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

Expression, purification, crystallization and preliminary X-ray diffraction analysis of Deg5 from *Arabidopsis thaliana*

Haitian Fan,[a,b] Wei Sun,[a] Zhe Sun,[a] Feng Gao[a] and Weimin Gong[a]*

[a]Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, People's Republic of China, and [b]Graduate University of Chinese Academy of Sciences, Beijing 100039, People's Republic of China

Correspondence e-mail: wgong@ibp.ac.cn

Received 16 April 2012
Accepted 23 May 2012

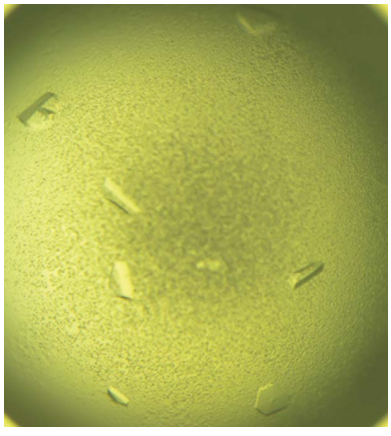
Arabidopsis thaliana Deg5 is an ATP-independent serine protease which resides on the luminal side of the thylakoid in chloroplasts. Deg5 and another Deg/HtrA-family protease, Deg8, have a synergistic function in the turnover of the D1 protein of photosystem II (PSII), which is prone to damage arising from high light exposure. An inactive mutant of the protein, Deg5^{S266A}, was overexpressed in *Escherichia coli*. After purification and crystallization, crystals that diffracted to 2.6 Å resolution were obtained. The crystals belonged to the monoclinic space group $C2$, with unit-cell parameters a = 109.1, b = 126.0, c = 83.3 Å, β = 102.9°, and contained three molecules in the asymmetric unit. The calculated Matthews coefficient and solvent content were 3.0 Å^3 Da^{-1} and 59.0%, respectively.

1. Introduction

High light exposure can lead to protein damage in photosynthetic organisms, and photosystem II (PSII) reaction centre protein D1 is the major target (Prasil *et al.*, 1992; Aro *et al.*, 1993). To maintain normal function of the electron-transfer chain, cells have evolved a quality-control system to prevent light-harvesting proteins from damage by repairing them or by degrading them and then synthesizing new copies; many chaperones and proteases are involved (Wickner *et al.*, 1999; Yamamoto, 2001; Maurizi, 2002). Deg/HtrA proteases are important participants among the various proteases in the chloroplast and are ATP-independent serine endopeptidases (Adam & Clarke, 2002; Clausen *et al.*, 2002). 16 types of Deg protease have been reported to be present in the chloroplast of *Arabidopsis thaliana* (Peltier *et al.*, 2002; Schubert *et al.*, 2002).

Deg proteases possess a chymotrypsin-like protease domain, which catalyzes the proteolysis of damaged proteins, and between zero and three PDZ domains, which mediate protein interactions (Pallen & Wren, 1997; Clausen *et al.*, 2002). In this work, Deg5 from *A. thaliana* was investigated, which consists of only one protease domain and no PDZ domains. Deg5 itself shows no proteolytic activity towards the modular substrate β-casein. However, it participates in the primary cleavage of the CD loop of PSII reaction centre protein D1 through forming a complex with another luminal protease, Deg8 (Sun *et al.*, 2007). Recently, Deg5 has also been found to be involved in the wounding-related disposal of PsbF apoprotein (Luciński *et al.*, 2011).

Although the functions of Deg5 have been characterized to some extent, knowledge of its structure is still lacking. Therefore, efforts were made to solve the crystal structure of Deg5. Here, we report the expression, purification, crystallization and preliminary crystallographic analysis of Deg5. This work paves the way for understanding the molecular mechanism of Deg5 and its involvement in photosynthesis.



2. Materials and methods

2.1. Cloning and expression

The plasmid encoding *A. thaliana* Deg5 (Ala73–Arg322; accession No. NP_567522) with its transit peptide deleted was a gift from Dr Lixin Zhang. The gene was inserted into vector pET28a (Novagen)

using *Nde*I and *Xho*I as double digestion sites with the additional N-terminal sequence MGSSHHHHHHSSGLVPRGSHMAS and the additional C-terminal sequence LSSVDKLAAALEHHHHHH. Because wild-type Deg5 is subject to degradation, the S266A mutant ($Deg5^{S266A}$) was constructed using a site-directed mutagenesis kit (Stratagene) with the primers 5′-CGATGCTGATATTAACTCAGGCAATGCCGGAGGGCCATTGC-3′ and 5′-GCAATGGCCCTCCGGCATTGCCTGAGTTAATATCAGCATCG-3′. The polymerase chain reaction (PCR) mixture (50 µl) consisted of 0.5 µ*M* forward and reverse primers, 0.2 m*M* dNTP, 50 ng plasmid and 1 U Phusion DNA polymerase (Finnzymes). PCR was performed in a thermocycler (Bio-Rad) using 5 min pre-incubation at 368 K followed by 16 cycles of 0.5 min at 368 K, 1 min at 328 K and 6 min at 345 K. The final extension time was 10 min at 345 K. After amplification, the sample was digested with 1 µl (10 U) *Dpn*I restriction endonuclease (New England Biolabs). Plasmids were transformed into *Escherichia coli* BL21 (DE3) expression strain. Bacteria were cultured at 310 K until the OD_{600} reached 0.8 and were then induced with 1 m*M* IPTG at 289 K for 20 h.

2.2. Protein purification

After induction, the cell pellets were spun down at 4000 rev min^{-1} (4670*g*) at 277 K for 30 min and resuspended in pre-chilled lysis buffer (20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* protease inhibitor PMSF) using 10 ml of buffer per 1 g of cells. The cells were homogenized by sonication and subsequently centrifuged at 20 000 rev min^{-1} (48 000*g*) at 277 K for 30 min. The supernatant was loaded onto a 5 ml pre-equilibrated Chelating Sepharose Fast Flow column (GE Healthcare) and washed with 20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl containing increasing concentrations of imidazole (0, 20, 50, 100 and 250 m*M*) at 277 K. The 250 m*M* imidazole elution was collected, concentrated to a small volume, loaded onto a 24 ml Superdex 200 10/300 GL column (GE Healthcare) and eluted using a buffer consisting of 20 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl at a rate of 0.5 ml min^{-1} at 277 K. For crystallization, the purified protein was concentrated to distinct concentrations (5, 10 and 15 mg ml^{-1}).

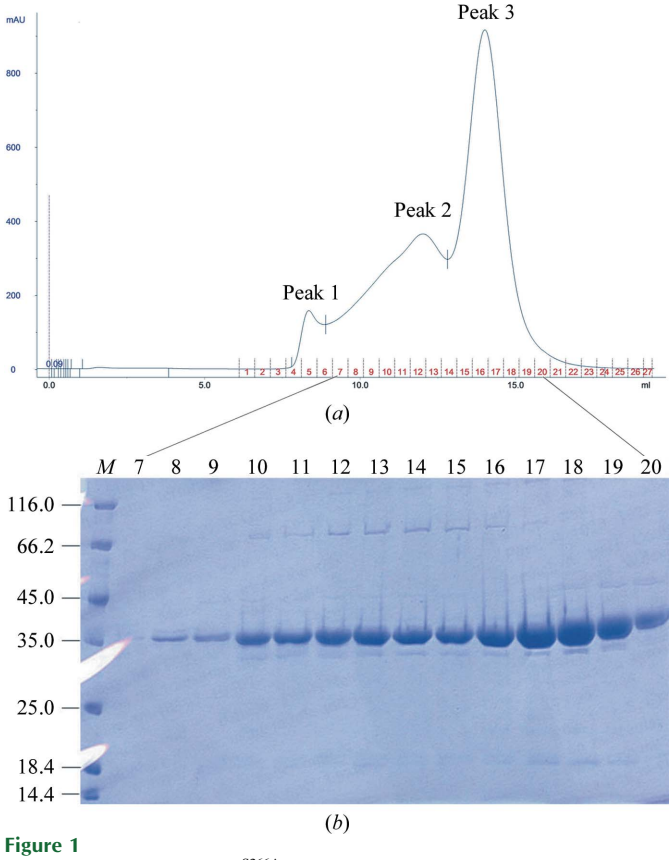


Figure 1
Purification of trimeric $Deg5^{S266A}$ by Superdex 200 10/300 GL FPLC. (*a*) Profiles from Superdex 200; peak 1 corresponds to the void volume and peaks 2 and 3 correspond to hexamers and trimers of $Deg5^{S266A}$, respectively. (*b*) SDS–PAGE verification of the samples from the size-exclusion column; fractions 7–20 are indicated. Lane *M*, molecular-mass marker (labelled in kDa).

2.3. Crystallization

Crystallization trials were set up by the sitting-drop vapour-diffusion method using the following kits: Crystal Screen, Crystal Screen 2, Index, PEG/Ion (Hampton Research) and Wizard I and II (Emerald BioSystems). Three drops consisting of 300 nl trimeric $Deg5^{S266A}$ (at 5, 10 and 15 mg ml^{-1}) and an equal volume of reservoir crystallization buffer were placed over one well of a sitting-drop plate (Swissci Triple Drop Plate UV Polymer). Subsequently, the plates were placed at 289 K for crystal growth. An initial hit was found using solution No. 14 of Crystal Screen [0.2 *M* calcium chloride dihydrate, 0.1 *M* Na HEPES pH 7.5, 28%(*v*/*v*) PEG 400].

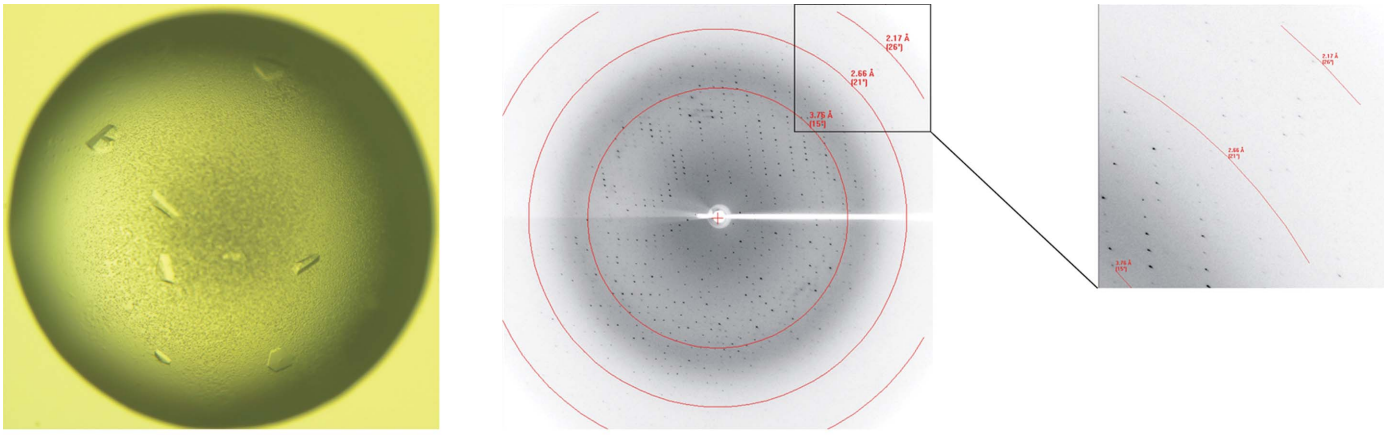
After the initial hit had been observed, we first reproduced the crystals by setting up a PEG 400 concentration gradient. A grid of different types and concentrations of precipitants was then set up to search for the optimal type and concentration of precipitant for crystal growth at a protein concentration of 10 mg ml^{-1}. The precipitants used were PEG 200, 400, 600, 1000, 1500, 3000, 3350, 4000, 6000, 8000, 10 000 and 20 000. Various concentrations of calcium chloride (0, 10, 20, 50, 100 and 200 m*M*) were also tried. The Na HEPES buffer was also changed to Tris–HCl buffer, leading to larger crystals. We used the hanging-drop vapour-diffusion method rather than the sitting-drop method during optimization.

2.4. Data collection and X-ray diffraction analysis

Data collection was performed on beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF) using a MAR225 CCD detector (MAR Research). Crystals were mounted on a nylon loop and flash-cooled in liquid nitrogen after soaking in reservoir solution supplemented with 20% glycerol as a cryoprotectant for 30 s. A data set was collected to a resolution of 2.6 Å. Diffraction data were then indexed, integrated and scaled employing the *HKL*-2000 software package (Otwinowski & Minor, 1997). Molecular replacement was carried out with *MOLREP* (Vagin & Teplyakov, 2010) from the *CCP*4 package (Winn *et al.*, 2011) using chain *A* (Asp24–Thr251) of *Thermotoga maritima* HtrA (PDB entry 1l1j; Kim *et al.*, 2003) as the search model, which has only one protease domain and shares 36% sequence identity with $Deg5^{S266A}$. The search model was modified by removing all water molecules and mutating all of the residues that differed from the $Deg5^{S266A}$ sequence to alanine before performing molecular replacement.

3. Results and discussion

A. thaliana $Deg5^{S266A}$ could be expressed in soluble form in *E. coli*. The molecular mass of the expressed $Deg5^{S266A}$ was approximately 31 kDa (Fig. 1*b*). After purification on a nickel column, most of the target protein could be eluted using a buffer containing 250 m*M* imidazole. $Deg5^{S266A}$ could form oligomers, which is consistent with other homologous Deg-family proteases (Krojer *et al.*, 2002; Kley *et al.*, 2011). According to the gel-filtration profile, peaks 2 and 3 correspond to approximate molecular masses of 190–220 and 90–



(*a*) (*b*)

Figure 2
Crystals and X-ray diffraction pattern of Deg5^{S266A}. (*a*) Typical crystals of Deg5^{S266A} from *A. thaliana*. (*b*) X-ray diffraction pattern of Deg5^{S266A}.

Table 1
X-ray diffraction data-collection and processing statistics for Deg5^{S266A}.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters (Å, °)	a = 109.1, b = 126.0, c = 83.3, β = 102.9
Wavelength (Å)	0.9792
X-ray beam dimensions (µm)	<130 × 40
Oscillation angle (°)	1
Space group	*C*2
Resolution (Å)	50.00–2.60 (2.69–2.60)
No. of reflections	33072 (3206)
Unique reflections	11024 (1106)
Completeness (%)	99.2 (97.8)
Average $I/\sigma(I)$	10.9 (2.3)
R_{merge}	0.094 (0.465)
Multiplicity	3.0 (2.9)

95 kDa, respectively, based on a previously determined calibration curve (Fig. 1*a*). Thus, a trimer is the dominant state and there is also a small amount of hexamer.

In the screening round, crystals were observed from wells corresponding to condition No. 14 of Crystal Screen [0.2 *M* calcium chloride dihydrate, 0.1 *M* Na HEPES pH 7.5, 28%(*v*/*v*) PEG 400] after one week of growth. Compared with the other two concentrations, protein at 10 mg ml^{-1} gave rise to larger crystals. Varying the type of precipitant and the concentration of calcium chloride did not have an obvious effect. However, replacing the Na HEPES pH 7.5 buffer by Tris–HCl pH 7.5 yielded larger crystals. In addition, increasing the drop volume and varying the ratio of protein to reservoir solution helped to obtained larger crystals (Fig. 2*a*) with dimensions of approximately 0.2 × 0.1 × 0.1 mm.

Finally, a crystal grown from conditions consisting of 0.2 *M* calcium chloride dihydrate, 0.1 *M* Tris–HCl pH 7.5, 25%(*v*/*v*) PEG 400 diffracted to 2.6 Å resolution at SSRF (Fig. 2*b*). The Deg5^{S266A} crystal belonged to the monoclinic space group *C*2, with unit-cell parameters a = 109.1, b = 126.0, c = 83.3 Å, β = 102.9°. Data-collection statistics are listed in Table 1. The molecular-replacement solution showed that the Deg5^{S266A} crystals contained a trimer of Deg5^{S266A} in the asymmetric unit. The corresponding Matthews coefficient was calculated to be 3.0 Å^3 Da^{-1} and the solvent content was 59.0% (Matthews, 1968). The molecular packing appeared to be reasonable. After preliminary refinement by *REFMAC* (Murshudov *et al.*, 2011) without model rebuilding R_{work} and R_{free} decreased to 38.4% and 45.4%, respectively, suggesting that the molecular replacement was successful.

We greatly thank Dr Lixin Zhang for providing the recombinant Deg5 plasmid. We also thank the staff at Shanghai Synchrotron Radiation Facility (SSRF) for diffraction data collection. This work was supported by the Ministry of Science and Technology (Grant No. 2011CB910503) and the National Natural Science Foundation of China (Grant No. 30721003).

References

Adam, Z. & Clarke, A. K. (2002). *Trends Plant Sci.* **7**, 451–456.
Aro, E. M., Virgin, I. & Andersson, B. (1993). *Biochim. Biophys. Acta*, **1143**, 113–134.
Clausen, T., Southan, C. & Ehrmann, M. (2002). *Mol. Cell*, **10**, 443–455.
Kim, D. Y., Kim, D. R., Ha, S. C., Lokanath, N. K., Lee, C. J., Hwang, H.-Y. & Kim, K. K. (2003). *J. Biol. Chem.* **278**, 6543–6551.
Kley, J., Schmidt, B., Boyanov, B., Stolt-Bergner, P. C., Kirk, R., Ehrmann, M., Knopf, R. R., Naveh, L., Adam, Z. & Clausen, T. (2011). *Nature Struct. Mol. Biol.* **18**, 728–731.
Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M. & Clausen, T. (2002). *Nature (London)*, **416**, 455–459.
Luciński, R., Misztal, L., Samardakiewicz, S. & Jackowski, G. (2011). *Plant Physiol. Biochem.* **49**, 311–320.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Maurizi, M. R. (2002). *Nature Struct. Biol.* **9**, 410–412.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pallen, M. J. & Wren, B. W. (1997). *Mol. Microbiol.* **26**, 209–221.
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.
Prasil, O., Adir, N. & Ohad, I. (1992). *The Photosystems: Structure, Function and Molecular Biology*, edited by J. Barber, pp. 295–348. Amsterdam: Elsevier.
Schubert, M., Petersson, U. A., Haas, B. J., Funk, C., Schröder, W. P. & Kieselbach, T. (2002). *J. Biol. Chem.* **277**, 8354–8365.
Sun, X., Peng, L., Guo, J., Chi, W., Ma, J., Lu, C. & Zhang, L. (2007). *Plant Cell*, **19**, 1347–1361.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Wickner, S., Maurizi, M. R. & Gottesman, S. (1999). *Science*, **286**, 1888–1893.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yamamoto, Y. (2001). *Plant Cell Physiol.* **42**, 121–128.