JOURNAL OF BACTERIOLOGY, Jan. 1989, p. 521–530
0021-9193/89/010521-10$02.00/0

Vol. 171, No. 1

A 20-Kilodalton Protein Is Required for Efficient Production of the *Bacillus thuringiensis* subsp. *israelensis* 27-Kilodalton Crystal Protein in *Escherichia coli*

LEE F. ADAMS, JONATHAN E. VISICK, AND H. R. WHITELEY*

Department of Microbiology, University of Washington, Seattle, Washington 98195

Received 18 August 1988/Accepted 20 October 1988

The 27-kilodalton (kDa) mosquitocidal protein gene from *Bacillus thuringiensis* subsp. *israelensis* has been cloned as a 10-kilobase (kb) *Hin*dIII fragment from plasmid DNA; efficient expression in *Escherichia coli* KM1 depends on a region of DNA located approximately 4 kb upstream (K. McLean and H. R. Whiteley, J. Bacteriol. 169:1017–1023, 1987). We have cloned the upstream DNA region and show that it contains a complete open reading frame (ORF) encoding a protein with a molecular mass of 19,584 Da. Sequencing of adjacent stretches of DNA revealed two partial ORFs: one has 55.2% identity in an overlap of 319 amino acids to the putative transposase of IS*231* of *B. thuringiensis* subsp. *thuringiensis*, and the other, a 78-codon partial ORF, may be the carboxyl terminus of the 67-kDa protein previously observed in maxicells of strain KM1. A 0.8-kb fragment containing only the 20-kDa protein gene greatly enhanced the expression of the 27-kDa protein in *E. coli*. The introduction of nonsense codons into the 20-kDa protein gene ORF abolished this effect, indicating that the gene product, not the mRNA or DNA, is required for the enhancement. The effect of the 20-kDa protein gene on various fusions of *lacZ* to the 27-kDa protein gene suggests that the 20-kDa protein acts after the initiation of translation of the 27-kDa protein gene.

During sporulation, *Bacillus thuringiensis* subsp. *israelensis* produces a parasporal inclusion that exhibits specific toxicity to nematocerous dipteran larvae such as mosquitoes and black flies (8, 10). Electron microscopy shows that the inclusions contain three different components (6, 14). Electrophoresis of the solubilized inclusions reveals four major polypeptides with molecular masses of 135, 128, 67, and 27 kilodaltons (kDa) (2, 35). These polypeptides are glycosylated (27) and are mosquitocidal; the 27-kDa peptide is also hemolytic and cytolytic (32). Many studies have also reported the presence of several minor crystal components: a 230-kDa species (probably an aggregate of the 135- or 128-kDa peptide or both; 38), a toxic 58-kDa polypeptide (derived from a 72-kDa peptide; 33), a 54-kDa peptide (probably a dimer of the 27-kDa peptide; 7), and a nontoxic 38-kDa peptide (probably a breakdown product of the 67-kDa peptide; 7, 19). The genes encoding the 135-, 128-, 72-, and 27-kDa peptides have been cloned from plasmid DNA, and the DNA sequences have been determined (29, 33, 39, 41).

We have previously reported the cloning of a 10-kilobase (kb) *Hin*dIII fragment which encodes the 27-kDa crystal peptide (24). Extracts of the *Escherichia coli* strain harboring the *Hin*dIII fragment were toxic to *Aedes aegypti* larvae and were hemolytic. Substantial levels of the 27-kDa peptide were found in extracts made from *E. coli* cells containing either the entire *Hin*dIII fragment or subclones that included both the crystal protein gene and a region of DNA located 4 kb upstream, but not in extracts of cells that contained the crystal protein gene alone. The upstream region could be present in *cis* or in *trans*, and Tn*5* mutagenesis showed that the required DNA region spanned approximately 0.8 kb. Maxicell analysis demonstrated that the 10-kb fragment encoded peptides of 67, 20, and 16 kDa in addition to the 27-kDa peptide; studies of strains containing subclones of this fragment suggested that the 0.8-kb region encoded the 20-kDa polypeptide. Measurement of the amount of 27-kDa protein gene-specific mRNA in strains containing either the 27-kDa protein gene alone or both the 27-kDa protein gene and the 0.8-kb upstream region indicated that the effect of the latter was posttranscriptional.

To better understand this mechanism of regulation, we have cloned the 0.8-kb DNA fragment required for high-level production of the 27-kDa crystal protein in *E. coli*. The present paper reports the DNA sequences of this fragment and of some of the adjacent DNA. The sequenced region contains a complete open reading frame (ORF) encoding a protein with a predicted molecular mass of 19,584 Da as well as two truncated ORFs. Complementation experiments showed that the presence of the 0.8-kb region in *trans* resulted in a greatly increased amount of 27-kDa peptide. The introduction of nonsense codons into the complete ORF abolished the effect of the fragment, thus proving that the gene product, and not the mRNA or the DNA, is required for the increased level of the 27-kDa peptide in *E. coli* extracts. This paper also presents data on fusions of the 27-kDa protein gene to *lacZ* and discusses the possible role of the 20-kDa peptide in the production of the 27-kDa peptide in *E. coli*.

MATERIALS AND METHODS

Media, bacterial strains, and plasmids. *E. coli* JM83 (36) and MC1000 (4) were maintained on L medium. *E. coli* MV1193 (37) and RZ1032 (*dut ung*; 17) were maintained on M9 medium (25). All fragments from the original 10-kb *Hin*dIII clone (24) were subcloned into pUC118 or pUC119 (37) or into a pACYC184 (5) derivative from which the 1.6-kb *Hin*cII-*Hin*dIII fragment had been removed.

DNA sequencing. Unidirectional nested deletions were created by the method of Henikoff (12). Single-stranded template was prepared by superinfection of insert-containing pUC118 or pUC119 with the kanamycin-resistant helper

* Corresponding author.

phage M13 KO7 (37) and overnight incubation with ampicillin (50 μg ml^{-1}) and kanamycin (70 μg ml^{-1}). Template purification (43) was followed by an additional ethanol precipitation in the presence of 2.5 M ammonium acetate (final concentration). DNA-sequencing reactions were performed according to Sanger et al. (28) using either Klenow fragment (New England BioLabs, Inc., Beverly, Mass.) according to the protocol of Williams et al. (43) or modified T7 polymerase (Sequenase; U.S. Biochemical Corp., Cleveland, Ohio). Reaction mixtures were loaded on 0.4-mm-thick, 40-cm-long 5 or 8% polyacrylamide–8 M urea gels and electrophoresed at 55 W constant power.

Cloning the 20-kDa protein gene into pUC118 and pACYC184. The 2.1-kb *Acc*I-*Acc*I fragment from pKM1 was blunted at both ends by filling in with Klenow fragment and was cloned into the *Sma*I site of pUC118 to generate pLA4. Orientation of the insert in pLA4 was determined by restriction with *Eco*RI. The 2.1-kb insert was then deleted unidirectionally from the 3′ end of the 20-kDa protein gene to 0.8 kb, generating pLA5. To place the 20-kDa protein gene into pACYC184, the 0.8-kb insert in pLA5, containing either the native or the mutated 20-kDa protein gene (see below), was isolated as a *Kpn*I (blunted)-*Hin*dIII fragment and was ligated to the larger, 2.4-kb *Hin*cII-*Hin*dIII fragment from pACYC184 (5), generating pLA15 and pLA15-89.

Detection of gene-specific mRNA. RNA samples were applied as dots on BA 83 nitrocellulose (Schleicher & Schuell, Inc., Keene, N.H.) with a Bio-Dot apparatus (Bio-Rad Laboratories, Richmond, Calif.). A single-stranded probe was made by Klenow extension, in the presence of [α-^{32}P]dATP and deoxynucleoside triphosphates, of universal primer bound to single-stranded pLA6, which contains the 0.25-kb fragment downstream of the *Eco*RI site within the 20-kDa protein gene (see Fig. 1). After cleavage with *Eco*RI to release the 0.25-kb fragment, the radiolabeled strand was purified by denaturation in 60 mM NaOH, electrophoresis through a 2% agarose gel, and electroelution onto NA 45 paper (Schleicher & Schuell). The probe was stripped from the paper with 300 mM NaOH and was neutralized prior to hybridization. Preparation of RNA samples and hybridization conditions were essentially as described by Maniatis et al. (23). RNA in hybridization assays was quantitated by scanning with a laser densitometer (Ultrascan; LKB Instruments, Inc., Rockville, Md.).

Alternatively, both 20-kDa and 27-kDa protein gene-specific mRNAs were detected by hybridization of complementary [γ-^{32}P]ATP-labeled oligonucleotides (see below) to 50 or 100 μg of total RNA and extension with murine leukemia virus reverse transcriptase (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) according to the instructions of the manufacturer. The reaction mixtures were treated with 0.1 N NaOH (final concentration) to hydrolyze the RNA and were then neutralized and precipitated with ammonium acetate and ethanol. After suspension in a small volume of 70% formamide, the cDNA was heated to 80°C and then loaded onto 1.5-mm-thick polyacrylamide gels containing 8 M urea. This method was also used to estimate the location of the transcriptional start sites of both genes.

Oligonucleotide synthesis and oligonucleotide-directed mutagenesis. Oligonucleotides were synthesized on a multiple-column DNA synthesizer (model 8600; BioSearch, Inc., San Rafael, Calif.). The 19-mer 5′-CCC TTT CTA ACA AAG TAG G-3′ was complementary to coding nucleotides 64 through 82 of the 20-kDa protein mRNA. The 16-mer 5′-GAG GGG TTT TCC ATG G-3′ was complementary to coding nucleotides 46 through 61 of the 27-kDa crystal protein mRNA (nucleotides 554 through 569 in Fig. 2 of reference 40). The mutagenic 34-mer 5′-CTG TTG TGA ATA TTT **ACT** AAA ACA CTC CAT TTT C-3′ was complementary to coding nucleotides 7 through 40 of the 20-kDa protein gene, except that AC replaced TA where indicated (boldface). This replacement resulted in the insertion of GT instead of TA at nucleotides 24 and 25, changing tyrosine 8 and lysine 9 to consecutive amber and ochre nonsense codons. The 34-mer was hybridized to single-stranded DNA from uracil-containing pLA5 (purified from *E. coli* RZ1032 cultures superinfected with M13 KO7), and the complementary strand was synthesized with T4 DNA polymerase in the presence of deoxynucleoside triphosphates and ATP (17). Mutations were confirmed by dideoxy sequencing with Sequenase as described above.

Isolation of 20-kDa protein and antibody production. The 20-kDa protein was isolated from *E. coli* JM83 harboring pLA5. In this construction, the 20-kDa protein gene is transcribed from the *lacZ* promoter of pUC118 and the protein is overproduced and accumulates as inclusions visible by phase microscopy. Substantial purification of inclusions was achieved by centrifuging (25,000 × *g*) sonicates of overnight cultures grown in L broth and washing once with 50 mM EDTA in 50 mM Tris (pH 7.5). To prepare antibodies, the harvested inclusions were electrophoresed through a 12% polyacrylamide–sodium dodecyl sulfate gel (18), and rabbits were immunized with the 20-kDa protein band cut from the gel (24).

Immunoblotting. *B. thuringiensis* subsp. *israelensis* crystals were purified by repeated washing and centrifugation through a Renografin gradient as described by McLean and Whiteley (24). Whole *B. thuringiensis* subsp. *israelensis* cell extracts (shown in Fig. 8) were prepared by treating cells with lysozyme (5 mg ml^{-1}), sonicating several times, and boiling in sample buffer containing 1.5% sodium dodecyl sulfate and 2.5 M urea. Protein samples from *B. thuringiensis* subsp. *israelensis* and *E. coli* were electrophoresed on 10% polyacrylamide–sodium dodecyl sulfate gels and were transferred to BA 83 nitrocellulose filters by the method of Towbin et al. (34). The filters were treated as described by Henning et al. (13) and were reacted with antibody to 20- or 27-kDa protein. Goat anti-rabbit alkaline phosphatase conjugate (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) was used to detect the antigen-antibody complexes.

***lacZ* fusions.** The 3.0-kb *Bam*HI-*Bam*HI fragment from the *lacZ*-coding sequence of pMC1871 (30) was inserted into subclones of pKM1 cleaved at the *Bam*HI site within the 27-kDa protein gene, yielding in-frame fusion plasmids pKM2-*lac* and pKM24-*lac*. These plasmids were maintained in *E. coli* MC1000 (with the *lac* operon deleted). Assays for β-galactosidase activity in extracts of these strains were performed as described by Miller (25). A fusion of the carboxy-terminal portion of the 27-kDa protein gene to the *lacZ* promoter and amino-terminal coding region, pKM25, was constructed by inserting the *Hae*III-*Eco*RI fragment of pKM1 into pUC119 and deleting the region between the *Bam*HI site in the vector and the *Bam*HI site within the 27-kDa protein-coding sequence. The *Hae*III site is 130 nucleotides upstream of the *Bam*HI site shown in Fig. 1; the *Hae*III-*Eco*RI fragment contains the complete 27-kDa protein-coding region but none of the known promoters.

Reagents and enzymes. Restriction endonucleases were from Boehringer Mannheim, New England BioLabs, Bethesda Research Laboratories, or American Allied Biochemical (Denver, Colo.). Exonuclease III and nuclease S1 were

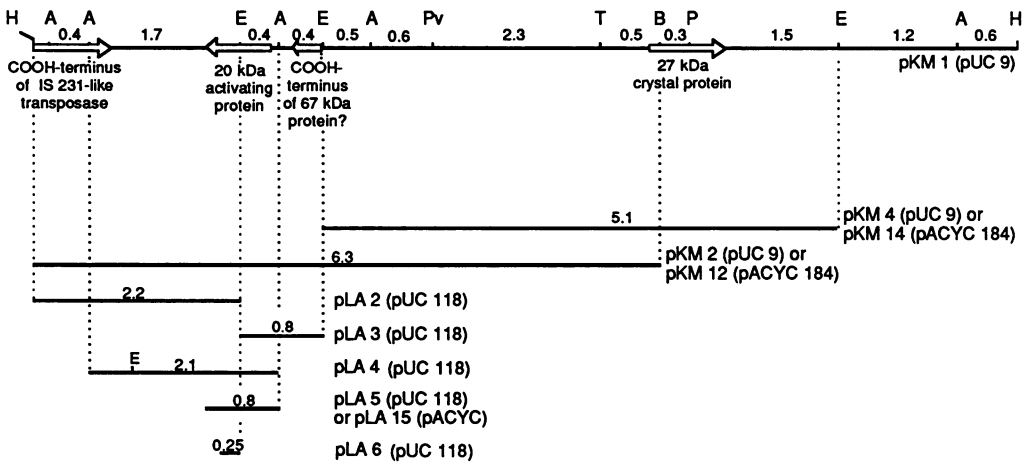

FIG. 1. Subclones of the 10-kb *Hin*dIII fragment (pKM1) used in this study. Lengths (in kilobases) are shown above each fragment; vectors are indicated in parentheses. Restriction sites: H, *Hin*dIII; A, *Acc*I; E, *Eco*RI; Pv, *Pvu*II; T, *Taq*I; B, *Bam*HI; P, *Pst*I. For clarity, not all *Acc*I and *Taq*I sites are shown.

from Bethesda Research Laboratories. T4 polynucleotide kinase, ligase, and DNA polymerase were from Boehringer Mannheim. [α-^{32}P]dATP (3,000 Ci mmol^{-1}; 10 μCi μl^{-1}), [γ-^{32}P]ATP (3,000 Ci mmol^{-1}; 10 μCi μl^{-1}), and [α-^{35}S]dATP (1,000 Ci mmol^{-1}; 17 μCi μl^{-1}) were obtained from New England Nuclear Corp., Boston, Mass.

RESULTS

Localization of the gene encoding the 20-kDa protein by DNA sequencing; analysis of flanking ORFs. The sequence was determined for the nearly 3,000 base pairs of DNA in pLA2 and pLA3, extending from the left *Hin*dIII site to the second *Eco*RI site from the left of the diagram shown in Fig. 1. This region contains one complete ORF flanked by two truncated ORFs (Fig. 2 and 3). The complete ORF (Fig. 2, boxed) spans the first *Eco*RI site from the left end and encodes a protein whose predicted molecular mass is 19,584 Da. This protein has no significant similarity to others in the National Biomedical Research Foundation library (data bank 16). A ribosome-binding site (RBS, Fig. 2; 31) precedes the presumed methionine start codon by seven nucleotides. At the 3′ end of the gene, a G-rich sequence (underlined) and potential stem and loop (double underlined) are located downstream from the TAA (ochre) stop codon.

A truncated, 78-codon ORF, oriented in the same direction as the 20-kDa protein gene, starts within the sequenced fragment at nucleotide 2 (Fig. 2) and terminates 13 nucleotides before a stem-and-loop structure (nucleotides 252 to 283, Fig. 2) located 235 nucleotides upstream of the 20-kDa protein translational start. The position of the 78-codon ORF suggests that it may be the carboxyl terminus of the 67-kDa protein observed in maxicells made from strains containing pKM1 (24). The orientations of both the 20-kDa protein gene and the upstream ORF are opposite that of the 27-kDa crystal protein gene (Fig. 1).

A second partial ORF extends rightward from the *Hin*dIII site at the left of Fig. 1 (nucleotide 2984) and contains 323 amino acids (Fig. 3). (Note that the nucleotide sequence shown in Fig. 3 is inverted relative to the sequence presented in Fig. 2.) Comparison of the deduced amino acid sequence with others in the National Biomedical Research Foundation library disclosed 20.9% identity in an overlap of 129 amino acids with the transposase of IS*4* and 22.3% identity in an overlap of 139 amino acids with the transposase of IS*186*. Greater similarity was found to the amino acid sequence of the putative transposase of IS*231* from *B. thuringiensis* subsp. *thuringiensis* berliner 1715 (22). Comparison of the ORF shown in Fig. 3 with the sequence from strain berliner 1715 showed 55.2% identity in an overlap of 319 amino acids. The initial and optimized scores for the comparison obtained by the FastP computer program (21) were 548 and 1,041, respectively, indicating a significant degree of homology. The alignment of these two sequences suggests that an additional 150 amino-terminal codons of the *B. thuringiensis* subsp. *israelensis* transposase should lie directly upstream of the *Hin*dIII site at nucleotide 2984 of Fig. 3.

The 20-kDa protein gene is required for efficient production of 27-kDa crystal protein in *E. coli*. To confirm that the 20-kDa peptide was involved in the production of the 27-kDa crystal peptide, complementation experiments were performed with plasmids of high (pUC118) and low (pACYC) copy number containing the gene coding for the 20-kDa peptide. The DNA inserted into these plasmids (pLA5 and pLA15, respectively) contained the 120 nucleotides between the *Acc*I site and the translational start of the gene encoding the 20-kDa peptide, the 20-kDa protein gene ORF, and 26 nucleotides beyond the potential stem-and-loop structure (see Materials and Methods).

Complementation of the recombinant strain KM14 (the 27-kDa crystal protein gene on pACYC; Fig. 1) with pLA5 (the 20-kDa protein gene on pUC118) demonstrated that the DNA fragment provided by the latter plasmid was necessary and sufficient to enhance production of the 27-kDa protein, as determined by immunoblotting (Fig. 4A). Virtually no 27-kDa peptide could be detected in *E. coli* harboring pKM14 alone or pKM14 and pUC118 (Fig. 4A). However, the amount of protein detected in extracts of the complemented strain KM14(pLA5) was less than that in *E. coli* harboring pKM1 (Fig. 4A). Phase microscopy demonstrated that the cells of the complemented strain contained inclusions, and immunoblotting showed that the 20-kDa peptide was produced in large amounts (Fig. 4C). Therefore, a lack of the 20-kDa peptide would not appear to be the cause of the relatively low production of 27-kDa peptide unless insolubility of the inclusions somehow inhibited the effect of 20-kDa

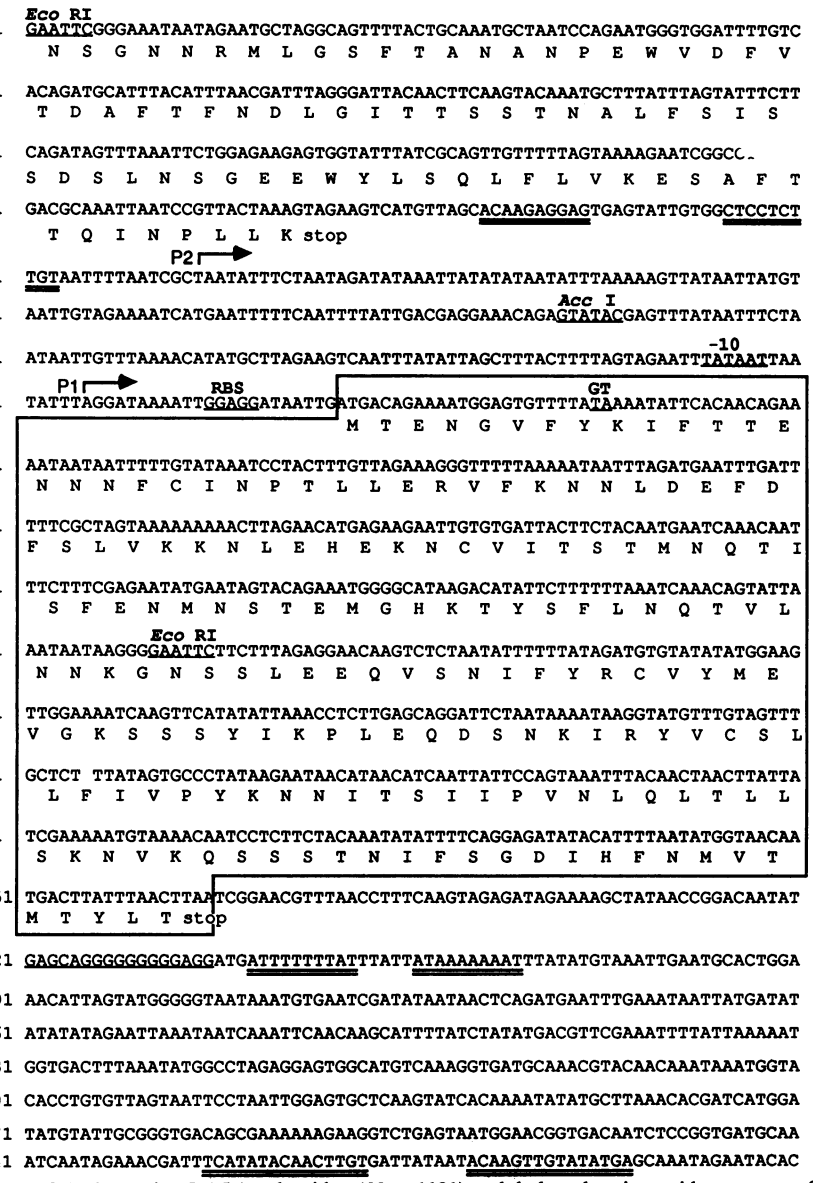

FIG. 2. Nucleotide sequence of the insert in pLA5 (nucleotides 403 to 1191) and deduced amino acid sequence of the 20-kDa protein gene (boxed). The nucleotide and deduced amino acid sequences of the carboxy-terminal region of the presumptive 67-kDa protein gene, from the *Eco*RI site at nucleotide 2, are also shown. A G-rich region following the 20-kDa protein gene is underlined, and potential stem-and-loop structures are double underlined. Start sites for transcription of the 20-kDa protein gene (P1 and P2) are indicated by arrows overhead; a potential ribosome-binding site (RBS) is underlined. The nucleotide changes in tyrosine 8 and lysine 9 to consecutive stop codons, yielding pLA5-89 (pLA15-89 in pACYC), are noted above the underlined region.

peptide. Instead, the lower copy number of the 27-kDa protein gene on pKM14 (pACYC vector) relative to that on pKM1 (pUC9 vector) seemed a more likely explanation.

To examine this possibility, we cloned the gene coding for the 20-kDa peptide into pACYC to yield pLA15 (see Materials and Methods). This plasmid was then transformed into *E. coli* KM4, which harbors the 27-kDa protein gene cloned on pUC9 (Fig. 1). Immunoblotting showed a significant increase in the amount of 27-kDa peptide, exceeding the amount detected in strain KM1 (Fig. 4A and B), whereas cells containing pKM4 alone or pKM4 and pACYC showed no detectable production of the 27-kDa peptide antigen (Fig. 4B). No 20-kDa peptide could be detected in *E. coli* strains containing pLA15 (Fig. 4C), probably because the plasmid does not contain a suitable promoter region for the 20-kDa protein gene. The enhancement observed in the complemented strain KM4(pLA15) suggests that a very small amount of the 20-kDa protein is sufficient for the effect and also that pLA15 must contain a fortuitous, albeit weak, *E. coli* promoter (see below).

Site-directed mutagenesis was used to determine whether the increase in the amount of 27-kDa peptide required the 20-kDa gene product rather than the DNA sequence or the mRNA. Tyrosine 8 and lysine 9 of pLA5 were changed to consecutive stop codons, yielding pLA5-89 (the mutant gene on pUC118); pLA15-89 (the mutant gene on pACYC) was constructed as described for pLA15 (see Materials and Methods). No detectable amount of the 27-kDa peptide was found in extracts of pKM14 complemented with pLA5-89 or of pKM4 complemented with pLA15-89 (Fig. 4A and B),

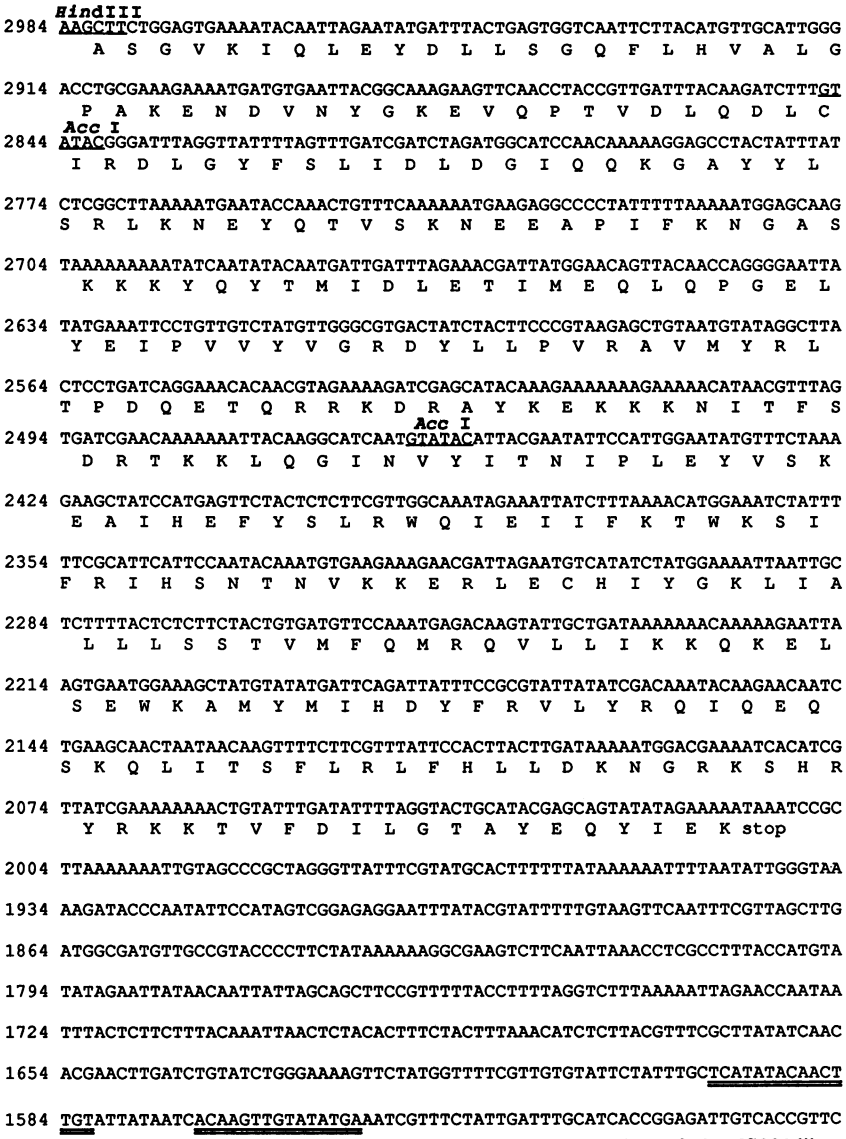

FIG. 3. Nucleotide and deduced amino acid sequences of the carboxy-terminal region of the IS*231*-like transposase. Nucleotides downstream of the putative transposase, including the stem and loop between nucleotides 1557 and 1596, are also shown in Fig. 2. (To display the sense strand of the putative transposase, the nucleotide sequence in this figure is inverted relative to that shown in Fig. 2; however, numbering of the nucleotides is the same in both figures.)

indicating that the 20-kDa gene product is required for expression.

To eliminate the possibility that the mutated 20-kDa protein genes in pLA5-89 and in pLA15-89 were blocked in transcription as well as in translation, a comparison was made of the amounts of 20-kDa protein-specific mRNA produced by the following strains: KM14(pLA5), KM14 (pLA5-89), KM4(pLA15), and KM4(pLA15-89). Hybridization with a single-stranded probe complementary to a downstream region (designed to detect full-length message) showed that measurable amounts of 20-kDa protein-specific mRNA were produced by all four strains (Fig. 5), although synthesis of this mRNA was reduced approximately fivefold in the strains containing the mutant genes. Similar reductions in the transcription of other genes have been ascribed to polar effects on transcription resulting from premature translation termination (1). As expected, large amounts of mRNA were synthesized from the clones in the pUC118 vector (pLA5 and pLA5-89), but synthesis of 20-kDa protein-specific mRNA from both of the pACYC-based plasmids (pLA15 and pLA15-89) was markedly less, perhaps reflecting the lack of a suitable *E. coli* promoter as well as the lower copy number of the parent plasmid. In spite of this low level of 20-kDa protein gene expression, pLA15 greatly enhanced production of the 27-kDa protein. If 20-kDa protein gene-specific mRNA were responsible for the enhancement, then pLA5-89, which produces about fivefold more mRNA than does pLA15, should also enhance 27-kDa protein synthesis. We observed, however, that neither of the strains carrying the mutant 20-kDa protein gene produced detectable amounts of 27-kDa crystal protein, suggesting that the enhancing effect is due to the peptide and not to the mRNA.

Timing of 20-kDa protein gene-specific RNA synthesis in *B. thuringiensis* subsp. *israelensis*. The timing of transcription of the 20- and 27-kDa protein genes and identification of the 5′

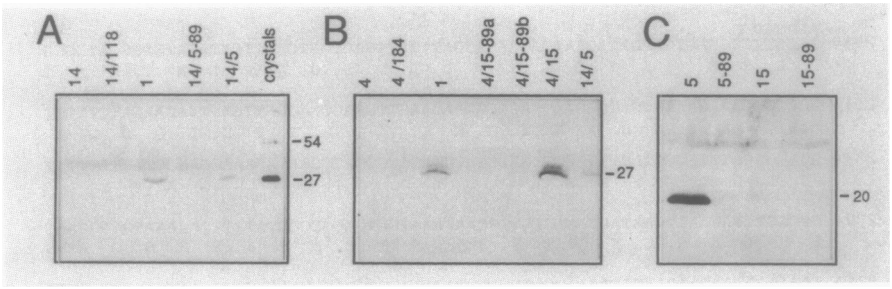

FIG. 4. Immunoblots showing expression of the 27-kDa (A and B) and 20-kDa (C) peptides in extracts from recombinant strains of *E. coli* JM83. The number above each lane refers to the plasmid content of the strain used in the assay. (A) Lane 14, pKM14 (27-kDa protein gene on pACYC184); lane 14/118, pKM14 complemented with pUC118; lane 1, pKM1 (27- and 20-kDa protein genes on pUC9); lane 14/5-89, pKM14 plus pLA5-89 (mutant 20-kDa protein gene on pUC118); lane 14/5, pKM14 plus pLA5 (20-kDa protein gene on pUC118); crystals, purified crystals from *B. thuringiensis* subsp. *israelensis*. (B) Lane 4, pKM4 (27-kDa protein gene on pUC9); lane 4/184, pKM4 plus pACYC; lanes 4/15-89a and 4/15-89b, pKM4 plus the mutant 20-kDa protein gene on pACYC (two isolates); lane 4/15, pKM4 plus pLA15 (20-kDa protein gene on pACYC); lane 14/5, pKM14 plus pLA5. (C) Lane 5, pLA5; lane 5-89, pLA5-89; lane 15, pLA15, lane 15-89, pLA15-89.

ends of the transcripts were determined by reverse transcriptase extension of oligonucleotides hybridized to the 5′ coding regions of the mRNA. No 20-kDa protein mRNA was detected before T_0, the onset of stationary phase (data not shown). A transcript appearing at T_0 (Fig. 6A) was mapped to a start site (P1 in Fig. 2) located approximately 23 ± 5 base pairs upstream from the translational start site. At T_2 (2 h post-stationary phase) (Fig. 6A), longer transcripts were observed which mapped to a position approximately 225 ± 15 base pairs upstream from the translational start (P2 in Fig. 2), near the upstream stem and loop described earlier. Surprisingly, 20-kDa protein mRNA synthesized in *E. coli* KM1 mapped uniquely to the position within the upstream stem and loop, and the bands were extremely faint (data not shown). Transcription of the 27-kDa protein gene in *B. thuringiensis* subsp. *israelensis* was first detected as a faint band at T_2 and more strongly at $T_{5.5}$ (Fig. 6B). Interestingly, these transcripts mapped to two sites: PB1, a *B. thuringiensis* subsp. *israelensis* start site identified earlier by Ward and Ellar (40), and PBS1, a site previously reported only for expression of the 27-kDa protein gene in *Bacillus subtilis* (42).

Detection of 20-kDa protein in *B. thuringiensis* subsp. *israelensis* crystals and cell extracts. Electrophoresis showed that solubilized preparations of purified crystals contained the 135-, 128-, 67-, and 27-kDa major crystal peptides, as well as the 38-kDa peptide and the 54-kDa dimer, but the 20-kDa peptide could not be detected in gels stained with Coomassie blue (Fig. 7A). Immunoblotting revealed the presence of the 20-kDa peptide in crystal preparations (Fig. 7B, lanes 4 to 6), but detection of this protein required electrophoresis of large amounts of solubilized crystals. In contrast, purified crystal preparations contained easily detectable amounts of 27-kDa peptide (Fig. 7B, lanes 1 to 3). Analysis of cell extracts at various time points during growth indicated that production of the 20-kDa antigen, beginning at ca. T_2 and continuing through $T_{8.5}$, was concurrent with 27-kDa crystal protein production (Fig. 8). However, the

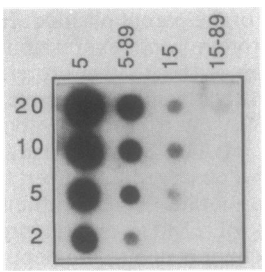

FIG. 5. Hybridization of single-stranded 20-kDa protein gene-specific probe to RNA samples from strains described in the legend to Fig. 4C. Numbers at left indicate the total RNA (in micrograms) applied to nitrocellulose.

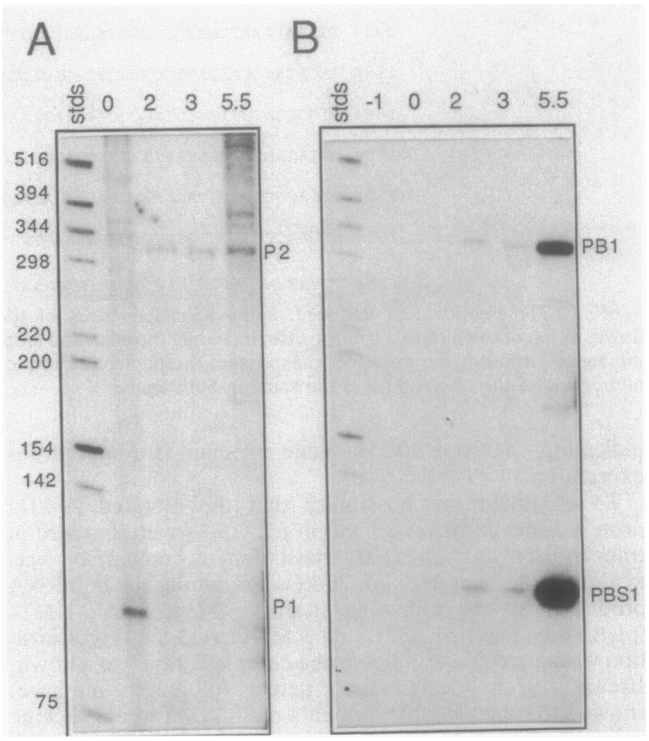

FIG. 6. Time of appearance of mRNA transcripts from (A) 20-kDa protein gene and (B) 27-kDa protein gene in *B. thuringiensis* subsp. *israelensis*. Numbers above each lane indicate the time post-stationary phase (in hours); numbers at left indicate the size of the standards (stds; in nucleotides). The transcripts made from the 20-kDa protein gene (P1 and P2) and 27-kDa protein gene (PBS1 and PB1) (42) are indicated.

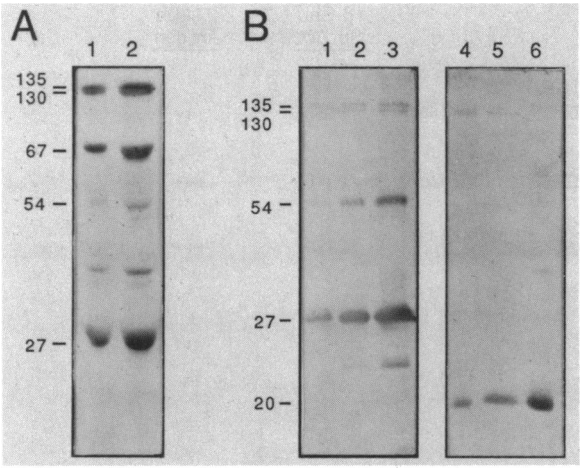

FIG. 7. Detection of the 27- and 20-kDa peptides in solubilized preparations of purified *B. thuringiensis* subsp. *israelensis* crystals. (A) *B. thuringiensis* subsp. *israelensis* crystal peptides stained with Coomassie blue. Lane 1, 50 μg of total protein; lane 2, 100 μg. (B) Immunodetection of 27-kDa (lanes 1 to 3) and 20-kDa (lanes 4 to 6) proteins in *B. thuringiensis* subsp. *israelensis* crystals. Lane 1, 2 μg of total protein; lane 2, 5 μg; lane 3, 10 μg; lane 4, 100 μg; lane 5, 200 μg. Lane 6 contains approx. 1 μg of crude extract from pLA5, which overproduces the 20-kDa protein. Molecular masses (in kilodaltons) of major crystal polypeptides are indicated at left of each panel.

amount of 20-kDa protein produced was substantially less than that of 27-kDa protein. A small amount of 20-kDa protein was also produced during late exponential phase (Fig. 8B), before the appearance of any 27-kDa protein.

The 20-kDa protein acts after the initiation of translation. McLean and Whiteley (24) showed that an increase in the steady-state level of mRNA produced from the 27-kDa crystal protein gene cannot account for the effect of the DNA fragment encoding the 20-kDa protein. The mechanism of action of the 20-kDa protein was further investigated by observing its effect on fusions of the 27-kDa protein to *lacZ*. First, a derivative of the *lacZ* gene lacking promoters and translational initiation signals was fused to the promoter region, the 5′-untranslated sequence, the ribosome-binding site, and the first 30 codons of the 27-kDa protein gene (pKM24-*lac*, Fig. 9). This translational fusion was expressed in *E. coli* at a level sufficient for immunodetection with antibody against the 27-kDa protein, which apparently reacts efficiently with the first 30 amino acids of the peptide. The substantial level of β-galactosidase activity in extracts of these cells (Fig. 9) further demonstrated that a considerable amount of the fusion peptide was being synthesized. When the 20-kDa protein was also present, either in *cis* (pKM2-*lac*) or in *trans* (pKM24-*lac*+pLA15), both immunodetection and measurement of β-galactosidase activity indicated no significant increase in the level of production of the fusion protein. These results confirm that the 20-kDa protein acts posttranscriptionally and also imply that it is not involved in the initiation of translation.

A second type of fusion was used to further examine these findings. The 5′-untranslated sequences and the first 30 codons of the 27-kDa protein gene were deleted, and the carboxy-terminal 217 codons which remained were fused to the promoter, the ribosome-binding site, and the first 16 codons of *lacZ* (pKM25, Fig. 9). Although this fusion did not encode active β-galactosidase, its product could be detected with 27-kDa protein-specific antibody. In the absence of the 20-kDa protein, the fusion protein was produced at a very low level, but the addition of the 20-kDa protein in *trans* (pKM25+pLA15) resulted in a substantial increase in the amount of protein (Fig. 9). Enhancement of 27-kDa protein production in the absence of any wild-type transcriptional or translational initiation signals clearly shows that the 5′ region of the 27-kDa protein gene is not required for the function of the 20-kDa protein. It seems likely, therefore, that the 20-kDa protein acts either during the synthesis of the 27-kDa protein or posttranslationally. In addition, the 20-kDa protein had no effect on synthesis of the 27-kDa crystal protein in an in vitro coupled transcription-translation system (data not shown), which also implies a posttranslational mode of action.

Pfannenstiel et al. (27) have reported that *B. thuringiensis* subsp. *israelensis* crystal proteins contain amino sugars. It seemed plausible, then, that the 20-kDa protein might be involved in the glycosylation of the 27-kDa protein either during or after translation, thereby increasing the net synthesis or stability of the 27-kDa protein. This idea was investigated by reacting wheat germ agglutinin conjugated to horseradish peroxidase with proteins transferred electrophoretically to nitrocellulose filters. All four major polypeptides from *B. thuringiensis* subsp. *israelensis* crystals were indeed found to be glycosylated, but there was no evidence that the 27-kDa crystal protein produced by recombinant *E. coli* strains contained amino sugars either in the presence of the 20-kDa peptide or in its absence (data not shown).

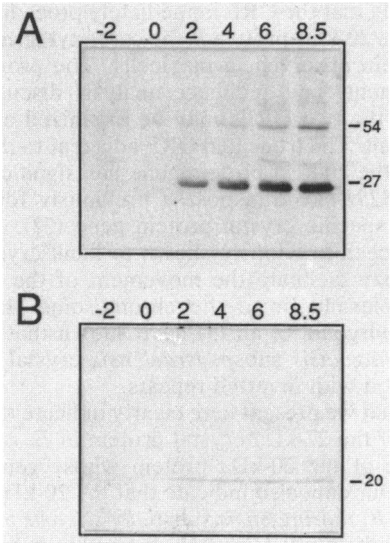

FIG. 8. Times of appearance of 27- and 20-kDa peptides in *B. thuringiensis* subsp. *israelensis*. Immunoblots to detect (A) 27-kDa and (B) 20-kDa peptides in whole-cell extracts from cells harvested at indicated times (in hours) post-stationary phase are shown. Lanes in panel B were loaded with seven times the amount of total protein in the corresponding lanes in panel A. Molecular masses (in kilodaltons) of immunostained proteins are indicated at right of each panel.

DISCUSSION

McLean and Whiteley (24) demonstrated that a region of DNA located 4 kb upstream of the 27-kDa crystal protein gene is necessary for efficient production of the latter in *E. coli*. Tn*5* insertions delimited a 0.8-kb required region; subcloning and maxicell analysis suggested that this region

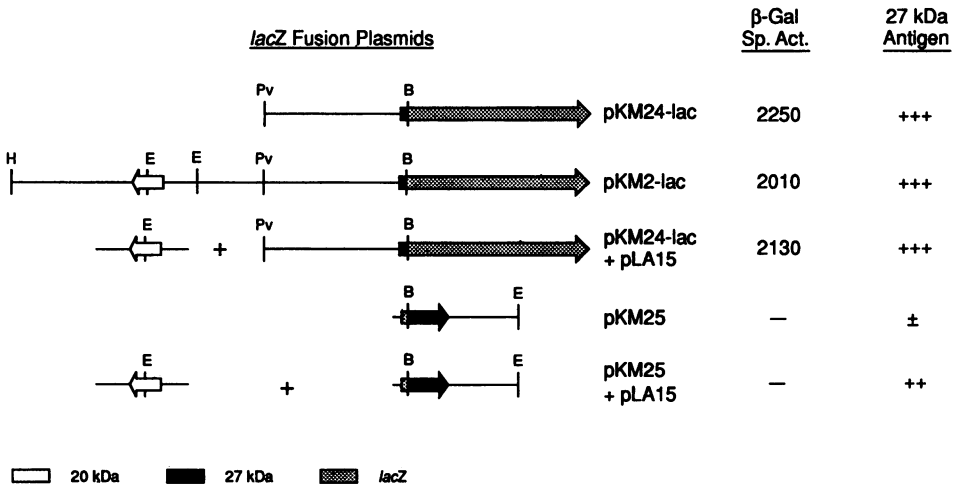

FIG. 9. Effect of the 20-kDa protein on fusions of the 27-kDa crystal protein gene to *lacZ*. Synthesis of the fusion products was measured by immunodetection with antibody specific to the 27-kDa protein; the products of the translational fusions pKM24-*lac* and pKM2-*lac* were also measured by β-galactosidase activity in cell extracts and are reported in nanomoles of *o*-nitrophenol formed per minute per milligram of protein. Three or more trials have been averaged in each case. Standard deviations are as follows: pKM24-*lac*, 190 U; pKM2-*lac*, 290 U; pKM24+pLA15, 95 U. Arrows representing the genes are drawn to scale. Only the restriction sites corresponding to sites in pKM1 are indicated; abbreviations are as defined in the legend to Fig. 1.

encoded a 20-kDa protein. The presence of this region either in *cis* or in *trans* greatly increased the amount of 27-kDa protein which could be detected in *E. coli* extracts, but the amount of 27-kDa protein mRNA increased only slightly, indicating that the effect of the 20-kDa protein is posttranscriptional. In this study, we have sequenced the upstream region and have identified an ORF that encodes a protein of 20 kDa. Analysis of strains carrying this ORF and the 27-kDa crystal protein gene showed a significant increase in the amount of the 27-kDa crystal peptide detected by immunoblot analysis. We have further shown, using site-directed mutagenesis, that it is the protein product of the ORF, not the mRNA or the DNA sequence, which is responsible for the effect. We have also observed that the 20-kDa protein mRNA appears 2 h earlier (T_0) than the 27-kDa protein mRNA (T_2) and that the two proteins are synthesized concurrently beginning at about T_2. A small amount of the 20-kDa protein is also detectable during exponential growth, before any 27-kDa protein has been synthesized. The timing of transcription and translation of the two genes is consistent with the idea that the 20-kDa protein enhances 27-kDa protein production.

Reverse transcriptase mapping of the 20-kDa protein gene promoter indicated that two major transcripts are produced in *B. thuringiensis* subsp. *israelensis*. The first transcript maps to a start site 23 nucleotides upstream of the translational start, is preceded by a consensus *E. coli* −10 sequence, and appears only at T_0. The second, less-abundant transcript appears at T_2 and persists through $T_{5.5}$. This transcript maps near or within the stem-and-loop structure lying 235 nucleotides upstream of the translational start of the 20-kDa protein gene. Curiously, *E. coli* KM1 uses this same start site exclusively, although the region lacks apparent consensus sequences for the −10 and −35 promoter regions recognized by the predominant form of *E. coli* RNA polymerase. Because the apparent start of the less-abundant transcript lies near the upstream stem and loop, two explanations are possible. (i) The actual sporulation transcriptional start for the 20-kDa protein gene in *B. thuringiensis* subsp. *israelensis* is located within or near the potential stem-and-loop structure. (ii) The reverse transcriptase used in the analysis to produce a cDNA copy of the 20-kDa protein stalled prematurely in the stem and loop, and the actual start site for the second transcript in *B. thuringiensis* subsp. *israelensis*, and the sole transcript in *E. coli*, lies farther upstream. Additional experiments will be needed to distinguish between these possibilities.

Earlier analyses (24) using maxicells demonstrated that pKM1 encodes polypeptides of 67 and 16 kDa in addition to the 20- and 27-kDa proteins. Experiments with subclones suggested that the 67-kDa protein, which may be the same as the mosquitocidal 67-kDa protein found in *B. thuringiensis* subsp. *israelensis* crystals, was encoded within a 2.5-kb segment immediately upstream from the 27-kDa protein (24). It seems likely that the ORF immediately preceding the gene coding for the 20-kDa protein is the carboxyl terminus of the 67-kDa protein observed in maxicells. The promoter-mapping experiments and sequence analysis discussed above suggest that the two ORFs may be organized as one transcriptional unit. The truncated ORF adjacent to the carboxyl terminus of the 20-kDa protein gene has significant homology to an IS*231*-like transposase previously found near a Lepidoptera-specific crystal protein gene (22). IS-like elements have been previously shown to flank crystal protein genes and may facilitate the movement of the latter from plasmid to plasmid or to the chromosome (15, 16, 20). Recently, Bourgouin et al. (3) have shown that one of the large *B. thuringiensis* subsp. *israelensis* crystal proteins is also associated with inverted repeats.

The data that we present here clearly indicate that efficient production of the 27-kDa crystal protein in *E. coli* requires the presence of the 20-kDa protein whose gene we have sequenced. Our data also indicate that the 20-kDa protein is expressed in *B. thuringiensis* subsp. *israelensis* and that it is associated with purified crystals from this subspecies. At present, however, we have no information regarding the physiological function of this protein in *B. thuringiensis* subsp. *israelensis*. The presence of the 20-kDa protein in crystals suggests a possible role for the protein in crystal synthesis (as a scaffolding protein, for example) or in the protection of one or more of the crystal proteins from proteolysis. In *B. subtilis*, at least, the 20-kDa protein may

not be required for high-level synthesis of the 27-kDa protein: Ward et al. (42) observed the production of 27-kDa protein inclusions in *B. subtilis* cells carrying the 27-kDa protein gene but lacking the 20-kDa protein-coding region. Furthermore, we were unable to detect any sequences homologous to the 20-kDa protein gene by hybridization of 20-kDa protein gene-specific probes to whole-cell DNA from *B. subtilis* 168S or *B. subtilis* BD170, although DNA from *B. thuringiensis* subsp. *israelensis* or *E. coli*(pLA5) showed very strong hybridization to the same probes (unpublished observations). While these data suggest that no homolog of the 20-kDa protein gene resides in *B. subtilis*, they do not preclude the presence of a functionally analogous protein lacking sufficient sequence homology for detection by this method. It is also possible that the 20-kDa protein could boost net synthesis of the 27-kDa protein in *B. subtilis* to still higher levels.

Although the effect of the 20-kDa protein on 27-kDa crystal protein expression in *E. coli* has been clearly demonstrated, the mechanism by which this occurs remains unclear. Most procaryotic systems appear to be regulated at the transcriptional level, but there are also several examples of posttranscriptional regulation in which a protein increases mRNA stability (11) or in which either a protein (9) or, in some cases, an RNA molecule (26) regulates the initiation of translation. However, our demonstration that it is the protein encoded by the 20-kDa protein gene which enhances production of the 27-kDa protein, the observation that this effect cannot be explained by changes in mRNA synthesis or accumulation (24), and our studies with fusions of the 27-kDa protein to *lacZ* seem to rule out these mechanisms. Instead, the 20-kDa protein appears to exert its effect after the initiation of translation, either during synthesis of the 27-kDa protein or after translation has been completed. Preliminary experiments suggest that the 20-kDa protein enhances the stability of the 27-kDa protein; we are presently examining possible mechanisms which might account for such stabilization.

ACKNOWLEDGMENTS

This research was supported by Public Health Service grant GM-20784 from the National Institute of General Medical Sciences. L. F. Adams is the recipient of Postdoctoral Fellowship GM 11715-02 from the National Institutes of Health. H. R. Whiteley is the recipient of Research Career Award K6-GM-442 from the National Institute of General Medical Sciences. This material is based in part on work supported by a National Science Foundation Graduate Fellowship (J. Visick).

We acknowledge the preliminary sequence data obtained by Kathy Tomczak and the high-resolution restriction map of the 10-kb fragment generated by Elizabeth Baker. We thank H. Ernest Schnepf for his advice and many helpful discussions.

LITERATURE CITED

1. **Adhya, S., and M. Gottesman.** 1978. Control of transcription termination. Annu. Rev. Biochem. **47**:967–996.
2. **Armstrong, J., G. Rohrmann, and G. Beaudreau.** 1985. Delta endotoxin of *Bacillus thuringiensis* subsp. *israelensis*. J. Bacteriol. **161**:39–46.
3. **Bourgouin, C., A. Delécluse, J. Ribier, A. Klier, and G. Rapoport.** 1988. A *Bacillus thuringiensis* subsp. *israelensis* gene encoding a 125-kilodalton larvicidal polypeptide is associated with inverted repeat sequences. J. Bacteriol. **170**:3575–3583.
4. **Casadaban, M., and S. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
5. **Chang, A., and S. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**:1141–1156.
6. **Charles, J.-F., and H. de Barjac.** 1982. Sporulation et cristallogenèse de *Bacillus thuringiensis* var. *israelensis* en microscopie électronique. Ann. Microbiol. (Inst. Pasteur) **133A**:425–442.
7. **Couche, G., M. Pfannenstiel, and K. Nickerson.** 1987. Structural disulfide bonds in the *Bacillus thuringiensis* subsp. *israelensis* protein crystal. J. Bacteriol. **169**:3281–3288.
8. **de Barjac, H.** 1978. Une nouvelle variété de *Bacillus thuringiensis* très toxique pour les moustiques: *Bacillus thuringiensis* var. *israelensis* sérotype 14. C.R. Acad. Sci. **286D**: 797–800.
9. **Fallon, A., C. Jinks, G. Strycharz, and M. Nomura.** 1979. Regulation of ribosomal protein synthesis in *Escherichia coli* by selective mRNA inactivation. Proc. Natl. Acad. Sci. USA **76**: 4311–4315.
10. **Goldberg, L., and J. Margalit.** 1977. A bacterial spore demonstrating rapid larvicidal activity against *Anopheles sergentii*, *Uranotaenia unguiculata*, *Culex univitattus*, *Aedes aegypti* and *Culex pipiens*. Mosq. News **37**:355–358.
11. **Gorski, K., J.-M. Roch, P. Prentki, and H. Krisch.** 1985. The stability of bacteriophage T4 gene 32 mRNA: a 5′ leader sequence that can stabilize mRNA transcripts. Cell **43**:461–469.
12. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
13. **Henning, U., H. Schwarz, and R. Chen.** 1979. Radioimmunological screening method for specific membrane proteins. Anal. Biochem. **97**:153–157.
14. **Insell, J., and P. Fitz-James.** 1985. Composition and toxicity of the inclusion of *Bacillus thuringiensis* subsp. *israelensis*. Appl. Environ. Microbiol. **50**:56–62.
15. **Kronstad, J., H. E. Schnepf, and H. R. Whiteley.** 1983. Diversity of locations for *Bacillus thuringiensis* crystal protein genes. J. Bacteriol. **154**:419–428.
16. **Kronstad, J., and H. R. Whiteley.** 1984. Inverted repeat sequences flank a *Bacillus thuringiensis* crystal protein gene. J. Bacteriol. **160**:95–102.
17. **Kunkel, T., J. Roberts, and R. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
18. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
19. **Lee, S., W. Eckblad, and L. 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.
20. **Lereclus, D., J. Ribier, A. Klier, G. Menou, and M. Lecadet.** 1985. A transposon-like structure related to the δ-endotoxin gene of *Bacillus thuringiensis*. EMBO J. **3**:2561–2567.
21. **Lipman, D., and W. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
22. **Mahillon, J., J. Seurinck, L. Van Rompuy, J. Delcour, and M. Zabeau.** 1985. Nucleotide sequence and structural organization of an insertion sequence element (IS*231*) from *Bacillus thuringiensis* strain berliner 1715. EMBO J. **4**:3895–3899.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **McLean, K., 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.
25. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Mizuno, T., M.-Y. Chou, and M. Inouye.** 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proc. Natl. Acad. Sci. USA **81**:1966–1970.
27. **Pfannenstiel, M., G. Muthukumar, G. Couche, and K. Nickerson.** 1987. Amino sugars in the glycoprotein toxin from *Bacillus thuringiensis* subsp. *israelensis*. J. Bacteriol. **169**:796–801.
28. **Sanger, F., S. Nicklen, and A. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA

74:5463–5467.
29. **Sen, K., G. Honda, N. Koyama, M. Nishida, A. Neki, H. Sakai, M. Himeno, and T. Komanao.** 1988. Cloning and nucleotide sequences of the two 130 kDa insecticidal protein genes of a *Bacillus thuringiensis* subsp. *israelensis*. Agric. Biol. Chem. **52:** 873–878.
30. **Shapira, S., J. Chou, F. Richaud, and M. Casadaban.** 1983. New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to *lacZ* gene sequences encoding an enzymatically active carboxy-terminal portion of β-galactosidase. Gene **25:**71–82.
31. **Shine, J., and L. Dalgarno.** 1975. Determination of cistron specificity in bacterial ribosomes. Nature (London) **254:**34–38.
32. **Thomas, W., and D. Ellar.** 1983. *Bacillus thuringiensis* var. *israelensis* crystal δ-endotoxin: effects on insect and mammalian cells in vitro and in vivo. J. Cell Sci. **60:**181–197.
33. **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.
34. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76:**4350–4354.
35. **Tyrell, D., L. Bulla, Jr., R. Andrews, Jr., L. Davidson, and P. Nordin.** 1979. Toxicity of parasporal crystals of *Bacillus thuringiensis* subsp. *israelensis* to mosquitoes. Appl. Environ. Microbiol. **38:**656–658.
36. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
37. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.
38. **Visser, B., M. van Workum, A. Dullemans, and C. Waalwijk.** 1986. The mosquitocidal activity of *Bacillus thuringiensis* var. *israelensis* is associated with M_r 230 000 and 130 000 crystal proteins. FEMS Microbiol. Lett. **30:**211–214.
39. **Waalwijk, C., A. Dullemans, M. van Workum, and B. Visser.** 1985. Molecular cloning and the nucleotide sequence of the M_r 28 000 crystal protein gene of *Bacillus thuringiensis* subsp. *israelensis*. Nucleic Acids Res. **13:**8207–8217.
40. **Ward, E., and D. Ellar.** 1986. *Bacillus thuringiensis* var. *israelensis* δ-endotoxin: nucleotide sequence and characterization of the transcripts in *Bacillus thuringiensis* and *Escherichia coli*. J. Mol. Biol. **191:**1–11.
41. **Ward, E., and D. Ellar.** 1987. Nucleotide sequence of a *Bacillus thuringiensis* var. *israelensis* gene encoding a 130 kDa delta-endotoxin. Nucleic Acids Res. **15:**7195.
42. **Ward, E., A. Ridley, D. Ellar, and J. Todd.** 1986. *Bacillus thuringiensis* var. *israelensis* δ-endotoxin: cloning and expression of the toxin in sporogenic and asporogenic strains of *Bacillus subtilis*. J. Mol. Biol. **191:**13–22.
43. **Williams, S., B. Slatko, L. Moran, and S. DeSimone.** 1986. Sequencing in the fast lane: a rapid protocol for [α-^{35}S]dATP dideoxy DNA sequencing. BioTechniques **4:**138–147.