Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction studies of the 51 kDa protein of the mosquito-larvicidal binary toxin from *Bacillus sphaericus*

C.-K. Chiou,[a]† E. W. Davidson,[b] T. Thanabalu,[c] A. G. Porter[c] and J. P. Allen[a]*

[a]Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287-1604, USA, [b]Department of Biology, Arizona State University, Tempe, AZ 85287-1501, USA, and [c]Institute of Molecular and Cell Biology, An Institute Affiliated to The National University of Singapore, 30 Medical Drive, Singapore 117609, Republic of Singapore

† Present address: The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA.

Correspondence e-mail: jallen@asu.edu

Received 2 November 1998
Accepted 22 February 1999

Certain strains of *Bacillus sphaericus* produce a highly toxic mosquito-larvicidal protein during sporulation which is active against vectors of dengue, encephalitis and malaria. This toxin is initially expressed as 51 and 42 kDa proteins and is converted to 43 and 39 kDa proteins, respectively, which form the active heterodimer complex. For a better understanding of the toxicity mechanism at the molecular level, the 51 kDa protein of the binary toxin of *B. sphaericus* strain 2297 was expressed as a glutathione-S-transferase fusion protein and purified by affinity chromatography. Protein crystals were grown from an amorphous precipitate in five months using the hanging-drop vapor-diffusion method. The protein crystals were dissolved and were found to be composed of a proteolytically modified 45.2 kDa derivative similar to the active form of this protein. The crystals form in space group $P4_32_12$ (or $P4_12_12$) and diffract to 2.6 Å, with unit-cell dimensions $a = b = 133.48$, $c = 69.76$ Å.

1. Introduction

Bacillus sphaericus produces insecticidal proteins during sporulation which are highly toxic to the larvae of certain mosquitoes: *Anopheles*, which are vectors of malaria, and *Culex quinquefasciatus*, which are vectors of certain encephalites and filariases. *B. sphaericus* has been used as a biopesticide to control mosquitoes, which transmit a wide range of diseases including malaria (Lacey & Undeen, 1986), which causes the death of over a million people annually (Nussenzweig & Long, 1994). The *B. sphaericus* toxins are deposited as a parasporal inclusion at sporulation (Payne & Davidson, 1984; Yousten & Davidson, 1982), are ingested by the larvae and the protoxins are dissociated in the alkaline environment of the midgut of the larvae. There, the protoxins are further activated by proteases, and the processed derivatives bind to specific receptors forming a complex on the brush-border membrane of the midgut of the larvae (Davidson, 1988; Nielsen-Leroux & Charles, 1992; Oei *et al.*, 1992).

Two major proteins, the 42 and 51 kDa proteins, are required for toxicity. These two proteins are homologous, with 25% identity and four conserved regions between their sequences (Baumann *et al.*, 1987, 1988). Although the detailed mechanism is not well understood, it has been established that the 51 kDa protein has a high affinity for specific regions of the gut epithelial cell and directs the 42 kDa protein to these sites as a tightly bound complex (reviewed in Porter *et al.*, 1993). The presence of the complex was deduced in physiological experiments, wherein measurable current changes on the cell membrane could be detected during patch-clamp experiments (Cokmus *et al.*, 1997). Though these studies implied that the binary toxin of *B. sphaericus* is a member of the class of pore-forming proteins, the lack of structural information or sequence homology for the toxin led us to determine its structure using X-ray diffraction methods. The crystallization and characterization of the 51 kDa protein of the binary toxin of *B. sphaericus* strain 2297 are reported here.

2. Materials and methods

The 51 kDa protein from *B. sphaericus* strain 2297 was cloned with glutathione-S-transferase fused to the N-terminus of the 51 kDa protein and expressed in *Escherichia coli* DH5α (Davidson *et al.*, 1990). A thrombin cleavage site is located between the two fusion partners for release of the 51 kDa protein. Briefly, cells harboring the 51 kDa toxin plasmid were grown overnight to stationary phase at 310 K in 2×YT broth (16 g tryptone, 10 g yeast extract, 5 g NaCl per liter) containing 100 µg ml^{-1} ampicillin. Cells were harvested by centrifugation at 277 K, 10000*g* for 10 min. The harvested *E. coli* cells were again resuspended in the same volume of fresh 2×YT media containing 0.1 m*M* isopropyl-β-D-thio-

Table 1
Native data statistics of the 51 kDa protein.

Space group	$P4_32_12$ (or $P4_12_12$)
Cell constants (Å)	
a	133.48
b	133.48
c	69.76
Resolution (Å)	2.6
Number of observations	48181
Unique reflections	16396
R_{merge}† (%)	8.2
Completeness (%)	82.3
Completeness in last 0.20 Å shell (%)	69.7

† $R_{merge} = \sum |I_i(h) - \langle I(h)\rangle| / \sum I(h)$, where $I_i(h)$ is the intensity measured for the *i*th reflection and $\langle I(h)\rangle$ is the mean intensity for multiply recorded reflections.

galactopyranoside (IPTG) to induce the expression of the 51 kDa protein at 310 K. After 2 h, the cells were centrifuged at 277 K, 10000*g* for 10 min. Cells were suspended and lysed by sonication in cold PBS buffer (10 g NaCl, 0.25 g KCl, 1.44 g Na_2HPO_4 and 0.25 g KH_2PO_4 per liter, pH 7.2). Next, 1% Triton X-100 was added to the suspension to release more expressed protein. The cell lysate was cleared by centrifugation at 10000*g*. The fusion protein was purified from the supernatant by affinity chromatography using 1 ml glutathione agarose beads (Pharmacia, Sweden), followed by several washes with PBS buffer. The beads were finally equilibrated in 3 ml thrombin digestion buffer (150 m*M* NaCl, 2.5 m*M* $CaCl_2$ and 50 m*M* Tris–HCl pH 8.0). To cleave the 51 kDa protein from the fusion partner, 10 µl (100 ng) thrombin (Sigma, USA) was added to the protein-coated beads, and the enzyme digestion was allowed to proceed at room temperature for 2 h. The 51 kDa protein was then separated by centrifugation of the beads at 500*g*. Typical protein yield was 1–2 mg per liter of culture. The purified protein corresponds to residues 4–448 of the 51 kDa protein plus a glycine 'kinker' sequence at the N-terminus, designed to increase accessibility of thrombin to the cleavage site.

The 51 kDa protein was dialyzed extensively into 50 m*M* Tris–HCl (pH 8.0) and concentrated to 17 mg ml^{-1} using a Centricon-30 filter (Amicon Inc, USA). Crystals were grown by the vapor-diffusion method using hanging drops. The crystallization drops were produced by mixing 10 µl of the concentrated protein and 10 µl of the reservoir solution, which consisted of 18% polyethylene glycol 8000 (Fluka), 0.2 m*M* $ZnCl_2$ and 0.1 *M* 2-(4-morpholino)-ethanesulfonic acid (Fisher Biotech) buffer (pH 6.5).

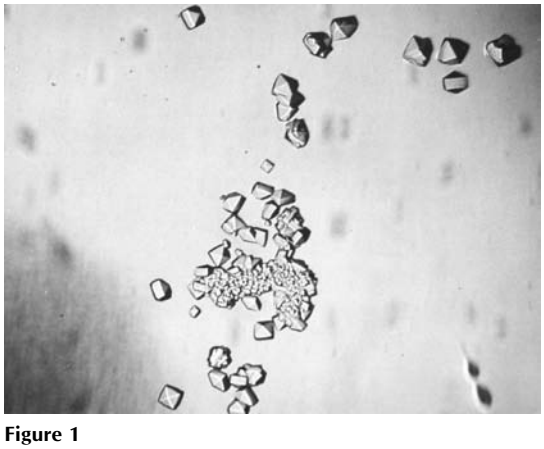

Figure 1
Crystals of the 51 kDa protein grown using the seeding method. The largest crystals are 0.2 mm in length and have well defined facets consistent with the tetragonal space group. Note that the crystals grow preferentially along the streaking path, which is approximately vertical.

3. Results and discussion

Crystals grew to a maximal size of 0.1 × 0.1 × 0.2 mm after approximately five months. A microseeding technique was applied successfully to reproduce the crystals (Thaller *et al.*, 1985; Stura & Wilson, 1990). Briefly, a cat whisker was dipped into the mother liquor of the crystal drop and streaked through a freshly made drop with equal volumes of concentrated protein and reservoir solution. The presence of crystals along the streak path was observed after approximately two months (Fig. 1).

Determination of the primary structure of the crystallized protein was performed by using both laser-desorption mass spectroscopy and amino-acid sequencing. Crystals were washed extensively with reservoir solution to completely remove the mother liquor. For the mass spectroscopy, a sinapinic acid matrix was prepared from a saturated solution in 0.1% trifluoroacetic acid containing 33% acetonitrile. Samples containing 20–50 µ*M* of protein were mixed with an equal volume of the matrix solution and 1 µl aliquots were dried on stainless sample pins and inserted into a Vestec Lasertec Research mass spectrometer. Analysis of the spectra yielded a molecular mass of 45240 Da. The amino-acid sequence of the first seven N-terminal residues of the crystallized 51 kDa protein was determined to be SLNYNLP by using automated Edman degradation on a Porton 2090E protein sequencer. This sequence corresponds to residues Ser29–Pro35 of the published toxin sequences (Baumann *et al.*, 1988). Based upon the N-terminus sequence and the measured molecular mass, the C-terminus was deduced to be at residue His421; thus, the crystallized protein is composed of a total of 393 residues. As was previously noted, the functional 51 kDa protein isolated from the native *B. sphaericus* strain 2362 exhibits a similar processing into a 45 kDa derivative during activation (Davidson *et al.*, 1990). Thus, the loss of the amino acids during the crystallization process leads to a derivative which is still fully functional and closely resembles the activated form of the protein. The proteolytic trimming in the crystallization solution could have resulted either from self-processing or cryptic proteases, *e.g.* from the culture bacteria or contaminant proteases in the thrombin solution. The processing may also have influenced the crystallization, since initially the 51 kDa protein precipitated in the drop. It is possible that the shortened fragment (45 kDa) was released into solution and grew into single crystals only after the N-terminal segment was cleaved from the expressed subunit.

Native crystals were tested for diffraction with a Rigaku R-AXIS IIC image-plate area detector mounted on a Rigaku RU-200HB rotating-anode X-ray generator at room temperature. The diffraction data were reduced using the R-AXIS software (Higashi, 1990). The 51 kDa crystals belong to space group $P4_32_12$ with $a = b = 133.48$ and $c = 69.76$ Å (Table 1). There is probably one monomer in the asymmetric unit, based upon the calculated solvent content of 61%. A native data set was collected to 2.6 Å resolution (Table 1). Heavy-metal derivatives of the 51 kDa crystals have been obtained and experiments are under way to determine the atomic resolution structure of the 51 kDa subunit of this binary toxin. Work is also in progress to determine the structure of the entire binary complex.

We thank Dr Daniel Brune for N-terminal protein sequencing and Dr Kemmons Tubbs for mass spectrometry. The authors are grateful to Dr Kathryn Ely for her critical reading of the manuscript and to Dr Pei-ting Lin and Paula Kovack for preparing the manuscript for publication. This work is supported by a grant from the National Institutes of Health GM 55031.

References

Baumann, L., Broadwell, A. H. & Baumann, P. (1988). *J. Bacteriol.* **170**, 2045–2050.

Baumann, P., Baumann, L., Bowditch, R. D. & Broadwell, A. H. (1987). *J. Bacteriol.* **169**, 4061–4067.
Cokmus, C., Davidson, E. W. & Cooper, K. (1997). *J. Invertebr. Pathol.* **69**, 197–204.
Davidson, E. W. (1988). *J. Med. Entomol.* **25**, 151–157.
Davidson, E. W., Oei, C., Meyer, M. & Bieber, A. L. (1990). *Can. J. Microbiol.* 840–878.
Higashi, T. (1990). *PROCESS: a Program for Indexing and Processing R-AXIS II Imaging Plate Data.* Rigaku Corporation, Tokyo, Japan.
Lacey, L. A. & Undeen, A. H. (1986). *Annu. Rev. Entomol.* **31**, 265–296.
Nielsen-Leroux, C. & Charles, J. (1992). *Eur. J. Biochem.* **210**, 585–590.
Nussenzweig, R. S. & Long, C. A. (1994). *Science*, **265**, 1381–1383.
Oei, C., Hindley, J. & Berry, C. (1992). *J. Gen. Microbiol.* **138**, 1515–1526.
Payne, J. M. & Davidson, E. W. (1984). *J. Invertebr. Pathol.* **43**, 383–388.
Porter, A. G., Davidson, E. W. & Liu, J.-W. (1993). *Microbiol. Rev.* **57**, 838–861.
Stura, E. A. & Wilson, I. A. (1990). *Methods*, **1**, 38–49.
Thaller, C., Eichele, G., Weaver, L. H., Wilson, E., Karlsson, R. & Jansonius, J. N. (1985). *Methods Enzymol.* **114**, 132–136.
Yousten, A. A. & Davidson, E. W. (1982). *Appl. Environ. Microbiol.* **44**, 1449–1455.