Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

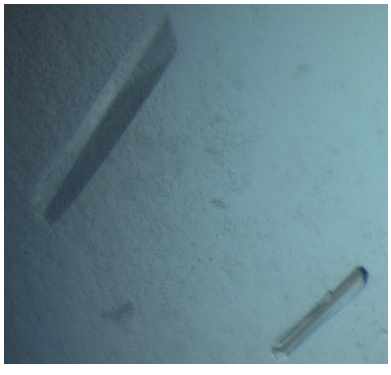
Crystallization and preliminary crystallographic analysis of β-L-arabinopyranosidase from *Streptomyces avermitilis* NBRC14893

Zui Fujimoto,[a] Hitomi Ichinose,[b] Koichi Harazono,[c] Mariko Honda,[b] Atsuko Uzura[c] and Satoshi Kaneko[b]*

[a]Protein Research Unit, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan, [b]Food Biotechnology Division, National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan, and [c]Research and Development Center, Nagase and Co. Ltd, 2-2-3 Murotani, Nishi-ku, Kobe, Hyogo 651-2241, Japan

Correspondence e-mail: sakaneko@affrc.go.jp

Received 2 March 2009
Accepted 7 May 2009



© 2009 International Union of Crystallography
All rights reserved

β-L-Arabinopyranosidase from *Streptomyces avermitilis* NBRC14893 is a monomeric protein consisting of a catalytic domain belonging to glycosyl hydrolase family 27, an unknown domain and a substrate-binding domain belonging to carbohydrate-binding module family 13. The complete enzyme (residues 45–658) has successfully been cloned and homologously expressed in the *Streptomyces* expression system. β-L-Arabinopyranosidase was crystallized by the sitting-drop vapour-diffusion method. The crystals diffracted to 1.6 Å resolution and belonged to space group $P2_12_12_1$, with unit-cell parameters a = 68.2, b = 98.9, c = 181.3 Å. The Matthews coefficient was calculated to be 2.38 Å^3 Da^{-1}.

1. Introduction

Arabinogalactan proteins (AGPs) are a family of proteoglycans that are localized on the cell surfaces of higher plants. They play important roles in root formation, promotion of somatic embryogenesis and attachment of pollen tubes (Seifert & Roberts, 2007). AGPs are characterized by the presence of large amounts of carbohydrates such as D-galactose and L-arabinose and the amino acids hydroxyproline, serine, threonine, alanine and glycine (Knox, 1995). Type II arabinogalactans and short oligosaccharides are the two types of carbohydrate that are attached to the AGP backbone. Type II arabinogalactans have β-1,3-linked D-galactosyl backbones in mono- or oligo-β-1,6-D-galactosyl and/or L-arabinosyl side chains. L-Arabinose and lesser amounts of other auxiliary sugars, such as D-glucuronic acid, L-rhamnose and L-fucose, are attached to the side chains primarily at nonreducing termini (Fincher *et al.*, 1974; Gaspar *et al.*, 2001). The complexity of AGPs complicates the structural and functional analysis of individual AGPs.

We are studying the glycoside hydrolases that degrade AGP sugar linkages. To date, we have cloned two enzymes: exo-β-1,3-galactanase and endo-β-1,6-galactanase. These hydrolyze β-1,3-galactan and β-1,6-galactan, respectively, which are the backbones of the type II arabinogalactan of the AGP sugar linkage (Ichinose *et al.*, 2005; Kotake *et al.*, 2004). In the course of research on AGP-degrading enzymes, we found that *Streptomyces avermitilis* NBRC14893 possesses both an exo-β-1,3-galactanase and an endo-β-1,6-galactanase (Ichinose *et al.*, 2006, 2008). Because these enzymes work more effectively when sugars such as L-arabinose and L-rhamnose are removed from the galactan backbone, we have focused on this bacterium as a source of hydrolases that degrade the sugar linkages of AGPs. We have cultivated this bacterium using gum arabic, a kind of AGP, as a carbon source and isolated a novel β-L-arabinopyranosidase. Because only one β-L-arabinopyranosidase (EC 3.2.1.88) with this enzymatic activity has been reported (Dey, 1973), we cloned the β-L-arabinopyranosidase gene from *S. avermitilis* NBRC14893. Analysis of the amino-acid sequence of this enzyme showed that it has a modular structure containing a catalytic module belonging to glycoside hydrolase family 27 (GH27), an unknown module of about 100 amino acids in length and a carbohydrate-binding module belonging to family 13 (CBM13). The classification of the protein

domains is based on the CAZY database (Cantarel *et al.*, 2008; http://www.cazy.org/).

The GH27 family includes α-galactosidase, α-*N*-acetylgalactosaminidase and isomalto-dextranase. To date, five crystal structures have been determined of GH27-family members. Among these five, the amino-acid sequence of β-L-arabinopyranosidase has the highest homology to that of rice α-galactosidase (40% identity, 54% similarity; Fujimoto *et al.*, 2003). The substrates of these enzymes, β-L-arabinopyranose and α-D-galactose, are differentiated by hydroxymethylation of the fifth C atom. The catalytic residues and the surrounding amino acids are well conserved.

In addition to the GH27 catalytic domain, β-L-arabinopyranosidase has two extra domains at the C-terminal end: an unknown domain and a CBM13 domain. The three-dimensional structure of a CBM13 domain was first determined for the plant galactose-binding lectin ricin B chain. The CBM13 domain has since been identified in many glycoside hydrolases such as xylanases (Boraston *et al.*, 2000; Fujimoto *et al.*, 2002). The crystal structure of β-L-arabinopyranosidase will clarify the function of CBM13, as well as the unknown domain, in β-L-arabinopyranosidase. In order to investigate the substrate-recognition mechanism of β-L-arabinopyranosidase, we therefore conducted crystallization trials on this enzyme.

2. Materials and methods

The gene encoding a putative β-L-arabinopyranosidase (SAV_2186; GenBank accession No. BAC69897) was amplified from *S. avermitilis* genomic DNA by PCR using Phusion DNA polymerase (Finnzymes, Espoo, Finland) and the following primers: forward, 5′-**CAT ATG** TTT CCC CCA CCG CGC TCG ACC-3′, and reverse, 5′-**AAG CTT** TTA CGT CAC CGT CCA CTT CTG GTT-3′. The amplified DNA was digested with *Nde*I (bold) and *Hin*dIII (bold) and cloned into the *S. cinnamoneus* metalloendopeptidase expression system (Hatanaka *et al.*, 2008) derived from pIJ702. The plasmid was transformed into *S. lividans* 1326, which served as the expression host. The transformants were grown in tryptic soy broth containing 50 µg kanamycin at 303 K for 3 d. The recombinant protein consisting of residues 45–658 was secreted at high levels in the culture broth. After removal of bacterial cells by centrifugation, the recombinant enzyme was purified using lactosyl-Sepharose as described previously (Ito *et al.*, 2004). The eluted protein was identified by SDS–PAGE (Fig. 1) and the relevant fractions were pooled and dialyzed against 20 m*M* Tris–HCl buffer pH 7.0.

The protein solution was concentrated to 5.7 mg ml^{-1} by ultrafiltration using a YM-10 membrane (Millipore, Massachusetts, USA) and filtered through a 0.1 µm membrane (Millipore, Massachusetts, USA). The protein buffer was exchanged for water during ultrafiltration prior to crystallization trials. Sparse-matrix crystal screening was performed using JCSG Core Suites I, II, III and IV (Qiagen, Hilden, Germany). Sitting-drop vapour-diffusion trials were set up in 96-well Intelli-Plates (Art Robbins Instruments, California, USA) at 293 K. The reservoirs contained 50 µl reservoir solution and each drop consisted of 0.3 µl protein solution and 0.3 µl reservoir solution. Within a few days, several rod-shaped crystals were observed under conditions consisting of 1.7%(*v*/*v*) PEG 400, 1.7 *M* ammonium sulfate, 15%(*v*/*v*) glycerol and 0.1 *M* HEPES buffer pH 7.5. Crystallization conditions were optimized manually by refinement of the pH using CrystalClear Strips 96-well sitting-drop plates (Douglas Instruments, Berkshire, England).

Native diffraction data were collected from a single crystal on beamline BL-17 of the Photon Factory (PF), High Energy Accelerator Research Organization, Tsukuba, Japan. Since the crystals were grown in a cryoprotectant condition containing 15% glycerol, it was possible to cool them directly without any soaking step. We first attempted to flash-cool the crystal in a loop, but the crystal exhibited high mosaicity. Therefore, a new crystal was mounted in a quartz glass capillary of 0.3 mm diameter. The mother liquor was removed from around the crystal and the capillary was placed into a cryostream at 95 K. This flash-cooling in a capillary method improved the mosaicity and the resolution of the data and reduced X-ray damage to the crystal. Similar results have been published previously (Yao *et al.*, 2004). Diffraction data were collected with 5 s exposures for 1° oscillations over a total of 360° at a wavelength of 0.983 Å with a Quantum Q270 CCD detector (ADSC, California, USA). Data were integrated and scaled using the programs *DENZO* and *SCALEPACK* from the *HKL*-2000 program suite (Otwinowski & Minor, 1997).

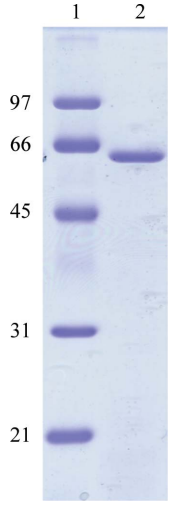


Figure 1
SDS–PAGE of recombinant β-L-arabinopyranosidase. The protein was subjected to SDS–PAGE on a 12% polyacrylamide gel and the bands were visualized by Coomassie Brilliant Blue R-250 staining. Lane 1, molecular-weight markers (1 µg each); lane 2, purified recombinant β-L-arabinopyranosidase (1 µg). Molecular weights are given in kDa at the side of the gel.

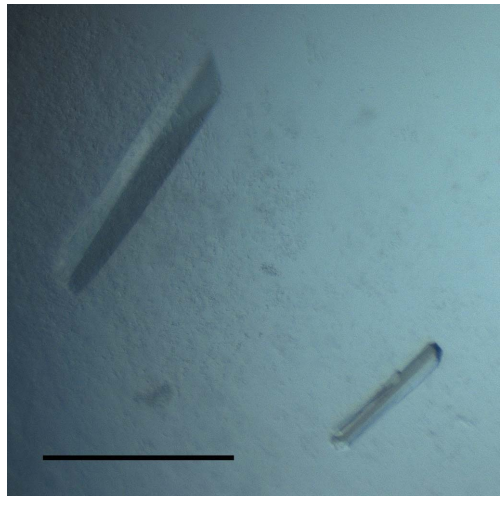


Figure 2
Crystals of β-L-arabinopyranosidase. The scale bar represents 500 µm.

Table 1
Data-collection statistics of the β-L-arabinopyranosidase crystal.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 68.2, b = 98.9, c = 181.3
Wavelength (Å)	0.983
Resolution range (Å)	50.0–1.6 (1.66–1.60)
R_{merge}†	0.081 (0.255)
Completeness (%)	98.8 (99.9)
Multiplicity	8.4 (8.3)
Average $I/\sigma(I)$	33.0 (9.9)
Unique reflections	160762 (16056)
Total reflections	1354822

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations i of reflection hkl.

3. Results and discussion

The optimized crystallization conditions of β-L-arabinopyranosidase were as follows: 1.7%(*v*/*v*) PEG 400, 1.7 *M* ammonium sulfate, 15%(*v*/*v*) glycerol and 0.1 *M* HEPES buffer pH 8.0. Rod-shaped crystals were obtained in a drop consisting of 3 µl protein solution and 2 µl reservoir solution equilibrated against 100 µl reservoir solution. The largest crystal grew to dimensions of 700 × 100 × 70 µm within two weeks (Fig. 2).

Using synchrotron radiation, the crystal diffracted to a maximum resolution of 1.6 Å. The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters a = 68.2, b = 98.9, c = 181.3 Å. The processing statistics of the collected data are summarized in Table 1. The Matthews coefficient was calculated to be 2.38 Å^3 Da^{-1} (Matthews, 1968), corresponding to a solvent content of 48.3% and an asymmetric unit containing two β-L-arabinopyranosidase molecules.

The molecular-replacement method was employed using the structure of rice α-galactosidase (PDB code 1uas) as a search model and structure analysis is ongoing. The structure-based interpretation will provide us with further information on how the modular structure of this enzyme functions towards the complex AGP substrate.

We would like to thank the scientists and staff at the Photon Factory. This work was financially supported in part by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science.

References

Boraston, A. B., Tomme, P., Amandoron, E. A. & Kilburn, D. G. (2000). *Biochem. J.* **350**, 933–941.
Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2008). *Nucleic Acids Res.* **37**, D233–D238.
Dey, P. M. (1973). *Biochim. Biophys. Acta*, **302**, 393–398.
Fincher, G. B., Sawyer, W. H. & Stone, B. A. (1974). *Biochem. J.* **139**, 535–545.
Fujimoto, Z., Kaneko, S., Momma, M., Kobayashi, H. & Mizuno, H. (2003). *J. Biol. Chem.* **278**, 20313–20318.
Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I. & Mizuno, H. (2002). *J. Mol. Biol.* **316**, 65–78.
Gaspar, Y., Johnson, K. L., McKenna, J. A., Bacic, A. & Schultz, C. J. (2001). *Plant Mol. Biol.* **47**, 161–176.
Hatanaka, T., Onaka, H., Arima, J., Uraji, M., Uesugi, Y., Usuki, H., Nishimoto, Y. & Iwabuchi, M. (2008). *Protein Expr. Purif.* **62**, 244–248.
Ichinose, H., Kotake, T., Tsumuraya, Y. & Kaneko, S. (2006). *Biosci. Biotechnol. Biochem.* **70**, 2745–2750.
Ichinose, H., Kotake, T., Tsumuraya, Y. & Kaneko, S. (2008). *Appl. Environ. Microbiol.* **74**, 2379–2383.
Ichinose, H., Yoshida, M., Kotake, T., Kuno, A., Igarashi, Y., Tsumuraya, Y., Samejima, M., Hirabayashi, J., Kobayashi, H. & Kaneko, S. (2005). *J. Biol. Chem.* **280**, 25820–25829.
Ito, S., Suzuki, R., Kaneko, S., Kawabata, Y., Kusakabe, I. & Hasegawa, T. (2004). *J. Biotechnol.* **110**, 137–142.
Knox, J. P. (1995). *FASEB J.* **9**, 1004–1012.
Kotake, T., Kaneko, S., Kubomoto, A., Haque, M. A., Kobayashi, H. & Tsumuraya, Y. (2004). *Biochem. J.* **377**, 749–755.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Seifert, G. J. & Roberts, K. (2007). *Annu. Rev. Plant Biol.* **58**, 137–161.
Yao, M., Yasutake, Y. & Tanaka, I. (2004). *Acta Cryst.* D**60**, 39–45.