Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic study of the functional form of the *Bacillus thuringiensis* mosquito-larvicidal Cry4Aa mutant toxin

Panadda Boonserm,[a]* Chanan Angsuthanasombat[a] and Julien Lescar[b]

[a]Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Nakornpathom 73170, Thailand, and [b]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore

Correspondence e-mail: mbpbs@mahidol.ac.th

Received 5 March 2004
Accepted 9 May 2004

The 65 kDa functional form of the mosquito-larvicidal Cry4Aa-R235Q mutant toxin has been crystallized. The crystals belong to space group $C222_1$, with unit-cell parameters a = 91.2, b = 202.1, c = 98.7 Å, and contain one molecule per asymmetric unit. The crystals diffract to ~2.9 Å using synchrotron radiation and a complete native data set has been collected. The structure has been solved using a molecular-replacement method with the Cry4Ba toxin protein as a search model.

1. Introduction

During sporulation, the Gram-positive bacterium *Bacillus thurigiensis* (Bt) expresses insecticidal crystal proteins known as δ-endotoxins, which are grouped into two major families: Cry and Cyt δ-endotoxins (Crickmore *et al.*, 1998). Depending on the bacterial isolates, these crystal proteins are toxic towards a variety of insect larvae of the orders Lepidoptera (moths and butterflies), Diptera (mosquitoes and blackflies), Coleoptera (beetles and weevils) and Hymenoptera (wasps and bees) (Schnepf *et al.*, 1998). Bt δ-endotoxins are synthesized as inactive protoxins that are processed by larval midgut proteases into their active form. The activated toxin then binds to a specific receptor on the midgut epithelial cells and inserts into the cell membrane, creating ion-leakage pores that cause cell swelling and eventually cell death by colloid osmotic lysis (Knowles, 1994).

Several three-dimensional crystal structures of Cry δ-endotoxins, namely Cry3Aa (Li *et al.*, 1991), Cry1Aa (Grochulski *et al.*, 1995), Cry1Ac (Li *et al.*, 2001), Cry3Bb (Galitsky *et al.*, 2001) and Cry2Aa (Morse *et al.*, 2001), have been elucidated. Recently, the crystallization of the 68 kDa active fragment of the dipteran-specific Cry4Ba was reported (Boonserm *et al.*, 2003). Active Cry toxins (60–70 kDa) are composed of three globular domains: a helix bundle (domain I), a domain containing anti-parallel β-strands with a Greek-key topology (domain II) and a β-sandwich with jelly-roll topology (domain III). Since the N-terminal domain I contains amphipathic α-helices surrounding a central hydrophobic α-helix that is long enough to span a lipid bilayer, it was proposed that after a large conformational change this domain could insert into the membrane of the host cell to form a pore (Li *et al.*, 1991; Grochulski *et al.*, 1995). This prediction was subsequently validated by expressing isolated domain I fragments or synthetic helices that were shown to penetrate the membrane and form ion-permeable pores (Walters *et al.*, 1993; Von Tersch *et al.*, 1994; Gazit *et al.*, 1998; Puntheeranurak *et al.*, 2004). Domain II plays a role in binding to receptors at the surface of the insect midgut. Site-directed mutagenesis targeting three exposed loops at the apex of domain II identified the residues that were involved in receptor binding of several insect species (Smedley & Ellar, 1996; Rajamohan *et al.*, 1996; Jurat-Fuentes & Adang, 2001), suggesting a role of these loop regions as the primary determinant of insect specificity. The C-terminal domain III could be involved in preserving the structural integrity of toxin molecules (Li *et al.*, 1991), in determining specificity (Lee *et al.*, 1995; Burton *et al.*, 1999) and in the regulation of ion-channel activity (Chen *et al.*, 1993; Wolfersberger *et al.*, 1996; Schwartz *et al.*, 1997).

Bacillus thuringiensis subsp. *israelensis* (Bti) is highly toxic towards the larvae of *Aedes*, *Culex* and *Anopheles*, which are vectors of dengue fever, filariasis and malaria, respectively (Roberts, 2002). This bacterium produces four major insecticidal proteins: Cry4Aa, Cry4Ba, Cry11Aa and Cyt1Aa (Schnepf *et al.*, 1998). Of these, Cry4Aa and Cry4Ba are most closely related, with about 55% sequence similarity. Although both proteins are toxic towards mosquito larvae, their level of toxicity towards the various mosquito species varies. In the absence of an experimental three-dimensional structure, putative loop residues from Cry4Aa (specific against *Culex* and to a lesser extent *Aedes* larvae) were exchanged with the corresponding residues from Cry4Ba (specific against *Aedes* and *Anopheles* larvae). Exchanging loop 3 of domain II resulted in a significant increase of toxicity towards *Culex*, with no significant decrease towards *Aedes* larvae (Abdullah *et al.*, 2003). It was reported

that both Cry4Aa and Cry4Ba protoxins were cleaved by trypsin into fragments of about 20 and 45 kDa, in addition to the removal of the C-terminal half of the toxins (Yamagiwa *et al.*, 1999; Angsuthanasombat *et al.*, 1993). These fragments were produced by trypsin cleavage at Arg235 and Arg203 of Cry4Aa and Cry4Ba, respectively. As previously reported, alanine substitution of Arg203, resulting in blocking the tryptic cleavage site of Cry4Ba, was shown to increase larvicidal activity (Angsuthanasombat *et al.*, 1993; Abdullah *et al.*, 2003). Additionally, elimination of the tryptic cleavage site at Arg235 by glutamine substitution (R235Q) led to retention of the high-level toxicity of the Cry4Aa mutant against *Aedes aegypti* larvae, indicating that the trypsin-site mutation had no adverse effect on the Cry4Aa toxin toxicity (Boonserm *et al.*, 2004).

Given the widespread emergence of resistance towards conventional pest-control agents, bacterial Cry proteins are of great interest for the development of new specific bio-insecticides. Thus, the availability of experimentally determined three-dimensional structures will be of great use in order to help in the the design of rationally modified toxins with altered specificity and improved potency. Here, we report the purification, crystallization, data collection and structure determination of the Cry4Aa-R235Q active mutant from Bti.

2. Materials and methods

2.1. Protein expression and solubilization

Escherichia coli JM109 cells harbouring a single amino-acid Cry4Aa mutant (see §3.1 below) were grown at 303 K in Luria–Bertani medium containing 100 µg ml^{-1} ampicillin until the OD_{600} of the culture reached 0.3–0.5. Protein expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.1 m*M* for 10 h and the culture was analysed by sodium dodecyl sulfate–10%(*w*/*v*) polyacrylamide gel electrophoresis (SDS–PAGE). *E. coli* cultures overexpressing the Cry4Aa single amino-acid mutant toxin as cytoplasmic inclusions were harvested by centrifugation and resuspended in cold distilled water. The cell suspension was then disrupted using a French pressure cell at 69 MPa. After centrifugation at 8000*g* at 277 K for 15 min, the pellets were washed three times in cold distilled water and resuspended by sonication. The protein concentrations of the partially purified inclusions were determined using a protein microassay reagent (Bio-Rad), with bovine serum albumin fraction V (Sigma) as a standard. Inclusions at a concentration of about 2 mg ml^{-1} were solubilized by incubation at 310 K for 1 h in 50 m*M* Na_2CO_3 pH 10.0. Solubilized protoxins were then separated from insoluble materials by centrifugation at 12 000*g* for 15 min.

2.2. Proteolytic activation and active-toxin purification

The 130 kDa solubilized Cry4Aa protoxins were mixed with trypsin (L-1-tosylamide-2-phenylethyl chloromethyl ketone-treated, Sigma) at an enzyme:protoxin ratio of 1:10(*w*/*w*) and incubated at 310 K for 16 h. Proteolysis was stopped by adding 1 m*M* tosyl-lysine chloromethyl ketone (TLCK). After analysis by SDS–PAGE, the trypsin-activated fraction was concentrated at 277 K using a Centriprep ultrafiltration device with a 30 kDa molecular-weight cutoff (Amicon). The protein was further purified by size-exclusion chromatography on an FPLC system (Superdex 200, Amersham Pharmacia Biotech) in 50 m*M* Na_2CO_3 pH 10.0 at a flow rate of 0.4 ml min^{-1}. Under these conditions, the 65 kDa Cry4Aa toxin elutes as a monomer, as shown by using BSA (67 kDa) as a marker. Eluted fractions containing the proteins were pooled and concentrated to 3–5 mg ml^{-1} by ultrafiltration as described above.

2.3. Crystallization and data collection

Crystals were grown using the hanging-drop vapour-diffusion technique. 5 µl precipitant solution was mixed with an equal volume of purified Cry4Aa at a concentration of 3–5 mg ml^{-1}. The drop was equilibrated against a reservoir containing 1 ml of the precipitant solution at 296 K. The purified protein in 50 m*M* Na_2CO_3 pH 10.0 was first mixed with solutions of 0.1 *M* Tris–acetate at pH values of between 7 and 9. Several precipitants, including salts, various molecular-weight polyethylene glycols (PEG), glycerol and 2-methyl-2,4-pentanediol (MPD), were tested. However, crystals only appeared in the conditions containing salts. Clusters of needle-like crystals were found in 0.1 *M* Tris–acetate pH 7.0 and 0.3–0.4 *M* Li_2SO_4. Microcrystals appeared after two weeks in 0.1 *M* Tris–acetate pH 7.0 and 0.2–0.3 *M* LiCl. However, in both cases crystals were too small and too poorly ordered for X-ray diffraction studies. Clusters of thin plate-shaped crystals were successively obtained using a protein concentration greater than 5 mg ml^{-1} in precipitant solution containing 0.1 *M* Tris–acetate pH 7.0 and 0.2–0.3 *M* KH_2PO_4. This condition was optimized by lowering the protein concentration to 3–5 mg ml^{-1}.

For data collection, crystals were soaked briefly in a cryoprotectant solution containing 20% 2-methyl-2,4-pentanediol (MPD), 10% polyethylene glycol (PEG) 400 and 0.1 *M* Tris–acetate and 0.3 *M* KH_2PO_4pH 7.0 before being mounted in a cryoloop and cooled to 100 K in a nitrogen-gas stream (Oxford Cryosystems). Measurements were made at an X-ray wavelength of 0.976 Å at ESRF beamline ID29 with an attenuated beam of dimensions 0.1 × 0.1 mm. Diffraction intensities were recorded on an ADSC Quantum IV CCD detector. The crystal-to-detector distance was set to 250 mm and the oscillation angle for each of the 180 images recorded was 1°. Integration, scaling and merging of the intensities were carried out using *MOSFLM* (Leslie, 1992) and *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Molecular replacement was carried out using *AMoRe* (Navaza, 1994).

3. Results and discussion

3.1. Protein activation and purification

It has been reported that the Cry4Aa protoxin is processed into two protease-resistant fragments of 20 and 45 kDa by the intramolecular cleavage of a 65 kDa intermediate and that these two fragments need to associate to exert toxicity (Yamagiwa *et al.*, 1999). The cleavage site is after Arg235 (Angsuthanasombat *et al.*, 1993). We have recently investigated the effect of intramolecular cleavage on the toxicity of the Cry4Aa toxin by constructing a Cry4Aa single amino-acid mutant (R235Q) that is devoid of this internal cleavage site and is thus resistant to proteolysis (Boonserm *et al.*, 2004). The R235Q single mutant still retains high toxicity towards *Aedes aegypti* larvae at a level comparable to that of the wild type (Boonserm *et al.*, 2004), indicating that the mutation at this trypsin-cleavage site has no adverse effect on the Cry4Aa toxicity. The active R235Q mutant of Cry4Aa was used for the present crystallographic study. After trypsin treatment of the R235Q active mutant, a major proteolytic fragment of 65 kDa was obtained that was resistant to further proteolysis. N-terminal amino-acid sequencing after trypsin activation indicated the first residue to be Gln5, which is putatively located before helix 1 in domain I on the basis of a sequence align-

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell (3.05–2.95 Å).

Space group	$C222_1$
Unit-cell parameters (Å)	
a	91.20
b	202.07
c	98.73
Resolution range (Å)	20.0–2.95
Mosaicity (°)	0.5
Measured reflections	138542 (13838)
Unique reflections	19646 (1896)
Data redundancy	7.1 (7.3)
Completeness (%)	100.0 (100.0)
Average $I/\sigma(I)$	15.8 (3.7)
$R_{\rm merge}$ (%)†	13.4 (47.4)

† $R_{\rm merge} = (\sum |I_i - \langle I \rangle| / \sum I_i) \times 100$, where I_i is an individual intensity observation, $\langle I \rangle$ is the mean intensity for that reflection and the summation is over all reflections.

ment with Cry3Aa. The trypsin-activated 65 kDa fragment was purified by gel filtration for subsequent crystallization trials.

3.2. Crystallization and data collection

The 68 kDa chymotrypsin-activated Cry4Ba toxin has been crystallized previously and well ordered crystals diffracting X-rays to 1.75 Å resolution were obtained (Boonserm *et al.*, 2003). A similar approach was used in the present study to promote the crystal growth of the Cry4Aa toxin. Salts were effective precipitants and plate-shaped crystals of Cry4Aa with approximate dimensions of 0.15 × 0.15 × 0.005 mm (Fig. 1) were obtained using 0.1 *M* Tris–acetate pH 7.0 and 0.2–0.3 *M* KH_2PO_4 as precipitants.

One such crystal was used to obtain a 100% complete native data set to 2.95 Å resolution. Data-collection statistics are summarized in Table 1. Assuming the presence of one Cry4Aa molecule per asymmetric unit, the value of $V_{\rm M}$ (Matthews, 1968) is 3.60 Å^3 Da^{-1}, giving a solvent content of 65%.

Figure 1
Crystals of Cry4Aa. Plate-shaped crystals (0.15 × 0.15 × 0.005 mm) can be seen in the crystallization drop.

3.3. Solution of the structure

A preliminary model of the structure was found by molecular replacement using the program *AMoRe* (Navaza, 1994). The rotation-function calculation was performed in the resolution range 20.0–4.0 Å using the Cry4Ba refined crystal structure as the search model (Li, 2004) and a Patterson integration radius of 30 Å. This model returned a weakly contrasted solution with a correlation coefficient for the structure-factor amplitudes of 0.160 (0.152 for the second highest solution) and a value of 0.27 for the intensities (0.24 for the second highest solution). This solution consistently appears as the first when varying the Patterson integration radius between 30 and 35 Å. The search model was then placed in the unit cell using the Crowther–Blow translation function with data between 9.0 and 4.0 Å resolution. This model returned a solution with a correlation coefficient of 0.157 and an *R* factor of 49.8% (0.117 and an *R* factor of 51.0% for the second highest peak). As a control, a systematic search using the Crowther–Blow translation function was carried out for the first 50 independent peaks of the rotation function in each of the space groups *C*222 and $C222_1$. This search unambiguously returned the same solution for space group $C222_1$ with the same relatively weak but significant contrast as stated above. This result is consistent with the low sequence identity between the search model Cry4Ba and Cry4Aa (35.6% for 516 aligned positions).

Rigid-body refinement yielded a correlation coefficient of 0.17 and an *R* factor of 49.6% for 7201 reflections between 9.0 and 4.0 Å resolution. Examination of the crystal packing did not reveal any steric hindrance from symmetry-related molecules. 1171 reflections randomly chosen between 20 and 2.9 Å (5% of the data) were set aside to monitor the progress of the refinement. Further refinement with each of the three protein domains treated as three independent rigid bodies followed by one cycle of molecular dynamics in torsion-angle space was carried out using *CNS* (Brünger *et al.*, 1998). This refinement returned an overall *R* factor of 34.1% for 13 814 reflections between 20 and 3.24 Å resolution and an $R_{\rm free}$ of 47.7%. Electron-density maps generated with phases from the partial model (516 residues) show several features compatible with the pattern of insertions and deletions as well as amino-acid substitutions between Cry4Aa and Cry4Ba. Refinement of the structure is now in progress using manual model building with the program *O* (Jones *et al.*, 1991) interspersed with cycles of electron-density map improvement and refinement with the programs *ARP*/*wARP*, *CNS* and *REFMAC* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). The current values for the *R* factor and $R_{\rm free}$ are 24.1 and 29.3% for all data in the resolution range 7–2.9 Å.

The authors gratefully acknowledge financial Support from the NTU *via* a startup grant, the Singapore BioMedical Research Council (to JL) and the Thailand Research Fund (to PB), as well as help and useful discussions with members of the ESRF/EMBL Joint Structural Biology Group, especially Dr Edward Mitchell.

References

Abdullah, M., Alzate, O., Mohammad, M., McNall, R. J., Adang, M. J. & Dean, D. H. (2003). *Appl. Environ. Microbiol.* **69**, 5343–5353.
Angsuthanasombat, C., Crickmore, N. & Ellar, D. J. (1993). *FEMS Microbiol. Lett.* **111**, 255–262.
Boonserm, P., Ellar, D. J. & Li, J. (2003). *Acta Cryst.* D**59**, 591–594.
Boonserm, P., Pornwiroon, W., Katzenmeier, G., Panyim, S. & Angsuthanasombat, C. (2004). *Protein Exp. Purif.* **35**, 397–403.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P. Grosse-Kunstleve, R. W., Jiang, J.-S., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Burton, S. L., Ellar, D. J., Jade, L. & Derbyshire, D. J. (1999). *J. Mol. Biol.* **287**, 1011–1022.
Chen, X. J., Lee, M. K. & Dean, D. H. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9041–9045.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Crickmore, N., Zeigler, D. R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J. & Dean, D. H. (1998). *Microbiol. Mol. Biol. Rev.* **62**, 807–813.
Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J. R., Pangborn, W. & English, L. (2001). *Acta Cryst.* D**57**, 1101–1109.
Gazit, E., La Rocca, P., Sansom, M. S. & Shai, Y. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12289–12294.
Grochulski, P., Masson, L., Borisova, S., Pusztai-carey, M., Schwartz, J. L., Brousseau, R. & Cygler, M. (1995). *J. Mol. Biol.* **254**, 447–464.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.*, A**47**, 110–120.
Jurat-Fuentes, J.-L. & Adang, M. J. (2001). *Appl. Environ. Microbiol.* **67**, 323–329.
Knowles, B. H. (1994). *Adv. Insect Physiol.* **24**, 275–308.
Lee, M. K., Young, B. A. & Dean, D. H. (1995). *Biochem. Biophys. Res. Commun.* **216**, 306–312.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**, 27–33.
Li, J. (2004). Personal communication.
Li, J., Carroll, J. & Ellar, D. J. (1991). *Nature* (*London*), **353**, 815–821.
Li, J., Derbyshire, D. J., Promdonkoy, B. & Ellar, D. J. (2001). *Biochem. Soc. Trans.* **29**, 571–577.
Matthews, B. W. (1968). *J. Mol. Biol.* **197**, 375–377.
Morse, R. J., Yamamoto, T. & Stroud, R. M. (2001). *Structure*, **9**, 409–417.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Puntheeranurak, T., Uawithya, P., Potvin, L., Angsuthanasombat, C. & Schwartz, J.-L. (2004). *Mol. Membr. Biol.* **21**, 67–74.
Rajamohan, F., Hussain, S. R., Cotrill, J. A., Gould, F. & Dean, D. H. (1996). *J. Biol. Chem.* **271**, 25220–25226.
Roberts, L. (2002). *Science*, **298**, 82.
Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D. R. & Dean, D. H. (1998). *Microbiol. Mol. Biol. Rev.* **62**, 775–806.
Schwartz, J.-L., Juteau, M., Grochulski, P., Cygler, M., Prefontaine, G., Brousseau, R. & Masson, L. (1997). *FEBS Lett.* **410**, 397–402.
Smedley, P. D. & Ellar, D. J. (1996). *Microbiology*, **142**, 1617–1624.
Von Tersch, M. A., Slatin, S. L., Kulesza, C. A. & English, L. H. (1994). *Appl. Environ. Microbiol.* **60**, 3711–3717.
Walters, F. S., Slatin, S. L., Kulesza, C. A. & English, L. H. (1993). *Biochem. Biophys. Res. Commun.* **196**, 921–926.
Wolfersberger, M. G., Chen, X. J. & Dean, D. H. (1996). *Appl. Environ. Microbiol.* **62**, 279–282.
Yamagiwa, M., Esaki, M., Otake, K., Inagaki, M., Komano, T. & Sakai, H. (1999). *Appl. Environ. Microbiol.* **65**, 3464–3469.