Crystallization and preliminary crystallographic analysis of recombinant VSP1 from *Arabidopsis thaliana*

Zhu-Bing Shi,[a] Hong-Hua Ge,[a] Ping Zhao[b] and Min Zhang[a]*

[a]School of Life Science, Anhui University, Hefei 230039, People's Republic of China, and [b]School of Life Sciences, University of Science and Technology of China, Hefei 230026, People's Republic of China

Correspondence e-mail: zhmin07@ustc.edu

VSP1 is a defence protein in *Arabidopsis thaliana* that may also be involved in control of plant development. The recombinant protein has been overexpressed in *Escherichia coli*, purified and crystallized using the sitting-drop vapour-diffusion method. The crystal diffracted to 1.9 Å resolution and a complete X-ray data set was collected at 100 K using Cu $K\alpha$ radiation from a rotating-anode X-ray source. The crystals belonged to space group $C2$. As there are no related structures that could be used as a search model for molecular replacement, work is in progress on experimental phasing using heavy-atom derivatives and selenomethionine derivatives.

1. Introduction

Based on their abundance and their patterns of accumulation and degradation, vegetative storage proteins (VSPs) were first described as temporary nitrogen-storage reserves in soybean (VSPα and VSPβ; Staswick, 1990, 1994). Although their amino-acid sequences are related to those of plant acid phosphatases, VSPα and VSPβ have weak enzymatic activity owing to a lack of the nucleophilic Asp residue that has been proposed to have a catalytic function (Leelapon *et al.*, 2004). VSP1 and VSP2 are *Arabidopsis thaliana* proteins that are named after soybean VSPs because of their sequence similarity and homology (their primary structures are approximately 40% identical to those of the soybean VSPs), similar expression profiles and immunogenic cross-reactivity (Berger *et al.*, 1995).

In addition to their postulated role as temporary storage reserves, *A. thaliana* VSP1 and VSP2 are likely to participate in plant development control and defence. It has been shown that a complex of VSP1 and FLOR1 (a leucine-rich repeat protein) interacts directly *in vitro* with AGAMOUS, a transcription factor that is required for the stamen and carpel determination of flowers (Gamboa *et al.*, 2001). The resistance of *A. thaliana* to insect attack and pathogens is related to the accumulation of large amounts of VSPs (Rojo *et al.*, 1999; Stotz *et al.*, 2000; Ellis & Turner, 2001; Berger *et al.*, 2002). In addition, the *vsp1* gene was induced by jasmonate, a plant hormone that is also involved in plant development and defence responses (Guerineau *et al.*, 2003). Furthermore, recombinant *A. thaliana* VSP2 displays direct anti-insect activity when incorporated into their diet (Liu *et al.*, 2005).

Amino-acid sequence analysis suggested that *A. thaliana* VSP1 and VSP2 belong to the bacterial nonspecific acid phosphatases, which are members of the haloacid dehalogenase (HAD) superfamily (Thaller *et al.*, 1998; Selengut, 2001). In the majority of members of the HAD superfamily the first Asp in the conserved DXDXT motif acts as the nucleophile during catalysis (Collet *et al.*, 1998) and the Mg^{2+} cation is also crucial for enzymatic activity (Allen & Dunaway-Mariano, 2004). The known three-dimensional structures of HAD-superfamily proteins indicate that the catalytic scaffold of the core domain is conserved. However, the presence and location of the cap domain that is responsible for active-site desolvation is quite diverse and probably results in the catalytic diversity (substrate specificity and reaction type) of the HAD phosphatases (Burroughs *et al.*, 2006).

VSP1 shares 86% sequence identity with VSP2. Interestingly, they have little sequence similarity ($\sim$8–14% identity) to other members
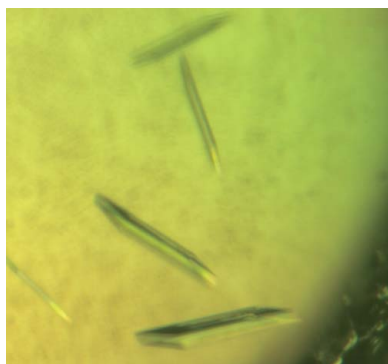
of the HAD superfamily and there are no homologous protein structures in the PDB. Crystallographic study of the *A. thaliana* VSPs should help in understanding the catalytic properties of these proteins and should also provide an insight into their function in the plant-defence or development-control systems. Here, we report the expression, purification, crystallization and preliminary X-ray studies of recombinant VSP1.

2. Methods

2.1. Cloning and expression

The VSP1 cDNA fragment corresponding to gene locus AT5g24780 was amplified by PCR from an *A. thaliana* cDNA library (Stratagene) with forward primer 5′-GGAATTC**CATATG**AAAAAT-CCTCTCACTTTCAC-3′ (the *Nde*I site is shown in bold) and reverse primer 5′-GAT**CTCGAG**AGAAGGTACGTAGTAGAGTG-3′ (the *Xho*I site is shown in bold). The PCR product was cloned into the expression vector pET-22b and the construct was transformed into *Escherichia coli* BL21 (DE3) strain. The construct contains an in-frame C-terminal His tag (Leu-Glu-His$_6$). The positive pET-22b-VSP1 expression plasmid was identified by restriction-endonuclease digestion and was further verified using DNA sequencing by Sangon Biotech (Shanghai, People's Republic of China). In order to remove the 15 N-terminal residues (the putative signal sequence), the protein was recloned with a new forward primer 5′-GGAATTC**CATATG**-GTCTCCCACGTCCAG-3′ (the *Nde*I site is shown in bold) and the same reverse primer. The recombinant cells were grown at 310 K in LB containing 100 μg ml^{-1} ampicillin until the OD_{600} reached 0.8. Expression of the recombinant protein was induced with 0.8 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) and the culture was incubated for an additional 24 h at 289 K.

2.2. Purification and activity assay

The cultured cells were harvested by centrifugation and lysed by sonication in a buffer consisting of 50 m*M* Tris–HCl pH 7.5, 1 *M* NaCl, 2 m*M* PMSF. Cell debris was removed by centrifugation at 15 000*g* for 30 min and the soluble fraction was applied onto an Ni–NTA His-Bind column (Novagen) equilibrated with binding buffer (50 m*M* Tris–HCl pH 7.5, 1 *M* NaCl). The column was eluted with a 20–200 m*M* imidazole gradient. Fractions containing the recombinant protein were pooled, desalted by dialysis and concentrated prior to activity assay and crystallization.

The presence and purity of the recombinant VSP1 was monitored by SDS–PAGE. The protein concentration was determined by a Bio-Rad protein assay with bovine serum albumin as a standard (Bradford, 1976).

Phosphatase activity was determined with *p*-nitrophenyl phosphate (*p*NPP) in 50 m*M* sodium acetate buffer pH 4.5 as described previously (Liu *et al.*, 2005). One unit of enzyme activity was defined as the amount of activity required to convert 1 μmol of substrate to product per minute at 310 K.

2.3. Crystallization and data collection

The recombinant VSP1 was buffer-exchanged into crystallization buffer (50 m*M* Tris–HCl pH 7.5, 20 m*M* NaCl) and concentrated to 12 mg ml^{-1} by centrifugal ultrafiltration (Millipore). Crystallization trials were carried out at 287 K by the sitting-drop vapour-diffusion method and the initial trials were performed using Crystal Screen kits I and II from Hampton Research (Jancarik & Kim, 1991). The 2 μl

sitting drops consisted of 1 μl protein solution and 1 μl reservoir solution and were equilibrated against 100 μl reservoir solution.

Preliminary X-ray analysis of the crystals was performed at 100 K using an R-AXIS IV^{++} image-plate detector and Cu *K*α radiation from a Rigaku FR-E rotating-anode generator. The crystals were cryoprotected by transferring them to reservoir solution containing an additional 10%(*v/v*) glycol and were cooled in a constant nitrogen stream during data collection. Diffraction data were collected from a single crystal using 1° oscillations with a crystal-to-detector distance of 110 mm and an exposure time of 300 s per image. The diffraction data were processed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997).

3. Results

The full-length cDNA of VSP1 contains 810 nucleotides and codes for a protein of 270 amino acids. When the full-length cDNA was constructed in pET-22b the recombinant protein could be over-expressed in *E. coli* BL21 (DE3). However, the expressed protein was insoluble even with low-temperature induction. Amino-acid sequence analysis suggested that the first 15 N-terminal residues of the protein were very hydrophobic and may be part of the signal sequence that locates the cellular position of the protein. After recloning and removal of the first 15 N-terminal residues, the recombinant VSP1 (residues 16–270) expressed well and was found in the supernatant of the bacterial extract. It could be further purified to apparent homogeneity *via* Ni-chelate chromatography. The purified protein showed a single protein band with a molecular mass of 29 kDa on SDS–PAGE, which corresponded well to the theoretical molecular weight of 28.8 kDa. A size-exclusion chromatography experiment indicated that VSP1 exists as a dimer in solution. The purified VSP1 showed acid phosphatase activity. The specific activity of the recombinant VSP1 was 16.5 U mg^{-1} and its maximum activity was obtained at pH 4.5 when *p*-nitrophenyl phosphate was used as substrate.

Initial crystals of VSP1 appeared in condition Nos. 40 [0.1 *M* sodium citrate tribasic dihydrate pH 5.6, 20%(*v/v*) 2-propanol, 20%(*w/v*) polyethylene glycol 4000] and 46 [0.2 *M* calcium acetate hydrate, 0.1 *M* sodium cacodylate trihydrate pH 6.5, 18%(*w/v*) polyethylene glycol 8000] of Crystal Screen I within a week. After optimization, the best crystals were grown by mixing 1.0 μl of
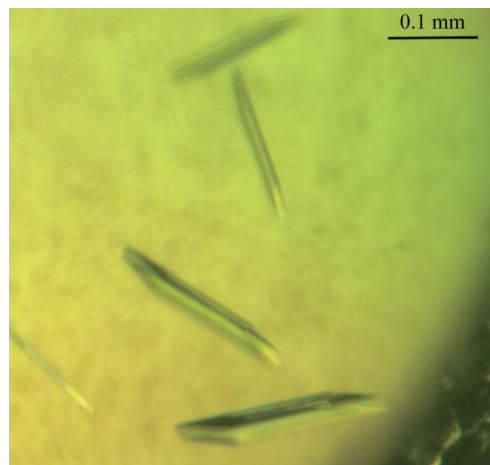

Figure 1
Crystals of recombinant VSP1 grown by the sitting-drop method. The crystals grew to dimensions of ∼0.02 × 0.1 × 0.15 mm.

Table 1
Summary of diffraction data collection and processing.

Space group	$C2$
Unit-cell parameters (Å, °)	$a = 123.1$, $b = 48.4$, $c = 85.6$,
	$\alpha = \gamma = 90$, $\beta = 116.3$
Wavelength (Å)	1.5418
Resolution range (Å)	30.00–1.90 (1.97–1.90)
No. of measured reflections	169773
No. of unique reflections	35624
Completeness (%)	99.1 (97.9)
Redundancy	4.8 (4.7)
R_{merge}†	0.058 (0.303)
Mean $I/\sigma(I)$	26.4 (5.3)
Matthews value (Å^3 Da^{-1})	2.3
Z	8
Solvent content (%)	45.8

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of reflection $I(hkl)$ and $I_i(hkl)$ is the intensity of an individual measurement of reflection $I(hkl)$.

10 mg ml^{-1} protein solution and an equal volume of reservoir solution containing 0.1 M sodium citrate tribasic dihydrate pH 5.6, 8%(v/v) 2-propanol, 20%(w/v) polyethylene glycol 4000. The rod-like crystals grew to their full dimensions (approximately $0.02 \times 0.1 \times 0.15$ mm) within one week (Fig. 1). A complete data set was collected to 1.9 Å resolution at 100 K. Data-collection and processing statistics are listed in Table 1. The VSP1 crystal belonged to the monoclinic space group $C2$, with unit-cell parameters $a = 123.1$, $b = 48.4$, $c = 85.6$ Å, $\beta = 116.3°$. Assuming the presence of two molecules per asymmetric unit, the Matthews coefficient (V_M) was calculated as 2.3 Å^3 Da^{-1}, which corresponds to 45.8% solvent content.

Since no related structures are available as search models to solve the VSP1 structure by molecular replacement, our efforts are now focused on experimental phasing using heavy-atom and selenomethionine derivatives.

References

Allen, K. N. & Dunaway-Mariano, D. (2004). *Trends Biochem. Sci.* **29**, 495–503.

Berger, S., Bell, E., Sadka, A. & Mullet, J. E. (1995). *Plant Mol. Biol.* **27**, 933–942.

Berger, S., Mitchell-Olds, T. & Stotz, H. U. (2002). *Physiol. Plant.* **114**, 85–91.

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

Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. (2006). *J. Mol. Biol.* **361**, 1003–1034.

Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G. & Van Schaftingen, E. (1998). *J. Biol. Chem.* **273**, 14107–14112.

Ellis, C. & Turner, J. G. (2001). *Plant Cell*, **13**, 1025–1033.

Gamboa, A., Paez, V. J., Acevedo, G. F., Vazquez, M. L. & Alvarez, B. R. E. (2001). *Biochem. Biophys. Res. Commun.* **288**, 1018–1026.

Guerineau, F., Benjdia, M. & Zhou, D. X. (2003). *J. Exp. Bot.* **54**, 1153–1162.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Leelapon, O., Sarath, G. & Staswick, P. E. (2004). *Planta*, **219**, 1071–1079.

Liu, Y. L., Ahn, J. E., Datta, S., Salzman, R. A., Moon, J. W., Huyghues-Despointes, B., Pittendrigh, B., Murdock, L. L., Koiwa, H. & Zhu-Salzman, K. (2005). *Plant Physiol.* **139**, 1545–1556.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rojo, E., Leon, J. & Sanchez-Serrano, J. J. (1999). *Plant J.* **20**, 135–142.

Selengut, J. D. (2001). *Biochemistry*, **40**, 12704–12711.

Staswick, P. E. (1990). *Plant Cell*, **2**, 1–6.

Staswick, P. E. (1994). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 303–322.

Stotz, H. U., Pittendrigh, B. R., Kroymann, J., Weniger, K., Fritsche, J., Bauke, A. & Mitchell-Olds, T. (2000). *Plant Physiol.* **124**, 1007–1017.

Thaller, M. C., Schippa, S. & Rossolini, G. M. (1998). *Protein Sci.* **7**, 1647–1652.