Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary crystallographic studies of the ligand-binding domain of a plant vacuolar sorting receptor

Sally W. Rogers,[a]‡ Buhyum Youn,[b]‡ John C. Rogers[a] and ChulHee Kang[b]*

[a]Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340, USA, and [b]School of Molecular Biosciences, Washington State University, Pullman, WA 99164-4660, USA

‡ These authors contributed equally to this article.

Correspondence e-mail: chkang@wsunix.wsu.edu

Received 30 June 2004
Accepted 26 August 2004

Vacuolar sorting receptor (VSR) proteins bind soluble protein ligands in a sequence-specific manner and target them to the lytic vacuole in plant cells. A VSR from *Arabidopsis thaliana*, AtBP80b, has been successfully purified after heterologous expression in *Drosophila* S2 cells. The AtBP80b protein (560 amino acids) was crystallized by the hanging-drop method with a PEG 400-based precipitant. Preliminary X-ray diffraction studies of an AtBP80b crystal showed that it belongs to the cubic space group $P2_13$ (or $P4_232$) and has unit-cell parameters $a = b = c = 145.9$ Å. Crystals of the VSR diffract beyond 2.5 Å resolution.

1. Introduction

In eukaryotic cells, soluble proteins enter the secretory pathway in the endoplasmic reticulum and then move to the Golgi complex. If they have no specific sorting information on them, they exit the Golgi and are secreted to the cell exterior. Therefore, sorting away from this 'bulk flow' (Denecke *et al.*, 1990; Phillipson *et al.*, 2001) requires both that the soluble protein carries a signal and that an integral membrane-receptor protein recognizes the signal and directs the soluble protein into a vesicle that will carry it to a different destination. In plant cells, sorting mechanisms are complex because proteins are sorted into pathways to two separate vacuoles (Okita & Rogers, 1996; Vitale & Galili, 2001). Within the plant Golgi apparatus, one receptor, termed BP80 (Hadlington & Denecke, 2000; Kirsch *et al.*, 1994), recognizes one type of sorting determinant and targets a soluble protein to the lytic vacuole pathway (Humair *et al.*, 2001; Jiang & Rogers, 1998; Paris *et al.*, 1997), while a different receptor recognizes a different sorting determinant and pulls a different soluble protein into the protein storage vacuole pathway (Okita & Rogers, 1996; Vitale & Galili, 2001).

Both types of sorting determinants are encoded by the polypeptide sequence, but differ in composition and position within the polypeptide (Matsuoka & Neuhaus, 1999). The structural details by which each receptor interacts with the correct sorting determinant is a central question in plant cell biology. Sorting determinants for the lytic vacuole pathway are sequence-specific and the best studied such determinants share a central Asn-Pro-Ile-Arg (NPIR) motif (Matsuoka & Neuhaus, 1999). This knowledge led to the biochemical purification of an ~80 kDa protein, termed BP80, from lysates of pea clathrin-coated vesicle membranes by means of an affinity column carrying the peptide SSSFADS**NPIR**PVTDRAASTYC (Kirsch *et al.*, 1994). This peptide, termed the proaleurain peptide, contains the vacuolar sorting determinants for a plant cysteine protease (Holwerda *et al.*, 1992). BP80 binding to the proaleurain peptide occurred with a K_d of approximately 40 nm and was pH-dependent, with optimal binding between pH 6.0 and 6.5, while binding was abolished at pH $\leq$ 4.5. Additionally, binding was sequence-specific: the peptide SRF**NPIR**LPT efficiently competed for binding, while the peptide SRFNPGRLPT, with an Ile-to-Gly mutation in the NPIR motif, did not compete (Kirsch *et al.*, 1994).

Further studies of BP80–proaleurain peptide interactions utilized a recombinant form of the BP80 lumenal portion (termed tBP80, lacking the transmembrane domain and cytoplasmic tail) that was expressed in *Drosophila* S2 cells, from which it was secreted and could be purified from the medium (Cao *et al.*, 2000). This tBP80 is comprised of three distinct domains: an N-terminal 'PA domain' (Cao *et al.*, 2000; Luo & Hofmann, 2001; Mahon & Bateman, 2000), a central region without defined homology to other proteins and a C-terminal domain comprised of three epidermal growth factor (EGF) repeats. Expression in a eukaryotic cell was required because the tBP80 requires chaperone-mediated folding in the endoplasmic reticulum to acquire a proper three-dimensional structure and to ensure correct placement of disulfide bonds among its 34 Cys residues. These results (Cao *et al.*, 2000) indicated that tBP80 had two separate proaleurain peptide ligand-binding sites, one NPIR-sequence-specific and the other not specific for NPIR.

Results from binding studies using a recombinant protein lacking the EGF-repeat domain indicated that NPIR-specific binding was defined by interactions of the PA domain and the central unique domain. The EGF repeats appeared not to participate in NPIR-specific binding, but rather affected the conformation of the other two domains in a favorable manner. Additionally, the EGF repeats participated in Ca^{2+}-dependent non-NPIR-specific binding. Limited proteolysis experiments demonstrated that a site between the PA and unique domains and a site between the unique and EGF-repeat domains were very protease-sensitive. Once one of the two sites was cut, the other became relatively protease resistant.

We wanted to determine the crystal structure of tBP80 as a first step towards understanding its ability to bind ligands in a specific manner. Because chromatographic purification in a previous study resulted in large losses and a low yield (Cao *et al.*, 2000), we chose to express a form with six His residues present at the C-terminus in order to facilitate its purification. Probably for technical reasons, expression of a tBP80 homolog from *Arabidopsis* termed truncated AtBP80b (Hadlington & Denecke, 2000) with a C-terminal His tag was successful, while we were not able to obtain a His-tagged tBP80. We therefore have expressed truncated AtBP80b and purified it to homogeneity in two steps. Here, we report its crystallization and initial diffraction studies.

2. Experimental procedures

2.1. Expression of truncated AtBP80b in *Drosophila* S2 cells

The lumenal domain of plasmid Z38123 (Paris *et al.*, 1997) encoding the BP80 homolog AtBP80b was amplified by PCR using primers 5′-CCCTCCGGAGCCAACTTTGCCTGAAC and 5′-GGGGGTACCGGAATCATGAAGCTTGG. This resulted in the coding sequence for the lumenal domain, comprised of 560 amino acids bracketed by *Kpn*I and *Bsp*EI restriction sites. The DNA sequence of the PCR product was determined to ensure its fidelity. The PCR product was then digested with *Kpn*I and *Bsp*EI and inserted into that interval in plasmid pMT/V5 (Invitrogen, Carlsbad, CA, USA) to give an in-frame protein fusion at its C-terminus with the plasmid-encoded sequence LESRGPFEGKPIPNPLLGLDSTGTGHHHHHH. Methods for the culture of *Drosophila* S2 cells, DNA transfection, selection of transformants and expression of recombinant protein have been described elsewhere (Cao *et al.*, 2000).

2.2. Purification of truncated AtBP80b

When cells reached a density of 1×10^7 ml^{-1}, recombinant protein expression was induced by the addition of copper sulfate to a final concentration of 500 µ*M*. 3 d after induction, the medium was separated from the cells by centrifugation for 10 min at 1000*g*. The medium was dialyzed extensively at 277 K against 0.1 *M* Na_3PO_4, 0.01 *M* Tris pH 8 and then passed over a column of Ni^{2+}–nitriloacetic acid agarose (Qiagen, Valencia, CA, USA) to select protein with the C-terminal His tag. After washing the column with the same buffer until $A_{280} \leq 0.01$, protein was eluted in 8–10 1 ml fractions with 0.1 *M* imidazole pH 7.7. Truncated AtBP80b-containing fractions were pooled and dialyzed overnight at 277 K against repeated changes of 20 m*M* MES, 50 m*M* NaCl pH 6.3. This material was then passed over a proaleurain peptide-affinity column (Kirsch *et al.*, 1994). The column was washed with the same buffer until $A_{280} \leq 0.01$. Bound protein was eluted in five 1 ml fractions of 0.1 *M* sodium acetate pH 4.0, which were then brought to neutrality with 1 *M* Tris base. Confirmation of the presence and purity of the protein was made by SDS–PAGE and immunoblot analysis using monoclonal antibody 17F9 (Cao *et al.*, 2000). The pooled fractions were then dialyzed against 20 m*M* Tris–HCl pH 7.4 at 277 K overnight.

2.3. Crystallization

The purified protein was concentrated to 12.0 mg ml^{-1} in 20 m*M* Tris buffer pH 7.5 using a YM30 membrane (Amicon). Crystallization trials were performed using the hanging-drop vapor-diffusion method at 277 K using 24-well VDX plates (Hampton Research). Each hanging drop was prepared by mixing equal volumes (2 µl each) of protein solution and reservoir solution, followed by placement over 0.5 ml reservoir solution. Initial crystallization conditions were established using the screening kits from Hampton Research Inc. (Crystal Screens I and II and PEG/Ion Screen) and Emerald BioStructures Inc. (Wizard I and II).

2.4. X-ray data collection and processing

The VSR crystal was flash-vitrified using the same reservoir solution as cryoprotectant. A crystal was picked up in a CryoLoop (0.5 mm diameter; Hampton Research) and immediately frozen in a nitrogen stream at 100 K. Diffraction data consisting of 200 images with 1° oscillation and extending to 2.5 Å were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 7-1 (MAR 345 image-plate detector) using the BLUE-ICE interface. Indexing, integration of images and scaling of the intensity data were performed using the programs *DENZO* and *SCALEPACK* from the *HKL* package (Otwinowski & Minor, 1997). The native crystal data extends to 2.6 Å with a completeness of 82%. The $R_{\rm merge}$ and $\langle I/\delta(I)\rangle$ values are 0.04 and 12.8, respectively, as summarized in Table 1 together with the data statistics.

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shell (2.71–2.66 Å).

X-ray wavelength (Å)	0.97 (SSRL BL 7-1)
Space group	$P2_13$ (or $P4_232$)
Unit-cell parameters (Å)	$a = b = c = 145.9$
Resolution (Å)	50.00–2.66
No. measured reflections	186569
No. unique reflections	24159
$R_{\rm merge}$† (%)	4.3 (66.1)
Average redundancy	7.87 (9.21)
Data completeness (%)	81.9 (24.1)
$\langle I/\sigma(I)\rangle$	12.8 (2.3)

† $R_{\rm merge} = \sum_h \sum_i |I(h)_i - \langle I\rangle| / \sum_h \sum_i I(h)_i$, where $I(h)$ is the intensity of reflection h, $\sum_h$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection h.

3. Results and discussion

The *Drosophila* S2 expression system and our purification protocol allowed us to obtain pure AtBP80b protein with yields of approximately 1–2 mg of homogeneous

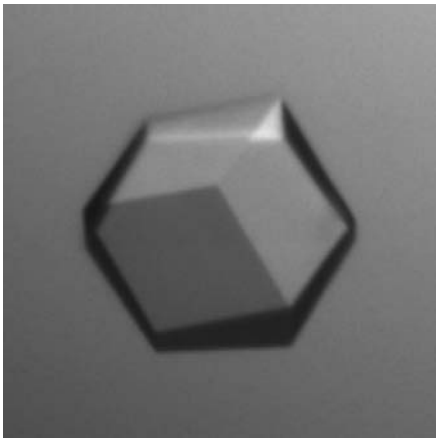

Figure 1
Picture of an AtBP80b crystal. This crystal was grown at 277 K. After two weeks of growth, the crystal grew to dimensions of 0.3 × 0.3 × 0.3 mm.

AtBP80 protein per litre of S2 cell-culture medium. Initially, small cube-shaped crystals were obtained from two different screening solutions: (i) 0.2 *M* trisodium citrate dehydrate, 0.1 *M* Tris–HCl pH 8.5, 30%(*v*/*v*) PEG 400 and (ii) 0.05 *M* cesium chloride, 0.1 *M* MES pH 6.5, 30%(*v*/*v*) Jeffamine M-600. These initial conditions were refined to optimize the crystallization by varying the concentration of the ingredients and the pH. The best crystals were obtained by mixing 2.0 µl protein solution (6 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.5) with an equal volume of reservoir solution containing 0.15 *M* tricitrate dehydrate, 0.1 *M* Tris–HCl pH 8.0, 30%(*v*/*v*) PEG 400. Under these optimized conditions, the crystals grew to dimensions of 0.3 × 0.3 × 0.3 mm after two weeks (Fig. 1) and the crystals belong to the cubic space group $P2_13$ (or $P4_232$), with unit-cell parameters $a = b = c = 145.9$ Å. The calculated V_M (Matthews coefficient) is 2.13 Å^3 Da^{-1}, with a solvent content of 42.14% (Matthews, 1968; Kantardjieff & Rupp, 2003), which corresponds to one or two molecules in the asymmetric unit with respect to the choice of space group.

In order to overcome the crystallographic phase problem, multiple-wavelength anomalous diffraction and multiple isomorphous replacement methods are being applied, since the molecular-replacement approach is not currently possible.

This research was supported by USDA grant No. 2003-35318-13672 to JR and CK. We thank Dr Nadine Paris for providing the AtBP80b expression construct.

References

Cao, X., Rogers, S. W., Butler, J., Beevers, L. & Rogers, J. C. (2000). *Plant Cell*, **12**, 493–506.
Denecke, J., Botterman, J. & Deblaere, R. (1990). *Plant Cell*, **2**, 51–59.
Hadlington, J. L. & Denecke, J. (2000). *Curr. Opin. Plant Biol.* **3**, 461–469.
Holwerda, B. C., Padgett, H. S. & Rogers, J. C. (1992). *Plant Cell*, **4**, 307–318.
Humair, D., Hernández Felipe, D., Neuhaus, J.-M. & Paris, N. (2001). *Plant Cell*, **13**, 781–792.
Jiang, L. & Rogers, J. C. (1998). *J. Cell Biol.* **143**, 1183–1199.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kirsch, T., Paris, N., Butler, J. M., Beevers, L. & Rogers, J. C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 3403–3407.
Luo, X. & Hofmann, K. (2001). *Trends Biochem. Sci.* **26**, 147–148.
Mahon, P. & Bateman, A. (2000). *Protein Sci.* **9**, 1930–1934.
Matsuoka, K. & Neuhaus, J.-M. (1999). *J. Exp. Bot.* **50**, 165–174.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Okita, T. W. & Rogers, J. C. (1996). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 327–350.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Paris, N., Rogers, S. W., Jiang, L., Kirsch, T., Beevers, L., Phillips, T. E. & Rogers, J. C. (1997). *Plant Physiol.* **115**, 29–39.
Phillipson, B. A., Pimpl, P., daSilva, L. L., Crofts, A. J., Taylor, J. P., Movafeghi, A., Robinson, D. G. & Denecke, J. (2001). *Plant Cell*, **13**, 2005–2020.
Vitale, A. & Galili, G. (2001). *Plant Physiol.* **125**, 115–118.