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

Expression, purification, crystallization and preliminary X-ray analysis of rice (*Oryza sativa* L.) Os4BGlu12 β-glucosidase

Sompong Sansenya, James R. Ketudat Cairns and Rodjana Opassiri*

Schools of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

Correspondence e-mail: opassiri@sut.ac.th

Received 3 November 2009
Accepted 8 January 2010

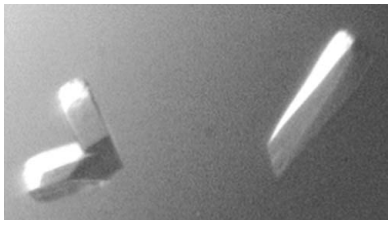
Rice (*Oryza sativa* L.) Os4BGlu12, a glycoside hydrolase family 1 β-glucosidase (EC 3.2.1.21), was expressed as a fusion protein with an N-terminal thioredoxin/His_6 tag in *Escherichia coli* strain Origami B (DE3) and purified with subsequent removal of the N-terminal tag. Native Os4BGlu12 and its complex with 2,4-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucopyranoside (DNP2FG) were crystallized using 19% polyethylene glycol (3350 or 2000, respectively) in 0.1 *M* Tris–HCl pH 8.5, 0.16 *M* NaCl at 288 K. Diffraction data sets for the apo and inhibitor-bound forms were collected to 2.50 and 2.45 Å resolution, respectively. The space group and the unit-cell parameters of the crystal indicated the presence of two molecules per asymmetric unit, with a solvent content of 50%. The structure of Os4BGlu12 was successfully solved in space group $P4_32_12$ by molecular replacement using the white clover cyanogenic β-glucosidase structure (PDB code 1cbg) as a search model.

1. Introduction

β-Glucosidases (EC 3.2.1.21) are found in all kingdoms of life (archaea, bacteria and eukaryotes). These enzymes hydrolyze the β-*O*-glycosidic bond at the anomeric C atom of glucose moieties at the nonreducing end of oligosaccharides or glycoside molecules (Esen, 1993). The functions of β-glucosidases in plants include oligosaccharide hydrolysis in cell-wall remodelling (Leah *et al.*, 1995; Hrmova *et al.*, 1996), the release of toxic aglycones and aromatic volatiles for defence against plant pathogens and herbivores (Poulton, 1990; Duroux *et al.*, 1998; Mizutani *et al.*, 2002), the release of phytohormones from inactive glycosides (Schliemann, 1984; Brzobohatý *et al.*, 1993), the release of monolignols from their glycosides in lignification (Dharmawardhana *et al.*, 1995) and the release of metabolic intermediates by cleaving glucose blocking groups from inactive glucosides (Barleben *et al.*, 1995).

The β-glucosidases that have been studied in plants are members of glycoside hydrolase families (GH) 1, 3 and 5, with GH1 β-glucosidases being the most prevalent (Henrissat, 1991; Opassiri *et al.*, 2007). The tertiary structures of GH1-family enzymes are $(\beta/\alpha)_8$ barrels with the catalytic acid/base and nucleophile carboxylic acid residues at the C-terminal ends of β-strands 4 and 7, respectively (Barrett *et al.*, 1995; Jenkins *et al.*, 1995; Henrissat *et al.*, 1995). The two glutamic acid residues which serve as catalytic amino acids are located in the highly conserved peptide motifs Thr-Phe-Asn-Glu-Pro (TFNEP) and Ile-Thr-Glu-Asn-Gly (ITENG), respectively. These amino-acid residues are located close to the bottom of the active site (Czjzek *et al.*, 2000; Henrissat *et al.*, 1995). The fundamental substrate specificity of these enzymes depends on the overall dimensions and geometry of the binding site and the distribution of the amino-acid residues involved in substrate recognition and binding (Czjzek *et al.*, 2000; Verdoucq *et al.*, 2004).

Recently, the *Os4bglu12* cDNA encoding a GH1 β-glucosidase was cloned from rice seedlings and recombinant thioredoxin/His_6-tagged Os4BGlu12 fusion protein was functionally expressed in *Escherichia coli* (Opassiri *et al.*, 2006). The deduced Os4BGlu12 N-terminal amino-acid sequence was identical to the N-terminal amino-acid



sequence of the previously purified cell-wall-bound rice β-glucosidase for 40 of 44 residues (Akiyama *et al.*, 1998).

The substrate preference of Os4BGlu12 shows both similarities and differences to that of Os3BGlu7, the structure of which is known (PDB codes 2rgl and 2rgm; Chuenchor *et al.*, 2008), and the two share 56% amino-acid sequence identity. Both enzymes efficiently hydrolyze β-(1,4)-linked oligosaccharides of 3–6 glucose residues, but differences were observed in their specificities for substrate chain lengths. The rates of hydrolysis of β-(1,4)-linked oligosaccharides by Os3BGlu7 increase with increasing chain length of the substrate, while those of Os4BGlu12 remain approximately constant (Opassiri *et al.*, 2004, 2006). Among the β-1,3-linked laminarioligosaccharides, Os4BGlu12 only hydrolyzes laminaribiose, while Os3BGlu7 also hydrolyzes laminaritriose, although at a lower rate than laminaribiose. Among the artificial *p*-nitrophenyl (*p*NP)-glycosides, both enzymes hydrolyze *p*NP β-D-glucoside (*p*NPGlc) and *p*NP β-D-fucoside most rapidly. Os4BGlu12 hydrolyzes *p*NP β-D-galactoside, *p*NP β-D-xyloside and *p*NP α-L-arabinoside at 45, 45 and 26% of the rate of *p*NPGlc, respectively, while Os3BGlu7 hydrolyzes these substrates relatively slowly. Os3BGlu7 also hydrolyzes *p*NP β-D-mannoside and *p*NP β-D-cellobioside, but Os4BGlu12 does not. In order to understand the structural basis for these substrate-specificity differences, we have set out to determine the structure of Os4BGlu12 for comparison with that of Os3BGlu7.

2. Materials and methods

2.1. Protein expression and purification

The cDNA (Genbank accession No. AAAA02014151, rice genome locus Os04g0474800) encoding the mature protein (precursor protein residues 25–510) of rice Os4BGlu12 β-glucosidase was cloned into pET32a(+)/DEST to express a soluble fusion protein in Origami B (DE3) *E. coli* as described previously (Opassiri *et al.*, 2006). This fusion protein contains an N-terminal thioredoxin followed by a His_6 tag, a thrombin cleavage site, an S-tag, an enterokinase site and a cloning-site linker of 22 amino-acid residues before the 486 amino-acid residues of mature Os4BGlu12. The recombinant clones were grown in LB medium containing 100 µg ml^{-1} ampicillin, 15 µg ml^{-1} kanamycin, 12.5 µg ml^{-1} tetracycline and were cultured at 310 K in an incubator shaker until the optical density at 600 nm reached 0.5–0.6. IPTG was then added to a final concentration of 0.4 m*M* to induce expression of the target proteins. After induction of protein expression, the cultures were incubated at 293 K for an additional 16 h before harvesting the cells. The culture broth was centrifuged at 277 K for 15 min at 823*g*. The cell pellet was resuspended in 5 ml freshly prepared extraction buffer [50 m*M* Tris–HCl pH 7.2, 200 µg ml^{-1} lysozyme, 1% Triton X-100, 1 m*M* phenylmethylsulfonylfluoride (PMSF) and 0.25 µg ml^{-1} DNase I] per gram and incubated at 298 K for 30 min. The supernatant containing soluble protein was separated from the cell debris by centrifugation at 8064*g* and 277 K for 10 min. The recombinant Os4BGlu12 fusion protein was purified by immobilized metal-ion affinity chromatography on cobalt (Co^{2+}-IMAC) resin, which was prepared by charging Sepharose 6 Fast Flow IMAC resin (GE Healthcare, Buckinghamshire, England) with 0.1 *M* Co^{2+} solution. The protein was loaded onto Co^{2+}-IMAC (1 ml cobalt resin per milligram of protein) pre-equilibrated with equilibration buffer (20 m*M* Tris–HCl pH 7.2 and 300 m*M* NaCl). The unbound protein was washed out with equilibration buffer containing 5 and 10 m*M* imidazole and the protein that was bound to the column was eluted by applying equilibration buffer containing 250 m*M* imidazole. All fractions from the purification steps were assayed for β-glucosidase activity with 1 m*M* *p*NP β-D-glucoside as described by Opassiri *et al.* (2003). The pure protein fractions were then pooled and concentrated and the buffer was changed to 20 m*M* Tris–HCl pH 7.2 in 30 kDa molecular-weight cutoff Centricon centrifugal filters (Millipore, Bedford, Massachusetts, USA) by centrifugation at 605*g* and 277 K. To remove the N-terminal thioredoxin/His_6 tag, the Os4BGlu12 fusion protein was cleaved with 0.1 µg enterokinase (New England Biolabs) per 100 µg of protein in 20 m*M* Tris–HCl pH 7.2 at 296 K for 16 h. The released fusion tag was then removed by adsorption to Co^{2+}-IMAC, as described for the fusion protein, and the flowthrough and 5 and 10 m*M* wash fractions containing unbound and weakly bound tag-free protein were collected. The fractions containing tag-free protein were pooled and concentrated in a 10 kDa molecular-weight cutoff Centricon centrifugal filter by centrifugation at 420*g* and 277 K. To assess the protein purity, all purification steps were analyzed by SDS–PAGE with a 12% separating gel using the method of Laemmli (1970). The concentration of protein in each fraction was assayed with the Bio-Rad protein assay kit (Bio-Rad, Richmond, California, USA) using bovine serum albumin as a standard.

2.2. Protein crystallization

The purified tag-free protein was concentrated to 3.5 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.2 in a 10 kDa molecular-weight Centricon filter at 420*g* and 277 K and the protein solution was filtered through an Ultrafree-MC 0.22 µm filter (Millipore) at 358*g* for 5 min. Conditions for the crystallization of native Os4BGlu12 were screened by the microbatch method with a systematic screen of monovalent and divalent salts (0.2 *M*) and pH (in the range 4.0–8.5 at 0.1 *M* buffer concentration) in the presence of 25%(*w*/*v*) polyethylene glycol (PEG) 4000. A promising condition [25%(*w*/*v*) PEG 4000 in 0.1 *M* Tris–HCl pH 8.5, 0.2 *M* NaCl] was optimized by hanging-drop vapour diffusion by varying the molecular weight of PEG and the concentrations of NaCl and protein. Hanging-drop vapour diffusion was performed at 288 K in drops consisting of 2 µl protein solution and 1 µl precipitant solution equilibrated against a reservoir containing 0.5 ml precipitant solution. Microseeding was performed to produce larger crystals with pre-incubation of the drop and reservoir for 1 h before streaking in the crystal seeds. The Os4BGlu12 protein solution was also mixed with 10 m*M* 2,4-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucopyranoside (DNP2FG) solution and the mixture was incubated at 277 K for 10 min. Crystallization of the complex was then performed as described for the native crystals.

2.3. Data collection and processing

X-ray data from Os4BGlu12 crystals were collected on beamline BL13B1 at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The crystals were transferred to cryoprotectant solution [composed of 23% PEG, 0.12 *M* Tris–HCl pH 8.5, 0.19 *M* NaCl and 28%(*v*/*v*) glycerol], soaked for 15 s and flash-vitrified in liquid nitrogen. During X-ray diffraction, the crystal-to-detector distance was 300 mm for the native Os4BGlu12 crystal and 280 mm for the Os4BGlu12–DNP2FG complex crystal. The wavelength of the X-ray beam was 1.00 Å for the collection of both data sets. The X-ray data sets were recorded on an ADSC Quantum 315 CCD detector. The crystals were maintained at 100 K in the nitrogen stream from an Oxford Cryosystems Cryostream during the entire X-ray data collection. The exposure time per frame was 20 and 15 s for crystals with and without the DNP2FG inhibitor, respectively. The diffraction images were indexed, integrated and scaled with *HKL*-2000 (Otwinowski & Minor, 1997). The data were analyzed with the *CCP*4 suite

of programs (Collaborative Computational Project, Number 4, 1994) and the initial native structure was solved by molecular replacement using the cyanogenic β-glucosidase from *Trifolium repens* (PDB code 1cbg; Barrett *et al.*, 1995) as the search model in the *MOLREP* program (Vagin & Teplyakov, 1997). The data set for Os4BGlu12 complexed with DNP2FG was solved by rigid-body refinement of the native Os4BGlu12 model against the data for the complex crystal with *REFMAC*5 (Murshudov *et al.*, 1997; Collaborative Computational Project, Number 4, 1994). 5% of the reflections were reserved from the refinement for calculation of the $R_{\rm free}$ values of the two data sets, with the reflections in the two $R_{\rm free}$ sets out to 2.50 Å resolution being the same.

3. Results and discussion

The thioredoxin/His_6-tagged Os4BGlu12 fusion protein was expressed in the soluble fraction in *E. coli* strain Origami B (DE3) cells. After purification, the Os4BGlu12 protein with the N-terminal tag removed was approximately 95% pure (Fig. 1) and had a yield of approximately 6 mg purified protein per litre of LB medium. Small crystals were observed using a precipitant consisting of 25%(*w*/*v*) PEG 4000 in 0.1 *M* Tris–HCl pH 8.5, 0.2 *M* NaCl within 17 d. Optimization of this condition was carried out by hanging-drop vapour diffusion with microseeding and native Os4BGlu12 crystals with dimensions of 0.120 × 0.025 × 0.020 mm were obtained using 19%(*w*/*v*) PEG 3350 in 0.1 *M* Tris–HCl pH 8.5, 0.16 *M* NaCl, while crystals of Os4BGlu12 with DNP2FG inhibitor with dimensions of up to 0.125 × 0.025 × 0.025 mm were obtained in 19%(*w*/*v*) PEG 2000 in 0.1 *M* Tris–HCl pH 8.5, 0.16 *M* NaCl (Fig. 2).

The crystals of native Os4BGlu12 and of Os4BGlu12 complexed with DNP2FG diffracted X-rays to 2.50 and 2.45 Å resolution, respectively. The statistics for the two X-ray data sets are summarized in Table 1. The diffraction patterns showed that both the native and the DNP2FG-complex crystals had similar unit-cell parameters of $a = b = 112.7$, $c = 182.8$ Å ($\alpha = \beta = \gamma = 90^\circ$) and $a = b = 114.1$, $c = 184.5$ Å, respectively, and had tetrahedral symmetry. The systematic absences were consistent with either space group $P4_12_12$ or $P4_32_12$. The $V_{\rm M}$ was found to be 2.46 Å^3 Da^{-1} with 49.98% solvent

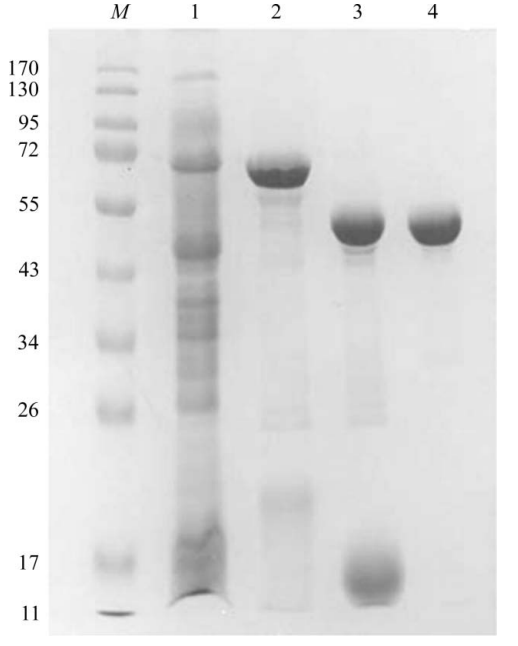


Figure 1
12% SDS–PAGE of Os4BGlu12 protein from the purification steps. Lane *M*, protein markers (kDa); lane 1, soluble protein extract of *E. coli* cells after Os4BGlu12 expression; lane 2, the N-terminal thioredoxin/His_6-tagged Os4BGlu12 fusion protein after the first Co^{2+}-IMAC column; lane 3, enterokinase digest of the fusion protein; lane 4, pure Os4BGlu12 after removal of the cleaved thioredoxin/His_6 tag by the second Co^{2+}-IMAC column.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	Os4BGlu12	Os4BGlu12–DNP2FG
Data collection		
Beamline	BL13B1	BL13B1
Detector	ADSC Quantum 315 CCD	ADSC Quantum 315 CCD
Crystal-to-detector distance (nm)	300	280
Wavelength (Å)	1.00	1.00
Exposure time (s)	20	15
Space group	$P4_32_12$	$P4_32_12$
Unit-cell parameters (Å)	$a = 112.6$, $b = 112.6$, $c = 182.7$	$a = 114.1$, $b = 114.1$, $c = 184.5$
Resolution range (Å)	30–2.50 (2.50–2.49)	30–2.45 (2.51–2.45)
No. of unique reflections	39533	43131
No. of observed reflections	408534	381316
Completeness (%)	100.0 (99.9)	100.0 (99.9)
Average redundancy per shell	9.7 (10.0)	8.4 (8.6)
$\langle I/\sigma(I)\rangle$	19.6 (6.4)	21.7 (4.4)
$R_{\rm merge}$† (%)	10.6 (41.7)	9.4 (49.9)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is the mean value of $I_i(hkl)$ for all i measurements.

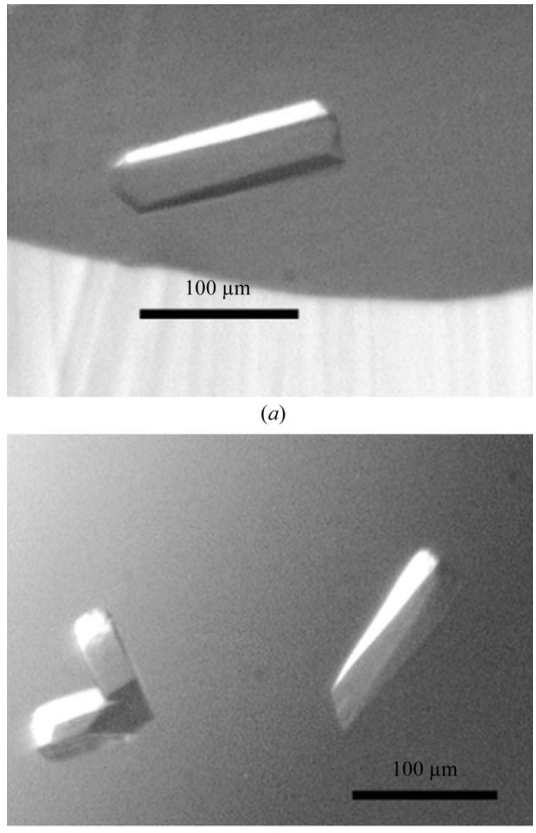


Figure 2
Single crystals of (*a*) native Os4BGlu12 and (*b*) Os4BGlu12 complexed with DNP2FG obtained after optimization, with maximum dimensions of 120 × 25 × 20 and 125 × 25 × 25 µm, respectively.

content for a solution with two molecules in the asymmetric unit. The structure of Os4BGlu12 was solved by molecular replacement using the *MOLREP* program (Vagin & Teplyakov, 1997) with the cyanogenic β-glucosidase from white clover (PDB code 1cbg; Barrett *et al.*, 1995) structure as the search model to give an initial $R_{\rm cryst}$ of 0.450, an $R_{\rm free}$ of 0.466 and a figure of merit (FOM) of 0.502. The space group of the two X-ray data sets was determined to be $P4_32_12$. The isomorphous crystals of Os4BGlu12 complexed with DNP2FG also contained two molecules per asymmetric unit, with a corresponding solvent content of 54.17% and a $V_{\rm M}$ of 2.68 Å^3 Da^{-1}. Rigid-body refinement of the native Os4BGlu12 structure with the Os4BGlu12–DNP2FG complex data yielded a solution with an initial $R_{\rm cryst}$ of 0.388, an $R_{\rm free}$ of 0.390 and an FOM of 0.794.

Currently, work is under way to refine the structures of native Os4BGlu12 and of the enzyme in complex with DNP2FG, together with optimization of the crystallization conditions to obtain higher resolution data. The structure will allow us to determine the structural differences that lead to the differences in oligosaccharide hydrolysis between Os4BGlu12 and other rice enzymes such as Os3BGlu6 (Seshadri *et al.*, 2009) and Os3BGlu7 (Chuenchor *et al.*, 2008).

We would like to thank Chun-Jung Chen, Phimonphan Chuankhayan, Buabarn Kuaprasert and the protein crystallography beamline staff at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The systematic screen for microbatch was provided by Jirandorn Yuvaniyama, Mahidol University, Bangkok, Thailand. This work was supported by grants from the Synchrotron Light Research Institute (SLRI), Thailand and the Thailand Research Fund. Data collection was carried out at the NSRRC, a national user facility supported by the National Science Council of Taiwan, Republic of China. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine of Taiwan.

References

Akiyama, T., Kaku, H. & Shibuya, N. (1998). *Phytochemistry*, **48**, 49–54.
Barleben, L., Ma, X., Koepke, J., Peng, G., Michel, H. & Stöckigt, J. (1995). *Biochim. Biophys. Acta*, **1747**, 89–92.
Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. & Hughes, M. A. (1995). *Structure*, **5**, 951–960.
Brzobohatý, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J. & Palme, K. (1993). *Science*, **262**, 1051–1054.
Chuenchor, W., Pengthaisong, S., Robinson, R. C., Yuvaniyama, J., Oonanant, W., Bevan, D. R., Esen, A., Chen, C. J., Opassiri, R., Svasti, J. & Ketudat Cairns, J. R. (2008). *J. Mol. Biol.* **377**, 1200–1215.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B. & Esen, A. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 13555–13560.
Dharmawardhana, D. P., Ellis, B. E. & Carlson, J. E. (1995). *Plant Physiol.* **107**, 331–339.
Duroux, L., Delmotte, F. M., Lancelin, J. M., Keravis, G. & Jay-Alleand, C. (1998). *Biochem. J.* **333**, 275–283.
Esen, A. (1993). *β-Glucosidases: Biochemistry and Molecular Biology.* Washington DC: American Chemical Society.
Henrissat, B. (1991). *Biochem. J.* **280**, 309–316.
Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P. & Davies, G. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 7090–7094.
Hrmova, M., Harvey, A. J., Wang, J., Shirley, N. J., Jones, G. P., Stone, B. A., Hoj, P. B. & Fincher, G. B. (1996). *J. Biol. Chem.* **271**, 5277–5286.
Jenkins, J., Lo Leggio, L. L., Harris, G. & Pickersgill, R. (1995). *FEBS Lett.* **362**, 281–285.
Laemmli, U. K. (1970). *Nature (London)*, **277**, 680–685.
Leah, R., Kigel, J., Svedsen, I. & Munndy, J. (1995). *J. Biol. Chem.* **270**, 15789–15797.
Mizutani, M., Nakanishi, H., Ema, J., Ma, S., Nogchi, E., Inohara-Ochiai, M., Fukuchi-Mizutani, M., Nakao, M. & Sakata, K. (2002). *J. Plant Physiol.* **130**, 2164–2176.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A. & Ketudat Cairns, J. R. (2004). *Biochem. J.* **379**, 125–131.
Opassiri, R., Ketudat Cairns, J. R., Akiyama, T., Wara-Aswapati, O., Svasti, J. & Esen, A. (2003). *Plant Sci.* **165**, 627–638.
Opassiri, R., Pomthong, B., Akiyama, T., Nakphaichit, M., Onkoksoong, T., Ketudat Cairns, M. & Ketudat Cairns, J. R. (2007). *Biochem. J.* **408**, 241–249.
Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A. & Ketudat Cairns, J. R. (2006). *BMC Plant Biol.* **6**, 33.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Poulton, J. E. (1990). *Plant Physiol.* **94**, 401–405.
Schliemann, W. (1984). *J. Plant Physiol.* **116**, 123–132.
Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B. & Ketudat Cairns, J. (2009). *Plant Physiol.* **151**, 47–58.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Verdoucq, L., Morinière, J., Bevan, D. R., Esen, A., Vasella, A., Henrissat, B. & Czjzek, M. (2004). *J. Biol. Chem.* **279**, 31796–31803.