C. fimi* XynC is a domain homologous with NodB proteins of *Rhizobium* spp. *R. flavefaciens* XynD contains family 11 and 16 catalytic domains. CBD, cellulose binding domain; a.a., amino acids.

mic space as a signal peptide. However, the molecular weight estimated by SDS-PAGE appeared to be similar to but slightly lower than that of the mature XynC deduced from the nucleotide sequence (66,146), and it is likely that the xylanase obtained here arose from a parental protein by partial proteolysis. Therefore, we analyzed proteins in the eluates from the DEAE-Toyopearl 650M column by Western blotting and zymogram analysis and compared them with the purified enzyme. In an active fraction eluted from the DEAE-Toyopearl column, two protein species with different molecular weights, i.e., 67,000 and 64,000, were found to be immunoreactive with the antibody raised against the purified enzyme (Fig. 6C) and to exhibit xylanase activity, as shown by zymogram analysis (Fig. 6B). The large protein detected in the eluate was apparently larger than the purified enzyme, and the small protein corresponded to the purified enzyme. Therefore, the former, which appeared to be a full-length integral protein, and the latter are referred to as XynC and XynC-II, respectively, in this study. For determination of the N-terminal amino acid sequence of the large protein, XynC, the fraction from the DEAE-Toyopearl column was concentrated and subjected to automated N-terminal sequencing after SDS-PAGE and elec-

trophotting onto a polyvinylidene difluoride membrane. The sequence of XynC, identified as Ala-Ala-Leu-Ile-Tyr-Asp-Asp-Phe-Glu-Thr, was completely identical to that of XynC-II, suggesting that XynC-II arose from XynC due to partial proteolysis in the C terminus of the parental protein.

Identification of XynC in the cellulosomal proteins of *C. thermocellum*. By Western blotting using the antiserum directed against XynC-II, a single immunoreactive band with an apparent molecular weight of 67,000 was detected in the cellulosomal proteins purified from *C. thermocellum* F1 by affinity digestion (Fig. 6C). The size of the immunoreactive protein was in good agreement with that of the full-length XynC produced by recombinant *E. coli* and the size calculated from the deduced amino acid sequence. This protein showed xylanase activity upon zymogram analysis (Fig. 6B). The profiles based on SDS-PAGE, zymogram analysis, and Western blotting suggest that XynC is one of the major components of the cellulosome. Therefore, we determined the N-terminal amino acid sequence of the major protein of the cellulosome with a molecular weight of 67,000. The identified sequence was Ala-Ala-Leu-Ile-Tyr-Asp-Asp-Phe-Glu-Thr, which was consistent with the amino acid sequences of XynC and XynC-II and the de-

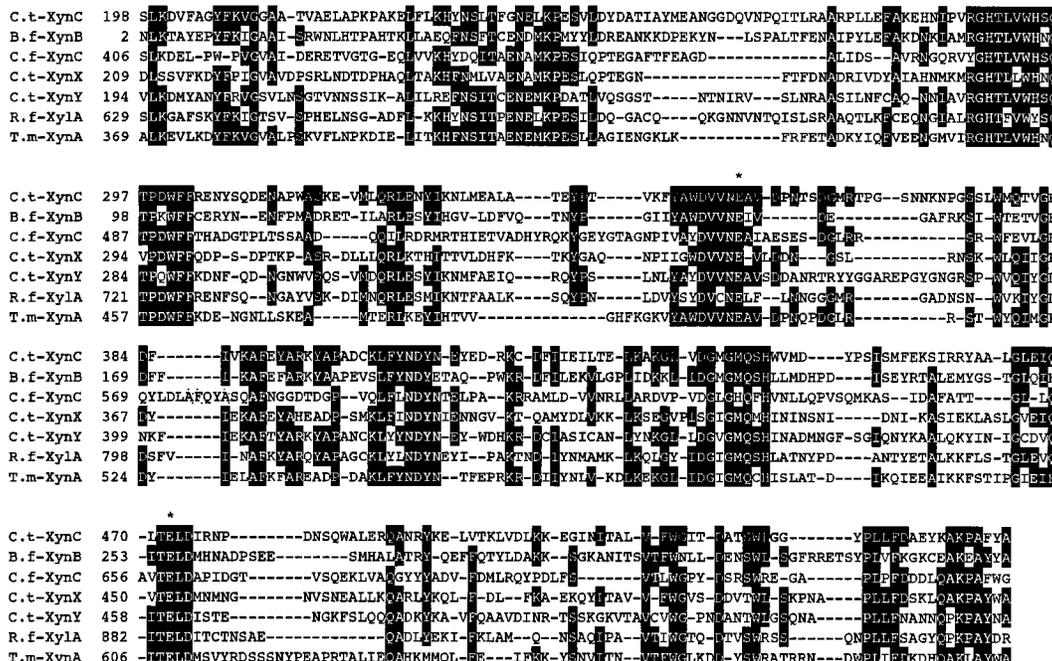


FIG. 3. Alignment of family 10 catalytic domains of *C. thermocellum* (C.t) XynC, *B. fibrisolvens* (B.f) XynB, *C. fimi* (C.f) XynC, *C. thermocellum* XynX, *C. thermocellum* XynY, *R. flavefaciens* (R.f) XynA, and *T. maritima* (T.m) XynA. Amino acids which are conserved in at least four of the seven sequences are shown in black boxes. -, gap left to improve alignment. Numbers at the start of the respective lines refer to amino acid residues; all sequences are numbered from Met-1 of the peptide. Glu residues involved in general acid-base catalysis are marked with an asterisk.

duced DNA sequence of *xynC*. These results indicate that the *xynC* gene is highly expressed in *C. thermocellum* F1 and its product is integrated into the cellulosome as a major component.

General characterization of XynC-II. The purified XynC-II had high specific activity toward oat-spelt xylan (557 U/mg) and low activity toward several substrates, i.e., 0.04 U/mg for PNPX, 0.30 U/mg for PNPC, 0.02 U/mg for PNPG, and 0.18 U/mg for CMC. The initial rates of reaction were measured at

60°C in various concentrations of xylan. From Lineweaver-Burk plots, the K_m and V_{max} values were estimated to be 3.9 mg/ml and 611 $\mu\text{mol}/\text{min}/\text{mg}$ of enzyme, respectively. The action of the enzyme on xylan and xylooligosaccharides was qualitatively analyzed. As shown in Fig. 7, XynC-II hydrolyzed xylan to yield mainly xylobiose and xylotriose, along with xylose as a minor product. When xylootetraose and larger xylooligosaccharides, i.e., xylopentaose to xylooctaose, were treated with the enzymes, xylobiose and xylotriose were produced as

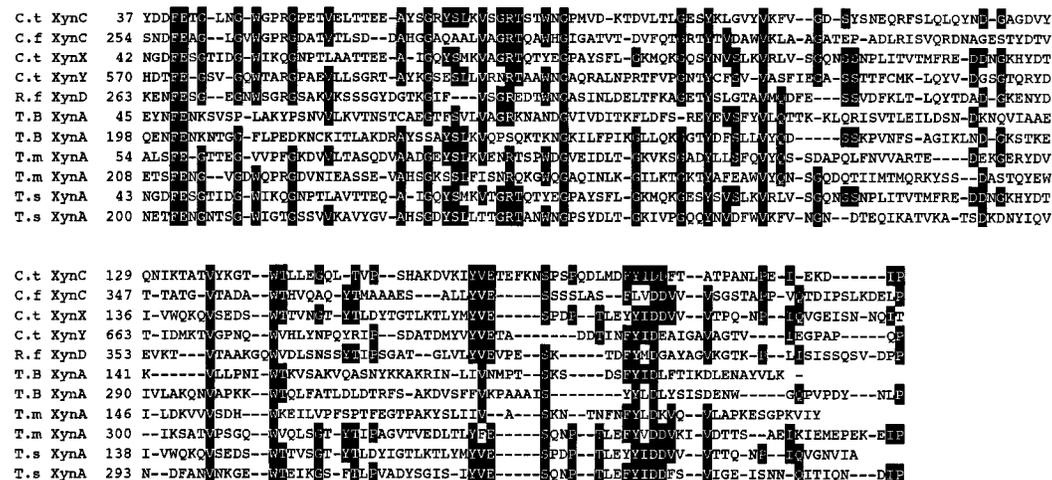


FIG. 4. Alignment of thermostabilizing domains of *C. thermocellum* (C.t) XynC, *C. fimi* (C.f) XynC, *C. thermocellum* XynX, *C. thermocellum* XynY, *R. flavefaciens* (R.f) XynD, *T. maritima* (T.m) XynA, thermophilic bacterial strain Rt8.B4 (T.B) XynA, and *T. saccharolyticum* (T.s) XynA. Amino acids which are conserved in at least six of the 11 sequences are shown in black boxes. -, gap left to improve alignment. Numbers at the start of the respective lines refer to amino acid residues; all sequences are numbered from Met-1 of the peptide.

XynC	557	WIPGDVNGDGRVNSDILMLKRYLLKSIDFPTPEGKIA
Ce1A	413	VVYGDVNGDGMVNSDITLTKRYLLKSVTNINREA
Ce1B	498	VPTGDVNGDGRVNSDVALKRYLLGLVENINKEAA
Ce1D	581	VLYGDVNGDGRVNSDITLTKRYVLLKAVSTLPSSKAEGA
Ce1E	411	ILYGDVNGDGRVNSDITLTKRYVLLRGIEEFPSPSGIIA
Ce1F	666	IMLGDVNFQGRVNSDYSRLKRYVTKSLEFDPPEHQKFIAA
Ce1G	499	VTYGDVNSDGMVNSDITLTKRYLLKKNPPASANMD
Ce1H	828	IKHGDVNFQAVNSDITLMLKRYLLKSLKELGTSEHEEKFKKA
Ce1J	1288	VVYGDVNSDGMVNSDITLTKRYVLLKGIIDNINLTA
Ce1S	675	KLYGDVNFQGRVNSDVAVALKRYVLRSGISINTDN
Ce1X		VKKGDVNFQGVNSDITLTKRYVLLKVVVDINSINVTN
XynY	730	VLLGDVNGDGNVNSDITLTKRSLVLRATLTDDAKAR
XynZ	426	TGLGDVNGDGNVNSDITLTKRSLVLRATLTDGALLR
LicB	269	PLKGDVNGDGHVNSDYSLLKRYVLRVTRFPPVGDQSV
CipA	1791	MWVGLVTKNSINLLVAEVIKCFNATKGSANYVEE
XynC	601	ADLNEDGRVNSDITLTKRYVLLKREL-COOH
Ce1A	449	ADVNRDCAVNSDITLTKRYVLLKSPHLPY-COOH
Ce1B	534	ADVNSGVNSDITLTKRYVLLKSELPPY-COOH
Ce1D	621	ADVNRDGRVNSDITLTKRYVLLKREKLPY-COOH
Ce1E	451	ADVNAELKINSDITLTKRYVLLKIDKFPED...
Ce1F	709	ADVNGNCRVNSDITLTKRYVLLKHEKFAEQ...
Ce1G	536	ADVNAELKINSDITLTKRYVLLKREKLPHTT-COOH
Ce1H	871	ADLNRRNKVNSDITLTKRYVLLKASEIPI-COOH
Ce1J	1323	ADLYFDGVNSDITLTKRYVLLKAEIDIPY...
Ce1S	711	ADLNEDGRVNSDITLTKRYVLLKREKLPYKN-COOH
Ce1X		ADLNEDGNVNSDITLTKRYVLLKREKLPYKN-COOH
XynY	766	ADVNRDGRVNSDITLTKRYVLLKREKLPYVAE...
XynZ	464	ADVNRDGRVNSDITLTKRYVLLKREKLPYVAE...
LicB	308	ADVNRDGRVNSDITLTKRYVLLKREKLPYVAE...
CipA	1824	LDLNRRCAINMQEIMIVKHKFGATSSDYDAQ-COOH

FIG. 5. Alignment of dockerin domains of XynC, Ce1A, Ce1B, Ce1D, Ce1E, Ce1F, Ce1G, Ce1H, Ce1J, Ce1S, Ce1X, CipA, LicB, XynY, and XynZ of *C. thermocellum*. Amino acids which are conserved in at least 7 of the 15 sequences are highlighted. Numbers at the start of the respective lines refer to amino acid residues; all sequences are numbered from Met-1 of the peptide.

end products accompanied by small amounts of xylose. By contrast, this enzyme was less active toward xylotriose and not active at all toward xylobiose. The enzyme activity was completely inhibited by HgCl₂, FeCl₃, and CuCl₂ and was partly inhibited by MnCl₂, AlCl₃, and *p*-chloromercuribenzoic acid at a concentration of 1 mM. The optimum pH for activity was found to be pH 5.5 when the enzyme activity was assayed by 10-min incubation at 60°C in Britton and Robinson's universal buffer solutions at various pHs. The enzyme was quite stable in the range of pH 4.0 to 11.0, when incubated at 25°C for 12 h in the same buffer solutions without the substrate. The effects of temperature on the activity and stability of the enzyme were examined. The optimum temperature for activity was found to be 80°C at pH 5.5. The enzyme was stable at 70°C for 10 min at pH 5.5 in the absence of the substrate; keeping the temper-

ature at 80°C for 10 min resulted in complete loss of enzyme activity.

DISCUSSION

The presence of xylanase activity has been often reported to be associated with the cellulosome of *C. thermocellum* (27, 35), although this bacterium is unable to grow on xylan and xylose (55). Recently, we have shown that the largest catalytic subunit, Ce1J, has xylanase activity and that its xylanase activity is ascribed to a family 44 catalytic domain (1, 2). However, the xylanase genes cloned from *C. thermocellum*, i.e., *xynY* and *xynZ*, could not be related to the major components of the cellulosome. Therefore, this is the first report about a xylanase gene encoding a major catalytic component of the *C. thermocellum* cellulosome.

As shown in Fig. 3, XynC comprises a single catalytic domain of family 10 in addition to the so-called thermostabilizing domain and a dockerin domain. The crystal structures of four different xylanases in family 10 have been resolved, and two catalytic residues were identified (11, 12, 23, 54). Residues Glu-754, identified as the nucleophile, and Glu-645, identified as the proton donor, in *C. thermocellum* XynZ are conserved as Glu-469 and Glu-352, respectively, in XynC (Fig. 3). Residues Trp-795 and Glu-721, which surround the nucleophile, His-723, which conditions the protonation state of the nucleophile, and Asp-641, which forms a buried salt bridge in XynZ (12), are also conserved in XynC and other enzymes in this family. The highest sequence identity to XynC is observed with XynB of *B. fibrisolvens* (44.2%) (14) rather than with the enzymes of *C. thermocellum*, i.e., XynX (38.4%) (GenBank accession no. M67438), XynY (43.5%) (15), and XynZ (34.5%) (21).

XynC has a dockerin domain in the C terminus of the peptide. Since the dockerin of XynC is highly homologous to many other docking domains conserved in the catalytic subunits of the *C. thermocellum* cellulosome (Fig. 5), it may be assumed to mediate docking of XynC to the scaffolding protein CipA. The presence of a dockerin domain allowed us to anticipate that XynC was a member of the cellulosome, and this enzyme was then identified in the cellulosomal proteins as a major component. Ali et al. fractionated the cellulosome of *C. thermocellum* NCIB 10682 into several subpopulations by ion-exchange chromatography (3). One such population with high activity on Avicel contained a subunit, J4, with strong xylanase activity, which appeared to be equivalent to the major component S9 or S10 of *C. thermocellum* YS, and the N-terminal amino acid

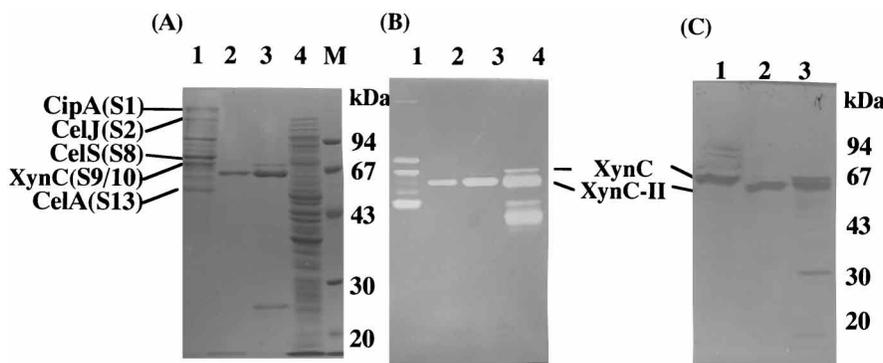


FIG. 6. Expression of XynC in *C. thermocellum* F1 and *E. coli*. Gels were stained with Coomassie brilliant blue (A) or stained for xylanase activity (B). XynC proteins in Western blot analysis were detected with a polyclonal mouse antiserum raised against the purified XynC-II (C). Lanes 1, cellulosomal proteins of *C. thermocellum* F1; lanes 2, purified XynC-II; lanes 3, proteins in an active fraction from a DEAE-Toyopearl column; lanes 4, whole-cell proteins of *E. coli* XL2-Blue; lane M, protein mass standards. The proteins corresponding to the major components specified in the *C. thermocellum* YS cellulosome (31) are shown to the left of panel A.

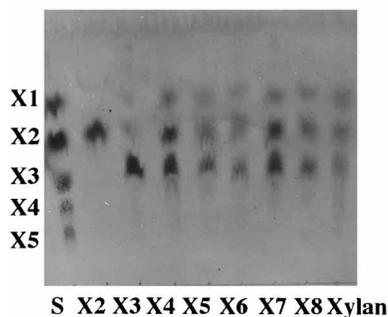


FIG. 7. Thin-layer chromatography of hydrolysis products from xylan and xylooligosaccharides. Each xylooligosaccharide (X2 to X8) (10 mg) was incubated with enzyme (0.1 U) for 2 h, and the hydrolysates were analyzed by thin-layer chromatography. S, authentic oligosaccharides; X1, xylose; X2, xylobiose; X3, xylotriose; X4, xylotetraose; X5, xylopentaose; X6, xylohexaose; X7, xyloheptaose; X8, xylooctaose.

sequence of J4 was completely identical to those of XynC reported in this study. These findings suggest that XynC is a major catalytic subunit of the cellulosomes from different strains of *C. thermocellum*. N-terminal amino acid sequence analysis of several components of the *C. thermocellum* F1 cellulosome disclosed that CelA and CelS were also contained in the cellulosome as major catalytic components (data not shown), indicating that the organization of the cellulosome of *C. thermocellum* F1 resembles those of *C. thermocellum* NCIB 10682 and YS.

Thermostabilizing domains are found mainly in the thermophilic xylanases of family 10. Exceptionally, *R. flavefaciens* XynA contains a thermostabilizing domain in addition to two distinct catalytic domains, i.e., family 11 and 16 domains (Fig. 2). Although *C. fimi* is a mesophilic bacterium, *C. fimi* XynC is optimally active at 60°C (10). The number and position of the thermostabilizing domain are variable in respective enzymes; e.g., a thermostabilizing domain occurs in the N terminus of *C. thermocellum* XynC, two occur in the N terminus of *T. saccharolyticum* XynA, and one occurs in the middle of *C. thermocellum* XynY (Fig. 2). Although removal of these domains from *C. thermocellum* XynY and *T. saccharolyticum* XynA reduced their thermal stabilities and optimal temperatures, the interaction between catalytic domains and thermostabilizing domains responsible for thermostabilization remains to be studied.

The main difficulty encountered for purification of this enzyme from the recombinant *E. coli* was the cleavage of the protein during cultivation and purification. As a result, we obtained the truncated enzyme, XynC-II, in a purified form. Since the N-terminal amino acid sequence of XynC is identical to that of XynC-II, it is apparent that proteolytic truncation occurs within the dockerin domain in the C terminus. Similar proteolysis within a dockerin region was observed in the recombinant CelD of *C. thermocellum* (47) and the recombinant CelA of *Clostridium cellulolyticum* expressed in *E. coli* (40). Therefore, these enzymes seem to contain fragile regions in the dockerins recognized by *E. coli* protease(s). On the other hand, Western blotting showed that the immunoreactive protein in the cellulosome had a molecular weight identical to that of XynC produced in recombinant *E. coli*. These findings suggest that XynC is not cleaved by *C. thermocellum* F1 protease(s) and further that it is not heavily glycosylated by this bacterium.

Family 10 enzymes exhibit in general a broad substrate specificity; e.g., Cex of *C. fimi* was first characterized as an exoglu-

canase having activity on crystalline cellulose, while it turned out to hydrolyze β -1,4-xyloside linkage more efficiently than β -1,4-glucoside linkage (19). XynC-II also exhibited a broad substrate specificity; i.e., it hydrolyzed xylan, PNPX, PNPC, PNPG, and CMC. However, the specific activity of this enzyme on CMC (0.18 U/mg) is lower than that on xylan (557 U/mg). Therefore, XynC could not contribute to the hydrolysis of the cellulose chain. On the other hand, since *C. thermocellum* cannot utilize xylan as a sole carbon source, xylanases do not have a function to supply this bacterium with usable saccharides for its growth. XynC 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 accessible unless xylan is hydrolyzed and removed.

ACKNOWLEDGMENTS

We are grateful to Y. Mitsuishi of the National Institute of Bioscience and Human Technology, Tsukuba, Japan, for the generous gift of xylooligosaccharides.

This investigation was partially supported by a Grant-in-Aid for Scientific Research (B) (04454076) from the Ministry of Education, Science, and Culture of Japan.

REFERENCES

- Ahsan, M. M., T. Kimura, S. Karita, K. Sakka, and K. Ohmiya. 1996. Cloning, DNA sequence, and expression of the gene encoding *Clostridium thermocellum* cellulase CelJ, the largest catalytic component of the cellulosome. *J. Bacteriol.* **178**:5732–5740.
- Ahsan, M. M., M. Matsumoto, S. Karita, T. Kimura, K. Sakka, and K. Ohmiya. Purification and characterization of the family J catalytic domain. *Biosci. Biotechnol. Biochem.*, in press.
- Ali, B. R. S., M. P. M. Romaniec, G. P. Hazlewood, and R. B. Freedman. 1995. Characterization of the subunits in an apparently homogeneous subpopulation of *Clostridium thermocellum* cellulosomes. *Enzyme Microb. Technol.* **17**:705–711.
- Bayer, E. A., E. Morag, and R. Lamed. 1994. The cellulosome—a treasure-trove for biotechnology. *Trends Biotechnol.* **12**:379–385.
- Béguin, P., P. Cornet, and J.-P. Aubert. 1985. Sequence of a cellulase gene of the thermophilic bacterium *Clostridium thermocellum*. *J. Bacteriol.* **162**:102–105.
- Béguin, P., and M. Lemaire. 1996. The cellulosome: an exocellular, multi-protein complex specialized in cellulose degradation. *Crit. Rev. Biochem. Mol. Biol.* **31**:201–236.
- Biely, P. 1985. Microbial xylanolytic systems. *Trends Biotechnol.* **3**:286–290.
- Cantor, C. R., and P. R. Schimmel. 1980. *Biophysical chemistry, part III. The behavior of biological macromolecules.* W. H. Freeman & Co., New York, N.Y.
- Choi, S. K., and L. G. Ljungdahl. 1996. Dissociation of the cellulosome of *Clostridium thermocellum* in the presence of ethylenediaminetetraacetic acid occurs with the formation of truncated polypeptides. *Biochemistry* **35**:4897–4905.
- Clarke, J. H., K. Davidson, H. J. Gilbert, C. M. G. A. Fontes, and G. P. Hazlewood. 1996. A modular xylanase from mesophilic *Cellulomonas fimi* contains the same cellulose-binding and thermostabilizing domains as xylanases from thermophilic bacteria. *FEMS Microbiol. Lett.* **139**:27–35.
- Derewenda, U., L. Swenson, R. Green, Y. Wei, R. Morosoli, F. Shareck, D. Kluepfel, and Z. S. Derewenda. 1994. Crystal structure, at 2.6-Å resolution, of the *Streptomyces lividans* xylanase A, a member of the family of β -1,4-D-glycanases. *J. Biol. Chem.* **269**:20811–20814.
- Dominguez, R., H. Souchon, S. Spinelli, Z. Dauter, K. S. Wilson, S. Chauvaux, P. Béguin, and P. M. Alzari. 1995. A common protein fold and similar active site in two distinct families of β -glycanases. *Nat. Struct. Biol.* **2**:569–576.
- Felix, C. R., and L. G. Ljungdahl. 1993. The cellulosome: the exocellular organelle of *Clostridium*. *Annu. Rev. Microbiol.* **47**:791–819.
- 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.
- 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.
- Fukumura, M., K. Sakka, K. Shimada, and K. Ohmiya. 1995. Nucleotide sequence of *Clostridium stercoararium xynB* gene encoding an extremely ther-

- mostable xylanase, and characterization of the translated product. *Biosci. Biotechnol. Biochem.* **59**:40–46.
17. **Gasparic, J., and J. Churacek.** 1978. Laboratory handbook of paper and thin-layer chromatography, p. 153–154. Ellis Horwood Ltd., Chichester, England.
 18. **Gerngross, U. T., M. P. A. Romaniec, T. Kobayashi, N. S. Huskisson, and A. L. Demain.** 1993. Sequencing of a *Clostridium thermocellum* gene (*cipA*) encoding the cellulosomal S₁-protein reveals an unusual degree of internal homology. *Mol. Microbiol.* **8**:325–334.
 19. **Gilkes, N. R., M. Claeysens, R. Aebersold, B. Henrissat, A. Meinke, H. D. Morrison, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, Jr.** 1991. Structural and functional relationships in two families of β -1,4-glycanases. *Eur. J. Biochem.* **202**:367–377.
 20. **Grépinet, O., and P. Béguin.** 1986. Sequence of the cellulase gene of *Clostridium thermocellum* coding for endoglucanase B. *Nucleic Acids Res.* **14**:1791–1799.
 21. **Grépinet, O., M.-C. Chebrou, and P. Béguin.** 1988. Nucleotide sequence and deletion analysis of the xylanase gene (*xynZ*) of *Clostridium thermocellum*. *J. Bacteriol.* **170**:4582–4588.
 22. **Hall, J., G. P. Hazlewood, J. P. Barker, and H. Gilbert.** 1988. Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. *Gene* **69**:29–38.
 23. **Harris, G. W., J. A. Jenkins, I. Connerton, N. Cummings, L. L. Leggio, M. Scott, J. P. Hazlewood, J. I. Lalurie, H. J. Gilbert, and R. W. Pickersgill.** 1994. Structure of the catalytic core of the family F xylanase from *Pseudomonas fluorescens* and identification of the xylopentaose-binding sites. *Structure* **2**:1107–1116.
 24. **Henrissat, B., and A. Bairoch.** 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**:695–696.
 25. **Joliff, G., P. Béguin, and J.-P. Aubert.** 1986. Nucleotide sequence of the cellulase gene *celD* encoding endoglucanase D of *Clostridium thermocellum*. *Nucleic Acids Res.* **14**:8605–8613.
 26. **Katsube, Y., Y. Hata, and H. Yamaguchi.** 1990. Estimation of xylanase active site from crystalline structure, p. 91–96. *In* M. Ikehara (ed.), Protein engineering. Japan Scientific Societies Press, Tokyo, Japan.
 27. **Kohring, S., J. Wiegel, and F. Mayer.** 1990. Subunit composition and glycosidic activities of the cellulase complex from *Clostridium thermocellum* JW20. *Appl. Environ. Microbiol.* **56**:3798–3804.
 28. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
 29. **Lamed, R., E. Setter, and E. A. Bayer.** 1983. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *J. Bacteriol.* **156**:828–836.
 30. **Lee, Y.-E., S. L. Lowe, B. Henrissat, and G. Zeikus.** 1993. Characterization of the active site and thermostability regions of endoxylanase from *Thermoanaerobacterium saccharolyticum* B6A-R1. *J. Bacteriol.* **175**:5890–5898.
 31. **Lemaire, M., and P. Béguin.** 1993. Nucleotide sequence of the *celG* gene of *Clostridium thermocellum* and characterization of its product, endoglucanase CelG. *J. Bacteriol.* **175**:3353–3360.
 32. **Lin, L.-L., and J. A. Thomson.** 1991. Cloning, sequence and expression of a gene encoding a 73kDa xylanase enzyme from the rumen anaerobe *Butyrivibrio fibrisolvens* H17c. *Mol. Gen. Genet.* **228**:55–61.
 33. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**:265–275.
 34. **Miller, G. L.** 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**:426–428.
 35. **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.
 36. **Morag, E., E. A. Bayer, and R. Lamed.** 1992. Affinity digestion for the near-total recovery of purified cellulosome from *Clostridium thermocellum*. *Enzyme Microb. Technol.* **14**:289–292.
 37. **Mori, Y.** 1992. Purification and characterization of an endoglucanase from the cellulosome (multicomponent cellulase complex) of *Clostridium thermocellum*. *Biosci. Biotechnol. Biochem.* **56**:1199–1203.
 38. **Navarro, A., M.-C. Chebrou, P. Béguin, and J.-P. Aubert.** 1991. Nucleotide sequence of the cellulase gene *celF* of *Clostridium thermocellum*. *Res. Microbiol.* **142**:927–936.
 39. **Neu, H. C., and L. A. Heppel.** 1965. The release of enzyme from *Escherichia coli* by osmotic shock and during the formation of spheroplast. *J. Biol. Chem.* **240**:3685–3692.
 40. **Pages, S., A. Belaich, C. Tardif, C. Reverbel-Leroy, C. Gaudin, and J.-P. Belaich.** 1996. Interaction between the endoglucanase CelA and the scaffolding protein CipC of the *Clostridium cellulolyticum* cellulosome. *J. Bacteriol.* **178**:2279–2286.
 41. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**:319–353.
 42. **Sakka, K., S. Furuse, and K. Shimada.** 1989. Cloning and expression in *Escherichia coli* of thermophilic *Clostridium* sp. F1 gene related to cellulose hydrolysis. *Agric. Biol. Chem.* **53**:905–910.
 43. **Sakka, K., S. Karita, K. Ohmiya, and K. Shimada.** 1994. Xylanases of *Clostridium stercoararium* and *Clostridium thermocellum* isolated in Japan, p. 32–38. *In* K. Shimada, S. Hoshino, K. Ohmiya, K. Sakka, Y. Kobayashi, and S. Karita (ed.), Genetics, biochemistry and ecology of lignocellulose degradation. Uni Publisher Co., Ltd., Tokyo, Japan.
 44. **Sakka, K., T. Shimanuki, and K. Shimada.** 1991. Nucleotide sequence of *celC307* encoding endoglucanase C307 of *Clostridium thermocellum*. *Agric. Biol. Chem.* **55**:347–350.
 45. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 46. **Schimming, S., W. H. Schwarz, and W. L. Staudenbauer.** 1992. Structure of the *Clostridium thermocellum* gene *licB* and encoded β -1,3-1,4-glycanase. *Eur. J. Biochem.* **204**:13–19.
 47. **Tokatlidis, K., S. Salamitou, P. Béguin, P. Dhurjati, and J.-P. Aubert.** 1991. Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. *FEBS Lett.* **291**:185–188.
 48. **Tomme, P., R. A. J. Warren, and N. R. Gilkes.** 1995. Cellulose hydrolysis by bacteria and fungi. *Adv. Microb. Physiol.* **37**:1–81.
 49. **Törröne, A., and J. Rouvinen.** 1994. Structural comparison of two major endo-1,4- β -xylanases from *Trichoderma reesei*. *Biochemistry* **34**:847–856.
 50. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**:4350–4354.
 51. **Wang, W. K., K. Kruus, and J. H. D. Wu.** 1993. Cloning and DNA sequence of the gene coding for *Clostridium thermocellum* cellulase Ss (CelS), a major cellulosome component. *J. Bacteriol.* **175**:1293–1302.
 52. **Wang, W. K., and J. H. D. Wu.** 1993. Structural features of the *Clostridium thermocellum* cellulase Ss gene. *Appl. Biochem. Biotechnol.* **39/40**:149–158.
 53. **Watson, M. E. E.** 1984. Compilation of published signal sequences. *Nucleic Acids Res.* **12**:5145–5164.
 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. **Wiegel, J., C. P. Mothershed, and J. Puls.** 1985. Differences in xylan degradation by various noncellulolytic thermophilic anaerobes and *Clostridium thermocellum*. *Appl. Environ. Microbiol.* **49**:656–659.
 56. **Winterhalter, C., P. Heinrich, A. Candussio, G. Wich, and W. Liebl.** 1995. Identification of a novel cellulose-binding domain within the multidomain 120kDa xylanase XynA of the hyperthermophilic bacterium *Thermotoga maritima*. *Mol. Microbiol.* **15**:431–444.
 57. **Wu, J. H. D., W. H. Orme-Johnson, and A. L. Demain.** 1988. Two components of an extracellular protein aggregate of *Clostridium thermocellum* together degrade crystalline cellulose. *Biochemistry* **27**:1703–1709.
 58. **Yagüe, E., P. Béguin, and J.-P. Aubert.** 1990. Nucleotide sequence and deletion analysis of the cellulase-encoding gene *celH* of *Clostridium thermocellum*. *Gene* **89**:61–67.
 59. **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.