Characterization of XYN10B, a modular xylanase from the ruminal protozoan *Polyplastron multivesiculatum*, with a family 22 carbohydrate-binding module that binds to cellulose

Estelle DEVILLARD*, Christel BERA-MAILLET†, Harry J. FLINT*, Karen P. SCOTT*, C. James NEWBOLD*, R. John WALLACE*, Jean-Pierre JOUANY† and Evelynne FORANO†¹

*Rowett Research Institute, Greenburn Road, Bucksburn, Aberdeen AB21 9SB, U.K., and †Unité de Microbiologie, Institut National de la Recherche Agronomique (INRA), Centre de Recherches de Clermont-Ferrand-Theix, 63122 Saint-Genès Champanelle, France

A new xylanase gene, *xyn10B*, was isolated from the ruminal protozoan *Polyplastron multivesiculatum* and the gene product was characterized. XYN10B is the first protozoan family 10 glycoside hydrolase characterized so far and is a modular enzyme comprising a family 22 carbohydrate-binding module (CBM) preceding the catalytic domain. The CBM22 was shown to be a true CBM. It showed high affinity for soluble arabinoxylan and is the first example of a CBM22 that binds strongly to celluloses of various crystallinities. The enzymic properties of XYN10B were

also analysed. Its optimal temperature and pH for activity were 39 °C and 7.0 respectively; these values being close to those of the ruminal ecosystem. The phylogenetic relationships between the XYN10B CBM22 or catalytic domain and related sequences from ruminal and non-ruminal bacteria and eukaryotes are reported. The *xyn10B* gene is shown to lack introns.

Key words: carbohydrate-binding module, glycoside hydrolase, rumen, protozoan.

INTRODUCTION

Enzymes that degrade cellulose and xylan present in plant cell walls are widely distributed among bacteria, fungi and plant species. Their distribution in the animal kingdom is thought to be much more limited, but there has been relatively little investigation of such enzymes in protozoa. The presence of xylanase genes has been demonstrated in the ruminal protozoan *Polyplastron multivesiculatum* [1,2]. Although certain ruminal protozoa were long suspected of contributing, along with anaerobic bacteria and fungi, to the breakdown of plant material in the rumen [3,4], these results provided the first clear-cut evidence that protozoal xylanolytic enzymes do exist, which are distinct from those of the bacteria they ingest.

The first xylanase described from *P. multivesiculatum*, XYNA, has a very simple organization, comprising a single catalytic module belonging to glycoside hydrolase family 11 (GH11) and an N-terminal signal peptide sequence [1]. The second *P. multivesiculatum* xylanase, XYN10B, reported in the present paper, has a more complex structure comprising a GH10 catalytic module preceded by a module that shares homology with carbohydrate-binding modules of family 22 (CBM22). CBMs have been classified into 32 families based on sequence similarities, according to the April 2003 update of the CAZy database (<http://afmb.cnrs-mrs.fr/CAZY/>). Many of these families bind only to crystalline cellulose; however, CBM22 contains members that bind to a range of single-stranded polysaccharides such as xylan and β -glucans [5–7]. These domains were originally identified as ‘thermostabilizing’ modules in GH10 xylanases from thermophilic bacteria [8,9]. Such modules have been found also in mesophilic bacteria such as *Cellulomononas fimi* [10] and *Bacillus* sp. [11], or in the rumen in

ruminococci, where they are associated with xylanases of GH11 rather than those of GH10 [12]. The results of the present study show for the first time that a module belonging to the CBM22 can bind to xylan, but also to crystalline cellulose. We also report the characterization of the enzymic properties of XYN10B, the first xylanase from GH10 to be isolated from a ruminal protozoan.

EXPERIMENTAL

Strains of *Escherichia coli* and ruminal micro-organisms

The ruminal protozoan *P. multivesiculatum* was obtained from the rumen content of a monofaunated sheep, containing this species as the sole protozoal species [1].

The plasmids isolated from the cDNA library were propagated in *E. coli* SOLR strain. The strain XL1-Blue and the strain BL21(DE3)pLys were used for standard cloning experiments and for protein expression experiments respectively. All the *E. coli* strains were grown in Luria–Bertani (LB) medium [13], containing appropriate antibiotics. Ruminal bacteria strains (*Fibrobacter succinogenes* S85, *Fibrobacter intestinalis* NR9, *Butyrivibrio fibrisolvens* A46, *Ruminococcus albus* 7, *Ruminococcus flavefaciens* 007 and *Prevotella ruminicola* 23) and fungal strains (*Piromyces* sp., *Neocallimastix frontalis*, *Neocallimastix* sp., *Caecomyces* sp., *Anaeromyces* sp. and *Orpinomyces joyonii*) were obtained and grown as described by Devillard et al. [1].

Isolation of clones from the *P. multivesiculatum* cDNA library

A cDNA library from *P. multivesiculatum* in lambda ZAPII had been previously constructed by Devillard et al. [1] and was

Abbreviations used: ANDE, affinity non-denaturing gel electrophoresis; CBM, carbohydrate-binding module; CBM22, family 22 CBM; GH, glycoside hydrolase family; IPTG, isopropyl- β -D-thiogalactoside; LB, Luria–Bertani; tXYN10B, truncated XYN10B.

¹ To whom correspondence should be addressed (e-mail forano@clermont.inra.fr).

The nucleotide sequence data reported for XYN10B has been submitted to the DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession number AJ252150.

screened for xylanolytic activity by plating in 0.7 % (w/v) soft agarose overlays, containing 0.1 % (w/v) oat spelt xylan and 2.5 mM isopropyl- β -D-thiogalactoside (IPTG). Xylan hydrolysis was detected by Congo Red staining. The cDNA inserts in xylanase-positive clones were then recovered in the form of pBluescript SK(–) phagemids (Stratagene), according to the manufacturer's instructions.

DNA sequencing and sequence analysis

The selected clone (accession number AJ252150) was sequenced full-length in both directions using an ABI Prism® BigDye® Terminator Cycle Sequencing Ready Reaction Kit (PerkinElmer Corporation, Norwalk, CT, U.S.A.) with M13 primers and internal primers. DNA sequence analyses were performed with the Internet facility site available at <http://www.infobiogen.fr>.

Protein-domain analysis was performed using ProDom (<http://prodes.toulouse.inra.fr/prodom/2002.1/html/home.php>). Multiple alignment of amino-acid sequences was obtained using the Fasta3 program (<http://www.ebi.ac.uk/fasta33>). Phylogenetic unrooted trees were constructed by first using a Psi-Blast analysis of the CBM22 carbohydrate-binding module or of the GH10 catalytic domain of XYN10B from *P. multivesiculatum* to obtain the closest protein sequences in the database. Secondly, all protein sequences were aligned using the multiple sequence alignment ClustalW 1.8 program (<http://searchlauncher.bcm.tmc.edu/multi-align>). Optimization of the alignment was performed manually before running the 3.5c PHYLIP package programs [14]: a distance matrix was generated with the Dayhoff Pam matrix using the ProtDist program following a Bootstrap analysis based on 100 replicates using the SeqBoot program. Finally, trees were obtained with the Neighbor joining analyses using Consense and Drawgram programs [14].

The two- and three-dimensional structure predictions of XYN10B were made using the PHD method [15] (<http://www.embl-heidelberg.de/predictprotein/>) and the SWISS-MODEL program [16] (<http://www.expasy.org/swissmod/SWISS-MODEL.html>) respectively.

DNA isolation and dot-blot hybridization

Plasmid-DNA isolation, transformation of *E. coli* and DNA modifications were carried out using standard techniques [13].

Genomic DNA extraction from *E. coli*, from *P. multivesiculatum* and from ruminal bacteria and fungi was performed as previously described by Devillard et al. [1].

Dot-blot hybridization conditions were identical with those described for the *xynA* gene by Devillard et al. [1]. The *xyn10B* probe was obtained by enzymic digestion of the recombinant plasmid using *Bam*HI and *Hind*III restriction endonucleases, which cut the plasmid at both ends of the insert. The insert was then recovered by electroelution from a 1 % (w/v) agarose gel [13]. The probe was radiolabelled with 32 P using random priming (Ready-to-Go; Amersham Biosciences) and used for hybridizing the membrane. The DNA spotting, hybridization and washing conditions of the membrane were carried out as described by Devillard et al. [1].

5'-end recovery

Amplification of genomic DNA fragments corresponding to the cloned cDNA was performed using the *Taq*/Pwo polymerase mixture (Expand High Fidelity PCR System; Roche Molecular Biochemicals) with the reaction buffers supplied by the

manufacturer. The PCR parameters were an initial denaturation step (94 °C for 5 min), 30 cycles of amplification (94 °C for 1 min, 52 °C for 2 min and 72 °C for 3 min), and a final elongation step (72 °C for 7 min).

Amplification of the 5' end of the gene was carried out by using the Universal Genome Walker™ kit (Clontech) according to the manufacturer's instructions. Genomic DNA (2.5 μ g) was digested separately with five different restriction endonucleases (*Pvu*II, *Dra*I, *Sca*I, *Stu*I and *Eco*RI) overnight at 37 °C. Adaptors were then ligated to the purified products of the digestion. An additional step consisting of digestion with the restriction enzyme *Bsr*GI was then performed to cut out the adaptor at the 3' end of the *xyn10B* gene. The digestion products were then used as templates for PCR amplification, which was carried out with a primer (melting temperature, 70 °C) specific for *xyn10B* coding region (purchased from Eurogentec, Saraing, Belgium) and a primer specific for the 5' end adaptor (Clontech kit). This amplification was performed using the *Tth* polymerase (Clontech), allowing a 'touchdown PCR': seven cycles of 30 s denaturation at 94 °C and 3 min annealing/elongation at 72 °C, followed by 32 cycles of 30 s denaturation at 94 °C and 3 min annealing/elongation at 70 °C. A final elongation of 4 min at 70 °C was performed.

Construction of His₆-*xyn10B* gene fusion, gene expression and preparation of protein extracts

The vector pET28a(+) (Novagen) was used for the construction and the expression of the hybrid proteins used in the present study. pET28a(+) carries six CAT codons encoding His₆.

Three different gene fusions were constructed. The first one encoded the protein XYN10B containing the CBM22, the catalytic module and the His₆ at the N-terminal. The second one encoded t1XYN10B containing the entire catalytic module and the His₆ at the C-terminal. The third one encoded t2XYN10B containing the catalytic domain lacking its first 21 amino acids that were predicted to form an α -helix and extended with His₆ at the N-terminal.

The following primers, containing restriction endonuclease recognition sites, were used for the PCR amplifications. For XYN10B, 5'-CGGGAGGTCATATGCTACAATATTGGCTG-AAGAAGA-3' and 5'-AGTGGATCCTCGAGTTATTCAACGT-TGTCAACAATAG-3'. For t1XYN10B, 5'-GTCACGGACCAT-GGGAAAAATACAAGAAAATATTTCT-3' and 5'-AGTGGAT-CCTCGAGTTCAACGTTGTCAACAATAGAAT-3'. For t2XYN10B, 5'-CGGGAGGTCATATGGGTACTGCTACTACAGTTG-CTG-3' and 5'-AGTGGATCCTCGAGTTATTCAACGTTGTC-AACAATAG-3'.

The PCR products obtained were cut with *Nde*I/*Xho*I (for XYN10B or t2XYN10B) and *Nco*I/*Xho*I (for t1XYN10B) and inserted into pET28a(+), treated with the same combinations of restriction endonucleases. These plasmids contain the His₆ codons in frame with parts of the *xyn10B* gene encoding the proteins of interest. The plasmids were then transformed into *E. coli* BL21(DE3)pLys (Stratagene) by an electroporation procedure [13].

To determine the optimal conditions of expression, different conditions of induction were tested. The strains were grown in LB medium supplemented with kanamycin (30 μ g/ml) for 4 h at 37 °C. The expression of proteins was induced by the addition of IPTG to a final concentration of 1 mM and the incubation was continued at 16 °C overnight, at 30 °C for 4 h, at 37 °C for 4 h or at 37 °C overnight. The cells were harvested (10 000 g for 20 min at 4 °C) and washed in 50 mM Na₂HPO₄, pH 6.6 before lysis. A freeze-thaw/sonication method was used for

soluble extract preparation: the cells were resuspended in 50 mM Na₂HPO₄, pH 6.6, in order to concentrate the culture 50 times, subjected to two freeze (−20 °C)–thaw [room temperature (20 °C)] cycles. Lysis was completed by sonication three times for 15 s with permanent cooling in ice (Vibra-Cell, Bioblock Scientific, Illkirch France). Cellular debris were removed by centrifugation (14 000 g for 10 min at 4 °C).

Protein analysis and enzymic assays

Determination of protein concentrations and SDS/PAGE were carried out as described previously by Devillard et al. [1].

For the Western blots, the proteins separated previously by electrophoresis were transferred on to nitrocellulose membranes (TransBlotTM; Bio-Rad), according to the manufacturer's instructions. The membrane was then blocked with a dry milk solution (10 % in 0.1 M PBS, pH 7.2) overnight at 4 °C. The membrane was incubated for 2 h in a solution of a horseradish peroxidase-conjugated antibody specific for His₆ (1/2000 dilution in a 2 % dry milk solution in 0.1 M PBS, pH 7.2) (Sigma). The membrane was washed twice for 5 min in PBS/Tween 20 [0.1 M PBS, pH 7.2 and 0.1 % (v/v) Tween 20] and once for 15 min in 0.1 M PBS. The detection of chemiluminescence (ECL[®]; Amersham Biosciences) was carried out according to the manufacturer's instructions.

The enzymic assays were carried out by incubating the protein extracts with 0.5 % (w/v) polysaccharides [insoluble oat spelts, birchwood and wood gum xylans, wheat arabinoxylan, β -glucan and medium-viscosity carboxymethylcellulose (all from Sigma), and Avicel PH101 (Fluka)] in 50 mM Na₂HPO₄, pH 7.0, at 39 °C for 30 min. The amount of enzyme extract and the incubation time were chosen to give conditions in which the measured activity was proportional to these two parameters. Reducing sugars released from polysaccharide hydrolysis were assayed as described in [1]. The optimum temperature of the xylanases was determined by incubating the protein extract at different temperatures (10–55 °C) for 30 min with 0.5 % (w/v) oat spelt xylan in 50 mM Na₂HPO₄, pH 6.6. The thermostability of the enzymes was also determined by incubating the crude cell extracts for different times at 60 °C in 30 mM Tris/HCl buffer, pH 6.6, containing 4 mM CaCl₂, before measuring their specific activities against oat spelt xylan. The optimal pH was determined by incubating the enzymic preparation in 0.5 % oat spelt xylan in 0.2 M Na₂HPO₄/0.1 M citric acid buffer (pH 4–9) at 39 °C for 30 min.

Native-affinity PAGE and binding assay

Affinity electrophoresis was used to determine binding specificity of XYN10B against different glycans following non-denaturing continuous PAGE (7.5 % polyacrylamide gels). Polysaccharides (wheat arabinoxylan, β -glucan, lichenan, laminarin and hydroxyethylcellulose) were included in the resolving gel solution at a final concentration of 0.1 % (w/v) and 0.2 % (w/v) before polymerization. Native gels without ligand were run simultaneously under the same conditions as controls. Electrophoresis was carried out in 25 mM Tris/HCl and 192 mM glycine, pH 8.1, at room temperature and at 4 °C, and 120 V. Ladder (pre-stained precision protein standards; Bio-Rad), BSA and t1XYN10B protein extracts were used as non-interacting controls of migration.

Binding assays were also carried out by incubating soluble protein extracts of XYN10B/pET/BL21 and t1XYN10B/pET/BL21 with different insoluble polysaccharides. Incubations were

performed for 2 h at 4 °C with several celluloses of various crystallinity: 2 % (w/v) *Acetobacter xylinum* cellulose, Avicel PH101, Sigmacell 101 (Sigma) or insoluble oat spelt xylan, in the presence or absence of 5 mM CaCl₂. The bacterial cellulose was obtained from *Ac. xylinum* cultures and purified as described in [17]. Insoluble xylan was obtained by washing four times in 50 mM Na₂HPO₄, pH 7.0. Incubations were then centrifuged for 5 min at 5000 g and the supernatants were analysed by Western blotting, as described above, for the presence of His₆ fusion proteins. The pellets were washed four times with 50 mM Na₂HPO₄, pH 7.0, before being analysed. Substrate-bound enzymes were associated with the pellet fractions, whereas unbound proteins were found in the supernatant.

Determination of cellulose crystallinity by NMR

Solid-state Cross-Polarization/Magic-Angle Spinning ('CP/MAS') ¹³C NMR spectra of cellulose were obtained on a 300 MHz Avance DSX spectrometer (Bruker BioSpin, Karlsruhe, Germany). Measurements were performed at 25 °C and using the conditions described in [18]. The crystallinity of the cellulose was assessed by measuring the ratio of the integrals of the peaks at 88 p.p.m. and 83 p.p.m., corresponding to the signals of the C4 of the glucose in the crystalline and non-crystalline regions of cellulose respectively [18]. The *Ac. xylinum* cellulose spectrum did not show any quantifiable peak at 83 ppm (non-crystalline), and was thus classified as highly crystalline. Avicel PH101 gave a crystalline to non-crystalline ratio of 1.4, whereas Sigmacell 101 cellulose with a ratio of 0.8 was more amorphous.

RESULTS

Cloning and sequence analysis of the *xyn10B* gene: proof of the gene origin

A cDNA library from the ruminal protozoan *P. multivesiculatum*, was previously constructed in lambda-ZAPII, and was screened for clones displaying xylanase activity [1]. Several positive clones were isolated and were analysed further, following *in vivo* excision. The study of one of these clones, harbouring a 1.8 kb insert and encoding the product referred to as XYN10B, is reported here.

Sequencing of the XYN10B-encoding insert (accession number AJ252150) revealed an 18 base polyadenylated region at the 3' end, preceded by a characteristic eukaryotic polyadenylation signal, AATAAA, at bases 1701–1706.

The extreme 5' end of the *xyn10B* gene, including the initiation codon ATG, was not present in the initial clone that was isolated from the library. Genome walking allowed amplification of 60 extra nucleotides at the 5' end of the known sequence. This sequence included an initiation codon (ATG) at the 5' end of a sequence encoding a typical signal peptide and this completed the sequence of the *xyn10B* gene. The AT content of the 5' untranslated region of *xyn10B* is 94.1 %, considerably higher than that of the coding region of the gene, 68.5 %. The complete *xyn10B* gene sequence is 1757 bases with an open reading frame that starts with an ATG initiation codon at base 60 and ends at base 1659 with a TAA stop codon.

In the *xyn10B*-coding region, the codon usage is heavily biased, as observed previously for protozoal genes, including the *xynA* gene from *P. multivesiculatum* previously studied in [1]. The codon usage is quite similar between the two *P. multivesiculatum* xylanase genes, except that ten codons are not used at all for *xyn10B* translation, whereas 19 codons were unused for *xynA* [1].

The alanine codon used preferentially in the *xyn10B* coding region is GCT, whereas it is not represented in *xynA*. Similarly, the cysteine amino acid is exclusively translated from the TGC codon for *xynA*, but from two codons for *xyn10B*. The codon usage for *xyn10B* was also compared with that of five ruminococcal xylanases and of six ruminal fungal xylanases (results not shown): for example, the codon CAT accounted for 100 % of histidine codons in *xyn10B* of *P. multivesiculatum* and for 67 % of histidine codons of *xynA* [1], whereas it is used in only 43 % and 48 % of histidine codons in xylanase genes from ruminal ruminococci and fungi respectively. The bias is very pronounced in favour of an A or a T for the first and the third base of the codon, which is consistent with reports for other ruminal protozoal genes [1,19], and presumably reflects the high AT content of the genome.

As *xyn10B* is one of the first genes from a ruminal protozoan to be studied, the presence of possible introns was assessed. This was carried out by sequencing, after PCR amplification, genomic DNA corresponding to the known cDNA sequence. Comparison of the genomic sequence with the cDNA sequence of *xyn10B* showed that this gene does not contain any intron sequence. The presence of a polyadenylation signal and a polyadenylated tail in the *xyn10B* cDNA is as expected for eukaryotic cDNA. Furthermore, the codon usage is biased in a manner typical of ciliate codon usage [19]. However, to substantiate further the protozoal origin of the gene, part of *xyn10B* was probed against DNA extracted from the major xylanolytic ruminal bacterial and fungal strains (see the Experimental section). The probe hybridized only with DNA extracted from *P. multivesiculatum*, whereas an oligonucleotide targeting 16 S- and 18 S-like rRNA genes hybridized with all the DNA samples (results not shown). This effectively rules out the possibility that *xyn10B* could have been derived from ruminal bacteria or fungi.

Gene product analysis

The *xyn10B* gene encodes a 533-amino-acid protein, with a predicted molecular mass of 60.3 kDa. The protein produced by *E. coli* shows xylanase activity and has an apparent molecular mass of about 60 kDa. A typical signal sequence is present at the N-terminus of XYN10B, with a predicted cleavage site between amino acids 20 and 21. The predicted molecular mass of the mature protein is thus 58.1 kDa. The catalytic domain (amino acids 180–533) was found by homology searches to be a GH10 xylanase catalytic module.

Like most of the glycoside hydrolases of microbial origin, XYN10B displays a modular structure and carries an N-terminal domain (amino acids 21–179) that belongs to CBM22. The two modules of XYN10B were not separated by a linker sequence, contrary to what is often found in modular enzymes. The percentage identity between the CBM22 of XYN10B from *P. multivesiculatum* and sequences available in databases is not very high. Maximal identity (approx. 30 %) was found with XYN1 from *Ruminococcus* sp. (TrEMBL accession number Q59790), or sequences from thermophilic micro-organisms, such as XYNA from *Thermoanaerobacterium saccharolyticum* (P36917) or XYN10B from *Clostridium thermocellum* (P51584). Identity with eukaryotic *Arabidopsis thaliana* putative CBM22 is about 24 %.

Enzymic properties of XYN10B

Several derivatives of XYN10B were constructed (Figure 1) in order to study the role of the CBM22 module, and the properties of these derivatives were compared. The *xyn10B* gene was subcloned

Figure 1 Domain organization of full and truncated XYN10B

Table 1 Xylan and glycan substrate specificity of *E. coli*/XYN10B proteic extracts

Reducing sugars are given as equivalent nanomoles of xylose for xylan substrates, and as equivalent nanomoles of glucose for glycan substrates.

Substrate	Specific activity (nmol · min ⁻¹ · mg ⁻¹)	Activity relative to wheat arabinoxylan (%)
Wheat arabinoxylan	1108 ± 99	100
Oat spelt xylan	801 ± 54	72
Birchwood xylan	740 ± 32	67
Wood gum xylan	566 ± 63	51
Barley β-glucan	< 10 ⁻²	< 10 ⁻⁵
Carboxymethylcellulose	< 10 ⁻²	< 10 ⁻⁵
Avicel PH101	< 10 ⁻²	< 10 ⁻⁵

and overexpressed in *E. coli* BL21(DE3)pLys using a pET28 vector, and the different proteins were produced as His₆ fusion proteins. Different attempts were made to purify the XYN10B derivatives, but only low amounts of pure protein were obtained, thus soluble proteic extracts of recombinant *E. coli* were used for enzymic assays and Western blot experiments.

The first derivative corresponded to His₆-tagged mature XYN10B, containing both the CBM22 and catalytic modules (Figure 1), and lacking the signal peptide. Two truncated enzymes containing only the catalytic domain were constructed and referred to as t1XYN10B and t2XYN10B. t1XYN10B contained the entire catalytic domain, whereas the first 21 amino acids of the catalytic domain, which are predicted to form an α-helix according to the PHD secondary structure prediction program [15] and the SWISS-MODEL program [16], were deleted from t2XYN10B (Figure 1).

XYN10B appeared to be specific for xylan degradation, the highest activity being found on soluble arabinoxylans (Table 1). The truncated enzymes displayed the same substrate specificity (results not shown). The activity of XYN10B was not increased in the presence of Ca²⁺ ions (results not shown).

The optimal temperature for XYN10B activity was found to be 39 °C (Figure 2A). This optimal temperature was conserved after truncation in both t1XYN10B and t2XYN10B (Figure 2A). In order to assess the thermostability of XYN10B and tXYN10B, recombinant *E. coli* soluble extracts were incubated for different times at 60 °C, before measuring their specific activity. XYN10B and its two derivatives proved not to be thermostable, as no more xylanase activity was detected after a 5 min treatment at 60 °C (results not shown). The same results were obtained in the presence of 5 mM CaCl₂. Thus the CBM22 of XYN10B does not confer any thermostability on the enzyme.

Figure 2 Enzymic properties of XYN10B, t1XYN10B and t2XYN10B

(A) Temperature effect on xylanase activity of XYN10B (◆) and t1XYN10B or t2XYN10B (□) in 50 mM Na_2HPO_4 buffer, pH 6.6. (B) pH effect on xylanase activity of XYN10B (◆), t1XYN10B (△) and t2XYN10B (○) in Na_2HPO_4 /citric acid buffer at varying pH, and at 39 °C. Enzyme activity was measured as μg of reducing sugars released in 30 min by xylan hydrolysis per μg of soluble proteins. Relative activity is expressed as a percentage of maximum activity $\pm$ S.D.

The effects of pH variation on the activity of XYN10B and tXYN10B were also studied. XYN10B remains active over a very large pH range of 4.5–8.7 (Figure 2B). Deletion of the CBM22 with conservation of the entire catalytic domain of XYN10B (t1XYN10B) has no effect on the activity at varying pH (Figure 2B). More surprising was the difference in optimal pH between XYN10B or t1XYN10B and t2XYN10B. The t2XYN10B protein is active in a more basic pH range (optimal pH 8.4) compared

with the non-truncated enzyme (optimal pH 7.0; Figure 2B). The deletion of the first α -helix of the catalytic domain of XYN10B induced a shift of optimal pH, probably due to an alteration of the three-dimensional folding of the xylanase.

Carbohydrate binding

The capacity of XYN10B to bind different polysaccharides was studied first in 2 h incubations of the enzyme extract with insoluble substrates [2% (w/v) oat spelt xylan, bacterial, Avicel and Sigmacell celluloses] and Western blot analysis of the substrate-bound and unbound proteins after centrifugation. Preparations of t1XYN10B were treated similarly as a control. The effect of Ca^{2+} on the affinity of XYN10B for these substrates was also tested, since these ions have been shown to increase the binding of a CBM from family 4 [20], which belongs to the same superfamily as CBM22 [21]. The anti-His₆ antibodies reacted with all the xylan and cellulose pellets, indicating that the CBM22 of XYN10B was able to bind to these polysaccharides (Figure 3), including celluloses of different crystallinity (see the Experimental section).

There was no evidence for any absolute requirement for Ca^{2+} in substrate binding, since binding occurred in the presence of the bivalent cation-chelating agent EDTA. Slightly more unbound XYN10B was recovered after binding to Sigmacell and oat spelt xylan in the presence of EDTA than in the presence of Ca^{2+} however, suggesting a slight stimulation by Ca^{2+} of binding to these substrates. No binding was obtained with t1XYN10B, indicating that the binding of XYN10B was specific.

To assess binding to soluble polysaccharides, affinity non-denaturing gel electrophoresis (ANDE) of XYN10B was performed. ANDE is based on the retarded mobility of a protein containing polysaccharide-binding domains during native PAGE if a binding substrate is included in the gel. Protein extracts from *E. coli* expressing XYN10B and t1XYN10B were separated by native PAGE containing different soluble polysaccharides (arabinoxylan, β -glucan, lichenan, laminarin and hydroxyethylcellulose) at a final concentration of 0.2% (w/v). As a reference, molecular-mass standards and BSA were also separated on the same gels. After electrophoresis, the proteins were transferred to a membrane and the His₆ tag was revealed by Western blotting. As shown in Figure 4, XYN10B interacts significantly with soluble arabinoxylan included in the gel,

Figure 3 Western blot analysis of binding of XYN10B-His₆ and t1XYN10B-His₆ to cellulose and xylan

C, control: Western blot of total protein extracts of *E. coli* producing XYN10B or t1XYN10B (not incubated); A, XYN10B-His₆ or t1XYN10B-His₆ extracts incubated with 2% (w/v) Avicel PH101; Ax, XYN10B-His₆ extracts incubated with 2% (w/v) *Ac. xylinum* cellulose; S, XYN10B-His₆ or t1XYN10B-His₆ extracts incubated with 2% (w/v) Sigmacell 101; X, XYN10B-His₆ or t1XYN10B-His₆ extracts incubated with 2% (w/v) insoluble oat spelt xylan. *E. coli* XYN10B-His₆ or t1XYN10B-His₆ soluble extracts were incubated with the appropriate substrate in the presence of 5 mM CaCl_2 or 10 mM EDTA, before centrifugation, and pellets (1) or supernatants (2) were separated by SDS/PAGE (10% gels), blotted and analysed with anti-His₆ antibodies (at 1/2000 dilution).

Figure 4 Native-affinity PAGE analysis of XYN10B and t1XYN10B

Protein extracts of *E. coli* producing XYN10B or t1XYN10B were separated in non-denaturing 7.5% polyacrylamide gels containing 0.2% (w/v) of wheat arabinoxylan (+). A gel without added polysaccharide (–), BSA and pre-stained precision protein standards (Bio-Rad) were used as non-interacting controls of migration (see the Experimental section).

as the xylanase migration was retarded by approx. 40 kDa, whereas t1XYN10B migration was unchanged. Similarly, XYN10B was found to interact with β -glucan and lichenan, although more weakly, as the enzyme migration was only retarded by approx. 20 kDa in the gel under these conditions (results not shown). No retardation of XYN10B was observed with laminarin and hydroxyethylcellulose substrates, as for the controls (results not shown). It should be noted that XYN10B and its truncated derivatives showed an apparent molecular mass 15 kDa higher under native conditions compared with denaturing conditions of electrophoresis.

Phylogeny

A phylogenetic tree was constructed for CBM22, which included the sequence of the module of XYN10B in addition to the sequences of 32 domains from 28 different glycoside hydrolases. Half of these sequences are from thermophilic micro-organisms, and 15 sequences are from mesophilic micro-organisms (*P. multivesiculatum*, *Ruminococcus* sp., *Ce. fimi*, *Cellulomonas pachnodae*, *Bacillus* sp., *Aeromonas caviae* and *Ar. thaliana*) (Figure 5A). The CBM22 of XYN10B appears most closely related to those found in the *Ar. thaliana* eukaryotic sequences in agreement with a recent study by Xie et al. [22]. The most closely related cluster includes sequences of CBM22 found in xylanases from ruminococci and in the xylanase XYN10B (XYNY) from *Cl. thermocellum*. The module of XYN10B from *P. multivesiculatum* is phylogenetically more distant from other bacterial sequences, such as those from *Caldicellulosiruptor* sp. or *Thermotoga* spp.

A second phylogenetic tree was constructed for the catalytic domain (GH10) of XYN10B from *P. multivesiculatum* (Figure 5B). This tree includes 32 sequences of GH10 xylanase domains isolated from either mesophilic or thermophilic micro-organisms. The sequence of the catalytic domain of XYN10B is most homologous with the catalytic domains of *R. flavefaciens* XYNA and *Cl. thermocellum* XYN10B (XYNY). This tree also shows that the other eukaryotic GH10 xylanase sequences are grouped into one cluster that also includes two prokaryotic sequences, XYN10A (XYNX) from *Ce. fimi* and XYN10L (XYND) from *F. succinogenes*. The sequence of the catalytic domain of XYN10B from *P. multivesiculatum* is thus divergent from all the other eukaryotic fungal or plant sequences analysed here (Figure 5B).

DISCUSSION

In the present paper, we report the first complete coding sequence for a glycoside hydrolase gene from a ruminal protozoan. Together with previous evidence from incomplete sequences encoding GH11 xylanases [1,2], this establishes that protozoa encode their own plant-cell-wall-degrading enzymes that are distinct from those of any ingested bacteria. XYN10B is the first example of a GH10 xylanase from a ruminal protozoan. The *xyn10B* gene shows a very high AT content in coding and non-coding regions, typical of ruminal protozoal genes [1,23], and also a heavily biased codon usage as for the other *P. multivesiculatum* xylanase gene [1]. The gene was shown not to contain any introns, as is generally reported for ciliate genes. In other primitive eukaryotes, like the ruminal fungi, the glycoside hydrolase genes also lack introns [24], whereas introns are common in genes coding for such enzymes in aerobic fungi [25]. *Xyn10B* encodes a modular xylanase with a GH10 catalytic module. The xylanase XYNA, previously described for *P. multivesiculatum*, is a GH11 xylanase [1]. We have now established that, like the ruminal bacteria and fungi, this ruminal protozoan produces the two main families of xylanase enzymes.

XYN10B also bears a CBM22. Such CBM22s have been reported in more than 50 sequences (CAZy, update April 2003); most of them are associated with a GH10 xylanase catalytic domain. Among these sequences, few are from mesophilic organisms (*Ce. fimi*, *Ce. pachnodae*, *Bacillus* sp., *Ae. caviae* and *Ar. thaliana*). Six sequences, all from ruminococcal species isolated from the rumen, show an association between the GH11 xylanase catalytic domain and a CBM22 module. Thus XYN10B from *P. multivesiculatum* is unusual compared with the other xylanases from ruminal micro-organisms.

In addition to a thermostabilizing role previously suggested, a new function has been more recently assigned to these CBM22 in enzymes from thermophilic micro-organisms. For XYN10B from *Cl. thermocellum*, XYNA from *Thermotoga maritima* and XYNA from *Caldibacillus cellulovorans* [5–7], it has been established that the module was involved in binding to substrates, such as xylan and β -glucans. The three dimensional structure of this domain has been determined for *Cl. thermocellum* XYN10B [7] and a phylogeny study of CBM22 was recently carried out [22]; it showed that five amino-acid residues have a critical role in substrate binding. They include three aromatic amino acids and two polar ones. These amino acids are present in one of the CBM22 of XYN10B from *Cl. thermocellum*, which binds to the substrate, and they are absent from the second one, which does not bind to the substrate. These three aromatic amino acids are conserved in the *P. multivesiculatum* XYN10B sequence (Trp⁸¹, Tyr¹²⁷ and Tyr¹⁵⁹) as well as the polar amino acids (Arg⁵⁰ and Glu¹⁶¹) and probably have a major role in the xylan-binding capability of XYN10B evidenced here. Furthermore, we also show in the present study that the CBM22 of XYN10B binds to cellulose, and it is the first time that a CBM from this family has been shown to bind to this substrate. Indeed, CBM22s have been shown to bind to soluble polysaccharides such as xylan or soluble β -1,4-linked polymers of glucose, including hydroxyethylcellulose and β -1,3-1,4-glucans (β -glucans, lichenan). The CBM of XYN10B demonstrated a good affinity for xylan, a weak affinity for mixed-linkage β -1,3-1,4-glucans and no affinity for β -1,3-glucans and hydroxyethylated cellulose, as shown by ANDE. In this last case, the ethanolic groups on the glucose polymer may prevent the CBM22 binding. Fontes et al. [8] have shown the CBM22 of XYN10B from *Cl. thermocellum* to have 100 times less affinity for β -glucan than for xylan. The differences in ligand specificity

Figure 5 Phylogenetic trees for CBM22s (A) and GH10 xylanase catalytic domains (B)

Phylogenetic relationships are based on ClustalW alignments, optimized manually, and analysed with the PHYLIP program. A total of 100 bootstrap replicates were performed, and the bootstrap values greater than 60 are indicated above each branch. Protein sequences are named by their species, gene names and SwissProt, TrEMBL or GenPept accession numbers (or ID codes).

within a single CBM family has been attributed to changes in the orientation of the surface aromatic residues [26]. In addition, in the present study, we showed for the first time that the CBM22 of XYN10B was able to bind to insoluble cellulose, and even to the more crystalline Avicel type of commercial cellulose as well as to the *Ac. xylinum* bacterial 100% crystalline cellulose. We also proved the specificity of the binding, as the truncated form of the enzyme, devoid of its CBM22, did not bind to this substrate. Some CBMs were shown to bind only to soluble forms of cellulose [27], whereas others bind to both soluble and non-soluble forms of the polysaccharide [28]. The mechanism of binding to either form is suggested to be dependent on the binding-site conformation (planar surface or cleft) [22,26]. Indeed, the CBM22 binding site was shown to be a shallow cleft probably interacting only with a single chain of cellulose and was thus thought not to bind to crystalline cellulose [7], in contrast with the GH10 CBM-binding site which interacts with the multiple planar chains of crystalline cellulose [28]. Determination of the tertiary structure of the CBM22 of *P. multivesiculatum* XYN10B, as well as detailed binding studies, would probably help our understanding of the crystalline cellulose binding ability of this module.

Finally, our results could reflect a slight effect of calcium on the binding of the CBM of XYN10B to xylan and Sigmacell. Abou-Hachem et al. [29] recently studied the effect of calcium on the binding of the CBM4-2 of a xylanase from *Rhodothermus marinus* [29], and identified the two calcium-binding sites of high and moderate affinity on the protein. By comparison of the structures of CBM4-2 and CBM22 [29], we found five amino acids, Thr³⁷, Glu⁴⁰, Lys⁶⁶, Ser⁷⁰ and Asp¹⁷², in XYN10B-CBM22 aligned with or very close to the amino acids identified in the XYN10B-CBM from *Cl. thermocellum*, which could be involved in calcium binding.

Deletion of its CBM22 has no effect on the enzymic properties of XYN10B, as previously shown for other xylanases containing a CBM [30]. Physico-chemical properties and substrate specificity are identical for XYN10B with or without (t1XYN10B) the CBM22, as is also the case for another GH10 xylanase associated with a CBM22 [8]. It should be pointed out that the second derivative of XYN10B, t2XYN10B, that lacks the first α -helix of the GH10 catalytic module displays an important shift of optimal pH for activity. This might be due to structural changes in the protein induced by this deletion. Indeed, it has been shown that the optimal pH of an enzyme might be changed by a single mutation of an amino acid outside the active site, as in the case of a GH11 xylanase of *Bacillus circulans* [31]. This result emphasises how careful one must be when deleting modules from modular enzymes, as deletion of a few amino acids in the catalytic module may induce quite severe changes in the enzyme activity. At the same time, this observation also illustrates the potential for protein engineering to manipulate the optimum pH. The dramatic effect of deleting the first 21 amino acids of the XYN10B GH10 catalytic domain upon optimal pH might indeed be explained through effects of conformational changes on the pIs of the active site glutamic residues as for the *B. circulans* GH11 [31].

Our results show that polysaccharide binding, and not stabilization, is the main function of the CBM22 in XYN10B, as is probably the case for the other mesophilic enzymes as well as for the thermophilic ones [6].

Ruminal protozoa are known to be predatory towards bacteria and fungi as well as ingesting plant material [32]. This raises the interesting possibility that certain protozoa might have acquired genes for cellulases and xylanases by horizontal gene transfer events that occurred in the rumen. However, the only other instances of CBM22s occurring in enzymes from ruminal micro-organisms are in Gram-positive bacteria belonging to

the genus *Ruminococcus*. Interestingly, in *Ruminococcus*, these domains occur immediately following a GH11, rather than a GH10, catalytic domain. This arrangement has not so far been reported for any group and is quite different from that found in *P. multivesiculatum* XYN10B. In addition, phylogenetic comparisons of *P. multivesiculatum* XYN10B do not clearly support such horizontal transfer within the rumen, as the closest related sequences are, for the catalytic domain, GH10 xylanases from ruminal bacteria, but, for the CBM22, sequences from the plant *Ar. thaliana*. In general, as has been pointed out previously [33], phylogenetic analyses appear consistent with a major role for horizontal transfer in the evolution of xylanolytic enzymes. Many of the key transfer events may have been very ancient, however. Thus, if we assume that plant cell wall degrading genes were acquired by certain groups of protozoa from bacterial or other sources, some of these transfer events may have occurred in the rumen, but equally some may have predated the evolution of ruminants.

We are grateful for joint support from Institut National de la Recherche Agronomique (INRA; France) and from the Scottish Office Agriculture, Environment and Fisheries Department (SOAEFD; U.K.) for this project. E. D. was the recipient of a joint INRA–RRI (Rowett Research Institute) fellowship and of a grant 'cotutelle de thèse' from the French Ministère de l'Éducation Nationale, de la Recherche et de la Technologie (MENRT). We are also very grateful to Maria Matulova (Laboratoire de Synthèse et Étude de Systèmes à Intérêt Biologique, UMR 6504 CNRS, Aubière, France) for performing the NMR experiments and to Joëlle Henry-Berger (Laboratoire de Physiologie Comparée et Endocrinologie Moléculaire, UMR 6547 CNRS, Aubière, France) for her help with the PHYLIP phylogeny program.

REFERENCES

- Devillard, E., Newbold, C. J., Scott, K. P., Forano, E., Wallace, R. J., Jouany, J. P. and Flint, H. J. (1999) A xylanase produced by the rumen anaerobic protozoan *Polyplastron multivesiculatum* shows close sequence similarity to family 11 xylanases from Gram-positive bacteria. *FEMS Microbiol. Lett.* **181**, 145–152.
- Takenaka, A., Goto, M., Kudo, H., Itabashi, H. and Cheng, K. J. (1999) The properties of fibrolytic enzymes from the rumen ciliate protozoa. In *Genetics, Biochemistry and Ecology of Cellulose Degradation* (Ohmiya, K., Hayashi, K., Sakka, K., Kobayashi, Y., Karita, S. and Kimura, T., eds.), pp. 551–557, Uni Publishers Co. Ltd, Tokyo, Japan.
- Williams, A. G. and Coleman, G. S. (1985) Hemicellulose degrading enzymes in rumen ciliate protozoa. *Curr. Microbiol.* **12**, 85–90.
- Coleman, G. S. (1985) The cellulase content of 15 species of entodiniomorphid protozoa, mixed bacteria and plant debris isolated from the ovine rumen. *J. Agric. Sci.* **104**, 349–360.
- Sunna, A., Gibbs, M. D. and Bergquist, P. L. (2000) The thermostabilising domain, XynA, of *Caldibacillus cellulosovorans* xylanase is a xylan binding domain. *Biochem. J.* **346**, 583–586.
- Meissner, K., Wassenberg, D. and Liebl, W. (2000) The thermostabilizing domain of the modular xylanase XynA of *Thermotoga maritima* represents a novel type of binding domain with affinity for soluble xylan and mixed-linkage β -1,3-/ β -1,4-glucan. *Mol. Microbiol.* **36**, 898–912.
- Charnock, S. J., Bolam, D. N., Turkenburg, J. P., Gilbert, H. J., Ferreira, L. M. A., Davies, G. J. and Fontes, C. M. G. A. (2000) The X6 "thermostabilizing" domains of xylanase are carbohydrate-binding modules: structure and biochemistry of the *Clostridium thermocellum* X6b domain. *Biochemistry* **39**, 5013–5021.
- Fontes, C. M. G. A., Hazlewood, G. P., Morag, E., Hall, J., Hirst, B. H. and Gilbert, H. J. (1995) Evidence for a general role for non-catalytic thermostabilizing domains in xylanases from thermophilic bacteria. *J. Bacteriol.* **307**, 151–158.
- Winterhalter, C., Heinrich, P., Candussio, A., Wich, G. and Liebl, W. (1995) Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the thermophilic bacterium *Thermotoga maritima*. *Mol. Microbiol.* **15**, 431–444.
- Clarke, J. H., Davidson, K., Gilbert, H. J., Fontes, C. M. G. A. and Hazlewood, G. P. (1996) A modular xylanase from mesophilic *Ce. fimi* contains the same cellulose-binding and thermostabilizing domains as xylanases from thermophilic bacteria. *FEMS Microbiol. Lett.* **139**, 27–35.

- 11 Blanco, A., Diaz, P., Zueco, J., Parascandola, P. and Pastor F. I. J. (1999) A multidomain xylanase from a *Bacillus* sp. with a region homologous to thermostabilizing domains of thermophilic enzymes. *Microbiology* **145**, 2163–2170
- 12 Zhang, J. X., Martin, J. and Flint, H. J. (1994) Identification of non-catalytic conserved regions in xylanases encoded by the xynB and xynD genes of the cellulolytic rumen anaerobe *Ruminococcus flavefaciens*. *Mol. Gen. Genet.* **245**, 260–264
- 13 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
- 14 Felsenstein, J. (1989) PHYLIP – Phylogeny Inference Package (version 3.2). *Cladistics* **5**, 164–166
- 15 Rost, B. and Sander, C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins* **19**, 55–72
- 16 Guex, N., Diemand, A. and Peitsch, M. C. (1999) Protein modelling for all. *Trends Biochem. Sci.* **24**, 364–367
- 17 Bibollet, X., Bosc, N., Matulova, M., Delort, A.-M., Gaudet, G. and Forano, E. (2000) ^{13}C and ^1H NMR study of cellulose metabolism by *Fibrobacter succinogenes* S85. *J. Biotechnol.* **77**, 37–47
- 18 Larsson, P. T., Wickholm, K. and Iversen, T. (1997) A CP/MAS ^{13}C NMR investigation of molecular ordering in celluloses. *Carbohydrate Res.* **302**, 19–25
- 19 Eschenlauer, S. C. P., McEwan, N. R., Calza, R. E., Wallace, R. J., Onodera, R. and Newbold, C. J. (1998) Phylogenetic position and codon usage of two centrin genes from the rumen ciliate protozoan, *Entodinium caudatum*. *FEMS Microbiol. Lett.* **166**, 147–154
- 20 Simpson, P., Jamieson, S., Abou-Hachem, M., Nordberg Karlsson, E., Gilbert, H. J., Holst, O. and Williamson, M. P. (2002) The solution structure of the CBM4-2 carbohydrate binding module from a thermostable *Rhodothermus marinus* xylanase. *Biochemistry* **41**, 5712–5719
- 21 Sunna, A., Gibbs, M. D. and Bergquist, P. L. (2001) Identification of novel β -mannan- and β -glucan-binding modules: evidence for a superfamily of carbohydrate-binding modules. *Biochem. J.* **356**, 791–798
- 22 Xie, H., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J., Raghothama, S., Fontes, C. M. G. A., Dias, F. M. V., Ferreira, L. M. A. and Bolam, D. N. (2001) *Clostridium thermocellum* Xyn10B carbohydrate-binding module 22-2: the role of conserved amino acids in ligand binding. *Biochemistry* **40**, 9167–9176
- 23 McEwan, N. R., Gatherer, D., Eschenlauer, S. C. P., McIntosh, F. M., Calza, R. E., Wallace, R. J. and Newbold, C. J. (2000) An unusual codon usage pattern in the ciliate family Ophryoscolecidae and its applications for determining the source of cloned DNA. *Anaerobe* **6**, 21–28
- 24 Chen, H., Li, X. L., Blum, D. L. and Ljungdahl, L. G. (1998) Two genes of the anaerobic fungus *Orpinomyces* sp. strain PC-2 encoding cellulases with endoglucanase activities may have arisen by gene duplication. *FEMS Microbiol. Lett.* **159**, 63–68
- 25 Knowles, J., Lehtovaara, P., Penttilä, M. and Teeri, T. (1987) Cellulase families and their genes. *Trends Biotechnol.* **5**, 255–261
- 26 Johnson, P. E., Brun, E., MacKenzie, L. F., Withers, S. G. and McIntosh, L. P. (1999) The cellulose binding domains from *Cellulomonas fimi* β -1,4-glucanase CenC bind nitroxide spin-labeled cellobiosaccharides in multiple orientations. *J. Mol. Biol.* **287**, 609–625
- 27 Gill, J., Rixon, J. E., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Hazlewood, G. P. and Gilbert, H. J. (1999) The type II and X cellulose-binding domains of *Pseudomonas* xylanase A potentiate catalytic activity against complex substrates by a common mechanism. *Biochem. J.* **342**, 473–480
- 28 Raghothama, S., Simpson, P. J., Szabo, L., Nagy, T., Gilbert, H. J. and Williamson, M. P. (2000) Solution structure of the CBM10 cellulose binding module from *Pseudomonas* xylanase A. *Biochemistry* **39**, 978–984
- 29 Abou-Hachem, M., Karlsson, E. N., Simpson, P., Linse, S., Sellers, P., Williamson, M. P., Jamieson, S., Gilbert, H. J., Bolam, D. N. and Holst, O. (2002) Calcium binding and thermostability of carbohydrate binding module CBM4-2 of Xyn10A from *Rhodothermus marinus*. *Biochemistry* **41**, 5720–5729
- 30 Sakka, K., Takada, G., Karita, S. Ohmiya, K. (1996) Identification and characterization of cellulose-binding domains in xylanase A of *Clostridium stercorarium*. *Ann. N.Y. Acad. Sci.* **782**, 241–251
- 31 Joshi, M. D., Sidhu, G., Pot, I., Brayer, G. D., Withers, S. G. and McIntosh, L. P. (2000) Hydrogen bonding and catalysis: a novel explanation for how a single amino acid substitution can change the pH optimum of a glycosidase. *J. Mol. Biol.* **299**, 255–279
- 32 Williams, A. G. and Coleman, G. S. (1997) The rumen protozoa. In *The Rumen Microbial Ecosystem* (Hobson, P. N. and Stewart, C. S., eds.), pp. 73–139, Blackie Academic and Professional, London
- 33 Garcia-Vallvé, S., Romeu, A. and Palau, J. (2000) Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. *Mol. Biol. Evol.* **17**, 352–361

Received 15 November 2002/4 April 2003; accepted 14 April 2003

Published as BJ Immediate Publication 14 April 2003, DOI 10.1042/BJ20021784