# Molecular cloning and expression in *Escherichia coli* of a *Trichoderma viride* endo-$\beta$-(1 → 6)-galactanase gene

Toshihisa KOTAKE*[1], Satoshi KANEKO†, Aya KUBOMOTO*, Md. Ashraful HAQUE*, Hideyuki KOBAYASHI† and Yoichi TSUMURAYA*

*Department of Biochemistry and Molecular Biology, Faculty of Science, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama 338-8570, Japan, and †Biological Function Division, National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan

A gene encoding endo-$\beta$-(1 → 6)-galactanase from *Trichoderma viride* was cloned by reverse transcriptase–PCR and expressed in *Escherichia coli*. The gene contained an open reading frame consisting of 1437 bp (479 amino acids). The deduced amino acid sequence of the protein showed little similarity with other known glycoside hydrolases. A signal sequence (20 amino acids) was found at the N-terminal region of the protein and the molecular mass of the mature form was calculated to be 50.488 kDa. The gene product expressed in *E. coli* as a recombinant protein fused with thioredoxin and $His_6$ tags had almost the same substrate specificity and mode of action as native enzyme purified from a commercial cellulase preparation of *T. viride*, i.e. recombinant enzyme endo-hydrolysed $\beta$-(1 → 6)-galacto-oligomers with a DP (degree of polymerization) higher than 3, and it could also hydrolyse $\alpha$-L-arabinofuranosidase-treated arabinogalactan protein from radish. It produced $\beta$-(1 → 6)-galacto-oligomers ranging from DP 2 to at least 8 at the initial hydrolysis stage and galactose and $\beta$-(1 → 6)-galactobiose as the major products at the final reaction stage. These results indicate that the cloned gene encodes an endo-$\beta$-(1 → 6)-galactanase. As far as we know, this is the first time an endo-$\beta$-(1 → 6)-galactanase has been cloned.

Key words: arabinogalactan protein, endo-$\beta$-(1 → 6)-galactanase, family 5 glycoside hydrolase, *Trichoderma viride*.

## INTRODUCTION

AGPs (arabinogalactan proteins) are a family of complex proteoglycans found in all tissues of higher plants and localized in cell walls, plasma membranes and the extracellular matrix. Although the functions of AGPs have not been identified clearly, several lines of evidence indicate that they are involved in many physiological events such as cell division, cell expansion and cell death [1,2]. For example, the glycosylphosphatidylinositol anchor signals found in genes for core proteins in *Arabidopsis* AGP suggest that AGPs play important roles in cell adhesion, morphogenesis and signal transduction as do mammalian glycosylphosphatidylinositol-anchored proteoglycans [3,4]. As an example of an AGP involved in a particular event, one may adduce tobacco stylar-transmitting tissue proteins (belonging to the AGP family), which have been reported to attract pollen tubes and stimulate their elongation [5]. The *Arabidopsis SOS5* gene, which encodes a protein with AGP- and fasciclin-like domains, is known to be essential for normal cell expansion, suggesting a role in cell-to-cell adhesion [6].

AGPs are characterized by large amounts of carbohydrate components rich in galactose (sugars in the present study are D series unless designated otherwise) and L-arabinose, and protein components (core proteins: generally $< 10\,\%$ of total weight) rich in hydroxyproline, serine, threonine, alanine and glycine. The reducing galactosyl residues in the carbohydrate chains are attached through O-glycosyl linkages to hydroxyproline and/or serine/threonine residues [1,2]. Since a large number of putative protein cores exist [7], it is difficult to separate particular AGP molecules from other AGP species in plant tissues, which make it difficult to elucidate the precise characteristics of individual AGPs. The carbohydrate moieties of AGPs have a common structure consisting of $\beta$-(1 → 3)-galactosyl backbones to which side chains of $\beta$-(1 → 6)-galactosyl residues are attached through O-6. L-Arabinose and lesser amounts of other auxiliary sugars such as glucuronic acid, 4-*O*-methyl-glucuronic acid, L-rhamnose and L-fucose are attached to the side chains, usually at non-reducing terminals [1,2,8].

In spite of the significant physiological interest in AGPs, until now not much research has been done on glycoside hydrolases, which cleave the sugar moiety of AGPs. We believe that it is important to study carbohydrate-degrading enzymes of AGPs because hydrolytic enzymes specific to particular sugar residues and type of glycosidic linkage provide useful tools for structural analysis of the sugar moieties of AGPs. An exo-$\beta$-(1 → 3)-galactanase (EC 3.2.1.145) from *Irpex lacteus* acts on $\beta$-(1 → 3)-galactosyl backbones of the sugar moieties of radish AGP and releases side chains consisting of consecutive, at least up to 20, (1 → 6)-linked $\beta$-galactosyl groups and their acidic derivatives substituted with single uronosyl residues at non-reducing terminals [9]. An endo-$\beta$-(1 → 6)-galactanase (no EC number) has been purified from *Aspergillus niger* as described by Brillouet et al. [10]. We have also shown that an enzyme purified from a commercial cellulase preparation of *Trichoderma viride* specifically endo-hydrolyses $\beta$-(1 → 6)-galactosyl side chains of radish AGPs treated previously with a fungal $\alpha$-L-arabinofuranosidase [11]. Despite their importance for the structural and functional analyses of the carbohydrate moieties of AGPs, these enzymes have not been cloned until now. In the present study, we report the isolation of a cDNA clone encoding a Tv6GAL

Abbreviations used: ABEE, *p*-aminobenzoic acid ethyl ester; AGP, arabinogalactan protein; DP, degree of polymerization; MALDI–TOF-MS, matrix-assisted laser-desorption ionization–time-of-flight MS; RACE, rapid amplification of cDNA ends; Tv6GAL, *Trichoderma viride* endo-$\beta$-(1 → 6)-galactanase.

[1] To whom correspondence should be addressed (e-mail kotake@molbiol.saitama-u.ac.jp).

The nucleotide sequence depicted in Figure 1 has been submitted to the DDBJ nucleotide sequence database under the accession no. AB104898.

[*T. viride* endo-$\beta$-(1 → 6)-galactanase] and the characterization of the recombinant enzyme expressed in *Esherichia coli*. This is the first report of the isolation of the gene for an endo-$\beta$-(1 → 6)-galactanase. It presents a novel member of glycoside hydrolase family 5.

## EXPERIMENTAL

### Oligosaccharides and polysaccharides

Debranched arabinan from sugar beet, $\beta$-(1 → 4)-galactan from lupin, $\beta$-glucan from barley and CM-cellulose 4 M were purchased from Megazyme (Wicklow, Ireland). Pustulan from *Umbilicaria papullosa* was from Calbiochem (San Diego, CA, U.S.A.). Chitosan from crab shells, guar gum, locust bean gum, laminarin from *Laminaria digitata* and xylan from birch wood were from Sigma (St. Louis, MO, U.S.A.). CM-curdlan was from Wako (Osaka, Japan). $\beta$-(1 → 3)(1 → 6)-Galactan from *Prototheca zopfii*, $\beta$-(1 → 3)-galactan, native and $\alpha$-L-arabinofuranosidase-treated AGPs from radish and $\beta$-(1 → 3)-, $\beta$-(1 → 4)- and $\beta$-(1 → 6)-galacto-oligosaccharides were prepared as described in [11]. Methyl $\beta$-glycoside of $\beta$-(1 → 6)-galactohexaose was a gift from Dr Kováč of the National Institutes of Health, NIDDK (National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, U.S.A.).

### Peptide sequencing of endo-$\beta$-(1 → 6)-galactanase

Endo-$\beta$-(1 → 6)-galactanase was purified by conventional chromatographic techniques from Onozuka R-10, a commercial cellulase preparation from *T. viride* as described previously [11]. The purified endo-$\beta$-(1 → 6)-galactanase was separated on SDS/PAGE [12], blotted on to a PVDF-Plus membrane (Osmonics, Moers, Germany) and subjected to an N-terminal amino acid analysis with a protein sequencer HP G1000A (Hewlett Packard, Palo Alto, CA, U.S.A.). The N-terminal amino acid sequence (35 residues) was DTTLTIDPTSNWGTWEGWGVSLAWWAKAFGNRDDL.

### Culture of *T. viride* and cDNA cloning by reverse transcription-PCR

*T. viride* IFO 31 137 was obtained from the Institute for Fermentation Osaka (Osaka, Japan). Mycelia of *T. viride* were cultured in liquid medium containing 2 % (w/v) malt extract, 2 % (w/v) glucose and 0.1 % (w/v) peptone (pH 6.0) at 25 °C for 2 days. The submerged culture was then inoculated in wheat bran medium containing 2 g of wheat bran and 3 ml of water and cultured at 25 °C for 7 days.

Total RNA was extracted from conidia of *T. viride*. The conidia were frozen in liquid nitrogen, homogenized with mortar and pestle and extracted with a kit of Isogen (Nippon Gene, Tokyo, Japan) according to the manufacturer's instructions.

Two degenerate primers, F-1 (5′-AAYTGGGGNACNTGGGARGG-3′) and F-2 (5′-ACNTGGGARGGNTGGGG-3′), were designed based on the N-terminal amino acid sequence, NWGTWEG and TWEGWG respectively of endo-$\beta$-(1 → 6)-galactanase. Single-strand cDNA was synthesized from 2 $\mu$g of total RNA from the conidia using a reverse transcriptase, ReverTra Ace-$\alpha$- (Toyobo, Osaka, Japan) and oligo(dT)-adaptor primer (Takara, Tokyo, Japan). The first PCR was performed with the degenerate primer F-1 and an adaptor primer M13M4 (Takara) using the single-strand cDNA as a template under the following conditions: 0.5 min denaturing at 94 °C, 0.5 min annealing at 55 °C and 1.5 min amplification at 72 °C with 30 cycles. The nested second PCR was performed with F-2 and M13M4 primers using diluted first PCR product as a template under the same conditions. The amplified cDNA fragment including the 3′-untranslated region of the endo-$\beta$-(1 → 6)-galactanase gene was subcloned into a pGEM T-Easy vector (Promega, Madison, WI, U.S.A.) and the nucleotide sequence was determined with an ABI PRISM 310 genetic analyzer (Applied Biosystems, Foster City, CA, U.S.A.). The cDNA for the 5′-region of the gene was cloned with the aid of a kit for rapid amplification of 5′-cDNA ends (5′-RACE kit; Invitrogen, Carlsbad, CA, U.S.A.) using internal specific primers R-2 (5′-GAGTCTGGCCATTGACTGC-3′) and R-3 (5′-GGCCAGATCATCTCGGTTG-3′) corresponding to the 277–259 and 231–213 nt respectively (Figure 1). The coding region for Tv6GAL was amplified with proofreading polymerase (KOD plus; Toyobo) and the nucleotide sequence was determined.

### Southern-blot analyses

For the Southern-blot analyses, an AlkPhos Direct Labeling and Detection kit (Amersham Biosciences, Piscataway, NJ, U.S.A.) was used. The genomic DNA was extracted from cultured mycelia of *T. viride* with a DNeasy Plant Mini kit (Qiagen, Hilden, Germany). The genomic DNA was digested with restriction enzymes, separated on a gel containing 0.7 % agarose and blotted on to nylon membrane (Hybond $N^+$; Amersham Biosciences). The cDNA probe was the 733 bp fragment (479–1211 in Figure 1) excised from *Tv6GAL* cDNA with restriction enzymes *Pst*I and *Bgl*II. Probe labelling, hybridization and signal detection were performed according to the manufacturer's instructions.

### Expression of endo-$\beta$-(1 → 6)-galactanase in *E. coli* and its purification

Partial *Tv6GAL* cDNA corresponding to $Asp^{21}$-$Gln^{479}$ of the protein was amplified with specific primers, EX-F-2 primer (5′-GGATCCATGGACACCACGCTTACCATC-3′) and EX-R-1 primer (5′-*GAGCTC*ATTGCAACACAACGC-3′), into which restriction sites for *Bam*HI (underlined) and *Sac*I (italic) respectively were introduced. The cDNA fragment was subcloned into a pGEM5zf + vector (Promega), its nucleotide sequence was confirmed, and it was inserted between the *Bam*HI and *Sac*I sites of the pET32a expression vector (Novagen, Madison, WI, U.S.A.). The construct was designed to express endo-$\beta$-(1 → 6)-galactanase fused to thioredoxin and $His_6$ tags at the N-terminus. The plasmid construct was introduced into a BL21 gold strain of *E. coli* (Novagen). The *E. coli* cells were grown at 10 °C, and the recombinant protein was induced by treatment with 0.5 mM isopropyl $\beta$-D-thiogalactopyranoside for 24 h. The recombinant Tv6GAL was put on a Chelating Sepharose FF column (Amersham Biosciences). The column was washed with 50 mM sodium phosphate buffer (pH 7.2) containing 50 mM imidazole, and the bound protein was then eluted with the same buffer containing 250 mM imidazole. The active fraction was dialysed against 50 mM sodium phosphate buffer (pH 7.2) and subjected to a second chelating chromatography with the same elution method. Approx. 730 $\mu$g of purified recombinant Tv6GAL was obtained from a 500 ml culture of *E. coli*, which was used for further enzymic characterization. Since the recombinant protein has an enterokinase and a thrombin cleavage site following N-terminal thioredoxin and $His_6$ tags, we attempted to cleave the thrombin site by digestion of the purified recombinant Tv6GAL (approx. 10 $\mu$g) with thrombin (0.1 unit; Novagen) at 20 °C for 24 h to split off the preceding sequence. The digest was examined for its specific activity using algal galactan as a substrate.

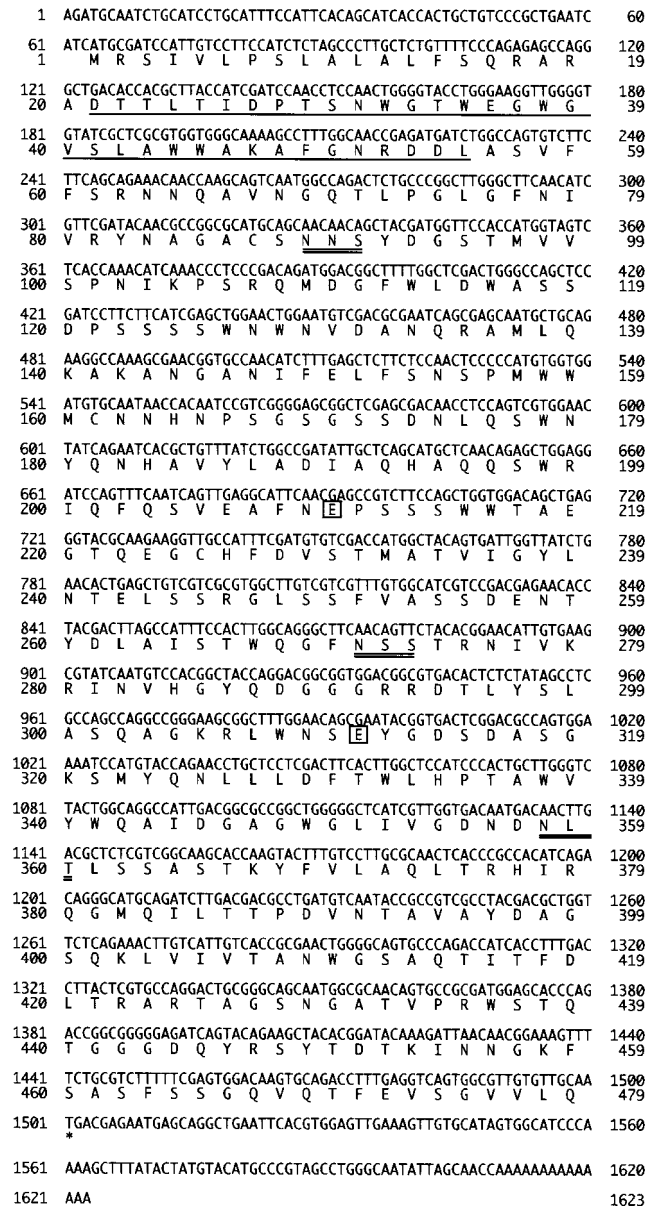

**Figure 1 Nucleotide sequence of the *Tv6GAL* gene and the deduced amino acid sequence**

Nucleotides are numbered from the first base of the cDNA clone. The underlined amino acid sequence denotes a peptide sequence corresponding to the N-terminal sequence of the native endo-$\beta$-(1 → 6)-galactanase. The peptide sequence preceding the N-terminal sequence represents the putative signal sequence composed of 20 amino acid residues. Doubly underlined amino acids represent possible N-glycosylation sites, and boxed glutamic acids are putative catalytic residues found in family 5 glycosyl hydrolases.

### Substrate specificity and mode of action of the recombinant Tv6GAL

Enzymic activity of the purified recombinant Tv6GAL was measured using reaction mixtures (total volume, 0.2 ml) consisting of the recombinant enzyme, 0.25 or 0.5 % (w/v) polysaccharide or 5 mM galacto-oligosaccharide, 0.01 % (w/v) BSA and 50 mM sodium acetate buffer (pH 4.3). The reaction mixtures were incubated at 25 °C for 0–6 h. The activity of the recombinant enzyme towards polysaccharides was determined as the increase in the reducing terminals of liberated sugars by the Somogyi–Nelson method [13,14]. Enzyme activity (1 unit) liberates 1 $\mu$mol of reducing sugars/min. Mono- and oligo-saccharides in enzymic hydrolysates were separated by TLC on Silica gel $60F_{254}$ (Merck, Darmstadt, Germany) using 7:1:2 (by vol.) propan-1-ol/ethanol/water as solvent and detected by charring after spraying TLC plates with 20 % (v/v) $H_2SO_4$/methanol.

The mode of action of the recombinant Tv6GAL was analysed using methyl $\beta$-glycoside of $\beta$-(1 → 6)-galactohexaose and $\alpha$-L-arabinofuranosidase-treated AGP from radish as substrates. The reaction mixture (total volume, 20 $\mu$l) containing 5 mM methyl $\beta$-glycoside of $\beta$-(1 → 6)-galactohexaose or 0.5 % (w/v) $\alpha$-L-arabinofuranosidase-treated AGP from radish, 18 m-units of the recombinant enzyme and 50 mM sodium acetate buffer (pH 4.3) was incubated at 25 °C for 0–23 h. At appropriate time intervals, a portion was withdrawn and inactivated by heating. The reducing sugars liberated were coupled at their reducing terminals with ABEE (*p*-aminobenzoic acid ethyl ester) by the method of Matsuura and Imaoka [15]. The ABEE-derivatized sugars were analysed on an HPLC system equipped with a TSKgel Amide-80 column (4.6 mm × 250 mm; Tosoh, Tokyo, Japan). The column was eluted with a linear gradient of $CH_3CN$/water from 74:26 to 58:42 (v/v), for 40 min at a flow rate of 1 ml/min and 40 °C. ABEE sugars were monitored by a fluorescence detector model RF-$10A_{XL}$ at 305 nm (excitation) and 360 nm (emission).

### MALDI–TOF-MS (matrix-assisted laser-desorption ionization–time-of-flight MS) analysis of recombinant Tv6GAL protein

MALDI–TOF-MS was performed with a KOMPACT MALDI IV tDE (Shimadzu, Kyoto, Japan). Recombinant Tv6GAL protein dissolved in 0.5 $\mu$l of water was crystallized by adding 0.5 $\mu$l of matrix solution containing 1 % sinapinic acid and 0.1 % trifluoroacetic acid and 0.5 $\mu$l of 1 % NaCl, which was then allowed to dry. For the mass calibration, BSA (molecular mass of 66.431 kDa) was used.

## RESULTS

### Nucleotide sequence of the endo-$\beta$-(1 → 6)-galactanase

Full-length cDNA for an endo-$\beta$-(1 → 6)-galactanase was isolated by 3′-RACE PCR using degenerate primers designed based on the N-terminal amino acid sequence determined for the native enzyme purified from a commercial cellulase preparation from *T. viride*, followed by a 5′-RACE PCR procedure with the internal specific primers. The cloned cDNA, designated as *Tv6GAL*, appeared to encode a polypeptide of 479 amino acids (molecular mass of 52.662 kDa) including a sequence which completely matched the N-terminal amino acid sequence of native endo-$\beta$-(1 → 6)-galactanase (Figure 1, underlined). The cDNA contained a putative signal sequence (20 amino acid residues) preceding the N-terminal sequence of the mature enzyme (459 amino acids, molecular mass of 50.488 kDa). Three putative N-glycosylation sites were found in the deduced protein sequence (Figure 1, double underlined). The calculated pI of 5.23 for mature Tv6GAL was similar to that (pI 5.4) determined for the native enzyme on isoelectric electrophoresis [11]. The entire amino acid sequence deduced for *Tv6GAL* did not show any strong similarity with known glycoside hydrolases, but shared 59 % similarity with a hypothetical protein from *Neurospora crassa* (accession no. EAA29705). The sequence showed low similarity with xylanases from *Erwinia chrysanthemi* (14.3 %, accession no. AAB53151) and from *Aeromonas caviae* (13.4 %, accession no. U86340), suggesting that Tv6GAL is a distant member of family 5 glycoside hydrolase.

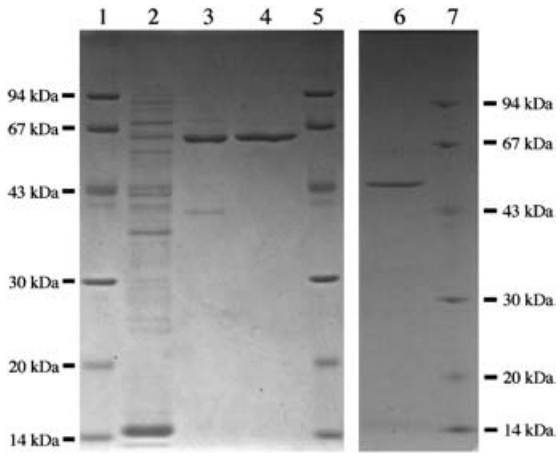


**Figure 2 SDS/PAGE of recombinant Tv6GAL at different purification steps**

The enzyme proteins (2 $\mu$g) obtained after different purification steps of the recombinant Tv6GAL were analysed by SDS/PAGE. Lanes 1, 5 and 7, molecular-mass markers; lane 2, lysate of *E. coli*; lanes 3 and 4, recombinant Tv6GAL purified once and twice respectively on a chelating column; lane 6, thrombin-digested recombinant Tv6GAL (observed and calculated molecular masses are 51 and 54.394 kDa respectively). Protein in the gel was stained with Coomassie Brilliant Blue R-250.

Southern-blot analyses were performed to determine the number of *Tv6GAL* genes in the *T. viride* genome. A labelled cDNA probe hybridized to the single restriction fragment of genomic DNA digested with *Sal*I, *Hin*dIII or *Xho*I. The *Bam*HI digested fragment showed two bands at positions 15.0 and 0.8 kb, which probably stem from an internal *Bam*HI site in the *Tv6GAL* gene (nt 659 in Figure 1) (results not shown). These results indicate that *Tv6GAL* exists as a single copy gene in the *T. viride* genome.

### Expression of recombinant Tv6GAL

To examine the characteristics of the gene product of *Tv6GAL*, partial cDNA encoding mature endo-$\beta$-(1 $\rightarrow$ 6)-galactanase ($Asp^{21}$-$Gln^{479}$) was cloned into the pET32a expression vector. The recombinant protein was then expressed in *E. coli* as a protein fused to thioredoxin and $His_6$ tags to allow its purification on chelating chromatography. By monitoring the enzymic activity during several purification procedures using algal $\beta$-(1 $\rightarrow$ 3)(1 $\rightarrow$ 6)-galactan from *P. zopfii* as the substrate, the recombinant Tv6GAL was purified into a homogeneous state by two successive chelating chromatographies. SDS/PAGE of the purified protein confirmed its homogeneity with an apparent molecular mass of approx. 64 kDa (Figure 2). However, this value is approx. 4 kDa less than the 68.296 kDa expected for the fused protein. We found a similar discrepancy for a specimen of the native enzyme purified from a commercial cellulase preparation [11] (molecular mass of approx. 47 kDa on SDS/PAGE) when comparing it with the expected value (50.488 kDa) deduced from the cDNA sequence. Therefore MALDI–TOF-MS was applied to determine accurate molecular masses of the native enzyme and the recombinant Tv6GAL. The obtained value for the native enzyme, 50.6874 kDa, was almost the same as the molecular mass (50.488 kDa) calculated for the native enzyme (Figure 3A). Moreover, the determined value for the recombinant protein, 68.1793 kDa, was in good agreement with the calculated molecular mass (68.296 kDa)

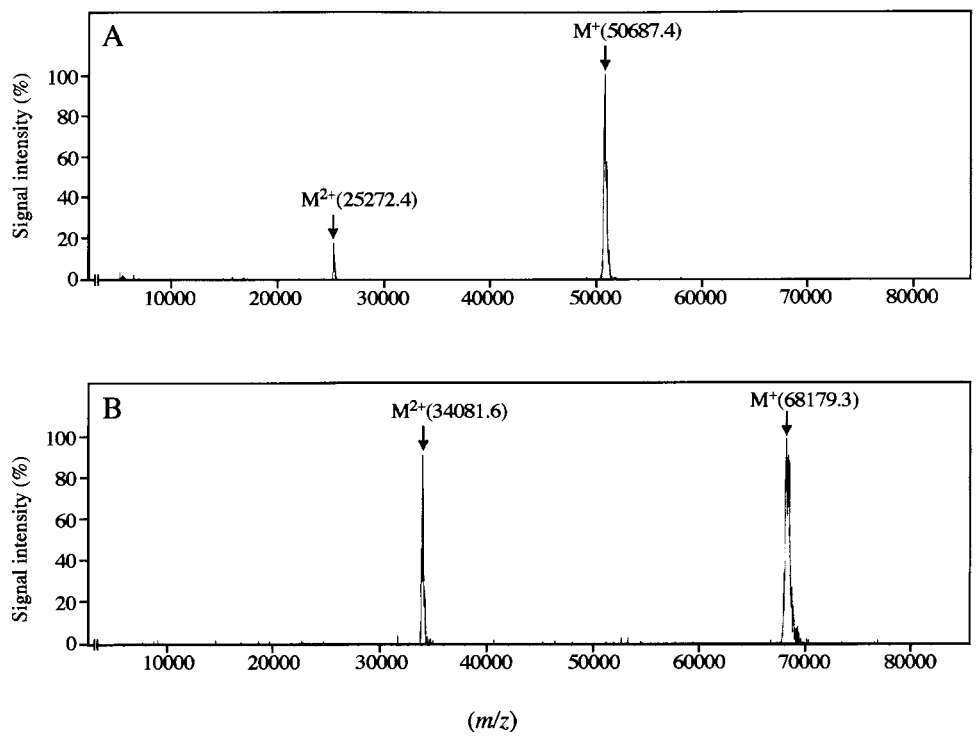


**Figure 3 MALDI–TOF-MS analyses of the native enzyme and recombinant Tv6GAL**

Molecular mass of the native enzyme (**A**) and the recombinant Tv6GAL (**B**) was analysed by MALDI–TOF-MS and observed mass values are shown above the peaks. The molecular ions ($M^+$, *m/z*), 50687.4 for the native enzyme and 68179.3 for the recombinant Tv6GAL are in good agreement with the calculated masses, 50488 for the native enzyme and 68296 for the recombinant Tv6GAL respectively.

**Table 1 Substrate specificity of recombinant Tv6GAL towards polysaccharides**

| Substrate | Relative activity (%)* |
|---|---|
| Algal galactan | |
| $\beta$-(1 → 3)(1 → 6)-Galactan from *P. zopfii* | 100 |
| AGP | |
| $\alpha$-L-arabinofuranosidase-treated AGP from radish | 85 |
| Native AGP from radish | 11 |
| Others | |
| $\beta$-(1 → 3)-Galactan | 0 |
| $\beta$-(1 → 4)-Galactan from lupin | 0 |
| CM-curdlan ($\beta$-(1 → 3)-glucan) | 0 |
| CM-cellulose | 0 |
| $\beta$-(1 → 3)(1 → 4)-Glucan from barley | 0 |
| Laminarin from *L. digitata* ($\beta$-(1 → 3)(1 → 6)-glucan) | 0 |
| Pustulan from *U. papullosa* ($\beta$-(1 → 6)-glucan) | 0 |
| Galactomannan from guar† | 0 |
| Galactomannan from locust bean† | 0 |
| Xylan from birchwood | 0 |
| Chitosan from crab shells† | 0 |
| Debranched arabinan ($\alpha$-(1 → 5)-L-arabinan) from sugar beet† | 0 |

* Activity is expressed as percentage of that towards algal galactan taken as 100.
† The enzyme was incubated with polymers at a concentration of 2.5 mg/ml and 5 mg/ml was employed for other polymers.

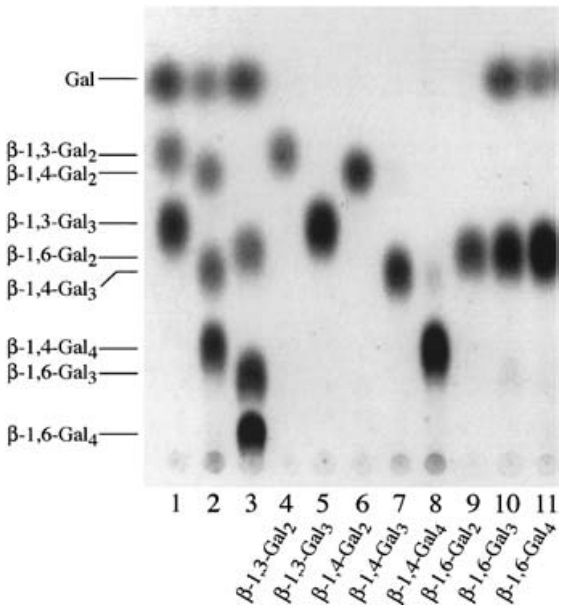


**Figure 4 Substrate specificity of the recombinant enzyme**

Hydrolysis products of various oligosaccharides by the action of the recombinant Tv6GAL were analysed on TLC. Lane 1, standard galactose and $\beta$-(1 → 3)-galacto-oligomers with DP 2 and 3; lane 2, galactose and $\beta$-(1 → 4)-galactooligomers with DP 2–4; lane 3, galactose and $\beta$-(1 → 6)-galacto-oligomers with DP 2–4. Localization of the standard sugars is indicated on the left side. Susceptibility of these oligomers to the enzyme was examined as indicated below each lane.

expected from the *Tv6GAL*/pET32a construct (Figure 3B). The lower molecular masses of the recombinant protein and the native enzyme on SDS/PAGE compared with those expected from the sequence are therefore attributable to the nature of Tv6GAL itself, and not to the elimination of peptide fragment(s) by post-translational processing.

The specific activity (7.8 units/mg of protein) of the recombinant Tv6GAL towards algal $\beta$-(1 → 3)(1 → 6)-galactan was approx. 1/20th of that (160 units/mg of protein) of the native Tv6GAL purified from a commercial cellulase preparation from *T. viride* [11]. Because this lower activity may be caused by the fused thioredoxin protein and/or the $His_6$ tags, the recombinant protein was digested with thrombin to remove the thioredoxin and $His_6$ tags. Certainly, thrombin hydrolysed its cleavage site giving a reduced molecular mass of 51 kDa (Figure 2) and raised the specific activity (213 units/mg of protein) of the recombinant Tv6GAL to a level comparable with that of the native enzyme purified from a commercial preparation, suggesting that the lower activity of the recombinant Tv6GAL can be attributed to the N-terminal thioredoxin protein and/or the $His_6$ tags.

### Substrate specificity of recombinant Tv6GAL

The substrate specificity of the recombinant Tv6GAL towards polysaccharides is summarized in Table 1. The recombinant Tv6GAL hydrolysed polymer substrates containing (1 → 6)-linked $\beta$-galactosyl residues such as algal galactan from *P. zopfii* and AGPs from radish. The $\alpha$-L-arabinofuranosidase-treated AGP from radish was a substrate for the recombinant Tv6GAL approx. eight times as good as the native AGP, suggesting that L-arabinofuranosyl residues attached to AGP prevent the enzyme from accessing $\beta$-(1 → 6)-galactosyl side chains of AGP. This has also been observed for the action of the native enzyme [11]. The recombinant enzyme failed to hydrolyse other polysaccharides such as CM-cellulose, galactomannan and laminarin that serve as substrates for other members of the family 5 glycoside hydrolase. The activity of the recombinant enzyme towards oligosaccharides was examined using $\beta$-(1 → 3)-, $\beta$-(1 → 4)- and $\beta$-(1 → 6)-galacto-oligosaccharides. The recombinant enzyme hydrolysed $\beta$-(1 → 6)-galacto-triose and -tetraose completely into galactose and $\beta$-(1 → 6)-galactobiose, whereas the other oligosaccharides did not serve as substrates (Figure 4). The fact that it does not hydrolyse $\beta$-(1 → 6)-galactobiose indicates that $\beta$-(1 → 6)-galactotriose is the minimum substrate for the recombinant enzyme. Together with the data on polysaccharides, these results confirm that the cloned cDNA encodes an endo-$\beta$-(1 → 6)-galactanase with substrate specificity identical with that of the native enzyme purified from a commercial cellulase preparation for our experiments [11].

### Mode of action

The mode of action of the recombinant enzyme was analysed using the methyl $\beta$-glycoside of $\beta$-(1 → 6)-galactohexaose and $\alpha$-L-arabinofuranosidase-treated AGP as substrates. The hydrolysis products released as reducing sugars by the recombinant enzyme were coupled with ABEE and monitored by the HPLC system. In the initial phase of the hydrolysis reaction, methyl $\beta$-glycoside of $\beta$-(1 → 6)-galactohexaose produced $\beta$-(1 → 6)-galacto-oligosaccharides with DP (degree of polymerization) 2–5 (Figure 5A). The only final products were galactose and $\beta$-(1 → 6)-galactobiose. The recombinant Tv6GAL also released a series of $\beta$-(1 → 6)-galacto-oligosaccharides with DP 2–8 at the initial hydrolysis stage of $\alpha$-L-arabinofuranosidase-treated AGP (Figure 5B) and galactose and $\beta$-(1 → 6)-galactobiose at the final stage as observed for the oligosaccharide substrate. These results indicate that the recombinant Tv6GAL acts in an endo-fashion on consecutive (1 → 6)-linked $\beta$-galactosyl sequences. Enzymic digestion of $\alpha$-L-arabinofuranosidase-treated AGP also produced sugars (as ABEE derivatives), which passed through the HPLC column at 2.7 min, faster than standard ABEE (at 3.3 min,

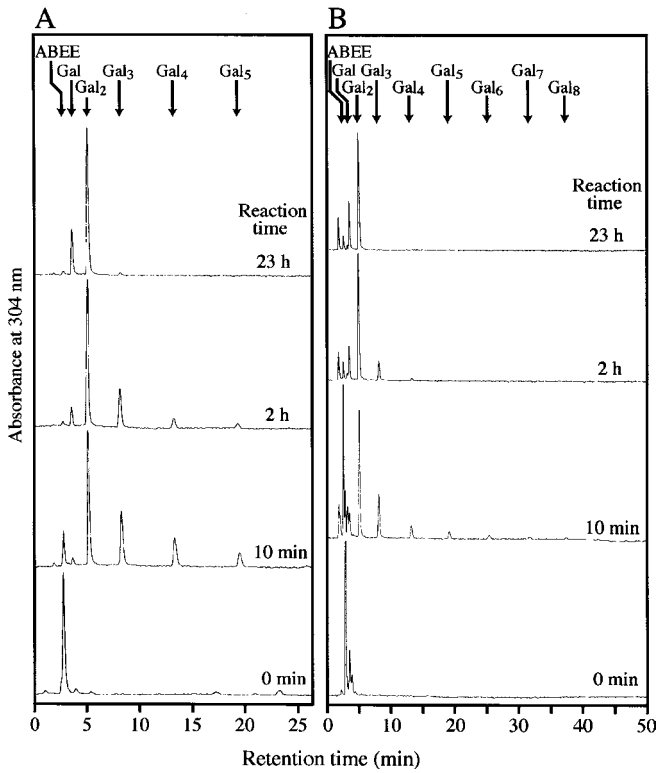


**Figure 5 Mode of action of recombinant Tv6GAL**

The mode of action of the recombinant enzyme was examined using the methyl $\beta$-glycoside of $\beta$-(1 → 6)-galactohexaose (**A**) and $\alpha$-L-arabinofuranosidase-treated AGP (**B**) as the substrates. Reducing sugars released from each substrate were derivatized with ABEE and analysed by HPLC equipped with an Amide-80 column. Arrows indicate the elution positions of standard ABEE and ABEE-derivatized galactose and $\beta$-(1 → 6)-galacto-oligomers with DP 2–4 and the consecutive DPs (DP 5–8) deduced for the remaining oligomers. The intensity of fluorescence at reaction times 10 min, 2 h and 23 h is scaled down to 40, 7 and 3 % (**A**), and 70, 14 and 7 % (**B**) respectively compared with the intensity at 0 min.

Figure 5B) after 10 min enzymic reaction, which seem to be a mixture of acidic sugars, namely 4-*O*-methyl-glucuronosyl-$\beta$-(1 → 6)-galacto-oligomers released from acidic side chains of $\alpha$-L-arabinofuranosidase-treated AGP concomitant with the neutral galacto-oligomers mentioned above. Release of both neutral and acidic galacto-oligomers has been observed in enzymic hydrolysate of radish AGP by the action of the native enzyme [11]. Thus the mode of action of the recombinant enzyme determined in the present study completely matches with that of the native endo-$\beta$-(1 → 6)-galactanase purified from a commercial cellulase preparation [11].

## DISCUSSION

Genomic information indicates that *Arabidopsis* possesses at least 40 different genes encoding core proteins in AGPs and AGP-like substances [16]. Because of overlapping expression of AGPs in plant tissues and their genetic redundancy, the fine structure including the sugar moieties of each AGP molecule has not been clarified. In addition, functional knock-out of AGPs by disruption or transcriptional suppression of core protein genes has not yet been achieved. Enzymes capable of hydrolysing $\beta$-(1 → 3)(1 → 6)-galactan backbones in AGPs should be useful in the investigation of the structure of the sugar moieties of AGPs. Furthermore, it is anticipated that the introduction of genes coding such enzymes into plants will cause structural modifications of AGPs *in vivo*, leading to an understanding of the molecular function of AGPs. We hope that this cloning of such an enzyme hydrolysing the carbohydrate moieties of AGPs is a first step in that direction.

Recently, another $\beta$-(1 → 6)-galactanase has been purified from *A. niger* and characterized by Luonteri et al. [17]. This enzyme has a molecular mass of 58 kDa and exhibits activity towards $\beta$-(1 → 6)-galactosyl residues of arabinogalactans from larch wood and Norway spruce. The entire amino acid sequence of the enzyme has not been determined, but the N-terminal sequence of 15 amino acids shows similarity to neither Tv6GAL nor the other members of the family 5 glycosyl hydrolase. It has been known that glycosyl hydrolases contain carbohydrate-binding modules in their N- and/or C-termini. However, the N-terminal sequence of $\beta$-(1 → 6)-galactanase from *A. niger* does not show any similarity to the known carbohydrate-binding modules, suggesting that Tv6GAL is different from $\beta$-(1 → 6)-galactanase of *A. niger*. The substrate specificity of the $\beta$-(1 → 6)-galactanase from *A. niger* is also different from that of Tv6GAL. The enzyme from *A. niger* possesses activity to release monomeric arabinose from black liquor $\alpha$-(1 → 5)-L-arabinan [17], whereas Tv6GAL does not hydrolyse $\alpha$-(1 → 5)-L-arabinan. Furthermore, the enzyme from *A. niger* does not produce $\beta$-(1 → 6)-galacto-oligosaccharides larger than galactobiose when acting on larch wood arabinogalactan, whereas Tv6GAL released a series of $\beta$-(1 → 6)-galacto-oligosaccharides with DP 2–8 during initial hydrolysis of $\alpha$-L-arabinofuranosidase-treated AGP. The properties of the enzyme from *A. niger* could not be simply compared with those of Tv6GAL, because reaction conditions and substrates employed for each enzyme are different and the intermediate products of the arabinogalactans hydrolysed by the enzyme from *A. niger* have not been shown. However, present results suggest that the enzyme from *A. niger* is distinct from Tv6GAL, at least with respect to its primary structure.

The present study provides a cDNA sequence for a fungal endo-$\beta$-(1 → 6)-galactanase gene, *Tv6GAL*. On the basis of hydrophobic cluster analysis, glycoside hydrolases were classified into more than 90 families [18,19]. Tv6GAL shares low similarity with other family 5 glycoside hydrolases such as xylanases from *E. chrysanthemi* (14.3 %) and *A. caviae* (13.4 %) suggesting that Tv6GAL can be classified as a distant member of family 5 glycoside hydrolase. This family is one of the major family of clan GH-A, which includes various kinds of enzymes such as endoglycosylceramidase (EC 3.2.1.123), $\beta$-mannosidase (EC 3.2.1.25), cellulase (EC 3.2.1.4), exo-$\beta$-(1 → 3)-glucanase (EC 3.2.1.58), licheninase (EC 3.2.1.73), endo-$\beta$-(1 → 6)-glucanase (EC 3.2.1.75), endo-$\beta$-(1 → 4)-mannanase (EC 3.2.1.78), endo-$\beta$-(1 → 4)-xylanase (EC 3.2.1.8) and $\beta$-(1 → 4)-cellobiosidase (EC 3.2.1.91). The enzymes belonging to family 5 glycoside hydrolase are inferred to have a $(\beta/\alpha)_8$ barrel structure and it is known that they operate via a retaining double-displacement mechanism, which is catalysed by two glutamic residues acting as an acid/base pair and a nucleophile. Tertiary structure prediction with the 3D-PSSM program (sbg.bio.ic.ac.uk/~3dpssm/, [20]) shows that the structure of Tv6GAL is most similar to a $\beta$-mannanase (EC 3.2.1.78) from *Trichoderma reesei*, which belongs to the glycoside hydrolase family 5 [21], and suggests that Tv6GAL also has a $(\beta/\alpha)_8$ barrel structure and can be classified as a member of family 5 glycoside hydrolase. Structural analyses for two family 5 enzymes, an exo-$\beta$-(1 → 3)-glucanase from *Candida albicans* [22] and a $\beta$-mannanase from *Thermomonospora fusca* [23], have revealed that eight amino acids encompassing the catalytic site are conserved in family 5 glycoside hydrolases. Although

the sequence similarity of Tv6GAL with the other members of family 5 is low, the eight conserved amino acids, $Arg^{81}$, $His^{164}$, $Asn^{209}$, $Glu^{210}$ (acid/base), $His^{284}$, $Tyr^{286}$, $Glu^{311}$ (nucleophile) and $Trp^{341}$ are found in Tv6GAL, corresponding to $Arg^{54}$, $Trp^{114}$, $Asn^{168}$, $Glu^{169}$ (acid/base), $His^{241}$, $Tyr^{243}$, $Glu^{276}$ (nucleophile) and $Trp^{306}$ in $\beta$-mannanase from *T. reesei* respectively. Together with the tertiary structure analysis, the conserved amino acids encompassing the catalytic site strongly suggest that Tv6GAL is a member of family 5 glycoside hydrolase.

Except for polymers having consecutive $\beta$-(1 → 6)-galactosyl groups, the recombinant enzyme does not hydrolyse any other polysaccharides that serve as substrates for other members of the family 5 group (Table 1), confirming the characteristic action of Tv6GAL among the group. To address the strict recognition of substrates by Tv6GAL and its conformational relationships with other members in the family 5 group, stereochemical analysis of the three-dimensional structure of Tv6GAL is required.

We are grateful to Dr B. Henrissat (Architecture et Fonction des Macromolecules Biologiques, CNRS, Marseille, France) for the classification analysis of Tv6GAL. We also thank Dr P. Kováč (the National Institutes of Health, NIDDK, U.S.A.) for providing an oligosaccharide substrate.

## REFERENCES

1 Fincher, G. B., Stone, B. A. and Clarke, A. E. (1983) Arabinogalactan-proteins: structure, biosynthesis, and function. Annu. Rev. Plant Physiol. **34**, 47–70
2 Nothnagel, E. A. (1997) Proteoglycans and related components in plant cells. Int. Rev. Cytol. **174**, 195–291
3 Majewska-Sawka, A. and Nothnagel, E. A. (2000) The multiple roles of arabinogalactan proteins in plant development. Plant Physiol. **122**, 3–9
4 Schultz, C. J., Johnson, K. L., Currie, G. and Bacic, A. (2000) The classical arabinogalactan protein gene family of *Arabidopsis*. Plant Cell **12**, 1751–1767
5 Wu, H.-M., Wang, H. and Cheung, A. Y. (1995) A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. Cell (Cambridge, Mass.) **82**, 395–403
6 Shi, H., Kim, Y., Guo, Y., Stevenson, B. and Zhu, J. K. (2003) The *Arabidopsis SOS5* locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. Plant Cell **15**, 19–32
7 Casper, Y., Johnson, K. L., McKenna, J. A., Bacic, A. and Schultz, C. J. (2001) The complex structures of arabinogalactan-proteins and the journey towards understanding function. Plant Mol. Biol. **47**, 161–176
8 Tsumuraya, Y., Ogura, K., Hashimoto, Y., Mukoyama, H. and Yamamoto, S. (1988) Arabinogalactan-proteins from primary and mature roots of radish (*Raphanus sativus* L.). Plant Physiol. **86**, 155–160
9 Tsumuraya, Y., Mochizuki, N., Hashimoto, Y. and Kovác, P. (1990) Purification of an exo-$\beta$-(1 → 3)-galactanase of *Irpex lacteus* (*Polyporus tulipiferae*) and its action on arabinogalactan-proteins. J. Biol. Chem. **265**, 7207–7215
10 Brillouet, J.-M., Williams, P. and Moutounet, M. (1991) Purification and some properties of a novel endo-$\beta$-(1 → 6)-galactanase from *Aspergillus niger*. Agric. Biol. Chem. **55**, 1565–1571
11 Okemoto, K., Uekita, T., Tsumuraya, Y., Hashimoto, Y. and Kasama, T. (2003) Purification and characterization of an endo-$\beta$-(1 → 6)-galactanase from *Trichoderma viride*. Carbohydr. Res. **338**, 219–230
12 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**, 680–685
13 Nelson, N. J. (1944) A photometric adaptation of the Somogyi method for the determination of glucose. J. Biol. Chem. **153**, 375–380
14 Somogyi, N. (1952) Notes on sugar determination. J. Biol. Chem. **195**, 19–23
15 Matsuura, F. and Imaoka, A. (1988) Chromatographic separation of asparagine-linked oligosaccharides labeled with an ultraviolet absorbing compound, *p*-aminobenzoic acid ethyl ester. Glycoconj. J. **5**, 13–26
16 Schultz, C. J., Rumsewicz, M. P., Johnson, K. L., Jones, B. J., Gaspar, Y. M. and Bacic, A. (2002) Using genomic resources to guide research directions. The arabinogalactan protein gene family as a test case. Plant Physiol. **129**, 1448–1463
17 Luonteri, E., Laine, C., Uusitalo, S., Teleman, A., Siika-Aho, M. and Tenkanen, M. (2003) Purification and characterization of *Aspergillus* $\beta$-D-galactanases acting on $\beta$-1,4- and $\beta$-1,3/6-linked arabinogalactans. Carbohydr. Pol. **53**, 155–168
18 Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280**, 309–316
19 Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **293**, 781–788
20 Kelley, L. A., MacCallum, R. M. and Sternberg, M. J. (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. J. Mol. Biol. **299**, 499–520
21 Sabini, E., Schubert, H., Murshudov, G., Wilson, K. S., Siika-Aho, M. and Penttila, M. (2000) The three-dimensional structure of a *Trichoderma reesei* $\beta$-mannanase from glycoside hydrolase family 5. Acta Crystallogr. D **56**, 3–13
22 Cutfield, S. M., Davies, G. J., Murshudov, G., Anderson, B. F., Moody, P. C., Sullivan, P. A. and Cutfield, J. F. (1999) The structure of the exo-$\beta$-(1,3)-glucanase from *Candida albicans* in native and bound forms: relationship between a pocket and groove in family 5 glycosyl hydrolases. J. Mol. Biol. **294**, 771–783
23 Hilge, M., Gloor, S. M., Rypniewski, W., Sauer, O., Heightman, T. D., Zimmermann, W., Winterhalter, K. and Piontek, K. (1998) High-resolution native and complex structures of thermostable $\beta$-mannanase from *Thermomonospora fusca* – substrate specificity in glycosyl hydrolase family 5. Structure **6**, 1433–1444

Received 29 July 2003/7 October 2003; accepted 17 October 2003
Published as BJ Immediate Publication 17 October 2003, DOI 10.1042/BJ20031145