Arabinanase A from *Pseudomonas fluorescens* subsp. *cellulosa* exhibits both an endo- and an exo- mode of action

Vincent A. McKIE*, Gary W. BLACK*‡, Sarah J. MILLWARD-SADLER*, Geoffrey P. HAZLEWOOD†, Judith I. LAURIE† and Harry J. GILBERT*§

*Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, and †Department of Cellular Physiology, The Babraham Institute, Babraham, Cambridge CB2 4AT, U.K.

Pseudomonas fluorescens subsp. *cellulosa* expressed arabinanase activity when grown on media supplemented with arabinan or arabinose. Arabinanase activity was not induced by the inclusion of other plant structural polysaccharides, and was repressed by the addition of glucose. The majority of the *Pseudomonas* arabinanase activity was extracellular. Screening of a genomic library of *P. fluorescens* subsp. *cellulosa* DNA constructed in Lambda ZAPII, for recombinants that hydrolysed Red-dyed arabinan, identified five arabinan-degrading plaques. Each of the phage contained the same *Pseudomonas* arabinanase gene, designated *arbA*, which was present as a single copy in the *Pseudomonas* genome. The nucleotide sequence of *arbA* revealed an open reading frame of 1041 bp encoding a protein, designated arabinanase A (ArbA), of M_r 39438. The N-terminal sequence of ArbA exhibited features typical of a prokaryotic signal peptide. Analysis of the primary structure of ArbA indicated that, unlike most *Pseudomonas* plant cell wall hydrolases, it did not contain linker sequences or have a modular structure, but consisted of a single catalytic domain. Sequence comparison between the *Pseudomonas* arabinanase and proteins in the SWISS-PROT database showed that ArbA exhibits greatest sequence identity with arabinanase A from *Aspergillus niger*, placing the enzyme in glycosyl hydrolase Family 43. The significance of the differing substrate specificities of enzymes in Family 43 is discussed. ArbA purifed from a recombinant strain of *Escherichia coli* had an M_r of 34000 and an N-terminal sequence identical to residues 32–51 of the deduced sequence of ArbA, and hydrolysed linear arabinan, carboxymethylarabinan and arabino-oligosaccharides. The enzyme displayed no activity against other plant structural polysaccharides, including branched sugar beet arabinan. ArbA produced almost exclusively arabinotriose from linear arabinan and appeared to hydrolyse arabino-oligosaccharides by successively releasing arabinotriose. ArbA and the *Aspergillus* arabinanase mediated a decrease in the viscosity of linear arabinan that was associated with a significant release of reducing sugar. We propose that ArbA is an arabinanase that exhibits both an endo- and an exo- mode of action.

INTRODUCTION

Arabinan is a pectic polysaccharide found primarily in seeds, fruits and vegetables. It consists of a backbone of α-1,5-linked L-arabinofuranosyl residues, some of which are substituted with α-1,2- and α-1,3-linked side chains of L-arabinose in the furanose conformation [1]. There is some evidence to suggest that in plant cell walls arabinan is generally linked to the rhamnopyranosyl units of rhamnogalacturonan via an α-1,2 linkage [2], although arabinan composed of only L-arabinose has been isolated from white mustard cotyledons [3]. The two major enzymes that hydrolyse arabinan, α-L-arabinofuranosidases (EC 3.2.1.55) and endo-α-1,5-arabinanases (EC 3.2.1.99), are found in phytopathogenic and saprophytic micro-organisms [4]. The former enzyme cleaves the arabinose side chains, allowing the latter enzyme to attack the arabinan backbone [5]. Although several plant cell wall hydrolases, such as cellulases, xylanases and mannanases, have been subjected to extensive analysis, there is a paucity of information on the structure and biochemistry of arabinan-degrading enzymes. The molecular structure of only a single endo-α-1,5-arabinanase has been described; that of arabinanase A (ABNA) from *Aspergillus niger*. This enzyme belongs to glycosyl hydrolase Family 43 and consists of a single catalytic domain [6].

Previous work by our group has shown that *Pseudomonas fluorescens* subsp. *cellulosa* expresses a wide range of plant cell wall hydrolases, including numerous cellulases, xylanases and mannanases, an arabinoxylan-specific arabinofuranosidase, an acetyl-xylan esterase and a cellodextrinase [7–11]. *Pseudomonas* enzymes that attack xylans and cellulose have a modular structure comprising a catalytic domain linked to a cellulose-binding domain (CBD) via hydroxyamino acid-rich sequences [8,12], while only the *P. fluorescens* endo-β-1,4-mannanase, of the enzymes described to date, comprises a single catalytic domain and lacks a CBD [11]. To further develop our knowledge of the plant cell wall hydrolases expressed by *P. fluorescens* subsp. *cellulosa*, and to increase our understanding of the structure/function relationship of endo-arabinanases, we have analysed the biochemical properties and molecular structure of a *Pseudomonas* endo-α-1,5-arabinanase. In this paper we demonstrate that *P.*

Abbreviations used: ABNA, arabinanase A from *Aspergillus niger*; ArbA, arabinanase A from *Pseudomonas fluorescens* subsp. *cellulosa*; CBD, cellulose-binding domain; CFE, cell-free extract; IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria broth; ORF, open reading frame; PC buffer, 50 mM phosphate/12 mM citrate buffer, pH 6.5.

‡ Present address: School of Health Sciences, University of Sunderland, Fleming Building, Wharncliffe Street, Sunderland SR2 3SD, U.K.

§ To whom correspondence should be addressed.

The nucleotide sequence of *arbA* will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession number Y10458.

fluorescens subsp. *cellulosa* hydrolyses arabinan and describe the primary sequence and catalytic properties of the major *Pseudomonas* arabinanase. The enzyme, which exhibits sequence identity with ABNA and is therefore classified as a glycosyl hydrolase Family 43 enzyme, does not contain a CBD. The predominant product produced from linear arabinan by the enzyme was arabinotriose, suggesting that it has an exo- mode of action.

MATERIALS AND METHODS

Bacterial strains, vectors and growth conditions

P. fluorescens subsp. *cellulosa* and *Escherichia coli* were grown in Luria broth (LB) at 37 °C. In arabinanase induction experiments the pseudomonad was cultured in LB, supplemented with the appropriate sugar or polysaccharide to a final concentration of 0.5 % (w/v), for 12 h, which equates to mid-exponential phase. Media used to culture *E. coli* were supplemented with ampicillin (100 μg/ml) and 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside to select for transformants and pMTL-derived recombinants respectively. To select for arabinanase-positive bacteria, LB-agar was supplemented with 0.2 % Red-dyed arabinan (Megazyme Pty). To induce transcription of genes cloned into pET vectors, *E. coli* cultures in mid-exponential phase were supplemented with isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, and induction was allowed to proceed for a further 3 h. Recombinant Lambda phage were grown on NZY medium seeded with the *E. coli* strain XL1-Blue (Stratagene). *E. coli* JM83 [13] and BL21(DE3) were used as the host organisms for the pMTL- and pET-based plasmids respectively. The vectors used in this work were Lambda ZAPII, pMTL22p, pMTL23p and pMTL6000 [14], and pET21a (Novagen).

Screening of a genomic library for arabinanase-positive phage

A genomic library of *P. fluorescens* subsp. *cellulosa* DNA, constructed in Lambda ZAPII [11], was plated out on NZY top agarose supplemented with Red-dyed arabinan to a density of 10 plaques/cm^2. After overnight incubation at 37 °C, arabinanase-expressing phage were surrounded by clear haloes in a red background. Selected phage were subjected to a second round of purification, after which all the phage screened were arabinanase-positive.

Recombinant DNA technology

Digestion of DNA with restriction endonucleases, ligation of DNA using T4 DNA ligase, agarose-gel electrophoresis, Southern hybridization and transformation of *E. coli* were carried out as described previously [7]. Plasmid and phage DNA were purified using Qiagen columns following the manufacturer's instructions. To sequence cloned *Pseudomonas* DNA, appropriate restriction fragments were subcloned into pMTL-based vectors and sequenced using the PRISM® ready reaction dye-deoxy terminator cycle sequencing kit supplied by Applied Biosystems Instruments (ABI) and the ABI 373 automatic sequencing model. Sequences were compiled and ordered using computer programs described by Staden [15]. The cloned DNA was sequenced completely in both strands.

To clone the full-length arabinanase gene, or a derivative encoding the mature enzyme, into pET21a, PCR was used to amplify the appropriate sequences containing the desired restriction sites at the 5′ and 3′ ends respectively. The 100 μl reactions contained 10 mM Tris/HCl buffer, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 500 μM dNTPs, 200 ng of DNA (pVM1), 250 μM primers and 2.5 units of DNA polymerase from *Thermus aquaticus* (supplied by ABI). The reactions were overlaid with 50 μl of mineral oil and then subjected to 30 cycles of 1 min at 94 °C, 1 min at 55 °C and 2 min at 72 °C. The primers used to amplify the full-length arabinanase gene were 5′ CTCCATATG-CCCACCCACCACCGATCACC 3′ and 5′ CTCAAGCTTTT-TCAACAGCACATCGACAG 3′ (primer A), while primer A and 5′ CTCCATATGGCCAAACAGGTCGATGTGCAC 3′ were used to amplify the region of the arabinanase gene between nucleotides 91 and 1041. The amplified DNAs were digested with *Nde*I and *Hin*dIII and cloned into *Nde*I/*Hin*dIII-restricted pET21a to generate pVM3 and pVM4, which contain the full-length and truncated forms of the arabinanase gene respectively.

Purification of the recombinant arabinanase

E. coli XL1-Blue harbouring pVM1 was cultured for 16 h in LB containing 100 μg/ml ampicillin. Cells harvested from 3 litres of culture were disrupted by sonication in 120 ml of 50 mM Tris/HCl buffer, pH 8.0, and a cell-free extract (CFE) was prepared by centrifuging at 20000 ***g*** for 20 min. Ammonium sulphate was added to 50 % saturation and the resultant precipitate containing arabinanase activity was recovered by centrifugation, redissolved in 10 mM Tris/HCl buffer, pH 8.0 (buffer A), dialysed against three changes of buffer A and loaded on to a column of DEAE-Trisacryl M (Biosepra; 30 mm × 200 mm), which was then eluted with a linear gradient containing 0–250 mM NaCl in buffer A. Arabinanase-containing fractions were pooled and arabinanase A (ArbA) was precipitated with ammonium sulphate (90 % saturation), redissolved in 1 ml of buffer A and applied to a column containing Bio-Gel P-60 (Bio-Rad Laboratories; 20 mm × 900 mm). Purified ArbA was obtained by eluting the column under gravity at a flow rate of 2 ml/h with buffer A.

Enzyme assays and protein analysis

The source of the enzyme was purified arabinanase, culture supernatants of *P. fluorescens* subsp. *cellulosa* or CFEs of appropriate recombinant *E. coli* strains. CFEs were prepared as described above. To evaluate the activity of the enzyme preparations against plant structural polysaccharides, 100 μl of the enzyme preparation was added to 0.65 ml of 50 mM phosphate/12 mM citrate buffer, pH 6.5 (PC buffer), containing 0.5 % (final concentration) of the appropriate polysaccharide, and the reaction was incubated for 20 min at 37 °C. The release of reducing sugar was detected by adding an equal volume of 3,5-dinitrosalicylic acid reagent, incubating at 100 °C for 15 min and measuring the A_{575} of the reaction [16]. To determine enzyme activity against arylglycosides, 1 ml of PC buffer containing the appropriate substrate at a concentration of 10 mM was incubated with 50 μl of the purified arabinanase, and the release of the chromophore was monitored at A_{420}. The arylglycosides evaluated were as follows: 4-nitrophenyl α-L-arabinofuranoside, 4-nitrophenyl α-L-arabinopyranoside and 4-nitrophenyl β-xylopyranoside, which were purchased from Sigma Chemical Co., and 2,4-dinitrophenyl β-xylobioside, which was a gift from Dr. Neil Hughes (Department of Chemistry, University of Newcastle, U.K.) [17]. One unit of enzyme generated 1 μmol of product/min at 37 °C.

The capacity of the arabinanase preparations to decrease the viscosity of linear arabinan and of carboxymethylarabinan was assessed using a Contraves low-shear 30-cuvette rotating viscometer attached to a 7015 X-Y Hewlett–Packard recorder. The substrate was dissolved in 5 ml of either PC buffer (for ArbA) or 200 mM sodium acetate buffer, pH 4.0 (for ABNA), to a final concentration of 10 % (w/v) and incubated with 0.4 unit of either

Pseudomonas ArbA or ABNA from *A. niger* (supplied by Megazyme Pty Ltd.) at 37 °C. At regular time intervals the viscosity of the solution was measured and the amount of reducing sugar generated was determined as described above.

HPLC was used to identify the products generated by *Pseudomonas* ArbA and *A. niger* ABNA from arabinan and arabino-oligosaccharides. An appropriate dilution of each of the arabinanases was incubated with 0.2% linear arabinan or 500 μg/ml arabino-oligosaccharides (supplied by Megazyme Pty Ltd.) in 50 mM Tris/HCl buffer, pH 7.5 (ArbA), or 200 mM sodium acetate buffer, pH 4.0. At regular time intervals, samples were removed and the enzyme was destroyed by heating at 100 °C for 5 min. The reactions were then loaded on to a Dionex PA1 column and the oligosaccharides were eluted with 100 mM NaOH containing a 0–500 mM gradient of sodium acetate over 25 min. The sugars were detected with a pulsed amperometer, and were identified and quantified by comparing their elution times and response magnitudes with those of standard amounts of arabinose and arabino-oligosaccharides, up to arabinoheptaose. A linear relationship was observed between the amount of standard injected and the detector response.

The capacity of the arabinanase to bind to Avicel and other forms of cellulose was evaluated as described by Ferreira et al. [12]. The method of Sedmak and Grossberg [18] was used to measure protein concentration, and SDS/PAGE was carried out as described by Laemmli [19].

RESULTS

Does *P. fluorescens* subsp. *cellulosa* express arabinanase activity?

To evaluate whether *P. fluorescens* subsp. *cellulosa* is capable of degrading arabinan, the bacterium was grown on solid medium containing Red-dyed linear arabinan for 36 h. The data (not shown) showed that the pseudomonad colonies were surrounded by clear haloes in a red background, indicating that the bacterium hydrolyses the polysaccharide. To establish under what conditions *Pseudomonas* expresses arabinanase activity, the bacterium was grown on media supplemented with various carbon sources. It was apparent that arabinose (3.7×10^{-2} units/ml of culture) and both linear (4.0×10^{-2} units/ml of culture) and substituted (4.3×10^{-2} units/ml of culture) arabinan induced the expression of arabinanases, while the plant structural polysaccharides carboxymethylcellulose, oat-spelt xylan, carob galactomannan and β-1,4-galactan did not. Glucose repressed the synthesis of arabinan-degrading enzymes [$(0–8) \times 10^{-4}$ units/ml of culture] in *P. fluorescens* subsp. *cellulosa* grown in media containing arabinose or linear arabinan. When culture supernatants of the pseudomonad, grown on media that induced arabinanase activity, were incubated with linear arabinan, the major product generated was arabinotriose (results not shown). In excess of 97% of the arabinanase activity was detected in culture supernatants, suggesting that the arabinan-degrading enzymes were primarily extracellular.

Isolation and characterization of *Pseudomonas* arabinanase genes

To isolate genes that encode arabinanases, a library of *P. fluorescens* subsp. *cellulosa* genomic DNA, constructed in Lambda ZAPII, was screened for arabinanase-expressing recombinants on media containing Red-dyed arabinan. From 20000 phage analysed, five exhibited significant arabinanase activity. The *Pseudomonas* DNA insert from one of the arabinanase-positive phage was subcloned into low (pMTL6000)- and high (pMTL23p)-copy-number *E. coli* vectors to generate the recombinants pVM1 and pVM2 respectively. A restriction map of the pseudomonad insert in the two plasmids is presented in Figure 1. *E. coli* cells harbouring either recombinant plasmid initially expressed arabinanase activity; however, on further culturing, *E. coli* containing pVM2 lost its arabinanase-positive phenotype, and deletions occurred in the recombinant plasmid. These data suggest that high-level expression of the *Pseudomonas* arabinanase was lethal to *E. coli*. The subcloned arabinanase gene and the encoded enzyme are designated *arbA* and ArbA respectively.

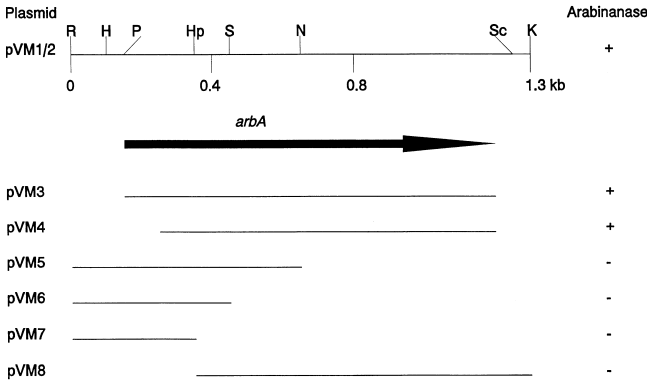

Figure 1 Restriction map of plasmids containing *arbA*

The positions of the cleavage sites for *Hin*dIII (H), *Hpa*I (Hp), *Kpn*I (K), *Nco*I (N), *Pst*I (P), *Eco*RI (R), *Sca*I (Sc) and *Stu*I (S) are displayed. The pseudomonad inserts in pVM1 and pVM2 were cloned into pMTL6000 and pMTL23 respectively. The plasmids pVM3 and pVM4 were generated by cloning PCR-amplified DNA, derived from pVM1, into pET21a, as described in the Materials and methods section. The other plasmids were generated by subcloning restriction fragments into pMTL vectors. The solid arrow shows the extent and orientation of *arbA*. The catalytic activity of ArbA encoded by the recombinant plasmids is shown (+/−).

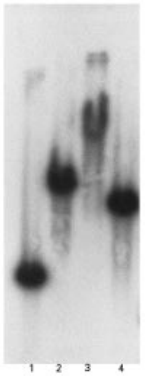

Figure 2 Southern hybridization of *P. fluorescens* subsp. *cellulosa* DNA with *arbA*

Pseudomonas DNA digested with *Eco*RI (lane 1), *Xho*I (lane 2), *Bam*HI (lane 3) and *Kpn*I (lane 4) was subjected to Southern hybridization using the pseudomonad insert from pVM1 as the probe.

To investigate whether *arbA* is a single-copy gene, *P. fluorescens* subsp. *cellulosa* chromosomal DNA was subjected to Southern hybridization, using *arbA* as the probe. The data (Figure 2) showed that there was a single copy of the arabinanase gene in the pseudomonad genome. Plaque hybridization experiments indicated that the other four arabinanase-positive phage isolated from the genomic library also contained *arbA*. These data suggest

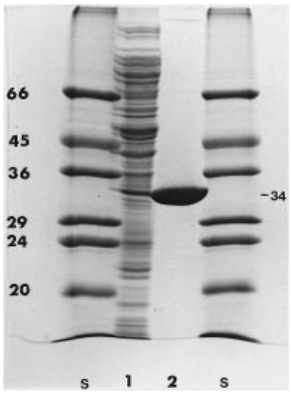

Figure 3 SDS/PAGE of recombinant ArbA

CFEs derived from *E. coli* harbouring pVM1 (lane 1) or ArbA (lane 2), purified as described in the Materials and methods section, and low-M_r standards from Sigma Chemical Co. (lane s), were subjected to SDS/PAGE using a 10% (w/v) polyacrylamide gel. The numbers indicate $10^{-3} \times M_r$ for the markers (left) and purified ArbA (right).

that *P. fluorescens* subsp. *cellulosa* contains a single arabinanase gene.

Characterization of ArbA

ArbA was purified from *E. coli* cells harbouring pVM1. The enzyme had an M_r of 34000 (Figure 3) and the purified enzyme had a specific activity 100-fold higher than that of the original cell-free extract. The N-terminal sequence of the arabinanase was KQVDVHDPVMTREGDTWYLF, and the K_m and k_{cat} values of the enzyme against linear arabinan were 3.75 mg/ml and 109 s^{-1} respectively. Analysis of the substrate specificity of ArbA showed that it hydrolysed only carboxymethylarabinan (k_{cat} 212 s^{-1}) and linear arabinan (k_{cat} 207 s^{-1}); the enzyme did not attack branched sugar beet arabinan containing α-1,3-arabinose side-chains or hydrolyse any of the other major plant structural polysaccharides such as cellulose (Avicel, amorphous cellulose and carboxymethylcellulose), xylan, galactan and mannan. In addition, ArbA did not hydrolyse 4-nitrophenyl α-L-arabinofuranoside, 4-nitrophenyl α-L-arabinopyranoside, 4-nitrophenyl β-xyloside or 2,4-dinitrophenyl β-xylobioside.

To investigate the nature of the products released from linear arabinan by ArbA, HPLC analysis was performed on the reaction products. The data (Figure 4) show clearly that the major product was arabinotriose, while *Aspergillus* ABNA initially released a mixture of arabino-oligosaccharides, which is typical of an endo-acting enzyme. Similarly, when ArbA was incubated with arabino-oligosaccharides, arabinotriose was again the major product (Figure 5). Thus arabino-octaose was cleaved initially to arabinopentaose and arabinotriose; arabinoheptaose to arabinotetraose and arabinotriose; arabinohexaose to arabinotriose; arabinotetraose to arabinotriose and arabinose; while arabinotriose was very slowly hydrolysed to arabinose and arabinobiose. The relative activity of the enzyme against the arabino-oligosaccharides increased with increasing size of the substrate up to arabinohexaose, after which the rate remained constant (Table 1). These data suggest that ArbA has a substrate-binding site that accommodates six arabinose units, with cleavage occurring between binding sites 3 and 4. The predominant release of arabinotriose from linear arabinan suggests that the enzyme displays significant exo-activity.

To establish whether ArbA has an exclusively exo-mode of

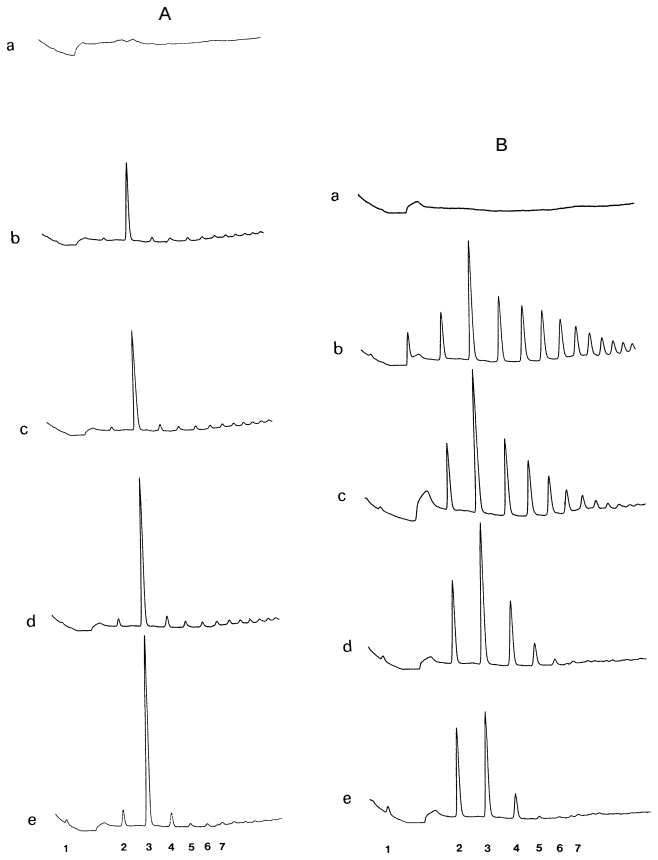

Figure 4 HPLC analysis of the products generated from linear arabinan by ArbA and ABNA

Linear arabinan (0.2%) was incubated with 0.0032 unit of ArbA (**A**) or 0.0023 unit of ABNA (**B**) in 1 ml of PC buffer (ArbA) or 200 mM sodium acetate buffer, pH 4.0 (ABNA). At 0 min (trace a), 2 min (b), 4 min (c), 8 min (d) and 16 min (e), samples were removed and subjected to HPLC analysis as described in the Materials and methods section. The elution times of the standards arabinose (1), arabinobiose (2), arabinotriose (3), arabinotetraose (4), arabinopentaose (5), arabinohexaose (6) and arabinoheptaose (7) are indicated. The data shown provide a qualitative assessment only of the products generated by the two enzymes, as the volume of the samples injected on to the HPLC column varied, to ensure that the peaks remained on the scale.

action, the enzyme was incubated with linear arabinan and carboxymethylarabinan, and the decrease in viscosity was compared with reducing sugar release. The data (Figure 6) show that the release of reducing sugar was associated with a decrease in viscosity. When the experiment was repeated with the well documented ABNA from *A. niger*, the data were very similar to those observed for ArbA. These results suggest that *Pseudomonas* ArbA exhibits an endo-mode of action and significant exo-activity, successively releasing arabinotriose units.

Nucleotide sequence of *arbA*

The *Pseudomonas* DNA in pVM1 was sequenced completely in both strands, revealing a single open reading frame (ORF) of 1041 bp encoding a protein of M_r 39438 (Figure 7). Several lines of evidence indicated that the ORF constituted *arbA*: (1) 7 bp upstream of the putative translation initiation codon of *arbA* is the sequence AGAGGAG, which is typical of prokaryotic ribosome-binding sequences; (2) plasmids containing just the complete 1041 bp ORF expressed arabinanase activity; (3) truncated

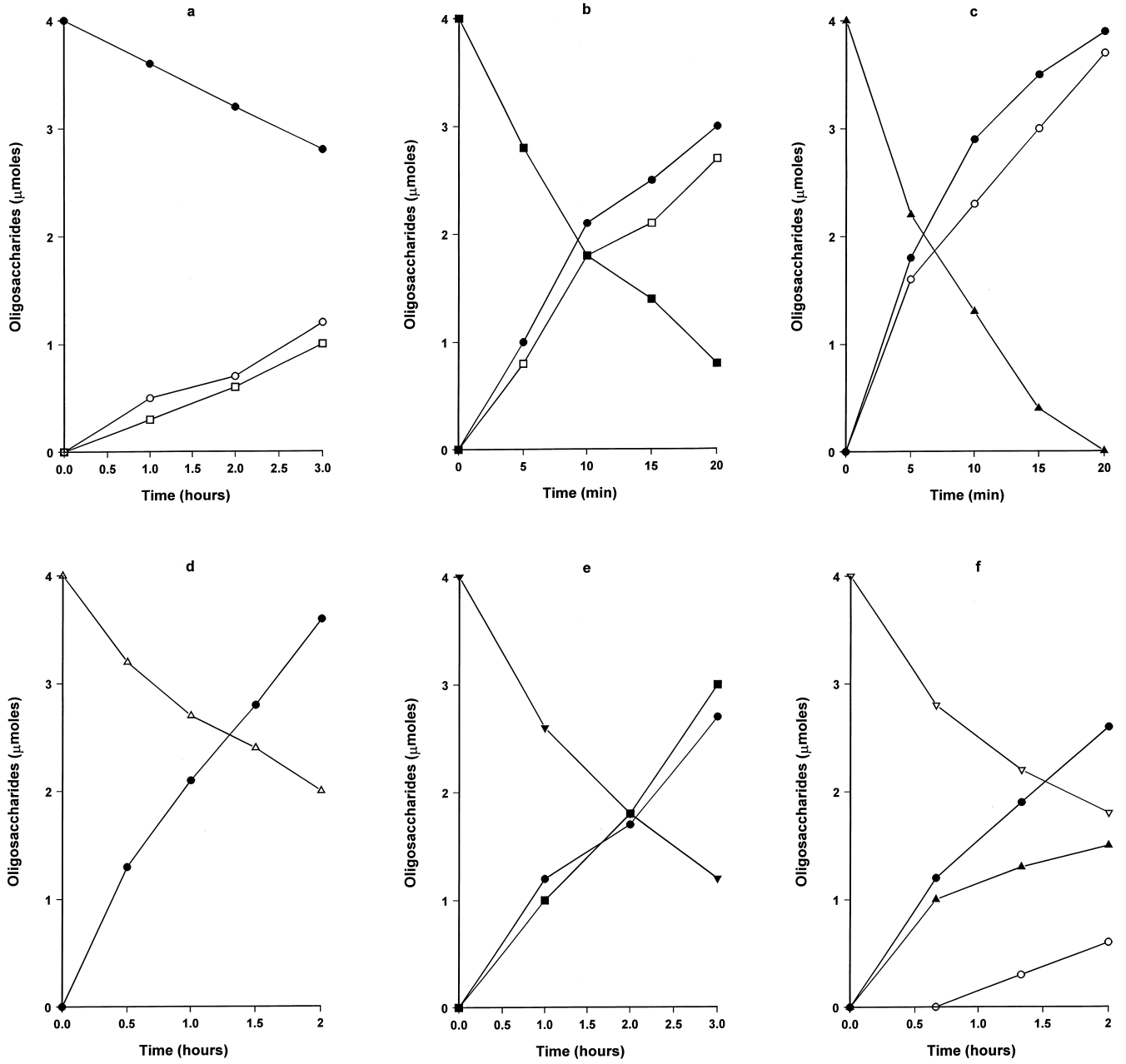

Figure 5 Analysis of the products generated by the action of ArbA on arabino-oligosaccharides

Arabino-oligosaccharides were incubated with the following amounts of ArbA: arabinotriose, 0.5 unit (**a**); arabinotetraose, 0.25 unit (**b**); arabinopentaose, 0.125 unit (**c**); arabinohexaose, 0.00125 unit (**d**); arabinoheptaose, 0.00125 unit (**e**); arabino-octaose, 0.00125 unit (**f**). At regular intervals aliquots were removed from the reactions and the products generated were analysed by HPLC for the presence of arabinose (□), arabiobiose (○), arabinotriose (●), arabinotetraose (■), arabinopentaose (▲), arabinohexaose (△), arabinoheptaose (▼) and arabino-octaose (▽).

derivatives of pVM1 lacking the complete ORF did not direct the synthesis of a functional arabinanase (Figure 1); (4) the N-terminal sequence of recombinant ArbA corresponded to residues 32–51 of the deduced primary structure of the arabinanase; and (5) the experimentally determined M_r of recombinant ArbA (34000; Figure 3) was similar to the predicted size of the mature protein (36438) derived from the nucleotide sequence data. The codon usage of *arbA* is similar to that of other *P. fluorescens* subsp. *cellulosa* plant cell wall hydrolase genes. The gene makes little use of the codons that modulate *E. coli* gene expression and, like other genes from *P. fluorescens* subsp. *cellulosa*, does not contain a high G/C content in the Wobble position [20]. These data question whether *P. fluorescens* subsp. *cellulosa* really belongs to the genus *Pseudomonas*.

The N-terminal 31 amino acids of ArbA were similar to bacterial signal peptides. The hydrophilic N-terminal 14 amino acids contained several basic amino acids, and were followed by a stretch of 17 small hydrophobic residues. The cleavage of the signal peptide in *E. coli*, between residues 31 and 32, is consistent with the known specificity of bacterial signal peptidases [21]. Inspection of the primary structure of ArbA does not reveal any serine- or glycine-rich linker sequences typical of modular plant

Table 1 Activity of ArbA against arabino-oligosaccharides

The relative activity of ArbA against the substrates was determined by the time taken to hydrolyse 50% of the substrate (4 μmol).

Arabino-oligosaccharide	Relative activity of ArbA (units)
Arabinotriose	1
Arabinotetraose	200
Arbinopentaose	700
Arabinohexaose	2000
Arabinoheptaose	2174
Arabino-octaose	1767

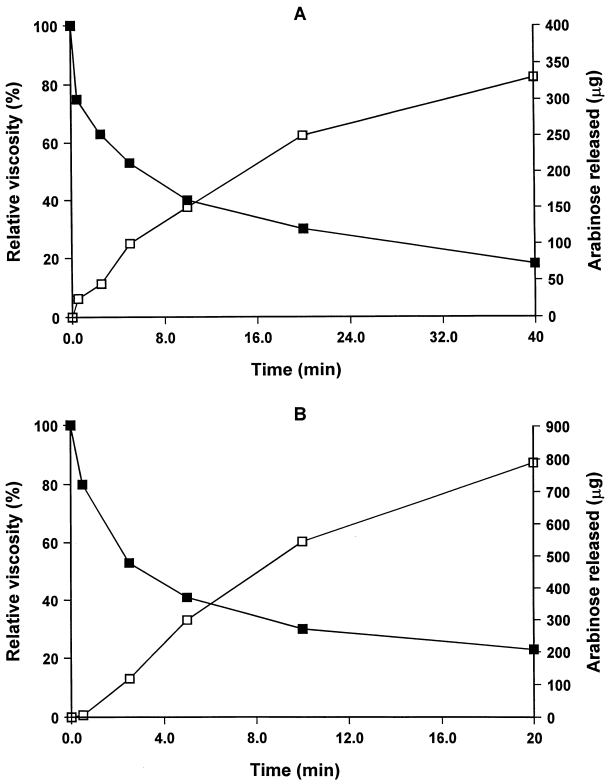

Figure 6 Activity of ArbA and ABNA on linear arabinan

Linear arabinan (10%, w/v) was incubated with ArbA (**A**) or ABNA (**B**), and the release of reducing sugar (□) and the decrease in the viscosity of the substrate (■) were determined as described in the Materials and methods section.

```
                                                                    -100
ATTGACACCGCACTATCCGCCGCTTCAGGCAAGGCGGGATAATTCAGGCTTGAGCTGGAT

GAGCAAATACTATTATATTGCAATATTGTCAGACAATATAAACAACAAAGCTTAACCCTG

                                                          +1
CCAAGACACTGCAGCGGTTTTATGAGAGGAGAGTTCCCATGCCCACCCACCACCCGATCA
                                                        M  P  T  H  H  P  I

CCCGCCAGCACTGGCACCACTCTTGGTTGTCCGCACTTGCACTGCTCTGTGCAAGCCTTG
T  R  Q  H  W  H  H  S  W  L  S  A  L  A  L  L  C  A  S  L

                    100
CCTGTGGTGCCAAACAGGTCGATGTGCACGATCCGGTCATGACCCGCGAAGGCGATACCT
A  C  G  A  K  Q  V  D  V  H  D  P  V  M  T  R  E  G  D  T

                                                                   300
GGTATTTGTTCAGTACCGGCCCCGGCATCACCATATACAGCTCCAAAGACCGCGTTAACT
W  Y  L  F  S  T  G  P  G  I  T  I  Y  S  S  K  D  R  V  N

GGCGCTACTCCGACCGCGCCTTTGGCACCGAGCCCACCTGGGCCAAACGTGTTTCGCCCT
W  R  Y  S  D  R  A  F  G  T  E  P  T  W  A  K  R  V  S  P

                                                  400
CCTTTGACGGGCATTTGTGGGCCCCGGATATTTACCAGCACAAAGGCCTGTTTTACCTCT
S  F  D  G  H  L  W  A  P  D  I  Y  Q  H  K  G  L  F  Y  L

ATTACTCCGTTTCGGCCTTTGGCAAAAATACCTCGGCGATTGGCGTAACGGTGAATAAAA
Y  Y  S  V  S  A  F  G  K  N  T  S  A  I  G  V  T  V  N  K

                    500
CGCTTAATCCGGCCTCCCCCGATTACCGCTGGGAAGACAAGGGTATCGTGATTGAATCCG
T  L  N  P  A  S  P  D  Y  R  W  E  D  K  G  I  V  I  E  S

                                                                   600
TTCCCCAGCGCGACCTGTGGAATGCGATTGATCCGGCGATCATTGCCGATGACCATGGCC
V  P  Q  R  D  L  W  N  A  I  D  P  A  I  I  A  D  D  H  G

AGGTGTGGATGAGTTTTGGTTCCTTCTGGGGCGGATTAAAACTCTTCAAACTCAATGATG
Q  V  W  M  S  F  G  S  F  W  G  G  L  K  L  F  K  L  N  D

                                                  700
ACCTGACCCGCCCTGCGGAGCCCCAGGAGTGGCATAGCATCGCCAAACTCGAACGCTCGG
D  L  T  R  P  A  E  P  Q  E  W  H  S  I  A  K  L  E  R  S

TACTGATGGATGACAGCCAGGCGGGCAGCGCGCAGATTGAAGCGCCGTTTATATTGCGCA
V  L  M  D  D  S  Q  A  G  S  A  Q  I  E  A  P  F  I  L  R

                    800
AGGGCGATTATTATTATTTGTTTGCCAGTTGGGGACTCTGCTGCCGCAAAGGTGATTCCA
K  G  D  Y  Y  Y  L  F  A  S  W  G  L  C  C  R  K  G  D  S

                                                                   900
CTTACCACCTGGTTGTGGGCCGCAGTAAACAGGTGACGGGGCCTTATCTGGATAAAACGG
T  Y  H  L  V  V  G  R  S  K  Q  V  T  G  P  Y  L  D  K  T

GCCGGGATATGAACCAGGGCGGCGGTAGCCTGTTGATCAAAGGCAACAAACGCTGGGTCG
G  R  D  M  N  Q  G  G  G  G  S  L  L  I  K  G  N  K  R  W  V

                                                 1000
GACTCGGCCATAACAGTGCCTATACCTGGGACGGCAAGGATTACCTGGTGTTACATGCCT
G  L  G  H  N  S  A  Y  T  W  D  G  K  D  Y  L  V  L  H  A

ATGAAGCCGCTGATAATTACCTGCAAAAACTGAAAATCCTCAACCTGCATTGGGATGGCG
Y  E  A  A  D  N  Y  L  Q  K  L  K  I  L  N  L  H  W  D  G

                   1100
AAGGTTGGCCACAAGTCGATGAAAAGGAATTGGATAGTTACATCAGCCAACGTTTGAAAT
E  G  W  P  Q  V  D  E  K  E  L  D  S  Y  I  S  Q  R  L  K

                                                                  1200
AACTGTTGGTTAATAAACTCGCGCGGTTATCAGTAAATAATCTGTCGATGTGCTGTTGAA

AGTCTTATGCTGGTAATCAGTACTTTTTTGCGGGCGAGGTACCC
```

Figure 7 Nucleotide sequence of *arbA*

The nucleotide sequence of *arbA* and the deduced primary structure of ArbA are shown. The experimentally determined N-terminal sequence of ArbA purified from recombinant *E. coli* is underlined.

cell wall hydrolases. Similarly, the enzyme does not contain regions that resemble non-catalytic domains, such as the Type 1 CBDs [22], that are prevalent in *Pseudomonas* cellulases and xylanases. These data, coupled with the observation that removal of sequences from either the N- or C-terminus of ArbA inactivated the enzyme (Figure 1), indicate that the arabinanase is non-modular, comprising only a single catalytic domain.

Comparison of the primary structure of ArbA with those of other proteins in the SWISS-PROT/GenBank/EMBL databases revealed that the enzyme exhibited 27% and 32% sequence identity respectively with ABNA from *A. niger* and the product of an ORF from *Bacillus subtilis* whose function has not been defined (Figure 8). Hydrophobic cluster analysis has shown that ABNA is a member of glycosyl hydrolase Family 43 [23]. It would appear, therefore, that ArbA is also a member of Family 43.

Figure 8 Alignment of ArbA with ABNA and the product of a *B. subtilis* ORF

The primary sequences of ArbA, ABNA [6] and the *B. subtilis* protein (Bsu; EMBL accession number Z75208) are aligned. Residues that exhibit identity in all three proteins are boxed. The positions of the amino acids at the end of each row, in the respective primary sequences, are shown.

Hyperexpression of *arbA* in *E. coli*

To increase the expression of *arbA* in *E. coli*, full-length *arbA*, and a derivative of the gene encoding the mature form of the arabinanase, were inserted into the expression vector pET21a to generate pVM3 and pVM4 respectively. Analysis of the capacity of *E. coli* strain BL21(DE3), harbouring pVM3 and pVM4, to express arabinanase activity showed that high levels of ArbA were produced only by cells containing the truncated form of *arbA* (Figure 9). The specific activities of ArbA in CFEs derived from *E. coli* cells harbouring pVM3 and pVM4 that had been induced with IPTG for 4 h were 0.1 and 21.8 units/mg of protein respectively.

DISCUSSION

Previous work has established that *P. fluorescens* subsp. *cellulosa* produces an extensive repertoire of cellulose- and hemicellulose-degrading enzymes. Data presented here clearly show that the pseudomonad expresses arabinanase activity and is therefore able to degrade the pectic polysaccharides of the cell wall matrix. Although arabinanase activity was induced by the substrate arabinan, and was subject to catabolite repression by glucose, the enzymes(s) was not synthesized in response to the addition of other plant structural polysaccharides. This is in contrast with *Pseudomonas* cellulases and xylanases; these enzymes are induced by both xylan and cellulose [24]. Presumably this reflects the origin of the different polysaccharides. Plant cell walls invariably contain both cellulose and xylan, and thus it is rational to suppose that the presence of either polysaccharide signals the close proximity of plant cell wall material, and hence induces the synthesis of the major plant structural polysaccharidases. In contrast, arabinan is not prevalent in all plant cell walls, and is often only loosely associated with the structural matrix. Thus, as arabinan does not always signal the presence of significant quantities of plant cell wall material, the synthesis of arabinanases and the major plant cell wall hydrolases, such as cellulases and xylanases, is not co-ordinated.

The isolation of only one arabinanase gene from the *P. fluorescens* subsp. *cellulosa* genomic library indicates that the bacterium synthesizes a single arabinanase. However, this conclusion must be viewed with some caution, as it is possible that the pseudomonad contains other arabinanase genes which are not expressed in *E. coli*, or that the regions of DNA containing these other genes may not be present in the genomic library if they are lethal to the enteric bacterium. However, given that arabinotriose, which is the major product generated by ArbA, is the predominant oligosaccharide produced from arabinan by cultures of *P. fluorescens* subsp. *cellulosa*, we suggest that this enzyme is, at least, the major arabinanase synthesized by the pseudomonad.

Analysis of the molecular architecture of ArbA suggests that it consists of a single catalytic domain; it does not contain a CBD, which are prevalent in other *P. fluorescens* subsp. *cellulosa* plant cell wall hydrolases [8–10]. It has been suggested [10] that the major role of CBDs is to sequester a range of plant cell wall hydrolases on to the cell wall matrix, promoting intimate enzyme–substrate and enzyme–enzyme associations which increase the efficiency of plant cell wall hydrolysis. That ArbA lacks a CBD supports the view, discussed above, that arabinan, which is the primary target for the enzyme, is not tightly associated with the plant cell wall and is thus readily accessible to the *Pseudomonas* arabinanase. The enzyme therefore does not require prolonged association with the plant cell wall to elicit efficient substrate hydrolysis, and hence does not contain a CBD.

Attempts to increase the level of expression of full-length *arbA*, by cloning the gene into either a high-copy-number plasmid or an efficient expression vector, resulted in low-level production of ArbA and plasmid instability. In contrast, the insertion of a

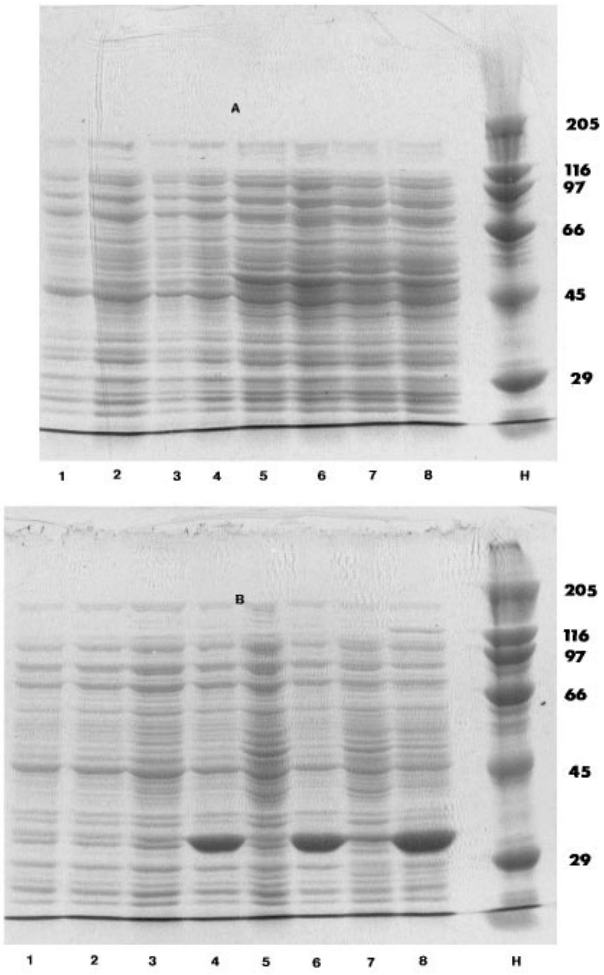

Figure 9 Expression of ArbA in *E. coli* strains harbouring pVM3 and pVM4

Parallel cultures of *E. coli* BL21(DE3) harbouring pVM3 (**A**) or pVM4 (**B**) were grown to mid-exponential phase (lanes 1 and 2). One of the two cultures of each strain was allowed to continue growing (lanes 3, 5 and 7), while the other was supplemented with 1 mM IPTG (lanes 4, 6 and 8). At 2 h (lanes 3 and 4), 4 h (lanes 5 and 6) and 19 h (lanes 7 and 8) after the mid-exponential point of the growth phase, samples were removed and CFEs were prepared, which were subjected to SDS/PAGE using a 7.5% (w/v) polyacrylamide gel together with high-M_r standards from Sigma Chemical Co. (lane H; numbers to the right are $10^{-3} \times M_r$).

truncated form of *arbA*, encoding mature ArbA, into an expression vector resulted in the production of large amounts of the arabinanase. These data suggest that the signal peptide of ArbA in some way causes inefficient synthesis of the enzyme in *E. coli*. Previous studies have shown that the signal peptide of the *P. fluorescens* subsp. *cellulosa* extracellular enzyme mannanase A adversely influenced expression of the polypeptide in *E. coli* [25]. It was suggested that the signal peptides of the *P. fluorescens* proteins were poorly recognized by *E. coli* signal peptidases, resulting in disruption of protein secretory process in the enteric bacterium, leading to cell death [25]. It is possible that high levels of full-length ArbA also disrupt protein secretion in *E. coli*.

Comparison of the amino acid sequence of ArbA with those of other glycosyl hydrolases indicated that the enzyme belongs to glycosyl hydrolase Family 43 [23]. This family contains an endo-acting arabinanase from an aerobic fungus, several bacterial xylosidases which also display some arabinofuranosidase activity against *p*-nitrophenyl α-arabinofuranoside, and a *B. subtilis* protein of unknown function. The presence of ArbA in a family containing an aerobic fungal enzyme supports the view that the repertoire of glycosidases expressed by a single organism is the result of extensive gene transfer, not only between different bacteria, but also between prokaryotes and lower eukaryotes. In addition, the presence of enzymes with distinct substrate specificities (the bacterial xylosidase/arabinofuranosidases hydrolyse xylobiose and xylan by an exo- mode of action [26,27], while ArbA and ABNA only attack arabinan [6]) supports the view that subtle changes in the structures of closely related proteins can lead to significant changes in catalytic specificity. Primary sequence alignments of ABNA, ArbA and the *B. subtilis* ORF revealed six carboxylic acid residues that are invariant in these proteins. The vast majority of glycosyl hydrolases contain two key catalytic carboxylic acid amino acids that mediate bond cleavage via general acid–base catalysis [28]. It is likely, therefore, that two of the six conserved aspartates and glutamates constitute the catalytic amino acids of the enzyme. Glycosyl hydrolases cleave glycosidic bonds via either a double- or single-displacement general acid–base mechanism, leading to retention and inversion respectively of the anomeric carbon [28]. The observation that xylanase B from *Bacillus pumilis* is an 'inverting enzyme' [29] indicates that ArbA will also catalyse cleavage of glycosidic bonds via a single-displacement mechanism. Unfortunately, attempts to monitor the stereochemical course of arabinohexaose cleavage by ArbA were unsuccessful due to the very rapid mutarotation of the furanose ring of arabinose.

Analysis of the substrate specificity of ArbA indicates that the enzyme has an exo- mode of action, successively releasing arabinotriose units from the polymeric substrate. However, this view is in contrast with the observation that the enzyme hydrolyses Red-dyed arabinan, a substrate that is normally only cleaved by endo-α-1,5-arabinanase activity. It is possible that ArbA is predominantly an exo-enzyme, but exhibits sufficient endo-activity to attack Red-dyed arabinan. Compelling evidence suggesting that ArbA exhibits significant endo- activity is derived from the observation that the ratio of reducing sugar release from, to the decrease in viscosity of, linear arabinan is very similar to that observed for the *Aspergillus* endo-α-1,5-arabinanase ABNA against the same substrate. Although it could be argued that ABNA is also an exo-acting enzyme, the initial hydrolysis of linear arabinan by the fungal enzyme to a mixture of oligosaccharides supports the view that it also has an endo-mode of action. It would appear, therefore, that ArbA exhibits both an endo- and an exo- mode of action. For this mechanism to be correct, the active site of the enzyme must undergo substantial conformational changes as it switches from an open cleft (endo- mode of action) to a tunnel-like structure which would confer an exo- mode of action on the arabinanase. Currently it is unclear whether substrates directly enter the active-site tunnels of typical exo-acting polysaccharidases, such as cellobiohydrolases, or whether the loops that form the tunnel can open, allowing random binding of the polysaccharide followed by an exo-mode of action [30], similar to the mechanism proposed for ArbA. The observation that cellobiohydrolases I and II from *Trichoderma reesei* appear to exhibit some endo-activity as they generate reducing ends within crystalline cellulose (the enzymes processively release cellobiose from the non-reducing end of cellulose chains), and the description of CenC from *Cellulomonas fimi* as a cellulase with semi-processive activity exhibiting both endo- and exo- activities, supports the view that the topology of the active sites of exo-acting enzymes is more flexible than was first thought [31,32]. Previous analysis of the mode of action of prokaryotic endo-α-1,5-arabinanases, pri-

marily from *Bacillus subtilus* [33–35], did not indicate an exo- mode of action. Thus it remains to be established whether other bacteria (and fungi) also produce arabinanases that exhibit an exo- mode of action, or whether ArbA is a unique enzyme.

We thank the Biotechnology and Biological Sciences Research Council (grant LE13/138) for financial support, and Noel Carter and Anil Wipat for releasing the *Bacillus* sequence to us prior to publication.

REFERENCES

1 Tanaka, M., Abe, A. and Uchida, T. (1981) Biochim. Biophys. Acta **658**, 377–386
2 Braithwaite, K. L. (1995) Ph.D. Thesis, University of Newcastle upon Tyne
3 Hwang, J.-T. and Suzuki, H. (1976) Agric. Biol. Chem. **40**, 2169–2175
4 Karimi, S. and Ward, O. P. (1989) J. Ind. Microbiol. **4**, 173–180
5 Van de Veen, P., Flipphi, M. J. A., Voragen, A. G. J. and Visser, J. (1992) Prog. Biotechnol. **7**, 497–500
6 Flipphi, M. J. A., Panneman, H., van der Veen, P., Visser, J. and de Graaff, L. H. (1993) Appl. Microbiol. Biotechnol. **40**, 318–326
7 Gilbert, H. J., Jenkins, G., Sullivan, D. A. and Hall, J. (1987) Mol. Gen. Genet. **210**, 551–556
8 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
9 Ferreira, L. M. A., Hazlewood, G. P. and Gilbert, H. J. (1991) Biochem. J. **279**, 793–799
10 Ferreira, L. M. A., Wood, T. M., Williamson, G., Faulds, C., Hazlewood, G. P., Black, G. W. and Gilbert, H. J. (1993) Biochem. J. **294**, 349–355
11 Braithwaite, K. L., Black, G. W., Hazlewood, G. P., Ali, B. R. S. and Gilbert, H. J. (1995) Biochem. J. **305**, 1005–1010
12 Ferreira, L. M. A., Durrant, A. J., Hall, J., Hazlewood, G. P. and Gilbert, H. J. (1990) Biochem. J. **269**, 261–264
13 Norrander, J., Kempt, T. and Messing, J. (1983) Gene **26**, 101–106
14 Chambers, S. P., Prior, S. E., Barstow, D. A. and Minton, N. P. (1988) Gene **68**, 139–149
15 Staden, R. (1980) Nucleic Acids Res. **16**, 3673–3694
16 Miller, G. L. (1959) Anal. Chem. **31**, 426–428
17 Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Béguin, P., Wakarchik, W. W., Kilburn, D. G., Miller, R. C., Warren, R. A. J. and Withers, S. G. (1992) J. Biol. Chem. **267**, 12559–12561
18 Sedmak, J. J. and Grossberg, J. R. (1977) Anal. Biochem. **79**, 544–552
19 Laemmli, U. K. (1970) Nature (London) **227**, 680–685
20 Ferreira, L. M. A., Hazlewood, G. P. and Gilbert, H. J. (1991) Biochem. J. **269**, 261–264
21 Perlman, D. and Halvorson, H. O. (1983) J. Mol. Biol. **167**, 391–409
22 Hall, J., Black, G. W., Ferreira, L. M. A., Millward–Sadler, S. J., Ali, B. R. S., Hazlewood, G. P. and Gilbert, H. J. (1995) Biochem. J. **309**, 749–756
23 Henrissat, B. and Bairoch, A. (1993) Biochem. J. **293**, 781–788
24 Hazlewood, G. P., Laurie, J. I., Ferreira, L. M. A. and Gilbert, H. J. (1992) J. Appl. Bacteriol. **72**, 244–251
25 Bolam, D. N., Hughes, N., Virden, R., Lakey, J. H., Hazlewood, G. P., Henrissat, B. and Gilbert, H. J. (1996) Biochemistry, **35**, 16195–16204
26 Gasparic, A., Martin, J., Daniel, A. S. and Flint, H. J. (1995) Appl. Environ. Microbiol. **61**, 2958–2964
27 Utt, E. A., Eddy, C. K., Keshav, K. F. and Ingram, L. O. (1991) Appl. Environ. Microbiol. **57**, 1227–1234
28 Withers, S. G. (1995) in Carbohydrate Bioengineering (Petersen, S. B., Svensson, S. and Pedersen, S., eds.), pp. 97–113, Elsevier Science B. V., The Netherlands
29 Kasumi, T., Tsumuraya, Y., Brewer, C. F., Kersters-Hilderson, H., Claeyssens, M. and Hehre, E. J. (1987) Biochemistry **26**, 3010–3016
30 Henrissat, B. and Davies, G. (1995) Structure **3**, 853–859
31 Stahlberg, J., Johansson, G. and Pettersson, G. (1993) Biochim. Biophys. Acta **1157**, 107–113
32 Tomme, P., Kwan, E., Gilkes, N. R., Kilburn, D. G. and Warren, R. A. J. (1996) J. Bacteriol. **96**, 4216–4223
33 Yoshihara, O. and Kaji, A. (1983) Agric. Biol. Chem. **47**, 1935–1940
34 Weinstein, L. and Albersheim, P. (1979) Plant Physiol. **63**, 425–432
35 Kaji, A. and Saheki, T. (1975) Biochim. Biophys. Acta **410**, 354–360

Received 1 October 1996/19 November 1996; accepted 3 December 1996