JOURNAL OF BACTERIOLOGY, Oct. 1984, p. 95–102
0021-9193/84/100095-08$02.00/0

Vol. 160, No. 1

Inverted Repeat Sequences Flank a *Bacillus thuringiensis* Crystal Protein Gene

JAMES W. KRONSTAD AND H. R. WHITELEY*

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

Received 10 May 1984/Accepted 25 July 1984

Two sets of inverted repeat DNA sequences, IR2150 and IR1750, were discovered flanking the crystal protein gene on the 75-kilobase plasmid of *Bacillus thuringiensis* subsp. *kurstaki* HD73. A restriction map of ca. 40 kilobases around the crystal protein gene was constructed, and the positions of the copies of IR2150 and IR1750 were determined. Three copies of IR2150 were found flanking the crystal protein gene in an inverted orientation, and one partial and three intact copies of IR1750 were found in both inverted and direct orientations around the gene. Hybridization experiments with fragments from within IR2150 and IR1750 demonstrated the presence of multiple copies of these sequences on the chromosome of *B. thuringiensis* subsp. *kurstaki* HD73 and also revealed a strong correlation between the presence of these sequences and the presence of the crystal protein gene on plasmids from 14 strains of *B. thuringiensis*.

During sporulation, *Bacillus thuringiensis* produces a crystalline inclusion composed of a single polypeptide, the crystal protein or delta endotoxin. Depending on the subspecies, the crystal protein is toxic to the larvae of lepidopteran or dipteran insects, and preparations of spores and crystals are used commercially as insecticides.

Some subspecies of *B. thuringiensis* contain more than one homologous crystal protein gene (9, 10; H. R. Whiteley, H. E. Schnepf, J. W. Kronstad, and H. C. Wong, *in* A. T. Ganesan and J. A. Hoch, ed., *Genetics and Biotechnology of Bacilli*, in press) and most subspecies contain numerous extrachromosomal elements (5, 10, 25). The crystal protein genes are found predominantly on large plasmids ranging in size from about 50 kilobases (kb) to greater than 225 kb (3, 5, 10); in some subspecies, the gene may be on the chromosome (6, 9, 10). The diversity of plasmid sizes may indicate a large amount of plasmid rearrangement, and in fact, curing experiments (3, 5) suggested that considerable rearrangements may occur, including deletions, insertions, and integration of plasmid DNA into the chromosome. Furthermore, hybridization experiments with plasmids from various subspecies revealed that common sequences are present in addition to those encoding the crystal protein and that the plasmids can be divided into classes by size and by homology (11). The presence of insertion sequences on *B. thuringiensis* plasmids could account for some of these observations. Two findings indicate that such sequences may be present: (i) electron microscopy demonstrated the presence of an inverted repeat structure on a 16.4-kb plasmid of *B. thuringiensis* subsp. *galleriae* (4); and (ii) when *Streptococcus faecalis* plasmid pAMβ1 was transferred into *B. thuringiensis* and subsequently reisolated, the plasmid contained a segment of DNA originating from a *B. thuringiensis* plasmid (12). The inserted sequence hybridized to large plasmids from several other subspecies of *B. thuringiensis* (12).

We cloned a crystal protein gene from a large plasmid in *B. thuringiensis* subsp. *kurstaki* (20), surveyed a number of strains for the distribution of this gene (10), and determined the DNA sequence of the promoter and coding region of the gene (26; Whiteley et al., in press). We have also cloned related crystal protein genes from other *B. thuringiensis* subsp. *kurstaki* strains and from other subspecies (25; Whiteley et al., in press). The present communication reports the discovery of two sets of inverted repeat sequences flanking a crystal protein gene on a 75-kb plasmid from *B. thuringiensis* subsp. *kurstaki* HD73. These sequences were found on plasmids carrying the crystal protein gene in 14 other strains (representing seven subspecies) as well as on the chromosome of *B. thuringiensis* subsp. *kurstaki* HD73.

MATERIALS AND METHODS

Bacterial strains and plasmids. The sources of *B. thuringiensis* strains and conditions for their growth were reported previously (10). The following *Escherichia coli* strains and corresponding plasmids were employed: HB101(pBR322), HB101(ColE1::Tn*5*) (obtained from David Garfinkel, Department of Microbiology and Immunology), JM83(pUC18) (obtained from Martin Yanofsky, Department of Microbiology and Immunology), JM83(pUC8) (24), and JM103 (M13mp7, M13mp8, M13mp9) (16). Conditions for maintenance and growth of *E. coli* strains have been previously described (10, 13).

Isolation of inverted repeat DNA. Inverted repeat sequences were detected by the method of Ohtsubo and Ohtsubo (18). In brief, plasmid DNA (1 to 6 μg/ml) was linearized by digestion with a restriction endonuclease (*Sal*I for the 75-kb plasmid of *B. thuringiensis* subsp. *kurstaki* HD73), denatured by boiling, and allowed to renature briefly at 68°C in the presence of 300 mM NaCl. The DNA was then added to S1 nuclease buffer (30 mM sodium acetate [pH 4.6], 4.5 mM $ZnSO_4$, 280 mM NaCl) containing 100 U of S1 nuclease (Bethesda Research Laboratories) per μg of DNA. After incubation for 40 min at 30°C, the S1 digestion was stopped by the addition of 750 mM ammonium acetate and 20 mM Na_2EDTA. The DNA was then precipitated with ethanol and analyzed by electrophoresis.

DNA purification. Plasmid DNA and total cell DNA from *B. thuringiensis* and plasmid DNA from *E. coli* were prepared as previously described (10, 20). In all cases, plasmid DNA was banded in cesium chloride-ethidium bromide gradients. Restriction endonuclease fragments or inverted repeat "snap back" DNAs for cloning, restriction enzyme digestion, or use as hybridization probes were prepared by electroelution from an agarose gel slice and concentrated by

* Corresponding author.

chromatography on a DE52 cellulose column (21). *B. thuringiensis* plasmids were separated by centrifugation through sucrose gradients (5 to 25%) as previously described (20).

Transfer of DNA to nitrocellulose and hybridization. The preparation of nitrocellulose filters and hybridization conditions were as described (10). In all cases, fragments to be used as hybridization probes were cloned into appropriate *E. coli* vectors. Fragments were then separated by electrophoresis and labeled to a specific activity of 1×10^7 to 2×10^7 cpm/μg with ^{32}P by the DNA polymerase catalyzed fill-in or nick translation reactions (14, 15). For the experiments reported here, hybridization conditions were designed to allow a ca. 25% base pair (bp) mismatch based on 37 mol% guanine plus cytosine for *B. thuringiensis* DNA (10, 26). The percent mismatch was calculated by the method of Howley et al. (7).

Restriction mapping and cloning. Digestion with restriction enzymes was performed as recommended by the manufacturers, New England Biolabs and Bethesda Research Laboratories. Plasmid samples and restriction endonuclease fragments were subjected to electrophoresis as previously described (10). Restriction enzyme maps were generated by standard techniques (13). In some instances, restriction endonuclease fragments were purified from agarose gels and digested with additional enzymes. The construction of recombinant plasmids was carried out by standard techniques (13). T4 DNA ligase was purchased from New England Biolabs. Photographic negatives of ethidium bromide-stained agarose gels were scanned with an LKB Bromma 2202 Ultroscan laser densitometer, and the areas of the resulting peaks were determined with an LKB Bromma 2220 recording integrator.

Immunoblotting and insect toxicity analysis. For antigenic analysis, *E. coli* strains were grown in 5 ml of L broth (17), and extracts were prepared by sonication in the presence of urea, sodium dodecyl sulfate, and beta-mercaptoethanol. Crystals from *B. thuringiensis* subsp. *kurstaki* HD73 were prepared as described previously (20). Protein samples were subjected to electrophoresis on 10 to 20% polyacrylamide slab gels containing sodium dodecyl sulfate (22), transferred to nitrocellulose (23), and incubated with antibody (raised against HD1-Dipel crystal protein [20]) and ^{125}I-protein A (New England Nuclear Corp.). Immune complexes on the nitrocellulose were detected by autoradiography with Kodak X-Omat AR film.

For toxicity analysis, neonate larvae of the tobacco hornworm were obtained from J. Truman and L. Riddiford (Dept. of Zoology, University of Washington, Seattle). *E. coli* cultures (in 150 ml of L broth) were grown overnight and centrifuged, the cells were washed in phosphate-buffered saline (2 mM KH_2PO_4, 8 mM Na_2HPO_4, 100 mM NaCl), suspended in 2 ml of the same buffer, and disrupted by sonication. A 0.2-ml sample of each extract (10 to 20 mg of total protein) was then layered on 5 ml of solid feed meal (19) in each of four glass scintillation vials and allowed to dry. One neonate larva was placed in each vial and observed for 10 days.

RESULTS

Cloning fragments of the 75-kb plasmid of *B. thuringiensis* subsp. *kurstaki* HD73. As previously described (25), the crystal protein gene from the 75-kb plasmid of *B. thuringiensis* subsp. *kurstaki* HD73 (henceforth this strain will be designated "HD73") was cloned by insertion of a 13.6-kb *Bgl*II-B fragment into the *Bam*HI site of pBR322 to yield the recombinant plasmid pJWK20. This plasmid was identified by hybridization with intragenic restriction fragments from the cloned crystal protein gene on plasmid pES1 (26) (Fig. 1A). Figure 2 shows the results of an immunoblot experiment which compares the mobilities of the crystal protein from purified HD73 crystals with antigen present in an extract of *E. coli* cells harboring pJWK20. The mobilities were identical. An extract of *E. coli* cells carrying pJWK20 was also applied to feed meal used to raise larvae of the tobacco hornworm, *Manduca sexta*. Larvae that were fed this meal died before the first instar as did larvae that were fed meal containing crystal protein from purified HD73 crystals. Larvae that were fed meal containing an extract of *E. coli* cells carrying pBR322 showed no retardation of growth compared with larvae that were fed untreated meal. These experiments provide evidence that pJWK20 contains a crystal protein gene.

The restriction map shown in Fig. 1B was generated by standard methods, primarily the analysis of double-digestion products, and represents about 40 kb of the 75-kb plasmid of HD73. Digestion of this plasmid with restriction enzyme *Bgl*II yielded six fragments, designated A through F; the largest fragment, *Bgl*II-A, which contains the unmapped portion of the plasmid, is not shown in Fig. 1B. *Bgl*II fragments C, D, E, and F were cloned by insertion into the *Bam*HI site of pBR322 to yield plasmids pJWK25 (*Bgl*II-C), pJWK22 (*Bgl*II-D), pJWK21 (*Bgl*II-E), and pJWK24 (*Bgl*II-F). For restriction mapping, two plasmids were constructed to contain most of the DNA present in the *Bgl*II-B fragment: plasmid pJWK014 was obtained by partial digestion of pJWK20 (*Bgl*II-B) with restriction enzyme *Hin*dIII followed by religation, and plasmid pJWK443 was generated by insertion of the 6.6-kb *Hin*dIII fragment of pJWK20 into the *Hin*dIII site of pBR322.

Inverted repeat DNA on the 75-kb plasmid of HD73. To test for the possible presence of inverted repeat sequences, the 75-kb plasmid was isolated by sucrose gradient sedimentation, linearized, denatured, allowed to renature briefly, and digested with S1 nuclease as described above. Electrophoretic analysis (Fig. 3, lane b) showed that there were two S1-resistant bands, presumably representing inverted repeat snap-back sequences. One band was ca. 2,150 bp in length ("IR2150") and the other was about 1,750 bp long ("IR1750"). As a control, ColE1::Tn*5* DNA, linearized at the single *Eco*RI site and subjected to the same treatment as the 75-kb plasmid, yielded the well-characterized (1) 1,535-bp inverted repeat DNA of Tn*5* (faint band in Fig. 3, lane a).

Mapping the position of inverted repeat DNA. As shown in Fig. 3, IR2150 and IR1750 snap-back fragments were well separated and therefore could be easily isolated from agarose gels. Restriction enzyme digestion of these isolated fragments as well as the fragments cloned from the 75-kb plasmid allowed the localization of the repeated DNA sequences as shown in Fig. 1D. A common 1.65-kb *Cla*I restriction fragment was found on IR2150 snap-back DNA (Fig. 4) and on recombinant plasmids pJWK22, pJWK21, and pJWK014 (Fig. 4, lanes b, c, and d, respectively). Densitometer scanning of the negative of Fig. 4 revealed that this 1.65-kb fragment was present in four copies in a *Cla*I digest of the 75-kb plasmid (Fig. 4, lane a). The restriction digestion analysis and the mapping data (Fig. 1B) indicated that one copy of IR2150 is present on each of the *Bgl*II-B, -D, and -E fragments of the 75-kb plasmid. The fourth 1.65-kb *Cla*I fragment is present on pJWK443 (Fig. 1B and C) and, as shown in the next section, is not homologous to IR2150. Similar experiments were performed comparing *Kpn*I digests of IR1750 DNA, DNA from plasmid pJWK25,

and DNA from the 75-kb plasmid (Fig. 4, lanes f, g, and h). Two copies of IR1750 are indicated by the restriction mapping data (Fig. 1B) and by densitometer scanning of the *Kpn*I band produced from IR1750 in the digest of pJWK25 (Fig. 4, lane g).

A more precise determination of the positions on the map of the ends of IR2150 and IR1750 was obtained by digestion of purified preparations of these snap-back sequences with various restriction enzymes. For example, the ends of IR2150 were mapped after digestion with *Eco*RI, whereas the position of IR1750 was mapped after digestion with *Cla*I (data not shown). The orientation of each repeated DNA sequence was determined from the order of internal restriction sites: the *Cla*I, *Eco*RI, and *Pvu*II sites of IR2150 and the *Kpn*I, *Pst*I, and *Cla*I sites of IR1750 (Fig. 1D). The third and fourth copies of IR1750 shown in Fig. 1D are discussed below.

Presence of IR2150 and IR1750 at other sites on the 75-kb plasmid. The 75-kb plasmid was analyzed for additional copies of IR2150 and IR1750 by hybridization with two cloned probes. The first was obtained by inserting the *Cla*I fragment contained within IR2150 on pJWK014 into the *Cla*I site of pBR322, and the second was constructed by inserting the *Kpn*I fragment contained within one of the copies of IR1750 on pJWK25 into the *Kpn*I site of pUC18. The positions of the cloned fragments are shown on the maps in Fig. 1C and D. These fragments were isolated, labeled with ^{32}P in vitro, and hybridized to restriction endonuclease

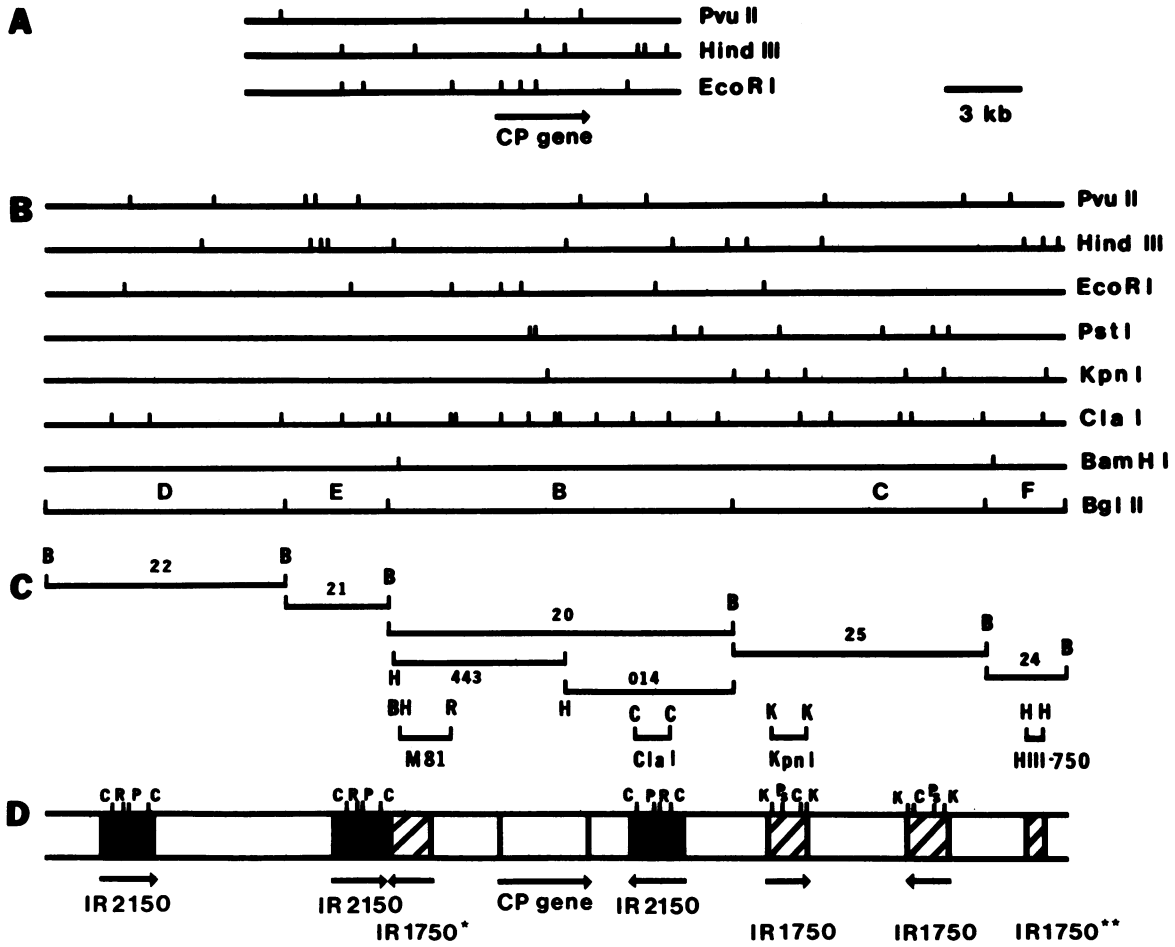

FIG. 1. Position of the crystal protein gene, IR2150 and IR1750, and cloned fragments on a restriction map of about 40 kb from the 75-kb plasmid of *B. thuringiensis* subsp. *kurstaki* HD73. (A) Restriction map showing the location and orientation of the crystal protein gene and intragenic restriction fragments on the recombinant plasmid pES1 (20, 26). The map shows the plasmid linearized at a *Sal*I site. The 1.1-kb *Hin*dIII fragment from entirely within the crystal protein gene was employed as a hybridization probe in experiments reported in this communication. (B) Restriction map of the 40 kb of DNA around the crystal protein gene from the 75-kb plasmid of *B. thuringiensis* subsp. *kurstaki* HD73. The map was constructed from the individual maps of the *Bgl*II fragments labeled B through F. The unmapped portion of the plasmid is the 35-kb *Bgl*II-A fragment. (C) Position of cloned fragments on the restriction map. Fragments labeled 22, 21, 20, 25, 24, 443, and 014 represent the insertions of recombinant plasmids as described in the text. Restriction enzyme sites are indicated by H for *Hin*dIII and B for *Bgl*II. Fragments labeled M81 (R, *Eco*RI; and BH, *Bam*HI) and HIII-750 were employed in hybridization experiments to position and orient IR1750* and IR1750**. The fragments labeled *Cla*I and *Kpn*I indicate the position of the insertions employed as hybridization probes specific for IR2150 and IR1750, respectively. (D) Diagram indicating the positions of the copies of IR2150 and IR1750. Pertinent restriction sites are indicated by the following letters: C, *Cla*I; R, *Eco*RI; P, *Pvu*II; K, *Kpn*I; and Ps, *Pst*I. The direction of transcription of the crystal protein gene is indicated by an arrow. Arrows also indicate the orientation of the copies of IR2150 and IR1750 relative to each other. The length of IR1750* has not been precisely determined.

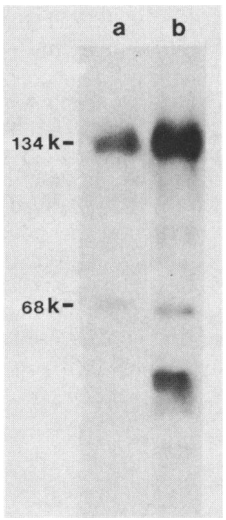

FIG. 2. Autoradiograph of a radioimmunoassay for polypeptides which react with antibodies to the crystal protein after sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transfer to nitrocellulose. (a) 100 μg of *E. coli* extract from cells carrying pJWK20; (b) 1 μg of solubilized *B. thuringiensis* subsp. *kurstaki* HD73 crystal protein. The numbers in the margin indicate size in daltons, and the bands below 134,000 may result from proteolytic processing of the 134,000-dalton polypeptide.

digests of total HD73 plasmid DNA and total cell DNA (the results obtained in the latter experiments are discussed in the next section).

Hybridization of the IR2150 probe to restriction digests of HD73 plasmid DNA (Fig. 5A) indicated the presence of three copies of this inverted repeat sequence. Thus, lane a of Fig. 5A shows the hybridization with three major *Bgl*II

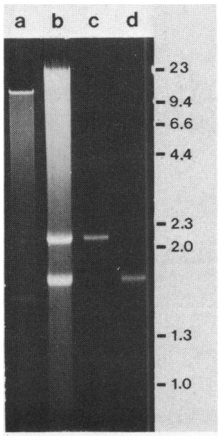

FIG. 3. Electrophoretic analysis of inverted repeat DNA from the 75-kb plasmid of HD73 and ColE1::Tn*5*. Photograph of a vertical 0.7% agarose gel stained with ethidium bromide showing the DNA fragments resulting from snap-back of inverted repeat sequences on the plasmids. Lane a, The 1,535-bp inverted repeat of Tn*5* present on ColE1 (the larger fragment at about 12 kb represents the linear plasmid); lane b, the 2,150- and 1,750-bp fragments (IR2150 and IR1750); lane c, the 2,150-bp fragment isolated from a preparative gel; and lane d, the 1,750-bp fragment isolated from a preparative gel. The kb size standards in the margin are the *Hin*dIII fragments of phage lambda and the *Rsa*I fragments of the replicative form of phage M13mp8.

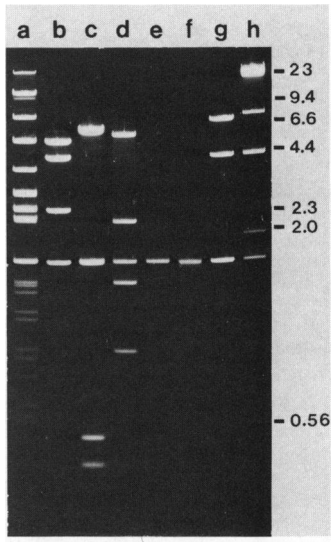

FIG. 4. Electrophoretic analysis of restriction digests of snap-back DNA and recombinant plasmids carrying fragments of the 75-kb plasmid. Photograph of a 0.7% vertical agarose gel stained with ethidium bromide. Lane a, *Cla*I digest of sucrose gradient-purified 75-kb plasmid DNA; lane b, *Cla*I digest of pJWK22 DNA; lane c, *Cla*I digest of pJWK21 DNA; lane d, *Cla*I digest of pJWK014 DNA; lane e, *Cla*I digest of gel-purified IR2150 snap-back DNA (lane c of Fig. 3); lane f, *Kpn*I digest of gel-purified IR1750 snap-back DNA (lane d of Fig. 3); lane g, *Kpn*I digest of pJWK25 DNA; and lane h, *Kpn*I digest of sucrose gradient-purified 75-kb plasmid DNA. The kb size standards in the margin are the *Hin*dIII fragments of phage lambda.

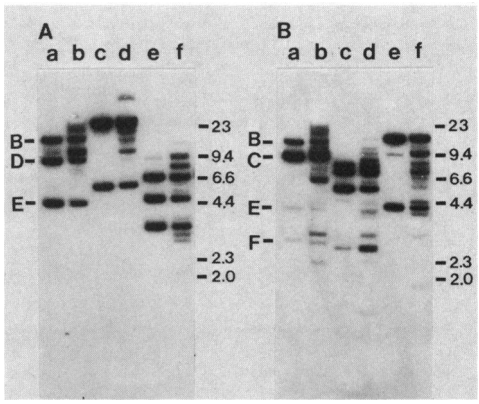

FIG. 5. Comparison of restriction endonuclease digests of HD73 plasmid DNA and HD73 total cell DNA by hybridization with IR2150 and IR1750 probes. (A) The autoradiograph resulting from hybridization of the ^{32}P-labeled *Cla*I fragment of IR2150 to the following digested DNA transferred to nitrocellulose. Lane a, plasmid DNA, *Bgl*II; lane b, total cell DNA, *Bgl*II; lane c, plasmid DNA, *Pst*I; lane d, total cell DNA, *Pst*I; lane e, plasmid DNA, *Hin*dIII; and lane f, total cell DNA, *Hin*dIII. (B) The autoradiograph resulting from hybridization of ^{32}P-labeled *Kpn*I fragment of IR1750 to the following digested DNAs transferred to nitrocellulose. Lane a, plasmid DNA, *Bgl*II; lane b, total cell DNA, *Bgl*II; lane c, plasmid DNA, *Pvu*II; lane d, total cell DNA, *Pvu*II; lane e, plasmid DNA, *Eco*RI; and lane f, total cell DNA, *Eco*RI. The kb size standards in the margin are the *Hin*dIII fragments of phage lambda. The letters in the left margins of (A) and (B) indicate the *Bgl*II fragments of the 75-kb plasmid.

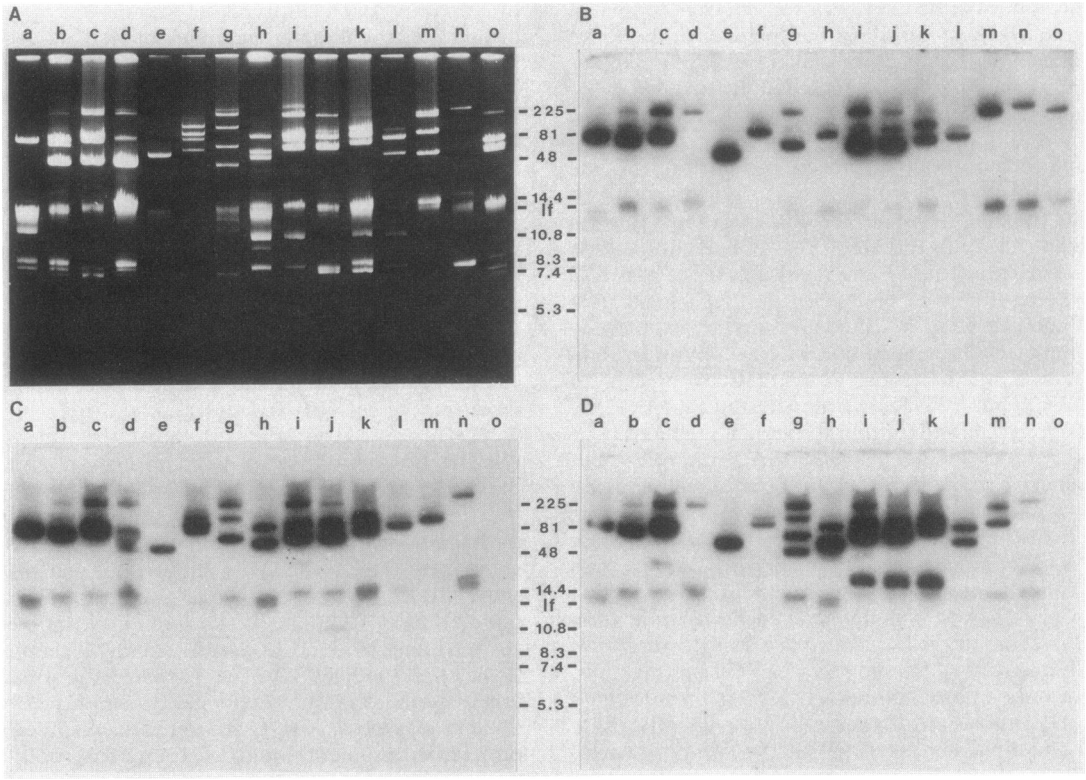

FIG. 6. Analysis of the plasmid DNA from 15 strains of *B. thuringiensis* hybridized with crystal protein gene probe, and with IR2150 and IR1750 probes. (A) Photograph of a 0.7% agarose gel stained with ethidium bromide, showing plasmid preparations purified by cesium chloride-ethidium bromide density centrifugation. The numbers in the center indicate size in kb as determined for *kurstaki* HD1 (5) and *thuringiensis* HD2 (5, 10). The position of linearized DNA fragments is indicated by lf. Lane a, *kurstaki* HD73; lane b, *kurstaki* HD1; lane c, *kurstaki* HD244; lane d, *kurstaki* HD1-Dipel; lane e, *sotto*; lane f, *darmstadiensis*; lane g, *tolworthi*; lane h, *thuringiensis* HD2; lane i, *thuringiensis* Berliner 1715; lane j, *thuringiensis* HD120; lane k, *thuringiensis* HD290; lane l, *thuringiensis* F; lane m, *morrisoni*; lane n, *galleriae*; and lane o, *alesti*. (B) Location of sequences homologous to a cloned crystal protein gene. Autoradiograph resulting from the transfer of the DNA in the gel of (A) to nitrocellulose and hybridization with ^{32}P-labeled 1.1-kb *Hin*dIII fragment from pES1 (Fig. 1A). The lane designations are as in (A). A previously undetected band (10) of greater than 225 kb was found in plasmid preparations from *thuringiensis* Berliner 1715 (lane i), *thuringiensis* HD290 (lane k), and *thuringiensis* F (lane l). The data in (B) show that this band does not carry a homologous crystal protein gene. (C) Location of sequences homologous to IR2150. Autoradiograph resulting from the transfer of the DNA in a gel identical to the one shown in (A) to nitrocellulose and hybridization with the ^{32}P-labeled *Cla*I fragment (Fig. 1C) of IR2150. The lane designations and size standards are as in (A). (D) Location of sequences homologous to IR1750. Autoradiograph resulting from the transfer of the DNA in a gel identical to the one shown in (A) to nitrocellulose and hybridization with the ^{32}P-labeled *Kpn*I fragment (Fig. 1C) of IR1750. The lane designations and size standards are as in (A).

bands (*Bgl*II-B, -D, and -E), lane c shows the hybridization of two major *Pst*I bands, and lane e shows the hybridization of three major *Hin*dIII bands. Hybridization to the *Bgl*II-C fragment was not detected, indicating that, at the stringency employed in these experiments (25% bp mismatch [7]), no homology exists between IR2150 and IR1750. The positions of the three copies of IR2150 deduced from these experiments and from the restriction mapping data discussed above are shown in Fig. 1D. The 75-kb plasmid band in HD73 has been reported to be a doublet with the more prominent member (described in this communication) carrying the crystal protein gene (3, 5, 10). The minor bands seen in lanes a and e of Fig. 5 may represent hybridization to fragments of the other, less prominent member of the doublet. The corresponding minor bands were visible on the ethidium bromide-stained agarose gel, demonstrating that the hybridization does not result from partly homologous or partial copies of IR2150 on the *Bgl*II-A, -C, or -F fragments. It does not seem likely that the minor bands resulted from partial digestion by restriction enzymes, since the same results were obtained when digests were supplemented with additional amounts of enzyme or incubated for a longer time, or both.

Parallel experiments with HD73 plasmid DNA and the IR1750 probe revealed strong hybridization with *Bgl*II-B and -C and faint hybridization with the *Bgl*II-E and -F fragments (Fig. 5B, lane a). It should be noted that neither of the two probes for the repeated sequences hybridized to the unmapped fragment, *Bgl*II-A (Fig. 5), or to the other plasmids in HD73 (shown in a later section). The lack of hybridization to the *Bgl*II-D fragment again indicates that there is no homology between IR2150 and IR1750. The results obtained with *Pvu*II digests (Fig. 5B, lane c; four major bands) and *Eco*RI (Fig. 5B, lane e; two major bands) suggested that the faint hybridization to *Bgl*II-E results from the presence of a third copy of the repeated sequence (named IR1750*) which overlaps *Bgl*II-B and -E and that there is a fourth copy (named IR1750**) of lower homology present on the *Bgl*II-F

fragment. The minor band in lane e of Fig. 5B may represent hybridization to a fragment of the other member of the 75-kb plasmid doublet. This band was also visible in minor amounts on the ethidium bromide-stained agarose gel.

To locate IR1750* more precisely, as well as to determine its orientation and approximate length, we performed various cross-hybridization experiments. A portion of the IR1750* sequence was cloned by insertion of a 2.0-kb *Bam*HI-*Eco*RI fragment from pJWK20 into M13mp8 to yield the recombinant phage pJWKM81 (Fig. 1C). Hybridization of this 2.0-kb fragment to *Pst*I and *Kpn*I digests of pJWK25 indicated homology with the ca. 500-bp *Kpn*I-*Pst*I fragment of the IR1750 copies on pJWK25 (data not shown), thereby establishing the orientation shown in Fig. 1D. Hybridization of the internal *Kpn*I fragment of IR1750 to digests of plasmids pJWK20 and pJWK21 demonstrated that IR1750* is at least 1,400 bp long (data not shown). For illustration, the length of IR1750* in Fig. 1D is shown as 1,750 bp, but exact determination of the size of this sequence will require a more detailed analysis. The faint hybridization with the *Bgl*II-E fragment (Fig. 5B, lane a) is probably due to the fact that only a small amount (less than 300 bp) of the total IR1750* sequence is present on this fragment.

The length of IR1750** was determined by hybridization of the *Kpn*I fragment of IR1750 to a *Hin*dIII digest of pJWK24. Homology was detected to a ca. 750-bp fragment, indicating that only a partial copy of IR1750 is present (data not shown). Hybridization of the ca. 750-bp *Hin*dIII fragment to *Pst*I and *Kpn*I digests of pJWK25 (data not shown) demonstrated that the end of IR1750 containing a *Cla*I site is present on pJWK24 as shown in Fig. 1D. Lastly, it is clear from the maps shown in Fig. 1B and D that considerable restriction site variations exist between the two copies of IR1750 located downstream from the crystal protein gene and IR1750* and IR1750**; e.g., IR1750* and IR1750** lack *Kpn*I sites.

Presence of IR2150 and IR1750 on HD73 chromosomal DNA. If sequences related to IR2150 and IR1750 are present on the chromosome, restriction digests of total cell DNA would contain not only the fragments arising from digestion of the 75-kb plasmid but one or more additional bands from the chromosome. Fig. 5A compares the hybridization of the IR2150 probe to *Bgl*II, *Pst*I, and *Hin*dIII digests of plasmid DNA (lanes a, c, and e) and total cell DNA (lanes b, d, and f). It is clear that digestion of total cell DNA with *Bgl*II yields at least two major bands in addition to those hybridized with plasmid DNA. Similarly, digestion of total cell DNA with *Pst*I yields three or four additional bands, and digestion with *Hin*dIII yields four to six additional bands. The enzymes used in these experiments do not cleave within the IR2150 sequence, and therefore, each additional band in digests of total cell DNA represents at least one copy of IR2150 located on the chromosome of HD73. The presence of partial digestion products appears unlikely, because the same results were obtained with samples that had been supplemented with additional amounts of enzyme or incubated for a longer time, or both. The variations in the number of fragments obtained with the different enzymes may reflect clustering of IR2150 sequences but it is clear that multiple copies are present on the chromosome.

Figure 5B shows a parallel experiment with the IR1750 probe. Compared with the digests of plasmid DNA, digests of total cell DNA revealed additional bands: about four for *Bgl*II (lane b), about four for *Pvu*II (lane d), and about seven for *Eco*RI (lane f). There are no *Eco*RI sites within the IR1750 sequences on the 75-kb plasmid. Therefore, digestion of total cell DNA with *Eco*RI should provide a minimum estimate of the number of copies of IR1750 on the chromosome. As in the analysis of the copies of IR1750 on the 75-kb plasmid, differences in hybridization intensity between the additional bands in a particular digest of total cell DNA may reflect sequence or size variation among copies of the repeated DNA.

Distribution of repeated DNA on plasmids from 15 additional *B. thuringiensis* strains. Figure 6A presents the plasmid profiles from 15 strains of *B. thuringiensis* representing the following eight subspecies: lanes a through d, *kurstaki*; lane e, *sotto*; lane f, *darmstadiensis*; lane g, *tolworthi*; lanes h through l, *thuringiensis*; lane m, *morrisoni*; lane n, *galleriae*; and lane o, *alesti*. The plasmid content and location of the crystal protein gene for these strains have been described in detail elsewhere (10) and are repeated here for the convenience of the reader. Fig. 6B shows the hybridization of a *Hin*dIII fragment from within the crystal protein gene of pES1 (Fig. 1A) to the plasmids shown in Fig. 6A after transfer to nitrocellulose. The crystal protein gene is present on large plasmids (greater than 45 kb) in each strain, and in some cases, the gene is on more than one plasmid in the same strain: e.g., *kurstaki* HD1 (lane b), *kurstaki* HD244 (lane c), *tolworthi* (lane g), *thuringiensis* Berliner 1715 (lane i), *thuringiensis* HD120 (lane j), and *thuringiensis* HD290 (lane k). Previously (10), we reported the presence of two homologous crystal protein genes on the 225-kb plasmid doublet of *kurstaki* HD1. We recently resolved the doublet on sucrose gradients, and we found that both genes were present on the smaller plasmid in the doublet (J. Kronstad, unpublished data).

Figure 6C and D presents the results of hybridization of the IR2150 and IR1750 probes, respectively, to blots of gels identical to the one shown in Fig. 6A. The following points can be made from the data presented. (i) Sequences related to IR2150 and IR1750 were found on the plasmids carrying the crystal protein gene in every strain except *alesti* (lane o, no hybridization). The plasmid carrying the crystal protein gene in *morrisoni* (lane m) only carries sequences related to IR1750. In Fig. 6C and D, hybridization to the 225-kb plasmids in *kurstaki* HD1 (lane b), *thuringiensis* HD120 (lane j), and *thuringiensis* HD290 (lane k) was weak, probably due to low amounts of these plasmids in the preparations. Upon longer exposure of the autoradiograph, these plasmids were found to hybridize to both IR2150 and IR1750 probes. (ii) Hybridization was found to several plasmids that do not carry this crystal protein gene; e.g., the IR2150 probe hybridized to the ca. 71-kb plasmid of *kurstaki* HD1 Dipel (Fig. 6C, lane d) and the about 90-kb plasmid of *morrisoni* (Fig. 6C, lane m), and the IR1750 probe hybridized to the 14.3-kb plasmid in *thuringiensis* Berliner 1715, HD120, and HD290 (Fig. 6D, lanes i, j, and k), the 45-kb plasmid in *thuringiensis* HD2 and F (Fig. 6D, lanes h and l), the 90-kb plasmid in *morrisoni* (Fig. 6D, lane m) and the 45-kb plasmid in *tolworthi* (Fig. 6D, lane g). It is interesting to note that the 14.3-kb plasmid of *thuringiensis* Berliner 1715 has been shown to carry sequences homologous to large plasmids in the same strain, in *kurstaki* HD1 and in a *subtoxicus* strain (11). The presence of IR1750 may explain these results. Minor hybridization bands, e.g., at ca. 11 kb in *kurstaki* HD73 (Fig. 6C, lane a), at ca. 9 kb in *thuringiensis* HD120 (Fig. 6C, lane j), and at ca. 38 kb in *kurstaki* HD244 (Fig. 6D, lane c), may result from deletions or rearrangements between copies of the repeated DNA. (iii) Both IR2150 and IR1750 probes hybridized strongly with linear DNA fragments present in the plasmid preparations. These fragments

are probably derived by breakage of large plasmids or the chromosome. (iv) Many plasmids (generally smaller than 45 kb) were found not to carry IR2150 or IR1750 sequences or the crystal protein gene. Thus, the data shown in Fig. 6B, C, and D indicate a strong correlation between the presence of the crystal protein gene and the presence of IR2150 and IR1750 in all strains except *alesti* HD4.

DISCUSSION

The crystal protein gene has been cloned from the 75-kb plasmid of *B. thuringiensis* subsp. *kurstaki* HD73, and a restriction map of ca. 40 kb around the gene has been constructed. Two sets of inverted repeat DNA sequences were found on the 75-kb plasmid, one of 1,750 bp (IR1750) and one of 2,150 bp (IR2150). One partial and three intact copies of IR1750 sequences were found within the mapped region, and both direct and inverted copies flank the crystal protein gene. IR2150 was found in three copies in the mapped region, but only inverted copies flank the crystal protein gene. Hybridization probes containing internal fragments from IR1750 or IR2150 revealed the presence of multiple copies of these sequences on the chromosome of HD73 and also on plasmids carrying the crystal protein gene in 14 strains representing seven *B. thuringiensis* subspecies. A strain of *alesti* was also analyzed and found not to carry either IR1750 or IR2150 on plasmid DNA. In addition, sequences homologous to IR1750 or IR2150 were found on a number of plasmids which did not hybridize with a probe derived from our cloned crystal protein gene, including the 14.3-kb plasmids of *thuringiensis* Berliner 1715, HD120, and HD290. These plasmids contain sequences homologous to IR1750 but not IR2150.

Previously, Lereclus et al. (11) reported sequence homology between the large plasmids from 11 *B. thuringiensis* strains, including *kurstaki* HD1 and *thuringiensis* Berliner 1715. In addition, they found conserved sequences between the 14.3-kb plasmid of *thuringiensis* Berliner 1715 and large plasmids in the same strain and in *kurstaki* HD1. Furthermore, homology was detected between large plasmids in *kurstaki* HD1, *thuringiensis* Berliner 1715, and *subtoxicus* and chromosomal DNA from a plasmidless mutant of *kurstaki* HD1. Our findings concerning the wide distribution of IR2150 and IR1750 sequences on large plasmids, on the 14.3-kb plasmid in *thuringiensis* Berliner 1715 (IR1750), and on the chromosome in *kurstaki* HD73 could account for the homology detected by Lereclus et al. (11). The presence of additional homologous sequences on the large plasmids in various *B. thuringiensis* strains has not been excluded.

The distribution of IR1750 and IR2150 sequences on the 75-kb plasmid and the chromosome of HD73, as well as their almost ubiquitous presence on large plasmids in 14 strains, suggests that they are, in fact, insertion sequences. Generally, insertion sequences can mediate the transposition of DNA that they flank in inverted or direct orientations (2). For example, Tn*9* is flanked by direct copies of IS1 (2). Thus, if IR1750 and IR2150 are insertion sequences, they could mobilize the crystal protein gene and flanking DNA sequences as well as provide regions of homology for recombination. It will be interesting to see whether crystal protein genes in other strains are flanked by the same complex organization of IR2150 and IR1750.

Lereclus et al. (12) also reported the in vivo insertion of a 4.5-kb segment of DNA, originating from the 81-kb plasmid of *B. thuringiensis* subsp. *kurstaki* KT_0, into the *Streptococcus faecalis* plasmid pAMβ1. *B. thuringiensis* subsp. *kurstaki* KT_0 has a similar plasmid profile (12) to HD73; the 81-kb plasmid of strain KT_0 may be equivalent to the 75-kb plasmid of HD73 described in this communication. However, the size of the insert, 4.5 kb, does not correspond to the size of IR1750 or IR2150.

In an extensive series of curing experiments, Gonzalez et al. (5) discovered that rearrangements among *B. thuringiensis* plasmids were fairly common. Typically, plasmids were lost with the concomitant appearance of smaller putative derivatives. Carlton and Gonzalez (3) and Gonzalez et al. (5) have also reported the integration of plasmids into the chromosome. The presence of regions of homology on the *B. thuringiensis* plasmids and chromosome, such as IR1750, IR2150, and the crystal protein gene, could account for the extensive recombination reported. Interestingly, during plasmid transfer among various *B. thuringiensis* strains and from *B. thuringiensis* to *Bacillus cereus*, chromosomal markers are transferred (3; S. A. Minnich, personal communication). It is possible that *B. thuringiensis* plasmids could integrate into the chromosome via IR1750 or IR2150 and mediate chromosome transfer in a fashion analogous to the F factor of *E. coli* (8).

ACKNOWLEDGMENTS

This research was supported by Public Health Service grants GM-20784 and GM-26100 from the National Institutes of General Medical Sciences and by a grant from the Cetus Corp. J.W.K. was supported by the Molecular and Cellular Biology Training Program, Public Health Service National Research Award GM-07270-08 from the National Institutes of Health. H.R.W. is a recipient of Public Health Service Research Career Award K6-GM-442 from the National Institutes of Medical Sciences.

We thank Ernie Schnepf, Ken McLean, and Bill Widner for criticisms during the preparation of this manuscript.

LITERATURE CITED

1. **Berg, D. E., and C. M. Berg.** 1983. The prokaryotic transposable element Tn*5*. Biotechnology **1**:417–435.
2. **Calos, M. P., and J. H. Miller.** 1980. Transposable elements. Cell **20**:579–595.
3. **Carlton, B. C., and J. M. Gonzalez, Jr.** 1984. The genetics and molecular biology of *Bacillus thuringiensis*. *In* D. A. Dubnau (ed.), The molecular biology of the bacilli, vol. 2. Academic Press, Inc., New York. (In press.)
4. **Debabov, V. G., O. I. Khlebalina, F. P. Galushka, and I. A. Sladkova.** 1980. Plasmids of *Bacillus thuringiensis* var. *galleriae*, strain 612. Mol. Biol. **14**:818–823.
5. **Gonzalez, J. M., Jr., H. T. Dulmage, and B. C. Carlton.** 1981. Correlation between specific plasmids and delta-endotoxin production in *Bacillus thuringiensis*. Plasmid **5**:351–365.
6. **Held, G. A., L. A. Bulla, Jr., E. Farrari, J. Hoch, A. I. Aronson, and S. A. Minnich.** 1982. Cloning and localization of the lepidopteran protoxin gene of *Bacillus thuringiensis* subsp. *kurstaki*. Proc. Natl. Acad. Sci. U.S.A. **79**:6065–6069.
7. **Howley, P. M., M. A. Israel, M. Law, and M. A. Martin.** 1979. A rapid method for detecting and mapping homology between heterologous DNAs. J. Biol. Chem. **254**:4876–4883.
8. **Hu, S., E. Ohtsubo, and N. Davidson.** 1975. Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*: structure of F13 and related F-primes. J. Bacteriol. **122**:749–763.
9. **Klier, A., F. Fargette, J. Ribier, and G. Rapoport.** 1982. Cloning and expression of the crystal protein genes from *Bacillus*

thuringiensis strain *berliner* 1715. EMBO J. **1**:791–799.
10. **Kronstad, J. W., H. E. Schnepf, and H. R. Whiteley.** 1983. Diversity of locations for *Bacillus thuringiensis* crystal protein genes. J. Bacteriol. **154**:419–428.
11. **Lereclus, D., M. M. Lecadet, J. Ribier, and R. Dedonder.** 1982. Molecular relationships among plasmids of *Bacillus thuringiensis*: conserved sequences through 11 crystaliferous strains. Mol. Gen. Genet. **186**:391–398.
12. **Lereclus, D., G. Menou, and M. M. Lecadet.** 1983. Isolation of a DNA sequence related to several plasmids from *Bacillus thuringiensis* after mating involving the *Streptococcus faecalis* plasmid pAMβ1. Mol. Gen. Genet. **191**:307–313.
13. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.
14. **Maniatis, T., R. C. Hardison, E. Lacy, J. Laver, C. O'Connel, D. Quon, G. K. Sim, and A. Efstratiadis.** 1978. The isolation of structural genes from libraries of eucaryotic DNA. Cell **15**:687–701.
15. **Maniatis, T., A. Jeffrey, and D. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage lambda. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
16. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
17. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.
18. **Ohtsubo, H., and E. Ohtsubo.** 1976. Isolation of inverted repeat sequences including IS1, IS2, and IS3 in *Escherichia coli* plasmids. Proc. Natl. Acad. Sci. U.S.A. **73**:2316–2320.
19. **Schesser, J. H., K. J. Kramer, and L. A. Bulla, Jr.** 1977. Bioassay for homogeneous parasporal crystal of *Bacillus thuringiensis* using the tobacco hornworm, *Manduca sexta*. Appl. Environ. Microbiol. **33**:878–880.
20. **Schnepf, H. E., and H. R. Whiteley.** 1981. Cloning and expression of the *Bacillus thuringiensis* crystal protein gene in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **78**:2893–2897.
21. **Smith, H. O.** 1980. Recovery of DNA from gels. Methods Enzymol. **65**:371–380.
22. **Studier, F. W.** 1973. Analysis of bacteriophage T7 early RNAs and proteins on slab gels. J. Mol. Biol. **79**:237–248.
23. **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. U.S.A. **76**:4350–4354.
24. **Viera, 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.
25. **Whiteley, H. R., J. W. Kronstad, H. E. Schnepf, and J. P. DesRosier.** 1982. Cloning the crystal protein gene from *Bacillus thuringiensis* in *E. coli*, p. 131–144. *In* J. A. Hoch, S. Chang, and A. T. Ganesan (ed.), Molecular cloning and gene regulation in bacilli. Academic Press, Inc., New York.
26. **Wong, H. C., H. E. Schnepf, and H. R. Whiteley.** 1983. Transcriptional and translational start sites for the *Bacillus thuringiensis* crystal protein gene. J. Biol. Chem. **258**:1960–1967.