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

Crystallization and preliminary X-ray data analysis of a DJ-1 homologue from *Arabidopsis thaliana* (*At*DJ-1D)

Kyung Hye Seo,[a,b] Ningning Zhuang,[a,b] Joon-Yung Cha,[a,b] Daeyoung Son[b,c]* and Kon Ho Lee[b,d]*

[a]Division of Applied Life Science (BK21 Program), Gyeongsang National University, Jinju 660-701, Republic of Korea, [b]PMBBRC, Gyeongsang National University, Jinju 660-701, Republic of Korea, [c]Department of Applied Biology, Gyeongsang National University, Jinju 660-701, Republic of Korea, and [d]Department of Microbiology, School of Medicine, Gyeongsang National University, Jinju 660-751, Republic of Korea

Correspondence e-mail: dyson@gnu.ac.kr, lkh@gnu.ac.kr

Received 6 October 2011
Accepted 24 November 2011

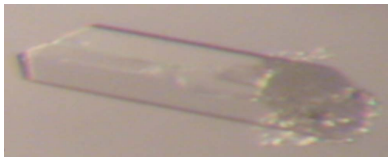
A DJ-1 homologue protein from *Arabidopsis thaliana* (*At*DJ-1D) belongs to the DJ-1/ThiJ/Pfpl superfamily and contains two tandem arrays of DJ-1-like sequences, but no structural information is available to date for this protein. *At*DJ-1D was expressed in *Escherichia coli*, purified and crystallized for structural analysis. A crystal of *At*DJ-1D was obtained by the hanging-drop vapour-diffusion method using 0.22 *M* NaCl, 0.1 *M* bis-tris pH 6.5, 21% polyethylene glycol 3350. *At*DJ-1D crystals belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 56.78$, $b = 75.21$, $c = 141.68$ Å, $\beta = 96.87^\circ$, and contained a trimer in the asymmetric unit. Diffraction data were collected to 2.05 Å resolution. The structure of *At*DJ-1D has been determined using the multiple-wavelength anomalous dispersion (MAD) method.

1. Introduction

The DJ-1/ThiJ/Pfpl superfamily includes a broad spectrum of proteins from most organisms which are characterized by having a highly conserved DJ-1 domain. DJ-1 (also known as PARK7) was first identified as a novel oncogene in mouse (Nagakubo *et al.*, 1997) and was later identified as an autosomal recessive gene of Parkinson's disease (PD; Bonifati *et al.*, 2003; Wilson *et al.*, 2003; Mo *et al.*, 2008). Human DJ-1 is known to be involved in many different biochemical functions such as oncogenesis (Nagakubo *et al.*, 1997), neuroprotection (Park *et al.*, 2005; Kim *et al.*, 2005), male fertility (Okada *et al.*, 2002), control of protein–RNA interactions (van der Brug *et al.*, 2008) and modulation of androgen receptor transcription activity (Takahashi *et al.*, 2001; Tillman *et al.*, 2007). Furthermore, DJ-1 possesses protease and chaperone activities (Olzmann *et al.*, 2004; Shendelman *et al.*, 2004) and plays an established role in cellular protection against oxidative stress (Canet-Avilés *et al.*, 2004; Gu *et al.*, 2009; Kim *et al.*, 2005; Martinat *et al.*, 2004; Meulener *et al.*, 2006). Other members of the DJ-1/ThiJ/Pfpl superfamily from other organisms have been also characterized. However, their precise biochemical functions have not been as fully studied as those of the human homologue.

In the model plant *Arabidopsis thaliana*, six genes encoding homologues of human DJ-1 have been identified: *At*3g14990 (*At*DJ-1A), *At*1g53280 (*At*DJ-1B), *At*4g34020 (*At*DJ-1C), *At*3g02720 (*At*DJ-1D), *At*2g38860 (*At*DJ-1E) and *At*3g54600 (*At*DJ-1F). Of these plant DJ-1 homologues, *At*DJ-1A and *At*DJ-1C have recently been characterized. *At*DJ-1A carries the functionally conserved DJ-1 amino-acid residues associated with PD conditions and confers protection against oxidative stress by cytosolic superoxide dismutase activation (Xu *et al.*, 2010). In contrast, *At*DJ-1C, which does not have the highly conserved cysteine, is a chloroplast-targeted protein and is essential for chloroplast development (Lin *et al.*, 2011). No functional information on the other plant homologue proteins is available to date.

Interestingly, these plant DJ-1 homologues carry two full-length DJ-1 polypeptides arranged in tandem. Considering that members of the DJ-1/ThiJ/Pfpl superfamily in bacteria and mammals contain only one DJ-1 domain (Wei *et al.*, 2007) and form a dimer [or sometimes higher oligomers; Wilson *et al.*, 2005; Canet-Avilés *et al.*, 2004; Du *et*



al., 2000; YhbO from *Escherichia coli* (B. Claude, C. Abergel & J.-M. Claverie, unpublished work; PDB entry 1oi4)], these plant homologues have a unique structure of two DJ-1 domains in a tandem array, which distinguishes them from the other DJ-1 superfamily members. It seems that this oligomerization is important for function (Wei *et al.*, 2007; Wilson *et al.*, 2005). In order to understand their unique structural arrangement of two DJ-1 domains in tandem, one of these six plant DJ-1 homologues, *At*DJ-1D (*At*3g02720), which has 25% sequence identity to human DJ-1, was selected for structural study. Here, we report the overexpression from *Escherichia coli*, purification and crystallization of *At*DJ-1D. A crystal of *At*DJ-1D diffracted to 2.05 Å resolution and the structure has been solved by the multiple-wavelength anomalous dispersion (MAD) method.

2. Materials and methods

2.1. Protein expression and purification

First-strand *At*DJ-1D cDNA was prepared *via* reverse transcriptase polymerase chain reaction using SuperScript II RT (Invitrogen Inc., Burlington, Ontario, Canada). Primers were designed using the forward primer 5′-GCGGATCCATGGCGAACTCGAGAACTG-TC-3′ and the reverse primer 5′-GCGTCGACTTAGAATGAAAC-CTGAATGCC-3′. The PCR product was digested with *Bam*HI and *Sal*I restriction enzymes and ligated into the *Bam*HI and *Sal*I restriction sites of a pET-28a expression plasmid (Novagen, Madison, Wisconsin, USA) to produce recombinant *At*DJ-1D with a His tag and a T7 tag at the N-terminus (MGSSHHHHHHSSGLVPRGS-HMNSMTGGQQMGRGS). The DNA sequence of the inserted *At*DJ-1D was confirmed by DNA sequencing. The calculated molecular weight of the tagged monomer was 45.2 kDa and the theoretical isoelectric point was 5.8.

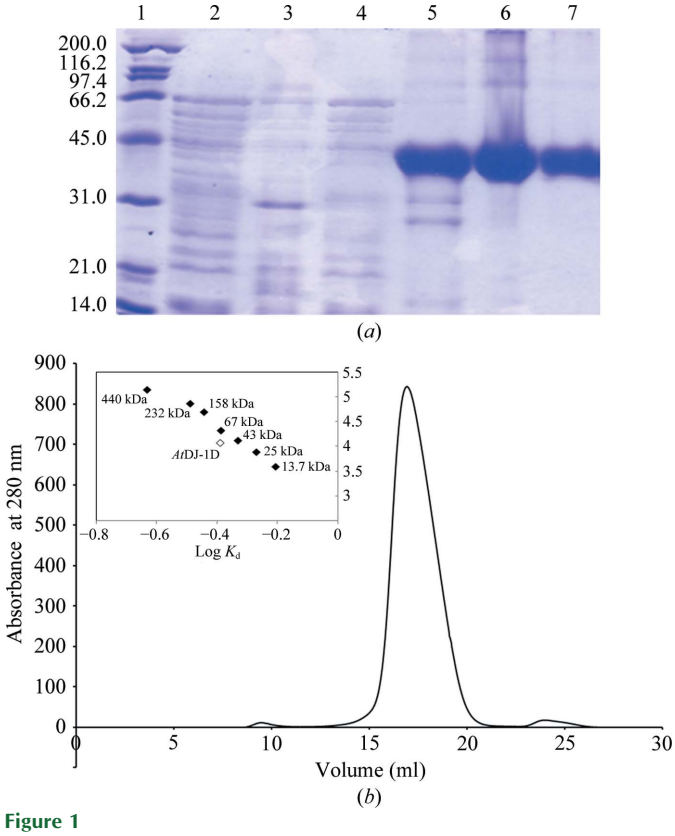


Figure 1
(*a*) SDS–PAGE analysis of *At*DJ-1D at different steps of the purification process. Lane 1, marker (labelled in kDa); lane 2, total cell lysate; lane 3, supernatant after cell lysis; lane 4, pellet after cell lysis; lane 5, proteins from an Ni–NTA affinity column; lane 6, proteins separated after anion-exchange chromatography; lane 7, proteins separated after Superdex 200 chromatography (8.5 µg protein used). (*b*) Superdex 200 size-exclusion chromatography of *At*DJ-1D. A single peak corresponding to a monomer was observed. The scale at the bottom indicates the elution volume. Inset, semi log plot of the molecular mass of the standard proteins used *versus* their log K_d values.

The plasmid was transformed into *Escherichia coli* strain BL21 (DE3) for protein expression. A 10 ml aliquot of an overnight culture was seeded into 1000 ml fresh M9 minimal medium (6 g Na_2HPO_4, 3 g KH_2PO_4, 1 g NH_4Cl, 0.5 g NaCl, 2 g glucose, 2 m*M* $MgSO_4$ and 0.1 m*M* $CaCl_2$ per litre of solution) containing 30 mg ml^{-1} kanamycin and grown to an OD_{600} of 0.6 at 303 K. The protein was expressed at 288 K for 3 d with 1 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation at 6520*g* for 6 min at 277 K, suspended in binding buffer (50 m*M* NaH_2PO_4, 0.5 *M* NaCl, 5 m*M* imidazole and 5 m*M* *β*-mercaptoethanol per litre of solution, pH 8.0) and disrupted by sonication. The crude lysate was centrifuged at 15 930*g* for 1 h at 277 K. The clear supernatant was then filtered (Qualitative filter paper, Advantec) and applied onto a column of Nickel Sepharose 6 Fast Flow (GE Healthcare, Sweden) pre-equilibrated with the binding buffer. The column was washed first with 20 column volumes of binding buffer and then with two column volumes of washing buffer (50 m*M* Tris–HCl pH 8.0, 0.5 *M* NaCl, 20 m*M* imidazole, 5 m*M* *β*-mercaptoethanol). The recombinant *At*DJ-1D was eluted with 50 m*M* Tris–HCl pH 8.0, 0.1 *M* NaCl, 0.3 *M* imidazole and 5 m*M* *β*-mercaptoethanol. Fractions containing *At*DJ-1D were pooled, concentrated and exchanged into 50 m*M* Tris–HCl pH 8.0, 10% glycerol and 1 m*M* DTT by ultrafiltration (Centriprep YM-30, Millipore Corporation, Bedford, Massachusetts, USA). The *At*DJ-1D was further purified by anion-exchange chromatography on a Resource 15Q column (GE Healthcare, Piscataway, New Jersey, USA) and eluted with a linear gradient of 0–500 m*M* NaCl in 50 m*M* Tris–HCl pH 8.0, 10% glycerol and 1 m*M* DTT. The fractions of the major peak were concentrated by ultrafiltration (Centriprep YM-30, Millipore Corporation, Bedford, Massachusetts, USA). The *At*DJ-1D was finally purified by Superdex 200 size-exclusion chromatography (GE Healthcare, Piscataway, New Jersey, USA) equilibrated with 20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 10% glycerol and 1 m*M* DTT. *At*DJ-1D eluted as a monomer and was sufficiently pure for crystallization as judged by SDS–PAGE (Fig. 1). The fractions were pooled and concentrated to a final concentration of 8.5 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0, 10% glycerol and 1 m*M* DTT by ultrafiltration (Microcon YM-30, Millipore Corporation, Bedford, Massachusetts, USA). All purification steps were carried out at 293 K with ice-cooled buffers. The protein concentration was determined by the Bradford assay using bovine serum albumin as a standard. The N-terminal His tag and T7 tag were not removed for crystallization.

2.2. Crystallization

For crystallization trials of *At*DJ-1D, Crystal Screen, Crystal Screen 2 and Index (Hampton Research, California, USA) and Wizard Screens I and II and Cryo I and II (Emerald BioStructures, Bainbridge Island, Washington, USA) and laboratory solutions were used. Crystallization was set up in 72-well plates at 291 K using the microbatch method under Al's oil (D'Arcy *et al.*, 2003). For each screen, equal volumes (1 µl) of protein solution (8.5 mg ml^{-1}) and reservoir solution were mixed and equilibrated under the oil. Crystals were initially produced in the condition Index solution No. 71 (0.2 *M* NaCl, 0.1 *M* bis-tris pH 6.5, 25% polyethylene glycol 3350).

Table 1
Data statistics.

Values in parentheses are for the highest resolution shell.

	SeMet derivative		
	Peak	Edge	Remote
Space group	$P2_1$		
Unit-cell parameters (Å, °)	a = 56.78, b = 75.21, c = 141.68, β = 96.87		
No. of chains in the asymmetric unit	3		
V_M (Å^3 Da^{-1})	2.38		
Solvent content (%)	48.43		
Wavelength (Å)	0.97954	0.97955	0.97186
Resolution (Å)	2.05 (2.17–2.05)	2.05 (2.12–2.05)	2.07 (2.14–2.07)
Unique reflections	73962 (6948)	73914 (6926)	73914 (6926)
Completeness (%)	99.2 (94.1)	99.1 (93.8)	99.5 (93.8)
R_{merge}† (%)	0.140 (0.497)	0.166 (0.733)	0.131 (0.510)
Multiplicity	3.6 (3.4)	3.6 (3.4)	3.6 (3.4)
$\langle I/\sigma(I)\rangle$	6.1 (4.6)	5.9 (4.1)	5.9 (4.1)
Figure of merit	0.45‡/0.80§		

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith intensity measurement of reflection hkl, including symmetry-related reflections, and $\langle I(hkl)\rangle$ is its average. ‡ As indicated by *SOLVE*. § As indicated by *RESOLVE*.

Further screening to find optimal crystallization conditions was performed by hanging-drop vapour-diffusion trials, varying the salt and precipitant concentrations and the volume of the drop. The best crystals were grown at 291 K using *At*DJ-1D protein (8.5 mg ml^{-1}) in 20 m*M* Tris–HCl pH 8.0, 10% glycerol and 1 m*M* DTT and reservoir solution consisting of 0.22 *M* NaCl, 0.1 *M* bis-tris pH 6.5, 21% polyethylene glycol 3350. The drop consisted of 3 µl protein solution and 4 µl of a mixture of 4 µl reservoir solution and 1 µl Index Screen solution No. 44 (0.1 *M* urea) as an additive. Rod-shaped *At*DJ-1D crystals appeared after 2–3 d and grew to maximum dimensions of 0.4 × 0.3 × 0.2 mm (Fig. 2).

2.3. X-ray data collection

Before data collection, crystals were transferred for about 30 min into a cryosolution consisting of mother liquor supplemented with 15% ethylene glycol. The crystals were flash-cooled in liquid nitrogen. X-ray diffraction data were collected on beamline 6C1 of Pohang Accelerator Laboratory, Republic of Korea using an ADSC Quantum 210 CCD detector. For multiple-wavelength anomalous dispersion (MAD), three data sets were collected at wavelengths of 0.97954 Å (peak), 0.97955 Å (edge) and 0.98736 Å (remote) at 100 K. 180° of data were collected at each wavelength. All diffraction images were indexed, integrated and scaled using the *HKL*-2000 suite (Otwinowski & Minor, 1997). Data statistics are shown in Table 1. The positions of Se atoms were determined using *SOLVE* (Terwilliger & Berendzen, 1999). Density modification was performed with *RESOLVE* (Terwilliger, 2001).

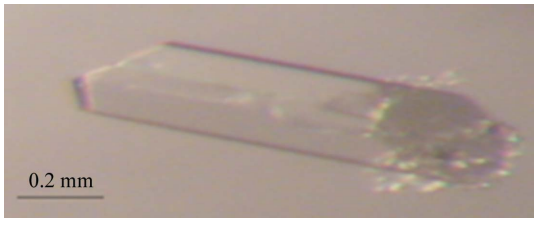


Figure 2
Crystal of *At*DJ-1D obtained by the hanging-drop method.

3. Results

We have established the expression, purification and crystallization of *At*DJ-1D. We collected a 2.05 Å resolution MAD data set at three wavelengths (peak, 0.97954 Å; edge, 0.97955 Å; remote, 0.98736 Å). Indexing indicated that the crystals of *At*DJ-1D belonged to the monoclinic space group $P2_1$, with unit-cell parameters a = 56.78, b = 75.21, c = 141.68 Å, β = 96.87° (Table 1). A solvent content of 48.4% and a Matthews coefficient V_M of 2.38 Å^3 Da^{-1} (Matthews, 1968) would indicate that the asymmetric unit contains three molecules. Phases were determined from the positions of 17 Se atoms located in *At*DJ-1D that were found from a total of 30 Se atoms (18 in *At*DJ-1D and 12 in the fusion tag) using *SOLVE* and *RESOLVE*. The structure of *At*DJ-1D was determined by the multiple-wavelength anomalous dispersion (MAD) method. The molecular weight of the protein was estimated to be about 45 kDa from Superdex 200 size-exclusion chromatography and that of the monomer of *At*DJ-1D was found to be approximately 45 kDa from SDS–PAGE, indicating that the protein exists as a monomer in solution (Fig. 1). However, *At*DJ-1D forms a trimer in the asymmetric unit in which the three monomers are related by noncrystallographic symmetry. It is not clear that the trimerization of *At*DJ-1D is related to its biological function. A detailed discussion of the refined structure will be published elsewhere.

We thank the staff of beamlines 6B, 6C and 4A at Pohang Accelerator Laboratory for help with data collection. This work was supported by the National Research Foundation of Korea/MEST Grant Nos. NRF-2010-C1AAA001-2010-0029084 (KHL) and 2009-0068189 (DS).

References

Bonifati, V. *et al.* (2003). *Science*, **299**, 256–259.
Brug, M. P. van der *et al.* (2008). *Proc. Natl Acad. Sci. USA*, **105**, 10244–10249.
Canet-Avilés, R. M., Wilson, M. A., Miller, D. W., Ahmad, R., McLendon, C., Bandyopadhyay, S., Baptista, M. J., Ringe, D., Petsko, G. A. & Cookson, M. R. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 9103–9108.
D'Arcy, A., Mac Sweeney, A., Stihle, M. & Haber, A. (2003). *Acta Cryst.* D**59**, 396–399.
Du, X., Choi, I.-G., Kim, R., Wang, W., Jancarik, J., Yokota, H. & Kim, S.-H. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 14079–14084.
Gu, L., Cui, T., Fan, C., Zhao, H., Zhao, C., Lu, L. & Yang, H. (2009). *Biochem. Biophys. Res. Commun.* **383**, 469–474.
Kim, R. H., Smith, P. D., Aleyasin, H., Hayley, S., Mount, M. P., Pownall, S., Wakeham, A., You-Ten, A. J., Kalia, S. K., Horne, P., Westaway, D., Lozano, A. M., Anisman, H., Park, D. S. & Mak, T. W. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 5215–5220.
Lin, J., Nazarenus, T. J., Frey, J. L., Liang, X., Wilson, M. A. & Stone, J. M. (2011). *PLoS One*, **6**, e23731.
Martinat, C., Shendelman, S., Jonason, A., Leete, T., Beal, M. F., Yang, L., Floss, T. & Abeliovich, A. (2004). *PLoS Biol.* **2**, e327.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Meulener, M. C., Xu, K., Thomson, L., Thompson, L., Ischiropoulos, H. & Bonini, N. M. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 12517–12522.
Mo, J.-S., Kim, M.-Y., Ann, E.-J., Hong, J.-A. & Park, H.-S. (2008). *Cell Death Differ.* **15**, 1030–1041.
Nagakubo, D., Taira, T., Kitaura, H., Ikeda, M., Tamai, K., Iguchi-Ariga, S. M. & Ariga, H. (1997). *Biochem. Biophys. Res. Commun.* **231**, 509–513.
Okada, M., Matsumoto, K., Niki, T., Taira, T., Iguchi-Ariga, S. M. & Ariga, H. (2002). *Biol. Pharm. Bull.* **25**, 853–856.
Olzmann, J. A., Brown, K., Wilkinson, K. D., Rees, H. D., Huai, Q., Ke, H., Levey, A. I., Li, L. & Chin, L.-S. (2004). *J. Biol. Chem.* **279**, 8506–8515.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Park, J., Kim, S. Y., Cha, G.-H., Lee, S. B., Kim, S. & Chung, J. (2005). *Gene*, **361**, 133–139.
Shendelman, S., Jonason, A., Martinat, C., Leete, T. & Abeliovich, A. (2004). *PLoS Biol.* **2**, e362.

Takahashi, K., Taira, T., Niki, T., Seino, C., Iguchi-Ariga, S. M. & Ariga, H. (2001). *J. Biol. Chem.* **276**, 37556–37563.
Terwilliger, T. C. (2001). *Acta Cryst.* D**57**, 1763–1775.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 1872–1877.
Tillman, J. E., Yuan, J., Gu, G., Fazli, L., Ghosh, R., Flynt, A. S., Gleave, M., Rennie, P. S. & Kasper, S. (2007). *Cancer Res.* **67**, 4630–4637.
Wei, Y., Ringe, D., Wilson, M. A. & Ondrechen, M. J. (2007). *PLoS Comput. Biol.* **3**, e10.
Wilson, M. A., Collins, J. L., Hod, Y., Ringe, D. & Petsko, G. A. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 9256–9261.
Wilson, M. A., Ringe, D. & Petsko, G. A. (2005). *J. Mol. Biol.* **353**, 678–691.
Xu, X. M., Lin, H., Maple, J., Björkblom, B., Alves, G., Larsen, J. P. & Møller, S. G. (2010). *J. Cell Sci.* **123**, 1644–1651.