Purification, identification and preliminary crystallographic studies of an allergenic protein from *Solanum melongena*

Abha Jain^{a,b} and Dinakar Masanu Salunke^{a,c*}

Received 29 October 2014

Accepted 13 January 2015

Keywords: Solanaceae; *Solanum melongena*; allergy.

^aRegional Centre for Biotechnology, 180 Udyog Vihar Phase 1, Gurgaon, Haryana 122 016, India, ^bManipal University, Madhav Nagar, Manipal, Karnataka 576 104, India, and ^cNational Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India. *Correspondence e-mail: dmsalunke@rcb.res.in

Solanum melongena (eggplant), a member of the Solanaceae family, is a widely cultivated vegetable crop and is commonly used as a food throughout the world. Allergic reactions caused by members of this family are well known. However, mechanistic analyses to understand their molecular basis have not been adequately explored. In order to address this issue, the 7S vicilin protein (SM80.1) of size 45 kDa was purified from seeds of *S. melongena* by ammonium sulfate fractionation and size-exclusion chromatography. Significant homology of SM80.1 to an allergy-related protein from *S. lycopersicum* was identified through a *BLAST* search. Crystallization attempts with purified protein using the hanging-drop vapour-diffusion method led to hexagonal-shaped crystals. The crystals diffracted to 2.21 Å resolution and belonged to space group *P*6₃22, with unit-cell parameters *a* = 117.9, *c* = 123.5 Å.

1. Introduction

The Solanaceae family is considered to be the third most important plant family and comprises many important plant genera such as *Datura*, *Capsicum*, *Solanum* and *Nicotiana* (Rigano *et al.*, 2013; Gerlach *et al.*, 2009; Rubatzky & Yamaguchi, 1997). Members of this family are widely utilized as important sources of food and medicine. For instance, *S. surattense* (Berkateli, Kantakari) has been reported to exhibit antimicrobial (Liu *et al.*, 2006), antifungal (Sheeba & Palanivel, 2013), antioxidant (Muruhan *et al.*, 2013), anti-diabetic (Gupta *et al.*, 2011) and antimalarial (Ramazani *et al.*, 2010) properties. Certain species of this family are rich in alkaloids and are therefore toxic in nature. Their toxicity varies from mild to acute. Pharmaceutically important proteins are also produced from edible Solanaceae plants (Rigano *et al.*, 2013).

Certain proteins from the Solanaceae family, such as patatin from raw potato (Schmidt *et al.*, 2002; Seppälä *et al.*, 2001) and Lyc e 1, Lyc e 2 and Lyc e 3 from tomato (Foetisch *et al.*, 2003; Westphal *et al.*, 2003; Willerroider *et al.*, 2003), have been reported to cause allergies. It has also been suggested that three proteins in the molecular mass range 60–71 kDa from eggplant cause allergies (Pramod & Venkatesh, 2008). It is observed that the major players responsible for causing allergies are storage proteins that are abundantly present in plant seeds. These seed storage proteins have been catalogued into different superfamilies and families (Breiteneder &

© 2015 International Union of Crystallography

Radauer, 2004). For example, the cupin superfamily includes the globulin seed storage proteins and can be divided into 7S vicilins and 11S legumins.

Understanding the allergenic properties of *S. melongena* seeds at the molecular level would be facilitated by comparative structural proteomics of these seeds using allergy as a functional screen. Such studies have also been initiated involving other seed proteomes (Chanana *et al.*, 2004; Gaur *et al.*, 2008, 2010; Gupta *et al.*, 2008; Qureshi *et al.*, 2006). Here, we report the identification, purification and crystallization of a 45 kDa protein from *S. melongena* which may have allergenic properties.

2. Materials and methods

2.1. Macromolecule production

2.1.1. Protein identification. Seeds of *S. melongena* (eggplant) were purchased from the Indian Agricultural Research Institute (IARI), New Delhi, India. 50 g of the seeds were washed with alcohol. The washed seeds were ground and delipidified using petroleum ether. The delipidated powder was homogenized with 50 mM Tris-HCl pH 8 containing 150 mM NaCl in the presence of a protease-inhibitor cocktail (Sigma) for 4 h. The aqueous extract thus obtained was transferred to 1 l polypropylene bottles, which were placed

into a JLA-8.1000 fixed-angle rotor and centrifuged at 7000 rev min⁻¹ (12 300g) using a Beckman Avanti J-26 XPI High Performance Centrifuge for 1 h at 277 K. The supernatant containing soluble proteins was collected by decantation and subjected to ammonium sulfate fractionation in the range 0–95%(w/v). The pellets obtained were dissolved in homogenization buffer and then dialyzed to remove salt. 15% SDS–PAGE identified segregated proteins in each fraction.

2.1.2. N-terminal protein sequencing. The proteins present in the different fractions were transferred onto polyvinylidene difluoride (PVDF) membrane and stained with Ponceau S dye. The relevant bands extracted from the membrane were used for N-terminal sequencing. Edman degradation was utilized to obtain the sequence on a PPSQ-33A protein sequencer (Shimadzu). The sequence was then used in the *BLAST* search engine (Altschul *et al.*, 1997) to identify possible allergenic protein(s) based on homology to proteins from the nonredundant (NR) protein database.

2.1.3. Purification. Based on the above strategy, the protein was selected and subjected to size-exclusion chromatography for purification. A manually packed Sephacryl S-200 column pre-equilibrated with homogenization buffer was used for purification. The column was connected to an ÄKTA Explorer FPLC system (GE Healthcare) and run in isocratic mode. The purified protein was then concentrated. The concentration was determined using the BCA protein concentration assay.

Figure 1
(a) SDS–PAGE profile showing the composition of the fractions obtained by ammonium sulfate fractionation of the aqueous seed extract. Lanes 1–5 correspond to 20, 40, 60, 80 and 95% ammonium sulfate fractions, respectively. Lane M contains low-molecular-weight marker (Sigma; labelled in kDa). (b) Alignment of the N-terminal sequence obtained for SM80.1 by Edman degradation with those of other homologous vicilins from the Solanaceae family (*S. lycopersicum*, *S. tuberosum* and *N. tomentosiformis*). Similar residues are marked in grey.

2.2. Crystallization

The purified protein was subjected to different crystallization conditions. Trials were set up with various commercially available screens from Hampton Research, including PEG/Ion, PEG/Ion 2, Crystal Screen, Crystal Screen 2 and Index. Protein concentrations ranging from 5 to 20 mg ml⁻¹ were explored along with a pH gradient from 7 to 8 in Tris-HCl buffer. A few hits were obtained in the Index screen. In order to improve the crystal quality, polyethylene glycols with diverse molecular weights from 3350 to 8000 Da were explored.

2.3. Data collection and processing

Diffraction data were collected from crystals on beamline BM14 at the European Synchrotron Radiation Facility at a wavelength of 0.97625 Å using a MAR CCD detector. No separate cryoprotectant was used during data collection, as sodium malonate, which was part of the crystallization buffer, is known to act as a cryoprotectant. The data were collected at 100 K with a crystal-to-detector distance of 232.11 mm and an oscillation range of 1°. Data were collected from a single crystal and were integrated and scaled with *iMosflm* (Battye *et al.*, 2011).

3. Results and discussion

3.1. Purification and molecular characterization

Delipidified seed powder was used to extract the proteins in homogenization buffer. Ammonium sulfate fractionation in the range 0–95% (w/v) resulted in the precipitation of different proteins based on their differential solubilities. Analyses of these fractions were performed by 15% SDS-PAGE along with molecular-weight marker (Fig. 1a).

Table 1

Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	24-well
Temperature (K)	298
Protein concentration (mg ml ⁻¹)	10
Buffer composition of protein solution	50 mM Tris-HCl pH 8
Composition of reservoir solution	1.5 M sodium malonate, 10% (w/v) PEG 3350
Volume and ratio of drop	2 + 2 µl
Volume of reservoir (µl)	500

A protein band from the 80% ammonium sulfate fraction corresponding to a molecular weight of 45 kDa was transferred onto PVDF membrane and N-terminal amino-acid sequencing was subsequently carried out to identify the protein (Fig. 1b). A *BLAST* search for the N-terminal 30-amino-acid sequence of the protein was carried out. The search revealed sequence homology of this protein to the 7S vicilin from *S. lycopersicum*, which is known to be an allergenic protein (Bässler *et al.*, 2009). It also showed homology to other vicilins from the Solanaceae family. However, no crystal structures of Solanaceae family 7S vicilins are known.

After dialysis, the 80% fraction was loaded onto a pre-equilibrated manually packed Sephacryl S-200 column for purification. Peak 1 in the chromatogram corresponded to a highly purified protein of molecular weight 45 kDa, as shown by 15% SDS-PAGE (Fig. 2). The purified protein was named SM80.1 and classified as a 7S vicilin from *S. melongena*.

3.2. Crystallization and characterization of protein crystals

Protein solution was prepared in 50 mM Tris-HCl buffer pH 8.0 for crystallization. After several attempts, diffraction-quality crystals were obtained by the vapour-diffusion method in drops composed of equal volumes (2 µl) of protein solution and reservoir solution equilibrated against 500 µl reservoir

Figure 2

(a) Size-exclusion chromatography of the 80% ammonium sulfate fraction on a Sephacryl S-200 column. The peak for purified SM80.1 is marked with an arrow. Lanes 1 and M in the SDS-PAGE profile (b) correspond to purified protein and molecular-weight marker (labelled in kDa), respectively.

Table 2

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BM14, ESRF
Wavelength (Å)	0.97625
Temperature (K)	100
Detector	MAR CCD
Crystal-to-detector distance (mm)	232.11
Rotation range per image (°)	1
Total rotation range (°)	72
Exposure time per image (s)	25.6
Space group	<i>P</i> 6 ₃ 22
<i>a</i> , <i>b</i> , <i>c</i> (Å)	117.9, 117.9, 123.5
α , β , γ (°)	90, 90, 120
Mosaicity (°)	0.57
Resolution range (Å)	42–2.21 (2.33–2.21)
Total No. of reflections	219350 (30488)
No. of unique reflections	25997 (3721)
Completeness (%)	100 (100)
Multiplicity	8.4 (8.2)
$\langle I/\sigma(I) \rangle$	20.6 (3.6)
$R_{\text{merge}}^{\dagger}$ (%)	6.5 (54.6)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th measurement of the intensity of reflection *hkl* and $\langle I(hkl) \rangle$ is the mean intensity of reflection *hkl*.

(a)

(b)

Figure 3

Hexagonal-shaped crystals of SM80.1 from *S. melongena* (a) and their diffraction pattern (b).

solution (Table 1). A 10 mg ml^{−1} protein concentration with 10% polyethylene glycol (3350 Da) and 1.5 M sodium malonate resulted in hexagonal-shaped crystals (Fig. 3) which belonged to space group *P*6₃22, with unit-cell parameters *a* = 117.9, *c* = 123.5 Å. Data were collected to 2.21 Å resolution, which gave approximately 100% completeness and a multiplicity of 8.4 with an R_{merge} of 6.5%. The data-collection and processing statistics are presented in Table 2. Calculation of the Matthews coefficient suggested the presence of one monomer per asymmetric unit with 53.73% solvent content. A BLAST search of the N-terminal 30-amino-acid sequence of the purified protein against the protein sequences in the PDB identified the structures of vicilin from Korean pine (*Pinus koraiensis*), almond (*Prunus dulcis*) and jack bean (*Canavalin ensiformis*) with relatively weak homology. It was thus recognized that the structure will need to be determined either by molecular replacement or by *ab initio* methods.

In summary, SM80.1 from *S. melongena* has been isolated, purified and crystallized and subjected to preliminary X-ray characterization. The protein shows homology to 7S allergenic vicilins. As SM80.1 shows weak homology to 7S vicilins for which three-dimensional structures are known, both the molecular-replacement and the *ab initio* approaches will be explored for structure determination, which will provide further insights into the correlation of protein structure with allergenic properties.

Acknowledgements

We thank the Department of Biotechnology, Government of India for financial support. We also thank the EMBL staff, Dr Hassan Belrhali and Dr Babu A. Manjasetty for providing support on the beamline and EMBL-DBT for providing access to the BM14 beamline at the ESRF.

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
- Bässler, O. Y., Weiss, J., Wienkoop, S., Lehmann, K., Scheler, C., Dölle, S., Schwarz, D., Franken, P., George, E., Worm, M. & Weckwerth, W. (2009). *J. Proteome Res.* **8**, 1111–1122.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Breiteneder, H. & Radauer, C. (2004). *J. Allergy Clin. Immunol.* **113**, 821–830.
- Chanana, V., Kaur, K. J. & Salunke, D. M. (2004). *Acta Cryst.* **D60**, 2100–2103.
- Foetisch, K., Westphal, S., Lauer, I., Retzek, M., Altmann, F., Kolarich, D., Scheurer, S. & Vieths, S. (2003). *J. Allergy Clin. Immunol.* **111**, 889–896.
- Gaur, V., Qureshi, I. A., Singh, A., Chanana, V. & Salunke, D. M. (2010). *Plant Physiol.* **152**, 1842–1850.
- Gaur, V., Sethi, D. K. & Salunke, D. M. (2008). *Acta Cryst.* **F64**, 32–35.
- Gerlach, J. Q., Bhavanandan, P. V., Haynes, A. P. & Joshi, L. (2009). *J. Bot.* **2009**, 560394.
- Gupta, P., Gaur, V. & Salunke, D. M. (2008). *Acta Cryst.* **F64**, 733–736.
- Gupta, R., Sharma, A. K., Dobhal, M. P., Sharma, M. C. & Gupta, R. S. (2011). *J. Diabetes*, **3**, 29–37.
- Liu, X., Huang, B., Lin, J., Fei, J., Chen, Z., Pang, Y., Sun, X. & Tang, K. (2006). *J. Plant Physiol.* **163**, 546–556.

- Muruhan, S., Selvaraj, S. & Viswanathan, P. K. (2013). *Asian Pac. J. Trop. Biomed.* **3**, 28–34.
- Pramod, S. N. & Venkatesh, Y. P. (2008). *J. Investig. Allergol. Clin. Immunol.* **18**, 59–62.
- Qureshi, I. A., Sethi, D. K. & Salunke, D. M. (2006). *Acta Cryst.* **F62**, 869–872.
- Ramazani, A., Zakeri, S., Sardari, S., Khodakarim, N. & Djadid, N. D. (2010). *Malar. J.* **9**, 124.
- Rigano, M. M., De Guzman, G., Walmsley, A. M., Frusciante, L. & Barone, A. (2013). *Int. J. Mol. Sci.* **14**, 2753–2773.
- Rubatzky, V. E. & Yamaguchi, M. (1997). *World Vegetables: Principles, Production and Nutritive Values*, pp. 532–576. London: Chapman & Hall.
- Schmidt, M. H., Raulf-Heimsoth, M. & Posch, A. (2002). *Ann. Allergy Asthma Immunol.* **89**, 613–618.
- Seppälä, U., Majamaa, H., Turjanmaa, K., Helin, J., Reunala, T., Kalkkinen, N. & Palosuo, T. (2001). *Allergy*, **56**, 619–626.
- Sheeba, E. & Palanivel, S. (2013). *Middle East. J. Sci. Res.* **16**, 122–126.
- Westphal, S., Kolarich, D., Foetisch, K., Lauer, I., Altmann, F., Conti, A., Crespo, J. F., Rodríguez, J., Enrique, E., Vieths, S. & Scheurer, S. (2003). *Eur. J. Biochem.* **270**, 1327–1337.
- Willeroider, M., Fuchs, H., Ballmer-Weber, B. K., Focke, M., Susani, M., Thalhamer, J., Ferreira, F., Wüthrich, B., Scheiner, O., Breiteneder, H. & Hoffmann-Sommergruber, K. (2003). *Int. Arch. Allergy Immunol.* **131**, 245–255.