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

Purification, crystallization and preliminary X-ray analysis of PPD6, a PsbP-domain protein from *Arabidopsis thaliana*

Michael Hall,* Thomas Kieselbach, Uwe H. Sauer and Wolfgang P. Schröder

Department of Chemistry, Umeå University, SE-90187 Umeå, Sweden

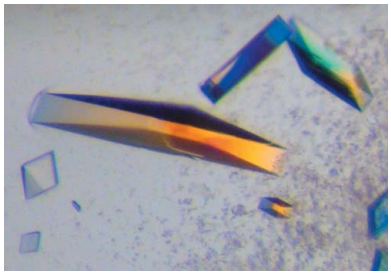
Correspondence e-mail: michael.hall@chem.umu.se

Received 9 September 2011
Accepted 11 October 2011

The PsbP protein is an extrinsic component of photosystem II that together with PsbO and PsbQ forms the thylakoid lumenal part of the oxygen-evolving complex in higher plants. In addition to PsbP, the thylakoid lumen contains two PsbP-like proteins (PPLs) and six PsbP-domain proteins (PPDs). While the functions of the PsbP-like proteins PPL1 and PPL2 are currently under investigation, the function of the PsbP-domain proteins still remains completely unknown. PPD6 is unique among the PsbP family of proteins in that it contains a conserved disulfide bond which can be reduced *in vitro* by thioredoxin. The crystal structure determination of the PPD6 protein has been initiated in order to elucidate its function and to gain deeper insights into redox-regulation pathways in the thylakoid lumen. PPD6 has been expressed, purified and crystallized and preliminary X-ray diffraction data have been collected. The crystals belonged to space group $P2_1$, with unit-cell parameters $a = 47.0$, $b = 64.3$, $c = 62.0$ Å, $\beta = 94.2^\circ$, and diffracted to a maximum d-spacing of 2.1 Å.

1. Introduction

In plants and cyanobacteria, the essential process of oxygenic photosynthesis takes place in the thylakoid membrane system of the chloroplast. Located in this membrane is photosystem II (PSII), a multisubunit protein–pigment complex that uses light energy to oxidize (split) water into dioxygen gas and protons. The PsbP protein is an extrinsic component of PSII, which together with the PsbO and PsbQ subunits constitutes the lumenal part of the oxygen-evolving complex in higher plants (Seidler, 1996). In addition to these well studied extrinsic components of PSII, the thylakoid lumen contains a number of PsbP-like proteins (Kieselbach *et al.*, 2000). The genome of the model plant *Arabidopsis thaliana* codes for ten members of the PsbP protein family, two true PsbP proteins (PsbP1 and PsbP2) and eight PsbP homologues, all of which are experimentally proven or predicted to be located in the thylakoid lumen (Schubert *et al.*, 2002; Peltier *et al.*, 2002). Two of the eight homologues which share the highest sequence identity to PsbP1 have been classified as PsbP-like proteins (PPL1 and PPL2), while the remaining six are referred to as PsbP-domain proteins (PPD). Analysis of *Arabidopsis* PPL1 and PPL2 knockout mutants revealed that in contrast to PsbP, these proteins are not directly involved in the process of oxygen evolution. Instead, PPL1 is required for efficient repair of photodamaged PSII, while PPL2 is needed for accumulation of the chloroplast NADH dehydrogenase complex (Ishihara *et al.*, 2007). To date, no function has been assigned for any of the six PPD proteins. Although the PsbP-like and PsbP-domain proteins exhibit very low sequence identity to PsbP, 10–25% in *Arabidopsis* and 14–33% in cyanobacteria, one would nevertheless expect that they retain the same overall structural characteristics (Ishihara *et al.*, 2007). The high-resolution crystal structure of PsbP from *Nicotiana tabacum* (1.6 Å resolution; PDB entry 1v2b; Ifuku *et al.*, 2004) and the recent crystal structure of cyanobacterial PsbP (2.8 Å resolution; PDB entry 2xb3; Michoux *et al.*, 2010) revealed that the core region of the PsbP protein consists of a six-stranded antiparallel β-sheet with one α-helix on either side of the sheet. Sato concluded from a phylogenetic analysis and structural modelling of the PsbP-family proteins that this fold is



likely to be conserved among the family members and that variations in the structures are located mainly in the N-terminal and loop regions (Sato, 2010). The PPD6 protein contains an interesting feature which is unique among the members of the PsbP family. The loop region between β-strands 3 and 4 contains two conserved cysteine residues separated by three amino acids. These two cysteines form a reactive disulfide bond, which can be reduced *in vitro* by *Escherichia coli* thioredoxin (Hall *et al.*, 2010). To date, no soluble thioredoxin has been identified in the thylakoid lumen (Cain *et al.*, 2009). However, the *hcf164* gene product is a thioredoxin-like protein anchored to the thylakoid membrane with its catalytic domain facing the lumenal side (Lennartz *et al.*, 2001). It has been suggested that stroma-located m-type thioredoxin could facilitate the transfer of reducing equivalents across the thylakoid membrane to the catalytic moiety of HCF164, which in turn could reduce target proteins in the thylakoid lumen (Motohashi & Hisabori, 2006). The redox regulation of PPD6 is probably an important aspect of its cellular function and distinguishes it from the other PsbP-family proteins, since they are not predicted to contain any conserved intramolecular disulfide bonds. In order to gain insight into the molecular details of the possible redox regulation of PPD6, we have expressed, purified and crystallized PPD6 from *A. thaliana*. Alongside biochemical studies and analysis of knockout lines, the crystal structure of PPD6 will allow us to elucidate its function and provide insight into the role which thylakoid lumenal redox pathways play in regulation of the photosynthetic process.

2. Materials and methods

2.1. Construct and expression

The coding sequence of PPD6 was obtained from The *Arabidopsis* Information Resource (TAIR ID At3g56650.1). The full-length PPD6 protein contains 262 amino acids, which include a 67-amino-acid thylakoid lumen transit peptide at the N-terminus. The gene sequence corresponding to the mature protein, lacking the transit peptide, was created by the removal of bases 1–201, which code for amino acids 1–67. A synthetic gene construct (GenScript, Piscataway, New Jersey, USA) coding for the mature PPD6 protein starting at amino acid Arg68 and ending with Leu262 (195 amino acids) was cloned into the pET15b expression vector (Novagen). The vector provides an N-terminal six-histidine tag and a thrombin protease recognition site, thus yielding a fusion protein of 216 amino acids. Cleavage with thrombin leaves four extra non-native amino acids (GSHM) at the N-terminus of the mature protein. The construct was transformed into the cysteine disulfide-promoting *E. coli* Rosetta Gami 2 strain (Novagen). Cultures were grown at 310 K in LB medium to an OD_{600} of 0.6–1.0, at which point protein expression was induced for 16 h at 293 K by addition of IPTG (isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 0.5 m*M*.

2.2. IMAC purification

Following induction, the cells were harvested from 1 l of culture by centrifugation for 10 min at 8000*g*. The cells were resuspended in lysis buffer (20 m*M* Na_2HPO_4 pH 7.4, 500 m*M* NaCl, 20 m*M* imidazole) and sonicated on ice for 6 × 10 s using a Branson Digital Sonifier 450 (Branson, Danbury, Connecticut, USA). The supernatant fraction was isolated by centrifugation for 30 min at 20 000*g* and 277 K. The soluble lysate was applied onto a 1 ml HisGraviTrap pre-packed nickel-affinity column (GE Healthcare, Uppsala, Sweden) equilibrated with ten column volumes of lysis buffer. Following binding of the target protein, the column was washed with ten column volumes of lysis buffer and finally target protein was eluted with three column volumes of elution buffer (20 m*M* Na_2HPO_4 pH 7.4, 500 m*M* NaCl, 500 m*M* imidazole). The six-histidine tag was removed by the addition of ten units of thrombin protease (GE Healthcare, Uppsala, Sweden) per milligram of target protein, followed by 16 h incubation at 293 K. Complete cleavage was confirmed by SDS–PAGE.

2.3. Size-exclusion chromatography

Nickel affinity-purified PPD6 was further purified by a subsequent size-exclusion chromatography step. Protein samples were separated on a HiPrep 16/60 Sephacryl S-100 HR column (GE Healthcare, Uppsala, Sweden) equilibrated with 20 m*M* 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 6.5, 150 m*M* NaCl at a maximum backpressure of 0.15 MPa at 293 K. Fractions containing PPD6 were collected, pooled and concentrated using a Vivaspin concentrator with a 10 kDa molecular-weight cutoff (GE Healthcare, Uppsala, Sweden). The protein concentration was determined by the Bradford assay using BSA as standard (Bradford, 1976) and the purity of PPD6 was assessed by SDS–PAGE. After removal of the six-histidine tag, PPD6 has a calculated molecular mass of 22.0 kDa, compared with 23.8 kDa for the full-length fusion protein. Aliquots of purified PPD6 were flash-cooled in liquid nitrogen and stored at 193 K until use.

2.4. Crystallization

Initial crystallization screening was performed in 96-well sitting-drop plates (MRC-Wilden, UK) using a Mosquito crystallization robot (TTP LabTech, Melbourn, UK) with two commercial screens: Crystal Screen and Crystal Screen 2 (Hampton Research, Aliso Viejo, California, USA). The drops comprised 100 nl protein solution at 13 mg ml^{-1} (in 20 m*M* HEPES pH 6.5, 150 m*M* NaCl) and 100 nl reservoir solution and were equilibrated at 291 K against a reservoir containing 80 µl of the respective crystallization buffer. Manual optimization of the growth conditions was performed using 24-well Linbro plates and hanging drops of 2–12 µl in volume equilibrated against 500 µl reservoir solution.

2.5. Data collection and processing

Crystals were mounted free-standing in cryoloops, vitrified in liquid nitrogen and stored in a cryodewar until use. No additional cryoprotection steps were required. Diffraction data were collected in two sweeps and were recorded using a MAR 165 CCD detector on beamline I911-2 at the MAX-lab synchrotron, Lund, Sweden. High-resolution and low-resolution data sets were collected by rotating the crystals by 2° per image for 180 frames. Indexing and integration was carried out with *XDS* (Kabsch, 2010) and the two data passes were combined with *POINTLESS* (Evans, 2006) and scaled using *SCALA* (Evans, 2006). The Matthews coefficient (Matthews, 1968) was calculated using the program *MATTHEWS_COEF* (Kantardjieff & Rupp, 2003) from the *CCP*4 suite (Winn *et al.*, 2011).

3. Results and discussion

The PPD6 protein could be successfully overexpressed and purified to homogeneity with a typical final yield of 10–15 mg protein per litre of bacterial culture. Initial crystal screening yielded small crystals in 0.1 *M* sodium cacodylate pH 6.5, 18%(*w*/*v*) PEG 8000, 0.2 *M* zinc acetate (condition No. 45 of Crystal Screen). During optimization of the crystallization conditions, the protein crystallized over a wide pH range (4–8) using sodium cacodylate as a buffering agent and a

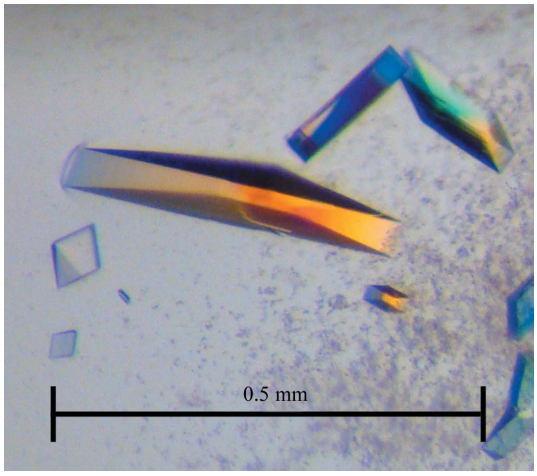


Figure 1
PPD6 crystals grown in 0.1 *M* HEPES pH 6.5, 20%(*w*/*v*) PEG 4000, 0.2 *M* zinc acetate at 291 K. Crystals typically grew to a maximum length of 0.5 mm after one week.

wide range of protein concentrations (5–15 mg ml^{-1}) and precipitant concentrations [12–22%(*w*/*v*) PEG 4000 or PEG 8000]. 0.2 *M* zinc acetate was required in all conditions; crystallization attempts with the zinc acetate replaced by magnesium acetate or cupric acetate did not result in any crystal growth. Across the entire pH range, crystals typically displayed rhomboidal or triangular morphology, with some thin plate crystals also being present. Performing crystallization without prior thrombin cleavage of the six-histidine tag from the recombinant PPD6 protein resulted in star-shaped crystals which could not be used for data collection owing to extensive smearing of the diffraction spots. After optimization of the crystallization conditions, the best diffracting crystals of PPD6 were grown at 291 K from 0.1 *M* HEPES pH 6.5, 20%(*w*/*v*) PEG 4000, 0.2 *M* zinc acetate using a 5 mg ml^{-1} protein solution. Triangular and rhomboidal crystals of up to 0.5 mm in length were obtained after approximately one week when growing crystals in drops containing 4 µl reservoir solution and 8 µl protein solution (Fig. 1). The crystals belonged to space group $P2_1$, with unit-cell parameters $a = 47.0$, $b = 64.3$, $c = 62.1$ Å, $\beta = 94.2^\circ$, and typically diffracted to a minimal *d*-spacing $d_{\rm min}$ of 2.1 Å ($d_{\rm min} = \lambda/2\sin\theta_{\rm max}$). Considering the calculated molecular weight of 22.0 kDa for the protein after thrombin cleavage and a unit-cell volume of 186 581 Å^3, this results in a Matthews coefficient $V_{\rm M}$ of 2.12 Å^3 Da^{-1} and a solvent content of 42% (Kantardjieff & Rupp, 2003; Matthews, 1968), assuming two PPD6 molecules per asymmetric unit. Data-processing statistics and crystal parameters are summarized in Table 1.

The presence of only a single methionine residue limits the possibility of using selenomethionine-labelled PPD6 protein for the collection of anomalous diffraction data. However, probably owing to bound Zn atoms, an anomalous signal with an anomalous correlation of >30% to 4 Å resolution was detected in the data, which can be used for Zn-SAD phasing. The determination of the PPD6 structure will improve our understanding of PPD6 function and the possible regulatory role of its intermolecular disulfide bond as well as the diverse functions of the PsbP-family proteins in general. Currently, work continues on finalizing the structure of PPD6 and performing structural analysis of the reduced-disulfide form of PPD6.

Table 1
Data-collection and processing statistics for a typical PPD6 crystal.

Values in parentheses are for the highest resolution shell.

No. of crystals	1
Beamline	I911-2, MAX-lab, Lund, Sweden
Detector	MAR 165 CCD
Crystal-to-detector distance (mm)	150
Wavelength (Å)	1.03796
Temperature (K)	100
Resolution range (Å)	28.55–2.10 (2.20–2.10)
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 47.0$, $b = 64.3$, $c = 62.1$, $\beta = 94.2$
No. of measured reflections	178417 (14094)
No. of unique reflections	21863 (2918)
Multiplicity	8.2 (4.8)
Completeness (%)	98.6 (90.8)
Mean $I/\sigma(I)$	30.1 (5.6)
$R_{\rm merge}$† (%)	3.6 (56.2)
$R_{\rm p.i.m.}$‡ (%)	1.4 (45.1)
Mosaicity (°)	0.33

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl) \times 100\%$. ‡ $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl) \times 100\%$, where $I_i(hkl)$ is the intensity of the *i*th observation of reflection *hkl*, $\langle I(hkl)\rangle$ is the weighted average intensity of all observations *i* of reflection *hkl* and *N* is the multiplicity.

This work was supported by the Swedish Research Council (to WPS; VR, 621-2008-3207), the Lars Hiertas Memorial Foundation (to MH) and the Carl Tryggers Stiftelse (to UHS; CTS 08:343).

References

Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Cain, P., Hall, M., Schröder, W. P., Kieselbach, T. & Robinson, C. (2009). *Plant Mol. Biol.* **70**, 273–281.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Hall, M., Mata-Cabana, A., Akerlund, H. E., Florencio, F. J., Schröder, W. P., Lindahl, M. & Kieselbach, T. (2010). *Proteomics*, **10**, 987–1001.
Ifuku, K., Nakatsu, T., Kato, H. & Sato, F. (2004). *EMBO Rep.* **5**, 362–367.
Ishihara, S., Takabayashi, A., Ido, K., Endo, T., Ifuku, K. & Sato, F. (2007). *Plant Physiol.* **145**, 668–679.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kieselbach, T., Bystedt, M., Hynds, P., Robinson, C. & Schröder, W. P. (2000). *FEBS Lett.* **480**, 271–276.
Lennartz, K., Plücken, H., Seidler, A., Westhoff, P., Bechtold, N. & Meierhoff, K. (2001). *Plant Cell*, **13**, 2539–2551.
Matthews, B. (1968). *J. Mol. Biol.* **28**, 491–497.
Michoux, F., Takasaka, K., Boehm, M., Nixon, P. J. & Murray, J. W. (2010). *Biochemistry*, **49**, 7411–7413.
Motohashi, K. & Hisabori, T. (2006). *J. Biol. Chem.* **281**, 35039–35047.
Peltier, J. B., Emanuelsson, O., Kalume, D. E., Ytterberg, J., Friso, G., Rudella, A., Liberles, D. A., Söderberg, L., Roepstorff, P., von Heijne, G. & van Wijk, K. J. (2002). *Plant Cell*, **14**, 211–236.
Sato, N. (2010). *Mol. Phylogenet. Evol.* **56**, 176–186.
Schubert, M., Petersson, U. A., Haas, B. J., Funk, C., Schröder, W. P. & Kieselbach, T. (2002). *J. Biol. Chem.* **277**, 8354–8365.
Seidler, A. (1996). *Biochim. Biophys. Acta*, **1277**, 35–60.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.