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

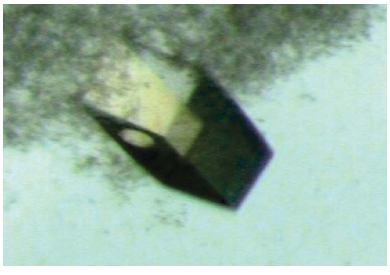
Crystallization and preliminary X-ray diffraction studies of two thermostable α-galactosidases from glycoside hydrolase family 36

M. Foucault,[a] H. Watzlawick,[b] R. Mattes,[b] R. Haser[a] and P. Gouet[a]*

[a]Institut de Biologie et Chimie des Protéines, CNRS–UCBL, UMR 5086, Laboratoire de Bio-Cristallographie IFR128 'BioSciences Lyon-Gerland', 7 Passage du Vercors, 69367 Lyon CEDEX 07, France, and [b]Institut für Industrielle Genetik, Universität Stuttgart, Allmandring 31, D-70569 Stuttgart, Germany

Correspondence e-mail: p.gouet@ibcp.fr

Received 28 October 2005
Accepted 20 December 2005



α-Galactosidases from thermophilic organisms have gained interest owing to their applications in the sugar industry. The α-galactosidases AgaA, AgaB and AgaA A355E mutant from *Geobacillus stearothermophilus* have been overexpressed in *Escherichia coli*. Crystals of AgaB and AgaA A355E have been obtained by the vapour-diffusion method and synchrotron data have been collected to 2.0 and 2.8 Å resolution, respectively. Crystals of AgaB belong to space group $I222$ or $I2_12_12_1$, with unit-cell parameters a = 87.5, b = 113.3, c = 161.6 Å. Crystals of AgaA A355E belong to space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b$ = 150.1, c = 233.2 Å.

1. Introduction

α-Galactosidases (α-D-galactoside galactohydrolases; melibiases; EC 3.2.1.22) are widely distributed in eukaryotes and bacteria. They catalyze the hydrolysis of α-galactose residues from oligosaccharides such as melibiose (galactose-α-1,6-glucose), raffinose (galactose-α-1,6-sucrose) and stachyose (galactose-α-1,6-raffinose) as well as from polymeric galactomannans (Fridjonsson, Watzlawick, Gehweiler & Mattes, 1999; Fridjonsson, Watzlawick, Gehweiler, Rohrhirsch *et al.*, 1999; Luonteri *et al.*, 1998; Margolles-Clark *et al.*, 1996; Shibuya *et al.*, 1995, 1997).

They are widely used in the sugar industry, especially for the elimination of D-raffinose from sugar-beet syrup in order to improve the yield of sucrose (Ganter *et al.*, 1988). This process requires high temperatures; thus, thermostable α-galactosidases have a greater potential than their mesophilic counterparts for use in industrial applications (Fridjonsson, Watzlawick, Gehweiler & Mattes, 1999).

Glycoside hydrolases have been divided into 87 families in the CAZy database according to both catalytic and homology criteria (Coutinho & Henrissat, 1999; Davies & Henrissat, 1995; Henrissat & Coutinho, 2001). α-Galactosidases are found in families 4, 27, 36 and 57. Family 27 has been extensively studied and four crystallographic structures have been determined: chicken *N*-acetylgalactosaminidase (PDB code 1ktb; Garman *et al.*, 2002), rice α-galactosidase (PDB code 1uas; Fujimoto *et al.*, 2003), the fungus *Trichoderma reesei* α-galactosidase (PDB code 1szn; Golubev *et al.*, 2004) and human α-galactosidase (PDB code 1r46; Garman & Garboczi, 2004). All these structures display the same $(\beta/\alpha)_8$-barrel topology and similar active sites. Their catalytic activity is based on a retention mechanism, which involves an aspartate as nucleophile and a proton donor. Little is known about the mechanism of α-galactosidases from families 4, 36 and 57. However, the crystallographic structure of the α-galactosidase from *Thermotoga maritima*, which belongs to family 36, has recently been deposited in the PDB with code 1zy9. The corresponding article has not yet been published, but it can be noticed that this structure, consisting of 564 residues, contains a central domain of 300 residues arranged with the same $(\beta/\alpha)_8$ topology as found in family 27.

The present α-galactosidases are encoded by the genes *AgaA* and *AgaB* from *Geobacillus stearothermophilus* strain KVE39 isolated from Icelandic hot springs (Ganter *et al.*, 1988). The proteins are named AgaA and AgaB, are composed of 729 amino acids each and share an identity of 97%. They belong to family 36, which also

includes α-*N*-acetylgalactosaminidases (EC 3.2.1.49), stachyose synthases (EC 2.4.1.67) and raffinose synthases (EC 2.4.1.82).

AgaA is of great interest for industrial applications because it is highly stable and completely active at 338 K under manufacturing conditions. AgaB has a lower affinity towards D-raffinose and its maximum activity is reached at 323 K. However, it is a powerful tool for the enzymatic synthesis of disaccharides, which are difficult to obtain *via* classical organic synthesis methodologies (Spangenberg *et al.*, 2000). Indeed, AgaB catalyzes transglycosylation reactions and produces 4-nitrophenyl-α-D-galactopyranosyl-(1,6)-α-D-galactopyranoside from 4-nitrophenyl-α-D-galactoside in high yield (Spangenberg *et al.*, 2000). Interestingly, a single mutant of AgaA, AgaA A355E, exhibits catalytic properties that are similar to those of AgaB (Mattes, personal communication). Here, we report the crystallization and preliminary crystallographic analysis of AgaB and AgaA A355E from *G. stearothermophilus*.

2. Materials and methods

2.1. Enzyme assay

α-Galactosidase activity was determined by measuring the rate of *para*-nitrophenyl-α-galactoside (pNPG) hydrolysis (4 mg ml^{-1}) in 0.1 *M* potassium phosphate buffer pH 6.5 at 298 K. One unit is defined as the amount of enzyme required to hydrolyze 1 µ*M* of pNPG per minute. The protein concentration was determined by the method of Bradford (1976) using bovine serum albumin as a standard.

2.2. Construction of the expression vectors

For expression of the α-galactosidase genes from the *G. stearothermophilus* in *Escherichia coli*, the genes were cloned in pBTac1 (Boehringer Mannheim). The coding sequence of *aga*B was PCR amplified from plasmid pCG3 (Watzlawick *et al.*, 1991) by using primer S512 (CGGAATTCTTATGGCGGTTACATAC) with an *Eco*RI linker and the reverse primer S511 (CATGTTAACTTATTGTTGAACAGC). Subsequently, the amplified gene fragment was modified by Klenow fill-in reaction and *Eco*RI digestion. The modified fragment was ligated to pAMG20, a derivative of pBTac1 with a deleted *Hin*dIII restriction site, which had been cut with *Eco*RI and *Sma*I. The final construct pAMG22 was used for expression of *aga*B in *E. coli* RM448.

The *aga*A gene was cloned in pBTac vector in a similar manner by using primer S510 (CGGAATTCTTATGTCAGTTGCATAC) and the reverse primer S511 for amplification of the coding *aga*A sequence from plasmid pCG1 (Ganter *et al.*, 1988) to produce pAMG21. Replacement of the codon for Ala355 by glutamic acid in the *aga*A gene was performed according to an oligonucleotide-directed *in vitro* mutagenesis protocol (Amersham Biosciences). A 687 bp *Mlu*I/*Hin*dIII DNA fragment from the M13 derivative, carrying the codon for glutamic acid (GAG) at amino-acid position AgaA 355, was replaced in pAMG21 to create pHWG8.

E. coli RM448 transformed with pAMG22 and pHWG8, respectively, was used for expression of the α-galactosidases AgaB and AgaA A355E mutant. Cells were grown for 12 h at 310 K in 200 ml 2×YT medium containing 100 µg ml^{-1} ampicillin. Cells were harvested by centrifugation at 5858*g* at 277 K for 15 min, washed and suspended in 10 m*M* potassium phosphate buffer pH 6.5 and stored at 253 K until use.

2.3. Purification

All purification procedures were performed at 291 K using an automated FPLC system (Amersham Biosciences).

Frozen recombinant *E. coli* cells were thawed and the cells were broken by passing them twice through a French press cell (Aminco, SLM Instruments Inc) at 6.9 MPa. The cell-free extract was obtained from the supernatant following centrifugation at 11 950*g* for 30 min at 277 K.

The proteins in the supernatant were fractionated by anion-exchange chromatography. Crude cell-free extract (50 mg) was applied onto a Superformance Fractogel EMD-DMAE 150-10 column (Merck) pre-equilibrated with 10 m*M* potassium phosphate buffer pH 6.5. The column was washed with the same buffer to elute unbound proteins. Bound proteins were eluted with an NaCl gradient (0–1 *M*) in the same buffer. 1 ml fractions were collected and tested for α-galactosidase activity. Active fractions were combined and further purified on a Mono-Q HR 5/5 column (Amersham Biosciences) applying the same buffer gradient.

For AgaB, a third purification step was necessary. Fractions displaying activity from the MonoQ separation were pooled and concentrated using Centricon YM-10 (Milllipore) ultrafiltration. The concentrated α-galactosidase was applied onto a Superdex 200 column (Hiload, 16 × 60 mm, Amersham Biosciences) and eluted with 10 m*M* potassium phosphate buffer pH 6.5 at a flow rate of 0.1 ml min^{-1}.

Active fractions were tested further for purity by SDS–PAGE. The fractions containing the purified protein were combined, dialyzed

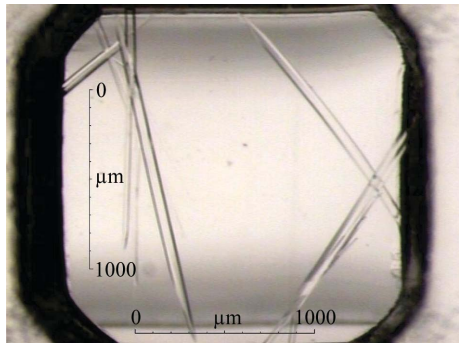


Figure 1
Crystals from α-galactosidase mutant AgaA A355E grown by the sitting-drop vapour-diffusion method in 20%(*w*/*v*) PEG 3350, 0.2 *M* sodium iodide pH 6.9 (condition No. 10 of PEG/Ion Screen, Hampton Research), 10%(*v*/*v*) ethylene glycol, 5%(*v*/*v*) polyvinyl pyrrolidone K15 (additive No. 20 from Additive Screen 2, Hampton Research) solution.

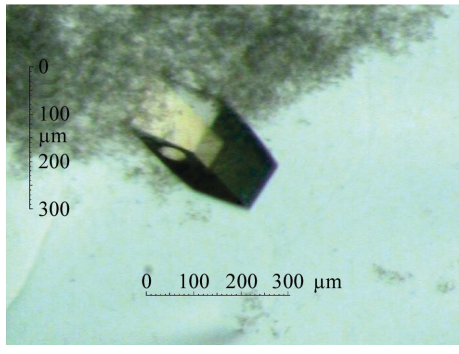


Figure 2
Crystal of the α-galactosidase AgaB grown by the hanging-drop vapour-diffusion method in 1.7 *M* ammonium sulfate, 20 m*M* HEPES pH 6.5 solution.

Table 1
X-ray data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	AgaB	AgaA A355E
Space group	$I222$ or $I2_12_12_1$	$P3_121$ (or $P3_221$)
Unit-cell parameters (Å, °)	a = 87.5, b = 113.3, c = 161.6, $\alpha = \beta = \gamma$ = 90	$a = b$ = 150.1, c = 233.2, $\alpha = \beta$ = 90, γ = 120
Wavelength (Å)	0.98	0.98
Resolution range (Å)	20.0–2.0	20.0–2.8
No. of observed reflections	234655	418365
No. of unique reflections	53028	74975
Multiplicity	4.4 (4.3)	5.6 (5.6)
Average $I/\sigma(I)$	14.9 (8.1)	17.7 (6.8)
Completeness (%)	97.3 (99.2)	99.5 (99.9)
R_{sym}† (%)	7.5 (18.8)	9.0 (25.3)

† $R_{sym} = \sum |I - \langle I \rangle| / \sum I$, where I represents the intensity of the reflection and $\langle I \rangle$ the averaged intensity.

against 10 m*M* HEPES buffer pH 7.0 and concentrated to 10 mg ml^{-1} by ultrafiltration on Centricon devices.

2.4. Crystallization conditions

Crystallization trials took place at 277 and 292 K by the vapour-diffusion method using a home-made ammonium sulfate screen and commercial screens from Hampton Research, Molecular Dimensions and Nextal.

No crystallization conditions were determined for AgaA, but condition No. 10 of Hampton Research PEG/Ion screen [20%(*w*/*v*) polyethylene glycol (PEG) 3350, 0.2 *M* sodium iodide pH 6.9] produced thin plates of AgaA A355E mutant at 292 K. After refinement of crystallization conditions, larger crystals of dimensions 1.2 × 0.1 × 0.1 mm were obtained in 8 × 12-well Greiner crystallization plates using a sitting-drop technique. Drops were made by mixing 1 µl protein solution at 13 mg ml^{-1} with 1 µl reservoir solution [20%(*w*/*v*) PEG 3350, 0.2 *M* sodium iodide pH 6.9, 10%(*w*/*v*) ethylene glycol, 5%(*v*/*v*) polyvinyl pyrrolidone K15; additive No. 20 from Additive Screen 2 of Hampton Research] (Fig. 1). They were equilibrated with 200 µl reservoir solution and incubated at 292 K for 1–2 weeks.

Only one crystal of suitable size for X-ray diffraction studies was obtained for AgaB (0.17 × 0.15 × 0.12 mm). It grew over a year at 292 K (the drop remained clear for ten months) in the hanging drop of a 24-well Limbro plate prepared by mixing 2 µl protein sample at 13 mg ml^{-1} with 2 µl reservoir solution consisting of 1.7 *M* ammonium sulfate, 20 m*M* HEPES pH 6.5 (Fig. 2).

2.5. X-ray diffraction data collection and processing

Crystals of AgaA A355E are very fragile and frequently break when they are touched with a cryoloop. They grow in a solution containing 10%(*v*/*v*) ethylene glycol and can be directly flash-cooled at 100 K in a nitrogen-gas cold stream for data collection. They diffract to 3.5 Å resolution on an in-house rotating-anode X-ray generator coupled with a MAR345 image-plate detector. The best data set was collected to 2.8 Å resolution at the European Synchrotron Radiation Facility (ESRF), Grenoble, France on beamline BM30A. The wavelength was 0.98 Å and 128 images with 0.7° rotation each were recorded on a MAR CCD 165 mm detector (MAR Research).

The data set obtained from the unique AgaB crystal was also collected on BM30A. The crystal was transferred for 1 min into a reservoir solution containing 30%(*v*/*v*) ethylene glycol for cryoprotection. 154 images with 1° rotation each were collected to a resolution of 2 Å at a wavelength of 0.98 Å. All data sets were processed and scaled with the *XDS* program package (Kabsch, 1993) and statistics are shown in Table 1.

3. Results and discussion

Expression of AgaB was carried out using the pBTac1 vector derivative pAMG22 in *E. coli* RM448. Disruption of the cells yielded active enzyme in the supernatant with a specific activity of 28 U mg^{-1}. After purification by anion-exchange chromatography, a sevenfold enrichment of AgaB was achieved. SDS–PAGE analysis displayed a predominant AgaB protein band and a second protein band which had been copurified during the anion-exchange chromatography. The latter protein was separated by gel-permeation chromatography on a Superdex 200 column. The final AgaB had been purified more than tenfold, with a final yield of 25%, and exhibited a specific activity of 300 U mg^{-1} at 298 K in 0.1 *M* potassium phosphate buffer pH 6.5.

The pBTac1 vector derivative pHWG8 was used for production of the AgaA A355E mutant in *E. coli* RM448. From the crude extract, exhibiting a specific activity of 30 U g^{-1}, the AgaA A355E protein was purified tenfold by anion-exchange chromatography, yielding a final product with a specific activity of 300 U mg^{-1}.

The homogeneity of the purified proteins was confirmed by SDS–PAGE. For crystallization experiments, the final products were dialyzed with HEPES buffer pH 7.0. Under these buffer conditions the enzymes exhibit only 50% of their activity under standard conditions.

AgaB and AgaA A355E from *G. stearothermophilus* KVE39 have been successfully crystallized using the vapour-diffusion method. Native data sets of AgaB and AgaA mutant were collected to 2 and 2.8 Å resolution, respectively, using synchrotron radiation. From the diffraction data processing, the crystal of AgaB was found to be orthorhombic and the space group was determined to be $I222$ or $I2_12_12_1$, with unit-cell parameters a = 87.5, b = 113.3, c = 161.6 Å. Assuming the presence of one monomer in the asymmetric unit, the Matthews coefficient (V_M; Matthews, 1968) is 2.5 Å^3 Da^{-1} and the solvent content is 51%. Complete data-collection and processing statistics are given in Table 1. Scaling and merging of the crystallographic data resulted in an overall R_{sym} of 7.5% and an R_{sym} in the highest resolution shell of 18.8%. The merged data were 97.3% complete to 2.0 Å resolution.

Crystals of AgaA A355E belong to the trigonal space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b$ = 150.1 Å, c = 233.2 Å. Scaling and merging of the best measured data set resulted in an overall R_{sym} of 9.0% and an R_{sym} in the highest resolution shell of 25.3%. The merged data were 99.5% complete to 2.8 Å resolution. Assuming the presence of four to six monomers in the asymmetric unit, a Matthews coefficient (V_M) of between 2.4 and 1.6 Å^3 Da^{-1} and a solvent content of between 49 and 33% were calculated for the crystal. These values fall within the range typical for protein crystals (Matthews, 1968).

Sequence identity between AgaA (or AgaB) and the α-galactosidase from *T. maritima* of known structure is low, with an identity of 24% on only 210 residues of the central domain. As a consequence, all attempts to solve the phase problem by molecular replacement, using either the full structure of the α-galactosidase from *T. maritima* or its central domain as a model, have failed [the molecular-replacement programs *AMoRe* (Navaza, 2001), *Phaser* (McCoy *et al.*, 2005) and the web server *CaspR* available at http://igs-server.cnrs-mrs.fr/ were tested].

Thus, the structure of either AgaB or AgaA A355E will need to be determined first using heavy-atom derivatives and/or SeMet derivatives. Soaking of AgaA A355E crystals in the presence of various heavy-atom salts as well as the overexpression of SeMet AgaA A355E and SeMet AgaB are in progress. The determined structure will be used as a model to solve the remaining structure by the molecular-replacement method.

References

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.

Coutinho, P. M. & Henrissat, B. (1999). *J. Mol. Microbiol. Biotechnol.* **1**, 307–308.

Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.

Fridjonsson, O., Watzlawick, H., Gehweiler, A. & Mattes, R. (1999). *FEMS Microbiol. Lett.* **176**, 147–153.

Fridjonsson, O., Watzlawick, H., Gehweiler, A., Rohrhirsch, T. & Mattes, R. (1999). *Appl. Environ. Microbiol.* **65**, 3955–3963.

Fujimoto, Z., Kim, W. D., Kaneko, S., Park, G. G., Momma, M., Kobayashi, H. & Mizuno, H. (2003). *Acta Cryst.* D**59**, 2289–2291.

Ganter, C., Bock, A., Buckel, P. & Mattes, R. (1988). *J. Biotechnol.* **8**, 301–310.

Garman, S. C. & Garboczi, D. N. (2004). *J. Mol. Biol.* **337**, 319–335.

Garman, S. C., Hannick, L., Zhu, A. & Garboczi, D. N. (2002). *Structure*, **10**, 425–434.

Golubev, A. M., Nagem, R. A., Brandao Neto, J. R., Neustroev, K. N., Eneyskaya, E. V., Kulminskaya, A. A., Shabalin, K. A., Savelev, A. N. & Polikarpov, I. (2004). *J. Mol. Biol.* **339**, 413–422.

Henrissat, B. & Coutinho, P. M. (2001). *Methods Enzymol.* **330**, 183–201.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Luonteri, E., Tenkanen, M. & Viikari, L. (1998). *Enzyme Microb. Technol.* **22**, 192–198.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.

Margolles-Clark, E., Tenkanen, M., Luonteri, E. & Penttila, M. (1996). *Eur. J. Biochem.* **240**, 104–111.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.

Shibuya, H., Kobayashi, H., Park, G. G., Komatsu, Y., Sato, T., Kaneko, R., Nagasaki, H., Yoshida, S., Kasamo, K. & Kusakabe, I. (1995). *Biosci. Biotechnol. Biochem.* **59**, 2333–2335.

Shibuya, H., Kobayashi, H., Sato, T., Kim, W. S., Yoshida, S., Kaneko, S., Kasamo, K. & Kusakabe, I. (1997). *Biosci. Biotechnol. Biochem.* **61**, 592–598.

Spangenberg, P., Andre, C., Dion, M., Rabiller, C. & Mattes, R. (2000). *Carbohydr. Res.* **329**, 65–73.

Watzlawick, H., Ganter, C., Stezowski, J. & Mattes, R. (1991). *Biochemical Engineering*, edited by M. Reuss, H. Chmiel, E.-D. Gille & H.-J Knackmuss, pp. 174–176. Stuttgart: Gustav Fisher.