Xylanase B from *Neocallimastix patriciarum* contains a non-catalytic 455-residue linker sequence comprised of 57 repeats of an octapeptide

Gary W. BLACK,* Geoffrey P. HAZLEWOOD,† Gang-Ping XUE,‡ Colin G. ORPIN† and Harry J. GILBERT*§

*Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, U.K.,

†Department of Cellular Physiology, AFRC Babraham Institute, Babraham, Cambridge CB2 4AT, U.K., and ‡CSIRO Division of Tropical Crops and Pastures, 306 Carmody Road, St Lucia, Queensland 4067, Australia

A *Neocallimastix patriciarum* cDNA library was screened for xylanase-expressing clones, which were distinct from the previously characterized *N. patriciarum xynA* cDNA encoding xylanase A. A single cDNA, designated *xynB*, which did not exhibit homology with *xynA*, was isolated. Northern-blot analysis of mRNA from Avicel-grown *N. patriciarum* showed that *xynB* hybridized to a 3.4 kb mRNA species. The nucleotide sequence of *xynB* revealed a single open reading frame of 2580 bp coding for a protein designated xylanase B (XYLB), of *M*_r 88066. The primary structure of XYLB was comprised of a 21-residue N-terminal signal peptide, followed by a 304-amino acid sequence that exhibited substantial homology with the catalytic domains of family F xylanases. The N-terminal domain was linked to a C-terminal 70-residue sequence by a putative linker region, comprising 12 tandem repeats of a sequence containing TLPG as the core sequence, followed by an

octapeptide XSKTLPGG where X can be S, K or N, which was repeated in tandem 45 times. Truncated derivatives of *xynB* encoding the N-terminal 338 residues directed the synthesis of a functional xylanase, confirming that the region of XYLB, which exhibited homology with family F xylanases, constitutes the catalytic domain. To investigate the catalytic properties of XYLB, the catalytic domain was fused to the *Escherichia coli* maltose-binding protein, and the fusion protein purified by amylose affinity chromatography. The purified enzyme hydrolysed oat, rye and wheat arabinoxylan releasing primarily xylobiose, xylotriose and some xylose. The XYLB fusion did not cleave any cellulosic substrates. The data presented in this report suggest that the multiple xylanases of *N. patriciarum* arose, not through the duplication of a single gene, but by the transfer of distinct xylanase-encoding DNA sequences into the anaerobic fungus. The possible origin of the *xynB* gene is discussed.

INTRODUCTION

Xylan, the major hemicellulose component of the plant cell wall, consists of a polymer of β -1,4-linked β -D-xylopyranose units substituted with arabinosyl, glucuronosyl and acetyl residues. A repertoire of microbial enzymes act in synergy to hydrolyse xylan: these include endo- β -1,4-xylanases (EC 3.2.1.8), which hydrolyse the xylan backbone, and a series of enzymes that cleave the side-chain constituents (Biely, 1985).

Xylanolytic micro-organisms generally express isoenzymic forms of endoxylanases that are encoded by multiple genes (Gilbert et al., 1988; Hazlewood et al., 1988; Flint et al., 1989; Clarke et al., 1991). In view of the similarity of the bond cleaved (β -1,4-glycosidic linkages) and the cross-specificity sometimes observed between cellulases and xylanases (Hall et al., 1988), the evolutionary relationship between these enzymes is an interesting question. Recently, sequence alignment and hydrophobic cluster analysis have been utilized to assign cellulases and xylanases to nine enzyme families (Gilkes et al., 1991). The xylanases were grouped into two enzyme families that did not contain true cellulases, suggesting that the two enzyme species have evolved from distinct ancestral genes.

The molecular architecture of endoxylanases conforms to three basic patterns: (i) small (20–30 kDa) single-domain enzymes (Ujiié et al., 1991); (ii) modular enzymes comprising multiple catalytic domains (Zhang and Flint, 1992); (iii) modular enzymes consisting of a catalytic domain fused to a non-catalytic cellulose-binding domain (CBD; Ferreira et al., 1990; Kellett et al., 1990; Sakka et al., 1993). The role of the CBD in enzyme function is controversial. There is some evidence to suggest that the domain

plays a direct role in the action of cellulases against crystalline cellulose (Maglione et al., 1992). However, the rationale for the acquisition of CBDs by other plant cell wall hydrolases is currently unknown.

Many hemicellulolytic/cellulolytic micro-organisms, comprising both eukaryotes and prokaryotes, reside in the bovine and ovine rumen. The rumen hemicellulolytic/cellulolytic fungi belong to the order Chytridiomycetes and include *Neocallimastix patriciarum* (Teunissen et al., 1992). The *Neocallimastix* cellulase system consists of a large aggregating multienzyme complex which exhibits very high activity towards crystalline cellulose (Wilson and Wood, 1992). Although the primary structure and molecular architecture of endoxylanases from aerobic fungi and several prokaryotes, including rumen bacteria, have been analysed, there is a paucity of information on the corresponding anaerobic fungal enzymes. The analysis of *N. patriciarum* xylanases is of particular interest as the data could provide important insights into the evolution of these enzymes. For example, is there evidence of horizontal transfer between rumen prokaryotes and eukaryotes, or have *Neocallimastix* and aerobic fungal xylanases evolved from the same progenitor sequences?

To elucidate the structure of the hemicellulases from *N. patriciarum* we have cloned cDNA sequences encoding *Neocallimastix* xylanases and investigated the primary structure and molecular architecture of the encoded enzymes. In a previous report we showed that the fungus synthesized a xylanase (XYLA) consisting of two identical catalytic domains and a small C-terminal repeated region (Gilbert et al., 1992). It remains to be elucidated whether multiple catalytic domains are a common feature of *Neocallimastix* plant cell wall hydrolases, as is the case

Abbreviations used: CBD, cellulose-binding domain; LB, Luria broth; 5'-RACE, rapid amplification of 5' cDNA ends; ORF, open reading frame.

§ To whom correspondence should be addressed.

The nucleotide sequence of the *xynB* gene has been submitted to the EMBL, GenBank and DDBH Nucleotide Sequence Databases under the accession no. X76919.

for the corresponding enzyme system expressed by *Caldocellum saccharolyticum* (Saul et al., 1990), and whether multiple *Neocallimastix* xylanases (if they exist) are a consequence of gene duplication or of the transfer of different xylanase genes into the lower eukaryote. To provide answers to these questions, we have carried out an extensive screen of a *Neocallimastix* cDNA library for xylanase-expressing clones. One xylanase-expressing clone, which was distinct from the XYLA cDNA (*xynA*), was isolated. The encoded enzyme showed no identity with XYLA and contains a single catalytic domain which is linked, via a putative 455 residue linker sequence consisting of repeating units of an octapeptide sequence, to a small non-catalytic C-terminal domain. The enzyme displayed characteristics typical of an endo- β -1,4-xylanase.

MATERIALS AND METHODS

Microbial strains, vectors and culture conditions

The *Escherichia coli* strain employed in this study was XL1-Blue (Stratagene). The vectors utilized were pMAL-p2 (New England Biolabs), pBluescript SK- (Stratagene), λ ZAPII (Stratagene) and pMTL22p (Chambers et al., 1988). *E. coli* recombinants containing *bla*⁺ plasmids were cultured in Luria broth (LB) supplemented with 100 μ g/ml ampicillin. To induce expression of xylanase activity in strains containing *malA-xynBcd* or *malA-xynBfl*, bacteria were cultured in LB supplemented with ampicillin and 0.1 mM isopropyl β -D-thiogalactopyranoside. Recombinant and xylanase-expressing strains were selected on LB-agar (2%) supplemented with 5-bromo-4-chloroindol-3-yl β -D-galactopyranoside (2 μ g/ml) and Blue Xylan (500 μ g/ml) respectively. *N. patriciarum* was cultured in anaerobic medium containing Avicel as sole carbon source (Kemp et al., 1984).

General recombinant DNA methodology

Plasmid DNA was purified from alkali-lysed *E. coli* cells using Qiagen columns (Hybaid) and following the protocol recommended by the manufacturer. *E. coli* transformation, agarose-gel electrophoresis, Southern-blot hybridization, Northern blots and modification of plasmid DNA using restriction endonucleases and T4 DNA ligase were as described by Sambrook et al. (1989). The construction and screening for xylanase-positive clones of a cDNA library of mRNA extracted from Avicel-grown *N. patriciarum* was conducted as described by Gilbert et al. (1992). The cDNA inserts expressing xylanase activity were excised into pBluescript SK- using VCS-M13 helper phage.

To sequence *xynB*, appropriate plasmid DNA was denatured by alkali, neutralized and further purified by spin dialysis (Murphy and Kavanagh, 1988). Dideoxy chain-termination sequencing (Sanger et al., 1980) of the resultant DNA was performed using the Sequenase sequencing kit (USB). Overlapping sequences were generated by creating nested deletions of the cDNA insert contained in pGB1 using exonuclease III. Sequences were compiled and ordered using the computer programs of Staden (1980).

The 5' end of *xynB* was extended using rapid amplification of 5' cDNA ends (5'-RACE). Employing *N. patriciarum* mRNA as template, single-stranded cDNA was synthesized by reverse transcription, using primer 3 (see Figure 3), and poly(A)⁺ tailed. Two rounds of PCR amplification were carried out using primer 2 (Figure 3) and an *SpeI* adaptor-(dT)₁₇ primer (CATCGATATCTAGATACTAG[T]₁₇). The DNA generated was then subjected to 30 cycles of PCR using the oligo(dT) primer and primer 1 (Figure 3), restricted with *HindIII* (*HindIII* site at the 5' end of

the *xynB* cDNA in the primer 1 sequence) and *SpeI*, and cloned into suitably restricted pBluescript. The 5' end of the cDNA insert in 25 of the resultant recombinant plasmids recovered was sequenced. The majority of the sequences (22) were identical with the 5' end of cloned *xynB*, and four of the plasmids contained cDNA inserts that extended the 5' sequence of *xynB*.

In order to construct a hybrid protein containing the catalytic region of XYLB fused to the *E. coli* maltose-binding protein, PCR was used to amplify the 5' end of *xynB* encoding the catalytic domain of XYLB. The primers 5'-CTCGGATCCG-CATCCGCCAATGCATATG-3' and 5'-CTCGTCGACT-CAGGTCTTGGCAGAAGATGAAGC-3' were employed to amplify the sequence between map positions 86 and 1054. The reaction conditions for the 30 cycles were as follows: annealing, 1 min at 56 °C; elongation, 2 min at 72 °C; denaturation, 1 min at 94 °C. The amplified DNA was restricted with *BamHI* and *SalI* and cloned into suitably digested pMAL-p2. The sequence between 995 and 2432 was amplified using primers 4 and 5 (Figure 3) as described above.

Assays

Cellulase- and xylanase-related activities were assayed at 37 °C in 50 mM potassium phosphate/12 mM citric acid buffer, pH 6.5, as described by Hazlewood et al. (1990) and Kellett et al. (1990). The h.p.l.c. analysis of xylan hydrolysis products was essentially as described by Kellett et al. (1990) except that rye, wheat and oat spelt xylan were employed as substrates (0.2%), and the sugars were eluted from the Dionex PA1 column with 100 mM NaOH containing a 0–75 mM sodium acetate gradient. The sugars were detected with a pulsed amperometer. Protein was determined as described by Sedmak and Grossberg (1977). SDS/PAGE was performed using the method of Laemmli (1970). Antiserum raised against XYLB fusion proteins was used in Western-blot analysis of *N. patriciarum* extracellular proteins as described by Hall et al. (1988).

Purification of fusion proteins

Approximately 50 ml of cell-free extract (derived from 1 litre of overnight culture of *E. coli* harbouring the *malA-xynB* fusion genes), prepared as described by Clarke et al. (1991), was allowed to percolate through an amylose column (25 mm $\times$ 100 mm) under gravity. The column was then washed with 10 vol. of 20 mM Tris/HCl buffer, pH 7.4, containing 200 mM NaCl and 1 mM EDTA (column buffer), and bound protein eluted with 10 mM maltose. The disaccharide was removed by dialysis against 4 $\times$ 1000 vol. of column buffer. The purified fusion protein, which was shown by SDS/PAGE to have an *M_r* of 73000, was subjected to N-terminal sequence analysis using a 470A gas-phase-sequencer equipped with a 120A on-line phenylthiohydantoin-derivative analyser (Applied Biosystems).

RESULTS AND DISCUSSION

Cloning of *xynB*

A cDNA library, constructed in λ ZAPII, using mRNA isolated from *N. patriciarum* cells grown on Avicel, was screened for xylanase-expressing clones that contained genes that were distinct from the previously described *xynA*. One recombinant, which hydrolysed oat spelt xylan but did not contain a DNA insert that hybridized to *xynA*, was isolated. The cDNA sequence and its encoded enzyme were designated *xynB* and XYLB respectively. The 2.8 kb *xynB* sequence was excised from λ ZAPII as a recombinant of pBluescriptSK-(pGB1), and was used to probe for *xynB* sequences in mRNA isolated from *N. patriciarum*

Figure 1 Northern-blot hybridization of *N. patriciarum* RNA

Total RNA was extracted from Avicel-grown *N. patriciarum* as described by Gilbert et al. (1992), and 15 μ g (lane 1) and 2 μ g (lane 2) of RNA was fractionated by denaturing agarose-gel electrophoresis using formaldehyde and formamide as the denaturants. The RNA, after transfer to Hybond N filters, was probed with 32 P-radiolabelled *xynB*. The filter was exposed to X-ray film for 16 h. The size of the major RNA species which hybridized to *xynB* is shown.

cultured on Avicel. Data presented in Figure 1 show that *xynB* hybridized predominantly to a 3.4 kb mRNA species, suggesting that pGB1 did not contain a full-length *xynB* cDNA. 5'-RACE was thus employed to extend the 5' region of *xynB* by a further 200 bp. Restriction maps of plasmids containing full-length and truncated derivatives of *xynB* are displayed in Figure 2.

Nucleotide sequence of *xynB*

The 2.8 kb *Neocallimastix* cDNA derived from pGB1 was sequenced in both strands (Figure 3). Translation of the nucleotide sequence revealed a single open reading frame (ORF) of 2580 bp encoding a polypeptide of M_r 88066. The assignment of the translation initiation codon is based on the presence of translational stop codons in all three phases upstream of the ATG start codon, and the very high A + T content (90% A + T)

Figure 2 Physical maps of recombinant plasmids containing full-length and truncated forms of *xynB*

The positions of the cleavage sites for *EcoRI* (E), *BclI* (B), *BstXI* (Bs), *HpaI* (Hp), *HindIII* (H), *XbaI* (Xa) and *XhoI* (X) are shown. Plasmids pGB3-6 contain 3' deleted forms of *xynB* inserted into pBluescript SK-. The position of the 3' nucleotide of truncated *xynB* in these plasmids is shown in the nucleotide sequence of *xynB* (Figure 3). The *xynB* insert in pGB7 consists of a 2.5 kb *EcoRI*-*XhoI* restriction fragment cloned into pMTL22p in phase with the vector's *lacZ'* gene. The inserts in pGB8 and pGB9 were cloned into the *Bam*HI and *SaI*I sites of pMAL-p2, such that the xylanase gene was in frame with *malA*. The solid arrow shows the extent and orientation of *xynB*. The catalytic activity of XYLB encoded by the recombinant plasmids is shown (+ or -).

upstream and downstream of the putative ORF [excluding the poly(A)⁺ tail], compared with an A + T content of 58% for the putative ORF. The overall G + C content of *Neocallimastix* DNA is approximately 15%, indicating that non-coding regions are generally very (A + T)-rich. This is consistent with the observation that other *Neocallimastix* cDNAs contain extremely (A + T)-rich sequences upstream and downstream of the respective ORFs. The bias in codon utilization is evident from the absence of 15 codons. There is a marked preference for T in the third position (50% of all codons end in T) and an exclusion of G from the wobble position. In general, the codon utilization of *xynB* is very similar to the three other *Neocallimastix* cDNAs previously sequenced (Gilbert et al., 1992; Reymond et al., 1992; Zhou et al., 1994).

Primary structure of XYLB

Inspection of the deduced primary structure of mature XYLB revealed several interesting features. The N-terminal region of the xylanase conforms to that of a signal peptide, consisting of an N-terminal hydrophilic basic region followed by a sequence of 13 predominantly hydrophobic or neutral amino acids. The presence of an N-terminal signal peptide is consistent with the extracellular location of microbial plant cell wall hydrolases.

The recent elucidation of the primary structures of a range of xylanases has provided an insight into the evolution of these enzymes. Using a combination of hydrophobic cluster analysis and sequence alignment studies, xylanases can be grouped into two enzyme families, F and G (Henrissat, 1992). Proteins within a family are structurally related and are thought to have evolved from a common ancestral sequence. Comparison of XYLB with sequences in the EMBL database revealed homology between the fungal enzyme and other xylanases belonging to family F (Figure 3). Family F consists predominantly of bacterial enzymes, and the multiple xylanases from filamentous fungi such as *Trichoderma* and *Aspergillus* are located in family G (Henrissat, 1992). This suggests that XYLB has not evolved from the progenitor sequences that gave rise to xylanases of aerobic lower eukaryotes, but is related to prokaryotic hemicellulases. Within family F are three enzymes derived from the rumen bacteria

Figure 3 Nucleotide sequence of *xynB*

The nucleotide sequence of *xynB* and the deduced primary structure of XYLB is shown. The reiterated sequence is boxed. The 3' nucleotide of each truncated derivative of *xynB* is indicated by a vertical arrow, and the designation of the plasmid encoding the derivative is also shown. The position of the primers 1-5 are indicated by overlining the appropriate sequences.

Butyrivibrio fibrisolvens (Henrissat, 1992) and *Ruminococcus flavefaciens* (Zhang and Flint, 1992). The homology between XYLB and rumen prokaryotic xylanases could be a consequence of the horizontal transfer of genes between rumen prokaryotes and lower eukaryotes, either when the organisms were resident in the rumen or before their colonization of the ruminant.

It is now well established that cellulolytic micro-organisms often express numerous isoforms of endoglucanases (Gilbert et al., 1987; Hazlewood et al., 1988). These proteins can be found in different enzyme families, suggesting that the acquisition of multiple enzymes is not the result of extensive gene duplication, but reflects the horizontal transfer of different genes into a single cellulolytic organism (Gilkes et al., 1991). Although several micro-organisms also express multiple forms of endoxylanases (Gilbert et al., 1988; Clarke et al., 1991; Lüthi et al., 1990), different isoforms of true endoxylanases from a single organism have, to date, all belonged to the same enzyme family. For example *Pseudomonas*, *Caldocellum* and *Butyrivibrio* xylanases are in family F, whereas the corresponding *Trichoderma reesei* enzymes are in family G (Henrissat, 1992). In contrast, the two *Neocallimastix* xylanases, XYLA (Gilbert et al., 1992) and XYLB, belong to families G and F respectively. In addition, *xynA* from *Ruminococcus flavefaciens* (Zhang and Flint, 1992) encodes a xylanase consisting of two discrete catalytic domains that belong to each of the two enzyme families. However, it is not clear whether *xynA* resulted from the transfer of a single gene encoding a multidomain enzyme or was created by the fusion of two discrete genes contained within the *Ruminococcus* genome. However, the presence of isoenzymic forms of xylanases belonging to families F and G in *Neocallimastix*, and probably *Ruminococcus*, suggests that some xylanolytic organisms have acquired xylanase isoenzymes through the transfer of multiple genes, and not exclusively by gene duplication.

The most noteworthy feature of XYLB is the primary structure between residues 335 and 790, designated B. This region contains 12 tandem repeats of a sequence which contains TLPG as its core sequence, which is followed by an octapeptide with a consensus sequence of XSKTLPGG, where X is either S, N or K, which is repeated in tandem 45 times. Confirmation that region B was not the result of a cloning artifact was obtained by the amplification of a 1.4 kb *N. patriciarum* genomic sequence, using primers 4 and 5, which flank the reiterated region of *xynB* (Figure 3). The precise role for region B is currently unknown. However, linker sequences rich in hydroxy amino acids and proline is a common feature of modular plant cell wall hydrolases. As the octapeptide repeat contains S, T and P, it is possible that this region constitutes a linker sequence, although the basis for the evolution of a 455-residue linker sequence, which contains a highly conserved repeated element, is not readily apparent. Most linker regions are small consisting of 10–50 residues (Gilkes et al., 1991). Although regular repeats of P and T are observed in several *Clostridium* and *Cellulomonas* enzymes, reiteration of an octapeptide has only been reported in one enzyme, xylanase A (XYLA) from *Ruminococcus flavefaciens* (Zhang and Flint, 1992). XYLA contains a 374-residue linker region comprised almost exclusively of D, N, Q and W. Within the sequence there is some reiteration of the octapeptide QQQNNDWN, although the reiteration of this sequence does not conform to the regular structure of the *Neocallimastix* enzyme.

Downstream of the repeated sequence is a 70-residue C-terminal region designated domain C. Small domains are also observed at the C-terminus of the other two *Neocallimastix* plant cell wall hydrolases, CELB and XYLA, that have been characterized to date (Gilbert et al., 1992; Zhou et al., 1994). Although the C-terminal domain is highly conserved in XYLA

and CELB, the corresponding sequence in XYLB exhibits little identity with XYLA or CELB. Similarly, searches of the SWISS protein database did not reveal any polypeptides that show extensive homology with the C-terminus of XYLB.

Functional analysis of XYLB

To test the assertion that the N-terminal 374-residue sequence of XYLB constitutes the catalytic domain, truncated derivatives of *xynB* were constructed, lacking 3' sequences, and the capacity of the encoded proteins to hydrolyse xylan was evaluated. The data (Figure 2) showed that a derivative of *xynB* lacking 1566 bp of 3' sequence (pGB5) directed the synthesis of a catalytically active xylanase, whereas removal of 1793 bp of 3' from *xynB* (pGB6) generated a derivative of the cDNA that did not encode a functional enzyme. *E. coli* containing a truncated form of *xynB* lacking 336 bp of 5' coding sequence did not display xylanase activity (pGB7). These data confirm that regions B and C are non-catalytic domains, and the catalytic centre of XYLB is located in the N-terminal region of the enzyme.

Non-catalytic domains are common features of modular plant cell wall hydrolases. The major role for these sequences, in addition to linker functions, are (i) as non-catalytic CBDs (Gilbert and Hazlewood, 1993) and (ii) as recognition sequences functional during complex-formation (Tokatlidis et al., 1991). Region C does not exhibit sequence identity with the numerous CBDs characterized to date. In addition, XYLB encoded by full-length *xynB* does not bind to either cellulose or xylan (results not shown). This suggests that domain C does not constitute a CBD. With regard to domain C constituting a docking sequence, it has been established that *Neocallimastix* spp. form large cellulase complexes. It is likely therefore that plant cell wall hydrolases from this fungus contain sequences that interact with other proteins to form the enzyme aggregates. However, the docking sequences of *Clostridium thermocellum* plant cell wall hydrolases, which interact with the non-catalytic scaffolding protein S₁ exhibit extensive sequence identity (Tokatlidis et al., 1991). The C-terminal region of XYLB does not show homology to CELB or XYLA, suggesting that the non-catalytic domains of XYLB and the other two enzymes fulfil distinct roles. It remains to be established whether the two C-terminal sequences constitute docking sequences which recognize different proteins, or whether they have more divergent roles in enzyme function.

Properties of XYLB

To evaluate the catalytic properties of XYLB attempts were made to purify the full-length enzyme. However, because of its low expression in *E. coli*, purification of the enzyme was not feasible. As an alternative strategy, hybrid genes, designated *malA-xynBfl* and *malA-xynBcd*, were constructed encoding the maltose-binding protein fused respectively to either full-length XYLB (MBP-FL) or the catalytic domain of XYLB (MBP-CD). When cell-free extracts, derived from *E. coli* harbouring *malA-xynBcd*, were incubated with amylose, a major protein of *M.* 73000 and a minor 42 kDa polypeptide were eluted from the polysaccharide with maltose (Figure 4). The purified protein fraction displayed xylanase activity, and the first ten residues of the N-terminal sequence of the 73 kDa polypeptide were identical with the corresponding region of the maltose-binding protein, in addition to having an *M_r* value similar to that predicted for MBP-CD. These data confirm the identity of the 73 kDa polypeptide as MBP-CD. Attempts to purify the MBP-FL fusion protein were less successful; only very low levels of protein could be recovered from the amylose matrix with maltose. Analysis of this fraction revealed weak xylanase activity and a protein of

Figure 4 SDS/PAGE of MBP-CD fusion protein purified by agarose affinity chromatography

The fusion enzyme and the maltose binding protein were purified as described in the Materials and methods section. The samples, which were subjected to SDS/PAGE, using a 10% (w/v) polyacrylamide gel, were as follows: protein purified from *E. coli* cells harbouring pGB8 (lane 1); cell-free extract derived from *E. coli* cells harbouring pGB8 (lane 2); protein purified from *E. coli* cells containing pMAL-p2 (lane 3); cell-free extract from *E. coli* containing pMAL-p2 (lane 4). The gel was stained with Coomassie Blue, and the M_r ($\times 10^{-3}$) of the MBP-CD fusion protein (which was subjected to N-terminal sequencing) is shown.

Figure 5 Western analysis of XYLB expressed by cultured *N. patriciarum*

N. patriciarum was cultured on soluble (1) and insoluble (2) xylan, and the extracellular protein subjected to Western analysis using antisera raised against MBP-CD as the probe. The M_r ($\times 10^{-3}$) of the major immunoreactive polypeptide is shown.

M_r 73 000, suggesting that MBP-FL was subjected to extensive processing in *E. coli*. The cleavage of non-catalytic domains of modular xylanases in *E. coli* has been described for both prokaryotic (Hall et al., 1989) and fungal enzymes (Gilbert et al., 1992). To investigate whether XYLB is processed in *N. patriciarum*, antisera raised against MBP-CD was used to probe extracellular proteins derived from the fungus cultured on xylan. The data (Figure 5) revealed a 93 kDa protein, which is similar to the predicted size of full-length XYLB, that cross-reacted with the antiserum. It would appear therefore that XYLB is neither proteolytically processed nor subjected to extensive glycosylation in *N. patriciarum*.

Evaluation of the catalytic properties of MBP-CD revealed that the enzyme rapidly hydrolysed both unsubstituted (oat spelts) and highly substituted (rye and wheat) forms of arabinoxylans (Table 1), releasing reducing sugar and

Table 1 Substrate specificity of XYLB

The concentrations of the polysaccharides and aryl β -glycosides employed as substrates were 2 mg/ml and 0.5 mM respectively. The enzyme (0.06 unit using *p*-nitrophenyl β -cellobioside as substrate) was incubated with the various substrates for 15 min at 37 °C. Assays were performed using the MBP-CD fusion protein. Enzyme activity is expressed as a percentage of the activity against *p*-nitrophenyl β -cellobioside. The specific activity of the xylanase using *p*-nitrophenyl β -cellobioside as the substrate was 45 units/mg of protein where 1 unit of enzyme activity is the amount of enzyme required to release 1 μ mol of xylose equivalents/min at 37 °C.

Substrate	Relative enzyme activity (%)
Avicel	0
Amorphous cellulose	0
β -glucan (β 1-4, β 1-3 with ratio of 3:1)	0
Lichenan (alternate β 1-4, β 1-3)	0
Cellobiose-celohexaose	0
Laminarin (β 1-3)	0
CM-cellulose	0
Methylumbelliferyl β -cellobioside	114
Methylumbelliferyl β -glucoside	0
<i>p</i> -Nitrophenyl β -arabinoside	0
<i>p</i> -Nitrophenyl β -cellobioside	100
<i>p</i> -Nitrophenyl β -glucoside	0
<i>p</i> -Nitrophenyl β -xyloside	0
<i>p</i> -Nitrophenyl β -xylobioside	0
Oat spelt xylan	435
Rye xylan	691
Wheat xylan	607

Table 2 Products generated from xylans incubated with MBP-CD

MBP-CD (0.12 unit) was incubated with 1.5 mg of substrate for 4 h at 37 °C after which time the release of reducing sugar had plateaued. Products generated are expressed as percentages of total sugars released. The detection and quantification of the products generated are described in the Materials and methods section.

Substrate	Products generated
Oat spelt xylan	Xylose (10%), xylobiose (45%), xylotriose (45%)
Rye xylan	Xylose (21%), xylobiose (44%), xylotriose (35%)
Wheat xylan	Xylose (13%), xylobiose (46%), xylotriose (41%)

promoting a rapid decline in substrate viscosity. The enzyme exhibited no activity against a range of cellulosic substrates, although it did hydrolyse several aryl β -glycosides. In the literature there is some evidence to suggest that specific endoxylanases exhibit significant arabinofuranosidase activity (Biely et al., 1992; Matte and Forsberg, 1992). However, when defining the substrate specificity of native plant cell wall hydrolases purified from the endogenous organism, there is always the risk that trace contaminants of other polysaccharidases can produce misleading results. Thus it is possible that the observation that some endoxylanases exhibit arabinofuranosidase activity could be a consequence of contamination of the xylanase preparations with arabinofuranosidase. To evaluate whether XYLB exhibits arabinofuranosidase activity, the purified *E. coli*-derived enzyme was incubated with highly substituted arabinoxylans from wheat and rye, and the products generated were analysed by h.p.l.c. The data revealed the presence of primarily xylose, xylobiose and xylotriose. There was no evidence that XYLB released arabinose from these substrates (Table 2). These data clearly show that XYLB from *N. patriciarum* is an endo-

xylanase with no arabinofuranosidase activity. It is expected that the cloning of genes encoding endoxylanases with apparent arabinofuranosidase into non-cellulolytic/xylanolytic hosts, such as *E. coli*, will establish whether the side-chain cleaving activity of these enzymes is a real phenomenon.

We thank the Agricultural and Food Research Council (grant LE13/138) for supporting this work and K. Gobius, L. M. Dierens and G. D. Simpson for technical support in the cloning of the cDNA.

REFERENCES

- Biely, P. (1985) Trends Biotechnol. **3**, 286–290
- Biely, P., Vrsanska, J. and Kucar, S. (1992) in Xylans and Xylanases (Visser, J., Beldman, G., Kusters-van Someren, M. A. and Voragen, A. G. J., eds.), pp. 81–96, Elsevier Science Publishers, Amsterdam
- Chambers, S. P., Prior, S. E., Barstow, D. A. and Minton, N. P. (1988) Gene **68**, 139–149
- Clarke, J. H., Laurie, J. I., Gilbert, H. J. and Hazlewood, G. P. (1991) FEMS Microbiol. Lett. **83**, 305–310
- Ferreira, L. M. A., Durrant, A. J., Hall, J., Hazlewood, G. P. and Gilbert, H. J. (1990) Biochem. J. **269**, 261–264
- Flint, H. J., McPherson, C. A. and Bisset, J. (1989) Appl. Environ. Microbiol. **55**, 1230–1233
- Gilbert, H. J. and Hazlewood, G. P. (1993) J. Gen. Microbiol. **139**, 187–194
- Gilbert, H. J., Jenkins, G., Sullivan, D. A. and Hall, J. (1987) Mol. Gen. Genet. **210**, 551–556
- Gilbert, H. J., Sullivan, D. A., Jenkins, G., Kellett, L. E., Minton, N. P. and Hall, J. (1988) J. Gen. Microbiol. **134**, 3239–3247
- Gilbert, H. J., Hazlewood, G. P., Laurie, J. I., Orpin, C. G. and Xue, G.-P. (1992) Mol. Microbiol. **6**, 2065–2072
- Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, Jr., R. C. and Warren, R. A. J. (1991) Microbiol. Rev. **55**, 303–315
- Hall, J., Hazlewood, G. P., Huskisson, N. S., Durrant, A. J. and Gilbert, H. J. (1988) Gene **69**, 29–38
- Hazlewood, G. P., Romaniec, M. P. M., Davidson, K., Grepinet, O., Béguin, P., Millet, J., Raynaud, O. and Aubert, J.-P. (1988) FEMS Microbiol. Lett. **51**, 231–236
- Hazlewood, G. P., Davidson, K., Laurie, J. I., Romaniec, M. P. M. and Gilbert, H. J. (1990) J. Gen. Microbiol. **136**, 2089–2097
- Henrissat, B. (1992) in Xylans and Xylanases (Visser, J., Beldman, G., Kusters-van Someren, M. A. and Voragen, A. G. J., eds.), pp. 97–111, Elsevier Science Publishers, Amsterdam
- Kemp, P., Lander, D. J. and Orpin, C. G. (1984) J. Gen. Microbiol. **130**, 27–37
- Kellett, L. E., Poole, D. M., Ferreira, L. M. A., Durrant, A. J., Hazlewood, G. P. and Gilbert, H. J. (1990) Biochem. J. **272**, 369–376
- Laemmli, U. K. (1970) Nature (London) **227**, 680–685
- Lüthi, E., Love, D. R., McAnulty, J., Wallace, C., Caughey, P. A., Saul, D. and Bergquist, P. L. (1990) Appl. Environ. Microbiol. **56**, 1017–1024
- Maglione, G., Matsushita, O., Russell, J. B. and Wilson, D. B. (1992) Appl. Environ. Microbiol. **58**, 3593–3597
- Matte, A. and Forsberg, C. W. (1992) Appl. Environ. Microbiol. **58**, 157–168
- Murphy, G. and Kavanagh, T. (1988) Nucleic Acids Res. **16**, 5198
- Reymond, P., Geourjon, C., Roux, B., Durand, R. and Fevre, M. (1992) Gene **110**, 57–63
- Sakka, K., Kojima, Y., Kondo, T., Karita, S., Ohmiya, K. and Shimada, K. (1993) Biosci. Biotechnol. Biochem. **57**, 273–277
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. (1980) J. Mol. Biol. **143**, 161–178
- Saul, D. J., Williams, L. C., Grayling, R. A., Chamley, L. W., Love, D. R. and Bergquist, P. L. (1990) Appl. Environ. Microbiol. **56**, 3117–3124
- Sedmak, J. J. and Grossberg, J. R. (1977) Anal. Biochem. **79**, 544–552
- Staden, R. (1980) Nucleic Acids Res. **16**, 3673–3694
- Teunissen, M. J., de Kort, G. V. M., Op den Camp, H. J. M. and Huis in't Veld, J. H. J. (1992) J. Gen. Microbiol. **138**, 1657–1664
- Tokatlidis, K., Salamitou, S., Béguin, P., Dhurjati, P. and Aubert, J.-P. (1991) FEBS Lett. **291**, 185–188
- Ujije, M., Roy, C. and Yaguchi, M. (1991) Appl. Environ. Microbiol. **57**, 1860–1862
- Wilson, C. A. and Wood, T. (1992) Enzyme Microb. Technol. **14**, 258–264
- Zhang, J.-X. and Flint, H. J. (1992) Mol. Microbiol. **6**, 1013–1023
- Zhou, L., Xue, G.-P., Orpin, C. G., Black, G. W., Gilbert, H. J. and Hazlewood, G. P. (1994) Biochem. J. **297**, 359–364