JOURNAL OF BACTERIOLOGY, Oct. 1988, p. 4732–4738
0021-9193/88/104732-07$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 170, No. 10

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

WILLIAM P. DONOVAN,* CATHY DANKOCSIK, 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 bipyramidal 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 rhomboid 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 mosquitocidal activity (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 NH_2-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-^{32}P]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

* Corresponding author.

```
        1   2   3   4   5   6   7   8   9   10
NH2- MET GLU ASP SER SER LEU ASP THR LEU SER

        11  12  13  14  15  16  17  18  19  20
        ILE VAL ASN GLU THR ASP PHE PRO LEU TYR
5' - ATT GTA AAT GAA ACA GAT TTT CCA TTA TAT

        21  22  23  24  25  26
        ASN ASN TYR THR GLU PRO- COOH
        AAT AAT TAT ACA GAA CC - 3'
```

FIG. 1. NH_2-terminal amino acid sequence of the 72-kDa protein and the sequence of the 47-mer *cryD*-specific oligonucleotide probe.

7.0) containing ampicillin (40 μg/ml). *B. megaterium* and *B. thuringiensis* subsp. *israelensis* cells were grown for 48 h at 30°C on DS agar plates (0.8% nutrient broth [Difco], 13 mM KCl, 1 mM $Ca(NO_3)_2$, 0.5 mM $MgSO_4$, 10 μM $MnCl_2$, 10 μM $FeSO_4$, 1.5% agar, pH 7) containing either tetracycline (10 μg/ml) or no antibiotic, respectively. After growth, cells were removed from the agar surface with a spatula, washed with deionized water, and suspended in deionized water at a concentration of 100 mg of cells (wet weight) per ml. *B. megaterium* and *B. thuringiensis* subsp. *israelensis* cell suspensions were mixed with an equal volume of 100 mM Tris (pH 7)–20 mM EDTA–10 mg of lysozyme per ml and incubated at 37°C for 60 min. SDS was added to a final concentration of 0.2%, and the mixtures were vortexed and centrifuged for 7 min in a Du Pont microfuge. The pelleted material, consisting of spores and insoluble material from lysed cells including insoluble crystal proteins, was suspended in 0.1% SDS–10 mM EDTA. A measured volume (usually 10 μl) of the lysozyme-treated suspensions and of the untreated *E. coli* cell suspensions were added to 3 volumes of preheated gel loading buffer (2% SDS, 5% 2-mercaptoethanol, 130 mM Tris hydrochloride [pH 6.8], 10% glycerol, 0.05% bromophenol blue) in a 0.5-ml microtube, incubated at 90°C for 7 min, and vortexed for 10 s, and measured volumes (usually 10 μl) of the mixtures were electrophoresed through an SDS gel.

DNA sequencing. DNA fragments were cloned into the M13 vectors mp18 and mp19 as described below. A total of 21 *cryD*-specific 17-mer primers (synthesized on an Applied Biosystems DNA synthesizer, model 380B) and one M13-specific primer (supplied by Bethesda Research Laboratories) were used to determine the complete DNA sequence of both strands of the *cryD* gene by the dideoxy method (23).

Toxicity measurements. *E. coli* cells were grown on LB agar plates containing ampicillin (40 μg/ml). *B. megaterium* and *B. thuringiensis* subsp. *israelensis* cells were grown on DS agar sporulation plates containing either tetracycline (10 μg/ml) or no antibiotic, respectively. After 48 h of growth at 30°C, cells (stationary phase or sporulated) were harvested, suspended in lysis buffer (1 mg of lysozyme per ml, 10 mM Tris hydrochloride [pH 7.5], 1 mM EDTA) and incubated at 37°C for 1 h. SDS was added to a final concentration of 0.1%, and the mixtures, consisting of free crystal proteins plus lysed cells and/or spores, were serially diluted. Dilutions were added to 50 ml of deionized water containing 20 *Aedes aegypti* fourth-instar larvae, and mortality was scored after 24 h. The 50% lethal dose (LD_{50}) values were determined by probit analysis with an eight-dose testing procedure and at least 60 larvae at each dose.

RESULTS

Isolation of the *cryD* gene. The 72-kDa crystal protein was purified from *B. thuringiensis* subsp. *israelensis* HD-567 by electroelution of the protein from SDS gels. Edman analysis of the purified protein yielded the NH_2-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 NH_2-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 *Hin*dIII restriction fragment of approximately 11 kilobases (kb) and to a unique *Eco*RI 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 *Eco*RI restriction fragments of HD-567 DNA into the *Eco*RI 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 *Eco*RI insert (Fig. 2, pEG214). The probe specifically hybridized to the 5.7-kb *Eco*RI fragment of pEG214 and also to a 1.1-kb *Dra*I fragment of pEG214 (Fig. 2). Sequencing of the 1.1-kb fragment revealed a long open reading frame that began with the NH_2-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 *Pvu*II-*Eco*RI fragment from pEG214 revealed that the *cryD* open reading frame extended through the *Eco*RI site. Therefore, the 3′ end of the *cryD* gene was not contained on the 5.7-kb *Eco*RI fragment.

The 5.7-kb *Eco*RI fragment hybridized, as expected, to an approximately 11-kb *Hin*dIII 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 *Hin*dIII 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*-*Pvu*II fragment and the 0.8-kb *Eco*RI 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 *Eco*RI fragment. This result indicates that plasmid pEG214 (5.7-kb *Eco*RI) contained a truncated form of the *cryD* gene, designated *cryD*Δ*76*, 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 NH_2-terminal methionine, a purine-rich sequence (AAAGGTGG) was found that probably serves as a ribosome-binding site. A 10-nucleotide inverted repeat ($\Delta G = -15.0$ 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 *Hin*dIII fragment (Fig. 2, pEG216) was similar to the restriction map recently reported by McLean and Whiteley (18) of a 10.7-kb *Hin*dIII 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 *Bam*HI-*Pvu*II fragment was subcloned from

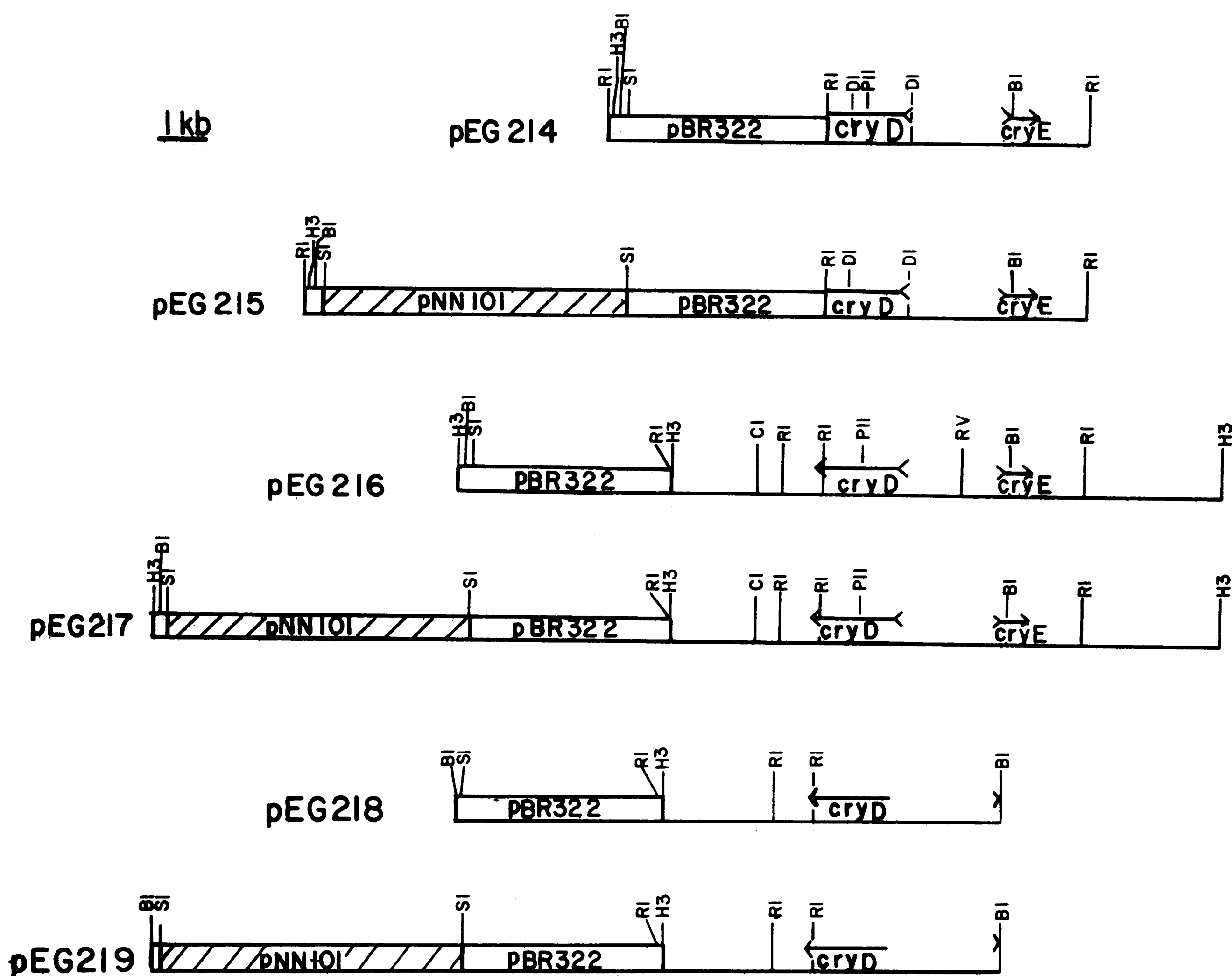

FIG. 2. Restriction maps of the *cryD*- and *cryE*-containing plasmids pEG214, pEG216, and pEG218 (*E. coli*) and pEG215, pEG217, and pEG219 (*E. coli-Bacillus*). The location and orientation of the *cryD* and *cryE* genes are indicated by arrows. Restriction sites: R1, *Eco*RI; H3, *Hin*dIII; B1, BamHI; S1, *Sal*I; D1, *Dra*I; Cl, *Cla*I; PII, *Pvu*II; and RV, *Eco*RV. Not all restriction sites are shown for each plasmid.

pEG216 (11.0-kb *Hin*dIII) and partially sequenced. The 2.6-kb fragment contained an open reading frame that was identical to at least the first 21 NH_2-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 *Hin*dIII fragment. Figure 2 (pEG216) shows that the gene for the 28-kDa protein, which we have designated *cryE*, 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 (*cryD*Δ*76 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). In order to measure the expression of the cloned *cryD* and *cryE* genes in a *Bacillus* species, the plasmids pEG215 (*cryD*Δ*76 cryE*$^+$) and pEG217 (*cryD*$^+$ *cryE*$^+$), capable of replicating in both *E. coli* and *B. megaterium*, were constructed by inserting the *Bacillus* vector pNN101 (Cam^r) into the unique *Sph*I sites of pEG214 and pEG216, respectively (Fig. 2). *B. megaterium* cells harboring pEG215 (*cryD*Δ*76 cryE*$^+$) (strain EG1324) did not contain detectable levels of a plasmid-encoded 72-kDa protein but did contain 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 *Sph*I 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 *Bam*HI-*Hin*dIII fragment from the 11.0-kb *Hin*dIII fragment into pBR322. The 6.7-kb fragment contained the complete *cryD* gene but only 21 NH_2-terminal codons of the *cryE* gene. The *Bacillus* vector pNN101 was inserted into the *Sph*I 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

```
        10        20        30        40        50        60
TTTAAAATAAAAAAATTCAATAAAAGGTGGAATGAATTATATGGAAGATAGTTCTTTAGA
                                          MetGluAspSerSerLeuAs

        70        8C        90       100       110       120
TACTTTAAGTATAGTTAATGAAACAGACTTTCCATTATATAATAATTATACCGAACCTAC
pThrLeuSerIleValAsnGluThrAspPheProLeuTyrAsnAsnTyrThrGluProTh

       130       140       150       160       170       180
TATTGCGCCAGCATTAATAGCAGTAGCTCCCATCGCACAATATCTTGCAACAGCTATAGG
rIleAlaProAlaLeuIleAlaValAlaProIleAlaGlnTyrLeuAlaThrAlaIleGl

       190       200       210       220       230       240
GAAATGGGCGGCAAAGGCAGCATTTTCAAAAGTACTATCACTTATATTCCCAGGTTCTCA
yLysTrpAlaAlaLysAlaAlaPheSerLysValLeuSerLeuIlePheProGlySerGl

       250       260       270       280       290       300
ACCTGCTACTATGGAAAAAGTTCGTACAGAAGTGGAAACACTTATAAATCAAAAATTAAG
nProAlaThrMetGluLysValArgThrGluValGluThrLeuIleAsnGlnLysLeuSe

       310       320       330       340       350       360
CCAAGATCGAGTCAATATATTAAACGCAGAATATAGGGGGATTATTGAGGTTAGTGATGT
rGlnAspArgValAsnIleLeuAsnAlaGluTyrArgGlyIleIleGluValSerAspVa

       370       380       390       400       410       420
ATTTGATGCGTATATTAAACAACCAGGTTTTACCCCTGCAACAGCCAAGGGTTATTTTCT
lPheAspAlaTyrIleLysGlnProGlyPheThrProAlaThrAlaLysGlyTyrPheLe

       430       440       450       460       470       480
AAATCTAAGTGGTGCTATAATACAACGATTACCTCAATTTGAGGTTCAAACATATGAAGG
uAsnLeuSerGlyAlaIleIleGlnArgLeuProGlnPheGluValGlnThrTyrGluGl

       490       500       510       520       530       540
AGTATCTATAGCACTTTTTACTCAAATGTGTACACTTCATTTAACTTTATTAAAAGACGG
yValSerIleAlaLeuPheThrGlnMetCysThrLeuHisLeuThrLeuLeuLysAspGl

       550       560       570       580       590       600
AATCCTAGCAGGGAGTGCATGGGGATTTACTCAAGCTGATGTAGATTCATTTATAAAATT
yIleLeuAlaGlySerAlaTrpGlyPheThrGlnAlaAspValAspSerPheIleLysLe

       610       620       630       640       650       660
ATTTAATCAAAAAGTATTAGATTACAGGACCAGATTAATGAGAATGTACACAGAAGAGTT
uPheAsnGlnLysValLeuAspTyrArgThrArgLeuMetArgMetTyrThrGluGluPh

       670       680       690       700       710       720
CGGAAGATTGTGTAAAGTCAGTCTTAAAGATGGATTGACGTTCCGGAATATGTGTAATTT
eGlyArgLeuCysLysValSerLeuLysAspGlyLeuThrPheArgAsnMetCysAsnLe

       730       740       750       760       770       780
ATATGTGTTTCCATTTGCTGAAGCCTGGTCTTTAATGAGATATGAAGGATTAAAATTACA
uTyrValPheProPheAlaGluAlaTrpSerLeuMetArgTyrGluGlyLeuLysLeuGl

       790       800       810       820       830       840
AAGCTCTCTATCATTATGGGATTATGTTGGTGTCTCAATTCCTGTAAATTATAATGAATG
nSerSerLeuSerLeuTrpAspTyrValGlyValSerIleProValAsnTyrAsnGluTr

       850       860       870       880       890       900
GGGAGGACTAGTTTATAAGTTATTAATGGGGGAAGTTAATCAAAGATTAACAACTGTTAA
pGlyGlyLeuValTyrLysLeuLeuMetGlyGluValAsnGlnArgLeuThrThrValLy

       910       920       930       940       950       960
ATTTAATTATTCTTTCACTAATGAACCAGCTGATATACCAGCAAGAGAAAATATTCGTGG
sPheAsnTyrSerPheThrAsnGluProAlaAspIleProAlaArgGluAsnIleArgGl

       970       980       990      1000      1010      1020
CGTCCATCCTATATACGATCCTAGTTCTGGGCTTACAGGATGGATAGGAAACGGAAGAAC
yValHisProIleTyrAspProSerSerGlyLeuThrGlyTrpIleGlyAsnGlyArgTh

      1030      1040      1050      1060      1070      1080
AAACAATTTTAATTTTGCTGATAACAATGGCAATGAAATTATGGAAGTTAGAACACAAAC
rAsnAsnPheAsnPheAlaAspAsnAsnGlyAsnGluIleMetGluValArgThrGlnTh

      1090      1100      1110      1120      1130      1140
TTTTTATCAAAATCCAAATAATGAGCCTATAGCGCCTAGAGATATTATAAATCAAATTTT
rPheTyrGlnAsnProAsnAsnGluProIleAlaProArgAspIleIleAsnGlnIleLe

      1150      1160      1170      1180      1190      1200
AACTGCGCCAGCACCAGCAGACCTATTTTTTAAAAATGCAGATATAAATGTAAAGTTCAC
uThrAlaProAlaProAlaAspLeuPhePheLysAsnAlaAspIleAsnValLysPheTh

      1210      1220      1230      1240      1250      1260
ACAGTGGTTTCAGTCTACTCTATATGGGTGGAACATTAAACTCGGTACACAAACGGTTTT
rGlnTrpPheGlnSerThrLeuTyrGlyTrpAsnIleLysLeuGlyThrGlnThrValLe

      1270      1280      1290      1300      1310      1320
AAGTAGTAGAACCGGAACAATACCACCAAATTATTTAGCATATGATGGATATTATATTCG
uSerSerArgThrGlyThrIleProProAsnTyrLeuAlaTyrAspGlyTyrTyrIleAr

      1330      1340      1350      1360      1370      1380
TGCTATTTCAGCTTGCCCAAGAGGAGTCTCACTTGCATATAATCACGATCTTACAACACT
gAlaIleSerAlaCysProArgGlyValSerLeuAlaTyrAsnHisAspLeuThrThrLe

      1390      1400      1410      1420      1430      1440
AACATATAATAGAATAGAGTATGATTCACCTACTACAGAAAATATTATTGTAGGGTTTGC
uThrTyrAsnArgIleGluTyrAspSerProThrThrGluAsnIleIleValGlyPheAl

      1450      1460      1470      1480      1490      1500
ACCAGATAATACTAAGGACTTTTATTCTAAAAAATCTCACTATTTAAGTGAAACGAATGA
aProAspAsnThrLysAspPheTyrSerLysLysSerHisTyrLeuSerGluThrAsnAs

      1510      1520      1530      1540      1550      1560
TAGTTATGTAATTCCTGCTCTGCAATTTGCTGAAGTTTCAGATAGATCATTTTTAGAAGA
pSerTyrValIleProAlaLeuGlnPheAlaGluValSerAspArgSerPheLeuGluAs

      1570      1580      1590      1600      1610      1620
TACGCCAGATCAAGCAACAGACGGCAGTATTAAATTTGCACGTACTTTCATTAGTAATGA
pThrProAspGlnAlaThrAspGlySerIleLysPheAlaArgThrPheIleSerAsnGl

      1630      1640      1650      1660      1670      1680
AGCTAAGTACTCTATTAGACTAAACACCGGGTTTAATACGGCAACTAGATATAAATTAAT
uAlaLysTyrSerIleArgLeuAsnThrGlyPheAsnThrAlaThrArgTyrLysLeuIl

      1690      1700      1710      1720      1730      1740
TATCAGGGTAAGAGTACCTTATCGCTTACCTGCTGGAATACGGGTACAATCTCAGAATTC
eIleArgValArgValProTyrArgLeuProAlaGlyIleArgValGlnSerGlnAsnSe

      1750      1760      1770      1780      1790      1800
GGGAAATAATAGAATGCTAGGCAGTTTTACTGCAAATGCTAATCCAGAATGGGTGGATTT
rGlyAsnAsnArgMetLeuGlySerPheThrAlaAsnAlaAsnProGluTrpValAspPh

      1810      1820      1830      1840      1850      1860
TGTCACAGATGCATTTACATTTAACGATTTAGGGATTACAACTTCAAGTACAAATGCTTT
eValThrAspAlaPheThrPheAsnAspLeuGlyIleThrThrSerSerThrAsnAlaLe

      1870      1880      1890      1900      1910      1920
ATTTAGTATTTCTTCAGATAGTTTAAATTCTGGAGAAGAGTGGTATTTATCGCAGTTGTT
uPheSerIleSerSerAspSerLeuAsnSerGlyGluGluTrpTyrLeuSerGlnLeuPh

      1930      1940      1950      1960      1970      1980
TTTAGTAAAAGAATCGGCCTTTACGACGCAAATTAATCCGTTACTAAAGTAGAAGTCATG
eLeuValLysGluSerAlaPheThrThrGlnIleAsnProLeuLeuLysEnd

      1990      2000      2010      2020      2030      2040
TTAGCACAAGAGGAGTGAGTATTGTGGCTCCTCTTGTAATTTTAATCGCTAATATTTCTA
     ---------->              <---------

      2050      2060      2070      2080      2090      2100
ATAGATATAAATTATATATAATATTTAAAAAGTTATAATTATGTAATTGTAGAAAATCAT
```

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

expected, no 28-kDa CryE protein (Fig. 4, lane 5). EG1323 cells had an LD_{50} (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 EG1323 and EG1316 cells (estimated from Coomassie stained SDS gels as in Fig. 4). Therefore, the LD_{50} value for the CryD protein against *Aedes aegypti* larvae was approximately 0.03 to 0.01 μg of protein per ml (0.005 × LD_{50} cells, wet weight).

EG1316(pEG217 *cryD*$^+$ *cryE*$^+$) and EG1323(pEG219 *cryD*$^+$ *cryE21*) 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(pEG215 *cryD*Δ*76* *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.

TABLE 1. Larvicidal activities of *cryD*- and *cryE*-containing strains

Strain	Plasmid(s)	LD_{50} (μg of cells/ml)[a]
B. thuringiensis subsp. *israelensis* HD-567	135, 105, 75, 68, 10.6, 4.9, 4.2, and 3.3 MDa	0.2–0.05
E. coli		
EG1315	pEG216 (*cryD*$^+$ *cryE*$^+$)	>60
EG1318	pEG214 (*cryD*Δ*76* *cryE*$^+$)	>60
B. megaterium		
EG1325	pEG220 (Cry$^-$)	>50
EG1324	pEG215 (*cryD*Δ*76* *cryE*$^+$)	>20
EG1316	pEG217 (*cryD*$^+$ *cryE*$^+$)	6–3
EG1323	pEG219 (*cryD*$^+$ *cryE21*)	7–3

[a] Values with > indicate that the strain showed no toxicity at the doses tested.

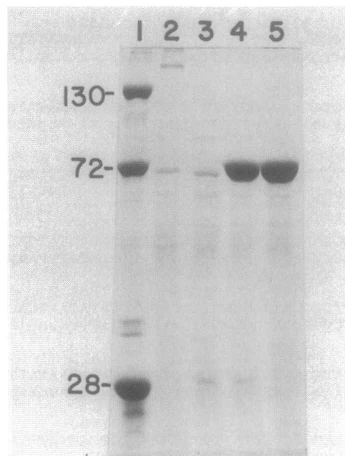

FIG. 4. Proteins synthesized by *cryD*- and *cryE*-containing bacterial strains. A Coomassie-stained SDS gel is shown. Each lane contains insoluble protein, as described in Materials and Methods, extracted from 700 μg (wet weight) of cells. Lane 1, *B. thuringiensis* subsp. *israelensis* HD-567. Lanes 2–5, Isogenic *B. megaterium* strains harboring various plasmids: lane 2, EG1325(pEG220 Cry$^-$); lane 3, EG1324(pEG215 *cryD*Δ*76 cryE*$^+$); lane 4, EG1316(pEG217 *cryD*$^+$ *cryE*$^+$); and lane 5, EG1323(pEG219 *cryD*$^+$ *cryE21*). Numbers indicate the 130-kDa, 72-kDa, and 28-kDa crystal proteins.

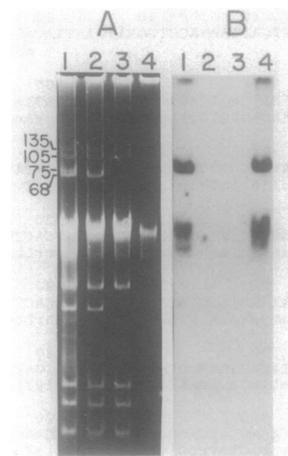

FIG. 5. Hybridization of the *cryD* gene to a 75-MDa plasmid. (A) Agarose gel displaying plasmids from derivatives of strain HD-567. Lane 1, HD-567; lane 2, HD-567 cured of the 75-MDa plasmid; lane 3, HD-567 cured of both the 75- and the 68-MDa plasmids; lane 4, a derivative of the strain shown in lane 3 after receiving the 75-MDa plasmid, by conjugation, from HD-567. (B) Southern blot of panel A probed with the radioactively labeled 2.7-kb *Eco*RV-*Eco*RI *cryD* fragment. The hybridizing DNA in the upper portion of the gel is the 75-MDa plasmid, and the hybridizing DNA in the lower portion of the gel is the diffuse band referred to in the text. Numbers indicate approximate plasmid sizes in megadaltons.

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 *Eco*RV-*Eco*RI *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 transcipient 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* subspp. *tenebrionis* and *san diego* (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

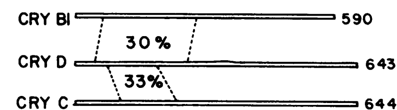

FIG. 6. Regions of sequence homology between the 72-kDa CryD protein and two *B. thuringiensis* crystal proteins. CryB1 is the 66-kDa lepidopteran- and dipteran-toxic protein of *B. thuringiensis* subsp. *kurstaki* (6). CryC is the 73-kDa coleopteran-toxic protein of *B. thuringiensis* EG2158 (Donovan et al., in press). Percentages and dashed lines denote the amount of homology and the extent of homology, respectively. Numbers to the right indicate the protein size in amino acids.

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 *cryD*Δ*76* allele. It is possible that the truncated form of CryD encoded by the *cryD*Δ*76* 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 *cryD*Δ*76* is necessary for stabilization of the *cryD* mRNA. Inverted repeats located at the 3′ end of the *malE* gene of *E. coli* have been shown to stabilize the *malE* 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 *Hin*dIII 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 *Eco*RI fragment (*cryD*Δ*76 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 *Hin*dIII 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 NH_2-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

1. **Angsuthanasombat, C., W. Chungjatupornchai, S. Kertbundit, P. Luxananil, C. Settasatian, P. Wilairat, and S. Panyim.** 1987. Cloning and expression of 130-kd mosquito-larvicidal delta-endotoxin gene of *Bacillus thuringiensis* var. *israelensis* in *Escherichia coli*. Mol. Gen. Genet. **208:**384–389.
2. **Aronson, A. I., W. Beckman, and P. Dunn.** 1986. *Bacillus thuringiensis* and related insect pathogens. Microbiol. Rev. **50:** 1–24.
3. **Bourgouin, C., A. Klier, and G. Rapoport.** 1986. Characterization of the genes encoding the haemolytic toxin and the mosquitocidal delta-endotoxin of *Bacillus thuringiensis israelensis*. Mol. Gen. Genet. **205:**390–397.
4. **Cheung, P. Y. K., D. Buster, and B. D. Hammock.** 1987. Lack of mosquitocidal activity by the cytolytic protein of the *Bacillus thuringiensis* subsp. *israelensis* parasporal crystal. Curr. Microbiol. **15:**21–23.
5. **Cheung, P. Y. K., and B. D. Hammock.** 1985. Separation of three biologically distinct activities from the parasporal crystal of *Bacillus thuringiensis* var. *israelensis*. Curr. Microbiol. **12:** 121–126.
6. **Donovan, W. P., C. C. Dankocsik, M. P. Gilbert, M. C. Gawron-Burke, R. G. Groat, and B. C. Carlton.** 1988. Amino acid sequence and entomocidal activity of the P2 crystal protein, an insect toxin from *Bacillus thuringiensis* var. *kurstaki*. J. Biol. Chem. **263:**561–567.
7. **Eckhardt, T.** 1978. A rapid method for the identification of plasmid DNA in bacteria. Plasmid **1:**584–588.
8. **Goldberg, L. J., and J. Margalit.** 1977. A bacterial spore demonstrating rapid larvicidal activity against *Anopheles sergentii*, *Uranotaenia unguiculata*, *Culex univitattus*, *Aedes aegypti* and *Culex pipiens*. Mosquito News **37:**355–358.
9. **Gonzalez, J. M., Jr., and B. C. Carlton.** 1982. Plasmid transfer in *Bacillus thuringiensis*, p. 85–95. *In* U. N. Streips, S. H. Goodgal, W. R. Guild, and G. A. Wilson (ed.), Genetic exchange: a celebration and a new generation. Marcel Dekker, New York.
10. **Gonzalez, J. M., Jr., and B. C. Carlton.** 1984. A large transmissible plasmid is required for crystal toxin production in *Bacillus thuringiensis* var. *israelensis*. Plasmid **11:**28–38.
11. **Held, G. A., Y. Huang, and C. Y. Kawanishi.** 1986. Effect of removal of the cytolytic factor of *Bacillus thuringiensis* subsp. *israelensis* on mosquito toxicity. Biochem. Biophys. Res. Commun. **141:**937–941.
12. **Herrnstadt, C., T. E. Gilroy, D. A. Sobieski, B. D. Bennett, and F. H. Gaertner.** 1987. Nucleotide sequence and deduced amino

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

13. **Hunkapiller, M. W., E. Lujan, F. Ostrander, and L. E. Hood.** 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. **91**:227–236.
14. **Hurley, J. M., L. A. Bulla, Jr., and R. E. Andrews, Jr.** 1987. Purification of the mosquitocidal and cytolytic proteins of *Bacillus thuringiensis* subsp. *israelensis*. Appl. Environ. Microbiol. **53**:1316–1321.
15. **Hurley, J. M., S. G. Lee, R. E. Andrews, Jr., M. J. Klowden, and L. A. Bulla, Jr.** 1985. Separation of the cytolytic and mosquitocidal proteins of *Bacillus thuringiensis* subsp. *israelensis*. Biochem. Biophys. Res. Commun. **126**:961–965.
16. **Krieg, V. A., A. M. Huger, G. A. Langenbruch, and W. Schnetter.** 1983. *Bacillus thuringiensis* var. *tenebrionis*, a new pathotype effective against larvae of Coleoptera. Z. Angew. Entomol. **96**:500–508.
17. **Lee, S. G., W. Eckblad, and L. A. Bulla, Jr.** 1985. Diversity of protein inclusion bodies and identification of mosquitocidal protein in *Bacillus thuringiensis* subsp. *israelensis*. Biochem. Biophys. Res. Commun. **126**:953–960.
18. **McLean, K. M., and H. R. Whiteley.** 1987. Expression in *Escherichia coli* of a cloned crystal protein gene of *Bacillus thuringiensis* subsp. *israelensis*. J. Bacteriol. **169**:1017–1023.
19. **Newbury, S. F., N. H. Smith, E. C. Robinson, I. D. Hiles, and C. F. Higgins.** 1987. Stabilization of translationally active mRNA by prokaryotic REP sequences. Cell **48**:297–310.
20. **Norton, N. B., K. A. Orzech, and W. F. Burke, Jr.** 1985. Construction and characterization of plasmid vectors for cloning in the entomocidal organism *Bacillus sphaericus* 1593. Plasmid **13**:211–214.
21. **Piggot, P. J., and C. A. M. Curtis.** 1987. Analysis of the regulation of gene expression during *Bacillus subtilis* sporulation by manipulation of the copy number of *spo-lacZ* fusions. J. Bacteriol. **169**:1260–1266.
22. **Queen, C., and L. J. Korn.** 1984. Analysis of biological sequences on small computers. DNA **3**:421–436.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Schnepf, H. E., H. C. Wong, and H. R. Whiteley.** 1985. The amino acid sequence of a crystal protein from *Bacillus thuringiensis* deduced from the DNA base sequence. J. Biol. Chem. **260**:6264–6272.
25. **Sekar, V., and B. C. Carlton.** 1985. Molecular cloning of the delta-endotoxin gene of *Bacillus thuringiensis* var. *israelensis*. Gene **33**:151–158.
26. **Sekar, V., D. V. Thompson, M. J. Maroney, R. G. Bookland, and M. J. Adang.** 1987. Molecular cloning and characterization of the insecticidal crystal protein gene of *Bacillus thuringiensis* var. *tenebrionis*. Proc. Natl. Acad. Sci. USA **84**:7036–7040.
27. **Thorne, L., F. Garduno, T. Thompson, D. Decker, M. Zounes, M. Wild, A. M. Walfield, and T. J. Pollock.** 1986. Structural similarity between the lepidoptera- and diptera-specific insecticidal endotoxin genes of *Bacillus thuringiensis* subsp. "*kurstaki*" and "*israelensis*." J. Bacteriol. **166**:801–811.
28. **Visser, B., M. van Workum, A. Dullemans, and C. Waalwijk.** 1986. The mosquitocidal activity of *Bacillus thuringiensis* var. *israelensis* is associated with Mr 230,000 and 130,000 crystal proteins. FEMS Microbiol. Lett. **30**:211–214.
29. **Von Tersch, M. A., and B. C. Carlton.** 1984. Molecular cloning of structural and immunity genes for megacins A-216 and A-19213 in *Bacillus megaterium*. J. Bacteriol. **160**:854–859.
30. **Waalwijk, C., A. M. Dullemans, M. E. S. van Workum, and B. Visser.** 1985. Molecular cloning and the nucleotide sequence of the Mr 28 000 crystal protein gene of *Bacillus thuringiensis* subsp. *israelensis*. Nucleic Acids Res. **13**:8207–8217.
31. **Ward, E. S., and D. J. Ellar.** 1987. Nucleotide sequence of a *Bacillus thuringiensis* var. *israelensis* gene encoding a 130 kDa delta-endotoxin. Nucleic Acids Res. **15**:7195.
32. **Ward, E. S., and D. J. Ellar.** 1988. Cloning and expression of two homologous genes of *Bacillus thuringiensis* subsp. *israelensis* which encode 130-kilodalton mosquitocidal proteins. J. Bacteriol. **170**:727–735.
33. **Ward, E. S., A. R. Ridley, D. J. Ellar, and J. A. Todd.** 1986. *Bacillus thuringiensis* var. *israelensis* delta-endotoxin: cloning and expression of the toxin in sporogenic and asporogenic strains of *Bacillus subtilis*. J. Mol. Biol. **191**:13–22.
34. **Whiteley, H. R., and H. E. Schnepf.** 1986. The molecular biology of parasporal crystal body formation in *Bacillus thuringiensis*. Annu. Rev. Microbiol. **40**:549–576.
35. **Wong, H. C., and S. Chang.** 1986. Identification of a positive retroregulator that stabilizes mRNAs in bacteria. Proc. Natl. Acad. Sci. USA **83**:3233–3237.
36. **Wu, D., and F. N. Chang.** 1985. Synergism in mosquitocidal activity of 26 and 65 kDa proteins from *Bacillus thuringiensis* subsp. *israelensis* crystal. FEBS Lett. **190**:232–236.
37. **Yamamoto, T., and R. E. McLaughlin.** 1981. Isolation of a protein from the parasporal crystal of *Bacillus thuringiensis* var. *kurstaki* toxic to the mosquito larva, *Aedes taeniorhynchus*. Biochem. Biophys. Res. Commun. **103**:414–421.
38. **Zuber, P., J. M. Healy, and R. Losick.** 1987. Effects of plasmid propagation of a sporulation promoter on promoter utilization and sporulation in *Bacillus subtilis*. J. Bacteriol. **169**:461–469.