*J. Microbiol. Biotechnol.* (2007), **17**(3), 481–489

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY



# Rapid Detection and Isolation of Known and Putative α-L-Arabinofuranosidase Genes Using Degenerate PCR Primers

**PARK, JUNG-MI, NAM SOO HAN, AND TAE-JIP KIM***

*Department of Food Science and Technology, School of Applied Life Science and Environment, Chungbuk National University, Cheongju 361-763, Korea*

Received: October 23, 2006
Accepted: November 28, 2006

**Abstract** α-L-Arabinofuranosidases (AFases; EC 3.2.1.55) are *exo*-type enzymes, which hydrolyze terminal nonreducing arabinose residues from various polysaccharides such as arabinan and arabinoxylan. Genome-wide BLAST search showed that various bacterial strains possess the putative AFase genes with well-conserved motif sequences at the nucleotide and amino acid sequence levels. In this study, two sets of degenerate PCR primers were designed and tested to detect putative AFase genes, based on their three highly conserved amino acid blocks (PGGNFV, GNEMDG, and DEWNVW). Among 20 *Bacillus*-associated species, 13 species were revealed to have putative AFase genes in their genome and they share over 67% of amino acid identities with each other. Based on the partial sequence obtained from an isolate, an AFase from *Geobacillus* sp. was cloned and expressed in *E. coli*. Enzymatic characterization has verified that the resulting enzyme corresponds to a typical AFase. Accordingly, degenerate PCR primers developed in this work can be used for fast, easy, and specific detection and isolation of putative AFase genes from bacterial cells.

**Keywords:** α-L-Arabinofuranosidase (AFase), PCR detection, degenerate primers

L-Arabinose is one of the main components of the hemicelluloses widely distributed in plant cell walls, where they are present in significant amounts as arabinoxylan, arabinan, and arabinogalactan. It has been rarely used and its physiological effects *in vivo* have received little attention in spite of its useful functional properties. Recently, Seri *et al.* [25] reported that L-arabinose selectively inhibits intestinal sucrase activity in an uncompetitive manner and suppresses the plasma glucose increase resulted from sucrose ingestion. Accordingly, L-arabinose has great merits as a functional sweetener and a food additive for good health to prevent obesity [31].

Arabinan-degrading enzymes have roughly been classified into the *exo*-acting α-L-arabinofuranosidases (EC 3.2.1.55) and the *endo*-1,5-α-L-arabinanases (EC 3.2.1.99) on the basis of their mode of action. AFases are typical *exo*-acting enzymes, which hydrolyze terminal nonreducing residues from arabinose-containing polysaccharides. These enzymes work in concert with other hemicellulases to completely degrade the hemicellulose backbone, which makes it an essential enzyme for the industrial production of arabinose [23]. To date, a number of different AFases were reported and classified into four glycoside hydrolase (GH) families: GH43, 51, 54, and 62, based on sequence similarity [5, 27]. In recent years, xylan-degrading enzymes have received much attention because of their practical applications in various agroindustrial processes, such as efficient conversion of hemicellulosic biomass to fuels and chemicals, delignification of paper pulp, digestibility enhancement of animal feedstock, clarification of juices, and production of value-added functional carbohydrate materials [29, 33]. In particular, AFase is one of the rate-limiting enzymes in xylan degradation for industrial purpose.

As the results of intensive worldwide genome projects, we can approximately find the distribution of any specific genes in nature and isolate them directly by polymerase chain reaction (PCR). Nevertheless, molecular cloning of the genes without their genetic information has remained a time-consuming and laborious procedure avoided by most researchers. At present, two major methods are used to detect or clone any members of a gene family based on similarity of their nucleotide sequences. One is colony or plaque hybridization, and the other is PCR technique using degenerate (having several mixed nucleotide bases at some positions) primers. If available, the latter is much a simpler, economic, and versatile technique compared with the former.

*Corresponding author
Phone: 82-43-261-3354; Fax: 82-43-271-4412;
E-mail: tjkim@cbnu.ac.kr

Therefore, degenerate PCR has proven to be a very powerful tool to find new genes or gene families [16–19].

In this study, therefore, (i) several highly conserved sequence blocks (conserved regions) were mined based on the putative AFase genes in genome databases, (ii) degenerate PCR primers were designed and applied to the detection of AFases, (iii) the distribution of AFase genes was investigated by bioinformatical and experimental approaches, and (iv) the specificity and efficacy of the degenerate primers were verified by molecular cloning and enzymatic characterization of a putative AFase gene from *Geobacillus* sp. KCTC3012 (*G*AFase).

## MATERIALS AND METHODS

### Bacterial Strains and Growth Conditions

Twenty *Bacillus*-related strains were purchased from the Korean Collection for Type Cultures (KCTC; Daejeon, Korea) and cultivated using their optimal growth media according to the instruction provided by KCTC. *Escherichia coli* MC1061 was used as a host for the cloning and expression of target genes and the resulting transformants were grown in Luria-Bertani (LB) medium containing ampicillin (100 μg/ml) at 37°C.

### Enzymes and Reagents

*Taq* DNA polymerase and restriction endonucleases were purchased from Roche Applied Science (Mannheim, Germany) or Takara Biomedical Inc. (Otsu, Japan). Modifying enzymes, including T4 DNA ligase, and calf intestinal alkaline phosphatase (CIP), were provided by Roche Applied Science. Agarose for gel electrophoresis and the GENECLEAN *Turbo* Nucleic Acid Purification Kit were provided by QBiogene (Carlsbad, U.S.A.). The *AccuPrep* Plasmid Extraction Kit, PCR Purification Kit, and Genomic DNA Extraction Kit were purchased from Bioneer Co. (Daejeon, Korea). Oligonucleotide primers used for degenerate PCR and DNA sequencing analyses were synthesized by Bioneer Co. The Ni-NTA superflow matrix from QIAGEN (Hilden, Germany) was used for the purification of recombinant AFase. Sugar beet (branched) arabinan and debranched arabinan were purchased from Megazyme International Ireland Ltd. (Wicklow, Ireland). Other chemicals and reagents were purchased from Sigma-Aldrich Co. (St. Louis, MO, U.S.A.), Merck (Darmstadt, Germany), or Duchefa Biochemie (Haarlem, The Netherlands).

### PCR Amplification by Using Degenerate Primers

To eliminate any experimental discrepancy in the purity and quantity of DNA templates, each genomic DNA was evenly isolated by using the *AccuPrep* Genomic DNA Extraction Kit and quantified by UV spectrophotometric method [30]. The amplification of genomic DNA was carried out in 50 μl of a solution containing *Taq* polymerase buffer [10 mM Tris-HCl, pH 8.3, 40 mM KCl, 1.5 mM $MgCl_2$), 0.2 mM dNTPs (dATP, dGTP, dCTP, and dTTP)], 50 ng of template DNA, 0.2 pmole of each primer, 2.5 units of *Taq* DNA polymerase (Roche Applied Science). PCR amplification was performed in the Px2 thermal cycler (Thermo-Hybaid, Middlesex, U.K.) as follows: an initial denaturation step for 1 min at 94°C followed by 30 cycles of 60 s at 94°C; 60 s at 52°C (optimized by temperature-gradient PCR experiments); 90 s at 72°C; and a final cycle of 10 min at 72°C.

### Cloning of PCR Products or *G*AFase gene

Each PCR-amplified product was recovered using the *AccuPrep* PCR Purification Kit and visualized through agarose gel electrophoresis by ethidium bromide staining. The resulting PCR product was extracted from sliced gel and directly ligated into a pMD18-T cloning vector (Takara Biomedical Inc.), according to the manufacturer's instructions. The ligation mixture was transformed into $CaCl_2$-competent *E. coli* MC1061 cells by the standard heat-shock protocol [24]. Transformants were grown on LB medium containing ampicillin (100 μg/ml) at 37°C. Recombinant plasmid with probable DNA fragment was extracted and confirmed by restriction endonuclease treatments and DNA sequencing analysis. For the constitutive overexpression of AFase, a commercial expression vector, pHCEIIB/NdeI (BioLeaders Co., Daejeon, Korea), was slightly modified and used in this study. Synthetic oligonucleotides, CT6HN (5'-GCTCGAGCACCATCACCATCACCATTA-3') and CT6HC (5'-AGCTTAATGGTGATGGTGATGGTGCTCGAGCT-GCA-3'), were annealed and inserted into the pHCEIIB/NdeI digested with NdeI and HindIII. The resulting construct was designated as pHCXHD, possessing six histidine residues fused at the C-terminus.

### Purification and Characterization of Recombinant *G*AFase

Recombinant *E. coli* harboring pHCXGAF was cultivated in LB broth containing ampicillin (100 μg/ml) at 37°C for 12 h. The cells were harvested and disrupted by sonication (VCX750, Sonics & Materials, Inc, Newtown, CT, U.S.A.) after resuspension in 50 mM sodium-acetate buffer (pH 6.0). *G*AFase with a six-histidines tag was simply purified by using nickel-nitrilotriacetic acid (Ni-NTA) column chromatography. The crude cell extract was loaded onto the column packed by Ni-NTA and washed twice with washing buffer (50 mM $NaH_2PO_4$, 300 mM NaCl, 20 mM imidazole, pH 7.0). Finally, pure *G*AFase was eluted with elution buffer (50 mM $NaH_2PO_4$, 300 mM NaCl, and 250 mM imidazole, pH 7.0). The molecular mass of purified *G*AFase was determined by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) as described by Laemmli [12]. Purified *G*AFase was incubated with

0.5% of each substrate, including sugar beet arabinan, debranched arabinan, arabinoxylan, or arabinogalactan, in a 50 mM sodium-acetate buffer (pH 6.0) at 60°C for 6 h to determine its hydrolytic modes of action. The reaction was stopped by boiling the mixture for 5 min. The resulting products were analyzed by thin-layer chromatography (TLC) with the solvent system of isopropyl alcohol/ethylacetate/water (3:1:1, v/v/v). The TLC plate (Silica gel 60 $F_{254}$, Merck) was dried and visualized by dipping into a solution containing 0.3% (w/v) *N*-(1-naphthyl)-ethylenediamine and 5% (v/v) $H_2SO_4$ in methanol and then heating for 10 min at 110°C.

### DNA Sequencing and Sequence Analysis

DNA sequencing was performed at the Genome Research Facility in Seoul National University (Seoul, Korea) using the 3730 DNA Analyzer (Applied Biosystems, Foster City, CA, U.S.A.) according to the manufacturer's instructions. The universal sequencing primers, M13 reverse and M13 forward, were used for the determination of the nucleotide sequences from the putative AFase gene fragments cloned. Sequence comparison against GenBank, EMBL, and SwissProt was carried out by using the BLAST [1] program available at the NCBI (National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov) Web site. Sequence alignment was done by using a local version of CLUSTAL W [28].

### Sequence Information

The nucleotide sequence of the *GA*Fase cloned and determined in this study was deposited in GenBank under the accession number of EF052863.

## RESULTS AND DISCUSSION

### Distribution of the Putative AFase Genes in Nature

To date, many AFase-type enzymes have been isolated and characterized from various microbial sources including *Bacillus* [3, 8, 9], *Streptomyces* [14], *Geobacillus* [4], *Clostridium* [34], *Cytophaga* [22], *Cellvibrio* [2], *Bifidobacterium* [15], and *Thermotoga* [32]. Specifically, AFase from *Geobacillus stearothermophilus* T-6 is a member of the AFases for well-investigated their protein structure and enzymatic properties [6]. As shown in Table 1, the amino acid sequence of *Geobacillus* AFase was aligned and compared with various other sequences available from the NCBI database using the BLAST searching program. Nevertheless, there has been some confusion in the classification of AFases, due to their complexity in substrate specificities. Recently, valuable genetic information has been available and easily accessible as the result of worldwide genome projects. According to the genome database, various microorganisms have AFase genes in their genome and share a broad range of similarity in the primary structure with each other. In this work, we proposed a wide distribution of putative AFase genes in the microbial world *via* both bioinformatical and experimental approaches.

**Table 1.** Comparison of deduced amino acid similarity between putative AFase genes found in databases.

| Microorganism sources | Protein ID[a] | Identity (%)[b] | Ref. |
|---|---|---|---|
| *Geobacillus stearothermophilus* | AAD45520 | 100 | [4] |
| *Geobacillus thermoleovorans* | ABD48560 | 96 | - |
| *Bacillus halodurans* | BAB05580 | 78 | - |
| *Bacillus licheniformis* | AAU24528 | 76 | - |
| *Bacillus subtilis* | CAB14832 | 71 | [8] |
| *Bacillus clausii* | BAD62939 | 69 | - |
| *Enterococcus faecium* | EAN10306 | 62 | - |
| *Rhizobium etli* | ABC93630 | 57 | - |
| *Agrobacterium tumefaciens* | AAL43920 | 54 | - |
| *Mesorhizobium loti* | BAB50453 | 54 | - |
| *Streptomyces lividans* | AAA61708 | 54 | [14] |
| *Kineococcus radiotolerans* | EAM76066 | 49 | - |
| *Reinekea* sp. | EAR07724 | 49 | - |
| *Arthrobacter* sp. | EAL97488 | 47 | - |
| *Bifidobacterium longum* | AAO84266 | 46 | [15] |
| *Roseiflexus* sp. | EAT28126 | 44 | - |
| *Solibacter usitatus* | EAM54431 | 40 | - |
| *Thermotoga maritima* | AAD35369 | 38 | [32] |
| *Cytophaga xylanolytica* | AAC38456 | 31 | [22] |
| *Cellvibrio japonicus* | AAK84947 | 30 | [2] |
| *Clostridium stercorarium* | AAC28125 | 29 | [34] |

[a]Each gene product can be identified from NCBI database by using its own number.

[b]Relative identities were determined by amino acid sequence alignment based on that of *Geobacillus stearothermophilus* T-6 AFase as 100 percent.

### Design of Degenerate PCR Primers

The hydrolysis of glycosidic bonds occurs *via* two known major mechanisms, giving rise to either an overall retention or inversion of the anomeric configuration. Both catalytic mechanisms are found in the different AFase families: GH51 and 54 were known to be retaining enzymes; GH43, which includes AFases as well as β-xylosidases, was shown to work *via* the inverting mechanism, whereas the stereochemistry of GH62 is not yet characterized [20]. Among the retaining AFases, type-A (AbfA) AFases were known to only hydrolyze small substrates, including short-chain arabinooligosaccharides, whereas type-B (AbfB) enzymes are able to hydrolyze polymeric substrates such as arabinoxylans in addition to small substrates. This classification of AFase has, however, been very ambiguous because of their unclear substrate specificities. Accordingly, the sequence similarity between putative AFases was considered as the most important factor for the enzyme classification in this study.

In order to determine the most suitable regions for the design of degenerate primers allowing the amplification of

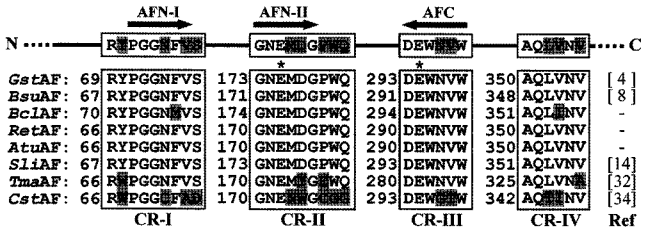


**Fig. 1.** Schematic representation of the location and orientation of three degenerate PCR primers based on the sequence alignment between putative AFases.

Amino acid residues in boxes represent the consensus sequences at the conserved regions (CRs), I, II, III, and IV, found in most putative AFases (the accession numbers of AFases are listed in Table 1). Variable residues are indicated as black letters on a dark gray background. Each number at the left side of the sequences corresponds to the position of the starting amino acid residue in each conserved region. Two putative catalytic glutamate residues [26] are marked by asterisks. Black arrows show the location and orientation of degenerate primers developed here.

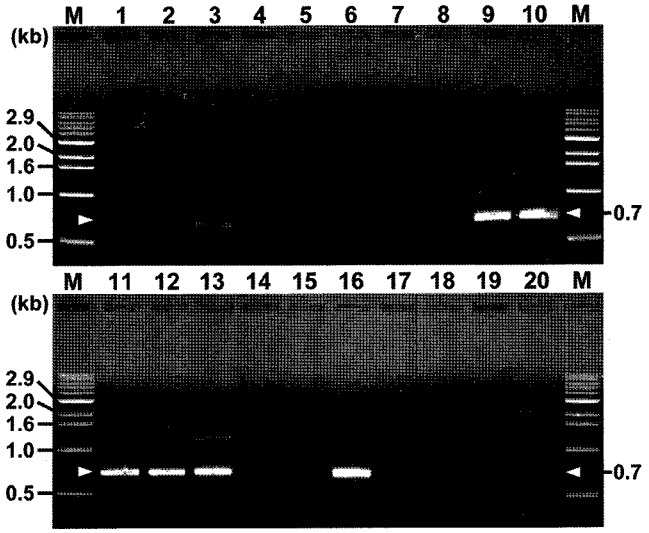


**Fig. 2.** PCR detection of putative AFase genes using degenerate primer set AF-I.

Efficacy and specificity of degenerate PCR detection was examined by using AF-I primer set and each bacterial genomic DNA template. After the degenerate PCR amplification, each sample was separated on 1% agarose gel. Lane M corresponds to 1 kb DNA marker (Bioneer Co.) and Lanes 1–20 show the PCR results from each bacterial template DNA with an individual number corresponding to that in Table 3. White arrow heads indicate the position of PCR-amplified fragments with an expected length (about 680 bp).

AFases, highly common sequence blocks should be found and analyzed from various putative AFase family genes. At first, the protein sequences of known and putative AFases were closely aligned and compared with each other by using the program CLUSTAL W. The alignment of AFase sequences revealed that 4 amino acid blocks (PGGNFV, GNEMDG, DEWNVW, and AQLVNV; shown in Fig. 1) are highly conserved in most sequences. Residues of PGG (in block I), GNE (block II), and DEW (block III) are highly conserved among known AFases, whereas block4 is relatively variable in sequences too short to design any stable primer with an adequate annealing temperature. In addition, the three-dimensional structure of hexameric *Geobacillus* AFase (AbfA in GH51) and its catalytic residues were recently identified as follows: Glu175 (acid/base) and Glu294 (nucleophile) are included in blocks II and III, respectively [26]. As a result, two degenerate primer sets (AF-I set with AFN-I/AFC and AF-II set with AFN-II/AFC) were designed for PCR detection of putative AFases (Table 2). Specific DNA fragments of 684 and 378 bp were expected to be obtained from any putative genomic AFase gene with each degenerate primer set AF-I and AF-II, respectively.

## Detection of AFase Genes Using Degenerate PCR

To evaluate the specificity and efficacy of degenerate primers, 20 different *Bacillus*-associated strains were purchased from KCTC and their genomic DNAs with appropriate purity and concentration were evenly isolated and prepared by using a genomic DNA extraction kit. Then, PCR conditions for specific detection were optimized for template and primer concentrations, annealing temperature, extension time, and numbers of cycles (described in Methods).

As shown in Figs. 2 and 3, amplified PCR products with appropriate size were specifically successfully detected from individual DNA templates. The specificity or efficacy of both primer sets depends on the microbial sources. For example, *Bacillus amyloliquefaciens* KCTC3002 showed the positive signal by primer set AF-I, whereas no amplification was detected by PCR with the primer set AF-II. Final results from degenerate PCR detection are summarized in Table 3, which suggested that 13 out of 20 *Bacillus* strains probably possess the AFase gene in their genome. According to the BLAST search of the genome database, *Bacillus cereus* and *B. thuringiensis*

**Table 2.** List of degenerate PCR primer sets for the detection of putative AFase genes.

| Set | Primer name | Primer sequence[a] | $T_m$ (°C) | Orientation | Degeneracy | Length (bp) |
|---|---|---|---|---|---|---|
| AF-I | AFN-I | 5'-CCNGGNGGNAAYTTYGT-3' | 48–58 | Forward | 256 | 684 |
| | AFC | 5'-CCANACRTTCCAYTCRTC-3' | 50–58 | Reverse | 32 | |
| AF-II | AFN-II | 5'-GGNAAYGARATGGAYGG-3' | 48–56 | Forward | 32 | 378 |
| | AFC | 5'-CCANACRTTCCAYTCRTC-3' | 50–58 | Reverse | 32 | |

[a]Abbreviations of degenerate nucleotides: N=A/G/C/T; R=A/G; Y=C/T.

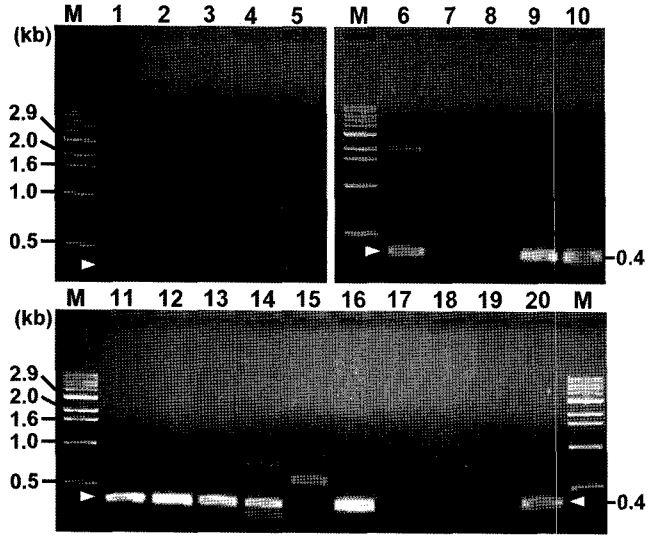


**Fig. 3.** PCR detection of putative AFase genes using degenerate primer set AF-II.

Efficacy and specificity of degenerate PCR detection were examined by using the AF-II primer set and each bacterial genomic DNA template. Experimental procedures and lane numbers are completely the same as those mentioned in Fig. 2. White arrowheads indicate the position of PCR-amplified fragments with an expected length (about 380 bp).

showed no hit for any AFase gene in their genome, which coincided with the PCR detection results in this work. On the other side, *B. subtilis*, *B. licheniformis*, and *Geobacillus* sp. were supposed to have the AFase gene (refer to Table 1), which was also consistent with the PCR detection results here.

In conclusion, primer set AF-II showed a better detection range and specificity for AFase than AF-I. As degenerate primers include mixed or variable nucleotide bases at some positions, high degeneracy in primer sequence can cause relatively low specificity with nonspecific bands. The primers with low degeneracy may reduce nonspecific amplification, but may have problems of narrow detection range. The degeneracy of primer AFN-I is much higher than that of other primers, AFN-II and AFC (Table 2). Accordingly, the combination of degeneracy in primers may result in some differences in specificity or detection efficacy [13]. Even though only a small part of strains were detected by the primer set AF-I, it has great merit to result in longer PCR products than those obtained by the primer set AF-II. The resulting fragment of 684 bp is able to cover 45–50% in length of the common AFase gene, which will give more information about the corresponding structural genes for the purpose of gene cloning. On the other hand, primer set AF-II can be

**Table 3.** Summarized results of the PCR detection using degenerate primer sets.

| Growth temp.[a] | Source of genomic DNA (KCTC/ATCC)[b] | Results with primer sets[c] | |
|---|---|---|---|
| | | AF-I[d] | AF-II |
| 30°C | 1. *Bacillus cereus* (KCTC1014/ATCC21768) | − | − |
| | 2. *Bacillus cereus* (KCTC3624/ATCC14579) | × | − |
| | 3. *Bacillus firmus* (KCTC3626/ATCC14575) | △ | − |
| | 4. *Bacillus megaterium* (KCTC3007/ATCC14581) | − | − |
| | 5. *Bacillus subtilis* (KCTC2217/ATCC33234) | − | + |
| | 6. *Bacillus subtilis* (KCTC1028/ATCC6051A) | − | + |
| | 7. *Bacillus thuringiensis* (KCTC3452/ATCC10792) | − | − |
| | 8. *Brevibacillus brevis* (KCTC3743/ATCC8246) | △ | × |
| | 9. *Paenibacillus alginolyticus* (KCTC3567/ATCC51185) | + | + |
| | 10. *Paenibacillus amylolyticus* (KCTC3455/ATCC9995) | + | + |
| | 11. *Paenibacillus macerans* (KCTC1822/ATCC8244) | + | + |
| | 12. *Paenibacillus polymyxa* (KCTC3008/ATCC8523) | + | + |
| | 13. *Bacillus circulans* (KCTC3004/ATCC21783) | + | + |
| 37°C | 14. *Bacillus amyloliquefaciens* (KCTC3002/ATCC23845) | − | + |
| | 15. *Bacillus coagulans* (KCTC3625/ATCC7050) | − | × |
| 45°C | 16. *Bacillus licheniformis* (KCTC1030/ATCC27811) | + | + |
| 55°C | 17. *Aneurinibacillus thermoaerophilus* (KCTC3741) | − | − |
| | 18. *Bacillus stearothermophilus* (KCTC2107/ATCC7953) | △ | − |
| | 19. *Geobacillus stearothermophilus* (KCTC1752/ATCC12980) | − | − |
| | 20. *Geobacillus* sp. (KCTC3012) | △ | + |

[a]According to the instructions from KCTC, each microorganism was grouped by its own growth temperature.
[b]If compatible, both KCTC and ATCC numbers are co-written in a parenthesis.
[c]Symbols: (+), correct-sized fragment was amplified and confirmed by sequence analysis; (△), correct-sized but faint fragment was amplified; (×), incorrect-sized fragment was amplified; (−), no fragment was detected.
[d]Degenerate PCR primer sets, AF-I and AF-II, are fully described in Table 2.

applied to the specific primary detection of AFase genes from a broad range of microbes. The combinational application of these primer sets can be more helpful to obtain further information about useful microbial AFases.

### Sequence Analysis of Putative AFase Genes

The PCR-amplified fragments, with the specific primer set AF-I or AF-II, were expected to have 684 bp encoding 228 amino acids or 378 bp encoding 126 amino acids, respectively. Ready-made T-cloning vector (pMD18-T) was used for the high throughput and efficient cloning of PCR products amplified by degenerate primers. The probable five AF-I and ten AF-II fragments were cloned and used as a template for DNA sequencing analysis. The nucleotide and the deduced amino acid sequences of detected fragments were determined and compared with each other by using the CLUSTAL W program. As shown in Table 4, 10 putative AFase gene fragments (by set AF-II) showed over 70% of amino acid sequence identities by sequence alignment. The relatively high similarity between them may be caused by the short length of each PCR product, but all the amplicons are fully supposed to be a part of the putative AFase gene.

For the verification of the genes in detail, all the deduced amino acid sequences were aligned with each other and the results were compared in Fig. 4. Highly conserved amino acid blocks I, II, and III (in Fig. 1) were obviously found in their sequence alignment. In addition to high sequence similarity between PCR fragments, the identification of conserved blocks strongly supports that they are likely to be putative AFase genes and both degenerate primer sets work very well for the detection of AFase family genes. Even though degenerate primer systems are known to have common problems of limited degeneracy, the primers designed here have high specificity for the PCR detection. Finally, the tools developed in this study can be used for the simple and fast detection and isolation of AFase genes from microbial sources.

### Cloning and Characterization of *GA*FAse Genes

Even though high sequence identities can give good evidence to prove the detection of AFase, simple cloning of putative AFases will be one of the final goals of this study. For the cloning of an AFase gene, genomic DNA libraries of *Geobacillus* sp. KCTC3012 were constructed with the plasmid pUC18 by using various restriction endonucleases, including EcoRI and KpnI. Two internal primers, GAF1 (5'-GATCGACCCATTTCATGACC-3') and GAF2 (5'-TTCTGTGGCG GCAATTGCAG-3'), were designed and synthesized, according to the nucleotide

**Table 4.** Pairwise sequence identity between putative AFases detected in this study.

| AFase genes[a] | Relative amino acid (nucleotide) sequence identities (%)[b] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *GAF* 3012 | *Bsu*AF 1028 | *Bsu*AF 2217 | *Bli*AF 1030 | *Bam*AF 3002 | *Bci*AF 3004 | *Pal*AF 3567 | *Pma*AF 1822 | *Ppo*AF 3008 | *Pam*AF 3455 |
| *Gst*AF | 93.0 (79.6) | 74.6 (72.0) | 74.6 (72.0) | 74.6 (72.0) | 74.6 (71.7) | 74.6 (70.3) | 73.7 (71.1) | 73.7 (70.6) | 71.9 (68.5) | 71.1 (69.0) |
| 3012 | | 71.9 (68.2) | 71.9 (68.2) | 73.7 (67.9) | 71.9 (70.6) | 75.4 (67.9) | 71.1 (71.1) | 72.8 (63.8) | 71.9 (67.9) | 71.9 (65.6) |
| 1028 | | | 100.0 (100.0) | 81.6 (74.1) | 86.8 (76.7) | 71.9 (66.6) | 71.9 (67.7) | 71.1 (66.3) | 70.2 (70.1) | 71.1 (69.2) |
| 2217 | | | | 81.6 (74.1) | 86.8 (76.7) | 71.9 (66.6) | 71.9 (67.7) | 71.1 (66.3) | 70.2 (70.1) | 71.1 (69.2) |
| 1030 | | | | | 80.7 (77.3) | 71.9 (66.6) | 71.1 (68.3) | 70.2 (67.3) | 72.8 (64.4) | 68.4 (70.8) |
| 3002 | | | | | | 71.9 (67.9) | 67.5 (65.7) | 69.3 (67.9) | 71.1 (65.8) | 70.2 (67.9) |
| 3004 | | | | | | | 70.2 (68.4) | 81.6 (77.0) | 82.5 (75.8) | 80.7 (75.5) |
| 3567 | | | | | | | | 70.2 (68.9) | 71.9 (68.6) | 70.2 (68.3) |
| 1822 | | | | | | | | | 87.7 (75.2) | 89.5 (82.8) |
| 3008 | | | | | | | | | | 83.3 (75.2) |

[a]The abbreviations of putative AFases: *Gst*AF, *Geobacillus stearothermophilus* T-6 AFase; *GA*F3012, *Geobacillus* sp. KCTC3012 AFase; *Bsu*AF1028, *Bacillus subtilis* AFase; *Bsu*AF2217, *B. subtilis* AFase; *Bli*AF1030, *B. licheniformis* AFase; *Bam*AF3002, *B. amyloliquefaciens* AFase; *Bci*AF3004, *B. circulans* AFase; *Pal*AF3567, *Paenibacillus alginolyticus* AFase; *Pma*AF1822, *P. macerans* AFase; *Ppo*AF3008, *P. polymyxa* AFase; *Pam*AF3455, *P. amylolyticus* AFase.

[b]Nucleotide sequence identities are written in parentheses.

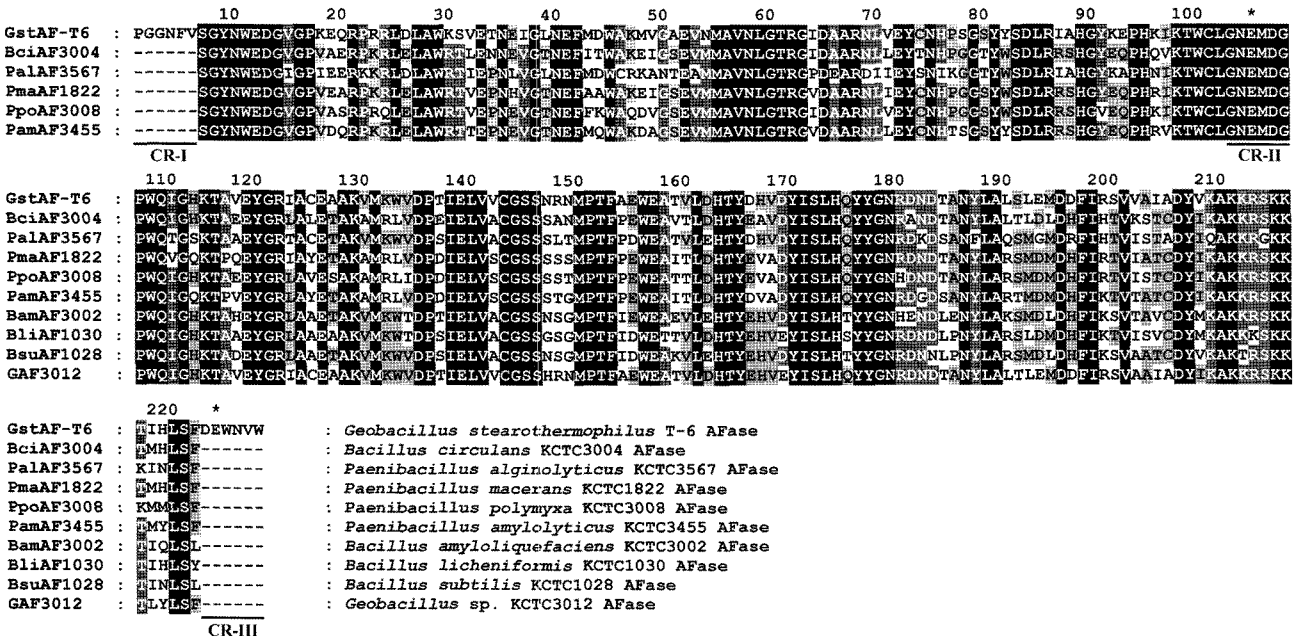


**Fig. 4.** Amino acid sequence alignment between putative AFases detected in this study.
Abbreviation of each AFase is identical to those in Table 4. Highly conserved amino acid sequence blocks in most AFases (as shown in Fig. 1) are underlined and marked as CR-I, II, and III, respectively. Putative catalytic residues are emphasized by asterisks. White letters on a black background show the fully conserved amino acid residues in all protein sequences.

sequence determined by PCR detection. Primers GAF1 and GAF2 were used for PCR amplification in pairs with the primers M13 forward (5'-GTTTTCCCAGTCACGACGTT-3') and M13 reverse (5'-AGCGGATAACAATTTCACAC-3'), respectively, to amplify the N- and C-terminal parts of the AFase gene from the genomic DNA libraries as templates. Each PCR-amplified fragment was directly cloned into the pMD18-T vector and then DNA sequences at both ends were determined. Based on the resulting sequence data, additional two primers, GAF-N (5'-TTTCATATGA-ACACGAAAAAAGCTAAAA-3') and GAF-C (5'-TTTC-TCGAGTTTCTTAGC CAAACGAATCAC-3'), were prepared and used to amplify the entire open reading frame of *GA*Fase. The amplified DNA digested with NdeI and XhoI was subcloned into the expression vector pHCXHD (described in Methods) and designated as pHCXGAF. The resulting *GA*Fase (Accession No. EF052863) shares 78.7% and 91.4% of sequence identity with the AFase from *Geobacillus stearothermophhilus* T-6 at the nucleotide and the amino acid levels, respectively. The method used in this study can be applied to the rapid isolation of an entire gene by PCR, even if only a short internal sequence of the gene is available.

As shown in Fig. 5A, the recombinant *GA*Fase (about 65 kDa), fused with a C-terminal 6xHis tag, was successfully overexpressed and easily purified to homogeneity by using Ni-NTA column chromatography. The hydrolyzing activity of the purified AFase was examined against various substrates containing arabinose residues (Fig. 5B). *GA*Fase showed much higher hydrolyzing activity on arabinan substrates than arabinoxylan, whereas arabinogalactan was rarely hydrolyzed.

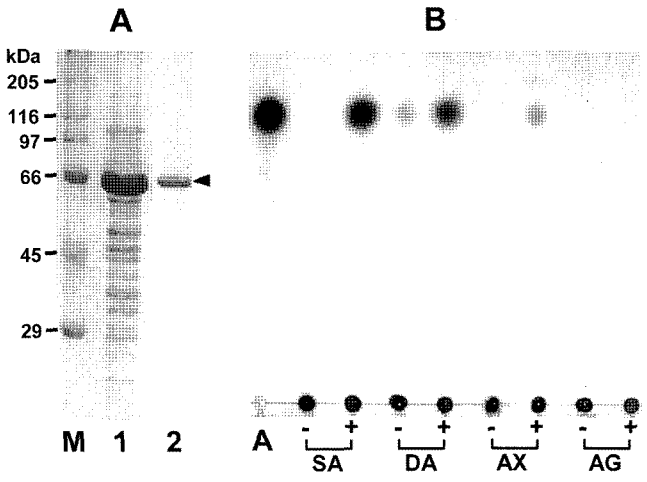


**Fig. 5.** Cloning, expression, and characterization of recombinant *GA*Fase.
The *Geobacillus* sp. KCTC3012 AFase gene was cloned and expressed in *E. coli*. Recombinant *GA*Fase with an arrowhead (65 kDa) was successfully expressed as shown in panel A: lane M, protein molecular weight marker (Sigma-Aldrich Inc.); lane 1, crude cell extract prepared from *E. coli* MC1061 harboring pHCXGAF; lane 2, *GA*Fase purified by Ni-NTA affinity chromatography. Recombinant *GA*Fase was reacted with various substrates (Panel B). As a result, L-arabinose was produced by the hydrolyzing activity of *GA*Fase: A, arabinose standard; SA, sugar beet (branched) arabinan; DA, debranched arabinan; AX, arabinoxylan; AG, arabinogalactan; (-), before the enzyme reaction; (+), after the reaction with *GA*Fase.

The relatively higher activity on sugar beet (branched) arabinan than the debranched one elucidated that *G*AFase corresponds to a typical *exo*-acting arabinose-producing enzyme, α-L-arabinofuranosidase.

Considering the specific hydrolyzing activities of AFases releasing functional sugar L-arabinose from hemicellulosic materials, the degenerate PCR primers developed here can be applicable to the high-throughput detection and isolation of various microbial AFases for industrial purposes.

## Acknowledgment

This work was supported by Korea Institute of Industrial Technology Evaluation & Planning (ITEP), The Ministry of Commerce, Industry & Energy (MOCIE), and Chungbuk National University Grant in 2004.

## References

1. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.
2. Beylot, M. H., V. A. McKie, A. G. Voragen, C. H. Doeswijk-Voragen, and H. J. Gilbert. 2001. The *Pseudomonas cellulosa* glycoside hydrolase family 51 arabinofuranosidase exhibits wide substrate specificity. *Biochem. J.* **358:** 607–614.
3. Degrassi, G., A. Vindigni, and V. Venturi. 2003. A thermostable α-arabinofuranosidase from xylanolytic *Bacillus pumilus*: Purification and characterisation. *J. Biotechnol.* **101:** 69–79.
4. Gilead, S. and Y. Shoham. 1995. Purification and characterization of α-L-arabinofuranosidase from *Bacillus stearothermophilus* T-6. *Appl. Environ. Microbiol.* **61:** 170–174.
5. Henrissat, B. and G. Davis. 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7:** 637–644.
6. Hovel, K., D. Shallom, K. Niefind, V. Belakhov, G. Shoham, T. Baasov, Y. Shoham, and D. Schomburg. 2003. Crystal structure and snapshots along the reaction pathway of a family 51 α-L-arabinofuranosidase. *EMBO J.* **22:** 4922–4932.
7. Jung, S. J., H. J. Kim, and H. Y. Kim. 2005. Quantitative detection of *Salmonella typhimurium* contamination in milk, using real-time PCR. *J. Microbiol. Biotechnol.* **15:** 1353–1358.
8. Kaneko, S., M. Sano, and I. Kusakabe. 1994. Purification and some properties of α-L-arabinofuranosidase from *Bacillus subtilis* 3-6. *Appl. Environ. Microbiol.* **60:** 3425–3428.
9. Kim, K. J., K. N. Kim, and Y. J. Choi. 2004. Characterization of the *arfA* gene from *Bacillus stearothermophilus* no. 236 and its protein product, α-L-arabinofuranosidase. *J. Microbiol. Biotechnol.* **14:** 474–482.
10. Kim, S. B., H. J. Lim, W. K. Lee, I. G. Hwang, G. J. Woo, and S. Ryu. 2006. PCR-based detection and molecular genotyping of enterotoxigenic *Clostridium perfringens* isolates from swine diarrhea in Korea. *J. Microbiol. Biotechnol.* **16:** 291–294.
11. Kunihiro, S., Y. Kawanishi, M. Sano, K. Naito, Y. Matsuura, Y. Tateno, T. Gojobori, Y. Yamagata, K. Abe, and M. Machida. 2002. A polymerase chain reaction-based method for cloning novel members of a gene family using a combination of degenerate and inhibitory primers. *Gene* **289:** 177–184.
12. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.
13. Linhart, C. and R. Shamir. 2002. The degenerate primer design problem. *Bioinformatics* **18:** 172–181.
14. Manin, C., F. Shareek, R. Morosoli, and D. Kluepfel. 1994. Purification and characterization of an α-L-arabinofuranosidase from *Streptomyces lividans* 66 and DNA sequence of the gene (abfA). *Biochem. J.* **302:** 443–449.
15. Margolles, A. and C. G. De Los Reyes-Gavilan. 2003. Purification and functional characterization of a novel α-L-arabinofuranosidase from *Bifidobacterium longum* B667. *Appl. Environ. Microbiol.* **69:** 5096–5103.
16. Martins, M. L., E. F. de Araujo, H. C. Mantovani, C. A. Moraes, and M. C. Vanetti. 2005. Detection of the *apr* gene in proteolytic psychrotrophic bacteria isolated from refrigerated raw milk. *Int. J. Food Microbiol.* **102:** 203–211.
17. Park, J. H., H. K. Park, B. C. Kang, E. S. Song, H. J. Jang, and C. M. Kim. 2006. Development of genus- and species-specific probe design system for pathogen detection based on 23S rDNA. *J. Microbiol. Biotechnol.* **16:** 740–747.
18. Park, K. S., Y. J. Bae, E. J. Jung, and S. J. Kang. 2005. RT-PCR-based detection of six garlic viruses and their phylogenetic relationships. *J. Microbiol. Biotechnol.* **15:** 1110–1114.
19. Park, S. H., H. J. Kim, and H. Y. Kim. 2006. Simultaneous detection of *Yersinia enterocolitica*, *Staphylococcus aureus*, and *Shigella* spp. in lettuce using multiplex PCR method. *J. Microbiol. Biotechnol.* **16:** 1301–1305.
20. Pitson, S. M., A. G. Voragen, and G. Beldman. 1996. Stereochemical course of hydrolysis catalyzed by arabinofuranosyl hydrolases. *FEBS Lett.* **398:** 7–11.
21. Regeard, C., J. Maillard, and C. Holliger. 2004. Development of degenerate and specific PCR primers for the detection and isolation of known and putative chloroethene reductive dehalogenase genes. *J. Microbiol. Methods* **56:** 107–118.
22. Renner, M. J. and J. A. Breznak. 1998. Purification and properties of ArfI, an α-L-arabinofuranosidase from *Cytophaga xylanolytica*. *Appl. Environ. Microbiol.* **64:** 43–52.
23. Saha, B. C. 2000. Alpha-L-arabinofuranosidases: Biochemistry, molecular biology and application in biotechnology. *Biotechnol. Adv.* **18:** 403–423.
24. Sambrook, J. and D. W. Russell. 2001. *Molecular Cloning: A Laboratory Manual*, pp. 116–118. 3rd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, U.S.A.
25. Seri, K., K. Sanai, N. Matsuo, K. Kawakubo, C. Xue, and S. Inoue. 1996. L-Arabinose selectively inhibits intestinal sucrase in an uncompetitive manner and suppresses glycemic response after sucrose ingestion in animals. *Metabolism* **45:** 1368–1374.

26. Shallom, D., V. Belakhov, D. Solomon, G. Shoham, T. Baasov, and Y. Shoham. 2002. Detailed kinetic analysis and identification of the nucleophile in α-L-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *J. Biol. Chem.* **277:** 43667–43673.
27. Sunna, A. and G. Antranikian. 1997. Xylanolytic enzymes from fungi and bacteria. *Crit. Rev. Biotechnol.* **17:** 39–67.
28. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.
29. Vakari, L., M. Tenkanen, J. Buchert, M. Ratto, M. Bailey, M. Siika-aho, and M. Linko. 1993. Hemicellulases for industrial applications, pp. 131–82. *In* J. N. Saddler (ed.), *Bioconversion of Forest and Agricultural Plant Residues*. CAB International, Oxford.
30. Wilfinger, W. W., K. Mackey, and P. Chomczynski. 1997. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *Biotechniques* **22:** 474–481.
31. Yoon, H. S., C. H. Kim, T. J. Kim, I. K. Keum, and N. S. Han. 2003. Novel functional sugar L-arabinose: Its functionality, uses and production methods. *Kor. J. Food Sci. Technol.* **35:** 757–763.
32. Yoon, H. S., I. K. Keum, N. S. Han, and C. H. Kim. 2004. Molecular cloning and characterization of a gene encoding α-L-arabinofuranosidase from *Thermotoga maritima*. *Food Sci. Biotechnol.* **13:** 244–247.
33. Zeikus, J. G., C. Lee, Y. E. Lee, and B. C. Saha. 1991. Thermostable saccharidases: New sources, uses, and biodesign, pp. 36–51. *In* G. F. Leatham and M. E. Himmel (eds.), *Enzymes in Biomass Conversion*. American Chemical Society, Washington, D.C.
34. Zverlov, V. V., W. Liebl, M. Bachleitner, and W. H. Schwarz. 1998. Nucleotide sequence of arfB of *Clostridium stercorarium*, and prediction of catalytic residues of α-L-arabinofuranosidases based on local similarity with several families of glycosyl hydrolases. *FEMS Microbiol. Lett.* **164:** 337–343.