Nucleotide Sequence of the Clostridium stercorarium \textit{xynB} Gene Encoding an Extremely Thermostable Xylanase, and Characterization of the Translated Product

Masayuki Fukumura, Kazuo Sakka, Kyo Shimada, and Kunio Ohmiya

Faculty of Bioresources, Mie University, Tsu 514, Japan
Received June 24, 1994

The nucleotides of the \textit{xynB} gene of \textit{Clostridium stercorarium} F-9 were sequenced. The structural gene consists of an open reading frame of 1161 bp encoding a xylanase (XynB) in family F of 387 amino acids with a molecular weight of 44,377. The molecular weight of the enzyme purified from a recombinant \textit{Escherichia coli} was around 41,000, smaller than the predicted value, on SDS-polyacrylamide gel electrophoresis due to the lack of 32 amino acids at the N-terminus. Intact XynB with a molecular weight of around 43,000 was immunologically detected in the total cell proteins of a recombinant \textit{E. coli} and \textit{C. stercorarium} F-9. The purified XynB was active toward xylan, carboxymethylcellulose, p-nitrophenyl-\(\beta\)-D-xlylopyranoside and p-nitrophenyl-\(\beta\)-D-celllobioside. The pH optimum was 7.0 and it was quite stable over the pH range of 5 to 12 at 4°C. This enzyme was optimally active at 80°C and retained about 50% of the original activity even after incubation at 100°C for 10 min.

Xylan, a major component of hemicellulose in plant cell walls, is the most abundant form of biomass second to cellulose and is now regarded as a natural resource that is convertible to biofuels, chemicals, and value-added compounds by microbial fermentation or enzymatic process. In recent years, thermostable xylanases have received considerable attention for potential application in biomass conversion strategies and further in bioleaching of pulp, i.e., mild and safe procedures for pulp improvement by xylanase treatment. Several xylanases and their genes have been characterized from thermophilic anaerobic bacteria such as \textit{C. stercorarium}, \textit{C. thermocellum}, and \textit{Caldocellum saccharolyticum}. Previously, we have cloned several genes related to xylan hydrolysis from \textit{C. stercorarium} F-9, which had been isolated in this laboratory, and sequenced the \textit{xynA} gene encoding a bifunctional protein with \(\beta\)-D-xylanase and \(\alpha\)-L-arabinofuranosidase activities and the \textit{xynA} gene encoding a family G xylanase (XynA). Immunological analysis has recently shown that the \textit{xynA} gene was predominantly expressed as a xylanase gene in \textit{C. stercorarium} F-9. Although some other xylanase genes were also cloned from a type strain, NCIB, of \textit{C. stercorarium}, these genes and the gene products have not yet been characterized in detail.

In this paper, the nucleotide sequence of the \textit{xynB} gene, another xylanase gene from \textit{C. stercorarium} F-9, is reported. We also describe some properties of the enzyme purified from a recombinant \textit{E. coli}.

Materials and Methods

Bacterial strains and plasmids. \textit{E. coli} strain M109 (recA1 supE44 endA1 hsdRI7 gyrA96 relA1 thrA1 (lac-proAB) [F'proD36 proAB-lacZAM15]) was used as a host strain for production of XynB and \textit{E. coli} XLI-Blue (supE44 hsdRI7 recA1 endA1 gyrA96 thiA1 lacI [F'proAB-lacZM15 Tn10 (tetR)]) was used as a host strain for preparation of single-stranded plasmid DNAs for nucleotide sequencing of the \textit{xynB} gene. The cloning vectors used for subcloning were pBlueScript II KS+ and KS (Stratagene Cloning Systems, La Jolla, U.S.A.). A plasmid pMF6 (Fig. 1) containing the \textit{xynB} gene is a derivative of pVK401.

Recombinant DNA techniques and nucleotide sequencing. Plasmid DNAs were purified from \textit{E. coli} by the method of Birnboim and Doly. \textit{E. coli} cells were transformed with the recombinant plasmids by the method of Hanahan. For preparing single stranded DNA plasmids, the 2.4-kbp Smal-SalI fragment of pMF6 was subcloned to pBlueScript II KS+ and KS- to construct pMF6-1 and pMF6-2, respectively (Fig. 1). A series of nested deletion mutants for nucleotide sequencing from pMF6-1 and pMF6-2 was constructed by exonuclease III-mung bean nuclease digestions using the protocol and reagents of Nippon Gene Co., Ltd. (Tokyo, Japan). The chain termination method was done with single-stranded DNA templates, dye-labeled custom primers (T3 and T7 primers), and Taq DNA polymerase, using a Dye Primer Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, U.S.A.) and products were analyzed on a model 373A automated DNA sequencer system (Applied Biosystems Inc.). Nucleotide and amino acid sequence analyses were performed using GENETYX computer software (Software Development Co., Ltd., Tokyo, Japan). Standard techniques described by Sambrook et al. were used for other DNA manipulations. Restriction enzymes, exonuclease III, and T4 DNA ligase were generously donated by Nippon Gene Co., Ltd.

Purification and characterization of XynB. \textit{E. coli} strain M109 (pMF6) cells were harvested from 1200 ml of LB broth, suspended in 20 mM Tris(hydroxymethyl)aminomethane-HCl (Tris-HCl) buffer (pH 8.0), and then disrupted by ultrasonication on ice. After removal of the cell debris by centrifugation at 4000 \times g for 10 min, the enzyme was precipitated by adding solid ammonium sulfate to 40% saturation, and dissolved in

![Fig. 1. Restriction Maps of pMF6 and its Derivatives.](image-url)

Plasmid construction is described in Materials and Methods. Thin lines correspond to vector plasmid DNAs and the box corresponds to the inserted DNA fragment. The coding region is depicted by oblique lines. Arrow (→) indicates the direction of the transcription of \textit{xynB} from a chloristrial promoter. Strategy for sequence of \textit{xynB} is indicated by arrows, PBS, pBlueScript, Plac, lac promoter.
Tris–HCl buffer. The enzyme solution was put on a DEAE-Toyopearl 650S column (2.2 × 20 cm, Tosoh Co., Tokyo, Japan) equilibrated with Tris–HCl buffer, and elution was done with a linear gradient of NaCl (0 to 0.4 M) in equilibration buffer. Fractions with xylanase activity were combined, dialyzed against 50 mM sodium phosphate–12 mM sodium citrate (PC) buffer (pH 6.3) containing 1 M ammonium sulfate, and then put on a butyl-Toyopearl column (1 × 10 cm, Tosoh) equilibrated with the same buffer. The enzyme was eluted with a descending gradient of ammonium sulfate (1.0 to 0 M). The active fractions were dialyzed against PC buffer containing 0.5 M NaCl and further fractionated on a gel filtration column, TSK-gel G3000SW (0.75 × 60 cm, Tosoh). The active fractions were combined and dialyzed against PC buffer. The purified enzyme thus obtained was used for characterization. N-terminal amino acid sequences of XynB were analyzed by a protocol of Applied Biosystems.

**Enzyme assays.** Xylanase activity was measured by 10-min incubation at 60°C in PC buffer or 20 mM phosphate buffer in the presence of 1% oat-spleet xylan (Fluka Ag, Buchs, Switzerland) unless otherwise stated. Reducing sugars released from the substrate were analyzed with the 3,5-dinitrosalicylic acid reagent as described by Miller. 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-celllobiosidase, and β-glucosidase activities were assayed at 60°C with p-nitrophenyl-β-d-xylopyranoside (PNPX, Sigma, St. Louis, U.S.A.), p-nitrophenyl-β-d-celllobioside (PNPC; Sigma), and p-nitrophenyl-β-d-glucopyranoside (PNPG; Sigma), respectively. One unit of these enzyme activities was defined as the amount of enzyme releasing 1 μmol of p-nitrophenol per min from each substrate. Carboxymethylcellulase (CMCase) activity was assayed as described previously. One unit of CMCase activity was defined as the amount of enzyme releasing 1 μmol of glucose equivalents per min from carboxymethylcellulose (CMC).

**Measurement of protein.** The protein in the enzyme solution was determined from the UV absorption at 280 nm. One milligram of XynB obtained by freeze-drying was dissolved in 1 ml of distilled water and an A280 value was measured.

**Isoelectric focusing electrophoresis.** Isoelectric focusing electrophoresis was done using 10-ml and 30-ml columns and Ampholine (pH 2.5–6.0) by the method of Doi.

**Preparation of antiserum and immunological detection of XynB.** Antiserum against the purified XynB was raised in BALB/c mouse. For immunological detection, protein samples were fractionated by SDS-polyacrylamide gel electrophoresis and transferred onto a nitrocellulose filter membrane using an electroblotting apparatus. Sartoblot II (Sartorius, Göttingen, Germany). Immunoreactive proteins were detected on Western blots by enzyme immunoassay using peroxidase conjugated...
to goat antimusite immunoglobulins (Tago Inc., Barlingame, U.S.A.) and 3,3'-diaminobenzidine tetrahydrochloride.

**Nucleotide sequence accession number.** The nucleotide sequence reported in this paper has been registered in the DDBJ, EMBL, and Gen Bank Nucleotide Sequence Databases with the accession number D12504.

**Results**

**Nucleotide sequence of xynB**

Figure 1 is a summary of the subclones used for sequencing xynB and the directions of sequencing. Figure 2 shows the xynB structural gene along with its flanking regions. There is an open reading frame composed of 1161 nucleotides encoding a protein of 387 amino acids with a predicted molecular weight of 44,377. It should be noted that no cysteine residue is contained in this enzyme. The ATG initiation codon at nucleotide position 187 is preceded by a putative Shine-Dalgarno sequence, AGGA, which allows base pairing between mRNA and the 3'-end of bacterial 16s rRNA.27 The reading frame is ended by the ochre stop codon TAA at position 1348. Possible promoter sequences, TGACA for the −35 region and TATAAT for the −10 region, which were found by primer extension (data not shown), with a 15-bp spacing between them were observed. These sequences show a high homology to the consensus promoter sequences for σ70 factor found in *E. coli* (TTGACA and TATAAT with a 17-bp spacing).28 The 26-bp palindromic sequence for mRNA hairpin formation, a possible ρ-independent transcription terminator,29 was found downstream from the TAA stop codon.

**Amino acid sequence homology between XynB and other xylanases**

The alignment of the amino acid sequence of XynB with those of xylanases hitherto reported disclosed that XynB contained a domain with high homology with that of XynA of alkalophilic *Bacillus* sp. strain C-125 (identity: 54%)30 and homology with other enzymes in family F,31 i.e., 41% identity with XynA of *C. saccharolyticum*,13 33% identity with Cex of *Cellulomonas fimii*,32 34% identity with XynZ of *C. thermocellum*,11 41% identity with CelB of *C. saccharolyticum*,10 30% identity with XylA of *Pseudomonas fluorescens* subspecies *cellulosa*,33 27% identity with XynA of *Butyrivibrio fibrisolvens*,34 and 34% identity with ORF4 of *C. saccharolyticum*.15 Therefore, the domain in XynB should be classified as a family F catalytic domain.31 As shown in Fig. 3, the alignment of the amino acid sequences of the catalytic domains classified into family F showed that the highly conserved stretches are localized in 8 regions (I–VIII). The moieties other than the catalytic domain of *C. stercorarium* XynB do not show any homology with other functional domains such as the cellulose-binding domain.

**Purification of XynB from a recombinant E. coli and detection of XynB in C. stercorarium F-9**

The XynB gene product was purified from the cellular fraction of *E. coli* JM109 (mpF6) by ammonium sulfate precipitation, and DEAE-Toyopearl 650S, butyl-Toyopearl 650S, and TSK-gel G3000SW column chromatographies. On DEAE-Toyopearl 650S column, 2 peaks with xylanase activity appeared. The major fraction, which was eluted with 0.2m NaCl, was further purified. The result of the purification of this enzyme is summarized in the Table. The final preparation gave a single band on SDS–polyacrylamide gel electrophoresis and the molecular weight of the enzyme was estimated to be around 41,000 (Fig. 4A, lane 3). This enzyme was denoted as XynB. This value was slightly smaller than that deduced from the nucleotide sequence (44,377). The N-terminal amino acid sequence of XynB was identified as Lys-Ala-Phe-Asn-Asp-Gln-Thr-Ser-Ala-Glu, which was found in the deduced amino acid sequence of XynB at amino acid positions 33 to 42 (Fig. 2). From the minor xylanase fraction from DEAE-column, two additional xylanases with molecular weights of around 43,000 and 41,000 were purified together as shown in Fig. 4A, lane 4, and denoted XynB-I and XynB-II, respectively. The N-terminal amino acid sequence of the large protein, XynB-I (Fig. 4A, lane 4), found by blotting it to PVDF membrane from the gel of SDS–polyacrylamide gel electrophoresis, was identified as Met-Asn-Lys-Phe-Leu, which starts from the first Met soon after the SD sequence (Fig. 2). The molecular weight of XynB-I, 43,000, from SDS–polyacrylamide gel electrophoresis was reasonably comparable to that of 44,377 deduced from the amino acid sequence (387 amino acid residues). The sequence of the smaller protein, XynB-II, was Lys-Ala-Phe-Asn-Asp, which was at 33–37 amino acids (Fig. 2). The sequence of XynB was consistent with that of XynB-II, indicating that XynB and XynB-II were the same protein. These results indicated that 32 amino acids were removed from the N-terminus of XynB by proteolysis, resulting in the formation of a protein species with a predicted molecular weight of 40,805.

To find whether the xynB gene is normally expressed by *E. coli* and *C. stercorarium* F-9, total cell proteins of *E. coli* (mpF6) and total cell proteins and extracellular proteins from a culture supernatant of *C. stercorarium* F-9 were analyzed by Western blotting, using antiserum prepared against the purified XynB (Fig. 4A, lane 5 and 4B). When *C. stercorarium* F-9 was grown on ball-milled cellulose, a single immunoreactive band with a molecular weight of

**Table Summary of Purification of XynB from E. coli JM109(mpF6)**

<table>
<thead>
<tr>
<th>Substrate: xylan</th>
<th>Total activity (units)</th>
<th>Total protein (mg)</th>
<th>Specific activity (U/mg)</th>
<th>Recovery (%)</th>
<th>Purification (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude enzyme</td>
<td>78,100</td>
<td>581</td>
<td>134</td>
<td>100</td>
<td>1.0</td>
</tr>
<tr>
<td>DEAE-TOYO PEARL</td>
<td>25,200</td>
<td>56.8</td>
<td>444</td>
<td>32.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Butyl-TOYO PEARL</td>
<td>13,000</td>
<td>12.6</td>
<td>1030</td>
<td>16.6</td>
<td>7.7</td>
</tr>
<tr>
<td>TSK-G3000SW</td>
<td>10,400</td>
<td>2.3</td>
<td>4460</td>
<td>13.3</td>
<td>33.3</td>
</tr>
</tbody>
</table>

Fig. 3. Alignment of Similar Regions of Primary Structures of the *C. stercorarium* F-9 Celluloxylanase (XynB; XynB-t), the *C. saccharolyticum* Xylanase (XynA; XynA-sa),13 (ORF4-sa),13 the *C. saccharolyticum* Exoglucanase (CelB; CelB-sa),14 the *C. thermocellum* Xylanase (XynZ; XynZ-th),11 the Akalophilic *Bacillus* sp. Strain C125 Xylanase (XynA; XynA-B),31 the *C. fimii* Exoglucanase (Cex; Cex-fim),29 the *P. fluorescens* subsp. Cellulosa Xylanase (XylA; XylA-B),34 and the *B. fibrisolvens* xylanase (XynA; XynA-f).35 Conserved regions are indicated by bold bars and numbers; (I–VIII). The numbers in front of each sequence note the first amino acids of individual lines. The amino acids conserved more than 7 out of 9 amino acids are shadowed.
about 43,000 was detected in the total cell proteins (Fig. 4B, lane 3) but not in the proteins grown on xylan (Fig. 4B, lane 4). On the other hand, a minor immunoreactive band with a molecular weight of around 43,000 (Fig. 4A, lane 5), which seemed to be a parental molecular species, was detected in addition to a major band (Fig. 4A, lane 5) corresponding to the purified XynB (Fig. 4A, lane 3 and 4B, lane 2). These supported the idea that the xynB gene product was mainly processed in *E. coli* but was synthesized in the intact state in *C. stercorarium* F-9.

**General characterization of XynB**
XynB has a higher activity toward CMC as well as xylan, and a lower activity toward PNPC and PNPX, although
xyN was cloned as a xylanase gene.\(^{40}\) The initial reaction rates measured at 60°C in various concentrations of xylan, CMC, PCNP, and PNPX for Lineweaver-Burk plots showed that the V\(_{\text{max}}\) was estimated to be 4.5 (µmol/min/µg) for xylan, 1.1 (µmol/min/µg) for CMC, 170 (nmol/min/µg) for PCNP, and 25 (nmol/min/µg) for PNPX, and the K\(_{\text{m}}\) to be 3.7 (mg/ml) for xylan, 13.7 (mg/ml) for CMC, 3.6 (mm) for PCNP, and 4.0 (mm) for PNPX. XyN was optimally active at pH 6.0 (Fig. 5A) and quite stable from pH of 5 to 12 at 4°C for 24 h (Fig. 5B). The molecular extinction coefficient of XyN of 0.1% at 280 nm was 1.5. This value was comparable to that (1.6) calculated by using 1576 M\(^{-1}\) cm\(^{-1}\) for tyrosine (\(\times 16\)) and 5225 M\(^{-1}\) cm\(^{-1}\) for tryptophan (\(\times 8\)) at 280 nm. The isoelectric point of XyN lacking in 32 amino acids at N-terminus was 4.2. XyN showed its maximum activity against xylan at 80°C from pH 5.8–6.4 but no activity at 100°C (Fig. 5C). The stability of the enzyme activity at the temperatures higher than 70°C was affected by incubation pHs while it was quite stable up to 70°C, i.e., it retained 50 and 58% of the original activity after incubation at 100°C for 10 min at pHs 6.1 and 6.4, respectively, but no activity after incubation at pH 5.8 (Fig. 5D). Enzyme activity was completely inactivated by Hg\(^{2+}\) ion but not affected by other metal ions, EDTA and p-chloromercuribenzoic acid.

**Discussion**

The xynB gene, which was cloned as a xylanase gene, conferred a hydrolytic activity toward CMC and PNPC in addition to xylan and PNPX. This wide range of substrate specificity is common to all the enzymes in family F.\(^{31}\) e.g., XynA of C. saccharolyticum\(^{46}\) and Cex of C. fimii\(^{35}\) were active on these 4 substrates above and XynZ of C. thermodenitrificans\(^{10}\) was active on p-nitrophenyl-β-d-glucopyranoside as well as xylan and PNPC but not on CMC. XynB of C. stercorarium also hydrolyzed CMC without reduction of viscosity of CMC solution (data not shown), supporting the idea that the enzymes in family F generally attack CMC in an exo-fashion.\(^{35}\)

It has been confirmed that the xynA gene is predominantly expressed as a xylanase gene in C. stercorarium F-9.\(^{80}\) On the other hand, XyN was inducibly synthesized as a cellular protein by addition of ball-milled cellulose but not xylan. Therefore, this enzyme might be important for using cellulose rather than xylan by hydrolyzing cellobiose-saccharides as an intracellular enzyme.

As shown in Fig. 3, the catalytic domains in family F, which consist of about 300 amino acids, have 8 well-conserved regions locally. Tull et al.\(^{36}\) identified the active site nucleophile as a glutamic acid residue in the absolutely conserved sequence, Thr-Glu-Leu-Asp (TELD) in Cex of C. fimii (the conserved region VI in Fig. 3). Enzymatic hydrolysis of cellulosic substrates presumably proceeds by general acid catalysis promoted by carboxylic acid residues (Asp or Glu) via a carbonium ion intermediate, which is stabilized either by a carboxyl group or by a histidine residue.\(^{37}\) Other than the glutamic acid residue in the TELD motif, 3 aspartic acid, 1 glutamic acid, and 2 histidine residues are absolutely conserved in the conserved regions II, III, IV, VI, and VII, and one of these amino acid residues might function as the active site in conjunction with the identified glutamic acid residue.

General characterization of XyN showed that this enzyme was extremely thermostable, i.e., it retained half of the initial enzyme activity even after incubation at 100°C for 10 min (Fig. 5) while the enzyme activity of XyN was not detected at 100°C. Therefore, it is apparent that XyN is denatured and inactivated by heating at 100°C and it recovers its activity by restoring the proper tertiary structure from the denatured state. One of the causes of irreversible thermal inactivation of enzymes is assigned to a combination of disulfide interchange and β-elimination of cysteine residues.\(^{38}\) Oshima stated that the absence of a cysteine residue in *Thermus thermophilus* proteins such as isopropylmalate dehydrogenase contribute to the thermal stability.\(^{39}\) The fact that this extremely thermostable enzyme, XyN, contains no cysteine residue supports the explanation about the role of cysteine residues in thermal stability and irreversible thermal inactivation of proteins.

Detailed analysis of thermal stability and irreversible inactivation of XyN will be reported elsewhere.

**Acknowledgments**

This research was in part supported by a Grant-in-Aid for Scientific Research on Priority Areas (Energy) No. 03203234 (1991) and No. 04202219 (1992, 93) from the Ministry of Education, Science, and Culture of Japan. The authors are grateful to Nagoya Seiraku Co., Ltd., Japan.

**References**