Purification, characterization and gene cloning of two α -L-arabinofuranosidases from *Streptomyces chartreusis* GS901

Noriki MATSUO^{*1,2}, Satoshi KANEKO^{†2}, Atsushi KUNO^{‡§}, Hideyuki KOBAYASHI[†] and Isao KUSAKABE^{||}

^{*}Central Research Laboratory, Godo Shusei Co. Ltd., Kamihongo, Matsudo, Chiba 271-0064, Japan, [†]National Food Research Institute, Ministry of Agriculture, Forestry, and Fisheries, Tsukuba, Ibaraki 305-8642, Japan, [‡]Department of Material and Biological Chemistry, Faculty of Science, Yamagata University, Yamagata 990-8560, Japan, [§]National Institute for Advanced Interdisciplinary Research, 1-1-4 Higashi, Tsukuba 305-8562, Japan, and ^{||}Institute of Applied Biochemistry, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan

α -L-Arabinofuranosidases I and II were purified from the culture filtrate of *Streptomyces chartreusis* GS901 and were found to have molecular masses of 80 and 37 kDa and pI values of 6.6 and 7.5 respectively. Both enzymes demonstrated slight reactivity towards arabinoxylan and arabinogalactan as substrates but did not hydrolyse gum arabic or arabinoxylo-oligosaccharides. α -L-Arabinofuranosidase I hydrolysed all of the α -linkage types that normally occur between two α -L-arabinofuranosyl residues, with the following decreasing order of reactivity being observed for the respective disaccharide linkages: α -(1 $\rightarrow$ 2) α -(1 $\rightarrow$ 3) α -(1 $\rightarrow$ 5). This enzyme cleaved the (1 $\rightarrow$ 3) linkages of the arabinosyl side-chains of methyl 3,5-di-*O*- α -L-arabinofuranosyl- α -L-arabinofuranoside in preference to the (1 $\rightarrow$ 5) linkages. α -L-Arabinofuranosidase I hydrolysed approx. 30% of the arabinan but hydrolysed hardly any linear arabinan. In contrast, α -L-Arabinofuranosidase II hydrolysed only (1 $\rightarrow$ 5)-arabinofuranobioside among the regioisomeric methyl arabinobiosides and did

not hydrolyse the arabinotrioside. Linear 1 $\rightarrow$ 5-linked arabinan was a good substrate for this enzyme, but it hydrolysed hardly any of the arabinan. Synergism between the two enzymes was observed in the conversion of arabinan and debranched arabinan into arabinose. Complete amino acid sequencing of α -L-arabinofuranosidase I indicated that the enzyme consists of a central catalytic domain that belongs to family 51 of the glycoside hydrolases and additionally that unknown functional domains exist in the N-terminal and C-terminal regions. The amino acid sequence of α -L-arabinofuranosidase II indicated that this enzyme belongs to family 43 of the glycoside hydrolase family and, as this is the first report of an *exo*-1,5- α -L-arabinofuranosidase, it represents a novel type of enzyme.

Key words: arabinan degradation, family 43, family 51, glycoside hydrolase.

INTRODUCTION

L-Arabinose residues are widely distributed in plant cell walls, where they are present in polymers such as arabinans, arabinoxylans, arabinogalactans and arabinogalactan proteins. The arabinans are composed mainly of α -L-arabinofuranosyl residues arranged into (1 $\rightarrow$ 5)-linked chains. Varying numbers of these residues are substituted at the O-2 and/or O-3 position by additional α -L-arabinofuranosides [1]. L-Arabinose is a major component of arabinan, which, although it has a sweet taste, is not readily absorbed by the body [2]. Seri et al. reported that it selectively inhibited the sucrase activity of pig intestinal mucosa in an uncompetitive manner [2]. They also showed that, in mice, L-arabinose suppressed the increase of blood glucose in a dose-dependent manner after the ingestion of sucrose, but did not suppress this increase after starch ingestion [2]. Furthermore, Sanai et al. reported that L-arabinose delays and decreases the digestion, absorption and net energy derived from sucrose when both are ingested simultaneously [3]. Thus, L-arabinose is useful in preventing postprandial hyperglycaemia in diabetic individuals when foods containing sucrose are ingested.

Arabinan-degrading enzymes have been classified on the basis of their mode of action, i.e. *endo*-acting or *exo*-acting. The

arabinan-degrading enzymes that act in an *endo*-fashion are called *endo*-1,5- α -L-arabinanases (EC 3.2.1.99) and those that act in an *exo*-fashion are called α -L-arabinofuranosidases [α -L-AFases (EC 3.2.1.55)]. The α -L-AFases have been extensively studied and reviewed [4–6]. α -L-AFases have been classified by Beldman et al. [6] on the basis of their substrate specificity, i.e. those that have very little or no activity on the arabinose-containing polymers have been classified as Arafur A and those active on polymers have been classified as Arafur B. However, this classification is too broad to define the substrate specificity of the α -L-AFases fully. Some of the α -L-AFase genes have been cloned and the amino acid sequences of the respective enzymes have been determined. Depending on their amino acid sequences, α -L-AFases have been classified into four of the glycoside hydrolase families (families 43, 51, 54 and 62). This classification is based on the amino acid sequence at their active sites rather than their substrate specificities [7,7a].

We have previously purified some α -L-AFases and elucidated their substrate specificities towards structurally defined substrates [8–14]. The substrate specificities of the α -L-AFases studied were so broad that we were not able to define them clearly. However, we have recently found very interesting α -L-AFases that have strict substrate specificity. These seem to be the most suitable

Abbreviations used: α -L-AFase, α -L-arabinofuranosidase (EC 3.2.1.55); PNP- α -L-Araf, *p*-nitrophenyl α -L-arabinofuranoside; PNP- α -L-Arap, *p*-nitrophenyl α -L-arabinopyranoside; PNP- β -D-Galp, *p*-nitrophenyl β -D-galactopyranoside; PNP- β -D-Xylp, *p*-nitrophenyl β -D-xylopyranoside; PNP, *p*-nitrophenol.

¹ To whom correspondence should be addressed (e-mail 23209331@people.or.jp).

² The first two authors contributed equally to this work.

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession numbers AB023625 (α -L-AFase I) and AB023626 (α -L-AFase II).

enzymes for studying structure–function relationships of α -L-AFases. In the present paper we report the purification, characterization, substrate specificities and the complete amino acid sequences of two different α -L-AFases from *Streptomyces chartreusis* GS901. One of these enzymes is novel because it hydrolyses the 1,5 linkages of arabinofuranose in an exclusively *exo*-manner.

MATERIALS AND METHODS

Substrates

p-Nitrophenyl α -L-arabinofuranoside (PNP- α -L-Araf), *p*-nitrophenyl α -L-arabinopyranoside (PNP- α -L-Arap), *p*-nitrophenyl β -D-galactopyranoside (PNP- β -D-Galp), *p*-nitrophenyl β -D-xylopyranoside (PNP- β -D-Xylp), and larch wood arabinogalactan were purchased from Sigma Chemical Company (St. Louis, MO, U.S.A.). Gum arabic was obtained from Wako Pure Chemical Industries (Osaka, Japan). Nihon Syokuhin Kakoh Co. (Fuji, Japan) supplied corn hull arabinoxylan. Debranched arabinan (linear 1 $\rightarrow$ 5-linked arabinan) was obtained from Megazyme Pty. (North Rocks, Australia). Sugar-beet arabinan and arabinoxyloligosaccharides were prepared by methods described previously [15–17]. Methyl 2-*O*-, methyl 3-*O*- and methyl 5-*O*- α -L-arabinofuranosyl- α -L-arabinofuranosides (arabinofuranobiosides) and methyl 3,5-di-*O*- α -L-arabinofuranosyl- α -L-arabinofuranoside (arabinofuranotriose) were synthesized by methods described previously [18,19].

α -L-AFase assay and measurement of protein

Each assay mixture contained 0.5 ml of a 2 mM PNP- α -L-Araf solution, 0.4 ml of McIlvaine buffer (0.2 M Na₂HPO₄/0.1 M citric acid) and 0.1 ml of enzyme solution. McIlvaine buffer, pH 5.5, and McIlvaine buffer, pH 7.0, were used for the assays of α -L-AFase I and α -L-AFase II respectively. The reactions were performed at 40 °C for 10 min, after which they were stopped by the addition of 1.0 ml of 0.2 M Na₂CO₃; the amount of *p*-nitrophenol (PNP) released was then determined at 408 nm (referred to hereafter as the standard method). One unit of enzyme activity is defined as the amount of enzyme that released 1 μ mol of PNP per min from PNP- α -L-Araf under these conditions. The protein concentration was determined by the method described by Lowry et al. [20] with BSA as a standard.

Micro-organisms and preparation of the crude enzyme

The strain GS901 was isolated from soil in Japan and was identified as *Streptomyces chartreusis* (results not shown). *S. chartreusis* GS901 was cultured in 500 ml shaking flasks containing 100 ml of medium consisting of 10 g/l arabinan, 3 g/l yeast extract, 23 mM (NH₄)₂HPO₄, 7.4 mM KH₂PO₄, 2.0 mM MgSO₄·7H₂O, 36 nM FeSO₄·7H₂O, 51 nM MnCl₂·4H₂O and 35 nM ZnSO₄·7H₂O. The pH was adjusted to 7.2. Preparations were precultivated on a rotary shaker (210 rev./min) at 30 °C for 72 h; 20 ml of this culture broth was used to inoculate 1 litre of the same medium in a 5-litre shaking flask. Cultivation was performed after incubation at 30 °C and 210 rev./min for 96 h. The culture broth was then filtered through filter paper (Toyo Roshi no. 2; Toyo Roshi Co., Tokyo, Japan). (NH₄)₂SO₄ was added to 70% satn. and the resulting mixture was kept at 4 °C for 16 h. The precipitate was collected by centrifugation (10000 g, 20 min) and dissolved in a small amount of 20 mM phosphate buffer, pH 6.5. The solution thus obtained was used as the crude enzyme preparation.

Purification of α -L-AFases

All purification procedures were performed at 6 °C.

Step 1

The crude enzyme was dialysed against 20 mM phosphate buffer, pH 6.5; the dialysed enzyme was applied to a column (25 mm $\times$ 190 mm) of CM-Sepharose CL-6B (Pharmacia, Uppsala, Sweden) equilibrated with 20 mM phosphate buffer, pH 6.5. The column was then washed with the same buffer and the enzyme was eluted from the column with a linear gradient of 0–0.25 M NaCl solution (total volume 500 ml) at a flow rate of 15 ml/h. The eluate was fractionated into 5 ml portions. α -L-AFase was separated into two different fractions: fraction I (α -L-AFase I, fraction tubes 52–64) and fraction II (α -L-AFase II, fraction tubes 78–84). Fractions I and II were concentrated by ultrafiltration (Nano Spin Plus membrane; Gelman Sciences Japan, Tokyo, Japan).

Step 2a

The concentrated α -L-AFase I was applied to a HiPrep Sephacryl S-200 HR16/60 column (16 mm $\times$ 600 mm) equilibrated with 0.15 M NaCl in 20 mM phosphate buffer, pH 6.0. An elution flow rate of 30 ml/h was used. The eluate was collected in 0.5 ml fractions. The fractions containing α -L-AFase activity (fraction tubes 9–25) were combined and dialysed against 20 mM phosphate buffer, pH 7.5, and subjected to step 3.

Step 2b

The concentrated α -L-AFase II was also applied to a HiPrep Sephacryl S-200 HR16/60 column (16 mm $\times$ 600 mm) equilibrated with 0.15 M NaCl in 20 mM phosphate buffer, pH 6.0. An elution flow rate of 30 ml per h was used. The eluate was collected in 0.5 ml fractions. The fractions with α -L-AFase activity (fraction tubes 73–88) were combined and dialysed against deionized water. The final preparation thus obtained was used as purified α -L-AFase II.

Step 3

The dialysed α -L-AFase I was applied to a Mono Q HR 5/5 column (5 mm $\times$ 50 mm) pre-equilibrated with 20 mM phosphate buffer, pH 7.5. After the column had been washed with the same buffer, the enzyme was eluted with a linear gradient of NaCl solution from 0 to 0.25 M (total volume 30 ml) at a flow rate of 0.5 ml/min. The eluate was collected in 0.5 ml portions. The active fractions (fractions 37–39) were combined and dialysed against deionized water. The final preparation thus obtained was used as purified α -L-AFase I.

SDS/PAGE

SDS/PAGE was performed with a 5–20% gradient gel (ATTO) as described by Laemmli [21]. The protein was stained with Coomassie Brilliant Blue R-250 and then destained with 7% (v/v) acetic acid. The molecular-mass of the enzyme was measured by SDS/PAGE with molecular-mass markers (Bio-Rad SDS/PAGE standard low; Bio-Rad). The N-terminal amino acid sequences of α -L-AFases I and II were determined with an

HP G1005A Protein Sequence System. Isoelectric focusing was performed as described previously [22]. Thus the protein was stained with Coomassie Brilliant Blue R-250 and an isoelectric focusing calibration kit (pH 3–10; Pharmacia) was used for calibration.

Enzymic properties

The effects of pH on the activity and stability of α -L-AFases were investigated with a series of McIlvaine buffers from pH 4.0 to 8.0 and Atkins–Pantin buffers (0.2 M boric acid/0.2 M KCl/0.2 M Na₂CO₃) from pH 8.0 to 10.5. The activities of the α -L-AFases were otherwise assayed under the conditions described for the standard method. For determination of the pH stability of the α -L-AFases, the enzymes were preincubated at various pH values in the absence of substrate at 30 °C for 120 min and the residual activity was then assayed by the standard method. The effects of temperature on the activity of the α -L-AFases were determined by using a series of water baths from 30 to 55 °C. With the exception of temperature, the assay conditions were the same as described for the standard method. For the temperature-stability measurement of the α -L-AFases, the enzymes were preincubated at various temperatures at pH 7.0 for 120 min and the residual activity was determined by the standard method.

Substrate specificities

For the determination of substrate specificities, various PNP-glycosides were used as substrates; in these cases the assay method was identical with that described for PNP- α -L-Araf. For non-chromogenic substrates, a solution was prepared containing 0.1 ml of α -L-AFase solution (0.04 unit of α -L-AFase I and 0.025 unit of α -L-AFase II), 0.4 ml of McIlvaine buffer (pH 5.5 for α -L-AFase I; pH 7.0 for α -L-AFase II; pH 6.5 for α -L-AFase I and II together) and 0.5 ml of 2% (w/v) substrate (arabinoxylan, arabinogalactan, gum arabic, arabinan, debranched arabinan, arabinoxylo-oligosaccharides or arabinofuranobiosides). After incubation for 0, 2, 4, 6, 8 or 24 h at 40 °C, the reaction was stopped by boiling for 5 min. The released L-arabinose was determined by HPLC under the following conditions: column, Shodex Sugar KS-801 (8 mm $\times$ 300 mm; Showa Denko, Tokyo, Japan); mobile phase, water; flow rate, 1.0 ml/min; column temperature, 80 °C. The eluted compounds were detected by refractive index. For the hydrolyses of arabinofuranotriose, 0.1 ml of 1% (w/v) arabinofuranotriose was used in the reaction mixture; after incubation for 8 or 24 h at 40 °C, the products were applied to a C₁₈ column (YMC-Pack ODS-AQ, 4.6 mm $\times$ 150 mm; YMC, Kyoto, Japan) and eluted with 1.7% CH₃CN at a flow rate of 0.5 ml per min. Sugars eluted from the column were detected with a refractive-index detector.

Cloning of α -L-AFase genes

A simple long-range inverse PCR method [23] improved for high G+C content genes [24] was used for the cloning of full-length α -L-AFase I and II genes (*afi* and *afii* respectively) without the construction of a genomic library. Genomic DNA was digested with *Bam*HI and *Bgl*II (for *afi*) or *Sal*I (for *afii*) and then self-ligated. An LA PCR Kit version 2 (Takara, Kyoto, Japan) was then used to amplify the long and GC-rich fragments of genomic DNA. DNA fragments were amplified with the following sense and anti-sense primers: sense primer for *afi*, 5'-GAG GTG AGG GGC GTG TAG-3'; anti-sense primer for *afi*, 5'-CTG CTG AGC GAC GAC TTC-3'; sense primer for *afii*, 5'-CCA

CGA CCC AGA TAC GGA TGC-3'; anti-sense primer for *afii*, 5'-CCT CGA CGC CAC CAC CTT CAC-3'. Each of the 25 amplification cycles consisted of a denaturation step at 98 °C for 1 min followed by annealing and primer extension steps at 72 °C for 10 min. After the analysis of the sequence of these fragments, we designed primers for amplification of the full-length *afi* and *afii* genes as follows: sense primer for *afi*, 5'-CCA TGG CAC GCA TCC GCT GGA GAT ACG GC-3'; anti-sense primer for *afi*, 5'-GGA TCC CCT CTG CTT GAG CCG CAG GAA CG-3'; sense primer for *afii*, 5'-CCA TGG GTA CCC GAG AGG CAG TCC GT-3'; anti-sense primer for *afii*, 5'-GGA TCC GGG GTC GCC GAC GAT CTC CTT-3'. The products of PCR obtained with these primers were subcloned into plasmid pCR2.1.

RESULTS

Purification and enzymic properties of α -L-AFases

Table 1 summarizes the purification procedures for α -L-AFases I and II from *S. chartreusis* GS901. Almost half of the α -L-AFase activity was not absorbed by the CM-Sepharose; however, two peaks, each containing α -L-AFase activity, were obtained by subsequent elution of the CM-Sepharose by NaCl solution. In the present study we purified the two enzymes that initially bound to the CM-Sepharose. Each of the purified enzymes could be resolved as a single band by SDS/PAGE when detected by staining with Coomassie Brilliant Blue R-250 (Figure 1). The molecular mass of α -L-AFases I and II were estimated to be 80 and 37 kDa respectively by SDS/PAGE. The pI values for α -L-AFases I and II were 6.6 and 7.5 respectively. The N-terminal amino acid sequence of α -L-AFase I was EDVTDYSITVDPA-AKGAAIDDTMYGVFFEDINRAADGGLYAELVQN and that for α -L-AFase II was APRPAPYENPLVRQRADPHIH-RHTDGRYY. The amino acid sequences of α -L-AFase I and II showed no similarity to each other. α -L-AFase I shared some similarity with a putative α -L-AFase from *S. coelicolor* (GenBank accession number AL031541), α -L-AFase I from *Bacteroides ovatus* (U15178) and α -L-AFase A from *Aspergillus niger* (L29005), which belongs to glycoside hydrolase family 51. No similarity of the N-terminal sequence of α -L-AFase II was found to any other α -L-AFase.

The pH and temperature activity and stability profiles of α -L-AFase I and α -L-AFase II were investigated. α -L-AFase I achieved maximal activity at pH 5.5. The enzyme was slowly inactivated above pH 8.5 and below pH 5.5. The maximum activity for α -L-AFase I occurred at 55 °C but inactivation started to occur at temperatures above 45 °C. α -L-AFase II achieved maximal activity at pH 7.0 and became slowly inactivated at pH values above 9.0 and below 5.0. The maximum activity for α -L-AFase II occurred at 50 °C but inactivation started to occur at temperatures above 40 °C.

Substrate specificities of α -L-AFases

The substrate specificities of α -L-AFases I and II from *S. chartreusis* GS901 were investigated by using PNP-glycosides, arabinose-containing polysaccharides, arabinoxylo-oligosaccharides and synthetic arabino-oligosaccharides as substrates. Both enzymes hydrolysed PNP- α -L-Araf but did not hydrolyse PNP- α -L-Arap, PNP- β -D-Xylp and PNP- β -D-Galp (results not shown). The results for the hydrolysis of arabinan, debranched arabinan, arabinoxylan, arabinogalactan and gum arabic by the α -L-AFases from *S. chartreusis* GS901 are shown in Table 2. Both enzymes released quite small amounts of arabinose from arabinan, linear arabinan, arabinoxylan and arabinogalactan with differing rates

Table 1 Summary of the purification of α -L-AFases I and II from *S. chartreusis* GS901

Step	Total activity (units)	Total protein (mg)	Specific activity (units/mg)	Recovery (%)	Purification (fold)
Culture filtrate	117	2042	0.0573	100	1
70% (NH ₄) ₂ SO ₄	94.9	436	0.218	81.1	3.8
CM Sepharose					
α -L-AFase I	16.8	8.01	2.10	14.4	36.3
α -L-AFase II	1.7	1.34	1.27	1.45	21.9
Sephacryl S-200					
α -L-AFase I	10.3	1.13	9.12	8.80	158
α -L-AFase II	1.44	0.456	3.16	1.23	54.6
Mono Q					
α -L-AFase I	7.69	0.742	10.4	6.57	182

Figure 1 SDS/PAGE of purified enzymes

Lanes 1 and 4, standard proteins (1 μ g each), including phosphorylase b (molecular mass 97 kDa), bovine serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa) and lysozyme (14.4 kDa); lane 2, purified α -L-AFase II (1 μ g); lane 3, purified α -L-AFase I (1 μ g).

Table 2 Activities of α -L-AFases from *S. chartreusis* GS901 with arabinose-containing polysaccharides

Substrate	Rate of hydrolysis (%)	
	α -L-AFase I	α -L-AFase II
Arabinoxylan	5.2	2.6
Arabinogalactan	2.2	2.4
Gum arabic	0	0
Arabinan	35.7	4.3
Debranched arabinan	0.9	54.1

of hydrolysis; however, they were ineffective against gum arabic. In addition, neither α -L-AFase I nor α -L-AFase II hydrolysed *O*- α -L-arabinofuranosyl-(1 $\rightarrow$ 3)-*O*- β -D-xylopyranosyl-(1 $\rightarrow$ 4)-D-xylopyranose or *O*- β -D-xylopyranosyl-(1 $\rightarrow$ 4)-[*O*- α -L-arabinofuranosyl-(1 $\rightarrow$ 3)]-*O*- β -D-xylopyranosyl-(1 $\rightarrow$ 4)-D-xylopyranose (results not shown). The degrees of hydrolysis of three

arabinofuranobiosides by the α -L-AFases were monitored over time (Figure 2). The arabinofuranobiosides were hydrolysed to arabinose and methyl α -L-arabinofuranoside by α -L-AFase I with the following decreasing order of reactivity: (1 $\rightarrow$ 2) > (1 $\rightarrow$ 3) > (1 $\rightarrow$ 5)-methyl α -L-arabinofuranobiosides (Figure 2A). In contrast, α -L-AFase II hydrolysed the (1 $\rightarrow$ 5)-methyl α -L-arabinofuranobioside, but not the (1 $\rightarrow$ 2) and (1 $\rightarrow$ 3) linkages (Figure 2B). In arabinofuranotrioxide, α -L-AFase I hydrolysed the (1 $\rightarrow$ 3) linkage to a greater extent than the corresponding (1 $\rightarrow$ 5) linkage, whereas the α -L-AFase II did not hydrolyse the trisaccharide (results not shown). The time course of the hydrolysis of arabinan by α -L-AFase I and II is shown in Figure 3(A). The upper limits for the conversion into arabinose by α -L-AFase I were 34.1% by α -L-AFase I and 4.3% by α -L-AFase II. Figure 3(B) shows the degree of hydrolysis of debranched arabinan by α -L-AFases I and II as it changed with time. The upper limits for the conversion into arabinose by α -L-AFase I and II were 0.7% and 54.1% respectively. When both enzymes were used together to hydrolyse arabinan and debranched arabinan, the degree of hydrolysis was higher than that obtained with the individual enzymes (Figure 3).

Cloning and nucleotide sequencing of the α -L-AFase I and α -L-AFase II genes of *S. chartreusis* GS901

In an attempt to elucidate structure-function relationships of α -L-AFases I and II from *S. chartreusis* GS901, we cloned the genes for these enzymes (designated *afi* and *afii* respectively). After the digestion of α -L-AFases I and II with V8 peptidase, the following sequences were obtained: α -L-AFase I, NTMVQAGDPSWHDYDLHVKATKKSGKEGFLVAFGV; α -L-AFase II, MANPWTLTGPIRLSTPEYDWE. Primers were designed on the basis of the deduced sequences of the peptide fragments. The products of the PCR were almost 1.9 kbp (*afi*) and 550 bp (*afii*) long and were used as probes for genomic Southern hybridization [25]; a 3.5 kbp *Bam*HI-*Bgl*II double-digested fragment (*afi*) and a 2.3 kbp *Sal*I fragment (*afii*) were detected with these probes. For long-range inverse PCR, the genomic DNA was digested with *Bam*HI and *Bgl*II (for *afi*) or *Sal*I (for *afii*), then self-ligated. High-melting-temperature primers were used for the amplification of G+C-rich DNA because the G+C content of the genomic DNA of *Streptomyces* species is approx. 70% [26]. After amplification, DNA fragments of 3.5 kbp (*afi*) and 2.3 kbp (*afii*) were obtained and these fragments contained unknown upper and lower regions of the *afi* and *afii* genes. From the sequence data of these fragments, a set of sense and anti-sense primers were designed and amplified to give full-length *afi* (AB023625) and *afii* (AB023626) genes. Single 2.4 kbp (*afi*) or 1.0 kbp (*afii*) fragments were obtained that corresponded to encoding of the entire *afi* or *afii* gene, as predicted from the molecular masses of α -L-AFases I and II. The PCR products were subcloned into pCR2.1 plasmids. Thus, by using appropriate primers with relatively high melting temperatures, the cloning of the *afi* and *afii* genes was successful by inverse PCR in spite the high G+C contents and even though the length of *afi* was greater than 2.4 kbp. The single open reading frames of α -L-AFase I and II were 2475 bp and 984 bp long and their G+C contents were 68.6% and 69.5% respectively. The ORFs of α -L-AFases I and II encoded 825 and 328 amino acid residues respectively. A site for the cleavage of a signal peptide (AHA/E or A), as described previously [24], was identified at the N-terminal end of mature α -L-AFases I and II. Glycoside hydrolase family analysis of the sequences indicated that α -L-AFase I from *S. chartreusis* GS901 has a central domain

Figure 2 Hydrolysis of arabinofuranobiosides

(A) α -L-AFase I. (B) α -L-AFase II. Symbols: ■, methyl 2-O- α -L-arabinofuranosyl- α -L-arabinofuranoside; ●, methyl 3-O- α -L-arabinofuranosyl- α -L-arabinofuranoside; ▲, methyl 5-O- α -L-arabinofuranosyl- α -L-arabinofuranoside. The experimental conditions used are described in the Materials and methods section. After incubation at 30 °C for 0, 2, 4, 6, 8 or 24 h, the reaction mixtures were analysed by HPLC.

(approx. residues 180–580) corresponding to family 51 and that α -L-AFase II belongs to family 43.

Similarities between α -L-AFases and other enzymes in terms of their amino acid sequences

The deduced amino acid sequences of α -L-AFases I and II were compared with sequences in the protein and nucleic acid database BLAST (on a web site of the National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/BLAST/>). α -L-AFase I resembled the following family 51 α -L-AFases, in order of decreasing similarity: putative α -L-AFase from *S. coelicolor* (GenBank accession number AL031541); ASDI from *Bact. ovatus* (U15178); ABFA from *A. niger* (L29005); α -L-AFase from *Thermotoga maritima* (AE001710); ABFA from *S. lividans* (U04630); α -L-AFase from *Arabidopsis thaliana* (AF149413); ARFII from *Cytophaga xylanolytica* (AF028019); α -L-AFase 1 from *Bacillus subtilis* (X89810); ABFA from *Bacillus* sp. D3 (Y16849); α -L-AFase 2 from *Bacillus subtilis* (Z99118, Z75208); ARFI from *Cytophaga xylanolytica* (AF028018); α -L-AFase from *Clostridium stercoarum* (AF002664); ASDII from *Bact. ovatus* (U15179). The α -L-AFase I with the strongest similarity (*S. coelicolor* putative α -L-AFase) was 85.2% similar to α -L-AFase I over the entire amino acid sequence. The residues from almost 250 to 450 showed similarity to all of the other family 51 α -L-AFases. Two glutamic residues, inferable as the acid–base catalyst and the catalytic nucleophile in the double-displacement mechanism of the retaining glycosidases, were found to be present in the strongly conserved regions (Figure 4, upper panels). In contrast, α -L-AFase II displayed slight similarity to the following family 43 glycanases in decreasing order of similarity: β -xylosidase from *Cochliobolus carbonum* (AF095243); arabinanase A from *Pseudomonas fluorescens* (Y10458); xylosidase/arabinosidase from *Selenomonas ruminantium* (AF040720). The region shown in Figure 4 (lower panel) was especially conserved in family 43 enzymes.

DISCUSSION

The complete amino acid sequences indicate that α -L-AFases I and II from *S. chartreusis* GS901 belong to families 51 and 43 of the glycoside hydrolases respectively. Family 51 includes thirteen

types of α -L-AFase and one type of endoglucanase (EC 3.2.1.4). This family is classified into the clan GH-A, which is the major clan of glycoside hydrolases, along with families 1, 2, 5, 10, 17, 26, 30, 35, 39, 42 and 53, on the basis of structural and mechanistic information. The enzymes belonging to clan GH-A are inferred as having a $(\beta/\alpha)_8$ -barrel structure and their enzyme mechanism is known as retaining owing to the double-displacement mechanism, which is catalysed by two glutamic residues acting as an acid–base pair and a nucleophile. The amino acid sequence of α -L-AFase I indicated that the residues from approx. 180 to 580 of α -L-AFase I belong to family 51. α -L-AFase I consists of 825 residues, indicating that α -L-AFase I possesses additional functional domains in both the N-terminal and the C-terminal regions. However, these domains did not show any similarities to known substrate-binding domains observed in many types of glycanases [27]. The similarity search of this C-terminal region (residues 580–825) showed a weak similarity to neuraminidase (M61732) and trans-sialidase (D50684, L38463, L38457, L38456) from *Trypanosoma cruzi*. These are enzymes using carbohydrate-based substrates, implying that this region should have a common or similar function in all of these enzymes; it might represent a novel kind of substrate-binding domain. In contrast, two glutamic residues (E³⁸⁰ and E⁴⁶⁷), which are completely conserved in the family 51 α -L-AFases, were found in the region considered to be the catalytic domain (Figure 4, upper panels). These glutamic residues are possibly responsible for catalysing the α -L-AFase activity, as reported previously [28].

The substrate specificity of α -L-AFase I from *S. chartreusis* GS901 can be summarized as follows: (1) the enzyme acts on arabinan but not on other arabinose-containing polysaccharides; (2) the enzyme hydrolyses arabinan but leaves most of the polymer intact, including most of the main-chain residues and the arabinose side chains; (3) the enzyme does not hydrolyse arabinoxylo-oligosaccharides; (4) the enzyme hydrolyses α -arabinofuranobiosides in the following decreasing order of reactivity: α -(1 $\rightarrow$ 2) > α -(1 $\rightarrow$ 3) > α -(1 $\rightarrow$ 5); (5) the enzyme hydrolyses the (1 $\rightarrow$ 3) linkage of arabinofuranotrioside in preference to the (1 $\rightarrow$ 5) linkage. The substrate specificity was quite similar to α -L-AFases from *A. niger* 5-16 [8,11] and *Streptomyces diastatochromogenes* 065 [12,29,30]. However, α -L-AFase I from *Aspergillus awamori* IFO4033 (putative family 51

Figure 3 Hydrolysis of arabinan and debranched arabinan

(A) Arabinan. (B) Debranched arabinan. Symbols: ■, α -L-AFase I; ●, α -L-AFase II; ▲, combination hydrolysis by α -L-AFase I and α -L-AFase II. The experimental conditions used are described in the Materials and methods section.

Figure 4 Highly conserved region in the sequences of α -L-AFases of families 51 and 43

(A) Conserved region of α -L-AFase I along with family 51 enzymes; (B) conserved region of α -L-AFase II along with family 43 enzymes. The strongly conserved sequences are aligned with those of α -L-AFases I and II. Identical and similar amino acid residues are enclosed in black boxes and open boxes respectively. It is inferred that family 51 α -L-AFases have two catalytic glutamic residues. Glutamic residues indicated by an asterisk are inferred as being the catalytic residues in family 51 enzymes [28]. The conserved region of α -L-AFase II is also identified along with family 43 *endo*-arabinanases. The amino acid residues indicated by an asterisk are possible catalytic residues. Abbreviations: AF I, the α -L-AFase I from *S. chartreusis* (AB023625); ATAF, α -L-AFase from *Arabidopsis thaliana* (AF149413); ASDI, ASDI from *Bact. ovatus* (U15178); ABFA (*A. niger*), ABFA from *A. niger* (L29005); SCAF, putative α -L-AFase from *S. coelicolor* (AL031541); AF II, α -L-AFase from *S. chartreusis* (AB023626); ABNA (*A. niger*), *endo*-arabinanase A from *A. niger* (L23430); BSABN, *endo*-arabinanase from *B. subtilis* (D85132); ABNA (*B. subtilis*), ABNA from *B. subtilis* (Z75208); ARBA, arabinanase A from *Ps. fluorescens* (Y10458).

from its N-terminal amino acid sequence) showed a different substrate specificity from the above type of α -L-AFases towards arabino-oligosaccharides [13]: it hydrolyses α -arabinofuranobiosides in the following decreasing order of reactivity: α -(1 $\rightarrow$ 5) > α -(1 $\rightarrow$ 2) > α -(1 $\rightarrow$ 3). In addition, this enzyme hydrolyses the (1 $\rightarrow$ 5) linkage of arabinofuranotrioside in preference to the (1 $\rightarrow$ 3) linkage. These results suggest that the substrate specificities of α -L-AFases belonging to family 51 are variable. It is apparent that the recognition of substrates by family-51 α -L-AFases is different; nevertheless, the enzymes have similar reactivity towards arabinose-containing polysaccharides. Unfortunately, it is not possible to discuss substrate specificities

relating to more specific structural aspects of these enzymes in any greater detail, because a crystal structure of any member of this family of enzymes has yet to be reported.

In contrast, at present family 43 of glycoside hydrolases includes 26 kinds of enzyme, including β -xylosidase (EC 3.2.1.37), α -L-AFase, bifunctional β -xylosidase/ α -L-AFase, *endo*-arabinanase and β -xylanase (EC 3.2.1.8). This family is classified into clan GH-F together with family 62. These enzymes are known as catalysing by the inverting mechanism that they perform. A preliminary crystal structure of arabinanase A obtained from *Ps. fluorescens* has been reported [31]. However, detailed structural information is not yet available. The substrate

specificity of α -L-AFase II from *S. chartreusis* GS901 can be summarized as follows. The enzyme acts completely differently from the types described above: (1) it cannot act on arabinose-containing polysaccharides, except for linear (1 $\rightarrow$ 5)-linked arabinan; (2) among regioisomeric α -L-arabinobiosides, it hydrolyses only (1 $\rightarrow$ 5)-linked disaccharide; (3) it cannot act on arabinoxylo-oligosaccharides. Thus the enzyme acts only on the (1 $\rightarrow$ 5) linkages of arabinofuranose and should therefore be called an *exo*-1,5- α -L-arabinofuranosidase, which is a novel kind of α -L-AFase because the enzyme releases arabinose only from 1 $\rightarrow$ 5-linked arabinofuranosyl polysaccharides or oligosaccharides. The *endo*-arabinanases also prefer 1,5-arabinan over arabinan as their substrate. Therefore the finding that α -L-AFase II is classified into family 43 of the glycoside hydrolases, together with arabinanases, is quite reasonable because both enzymes preferentially hydrolyse linear 1,5-arabinan. The modes of action of arabinanases have been studied and the known arabinanase from *A. niger* is exclusively of the *endo*-type [32]. In contrast, arabinanase A from *Ps. fluorescens* hydrolyses linear 1,5-arabinan in both an *endo*-manner and an *exo*-manner [33]. Therefore an investigation of the structure–function relationships of α -L-AFase II from *S. chartreusis* GS901 would be of great interest in understanding the differences between the mechanisms involved in the *endo*- and *exo*-hydrolases.

The glycoside hydrolase families are classified by their amino acid similarities. It is therefore reasonable to try to relate this classification to their substrate specificities. However, at present the sequence data and structural information available for α -L-AFases are quite limited in comparison with those for other glycosidases. However, α -L-AFases I and II from *S. chartreusis* GS901 seem to be good models for investigating structure–function relationships of the α -L-AFases. In the future we wish to provide additional information that should be of use in further determinations of structure–function relationships of the α -L-AFases.

We thank Dr. Bernard Henrissat (AFMB-CNRS IFRI, Marseille, France) for glycoside hydrolase families analysis of the sequence of α -L-AFases I and II.

REFERENCES

- Bacic, A., Harris, P. J. and Stone, B. A. (1988) in *The Biochemistry of Plants*, vol. 14 (Preiss, J., ed.), pp. 297–371, Academic Press, San Diego.
- Seri, K., Sanai, K., Matsuo, N., Kawakubo, K., Xue, C.-Y. and Inoue, S. (1996) *Metab. Clin. Exp.* **45**, 1368–1374.
- Sanai, K., Seri, K. and Inoue, S. (1997) *J. Jpn. Soc. Nutr. Food Sci.* **50**, 133–137.
- Dekker, R. F. H. and Richards, G. N. (1976) *Adv. Carbohydr. Chem. Biochem.* **32**, 277–352.
- Kaji, A. (1984) *Adv. Carbohydr. Chem. Biochem.* **42**, 283–394.
- Beldman, G., Schols, H. A., Piston, S. M., Searle-van Leeuwen, M. J. F. and Voragen, A. G. J. (1997) *Adv. Macromol. Carbohydr. Res.* **1**, 1–64.
- Henrissat, B. and Bairoch, A. (1996) *Biochem. J.* **316**, 695–696.
- Coutinho, P. M. and Henrissat, B. (1999) Carbohydrate-Active Enzymes Server at URL: <http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html>
- Kaneko, S., Shimasaki, T. and Kusakabe, I. (1993) *Biosci. Biotechnol. Biochem.* **57**, 1161–1165.
- Kaneko, S., Sano, M. and Kusakabe, I. (1994) *Appl. Environ. Microbiol.* **60**, 3425–3428.
- Kaneko, S. and Kusakabe, I. (1995) *Biosci. Biotechnol. Biochem.* **59**, 2132–2133.
- Kaneko, S., Ishii, T., Kobayashi, H. and Kusakabe, I. (1998) *Biosci. Biotechnol. Biochem.* **62**, 695–699.
- Kaneko, S., Higashi, K., Yasui, T. and Kusakabe, I. (1998) *J. Ferment. Bioeng.* **85**, 519–521.
- Kaneko, S., Arimoto, M., Ohba, M., Kobayashi, H., Ishii, T. and Kusakabe, I. (1998) *Appl. Environ. Microbiol.* **64**, 4021–4027.
- Kaneko, S., Kuno, A., Matsuo, N., Ishii, T., Kobayashi, H. and Kusakabe, I. (1998) *Biosci. Biotechnol. Biochem.* **62**, 2205–2210.
- Kusakabe, I., Yasui, T. and Kobayashi, T. (1975) *Nippon Nogeikagaku Kaishi* **49**, 295–305.
- Kusakabe, I., Ohgushi, T., Yasui, T. and Kobayashi, T. (1983) *Agric. Biol. Chem.* **47**, 2713–2723.
- Yoshida, S., Kusakabe, I., Matsuo, N., Shimizu, K., Yasui, T. and Murakami, K. (1990) *Agric. Biol. Chem.* **54**, 449–457.
- Kawabata, Y., Kaneko, S., Kusakabe, I. and Gama, Y. (1995) *Carbohydr. Res.* **267**, 39–47.
- Kaneko, S., Kawabata, Y., Ishii, T., Gama, Y. and Kusakabe, I. (1995) *Carbohydr. Res.* **268**, 307–311.
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275.
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
- Manabe, T., Tachi, K., Kojima, K. and Okuyama, T. (1979) *J. Biochem. (Tokyo)* **85**, 649–59.
- Triglia, T., Peterson, M. G. and Kemp, D. J. (1988) *Nucleic Acids Res.* **16**, 8186.
- Kuno, A., Shimizu, D., Kaneko, S., Koyama, Y., Yoshida, S., Kobayashi, H., Hayashi, K., Taira, K. and Kusakabe, I. (1998) *J. Ferment. Bioeng.* **86**, 434–439.
- Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
- Wright, F. and Bibb, M. J. (1992) *Gene* **113**, 55–65.
- Tomme, P., Warren, R. A. J., Miller, Jr., R. C., Kilburn, D. G. and Gilkes, N. R. (1995) in *Enzymatic Degradation of Insoluble Carbohydrates* (Saddler, J. N. and Penner, M. H. eds.) (ACS Symposium Series no. 618), pp. 142–163, American Chemical Society, Washington, DC.
- Zverlov, V. V., Liebl, W., Bachleitner, M. and Schwarz, W. H. (1998) *FEMS Microbiol. Lett.* **164**, 337–343.
- Higashi, K., Kusakabe, I., Yasui, T., Ishiyama, T. and Okimoto, Y. (1983) *Agric. Biol. Chem.* **47**, 2903–2905.
- Higashi, K. (1983) *Bull. Fac. Agric. Tamagawa University* **23**, 69–118.
- Scott, M., Pickersgill, R. W., Hazlewood, G. P., Gilbert, H. J. and Harris, G. W. (1999) *Acta Crystallogr. D* **55**, 544–546.
- Pitson, S. M., Voragen, A. G., Vincken, J. P. and Beldman, G. (1997) *Carbohydr. Res.* **303**, 207–218.
- McKie, V. A., Black, G. W., Millward-Sadler, S. J., Hazlewood, G. P., Laurie, J. I. and Gilbert, H. J. (1997) *Biochem. J.* **323**, 547–555.

Received 26 August 1999/21 October 1999; accepted 30 November 1999