Danilo Elton Evangelista,^a Andre Schutzer de Godoy,^a Fernando Fonseca Pereira de Paula,^b Flavio Henrique-Silva^{b*} and Igor Polikarpov^{a*}

^aInstituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-carlense 400, 13566-590 São Carlos-SP, Brazil, and
^bDepartamento de Genética e Evolução, Universidade Federal de São Carlos, Rodovia Washington Luís, Km 235, 13565-905 São Carlos-SP, Brazil

Correspondence e-mail: dfhs@power.ufscar.br, ipolikarpov@ifsc.usp.br

Received 8 September 2013

Accepted 22 January 2014

Expression, purification, crystallization and preliminary X-ray diffraction analysis of the pectin methylesterase from the sugar cane weevil *Sphenophorus levis*

Pectin methylesterase removes the methyl groups from the main chain of pectin, the major component of the middle lamella of the plant cell wall. The enzyme is involved in plant cell-wall development, is part of the enzymatic arsenal used by microorganisms to attack plants and also has a wide range of applications in the industrial sector. Therefore, there is a considerable interest in studies of the structure and function of this enzyme. In this work, the pectin methylesterase from *Sphenophorus levis* was produced in *Pichia pastoris* and purified. Crystals belonging to the monoclinic space group C2, with unit-cell parameters $a = 122.181$, $b = 82.213$, $c = 41.176$ Å, $\beta = 97.48^\circ$, were obtained by the sitting-drop vapour-diffusion method and an X-ray diffraction data set was collected to 2.1 Å resolution. Structure refinement and model building are in progress.

1. Introduction

Pectic substances are the major component of the middle lamella of the plant cell wall, contributing to the firmness, structural support and protection of plant tissues. Pectin is mainly composed of galacturonic acid units linked by α -1,4-glycosidic bonds, with the carboxylate groups partially converted to methyl esters (Gummadi & Panda, 2003). This polysaccharide can be degraded by enzymes from the pectinase family, including pectin methylesterase (PME; EC 3.1.1.11), which catalyses the demethylation of the main chain of pectin (Gummadi & Panda, 2003).

There are many biological processes in which PMEs are involved, *e.g.* development of the plant cell wall, fruit ripening, abscission, senescence (Jayani *et al.*, 2005) and regulation of pollen-tube growth (Bosch *et al.*, 2005). PMEs are also part of the arsenal of enzymes used by several phytopathogenic microorganisms to attack plants; consequently, PMEs contribute to the virulence of these pathogens (Rogers *et al.*, 2000). Pectinases also have considerable importance for industrial applications, such as plant-biomass hydrolysis, enzymatic methods of protoplast isolation, oil extraction, coffee and tea fermentation, and retting and degumming of fibres in the textile industry (Kashyap *et al.*, 2001). PMEs are widely applied for the extraction and clarification of fruit juices as well as the protection and improvement of several processed fruits and vegetables (Jayani *et al.*, 2005). For these reasons, there is a constant search for new sources of these enzymes.

Pectinases, as well as other plant cell-wall degrading enzymes (PCWDEs), from plants, bacteria, archaea and fungi have been extensively studied. However, studies involving genomic analysis, transcriptomes, molecular cloning, recombinant protein expression and protein sequencing revealed that some phytophagous insects can also endogenously produce PCWDEs (Watanabe *et al.*, 1998; Girard & Jouanin, 1999; Watanabe & Tokuda, 2001, 2010; Allen & Mertens, 2008; Celorio-Mancera *et al.*, 2009; Willis *et al.*, 2011; Calderón-Cortés *et al.*, 2012; Kirsch *et al.*, 2012).

Sphenophorus levis, popularly known as the sugarcane weevil, is a pest of sugarcane cultivars, which are among the most important agricultural plants in Brazil. The adult beetle deposits its eggs in the sugarcane base and, after hatching, the larvae feed themselves on the stem and rhizome, building galleries and causing plant death (Cerdeira *et al.*, 1999). In spite of their economic importance there are no effective methods for agricultural control of these coleopterans.

© 2014 International Union of Crystallography
All rights reserved

Aiming to identify new molecular targets for the biotechnological control of this insect, the Laboratory of Molecular Biology from the Federal University of São Carlos, São Paulo, Brazil, built a cDNA library of *S. levis* larvae (unpublished data). Bioinformatics analysis of the library allowed us to identify a PME (NCBI accession No. KF697077) that we named SL-PME.

PMEs are rarely present in the animal kingdom and this is one of the reasons why they have been almost exclusively studied from plant, bacteria, archaea and fungi. Furthermore, despite a large number of identified genes, there are only six different PMEs with known three-dimensional structures deposited in the Protein Data Bank (PDB). These PMEs belong to *Erwinia chrysanthemi* (Jenkins *et al.*, 2001), *Daucus carota* (Johansson *et al.*, 2002), *Arabidopsis thaliana* (Hothorn *et al.*, 2004), *Solanum lycopersicum* (Di Matteo *et al.*, 2005), *Dickeya dadantii* (Fries *et al.*, 2007) and *Yersinia enterocolitica* (Boraston & Abbott, 2012), all with 24–35% sequence identity to SL-PME. Among all classes of pectinases deposited in the PDB, there is not a single enzyme from an animal source. Since the level of amino-acid sequence homology between publicly available PME structures and SL-PME is quite low, and therefore obtaining precise structural information is not possible using homology modelling, we set out to determine the crystallographic structure. Here, we present the first report of X-ray diffraction studies of a recombinant pectinase of animal origin.

2. Materials and methods

2.1. Construction of the expression plasmid

The SL-PME ORF (open reading frame) is composed of 1158 bp encoding a 386-amino-acid polypeptide, which includes a putative 16-residue signal sequence. The ORF, excluding the signal peptide sequence, was obtained by PCR using the following primers, which include sites for the specified restriction enzymes (shown in bold): PME_Forward 5'-AA**ACTG**CAGGAACCAAAATCCACCAGG-3' (*Pst*I) and PME_Reverse 5'-AAAAG**CGG**CCGCTGCTCCACTTCCGGTATTGGC-3' (*Not*I). The restriction-enzyme sites were selected for directional cloning into the pPICZαB expression vector between the α-factor secretion signals and the C-terminal c-Myc epitope and His tag. After export, the 89 residues from the α-factor signal peptide are expected to be cleaved and the secreted protein will thus begin with residue 17 of the SL-PME at the N-terminus and contain residues 17–386 of SL-PME followed by 15 residues from the c-Myc epitope and six residues of the His tag at the C-terminus. Expression of the inserted sequence is driven by a promoter from the alcohol oxidase gene AOX1, which is induced by methanol. After separation of the amplified product in a 1% agarose gel, the amplicon was collected and purified using a PCR purification kit (Promega). Amplified DNA and vector were purified and digested with selected restriction enzymes for subsequent ligation to the corresponding sites. The constructed recombinant plasmid was transformed in chemically competent *Escherichia coli* cells (DH5α) for multiplication and then analysed by sequencing.

2.2. Expression and purification

In silico analysis of the predicted amino-acid sequence of SL-PME revealed one putative N-glycosylation site as well as one putative O-glycosylation site. In order to complete these post-translational modifications, the recombinant enzyme was expressed in a eukaryotic system utilizing the EasySelect *Pichia* Expression Kit (Invitrogen). 10 µg of the recombinant plasmid, linearized by the *Pme*I restriction enzyme (Fermentas), was used to transform competent *P. pastoris*

Figure 1

Expression and purification of recombinant SL-PME. (a) SDS-PAGE analysis of the recombinant SL-PME expression during 144 h in BMMY supplemented with 0.75% methanol. Lane M, Protein Marker, Broad Range (New England Biolabs); the other lanes are labelled with time of methanol-induced protein expression. (b) SDS-PAGE analysis of the recombinant SL-PME purified by Ni²⁺ affinity chromatography. Lane M, Protein Marker, Broad Range; the other lanes are labelled with the imidazole concentration at which the purified enzyme was collected. A refers to the first 5 ml fraction collected and B refers to the second and last 5 ml fraction collected. In both figures, the arrow indicates recombinant SL-PME.

(KM71H) by electroporation (1.5 kV, 25 µF, 200 Ω) in a 2 mm cuvette. The transformed yeast cells were cultivated in solid yeast extract peptone dextrose with 1 M sorbitol (YEPDS) and 500 µg ml⁻¹ zeocin. Positive recombinant yeasts were identified by PCR using the vector primers.

For protein expression, the recombinant cells were grown in 500 ml of BMGY (buffered complex medium containing glycerol) for 24 h at 303 K and 180 rev min⁻¹ (OD_{600 nm} = 6). The cells were harvested by centrifugation (5000g, 10 min) and resuspended in 100 ml of BMMY (buffered complex medium containing methanol) and incubated for 144 h at 301 K under constant agitation at 180 rev min⁻¹. Methanol was added every 24 h to a final concentration of 0.75%. The recombinant enzyme was purified directly from the culture supernatant using Ni-NTA superflow affinity resin columns (Qiagen) and eluted in buffer (10 mM Tris, 100 mM NaCl, 50 mM NaH₂PO₄ pH 8.0) with increasing imidazole concentration (10, 25, 50, 75, 100 and 250 mM). The fractions were analysed by 12% SDS-PAGE (Sambrook *et al.*, 1989) to identify fractions containing pure SL-PME.

2.3. Enzymatic assay of recombinant SL-PME activity

To verify the activity of the recombinant SL-PME, a gel matrix [100 mM sodium acetate buffer pH 5.0, 1% agar and 1% polygalacturonic acid 75% esterified from apple (Sigma)] was prepared. Samples of culture medium were deposited on the gel surface. The plate was incubated at 320 K for 4 h, stained with 0.2% (w/v) Congo Red (Vetec) for 20 min and washed with 1 M NaCl solution.

2.4. Crystallization

Imidazole was diluted to 7 mM in the purified SL-PME by applying 50 ml washing buffer (20 mM HEPES pH 8.0, 20 mM NaCl) and concentrating with a 10 000 MWCO centrifugal filter concentrator (Millipore) centrifuged at 4000g to a final concentration of 15.87 mg ml⁻¹. The protein concentration was determined by measuring UV absorbance at 280 nm with the use of a NanoDrop 2000 (Thermo Scientific) and applying the theoretical extinction coefficient calculated based on the protein sequence. For crystallization screening, the sitting-drop vapour-diffusion method was used, in which 1 µl protein solution was mixed with 1 µl reservoir solution at 292 K. Initial crystallization screens were carried out using the automated robotic system Honey Bee 961 Dispensing System (Digilab) and commercial crystallization kits (Index and SaltRx from Hampton Research and The PEGs and PEGs II Suite from Qiagen).

2.5. X-ray diffraction data collection

Crystals were mounted in loops from Hampton Research, using the reservoir solution complemented with 10% (v/v) ethylene glycol as a cryoprotectant. The crystals were then flash-cooled to 100 K in a cold nitrogen stream. X-ray diffraction data were collected on a Bruker APEX DUO single-crystal diffractometer system equipped with a kappa goniometer and APEX II CCD detector. The exposure time was 20 s per frame for data collection with a -20° detector offset angle. The point-group determination and data-collection strategy were performed with the *PROTEUM2* software. Diffraction data were integrated to 2.1 Å resolution with *SAINT* and scaled with *SADABS* (Bruker AXS, Madison, Wisconsin, USA). Statistics of the data collection and space-group determination were obtained with *XPREF* (Bruker AXS). The solvent content and Matthews coefficient were calculated using *phenix.xtriage* (Adams *et al.*, 2010). The

protein structure was solved with *Phaser* (McCoy *et al.*, 2007) by using an ensemble of models generated with *phenix.ensemble* and *phenix.sculptor* (Adams *et al.*, 2010).

3. Results and discussion

Recombinant SL-PME was successfully produced and purified with a yield of 830 mg l⁻¹ (Fig. 1). The enzyme was active as shown by the presence of a halo of modified substrate in the agar-plate assay (Fig. 2). The enzyme was submitted to extensive crystallization trials. Crystals appeared after 10 d in the original sitting-drop screen plates using the Index kit with 25% (w/v) PEG 3350, 0.1 M citric acid pH 3.5 as a crystallization solution and a 1:1 (v:v) protein-to-crystallization

Figure 3
Crystals of recombinant SL-PME. SL-PME crystals grown after 10 d in 25% (w/v) PEG 3350, 0.1 M citric acid pH 3.5. The ratio of crystallization solution: protein was 1:1 (v:v).

Figure 2
Enzymatic activity of recombinant SL-PME. The plate contained 1% 70–75% esterified pectin (Sigma), 1% agar and 100 mM sodium acetate buffer pH 5.0. A, Spot containing 20 µl of the cell-free culture medium after 144 h of induction with methanol. B, Spot containing 20 µl of the cell-free culture medium at 0 h of induction with methanol.

Figure 4
Diffraction pattern of recombinant SL-PME crystal. The 2.12 Å resolution ring which is close to the maximum resolution of the diffraction data set is shown.

Table 1

X-ray data-collection statistics for recombinant SL-PME.

Values in parentheses are for the highest resolution shell.

Detector	APEX II CCD
Wavelength (Å)	1.54166
Rotation range per frame (°)	0.5
Crystal-to-detector distance (mm)	50.0
Data-collection temperature (K)	100
Resolution range (Å)	36.1–2.1 (2.19–2.10)
Space group	C121
Unit-cell parameters (Å, °)	$a = 122.181$, $b = 82.213$, $c = 41.176$, $\beta = 97.48$
Unique reflections	22383 (1958)
Multiplicity	8.27 (4.78)
Completeness (%)	96.62 (85.43)
Mean $I/\sigma(I)$	8.93 (1.77)
Wilson B factor	14.85
R_{merge}	0.1578 (0.5264)
Solvent content (%)	54.7

solution ratio (Fig. 3). A blade-shaped crystal with dimensions $1.0 \times 0.2 \times 0.1$ mm was of X-ray diffraction quality (Fig. 4) and the resolution of the diffraction data extended to 2.1 Å.

The data set was indexed assuming a monoclinic C -centred lattice. After integration, reflections were scaled and merged into 22 383 total reflections, with a total R_{merge} of 15.8% and completeness of 96.6%. Based on the scaling, the space group was identified as $C2_1$, with unit-cell parameters $a = 122.181$, $b = 82.213$, $c = 41.176$ Å, $\beta = 97.48^\circ$ (Table 1). Assuming one enzyme molecule with a molecular mass of 50 kDa in the asymmetric unit, the Matthews coefficient was calculated as $2.71 \text{ Å}^3 \text{ Da}^{-1}$. This corresponds to 54.7% solvent content for the crystal (Matthews, 1968). The structure was solved with *Phaser* (McCoy *et al.*, 2007), using an ensemble of models selected by *HHpred* (Söding *et al.*, 2005) and based on the four PDB entries which had the highest scores in the hidden Markov model homology search: 3grh, 1xg2, 1gq8 and 3uw0 (Eklöf *et al.*, 2009; Di Matteo *et al.*, 2005; Johansson *et al.*, 2002; Boraston & Abbott, 2012). The PDB entries have 35, 25, 24 and 31% sequence identity, respectively, to the SL-PME amino-acid sequence. Molecular replacement gave a unique solution in the identified space group (rotation-function Z -score = 9.4, translation-function Z -score = 19.9 and log-likelihood gain = 180). Structure refinement and model building are in progress.

This work was supported by FAPESP and CNPq.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Allen, M. L. & Mertens, J. A. (2008). *J. Insect Sci.* **8**, 27.
- Boraston, A. B. & Abbott, D. W. (2012). *Acta Cryst.* **F68**, 129–133.
- Bosch, M., Cheung, A. Y. & Hepler, P. K. (2005). *Plant Physiol.* **138**, 1334–1346.
- Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H. & Oyama, K. (2012). *Annu. Rev. Ecol. Evol. Syst.* **43**, 45–71.
- Celorio-Mancera, M. de la P., Greve, L. C., Teuber, L. R. & Labavitch, J. M. (2009). *Arch. Insect Biochem. Physiol.* **70**, 122–135.
- Cerda, H., Fernandez, G., Lopez, A. & Varga, J. (1999). *Fla. Entomol.* **82**, 103–112.
- Di Matteo, A., Giovane, A., Raiola, A., Camardella, L., Bonivento, D., De Lorenzo, G., Cervone, F., Bellincampi, D. & Tsernoglou, D. (2005). *Plant Cell*, **17**, 849–858.
- Eklöf, J. M., Tan, T.-C., Divne, C. & Brumer, H. (2009). *Proteins*, **76**, 1029–1036.
- Fries, M., Ihrig, J., Brocklehurst, K., Shevchik, V. E. & Pickersgill, R. W. (2007). *EMBO J.* **26**, 3879–3887.
- Girard, C. & Jouanin, L. (1999). *Insect Biochem. Mol. Biol.* **29**, 1129–1142.
- Gummadi, S. N. & Panda, T. (2003). *Process Biochem.* **38**, 987–996.
- Hothorn, M., Wolf, S., Aloy, P., Greiner, S. & Scheffzek, K. (2004). *Plant Cell*, **16**, 3437–3447.
- Jayani, R. S., Saxena, S. & Gupta, R. (2005). *Process Biochem.* **40**, 2931–2944.
- Jenkins, J., Mayans, O., Smith, D., Worboys, K. & Pickersgill, R. W. (2001). *J. Mol. Biol.* **305**, 951–960.
- Johansson, K., El-Ahmad, M., Friemann, R., Jörnvall, H., Markovic, O. & Eklund, H. (2002). *FEBS Lett.* **514**, 243–249.
- Kashyap, D. R., Vohra, P. K., Chopra, S. & Tewari, R. (2001). *Bioresour. Technol.* **77**, 215–227.
- Kirsch, R., Wielsch, N., Vogel, H., Svatoš, A., Heckel, D. G. & Pauchet, Y. (2012). *BMC Genomics*, **13**, 587.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Rogers, L. M., Kim, Y.-K., Guo, W., González-Candelas, L., Li, D. & Kolattukudy, P. E. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 9813–9818.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press.
- Söding, J., Biegert, A. & Lupas, A. N. (2005). *Nucleic Acids Res.* **33**, W244–W248.
- Watanabe, H., Noda, H., Tokuda, G. & Lo, N. (1998). *Nature (London)*, **394**, 330–331.
- Watanabe, H. & Tokuda, G. (2001). *Cell. Mol. Life Sci.* **58**, 1167–1178.
- Watanabe, H. & Tokuda, G. (2010). *Annu. Rev. Entomol.* **55**, 609–632.
- Willis, J. D., Oppert, B., Oppert, C., Klingeman, W. E. & Jurat-Fuentes, J. L. (2011). *J. Insect Physiol.* **57**, 300–306.