Molecular Characterization of a Gene Encoding a 72-Kilodalton Mosquito-Toxic Crystal Protein from *Bacillus thuringiensis* subsp. *israelensis*

WILLIAM P. DONOVAN,* CATHY DANKOCISK, AND M. PEARCE GILBERT

Ecogen Inc., 2005 Cabot Boulevard West, Langhorne, Pennsylvania 19047

Received 12 May 1988/Accepted 13 July 1988

A gene encoding a 72,357-dalton (Da) crystal protein of *Bacillus thuringiensis* var. *israelensis* was isolated from a native 75-MDa plasmid by the use of a gene-specific oligonucleotide probe. *Bacillus megaterium* cells harboring the cloned gene *(cryD)* produced significant amounts of the 72-kDa protein (CryD), and the cells were highly toxic to mosquito larvae. In contrast, *cryD*-containing *Escherichia coli* cells did not produce detectable levels of the 72-kDa CryD protein. The sequence of the *cryD* protein, as deduced from the sequence of the *cryD* gene, was found to contain regions of homology with two previously described *B. thuringiensis* crystal proteins: a 73-kDa coleopteran-toxic protein and a 66-kDa lepidopteran- and dipteran-toxic protein of *B. thuringiensis* subsp. *kurstaki*. A second gene encoding the *B. thuringiensis* subsp. *israelensis* 28-kDa crystal protein was located approximately 1.5 kilobases upstream from and in the opposite orientation to the *cryD* gene.

Certain varieties of *Bacillus thuringiensis* synthesize parasporal crystals composed of proteins that have been shown to be toxic to the larvae of specific insects. *B. thuringiensis* subsp. *kurstaki* as well as other varieties produces a bipyrimal crystal composed of one or more related proteins of approximately 130 kilodaltons (kDa) which are toxic to lepidopterans (caterpillars) (for recent reviews, see references 2 and 34) and also a cuboidal crystal composed of a 66-kDa protein that is toxic to both lepidopteran and dipteran (mosquito, black fly) insects (6, 37). Other subspecies of *B. thuringiensis* have been identified which produce rhomboids crystals composed of a 73-kDa protein that is specifically toxic to coleopteran (beetle) larvae (12, 16; W. P. Donovan, J. M. Gonzalez, Jr., M. P. Gilbert, and C. Dankocsik, Mol. Gen. Genet., in press).

*B. thuringiensis* subsp. *israelensis* synthesizes an irregularly shaped parasporal crystal that is highly toxic to certain dipteran larvae (8). The complex crystal is composed of at least three major proteins of approximately 130 kDa, 70 kDa, and 28 kDa. The genes for the 130-kDa and the 28-kDa crystal proteins have been cloned and their nucleotide sequences have been reported (1, 25, 30, 31). These cloning experiments have indicated that the 130-kDa and the 28-kDa *B. thuringiensis* subsp. *israelensis* crystal proteins are mosquito toxic. However, other researchers have reported that the 28-kDa protein has little or no mosquitoactivity (4, 5, 11, 14, 15, 28). Cloning experiments have revealed that *B. thuringiensis* subsp. *israelensis* contains more than one gene for the 130-kDa protein (3, 32). To our knowledge there have been no reports concerning the cloning of the gene for the 70-kDa crystal protein.

We report here the isolation and complete nucleotide sequence of a *B. thuringiensis* subsp. *israelensis* gene, which we have designated *cryD*, encoding a 72-kDa crystal protein. Bioassay determinations with *Bacillus megaterium* cells harboring the cloned *cryD* gene demonstrated that the CryD protein is highly toxic to mosquito larvae. Sequence comparisons are presented which reveal that the CryD protein is related to two other *B. thuringiensis* crystal proteins that have distinct entomocidal activities.

**MATERIALS AND METHODS**

**Bacterial strains and plasmids.** Strain HD-567 of *B. thuringiensis* subsp. *israelensis* serotype 14 (NRRL B-18304, Peoria, Ill.), obtained from the collection of H. T. Dulmage, Cotton Insects Research, U.S. Department of Agriculture, S.E.A., Brownsville, Tex., was the source of the 72-kDa crystal protein and of the *cryD*-containing DNA. *B. megaterium* VT1660 (29) was used as a host for pNN101 (20) plasmid derivatives. *Escherichia coli* HB101 was used as a host for pBR322 derivatives. *E. coli* JM101 was the host for the sequencing vectors M13mp18 and M13mp19 and their derivative phages.

**Protein purification and NH₄-terminal amino acid sequence determination.** The methods for purifying crystal proteins have been described previously (6). A hot sodium dodecyl sulfate (SDS)-2-mercaptoethanol solution was used to solubilize crystal proteins from a sporulated culture of *B. thuringiensis* subsp. *israelensis* HD-567. The 72-kDa protein was purified from SDS gels by the procedure of Hunkapiller et al. (13). After precipitation with acetone (1:1, vol/vol), the 72-kDa protein was subjected to automated Edman degradation in an Applied Biosystems Gas-Phase Sequenator (model 470A) and analyzed on a DuPont Zorbax C18 column in a Hewlett-Packard high-pressure liquid chromatograph (model 1090) with 1040 diode array detection.

**Cloning.** The methods for constructing plasmid libraries enriched for size-specific DNA restriction fragments and for using synthetic oligonucleotides as gene-specific hybridization probes have been described before (6). The *cryD*-enriched plasmid library was transformed into *E. coli* HB101 cells, and ampicillin-resistant colonies were selected. These colonies were used in colony hybridization experiments with the *cryD*-specific 47-mer oligonucleotide probe that had been radioactively labeled at its 5' terminus with phage T4 kinase and [γ-³²P]ATP.

**Preparation of samples for protein gels.** *E. coli* cells were grown for 48 h at 30°C on LB agar plates (1% tryptone, 0.5% yeast extract [both from Difco], 0.5% NaCl, 1.5% agar, pH 7.3) for 36 h.
of the purified protein yielded the NH₂-terminal sequence shown in Fig. 1. Based on the sequence, a gene-specific 47-mer oligonucleotide probe was designed (Fig. 1). To determine the sizes of restriction fragments containing at least the NH₂-terminal region of the gene for the 72-kDa protein, the 47-mer probe was radioactively labeled and used in DNA blot hybridization experiments with total restriction enzyme-digested DNA from strain HD-567. The probe specifically hybridized to a unique HindIII restriction fragment of approximately 11 kilobases (kb) and to a unique EcoRI fragment of approximately 6 kb at a hybridization temperature of 47°C (data not shown).

A recombinant plasmid library was constructed by ligating size-selected, 5- to 7-kb EcoRI restriction fragments of HD-567 DNA into the EcoRI site of the E. coli vector pBR322. Transformed E. coli colonies containing recombinant plasmids were hybridized at 47°C with the labeled probe. The probe hybridized strongly to one colony (EG1318) which contained a plasmid (pEG214) that consisted of pBR322 plus a 5.7-kb EcoRI insert (Fig. 2, pEG214). The probe specifically hybridized to the 5.7-kb EcoRI fragment of pEG214 and also to a 1.1-kb Dral fragment of pEG214 (Fig. 2). Sequencing of the 1.1-kb fragment revealed a long open reading frame that began with the NH₂-terminal sequence, as previously determined by Edman analysis, for the 72-kDa protein. We have designated this open reading frame cryD. The location and orientation of the cryD gene are shown in Fig. 2 (pEG214). Sequencing of a 0.8-kb PvuII-EcoRI fragment from pEG214 revealed that the cryD open reading frame extended through the EcoRI site. Therefore, the 3' end of the cryD gene was not contained on the 5.7-kb EcoRI fragment.

The 5.7-kb EcoRI fragment hybridized, as expected, to an approximately 11-kb HindIII fragment of HD-567 DNA (not shown). The 5.7-kb fragment was used as a probe in colony hybridization experiments to isolate a recombinant plasmid (pEG216) consisting of pBR322 plus the 11.0-kb HindIII fragment (Fig. 2). The 11.0-kb fragment contained approximately 2.8 kb and 3.0 kb on either side of the 5.7-kb fragment (pEG216, Fig. 2). Sequencing of the 2.1-kb Cla-PvuII fragment and the 0.8-kb EcoRI fragment from pEG216 revealed that the 11-kb fragment contained the complete cryD gene. The cryD open reading frame was terminated by a translation stop codon located 76 codons beyond one end of the 5.7-kb EcoRI fragment. This result indicates that plasmid pEG214 (5.7-kb EcoRI) contained a truncated form of the cryD gene, designated cryDΔ76, that lacked 76 COOH-terminal codons. The complete sequence of the cryD gene and the deduced sequence of the CryD protein are shown in Fig. 3. The cryD gene encoded a protein of 72,357 Da (643 amino acids). Ten nucleotides upstream from the NH₂-terminal methionine, a purine-rich sequence (AAAGGGTTG) was found that probably serves as a ribosome-binding site. A 10-nucleotide inverted repeat (ΔG = −15.0 kcal/mol) was located 33 nucleotides downstream from the cryD open reading frame (Fig. 3).

Identification of a crystal protein gene adjacent to cryD. We observed that the restriction map of the 11-kb HindIII fragment (Fig. 2, pEG216) was similar to the restriction map recently reported by McLean and Whiteley (18) of a 10.7-kb HindIII fragment from B. thuringiensis subsp. israelensis serotype H-14. The 10.7-kb fragment had been reported to contain the gene for the 28-kDa B. thuringiensis subsp. israelensis crystal protein (18, 30). To determine whether the 11.0-kb fragment described in this report contained a similar gene, a 2.6-kb BamHI-PvuII fragment was subcloned from
pEG216 (11.0-kb HindIII) and partially sequenced. The 2.6-kb fragment contained an open reading frame that was identical to at least the first 21 NH₂-terminal codons of the gene for the 28-kDa protein (30), indicating that the gene encoding the 28-kDa protein was also located on the 11-kb HindIII fragment. Figure 2 (pEG216) shows that the gene for the 28-kDa protein, which we have designatedcryE, was located approximately 1.5 kb upstream from the cryD gene and was oriented in the opposite direction.

Expression of cryD and cryE in E. coli and B. megaterium. E. coli cells harboring either pEG214 (cryDΔ76 cryE⁺) (strain EG1318) or pEG216 (cryD⁺ cryE⁺) (strain EG1315) did not contain detectable levels of any plasmid-encoded 72-kDa or 28-kDa proteins (data not shown), and the cells were not toxic to mosquito larvae (Table 1). B. megaterium cells harboring pEG217 (cryD⁺ cryE⁺) (strain EG1316) contained significant amounts of a plasmid-encoded 72-kDa protein and minor amounts of a plasmid-encoded 28-kDa protein (Fig. 4, lane 3). These cells were not toxic to mosquito larvae (Table 1). B. megaterium cells harboring pEG217 (cryD⁺ cryE⁺) (strain EG1316) contained significant amounts of a plasmid-encoded 72-kDa protein and minor amounts of a plasmid-encoded 28-kDa protein (Fig. 4, lane 4), and the cells were highly toxic to mosquito larvae (Table 1). B. megaterium cells (strain EG1325) harboring pEG220 (Cry⁻) (constructed by ligating the Bacillus vector pNN101 into the SpF site of the E. coli vector pBR322) were not toxic (Table 1).

To further evaluate the toxicity of the 72-kDa protein, plasmid pEG218 (cryD⁺ cryE21) (Fig. 2) was constructed by subcloning a 6.7-kb BamHI-HindIII fragment from the 11.0-kb HindIII fragment into pBR322. The 6.7-kb fragment contained the complete cryD gene but only 21 NH₂-terminal codons of the cryE gene. The Bacillus vector pNN101 was inserted into the Spfl site of pEG218, resulting in the E. coli-Bacillus shuttle plasmid pEG219 (cryD⁺ cryE21) (Fig. 2). B. megaterium cells harboring pEG219 (strain EG1323) synthesized significant amounts of the CryD protein and, as
expected, no 28-kDa CryE protein (Fig. 4, lane 5). EG1323 cells had an LD₉₀ (6 to 3 µg of cells [wet weight] per ml) similar to that of EG1316 (cryD⁺ cryE⁺) cells (Table 1). The CryD protein represented 0.5% of the wet weight of EG1316 and EG1317 cells (estimated from Coomassie stained SDS gels as in Fig. 4). Therefore, the LD₉₀ value for the CryD protein against Aedes aegypti larvae was approximately 0.03 to 0.01 µg of protein per ml (0.005 × LD₉₀ cells, wet weight).

EG1316(pEG217 cryD⁺ cryE⁺) and EG1323(pEG219 cryD⁺ cryE⁺) cells were significantly inhibited in their ability to form spores (less than 5% spore formation), and the unsporulated cells usually contained one or more phase-bright inclusions (not shown). EG1324(pEG220 cryD⁺ cryE⁺) cells were similarly inhibited in spore formation; however, the unsporulated cells did not contain inclusions.

EG1324(pEG220 Cry⁺) cells formed approximately 80% spores, and the unsporulated cells did not contain inclusions.

FIG. 3. DNA sequence of cryD. The sequence begins with the Drai site and ends 360 nucleotides beyond the EcoRI site, as shown in Fig. 2 (pEG214 and pEG216). Arrows denote the inverted repeat described in the text.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Plasmid(s)</th>
<th>LD₉₀ (µg of cells/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. thuringiensis subsp. israelensis HD-567</td>
<td>135, 105, 75, 68, 10.6, 4.9</td>
<td>0.2-0.05</td>
</tr>
<tr>
<td>E. coli</td>
<td>EG1315</td>
<td>pEG216 (cryD⁺ cryE⁺)</td>
</tr>
<tr>
<td></td>
<td>EG1318</td>
<td>pEG214 (cryDA76 cryE⁺)</td>
</tr>
<tr>
<td>B. megaterium</td>
<td>EG1325</td>
<td>pEG220 (Cry⁺)</td>
</tr>
<tr>
<td></td>
<td>EG1326</td>
<td>pEG215 (cryD⁺ cryE⁺)</td>
</tr>
<tr>
<td></td>
<td>EG1316</td>
<td>pEG217 (cryD⁺ cryE⁺)</td>
</tr>
<tr>
<td></td>
<td>EG1323</td>
<td>pEG219 (cryD⁺ cryE⁺)</td>
</tr>
</tbody>
</table>
In each case the presence of inclusions in recombinant *B. megaterium* cells corresponded with the presence of the CryD protein, and therefore the inclusions are most likely aggregations of the CryD protein.

**Plasmid location of the cryD and cryE genes.** Strain HD-567 contains native plasmids of approximately 3.3, 4.2, 4.9, 10.6, 68, 75, 105, and 135 MDa (10). To determine whether any of these plasmids carried the cryD gene, the plasmids of this strain were electrophoretically size fractionated on an agarose gel (7, 9) and transferred to a nitrocellulose filter, and the filter was hybridized at moderate stringency (65°C) with the radioactively labeled 2.5-kb *EcoRV-EcoRI* cryD fragment from pEG216. The results of this analysis are shown in Fig. 5. The cryD fragment specifically hybridized to a 75-MDa plasmid from strain HD-567 (Fig. 5A and B, lane 1), indicating that this plasmid carried the cryD gene and also the cryE gene. This result confirms and extends the findings of Gonzalez and Carlton (10), who reported that the 75-MDa plasmid was necessary for crystal formation in *B. thuringiensis* subsp. *israelensis*. The cryD fragment also hybridized to a diffuse band of DNA from strain HD-567 (Fig. 5B, lane 1). This band of cryD-hybridizing DNA was not observed in derivatives of strain HD-567 that had been cured of the 75-MDa plasmid (Fig. 5A and B, lanes 2 and 3), suggesting that the DNA was derived from the 75-MDa plasmid. To further demonstrate that the diffuse band of hybridizing DNA was derived from the 75-MDa plasmid, the 75-MDa plasmid was transferred by conjugation back into a derivative of HD-567 that had been previously cured of this plasmid. As expected, the resulting transcient acquired, in addition to the cryD-hybridizing 75-MDa plasmid, a diffuse band of hybridizing DNA (Fig. 5A and B, lane 4).  

**Homologies between the CryD protein and other crystal proteins.** The computer search program of Queen and Korn (22) was used to compare the sequence of CryD with the reported sequences of six other *B. thuringiensis* crystal proteins. Two crystal proteins were found to be homologous with CryD. The 66-kDa lepidopteran- and dipteran-toxic CryB1 protein of *B. thuringiensis* subsp. *kurstaki* (6) contained a sequence of 215 amino acids (residues 61 to 275) that was 30% homologous to a sequence of 211 amino acids (residues 45 to 255) in CryD (Fig. 6). The 73-kDa coleopteran-toxic CryC protein of BT strain EG2158 and *B. thuringiensis* subsp. *tenreciensis* and *sandiego* (12, 26; Donovan et al., in press) contained a sequence of 124 amino acids (residues 107 to 230) that was 33% homologous to a sequence of 114 amino acids (residues 76 to 189) in CryD (Fig. 6). Interestingly, the CryC and CryD proteins contained similar numbers of amino acids, 644 and 643, respectively. CryD shared no significant regions of homology with either of the other two major crystal proteins of *B. thuringiensis* subsp. *israelensis*, the 130-kDa protein (31) or the 28-kDa protein (30). No significant similarities were detected between CryD and a cloned *B. thuringiensis* subsp. *israelensis* gene that potentially encoded a 72-kDa mosquito-toxic protein (27) or between CryD and a 130-kDa lepidopteran-toxic protein of *B. thuringiensis* subsp. *kurstaki* (24).

**DISCUSSION**

We have described the cloning and characterization of a unique *B. thuringiensis* subsp. *israelensis* crystal protein...
gene, the cryD gene, encoding a protein of 72,357 Da. B. megaterium cells harboring the cloned cryD gene contained crystallike inclusions composed of the 72-kDa CryD protein, and the protein was highly toxic to mosquito larvae. Our finding that the CryD protein is mosquito-toxic is in agreement with the findings of previous researchers, who reported that B. thuringiensis subsp. israelensis crystal proteins of 65 to 68 kDa, presumably similar to the 72-kDa CryD protein described here, had mosquito toxicity (15, 17, 36).

The CryD protein could not be detected in B. megaterium cells harboring the cloned cryDA76 allele. It is possible that the truncated form of CryD encoded by the cryDA76 strain (missing 76 COOH-terminal amino acids) is unstable and rapidly degraded. Another possibility is that the 10-nucleotide inverted repeat found at the 3’ end of the cryD gene and missing from cryDA76 is necessary for stabilization of the cryD mRNA. Inverted repeats located at the 3’ end of the maIE gene of E. coli have been shown to stabilize the maIE mRNA (19), and Wong and Chang (35) have demonstrated that an inverted repeat found at the 3’ end of a 130-kDa lepidopteran-toxic crystal protein gene from B. thuringiensis subsp. kurstaki served to stabilize the upstream mRNA.

The cloned 11-kb HindIII fragment containing the cryD gene was also found to contain the gene (cryE) for the 28-kDa crystal protein of B. thuringiensis subsp. israelensis. McLean and Whiteley (18) reported that expression in E. coli of a cloned B. thuringiensis subsp. israelensis gene for a 28-kDa protein required a 0.8-kb segment of DNA that was located approximately 4 kb upstream from the cloned gene. The cloned gene is most likely the same as the cryE gene described in this report. Our data indicate that the 0.8-kb segment of DNA should be located immediately downstream from the cryD gene. B. megaterium cells harboring the cloned B. thuringiensis subsp. israelensis 5.7-kb EcoRI fragment (cryDA76 cryE+), which lacks the 3’ end of the cryD gene as well as DNA sequences downstream from the cryD gene, synthesized apparently identical amounts of the 28-kDa CryE protein as B. megaterium cells harboring the cloned 11-kb HindIII fragment (cryD+ cryE+), which contains the complete cryD gene and approximately 3 kb of DNA downstream from the cryD gene. Therefore, unlike expression in E. coli, expression of the cryE gene in B. megaterium does not appear to require DNA sequences downstream from the cryD gene. Nevertheless, cryE-containing B. megaterium cells produced very little of the CryE protein. Ward et al. (33) have reported that a cloned B. thuringiensis subsp. israelensis gene encoding a 27-kDa crystal protein, most likely identical to the CryE protein reported here, was highly expressed in Bacillus subtilis cells. This finding, plus the fact that B. thuringiensis subsp. israelensis synthesizes large amounts of the CryE protein, makes the low level of synthesis of this protein by cryE-containing B. megaterium cells somewhat puzzling.

B. megaterium cells harboring multiple copies of cryD and cryE were significantly inhibited in their ability to form spores. We had previously found that B. megaterium cells harboring multiple copies of either the cloned lepidopteran- and dipteran-toxic cryB1 crystal protein gene from B. thuringiensis subsp. kurstaki (6) or the cloned coleopteran-toxic cryC crystal protein gene from B. thuringiensis EG2158 (Donovan et al., in press) were severely or moderately inhibited, respectively, in their ability to form spores. A possible explanation for the observed inhibition, suggested by previous findings with cloned B. subtilis sporulation-specific genes (21, 38), is that the promoters for the crystal genes, when present in the cell in high copy number, titrate a transcription factor necessary for sporulation. Furthermore, the finding that the cryD and cryE genes were apparently not expressed in E. coli cells suggests that these genes require some sporulation-specific transcription factor(s) for efficient expression.

The dipteran-toxic CryD protein was found to contain partial sequence homology with two B. thuringiensis crystal proteins, the lepidopteran- and dipteran-toxic CryB1 protein (6) and the coleopteran-toxic CryC protein (12, 26; Donovan et al., in press). The CryD protein was not homologous with other B. thuringiensis subsp. israelensis proteins that have been implicated in dipteran toxicity, the 130-kDa crystal protein (31) and the 28-kDa crystal protein (30). Knowledge of the presence or absence of such homologies may be useful in understanding the mode of action of these toxins.

ACKNOWLEDGMENTS

We thank W. Lane of the Harvard University Microchemistry Facility for protein NH2-terminal amino acid sequence determinations. C. Gawron-Burke, B. Carlton, and J. McIntyre of Ecogen Inc. for critical reading of the manuscript, K. Johnson and T. Johnson of Ecogen Inc. for insect toxicity assays, and J. M. Gonzalez, Jr., of Ecogen Inc. for providing derivatives of strain HD-567.

LITERATURE CITED

acid sequence of a coleopteran-active delta-endotoxin gene from Bacillus thuringiensis subsp. san diego. Gene 57:37–46.


28. Visser, B., M. van Workum, A. Dulleman, and C. Waalwijk. 1986. The mosquitoicidal activity of Bacillus thuringiensis var. israelensis is associated with Mr 230,000 and 130,000 crystal proteins. FEMS Microbiol. Lett. 30:211–214.


