Crystal Structure of an Exo-1,5-α-L-arabinofuranosidase from Streptomyces avermitilis Provides Insights into the Mechanism of Substrate Discrimination between Exo- and Endo-type Enzymes in Glycoside Hydrolase Family 43

Exo-1,5-α-L-arabinofuranosidases belonging to glycoside hydrolase family 43 have strict substrate specificity. These enzymes hydrolyze only the α-1,5-linkages of linear arabinan and arabino-oligosaccharides in an exo-acting manner. The enzyme from Streptomyces avermitilis contains a core catalytic domain belonging to glycoside hydrolase family 43 and a C-terminal arabinan binding module belonging to carbohydrate binding module family 42. We determined the crystal structure of intact exo-1,5-α-L-arabinofuranosidase. The catalytic module is composed of a 5-bladed β-propeller topologically identical to the other family 43 enzymes. The arabinan binding module had three similar subdomains assembled against one another around a pseudo-3-fold axis, forming a β-trefoil-fold. A sugar complex structure with α-1,5-α-L-arabinofuranotriose revealed three subsites in the catalytic domain, and a sugar complex structure with α-arabinofuranosyl azide revealed three arabinose-binding sites in the carbohydrate binding module. A mutagenesis study revealed that substrate specificity was regulated by residues Asn-159, Tyr-192, and Leu-289 located at the aglycon side of the substrate-binding pocket. The exo-acting manner of the enzyme was attributed to the strict pocket structure of subsite −1, formed by the flexible loop region Tyr-281–Arg-294 and the side chain of Tyr-40, which occupied the positions corresponding to the catalytic glycon cleft of GH43 endo-acting enzymes.

1-L-Arabinose residues are widely distributed in plant cell walls, where they are present in polymers such as arabinans, arabinoxylans, arabinoxylans, and arabinoxylan arabinofuranohydrolases (1). Research on plant cell walls is becoming a necessity because worldwide attention has now focused on bioethanol production to combat global warming and to improve global energy security. Because of competition between food and fuel, lignocellulose is expected to be used as a material for fuel ethanol production in the future. Generally, lignocellulose contains cellulose, which makes up ~40% of the total amount of cell wall components, together with ~20% hemicellulose, which is mainly composed of pentoses such as xylose and arabinose (2). Hemicelluloses often become bad factors in bioethanol production because the efficiency of ethanol conversion from pentoses is significantly lower than that from hexoses (3, 4).

In contrast, 1-arabinose is used as a functional sugar in the food industry. This sugar has a sweet taste and selectively inhibits its intestinal sucrase activity in a noncompetitive manner and consequently suppresses plasma glucose increase due to sucrose ingestion (5–7). Therefore, 1-arabinose may also be useful in preventing excess sucrose utilization.

Because the structure of L-arabinose-containing polysaccharides is highly variable and complex, a wide variety of α-L-arabinofuranosidas (EC 3.2.1.59) that have various substrate specificities are necessary for the hydrolysis of such polysaccharides and for the production of L-arabinose. We have previously purified some α-L-arabinofuranosidas and elucidated their substrate specificities toward structurally defined substrates (8–14). The α-L-arabinofuranosidas studied have broad substrate specificities; however, α-L-arabinofuranosidase II from Streptomyces chartreusis (15) has strict substrate specificity. It hydrolyzed only the α-1,5-linkages of linear arabinan and arabinooligosaccharides in an exo-acting manner and was subsequently designated as an exo-1,5-α-L-arabinofuranosidase (15). Glycoside hydrolases are classified into 118 families according to the similarity of their amino acid sequences, which imply both structural and mechanistic relationships (16, 17).

The abbreviations used are: GH, glycoside hydrolase; AKBMB42, carbohydrate binding module of A. kawachi α-L-arabinofuranosidase; AaFbA, α-L-arabinofuranosyl azide; BSAr43A, B. subtilis arabinan endo-α-1,5-L-arabinosidase; BSAKH-m2,3, B. subtilis arabinobioxyran arabinofuranohydrolase; CBM, carbohydrate binding module; JA/A43A, C. japonicus endo/exo-1,5-α-L-arabinanase; HPAEC-PAD, high performance anion-exchange chromatography with pulsed amperometric detection; PNP-α-L-Araf, p-nitrophenyl α-L-arabinofuranoside; SaaAr43A, S. avermitilis exo-1,5-α-L-arabinofuranosidase; SABMB42, carbohydrate binding module of S. avermitilis exo-1,5-α-L-arabinofuranosidase; Se-Met, selenomethionine; XynB3, G. stearothermophilus β-xidosidase.

Received for publication, July 14, 2010, and in revised form, August 11, 2010. Published, JBC Papers in Press, August 25, 2010, DOI 10.1074/jbc.M110.164251

Zui Fujimoto†, Hitomi Ichinose†, Tomoko Maehara*, Mariko Honda†, Motomitsu Kitaoka‡, and Satoshi Kaneko§

From the †Protein Research Unit, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan and the ‡Food Biotechnology Division, National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan.

* This work was supported, in part, by a grant-in-aid (Development of Biomass Utilization Technologies for Revitalizing Rural Areas) from the Ministry of Agriculture, Forestry, and Fisheries of Japan.

† This on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1.

The atomic coordinates and structure factors (codes 3AKF, 3AKG, 3AKH, and 3AKI) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Both authors contributed equally to this work.

To whom correspondence should be addressed. Tel.: 81-29-838-8022; Fax: 81-29-838-7996; E-mail: sakaneko@affrc.go.jp.
furanosidase indicates that it is a novel enzyme belonging to family 43 (GH43) (15).

GH43 is a family composed of a wide variety of enzymes, including β-xylanosidase (EC 3.2.1.37), α-L-arabinofuranosidase (exo-1,5-α-L-arabinofuranosidase), bifunctional β-xylanosidase/α-L-arabinofuranosidase, endo-α-L-arabinanase (E.C. 3.2.1.99), endo-β,1,4-xylanase (EC 3.2.1.8), and exo-β,1,3-galactanase (EC 3.2.1.145). Therefore, a detailed functional characterization using a recombinant enzyme with its mutants would be interesting because enzymes belonging to the same family have a common polypeptide folding and identical catalytic mechanism. Especially, a comparison of the structures of exo-1,5-α-L-arabinofuranosidase and endo-α-L-arabinanase will provide difference of the substrate recognition mechanisms of exo- and endo-type enzymes. In the previous work we succeeded in the heterogeneous expression of exo-1,5-α-L-arabinofuranosidase from *Streptomyces avermitilis* (SaAraf43A) (18). SaAraf43A is composed of an N-terminal GH43 catalytic domain and a C-terminal carbohydrate binding module family 42 (CBM42) domain. Of enzymes that contain CBM42 domains, *Aspergillus kawachii* α-L-arabinofuranosidase, which belongs to the GH54 family, has been extensively studied, and its crystal structures have been determined (19, 20). In this enzyme the CBM42 domain is the substrate recognition domain that specifically binds to α-L-arabinofuranose. Structural analyses of our enzyme will clarify the function of the CBM42 domain in SaAraf43A. These analyses will also elucidate the efficient catalytic mechanism of the multidomain glycosidase toward the recalcitrant substrate. Given this importance, crystallization trials of SaAraf43A, a modular enzyme of GH43 and CBM42, were performed (21).

In this article we analyzed the crystal structure of SaAraf43A. A mutagenesis study of SaAraf43A provides insights into the mechanism of substrate discrimination between exo- and endo-type enzymes in glycoside hydrolase family 43.

### EXPERIMENTAL PROCEDURES

**Substrates**—p-Nitrophenyl α-L-arabinofuranosidase (PNP-α-L-Araf) was purchased from Sigma. α-1,5-L-Arabinofuranosyl-gosaccharides with degrees of polymerization ranging from 2 to 5 were obtained from Megazyme (Wicklow, Ireland). Methyl 2-O-, methyl 3-O-, and methyl 5-O-α-L-arabinofuranosyl-α-L-arabinofuranosides (α-L-arabinofuranobioses) were prepared as previously reported (22).

α-L-Arabinofuranosyl azide (ArafAz) was synthesized as follows. In 50 ml of *N*,*N*-dimethylformamide, 5 g of 2,3,6-tribenzoyl-α-L-arabinofuranosyl bromide (23) (9.5 mmol) and 5 g of sodium azide (77 mmol) were dissolved, and the reaction was allowed to proceed at 95 °C for 3.5 h. Then the solution was rotary evaporated to remove *N*,*N*-dimethylformamide. The residue was dissolved in 100 ml of CH$_2$Cl$_2$, filtered, and washed with aqueous NaHCO$_3$, dried over Na$_2$SO$_4$, and evaporated into syrup. The syrup was then applied to a silica gel dry chromatography (3-cm diameter × 20 cm) with hexane-ethyl acetate (3:2 by volume) as solvent. Fractions containing the main product were collected, and 2,3,6-tribenzoyl-α-L-arabinofuranosyl azide (ArafAz) was crystallized from methanol. Finally, 2.2 g of 2,3,6-tribenzoyl-α-L-arabinofuranosyl azide was crystallized from methanol. 2.2 g of 2,3,6-tribenzoyl-α-L-arabinofuranosyl azide was crystallized from methanol. 2.2 g of 2,3,6-tribenzoyl-α-L-arabinofuranosyl azide was crystallized from methanol. Finally, 2.2 g of 2,3,6-tribenzoyl-α-L-arabinofuranosyl azide was crystallized from methanol.

### TABLE 1

| Data collection | Native | Se-Met (peak) | Se-Met (low remote) | Se-Met (edge) | Se-Met (high remote) | α-1,5-α-L-arabinobiose complex | α-1,5-α-L-arabinotriose complex | ArafAz complex |
|----------------|--------|---------------|---------------------|--------------|---------------------|---------------------------|-------------------------------|-----------------|
| Unit-cell parameters (Å) | $a = 41.0$ | $a = 41.0$ | $b = 91.5$ | $b = 39.0$ | $c = 93.5$ | $a = 41.1$ | $a = 41.2$ | $a = 41.0$ |
| Resolution (Å) | 75.8-2.2 | 75.8-2.2 | 50.0-2.2 | 50.0-2.2 | 50.0-2.2 | 50.0-1.8 | 50.0-1.7 | 100.0-2.2 |
| Complete (%) | 0.081 (0.265) | 0.073 (0.223) | 0.069 (0.221) | 0.068 (0.234) | 0.065 (0.235) | 0.056 (0.267) | 0.055 (0.257) | 0.074 (0.281) |
| Multiplicity | 135.1 | 135.1 | 135.1 | 135.1 | 135.1 | 135.1 | 135.1 | 135.1 |
| Average J/Å (f) | 44.4 (3.1) | 51.8 (17.5) | 52.4 (17.1) | 67.2 (15.7) | 51.9 (16.4) | 52.7 (9.6) | 53.7 (9.5) | 38.1 (7.5) |
| Unique reflections | 26,313 (2,601) | 26,208 (2,528) | 26,086 (2,511) | 26,617 (2,529) | 26,376 (2,529) | 47,867 (4,726) | 56,484 (5,492) | 34,964 (3,416) |
| Observed reflections | 182,571 | 354,823 | 352,391 | 355,344 | 358,318 | 663,514 | 823,137 | 308,745 |

### Structure refinement

| Resolution (Å) | 75.8-2.2 | 75.8-2.2 | 50.0-2.2 | 50.0-2.2 | 50.0-2.2 | 50.0-1.8 | 50.0-1.7 | 100.0-2.2 |
|----------------|--------|---------------|---------------------|--------------|---------------------|---------------------------|-------------------------------|-----------------|
| Rfree | 0.208 (0.304) | 0.191 (0.240) | 0.254 (0.357) | 0.263 (0.396) | 0.221 (0.311) | 0.220 (0.294) | 0.231 (0.267) | 0.267 |
| r.m.s.d. from ideal | 0.012 | 0.025 | 0.011 | 0.010 | 0.010 | 0.122 (0.244) | 0.133 (0.244) | 0.244 |
| Bond lengths (Å) | 1.390 | 2.184 | 1.390 | 2.184 | 1.390 | 2.184 | 1.390 | 2.184 |

### r.m.s.d., root mean square deviation.

Data collection and structure refinement statistics of the SaAraf43A heterologous expression of exo-1,5-α-L-arabinofuranosidase and endo-type enzymes. In the previous work we succeeded in the difference of the substrate recognition mechanisms of exo- and endo-type enzymes in glycoside hydrolase family 43. A mutagenesis study of SaAraf43A provides insights into the mechanism of substrate discrimination between exo- and endo-type enzymes in glycoside hydrolase family 43.
Exo-1,5-α-L-arabinofuranosidase from S. avermitilis

Protein Expression and Mutant Generation—Recombinant SaAraf43A was expressed in Escherichia coli BL21gold (DE3) (Stratagene, La Jolla, CA) using an expression vector pET30 (Novagen, Madison, WI) purified as described previously (18). Amino acid substitutions of SaAraf43A were generated by inverse PCR using pET30/SaAraf43A as template DNA and the appropriate primers (supplemental Table S1). Mutations were confirmed by DNA sequencing. The plasmids were transformed into E. coli BL21gold (DE3). Expression and purification of wild type and mutants were carried out in the same way as for wild type SaAraf43A (18).

Crystallization, Data Collection, and Structure Determination—Crystallization procedures have been reported previously (21). SaAraf43A was crystallized by the sitting-drop vapor-diffusion method with the precipitant solution composed of 0.8 M sodium citrate, 0.2 M sodium chloride, and 0.1 M Tris, pH 7.0. Crystals with maximum dimensions of 200 × 50 × 20 μm were consistently obtained using 100 μl of the reservoir solution with a drop consisting of 2–3 μl of protein solution and 2 μl of reservoir solution at 293 K. Selenomethionine (Se-Met)-labeled SaAraf43A was produced using the E. coli B834 (DE3) methionine auxotroph and crystallized in the same condition as with the native enzymes. Diffraction experiments were conducted at Photon Factory, High Energy Accelerator Research Organization, Tsukuba, Japan and at the large synchrotron radiation facility (SPring-8), Japan Synchrotron Radiation Research Institute, Harima, Japan. Approximately equal volumes of the reservoir solution containing 10% glycerol was added into the crystal drop, scooped in a nylon loop, and flash-frozen in a nitrogen stream at 95 K. Diffraction data were collected with ADSC CCD detectors. Data were integrated and scaled using the program DENZO and Scalepack in the HKL-2000 program suite (24). SaAraf43A crystals diffracted to ~2.0 Å resolution (space group P212121).

Structural analysis of SaAraf43A was conducted through the multilwavelength anomalous dispersion method using Se-Met-labeled SaAraf43A crystals. Four selenium atom positions were determined, and initial phases were calculated using the program SOLVE/RESOLVE (25). Manual model rebuilding, introduction of water molecules, and molecular refinement were conducted using Coot and Refmac5 (26, 27). One sodium ion and 1 chloride ion were added into the model. For the analyses of ligand binding structures of SaAraf43A, α-1,5-L-arabinofuranobiose, α-1,5-L-arabinofuranotriose, or AraAz was dissolved in the crystallization precipitant to the concentration of 5% (w/v); 1 μl of the ligand solution was added to the crystal drops and incubated for 20–30 min, and the crystals were subjected to diffraction experiments. Structural determination was conducted through the molecular replacement method using the ligand-free structure as the starting model. Data collection and structure refinement statistics are given in Table 1. Structural drawings were prepared by the program PyMol (DeLano Scientific LLC, Palo Alto, CA).

Substrate Specificity—To evaluate the catalytic efficiency of SaAraf43A toward α-1,5-linked L-arabinofuranooligosaccharides such as α-1,5-L-arabinofuranobiose, α-1,5-L-arabinofuranotriose, α-1,5-L-arabinofuranotetraose, and α-1,5-L-arabinofuranopentaose, 0.5 nM enzyme was incubated with 10 μM substrate in McIlvaine buffer (0.2 M Na2HPO4, 0.1 M citric acid), pH 6.0, for up to 120 min at 30 °C. After regular time intervals, 100-μl aliquots were taken, and the reaction was stopped by boiling for 5 min. The amount of each undegraded substrate was quantified by high performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) using L-fucose as an internal standard, as previously reported (28). Analysis of the samples was performed using CarboPacTM PA1 column (Dionex, Sunnyvale, CA) as described previously (18).

Substrate specificity of the mutants was analyzed using methyl 2-O-, 3-O-, and 5-O-α-L-arabinofuranosyl-α-L-arabinofuranosides. Briefly, the enzyme (4 nM–4 μM) was incubated with substrate (10 μM) in McIlvaine buffer, pH 6.0, at 30 °C. After regular time intervals, the reaction was stopped by boiling for 5 min, and then the amounts of undegraded sub-
The catalytic activity of SaAraf43A and its mutants was assayed using 100 nmol to 1 μmol of enzyme incubated with 0.25–5 mM PNP-L-Ara in McIlvaine buffer, pH 6.0, in a total volume of 0.5 ml at 30 °C for up to 30 min. The amount of PNP released was determined at 400 nm with an extinction coefficient of 19,608 M\(^{-1}\) cm\(^{-1}\). The assay was performed in triplicate.

RESULTS AND DISCUSSION

Overall Structure of SaAraf43A—The crystal structure of SaAraf43A was determined by the multiwavelength anomalous dispersion method using Se-Met derivative data. Successively, native and three ligand complex structures, SaAraf43A/1,5-L-arabinofuranobiose, SaAraf43A/1,5-L-arabinofuranotriose, and SaAraf43A/Ara Az, were determined. Structure refinement statistics are summarized in Table 1. The quality and accuracy of the final structures were further demonstrated in that more than 98% of their residues fall within the common regions of the Ramachandran stereocchemistry plot. Recombinant SaAraf43A molecule is composed of a single polypeptide chain of 468 amino acids (0–467), where N-terminal Met-0 and C-terminal LEHHHHHH467 were derived from the expression vector and purification tag. The N-terminal 9 residues Met-0—Val-8 and the C-terminal 12 residues Leu-456—His-467 were not identified because of lack of electron density. The final model consisted of 1 SaAraf43A molecule accompanied with 1 sodium ion, 1 chloride ion, and several glycerol molecules.

SaAraf43A is composed of 2 distinct domains (Fig. 1). The N-terminal catalytic domain is composed of a 5-bladed \(\beta\)-propeller that is built of 5 radially oriented “blades” (marked here as I–V, Fig. 1A), distributed almost equally around a full circle. Such a fold was first reported for tachylectin (29) and was found in three glycoside hydrolase families, GH32, GH43, and GH68, represented by invertase, \(-L\)-arabinanase, and levansucrase, respectively. The C-terminal domain is made of three repeated-peptide segments referred to as subdomains A, B, and C (Figs. 1B and 2). A combination of the three subdomains results in a fold similar to the “\(\beta\)-trefoil fold” proposed by Murzin et al. (30). The \(\beta\)-trefoil fold is observed in two families, CBM13 and CBM42.

Sugar complex crystals were prepared by adding the 1,5-L-arabinofuranosidase from S. avermitilis OCTOBER 29, 2010• VOLUME 285 • NUMBER 44 JOURNAL OF BIOCHEMISTRY 34137
**Exo-1,5-α-L-arabinofuranosidase from S. avermitilis**

FIGURE 3. Sugar binding structures in SaCBM42. A–C, bound sugars in subdomains α, β, and γ in the L-arabinotriose complex are shown. D, bound \( \text{Ara} \) in subdomain \( \beta \) of the \( \text{Ara} \) complex is shown. Residues from the crystallographic symmetry-related molecules are shown in white. Hydrogen bonds are shown as cyan dashed lines. 2Fo – Fc, electron density is shown for the bound sugars contoured at the 1.5 level. Carbon atoms are numbered for the arabinofuranosyl moiety at the binding site of subdomain α.

Waals contact with bound sugars in subdomain α. Similar hydrogen bond network and van der Waals contacts were observed in subdomain γ.

The relative positions of the 1-arabinofuranosyl moiety at the non-reducing end of the bound sugars and their binding manners were quite conserved between two subdomains as well with those of AkCBM42, whereas the positions of the reducing-end 1-arabinofuranosyl moiety differed between the two subdomains. This difference was caused by lesser contacts with the binding site and the close contact with symmetry-related molecules. The electron density for the reducing-end 1-arabinofuranose was rather vague, and they were observed to be bound nonspecifically by the crystallographic circumstances. Therefore, the sugar-binding site of SaCBM42 basically recognized only one moiety of \( \alpha, 1,5, \text{L-arabinofuranooligosaccharides} \).

In subdomain \( \beta \) of CBM42, His-381 and Asp-400 hydrogen-bonded to one glycerol molecule (Fig. 3B). This glycerol might have originated from the cryoprotectant. In contrast, in the complex structure with \( \text{Ara} \), the bound \( \text{Ara} \) molecule was observed in subdomain \( \beta \) as well as in subdomains α and γ of SaCBM42. Therefore, subdomain \( \beta \) would have the potential to bind 1-arabinooligosaccharides. The side chains of Tyr-431 and Glu-432 in the symmetry-related molecule were considered to prevent the sugars from binding to the crystal through the sugar-soaking method. The binding manner of \( \text{Ara} \) was similar to those of the other subdomains, and four hydropon bonds were present between the protein and the sugar (Fig. 3D).

The O5 atoms of the bound 1-arabinofuranoses were located at the bottom of the binding site, recognized by two hydrogen bonds. That the O5 atom was buried indicated the sugar-binding sites of SaCBM42 specifically recognizes the terminal 1-arabinofuranosyl moiety of the 1-arabinofuranooligosaccharides and cannot bind the linear 1,5,1-arabinooligosaccharides across the two binding sites. This specificity agreed with the exo-type catalytic mechanism of the enzyme and the binding specificity of SaCBM42 that bound terminal and branched 1-arabinofuranosyl residues, such as arabinan and arabinoxylan, described in the previous paper (18). SaCBM42 would play a role of finding the 1-arabinofuranosyl residues among the insoluble substrate and aid catalysis by the catalytic module. CBM13 also has a β-trefoil fold and is separated into three homologous subdomains, α, β, and γ. We have shown that CBM13 of *S. avermitilis* \( \beta \)-1-arabinofuranosidase or *Streptomyces olivaceoviridis* endo-β-1,4-xylanase also has three sugar-binding sites (31, 32). Each subdomain has one sugar-binding site. In these enzymes, three sugar-binding sites were believed to enhance the possibility of substrate binding. Similarly, the meaning of three binding sites of SaCBM42 should also be increasing the probability of insoluble substrate binding possessing 1-arabinosyl terminus or side chains.

**Structure and Sugar Binding Manner of the Catalytic Module**—The 5-bladed \( \beta \)-propeller fold is a common structure for GH43, which includes inverting enzymes and is classified into clan GH-F together with GH62 (33). Families GH32 and GH68 also share a similar five-bladed \( \beta \)-propeller fold, but they are composed of retaining enzymes and are classified into clan GH-J. GH43 contains both endo- and exo-acting enzymes and has
broad substrate specificities. A structural comparison of the catalytic cleft elucidated the difference in a substrate binding manner (Fig. 5). The active site of SaAraf43A is located on the central cavity of the β-propeller fold, forming a substrate binding pocket (Fig. 5A). This is very typical for the exo-mode action, by which SaAraf43A releases a single L-arabinose unit from the non-reducing end of linear arabinan or arabinooligosaccharides by cleaving the α-1,5-linkage. Geobacillus stearothermophilus β-xylosidase (XynB3) is also an exo-acting enzyme and has a catalytic pocket at the center of the β-propeller fold like SaAraf43A; however, the structures surrounding its catalytic pocket were not conserved (Fig. 5D) (34). In contrast, the endo-type enzymes Bacillus subtilis arabinan endo-α-1,5-L-arabinofuranosidase (BsArb43A) and Cellvibrio japonicus endo/ exo-α-1,5-L-arabinanase (CjArb43A) possess a catalytic cleft across the surface of the β-propeller fold. The clefts of the two enzymes lie in almost the same orientation (Fig. 5B and C) (33, 35) and are deep enough so that the linear substrate could dock through the catalytic center. The length of the cleft of BsArb43A was as long as the 1-arabinopentaose, and multiple subsites specific for arabinooligosaccharides might enable the endo-mode action of the enzyme. In the case of SaAraf43A, the longer loop region Tyr-281–Arg-294 (Asp-283—Asn-290 is shown in magenta in Figs. 1 and 5A) as well as the side chain of Tyr-40 occupied the positions corresponding to the glycon side of the cleft of BsArb43A. This hindrance probably conferred an exo-mode manner on SaAraf43A. B. subtilis arabinoxylan arabinofuranohydrolase (BsAXH-m2,3) also liberated L-arabinose from the substrate, although it is a side-chain-releasing enzyme (36). The xylotetraose-bound structure of BsAXH-m2,3 showed a different orientation of the sugar backbone in comparison with the substrate-bound structures of SaAraf43A or CjArb43A (Figs. 5, A, C, and E). BsAXH43A had a short loop region connecting β-strands 2 and 3 in blade I. In contrast, the corresponding loop region of CjArb43A and SaAraf43A is much longer, and therefore, the linear arabinan backbone could not bind the active site in the same orientation as BsAXH-m2,3. Thus, GH43 enzymes could change their substrate specificities by changing the structure of the loop region of the catalytic domain, yielding a variety of enzymatic activities.

Like other GH43 members, the three important acidic residues, Asp-20, Asp-135, and Glu-196, were conserved in SaAraf43A and were located at equivalent positions in the depth of the catalytic pocket (Figs. 5 and 6). The α-1,5-L-arabinofuranotriose complex structure of SaAraf43A revealed the recognition mechanism of the L-arabinofuranosyl substrate, where three L-arabinofuranosyl moieties were observed (Figs. 4B and 6). The three moieties seemed to be divided into

FIGURE 4. Schematic drawings of the sugar-binding structures in subdomain α of SaCBM42 (A) and catalytic pocket of SaAraf43A (B). Hydrogen bonds between the sugars and proteins were shown as dashed lines with their bond distances (Å). The figure was drawn by the program LIGPLOT (40).

Exo-1,5-α-L-arabinofuranosidase from S. avermitilis
L-arabinose and α-1,5-L-arabinofuranobiose. The non-reducing-end L-arabinofuranose was located at the bottom of the catalytic pocket, whereas the two reducing-end L-arabinofuranose residues were positioned at the outside of the pocket; their positions were called subsite $-1$ and subsites $+1$ and $+2$, respectively, where the boundary was proximal to the side chain of Glu-196. The electron density of L-arabinose at subsite $-1$ or subsite $+2$ was rather obscure, and the average B-factor values were 45.9, 29.9, and 43.9 Å$^2$ for the 3 moieties at subsites $-1$, $+1$, and $+2$. The bound ligand, therefore, could not be identified as a product or a substrate, but the structure of L-arabinose at subsite $-1$ was identified as α-furanoside for the present model. The side chain of Glu-196, which is 1 of the 3 conserved acidic residues, was proximal to the O1 atom of the L-arabinose moiety at subsite $-1$, or the O5 atom at subsite $+1$. Therefore, Glu-196 was considered to be the catalytic acid of the enzyme. Subsite $-1$ consisted of Asp-20, Thr-36, Tyr-40, Trp-76, Leu-134, Asp-135, His-260, and Arg-294. Five hydrogen bonds were observed: between the O2 and Asp-135-O$_{\text{H}11002}$ atoms, between the O3 and Asp-135-O$_{\text{H}11001}$ atoms, and between the O5 atom and the 3 protein atoms Asp-20-O$_{\text{H}11002}$, Arg-294-N$_{\text{H}11001}$, and Arg-294-N$_{\text{H}11002}$. The indole group of Trp-76 provided a platform contact involving the C3-C4-C5 plane. Trp-76 was conserved in α-L-arabinanases of GH43, such as CjArb43A and BsArb43A, and subsite $-1$ could be mentioned to have similar structure among GH43s, although Tyr-40 was not conserved in CjArb43A and BsArb43A. SaAraf43A hydrolyzes α-1,5-linked arabinan in the exo-acting manner and liberates L-arabinose. The exo-acting mechanism was expressed probably because subsite $-1$ was completely surrounded by amino acid residues including the loop region Tyr-281-Arg-294, as mentioned above, and only one L-arabinosyl moiety could be docked in the glycon subsite.

The electron density of the L-arabinofuranosyl moiety at subsite $+1$ was observed clearly, and it has the envelope 1E conformation (Fig. 6). Subsite $+1$ was surrounded by Phe-132, Leu-134, Asn-159, Thr-160, Tyr-192, Val-194, Glu-196, Ala-215, Thr-216, and Leu-289. Hydrophobic interactions were shown to be important in the formation of subsite $+1$. The side chains of Tyr-192 and Val-194 built one side of the cleft, whereas three aromatic side chains, Trp-76, Trp-100, and Phe-132, formed...
the wall on the other side. Hydrogen bonds were observed between the O1 and O3 atoms and the Asn-159-N82 atom and the O5 and Glu-196-O61 atoms. The position of the O5 atom corresponded to the scissile α-1,5-glycoside bond of the catalysis. Another α-1,5-glycoside bond between the l-arabinofuranosyl moiety subsites +1 and +2 was held by the Tyr-192, and its O5 atom was hydrogen-bonded by the Asn-159-N82 atom. The structure of subsite +1 would be important for the substrate specificity against the α-1,5-glycosidic bond. Three bound water molecules were clearly observed, which also mediated hydrogen network to the protein.

The l-arabinofuranosyl moiety at subsite +2 was embedded on the indole ring of Trp-100 and was partially supported by the aromatic rings of Phe-132 and Tyr-192. Including Asn-159, 4 residues in total constructed subsite +2. When the catalytic efficiency of SaAraf43A against α-1,5-linked l-arabinofuranooligosaccharides was investigated, the activity of the enzyme increased according to the degree of polymerization of the substrate, namely $k_{cat}/K_m$ values of 426 min$^{-1}$mM$^{-1}$ for arabinobiose increased to 5,556 min$^{-1}$mM$^{-1}$ for arabinotriose. The hydrolysis rates between arabinofuranose and arabinopentaose were almost the same, implying that the enzyme contains three significant l-arabinose-binding subsites (Table 2). Indeed, the sugar binding structure showed that the active site is composed of three subsites (Figs. 4B and 6). Subsite +2 was located on the surface of the catalytic pocket, and therefore, the activity of SaAraf43A against α-1,5-linked arabinofuranose and arabinopentaose would be almost at the same level.

The O2 and O3 atoms of the arabinofuranose at subsite +2 in the α-1,5-l-arabinofuranosidase complex structure of SaAraf43A were exposed to solvent, whereas those atoms at subsites +1 and −1 were not exposed. The structure indicated that SaAraf43A could not degrade decorated arabinan when its terminal or second l-arabinose was substituted by the sugar side chain, and this is why SaAraf43A did not show hydrolyzing activity against decorated arabinan, reported previously (18).

### TABLE 2

| Degrees of polymerization of α-1,5-linked l-arabinofuranooligosaccharides | $k_{cat}/K_m$ | Log $k_{cat}/K_m$ |
|---|---|---|
| 2 | 426 ± 0 | 2.62 |
| 3 | 5,556 ± 692 | 3.74 |
| 4 | 10,423 ± 1,443 | 4.02 |
| 5 | 9,818 ± 1,590 | 3.99 |

### TABLE 3

| SaAraf43A | $k_{cat}$ | $k_{cat}/K_m$ | Relative to WT |
|---|---|---|---|
| Wild type | 65.5 ± 0.4 | 0.5 ± 0.0 | 13.0 |
| D20A | ND | ND | ND |
| D20N | 0.1 ± 0.0 | 0.1 × 10$^{-2}$ ± 0.0 × 10$^{-2}$ | 0.9 × 10$^{-2}$ ± 0.0 × 10$^{-2}$ |
| D135A | ND | ND | 0.7 × 10$^{-2}$ |
| D135N | ND | ND | 197.1 |
| E196A | ND | ND | 1.5 |
| N195A | 33.5 ± 1.0 | 1.7 ± 0.1 | 248.4 |
| N195L | 273.2 ± 17.1 | 1.1 ± 0.1 | 24.8 |
| Y192A | 63.9 ± 13.1 | 11.5 ± 0.4 | 5.6 |
| E209A | 37.2 ± 3.4 | 40.7 ± 4.0 | 0.9 |

---

**Exo-1,5-α-l-arabinofuranosidase from S. avermitilis**

**Mutagenesis Study of SaAraf43A**—To verify the catalytic residues of SaAraf43A, some mutants were generated by PCR at the putative catalytic centers, Asp-20, Asp-135, and Glu-196. These residues hydrogen bonded to the l-arabinofuranosyl moiety at subsite −1 in the α-1,5-l-arabinotriose complex structure and were conserved among GH43 enzymes. Glu-196 was considered to be the catalytic acid of the enzyme as the side chain of Glu-196 was located close to the O1 atom of the l-arabinose moiety at subsite −1. Enzymatic analyses showed that the catalytic activities of the single amino acid variants, D20A, D135A, D135N, and E196A, vanished, and these residues were proven to be indispensable for catalysis (Table 3). However, only D20N maintained an extremely weak activity; mutation of this amino acid reduced the values of $k_{cat}$ and $K_m$, suggesting that the mutant has an increased affinity for a substrate but hardly released a product. The Asp-20-O62 atom hydrogen-bonded to the O5 atom of l-arabinofuranosyl moiety at subsite −1, and the Asp-20-O62 atom was 4.4 Å away from the anomeric C1 atom. However, there was one bound water molecule hydrogen bonded to the Asp-20-O62 atom and located close to the C1 atom (2.7 Å distance; Fig. 4B). Catalysis by family 43 enzymes occurs with inversion of the anomeric configuration (37), and this water could be considered to play as a catalytic nucleophile, representative for the inversion mechanism. When Asp-20 was substituted by asparagine, nucleophilicity of the water would decrease, but it was considered to maintain the attacking ability, resulting to the weak activity of the mutant. Mutagenesis studies of CjArb43A suggested that Glu-221 functions as a catalytic acid, Asp-38 functions as a catalytic base, and Asp-158 functions in $K_m$ modulation and in maintaining the correct orientation of the general acid residue (33). Our analyses on SaAraf43A provided equivalent results.

The substrate recognition mechanism of exo-1,5-α-l-arabinofuranosidase was also assessed by mutagenesis. According to the above structural analysis, three amino acids, Asn-159, Tyr-192, and Leu-289, were selected for the mutagenesis study. First, kinetic analysis with PNP-α-l-Araf was performed (Table 3). The $K_m$ values of N159A and N159L mutants were 1.7 and 1.1 mM, respectively, which were almost similar to the $K_m$ value of the wild type enzyme. However, the $k_{cat}$ values of the mutants were 5–40 times larger than that of the wild type. The hydrolysis activities of the N159A and N159L mutants were higher than those of the wild type enzyme, suggesting that Asn-159 has no considerable effect on substrate binding and probably discriminates l-arabinofuranose from the other sugars. The muta-
tion of Asn-159 vanished two hydrogen bonds to the substrate but created hydrophobic environment at subsite +1, which would be suitable for the binding of hydrophobic substances such as PNP. Hydrophobicity produced by leucine or alanine might enhance the hydrolytic activity against the PNP-α-1,5-Araf. Unlike Asn-159 mutants, the $k_{\text{cat}}$ values of Y192A and L289A were increased; however, the $K_m$ values of the mutants were 20–40 times larger than that of the wild type. The activities of Y192A and L289A were lower than that of the wild type enzyme, suggesting that Tyr-192 and Leu-289 were involved in substrate binding. As shown by the SaAraf43A/α-1,5-L-arabinofuranosyl complex structure, Tyr-192 contacts with the plane of the α-1,5-glycoside bond between the L-arabinofuranosyl moieties of subsites 1 and 2, but the side-chain Leu-289 was located away from the bound sugars with distances of more than 6 Å. Thereby, the position of Leu-289 was examined in the other structural models, ligand-free and sugar complex structures with α-1,5-L-arabinofuranobiose and AraAz (Fig. 7). The region of Asp-283–Asn-290 formed an extended loop, and Pro-288—Asn-290 formed a 310-helix on the tip of the loop (Fig. 1). The position of the loop point was different between the structures, and the largest shift of 2.6 Å was observed for the main chain of Asp-287. The B-factor values for this region were higher than the average values, and this loop was considered to be flexible. The side chain of Leu-289 was considered to possibly move toward subsites 1 and 1, and the estimated distance to the bound sugars was minimized to 4.2 Å as judged from the superposed model, which could be considered as a hydrophobic contact. Therefore, this loop region, containing Leu-289, was shown to bind the substrate by capping the subsites and also to play a role in uptake the substrate by taking an open form.

Second, the detailed substrate specificity of the enzymes was characterized by using various L-arabinofuranobiose with different types of glycosidic bond. The hydrolysis rate of three kinds of regioisomer of biosides by wild type and mutants are shown in Table 4. Interestingly, unlike against PNP-α-1,5-Araf, the activity of all the mutants against α-1,5-linked L-arabinofuranobiose decreased compared with the wild type. Although Asn-159 mutants exhibited higher activities toward PNP-α-1,5-Araf than the wild type enzyme, their activity toward α-1,5-linked L-arabinofuranobiose was lower than the wild type enzyme. Asn-159 is considered in the form making a suitable space for the fitting of the natural substrate such as α-1,5-linked L-arabinofuranobiose by forming hydrogen bonds at subsite +1. For Y192A and L289A mutants, a low activity toward both PNP-α-1,5-Araf and α-1,5-linked L-arabinofuranobiose was observed; however, improved substrate specificities are shown in Table 4 and Fig. 8. The mutants showed activity against α-1,2-linked L-arabinofuranobiose.

### Table 4

| SaAraf43A       | Activity $k_{\text{cat}}/K_m$ | α1,5 | α1,2 | α1,3 |
|----------------|-------------------------------|------|------|------|
| Wild type      | 3816 ± 434                    | ND   | ND   | ND   |
| N159A          | 107 ± 7                       | ND   | ND   | ND   |
| N159L          | 11 ± 0                        | ND   | ND   | ND   |
| Y192A          | 669 ± 38                      | 0.5 ± 0 | 0.0 | ND   |
| L289A          | 4 ± 0                         | 0.5 ± 0 | 0.0 | ND   |

FIGURE 7. Superimposed model of ligand-free (pale pink) and sugar complex structures of SaAraf43A with α-1,5-L-arabinobiose (yellow), α-1,5-L-arabinotriose (magenta), and AraAz (orange) around subsite +1 of the catalytic domain. Bound sugars are shown from the SaAraf43A/α-1,5-L-arabinofuranobiose complex structure.

FIGURE 8. HPAEC-PAD analysis of hydrolysis products of α-1,2-linked L-arabinofuranobiose generated by the SaAraf43A mutants (A) Y192A and (B) L289A. Fuc, L-fucose; Ara, L-arabinose; Ara(1→2)Ara, α-1,2-linked L-arabinofuranobiose.
in addition to their activity against α-1,5-linked L-arabinofuranobiose. As described above, Tyr-192 and Leu-289 regulate the substrate specificity of SaAraf43A. Tyr-192 is considered to fix the substrates at subsites +1 and +2 by hydrophobic interaction with its aromatic ring. In the Y192A mutant, strictness of the subsite +1 decreased, and α-1,2-linked L-arabinofuranobiose could be bound at subsite −1 and +1. In contrast, Leu-289 did not show the direct interaction with substrates, but it would have an important role to fix the substrates at both subsites −1 and +1, as mentioned previously. L289A mutation would decrease rigidity of the substrates around these subsites. Therefore, the difference between the endo- and exo-modes of catalysis was basically different in different enzymes.

BsArb43A, which showed endo-type catalytic properties, would have an important role to fix the substrates at both sub-sites by binding the L-arabinose moiety of an insoluble substrate complex, the position of the xylose at subsite +1 is maintained by a large number of hydrogen bonds, whereas the aglycon xylose unit at subsite +1 was bound much less tightly (34). In the case of SaAraf43A, the residues surrounding subsite +1, such as Tyr-192 and Leu-289, would cause tight substrate specificity.

This study provides a good model for investigating the hydrolytic mechanism of the modular enzyme. As the function of SaCBM42 has been demonstrated biochemically (18), our structural analysis showed that SaCBM42 assists in the catalysis by binding the l-arabinose moiety of an insoluble substrate through three arabinan-binding sites. This study also provides a good model for investigating the mode of action of exo- and endo-type enzymes because the GH43 family includes different types of arabinan-degrading enzymes. For example, amino acids involved in aglycon subsites +1 and +2 and their SaAraf43A structures were not conserved in CjArb43A or BsArb43A, which showed endo-type catalytic properties. Therefore, the difference between the endo- and exo-modes of action against arabinan or arabinooligosaccharides in GH43 was attributed not only to the structure of the glycon site but also to the fact that the overall structures of the catalytic clefts were basically different in different enzymes.

Acknowledgments—We thank the beamline researchers at Photon Factory and SPring-8 for the use of synchrotron radiation. We thank Dr. Ryoji Mizuno for technical support for mutant construction.

REFERENCES
1. Carpita, N. C., and Gibeaut, D. M. (1993) Plant J. 3, 1–30
2. Saha, B. C. (2003) J Ind. Microbiol. Biotechnol. 30, 279–291
3. Aristiidou, A., and Penttilä, M. (2000) Curr. Opin. Biotechnol. 11, 187–198
4. Skoog, K., and Hahn-Hägerdal, B. (1988) Enzyme Microb. Technol. 10, 66–80
5. Seri, K., Sanai, K., Matsuo, N., Kawakubo, K., Xue, C., and Inoue, S. (1996) Metabolism 45, 1368–1374
6. Sanai, K., Seri, K., and Inoue, S. (1997) J. Jpn. Soc. Nutr. Food Sci. 50, 133–137
7. Osaki, S., Kimura, T., Sugimoto, T., Hizukuri, S., and Iritani, N. (2001) J. Biol. Chem. 276, 39–47
8. Kaneko, S., Shimasaki, T., and Kusakabe, I. (1993) Biosci. Biotechnol. Biochem. 57, 1161–1165
9. Kaneko, S., Sano, M., and Kusakabe, I. (1994) Appl. Environ. Microbiol. 60, 3425–3428
10. Kaneko, S., and Kusakabe, I. (1995) Biosci. Biotechnol. Biochem. 59, 2132–2133
11. Kaneko, S., Kuno, A., Matsuo, N., Ishii, T., Kobayashi, H., Hayashi, K., and Kusakabe, I. (1998) Biosci. Biotechnol. Biochem. 62, 2205–2210
12. Kaneko, S., Higashi, K., Yasui, T., and Kusakabe, I. (1998) J. Ferment. Bioeng. 85, 518–520
13. Kaneko, S., Arimoto, M., Ohba, M., Kobayashi, H., Ishii, T., and Kusakabe, I. (1998) Appl. Environ. Microbiol. 64, 4021–4027
14. Kaneko, S., Kuno, A., Kobayashi, H., and Kusakabe, I. (1998) Biosci. Biotechnol. Biochem. 62, 695–699
15. Matsuo, N., Kaneko, S., Kuno, A., Kobayashi, H., and Kusakabe, I. (2000) Biochem. J. 346, 9–15
16. Henris tat, B. (1991) Biochem. J. 280, 309–316
17. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) Nucleic Acids Res. 37, D233–D238
18. Ichinose, H., Yoshida, M., Fujimoto, Z., and Kaneko, S. (2008) Appl. Microbiol. Biotechnol. 80, 399–408
19. Miyanaga, A., Koseki, T., Matsuzawa, H., Wakagi, T., Shoun, H., and Fushinobu, S. (2004) J. Biol. Chem. 279, 44907–44914
20. Miyanaga, A., Koseki, T., Miwa, Y., Mese, Y., Nakamura, S., Kuno, A., Hirabayashi, J., Matsuzawa, H., Wakagi, T., Shoun, H., and Fushinobu, S. (2006) Biochem. J. 399, 503–511
21. Fujimoto, Z., Ichinose, H., and Kaneko, S. (2008) Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 64, 1007–1009
22. Kawabata, Y., Kaneko, S., Kusakabe, I., and Gama, Y. (1995) Carbohydr. Res. 267, 39–47
23. Berlin, W., and Sauer, B. (1996) Anal. Biochem. 243, 171–175
24. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
25. Terwilliger, T. C. (2005) Methods Enzymol. 374, 22–37
26. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132
27. Marshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. D Biol. Crystallogr. 53, 240–255
28. Ichinose, H., Yoshida, M., Kotake, T., Kuno, A., Igarashi, K., Tsumuraya, Y., Samejima, M., Hirabayashi, J., Kobayashi, H., and Kaneko, S. (2005) J. Biol. Chem. 280, 25820–25829
29. Beisel, H. G., Kawabata, S., Ivanaga, S., Huber, R., and Bode, W. (1999) EMBO J. 18, 2313–2322
30. Murzin, A. G., Lesk, A. M., and Chothia, C. (1992) J. Mol. Biol. 223, 531–543
31. Ichinose, H., Fujimoto, Z., Honda, M., Harazono, K., Nishimoto, Y., Uzura, A., and Kaneko, S. (2009) J. Biol. Chem. 284, 25097–25106
32. Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I., and Mizuno, H. (2002) J. Mol. Biol. 316, 65–78
33. Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., Taylor, E. J., Gilbert, H. J., and Davies, G. J. (2002) Nat. Struct. Biol. 9, 665–668
34. Brix, C., Ben-David, A., Shallow-Shemtov, D., Leon, M., Nefiein, K., Shoham, G., Shoham, Y., and Schomburg, D. (2006) J. Mol. Biol. 359, 97–109
35. Proctor, M. R., Taylor, E. J., Nurizzo, D., Turkenburg, J. P., Lloyd, R. M., Vardakou, M., Davies, G. J., and Gilbert, H. J. (2005) Proc. Natl. Acad. Sci. U.S.A. 102, 2697–2702
36. Vandermarliere, E., Bourgeois, T. M., Winn, D. M., van Campenhout, S., Volckaert, G., Delcour, J. A., Strelkov, S. V., Rabijns, A., and Courtin, C. M. (2006) J. Mol. Biol. 360, 531–543
37. Pitson, S. M., Vooragen, A. G., and Geldman, G. (1996) FEBS Lett. 398, 7–11
38. Shallow-Shemtov, D., Leon, M., Bravman, T., Ben-David, A., Zaide, G., Belakov, V., Shoham, G., Schomburg, D., Baasov, T., and Shoham, Y. (2005) Biochemistry 44, 387–397
39. Kabsch, W., and Sander, C. (1983) Biopolymers 22, 2577–2637
40. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) Protein Eng 8, 127–134