Polyhedrin Gene of *Bombyx mori* Nuclear Polyhedrosis Virus

KOSTAS IATROU,* KENICHI ITO, AND HALINA WITKIEWICZ

Department of Medical Biochemistry, Faculty of Medicine, The University of Calgary, Calgary, Alberta
T2N 4N1, Canada

Received 18 October 1984/Accepted 14 January 1985

A portion of the genome of the nuclear polyhedrosis virus of *Bombyx mori* has been cloned. This part of the viral genome contains the gene encoding the viral occlusion body protein, polyhedrin. The polyhedrin gene has been sequenced in its entirety together with some of its 5' and 3' flanking sequences. The primary structure of polyhedrin predicted from the nucleotide sequence of the gene was found to be somewhat different from the one reported previously for the authentic protein (E. A. Kozlov, T. L. Levitina, N. M. Gusak, and S. B. Serebryani, *Bioorg. Khim.*, 7:1008-1015, 1981; S. B. Serebryani, T. L. Levitina, M. L. Kautzman, Y. L. Radavski, N. M. Gusak, M. N. Ovander, N. V. Sucharenko, and E. A. Kozlov, *J. Invertebr. Pathol.*, 30:442-443, 1977). Comparison of the primary structures of the polyhedrin of the nuclear polyhedrosis virus of *B. mori* with that of *Autographa californica* suggests that considerable selective pressure has been exercised at the protein level during evolution. Nucleotide sequence comparisons of the two structural genes reveal that the coding sequences have diverged significantly through the accumulation of silent and replacement substitutions. In contrast, a remarkable degree of sequence conservation was found to exist in the domains corresponding to the 5' and 3' noncoding regions of the polyhedrin mRNAs.

Nuclear polyhedrosis viruses (NPVs) represent a group of insect viruses (family *Baculoviridae*, genus *Baculovirus*) whose virions are embedded into polyhedron-shaped occlusion bodies or polyhedra in the nuclei of host cells (for extensive reviews see references 4, 13, and 19). The genetic information of baculoviruses is stored in the form of a relatively large molecule of covalently closed double-stranded DNA ranging in size from 50×10^6 to 100×10^6 daltons.

NPVs have been recognized for many years as the infectious agents that cause acute diseases in a wide range of insects, including the silkworm, *Bombyx mori* (4), and the first tissue culture propagation of the NPV of *B. mori* (BmNPV) was reported as early as 1935 (47). Occlusion bodies were shown to be made almost exclusively of a single protein, polyhedrin, which can be solubilized by treatment of the polyhedra with weak alkaline solutions (3). In *B. mori*, about 350 nucleocapsids enveloped singly or in multiples of two to five are occluded in each occlusion body.

Although NPVs may infect insects other than their normal hosts, their host range is usually restricted to different species or (at most) families within an order. Only very rarely can the host range extend to different insect orders or to different arthropod classes (9). There has only been one case in which infection of a vertebrate cell culture was caused by an NPV (21), but that case has not been studied any further. In contrast, repeated attempts by other investigators to propagate NPVs in a wide variety of vertebrate cells were unsuccessful (18, 25, 46). Because of their restricted host range and their effectiveness in infecting their hosts, baculoviruses have been considered as potential biological pest control agents (33, 45). This has in turn stimulated the molecular characterization of some NPVs, particularly the NPV of *Autographa californica* (AcNPV) (13).

A number of mutations have been found in AcNPV which result in the formation of abnormal, few, or no occlusion bodies in the nuclei of infected cells (7, 12, 27, 53). In addition, it has been established that extensive serial pas-

sage of NPVs in cell cultures results in a dramatic decrease in virus occlusion (14, 22, 31). In such cases the production of nonoccluded virus usually remains unaffected, and non-occluded virus becomes the predominant viral form in the infected cells. On the basis of these observations it has been concluded that occlusion body formation and, presumably, polyhedrin synthesis are not required for viral propagation in laboratory tissue cultures.

With the demonstration that polyhedrin is a protein encoded by a viral gene (50) and with the advent of recombinant DNA technology, it became apparent that the polyhedrin gene might represent a site of the viral genome amenable to genetic manipulation, such as the introduction of foreign genes into the DNA of the virus. Experiments in which the structure of the polyhedrin gene from spontaneous, chemically induced, or in vitro-constructed mutants was examined showed mutational changes in the gene or its immediate vicinity (8, 11, 40), thus supporting the hypothesis that the polyhedrin gene represents a convenient target for genetic engineering. Finally, polyhedrin is synthesized in the host cells late in infection in large amounts and this in turn reflects a corresponding abundance in the amount of accumulated polyhedrin mRNA (1, 42), probably achieved through high rates of mRNA synthesis. This suggested that heterologous genes might be appropriately introduced into the genomes of baculoviruses and expressed effectively under the control of the polyhedrin promoter. The use of AcNPV as a vector for the expression of foreign genes under polyhedrin promoter control was successfully attempted recently and resulted in the production of human β -interferon and *Escherichia coli* β -galactosidase in large amounts in insect tissue culture cells (34, 41).

Unfortunately, the molecular characterization of BmNPV has lagged far behind that of AcNPV. Although some controversial reports on the nature of the viral genome have appeared in the literature (reviewed in reference 19), no accurate genetic map or measurements of the size of BmNPV DNA have been reported due to the lack of published work based on current cloning methodology. We are interested in the mechanisms regulating chorion gene expression in the

* Corresponding author.

silkmoth *B. mori* (for a recent review see reference 16) and feel that BmNPV may be used effectively as a vehicle for introducing authentic or mutagenized silkmoth genes into the tissues in which they are normally expressed in order to study their expression *in vivo*. To test models of differential regulation of chorion gene expression (24), we would like to recombine well-characterized chorion genes into the genome of BmNPV, using the vicinity of the polyhedrin gene as a target site, and to infect follicular cells *in vivo* with the recombinant virus. As a first step in this direction we report here the localization, cloning, and complete nucleotide sequence of the polyhedrin gene of BmNPV and its vicinity. The primary structure of polyhedrin predicted from the nucleotide sequence of the gene is not identical to that derived from the protein itself (26, 39). Comparison of the polyhedrin genes of BmNPV and AcNPV indicates that although a large degree of point mutations were fixed in some parts of the genes during evolution, other domains of the genes and the proteins themselves have been subjected to considerable selective pressure.

MATERIALS AND METHODS

Cells and virus. *B. mori* tissue culture cells (Bm-5 [17]) and BmNPV passed once in Bm-5 cells were kindly supplied by J. L. Vaughn, Insect Pathology Laboratory, U.S. Department of Agriculture, Beltsville, Md. Cells were maintained at 28°C in complete IPL-41 growth medium (52) supplemented with 0.24 μ M ZnSO₄ and 16 μ M AlK(SO₄)₂ in the absence of antibiotics. The cells were subcultured weekly at a seeding density of 0.2×10^6 cells per ml (usually a 1:5 dilution of a 1-week-old culture). Infection of the cells with BmNPV was carried out by removing the medium from 2×10^7 cells and replacing it with 5 ml of a nonoccluded virus inoculum of tissue culture origin containing 1×10^7 to 5×10^7 PFU per ml. After 1 h the viral inoculum was removed and replaced with 25 ml of fresh medium containing gentamycin (50 μ g/ml). Cells were collected by low-speed centrifugation 3 to 4 days later when more than 98% of them exhibited a large number of viral occlusion bodies in their nuclei. Cellular pellets were used as the source of occluded virus, whereas the supernatants were used as the source of nonoccluded virus and as inocula for further infections.

***A. californica* polyhedrin gene probe.** A cloned *Hind*III fragment of the polyhedrin gene of AcNPV (*Hind*III-V) in pBR322 (clone pM5 *Hind*V) was kindly supplied by Eric B. Carstens, Department of Microbiology and Immunology, Queen's University, Kingston, Ontario, Canada. This cloned fragment is 937 base pairs long and contains the coding region of the gene downstream from amino acid number 83 (nucleotide 251 of the polyhedrin gene sequence published in reference 23) and almost the entire 3' noncoding part of the gene (E. B. Carstens, unpublished data).

Preparation of occlusion bodies. At the end of the infection period cells were pelleted by centrifugation and washed once with phosphate-buffered saline (10). Cellular pellets were solubilized with 0.4% sodium dodecyl sulfate–10 mM Tris-HCl (pH 7.8) by gentle rocking for 2 h (approximately 5 ml of solution per 4×10^7 cells), and the cellular suspension was layered on a 30-ml cushion of 65% (wt/vol) sucrose in 10 mM Tris (pH 7.8)–10 mM EDTA. After centrifugation at $110,000 \times g$ (24,000 rpm) in the SW27 rotor of a Beckman ultracentrifuge for 4 h at 15°C, the supernatant was removed partly by aspiration and finally by decanting, and the pelleted occlusion bodies were suspended in a small volume of distilled water with the aid of a Pasteur pipette. Finally, the occlusion body suspension was made 0.25 M with respect to

NaCl and centrifuged at $17,000 \times g$ (12,000 rpm) in the SS-34 rotor of a Sorvall centrifuge. The pelleted occlusion bodies were washed twice with 0.25 M NaCl as described above, and the final pellets were stored at -20 to -70°C until used for the extraction of viral DNA.

Viral DNA isolation. Occlusion bodies were solubilized in 0.1 M Na₂CO₃–10 mM EDTA–0.1 M NaCl (pH 10.8) (occlusion bodies from 2.5×10^7 cells per ml of buffer) with gentle swirling at room temperature over a period of 2 h. At the end of the process the volume of the solution was increased by 50% by the addition of distilled water, and the solution was finally made 1% with respect to sodium dodecyl sulfate. A small amount of insoluble matrix was removed by centrifugation, and the supernatant was extracted exhaustively with phenol and chloroform-isoamyl alcohol (24:1, vol/vol). After the extractions, the aqueous phase was concentrated to a volume equal to that of the Na₂CO₃ solution used at the beginning of the process, made 0.25 M with respect to ammonium acetate, and spun at $30,000 \times g$ (17,000 rpm) in a Sorvall centrifuge. The supernatant was precipitated with 2.5 volumes of ethanol. After being washed with 70% ethanol, the DNA pellets were dissolved in 5 mM Tris-HCl (pH 7.8)–0.1 mM EDTA and stored at 4 or -40°C .

The Na₂CO₃-insoluble matrix was also partially solubilized in a mixture of equal volumes of 0.1 M Na₂CO₃–10 mM EDTA–0.1 M NaCl (pH 10.8) and 7 M urea–0.135 M NaCl–10 mM Tris–(pH 7.5)–1 mM EDTA–2% sodium dodecyl sulfate (occlusion bodies from 5×10^7 cells per ml of buffer) for 3 to 4 h at room temperature with swirling. After extractions and precipitations as described above, the DNA was pelleted by centrifugation and stored in the cold.

The distributions of DNA in the fraction solubilized in carbonate solution alone and in the partially insoluble matrix were 66% and 33%, respectively. The yield of DNA was in the range of 1.2 μ g per 10^6 cells.

Other methods. Restriction enzyme digestions were performed with the buffers recommended by the suppliers. After electrophoresis on agarose gels, the resolved fragments were immobilized on a membrane support (Gene-Screen; New England Nuclear Corp., Boston, Mass.) as previously described (44) except that two 15-min treatments of the gels with 0.25 N HCl were included before alkaline denaturation to depurinate the DNA into sizes of approximately 1 to 2 kilobases (kb) (51). Nick translations and Southern hybridizations were as described previously (24). Probes were nick translated to specific activities of 2×10^7 to 4×10^7 cpm Cerenkov per μ g. Hybridizations were performed at 70°C for 14 to 16 h with 400,000 cpm Cerenkov of each probe per ml of hybridization solution. DNA fragments were isolated from gel slices by electroelution (20), but no bovine serum albumin was used. Cloning of various restriction fragments of the viral genome was done as described previously (24) by using equimolar mixtures of linearized and dephosphorylated plasmid pUC9 (48) and purified restriction fragments isolated by electroelution. After the transfection of competent *Escherichia coli* cells (a strain derivative of *E. coli* K-12 RRI having the genotype *leu pro thi strA hsd r⁻ m⁻ lacZ Δ M15 F' lac^R pro⁺* [38]), transformants were selected on ampicillin-containing plates. Ampicillin-resistant colonies were further plated on plates containing 5-bromo-4-chloro-3-indolyl- β -D-galactoside and isopropyl- β -D-thiogalactoside (30), and colorless colonies were picked from those plates for further biochemical analysis. The clones of the *Eco*RI fragment of BmNPV DNA inserted in pUC9 in two orientations have been designated Bmp/pR5 and Bmp/pR8, whereas those of the *Pst*I fragment

FIG. 1. Southern hybridizations of BmNPV DNA restriction digests. BmNPV DNA (0.15 to 0.22 μg per digest) was digested with a variety of restriction enzymes, and the digests were resolved on 0.7% agarose gels in the presence of ethidium bromide (0.5 $\mu\text{g}/\text{ml}$). After electrophoresis the gels were photographed (middle portion of each panel), and the digests were subsequently Southern transferred to membrane sheets, hybridized to nick-translated clone pM5 *Hind*V containing part of the AcNPV polyhedrin gene, and autoradiographed (left portion of each panel). After autoradiography and without any prior melting of the hybridized AcNPV polyhedrin gene sequences, the same filters were probed with total nick-translated BmNPV DNA and autoradiographed (right portion of each panel). The order of the digests was as follows: (upper panel) 1, *Ava*I; 2, *Bal*I; 3, *Bam*HI; 4, *Cl*aI; 5, *Eco*RI; 6, *Eco*RV; 7, *Hind*III; 8, *Nde*I; and (lower panel) 1, *Nru*I; 2, *Pst*I; 3, *Pvu*I; 4, *Pvu*II; 5, *Sal*I; 6, *Sph*I. For size markers (M), a mixture of *Hind*III-digested and *Eco*RI-digested λ DNA was used on each ethidium bromide-stained gel. For size markers on the narrow strips of the ethidium bromide-stained gels, a mixture of *Hind*II-*Hind*III double digest of λ DNA and *Hae*III-digested pMB9 DNA was used. The numbers at the right side of each panel indicate the length in kilobases of some of the size markers described above. Asterisks on the bands of the ethidium bromide-stained gels indicate the fragments hybridizing to the AcNPV polyhedrin gene probe.

have been designated Bmp/pP3 and Bmp/pP14. ^{32}P end labeling of plasmid DNA preparations with T4 polynucleotide kinase after linearization and dephosphorylation has been described previously (24). Restriction enzyme mapping was carried out by the partial digestion method with single-end-labeled fragments of cloned viral DNA (43). Nucleotide sequencing was done by the chemical method of Maxam and

Gilbert (29), and chemical reactions were analyzed on 85-cm-long 6% polyacrylamide gels. ^{32}P labeling at the 3' termini of restriction fragments was performed by repair synthesis of 3' recessed ends by using the Klenow fragment of *E. coli* DNA polymerase in a 20- μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 5 mM MgCl_2 , 10 mM β -mercaptoethanol, 70 μM each of dATP, dGTP, and TTP,

FIG. 2. Cloned portions of BmNPV. The 9.6-kb *EcoRI* and 10-kb *PstI* fragments of BmNPV cloned in plasmid pUC9 are represented by the top line. Some of the enzymes relevant to the sequences of the polyhedrin gene are also presented. In the lower line an expanded version of the *SalI*-*BamHI* fragment containing the polyhedrin gene is presented together with the cleavage sites used for generating end-labeled fragments for nucleotide sequence determination in the polyhedrin gene region. The designations 5'U and 3'U symbolize the 5' and 3' untranslated regions of the polyhedrin mRNA, and the end of the mRNA has been tentatively placed 30 nucleotides downstream from the first AATAAA polyadenylation signal (see also the text).

100 μ Ci of α -[32 P]dCTP (ca. 3,000 Ci/mmol), 10 U of Klenow fragment, and 5 to 20 μ g of the DNA digest. Sites used for 5' or 3' end labeling are listed in order from left to right in the lower part of Fig. 2, and their locations are identified by reference to the numbered DNA sequence shown in Fig. 3 as follows: *SalI* at -192, *XbaI* at 147, *NdeI* at 595, and *HindIII* at 1177. In all cases sequencing was performed in both directions.

RESULTS

To identify the part of the genome of BmNPV that contains the polyhedrin gene, viral DNA was isolated and restricted with a large number of restriction enzymes with hexanucleotide recognition sites. After electrophoresis of the restricted DNA, the resolved fragments were transferred to a membrane filter and subsequently hybridized to a DNA clone containing a large portion of the coding sequences and some 3' noncoding sequences of the polyhedrin gene of AcNPV (*HindIII*-V; see above for details). Typical results from such hybridizations are shown in Fig. 1. Through the hybridizations of the BmNPV digests to the AcNPV probe (left hybridization panels of Fig. 1), the sizes of the restriction fragments that included the BmNPV polyhedrin gene and the flanking sequences that are homologous to the hybridization probe were determined. Thus, several restriction enzymes were identified that produced single hybridizing fragments whose lengths were within the size range that may be cloned relatively easily in *E. coli* and also allowed the prediction to be made that the entire polyhedrin gene may be contained within them.

To determine unequivocally the position of the polyhedrin gene-containing fragments when bands were closely spaced and to visualize the smaller restriction fragments of the BmNPV genome that might have escaped detection in the ethidium bromide-stained gels, the same immobilized digests were subsequently hybridized to radioactive BmNPV DNA. Through such hybridizations (right hybridization panels of Fig. 1), all restriction fragments of a size equal to or greater than 500 base pairs were detected. Based on this informa-

tion, we were able to deduce a length of 125 ± 4 kb for the genome of BmNPV.

The *EcoRI* and *PstI* fragments of BmNPV DNA hybridizing to the AcNPV polyhedrin gene probe (approximately 9.6 and 10 kb, respectively; Fig. 1) were cloned and mapped in detail. Part of this mapping information is shown in Fig. 2. The combined *EcoRI* and *PstI* cloned fragments were found to span a total length of 11.8 kb of the viral genome. On the basis of the complete restriction mapping information derived from these clones as well as from additional Southern hybridizations (data not shown), we were able to determine the approximate location of the polyhedrin gene in the middle of the 11.8 kb of the cloned viral DNA and to initiate its sequencing.

The complete nucleotide sequence of the polyhedrin gene with some of its 5' and 3' flanking sequences is shown in Fig. 3. Upon inspection of the 2,060-nucleotide-long sequenced part of the cloned DNA, an open reading frame was identified in one of the DNA strands that resulted in the translation of a 244-amino acid-long polypeptide that, with few exceptions (see below), was identified as the polyhedrin of BmNPV based on comparisons with the published sequence of the authentic protein (26, 39) and with the predicted amino acid sequence of the AcNPV polyhedrin (23). The exact localization of the polyhedrin gene and its transcriptional orientation in the cloned DNA, as deduced from the determination of the primary structure of the gene, are shown in the lower part of Fig. 2 along with the restriction sites used as the major starting points for the nucleotide sequence analysis.

The sequenced DNA was also scanned for the presence of other open reading frames. Two more putative open reading frames were detected at the beginning and at the end of the sequenced DNA. The first one, at the beginning of the sequence, is read in the same orientation as the polyhedrin gene, comprises 82 amino acids, and terminates at nucleotide -260. The second one, at the end of the sequence, is read in the opposite orientation (from the complementary strand). This open frame translates into 239 amino acids and terminates at nucleotide 768 of the sequence shown in Fig. 3. The consequences of the possible presence of one or two

-562 -552 -542 -532 -522 -512 -502
 CUGCCACTA TTAATGAAT TAAATATACC AATTTTAAAR AACGACAGCA AAGAARCAAT TGATGAAAG
 -492 -482 -472 -462 -452 -442 -432
 AATGCGTAGA AGGAAATAT AATGTCATGC ACATGCTGAA CAGCAAGATC AATATGCTC CGTGCAATCA
 -422 -412 -402 -392 -382 -372 -362
 AAAAATATTG GCGATTTTGA AAAGAACAAC TGCACGCGCG CGGTATGTAC AGGAAGAGGT TTATACATAA
 -352 -342 -332 -322 -312 -302 -292
 CTGTACATT GCAAAAGTGG TTTCGTGTATC CAATGTGAA AACCGATGT TAATCAAGGC TGTGACACAT
 -282 -272 -262 -252 -242 -232 -222
 TTTTACAAAT AGCAATCCAA GTGTGTGGGT GAATCATGC ATCTTTTAAAT CAATCCCAAA GATGTGTATA
 -212 -202 -192 -182 -172 -162 -152
 AACCAACAAA CTGCAAAARA ATGAAGACTG TCGACAGCTC CTGTTCGTTT GCTGGCACTC GCAAGGCTCT
 -142 -132 -122 -112 -102 -92 -82
 CAATCTATCT TGATATATT GAATAATAAA ACAATATATA ATGTCAAAAT TGTTTTTAT TAACGATACA
 -72 -62 -52 -42 -32 -22 -12
 AATGGAAATA ATAACTATCT CGCAATAAAA TAAGTATTTT ACTGTTCG TAAACGTTTT GTAATAAAAA
 -2
 AACCTATAA T
 15 30 45 60
 ATG CCG AAT TAT TCA TAC AAC CCC ACC ATC GGG CGT ACT TAC GTG TAC GAC AAT AAA TAT
 Met Pro Asn Tyr Ser Tyr Asn Pro Thr Ile Gly Arg Thr Tyr Val Tyr Asn Asn Lys Tyr
 75 90 105 120
 TAC AAA AAC TTG GGC GGT CTC ATC AAA GCC AAG CCG AAG CAC CTA ATC GAA CAT
 Tyr Lys Asn Leu Gly Gly Leu Ile Lys Asn Ala Lys Arg Lys Lys Lys Leu Ile Gly His
 135 150 165 180
 GAA AAA GAG GAG AAG CAA TGG BAT CTT CTA GAC AAC TAC ATG GTT GCC GAA GAT CCC TTT
 Glu Lys Glu Gly Lys Asn Gln Tyr Trp Asn Leu Asn Tyr Met Val Ala Gly Asn Val Lys Pro
 195 210 225 240
 TTA GGA CCG GGC AAA AAC CAA AAA CTT ACC CTT TTT AAA GAG GTT CCG AAT GTG AAA CCC
 Leu Gly Pro Gly Lys Asn Gln Lys Leu Thr Leu Phe Lys Glu Val Arg Asn Val Lys Pro
 255 270 285 300
 BAT ACC ATG AAG TTA ATC GTC AAG TGG ACG GGC AAA GAG TTT CTG CGT GAA ACT TGG ACC
 Asp Thr Met Lys Leu Ile Val Asn Trp Ser Gly Lys Glu Phe Leu Arg Glu Thr Trp Thr
 315 330 345 360
 CGT TTT GTT GAG GAC AGC TTC CCC ATT GTA AAC GAC CAA GAG GTG ATG GAC GTG TAC CTC
 Arg Phe Val Glu Asp Ser Phe Pro Ile Val Asn Asp Gln Glu Val Met Asp Val Tyr Leu
 375 390 405 420
 GTC GCC AAC CTC AAA CCC ACA CCG CCC AAC AGG TGC TAC AAG TTC CTC GCT CAA CAC GCT
 Val Ala Asn Leu Lys Pro Thr Arg Pro Asn Arg Cys Tyr Lys Phe Leu Ala Gln His Ala
 435 450 465 480
 CTT AGG TGG GAC GAA GAC TAC GTG CCC CAC GAA GTA ATC AGA ATT ATG GAG CCA TCC TAC
 Leu Arg Trp Asp Gly Lys Asn Tyr Val Pro His Glu Val Ile Arg Ile Met Glu Pro Ser Tyr
 495 510 525 540
 GTG GGC ATG AAC AAC GAA TAC AGA ATT AGT CTG GCT AAA AAG GGC GGC TGC CCA ATC
 Val Gly Met Asn Asn Glu Tyr Arg Ile Ser Leu Ala Lys Lys Gly Gly Gly Cys Pro Ile
 555 570 585 600
 ATG AAC ATC CAC ACG GAG TAC ACC AAC TCG TGT GAG TCG TTT GTG AAC CCG GTC ATA TGG
 Met Asn Ile His Ser Glu Tyr Thr Asn Ser Phe Glu Ser Phe Val Asn Arg Val Ile Trp
 615 630 645 660
 GAG AAC TTC TAC AAA CCC ATC GTT TAC ATC GGC ACA GAC TCT GGC GAA GAA GAG AAC ATC
 Glu Asn Phe Tyr Lys Pro Ile Val Tyr Ile Gly Thr Asp Ser Ala Glu Glu Glu Ile
 675 690 705 720
 CTA ATT GAG GTT TCT CTC GTT TCT AAA ATA AAG GAG TTT GCA CCA GAC GCG CCT CTG TTC
 Leu Ile Glu Val Ser Leu Val Phe Lys Ile Lys Glu Phe Ala Pro Asp Ala Pro Leu Phe
 735
 ACT GGT CCG GCA TAT TAA
 Thr Gly Pro Ala Tyr Ter
 748 758 768 778 788 798 808
 AACACTATAC ATTTGTTATTA GTACATTTAT TAAGCTTTAG ATTCTGTACG TTGTTGATTT ACAGACAATT
 818 828 838 848 858 868 878
 GTTGTACGTA TTTTAATAAC TCATTAAAT TATAATCTTT AGGGTGGTAT GTTAGAGCCA AATCAAAATG
 888 898 908 918 928 938 948
 ATTTTCAGCG TCTTTGTATC TGAATTTAAA TATTAATACC TCAATAGATT TGTAAATAG GTTTTGATTTG
 958 968 978 988 998 1008 1018
 GTTCAAAACA AGGGTGTGTT TTGCAAAACG ATGCTGTGAC TATCTAATGG ATTTTCGCTC AACACCAACAC
 1028 1038 1048 1058 1068 1078 1088
 GACTTCCCAA ATCTGTAGC AGCACTATG TTTGTGTGT ATTCGTTGT GTTTGTGTTT GTAATAAAGA
 1098 1108 1118 1128 1138 1148 1158
 TTCACGTGCT TTCAAAATAT TATGCGCTTT TGATTTGTT TCACTACTGT CGTTAGTGT CAATTTGACTC
 1168 1178 1188 1198 1208 1218 1228
 GACBTAAACA CGTTAATAAA AGCTTGGACA TATTTAAACAT CCGGCGCGTT AGCGCATTA TTGCGCGCT
 1238 1248 1258 1268 1278 1288 1298
 CGTCCCAAC CTGCGTGTTA GAGTGTCTT CCGAAGACGA TTTTCCATA CCGACACGAC GCCTATTAT
 1308 1318 1328 1338 1348 1358 1368
 TGTGTGACT AACACCTCCG CBATCAATTT TTTAGTTGTT GAGCTTTTCC GAATTTATTC TGATTGCGGA
 1378 1388 1398 1408 1418 1428 1438
 CGTTTTTGGG CGGGTTTCAA TCTACTGTG CCCBATTITA ATTCACAGAA TACBTAGAA AGCBATGGTG
 1448 1458 1468 1478 1488
 CAGCGCGTGG TAACATTTCAC ACCGCAAAAT CTACTAATGG TGGCTGTAAT G

FIG. 3. Nucleotide sequence of the polyhedrin gene and surrounding region. The first nucleotide of the initiator ATG codon has been designated number 1. For assignments of cap site, polyadenylation signals, and other consensus sequences in the promoter region, see the text.

structural genes in the immediate vicinity of the polyhedrin gene are discussed below.

DISCUSSION

A prerequisite to the successful genetic manipulation of a particular site in a genome is the detailed characterization of the target site. We are predicting that, in a manner similar to that shown to occur in AcNPV (8, 11, 27, 34, 40, 41), direct or indirect mutational alteration, inactivation, and even removal of the polyhedrin gene from the BmNPV genome should bear no genetic consequences in terms of viability and virulence of the nonoccluded form of the virus in

laboratory animals (silkmoths) and tissue culture cells. Therefore, as the first step toward introducing foreign gene sequences into the genome of BmNPV, we identified, isolated by molecular cloning, and characterized by sequence analysis the gene encoding the viral protein polyhedrin.

Genome size of BmNPV and polyhedrin gene cloning. In the course of our preliminary characterization of the genome of BmNPV by restriction analysis and hybridization aimed at the identification of restriction fragments containing the polyhedrin gene, we had the opportunity to firmly establish the size of the viral DNA. Over the past 30 years, various values have been published on the size of the circular double-stranded BmNPV genome, ranging from 3 to 180 kb (see reference 19 for a review). Through our restriction digestion analysis (Fig. 1 and other data not shown), we calculated a length of 125 ± 4 kb for the viral genome a size very similar to that established for the genome of AcNPV (32, 49). No restriction pattern polymorphisms (restriction fragments appearing in submolar quantities) were detected when BmNPV DNA that had been serially passed five

Pro	Asn	Tyr	Ser	Tyr	Asn	Pro	Thr	Ile	Gly	Arg	Thr	Tyr	Val	Tyr	Asp	Asn	Lys	Tyr	Tyr
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Lys	Asn	Leu	Gly	Gly	Leu	Ile	Lys	Asn	Ala	Lys	Arg	Lys	Lys	His	Leu	Ile	Glu	His	Glu
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Lys	Glu	Glu	Lys	Gln	Trp	Asp	Leu	Leu	Asn	Asn	Tyr	Met	Val	Ala	Glu	Asp	Pro	Phe	Leu
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Gly	Pro	Gly	Lys	Asn	Gln	Lys	Leu	Thr	Leu	Phe	Lys	Glu	Val	Arg	Asn	Val	Lys	Pro	Asp
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Thr	Met	Lys	Leu	Ile	Val	Asn	Trp	Ser	Gly	Lys	Glu	Phe	Leu	Arg	Glu	Thr	Trp	Thr	Arg
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Phe	Val	Glu	Asp	Ser	Phe	Pro	Ile	Val	Asn	Asp	Gln	Glu	Val	Met	Asp	Val	Tyr	Leu	Val
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ala	Asn	Leu	Lys	Pro	Thr	Arg	Pro	Asn	Arg	Cys	Tyr	Lys	Phe	Leu	Ala	Gln	His	Ala	Leu
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Arg	Trp	Asp	Glu	Asp	Tyr	Val	Pro	His	Glu	Val	Ile	Arg	Ile	Met	Glu	Pro	Ser	Tyr	Val
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Gly	Met	Asn	Asn	Glu	Tyr	Arg	Ile	Ser	Leu	Ala	Lys	Lys	Gly	Gly	Gly	Cys	Pro	Ile	Met
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Asn	Ile	His	Ser	Glu	Tyr	Thr	Asn	Ser	Phe	Glu	Ser	Phe	Val	Asn	Arg	Val	Ile	Trp	Glu
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Asn	Phe	Tyr	Lys	Pro	Ile	Val	Tyr	Ile	Gly	Thr	Asp	Ser	Ala	Glu	Glu	Glu	Glu	Ile	Leu
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ile	Glu	Val	Ser	Leu	Val	Phe	Lys	Ile	Lys	Glu	Phe	Ala	Pro	Asp	Ala	Pro	Leu	Phe	Thr
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Gly	Pro	Ala	Tyr	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

FIG. 4. Comparison of two BmNPV polyhedrin sequences. The polyhedrin sequence predicted from the gene sequence (upper line) is compared to that previously reported for the purified protein (lower line; [26, 39]). Dots in the lower line indicate the same amino acids as in the top one and blocks of identical sequences are boxed. Notice the change in reading frame occurring at amino acid 114 and its restoration at residue 145.

Bm	Pro	Asn	Tyr	Ser	Tyr	Asn	Pro	Thr	Ile	Gly	Arg	Thr	Tyr	Val	Tyr	Asp	Asn	Lys	Tyr	Tyr
Ac	.	Asp	.	.	.	Arg	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bm	Lys	Asn	Leu	Gly	Gly	Leu	Ile	Lys	Asn	Ala	Lys	Arg	Lys	Lys	His	Leu	Ile	Glu	His	Glu
Ac	.	.	.	.	Ala	Val	.	.	.	.	.	.	.	.	.	Phe	Ala	.	.	.
Bm	Lys	Glu	Glu	Lys	Gln	Trp	Asp	Leu	Leu	Asp	Asn	Tyr	Met	Val	Ala	Glu	Asp	Pro	Phe	Leu
Ac	Ile	.	.	Ala	Thr	Leu	.	Pro	.	.	.	.	Leu	.	.	.	.	.	.	.
Bm	Gly	Pro	Gly	Lys	Asn	Gln	Lys	Leu	Thr	Leu	Phe	Lys	Glu	Val	Arg	Asn	Val	Lys	Pro	Asp
Ac	.	.	.	.	.	.	.	.	.	.	.	.	.	Ile	.	.	.	.	.	.
Bm	Thr	Met	Lys	Leu	Ile	Val	Asn	Trp	Ser	Gly	Lys	Glu	Phe	Leu	Arg	Glu	Thr	Trp	Thr	Arg
Ac	.	.	.	.	Val	.	Gly	.	Lys	.	.	.	.	Tyr	.	.	.	.	.	.
Bm	Phe	Val	Glu	Asp	Ser	Phe	Pro	Ile	Val	Asn	Asp	Gln	Glu	Val	Met	Asp	Val	Tyr	Leu	Val
Ac	.	Met	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Phe	.	.
Bm	Ala	Asn	Leu	Lys	Pro	Thr	Arg	Pro	Asn	Arg	Cys	Tyr	Lys	Phe	Leu	Ala	Gln	His	Ala	Leu
Ac	Val	.	Met	Arg	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bm	Arg	Trp	Asp	Glu	Asp	Tyr	Val	Pro	His	Glu	Val	Ile	Arg	Ile	Met	Glu	Pro	Ser	Tyr	Val
Ac	.	Cys	.	.	.	.	.	.	.	Asp	.	.	.	.	Val	.	.	.	.	.
Bm	Gly	Met	Asn	Asn	Glu	Tyr	Arg	Ile	Ser	Leu	Ala	Lys	Lys	Gly	Gly	Gly	Cys	Pro	Ile	Met
Ac	.	Ser	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bm	Asn	Ile	His	Ser	Glu	Tyr	Thr	Asn	Ser	Phe	Glu	Ser	Phe	Val	Asn	Arg	Val	Ile	Trp	Glu
Ac	.	Leu	.	.	.	.	.	.	.	.	Gln	.	Ile	Asp	.	.	.	.	.	.
Bm	Asn	Phe	Tyr	Lys	Pro	Ile	Val	Tyr	Ile	Gly	Thr	Asp	Ser	Ala	Glu	Glu	Glu	Glu	Ile	Leu
Ac	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Bm	Ile	Glu	Val	Ser	Leu	Val	Phe	Lys	Ile	Lys	Glu	Phe	Ala	Pro	Asp	Ala	Pro	Leu	Phe	Thr
Ac	.	.	.	.	.	.	.	.	Val	.	.	.	.	.	.	.	.	.	.	.
Bm	Gly	Pro	Ala	Tyr	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ac	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

FIG. 5. Comparison of the polyhedrin sequences of BmNPV and AcNPV. The BmNPV polyhedrin sequence predicted from the sequence of the gene is compared with that predicted from the corresponding gene of AcNPV (23). Dots in the sequence of AcNPV indicate the same amino acids as those in the BmNPV sequence, and blocks of identical sequences are boxed.

times through Bm-5 cells and restricted with enzymes that produce reasonably small numbers of well-separated DNA fragments was analyzed on agarose gels.

After hybridizations of Southern transferred restriction digests of BmNPV DNA to the probe (937 base pairs) containing part of the AcNPV polyhedrin gene, single restriction fragments were found to hybridize to the probe in a number of different digests (Fig. 1). These fragments presumably contain the corresponding gene of BmNPV. Two of them, a 9.6-kb *EcoRI* and a 10-kb *PstI* fragment (Fig. 2), were cloned, and the approximate locations of the sequences hybridizing to the AcNPV probe were determined. After nucleotide sequence determination (Fig. 3), the hybridizing portion of BmNPV DNA was identified as the polyhedrin gene. This identification was based on the amino acid sequence of the polypeptide that resulted from the conceptual translation of the corresponding RNA sequence after comparisons with the published sequence of the authentic BmNPV polyhedrin (26, 39) and with that derived from the polyhedrin gene sequences of AcNPV (23).

Nucleotide sequences of the polyhedrin gene. The sequences determined for the polyhedrin gene of BmNPV and its vicinity (Fig. 3) include 571 nucleotides of flanking and mRNA noncoding sequences to the 5' end of the ATG initiator codon (nucleotide 1 in Fig. 3), a coding sequence of 738 nucleotides (including initiation and termination codons), and 751 nucleotides of mRNA noncoding and flanking regions downstream from the 3' end of the TAA terminator (nucleotide 738 in Fig. 3). Since no information is yet available on the BmNPV polyhedrin mRNA sequences, the cap addition site has been inferred at position -57 by homology to the cap site reported recently for the polyhedrin gene of AcNPV (23, 42). DNA sequences similar to the

consensus TATA and CCAAT boxes that have been shown to represent important elements of eucaryotic gene promoters (6) were found in the 5' flanking region of the BmNPV polyhedrin gene 28 and 63 nucleotides upstream from the putative cap site, respectively (Fig. 3, positions -85 and -120). As with the AcNPV polyhedrin gene (23), the observed TATA box deviates from the established canonical sequence (TACAAA for TATAAA). Deviations of this type, however, are not unusual, particularly in genes of viral origin (2). An extra set of TATA and CAAT sequences was also observed at positions -116 and -151, respectively (Fig. 3). Of course, the contributions of any of the signals mentioned above to polyhedrin gene promoter function are speculative and may only be deduced through transcriptional studies.

The coding portion of the polyhedrin gene of BmNPV is not interrupted by intervening sequences. This is also true with the polyhedrin genes of AcNPV (23, 42) and of the NPV of *Orgyia pseudotsugata* (OpNPV [36]), whose N-terminal polyhedrin sequence was found to be very similar to that of BmNPV polyhedrin (37). In terms of codon usage, the TAC codon for tyrosine was found to be used at a frequency of 13/16, whereas the AAC codon for asparagine occurred at a frequency of 15/18 (the frequencies of the corresponding codons in the AcNPV polyhedrin gene are 12/15 and 12/14, respectively [23]). Except for these two cases, no other strong codon usage preferences were noted.

The polyhedrin mRNAs of AcNPV and OpNPV have been shown to be polyadenylated (36, 50), and most likely the same is true for the polyhedrin mRNA of BmNPV. Because the published sequence of the AcNPV polyhedrin gene does not extend to the point of the mRNA polyadenylation site, this site cannot be inferred on the gene for

FIG. 6. Sequence comparison between the BmNPV and AcNPV polyhedrin genes and their surrounding sequences. (A) The 5' flanking sequences of the two genes which have diversified considerably are compared after gaps were introduced to maximize homologies. Blocks of two or more identical nucleotides are boxed, and dots indicate nucleotide differences. (B) The comparison begins at nucleotide -76 of the sequence of the BmNPV gene and continues to the last known nucleotide of the AcNPV polyhedrin gene sequence (23). Except for a blank introduced after nucleotide -1 of the BmNPV polyhedrin gene sequence to accommodate an extra A residue present in the gene of AcNPV polyhedrin, no sequence rearrangements were necessary for maximum homology. Nucleotide differences are indicated by dots.

BmNPV polyhedrin by homology. However, the sequence AATAAA, shown to be present in the 3' noncoding region of almost all polyadenylated eucaryotic mRNAs approximately 25 to 30 nucleotides upstream from the site of polyadenylate addition (35), occurs at position 1081 of our gene (Fig. 3). It should be noted that a second AATAAA sequence is present approximately 90 nucleotides downstream from the first one (position 1174). An unambiguous answer as to the location of the polyadenylation site may only be derived by S1 nuclease protection experiments (5). Protection experiments of this kind to determine the length of AcNPV polyhedrin mRNA have indicated a cytoplasmic mRNA size of approximately

1,200 nucleotides, excluding the polyadenylate tail (42). Assuming a similar mRNA size for BmNPV polyhedrin, that value would place the end of the BmNPV mRNA at around 60 nucleotides downstream from the first polyadenylation signal.

Amino acid sequence of polyhedrin. When the amino acid sequence predicted from the gene primary structure (Fig. 3) was compared to that of the authentic protein (26, 39), a small number of discrepancies were noted which prevented the absolute matching of the two sequences. The differences are shown in Fig. 4 and are summarized as follows (the numbers in Fig. 4 and those mentioned below correspond to

our sequence and are exclusive of the initiating methionine residue): (i) amino acid substitutions in dipeptides at positions 39-40 (His-Glu for Glu-His) and 213-214 (Ser-Ala for Ala-Ser); (ii) an amino acid substitution at position 218 (Glu for Gln); (iii) the presence of a single valine residue at position 114 instead of two, resulting in a shift of the amino acid reading frame by one residue; (iv) the presence of an extra amino acid residue at position 144 (Glu) missing from the published polyhedrin sequence (26, 39), which results in the restoration of the colinear amino acid reading frame; and (v) amino acid substitutions at positions 142 (Trp for Gln) and 143 (Asp for Asn) within the region in which the reading frame shift has occurred.

Of the observed differences, those described in (i) above are positional. Of the remaining amino acid substitutions, those at positions 143 and 218 may be explained by the accumulation of single point-mutational events. The amino acid substitution at position 142, however, requires the accumulation of at least two point mutations on the particular codon. More complex mechanisms should, of course, be used to explain the deletion and insertion of the codons for the amino acid residues at positions 114 and 144. These differences are obviously more than what one would expect from simple variants of the same virus. Since we have confirmed our nucleotide sequence analysis and since the sequenced moiety of the cloned fragment was the only portion of the BmNPV genome that hybridized to the AcNPV polyhedrin gene probe, we feel that a reexamination of the polyhedrin protein sequencing data may be appropriate.

Polyhedrins of BmNPV and AcNPV. The sequence of BmNPV polyhedrin predicted by the primary structure of the gene was also compared with that predicted from the corresponding sequences of AcNPV (Fig. 5). No rearrangements were required to align the two amino acid sequences, which are clearly highly homologous. The degree of divergence between the two sequences is 13.9% (34 amino acid substitutions in a total of 244 residues). The collective result of the amino acid substitutions appears to be neutral, since the net change in charge effected by them amounts to one and the overall degree of hydrophobicity remains unchanged. In addition, computerized predictions of the secondary structures of BmNPV and AcNPV polyhedrins (15, 28) reveal that although some regional differences may occur, the overall distributions of α -helical and β -pleated sheet structures within the two molecules were identical (data not shown). In this respect it is also worth mentioning that BmNPV polyhedrin has 11 arginine residues, all of which are also present in the polyhedrin of AcNPV, which has a total of 13 arginine residues. Of the codons for the 11 arginine residues shared by the two sequences, 5 have been altered by double point mutations resulting in codons again specifying arginine residues. Thus, it appears that the changes in the polyhedrin molecule during evolution occurred under considerable selective pressure. This is further corroborated by the results obtained when the N-terminal amino acid sequences of the polyhedrins from two types of OpNPV were compared with those of the polyhedrin of BmNPV (37). A sequence identity of 88.9% was found in the comparisons involving the first 36 residues of a multicapsid OpNPV, whereas the sequence homology for 34 residues of a unicapsid OpNPV was 85.3% (the corresponding values for the comparison between BmNPV and AcNPV polyhedrins were 88.2 and 86.8% for the first 34 and 36 residues, respectively [Fig. 5]).

Sequence homology between the polyhedrin genes of BmNPV

and AcNPV. The nucleotide sequences determined for the polyhedrin gene of BmNPV were compared with those of the polyhedrin gene of AcNPV. This comparison indicated that major portions of the two sequences may be completely correlated (Fig. 6).

For the protein-encoding regions of the genes (shown to have diverged by 13.9% at the amino acid level [Fig. 5]), a 22.2% divergence was determined (164 nucleotide substitutions in a total of 738). Of the observed changes, 110 (two-thirds) represented silent substitutions resulting in the appearance of synonymous codons, whereas 54 (one-third) were replacement substitutions leading to amino acid changes. From these results it is apparent that a large number of silent substitutions have been fixed in the coding portion of the genes during evolution and that the two nucleotide sequences have diverged somewhat more than the corresponding amino acid sequences. This indicates, in turn, that the protein itself has been under more stringent selective pressure than the corresponding DNA sequence.

In contrast to the divergence found in the coding regions of the two polyhedrin genes, the nucleotide sequences of the 5' noncoding regions and those of a small portion of the 5' flanking regions show a remarkable conservation. In fact, except for an extra nucleotide that is present in the 58-nucleotide-long 5' noncoding region of the AcNPV gene just before the ATG initiation codon (Fig. 6), the two noncoding sequences are identical. Even this difference may not be significant, since in another isolate of AcNPV the extra nucleotide was not observed (E. B. Carstens, unpublished data). This remarkable length and nucleotide sequence conservation in the 5' noncoding regions strongly suggests that capping of the two mRNAs may occur at the same position (-57 in the sequences shown in Fig. 6). The sequence homology continues farther upstream for another 20 nucleotides (19 identical residues out of 20) and stops at nucleotide -76 just before the beginning of the presumed TATA boxes.

A similarly striking homology was also observed in the 3' noncoding regions of the genes. Only 5 nucleotide substitutions were observed in the 207 nucleotides of the sequences that were compared (2.4% sequence divergence). Unfortunately, because more sequence information on the AcNPV polyhedrin gene was not available, we could not extend the comparison to cover the entire 3' noncoding region and flanking gene sequences. Taking into consideration, however, that mutations probably occurred uniformly throughout the two genomes (including the polyhedrin genes), it is obvious that, in contrast to the point mutations that have been fixed in the coding regions, only a very small number of substitutions were allowed to be fixed in the 5' and 3' noncoding regions and in the immediate 5' flanking regions of the genes during evolution. This of course should be indicative of the degree of selective pressure which the different portions of the genes are under because of the functions they have to accomplish.

No significant homology can be recognized at first glance in the 5' flanking gene sequences upstream from nucleotide -76 (Fig. 6, upper panel). With the aid of a dot matrix program, we scanned the two sequences for the presence of homologies that may occur without respect to position. As a result of this search we were able to construct different versions of sequence alignments. One of these alignments is presented in Fig. 6 (upper panel). Although the overall sequence identity is only 38.8%, several blocks of common sequences may be discerned (boxes in Fig. 6). Except for the large gap of nine nucleotides introduced at the 3' end of the BmNPV sequence, only three single nucleotide gaps have

been introduced in the sequence of AcNPV to allow for the alignment. We would like to emphasize not the nucleotide sequences as such but the fact that the distances among all of the observed homology blocks are the same (± 1 nucleotide) in both flanking sequences. At this point we can only speculate on the significance of these blocks. Transcriptional experiments would be needed to determine whether they participate in the regulation of polyhedrin gene activity. Despite our uncertainties, sequences of this type should be considered, particularly when molecular engineering in the vicinity of the polyhedrin gene is to be undertaken.

Is the polyhedrin gene immediately flanked by other viral genes? Since we may need to insert a fragment of foreign genetic material such as *B. mori* genomic DNA in the vicinity of the polyhedrin gene while preserving the ability of the virus to function properly (including the formation of polyhedra), it is important not only to know the exact limits of the polyhedrin gene but also what lies next to it. In this respect we have noticed the two open reading frames at the two ends of the sequenced portion of the cloned DNA. More extensive studies are needed to determine whether the observed open reading frames are parts of adjacent structural genes. Such studies are now in progress.

ACKNOWLEDGMENTS

We thank J. L. Vaughn and R. Stone of the Insect Pathology Laboratory for the gifts of Bm-5 cells and initial inocula of BmNPV as well as for their advice for the establishment of the cells in our laboratory, E. B. Carstens for supplying us with the cloned HindIII-V fragment of the AcNPV genome and for communicating to us his sequences of the AcNPV polyhedrin gene before publication; G. Chaconas of the Cancer Research Laboratory, University of Western Ontario, for making available to us the computer program on protein secondary structure predictions, J. C. States of the Department of Medical Biochemistry, University of Calgary, for his gift of pUC9 and host cells, D. McKay of the Department of Medical Biochemistry, University of Calgary, for providing us with some of the computer software used for the analysis of our nucleic acid sequences, B. Pinder and R. Haselden for photography, Anne Vipond for technical assistance, and Susan Carlson for secretarial assistance.

This work was supported by grants from the Cancer Grants Program of the Alberta Heritage Savings Trust Fund (Alberta Cancer Board) and the Medical Research Council of Canada (to K. Iatrou) and by a postdoctoral fellowship from the Alberta Heritage Foundation for Medical Research (to H.W.).

LITERATURE CITED

- Adang, M. J., and L. K. Miller. 1982. Molecular cloning of DNA complementary to mRNA of the baculovirus *Autographa californica* nuclear polyhedrosis virus: location and gene products of RNA transcripts found late in infection. *J. Virol.* **44**:782-793.
- Baker, C. C., J. Herisse, G. Courtois, F. Galibert, and E. Ziff. 1979. Messenger RNA for the Ad2 DNA binding protein: DNA sequences encoding the first leader and heterogeneity at the mRNA 5' end. *Cell* **18**:569-580.
- Bergold, G. H. 1947. Die Isolierung des Polyeder-Virus und die Natur der Polyeder. *Z. Naturforsch.* **26**:122-143.
- Bergold, G. H. 1953. Insect viruses. *Adv. Virus Res.* **1**:91-139.
- Berk, A. J., and P. A. Sharp. 1978. Spliced early mRNA of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* **75**:1274-1278.
- Breathnach, R., and P. Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**:349-383.
- Brown, M., P. Faulkner, M. A. Cohran, and K. L. Chung. 1980. Characterization of two morphology mutants of *Autographa californica* nuclear polyhedrosis virus with large cuboidal inclusion bodies. *J. Gen. Virol.* **50**:309-316.
- Carstens, E. B. 1982. Mapping the mutation site of an *Autographa californica* nuclear polyhedrosis virus polyhedron morphology mutant. *J. Virol.* **43**:809-818.
- Couch, J. A., S. M. Martin, G. Tompkins, and J. Kinney. 1984. A simple system for the preliminary evaluation of infectivity and pathogenesis of insect virus in a nontarget estuarine shrimp. *J. Invertebr. Pathol.* **43**:351-357.
- Dulbecco, R., and M. Vogt. 1954. Plaque formation and isolation of pure lines with polyomyelitis viruses. *J. Exp. Med.* **99**:167-182.
- Duncan, R., K. L. Chung, and P. Faulkner. 1983. Analysis of a mutant of *Autographa californica* nuclear polyhedrosis virus with a defect in the morphogenesis of the occlusion body macromolecular lattice. *J. Gen. Virol.* **64**:1531-1542.
- Duncan, R., and P. Faulkner. 1982. Bromodeoxyuridine-induced mutants of *Autographa californica* nuclear polyhedrosis virus defective in occlusion body formation. *J. Gen. Virol.* **62**:369-373.
- Faulkner, P. 1981. Baculovirus, p. 3-37. In E. W. Davidson (ed.), *Pathogenesis of Invertebrate Microbial Diseases*. Allanheld, Osmun & Co., Totowa, N.J.
- Fraser, M. J., and W. F. Hink. 1982. The isolation and characterization of the MP and FP plaque variants of *Galleria mellonella* nuclear polyhedrosis virus. *Virology* **117**:366-378.
- Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* **120**:97-120.
- Goldsmith, M. R., and F. C. Kafatos. 1984. Developmentally regulated genes in the silkworms. *Annu. Rev. Genet.* **18**:443-487.
- Grace, T. D. C. 1967. Establishment of a line of cells from the silkworm *Bombyx mori*. *Nature (London)* **216**:613.
- Groner, A., R. R. Granados, and J. P. Burand. 1984. Interaction of *Autographa californica* nuclear polyhedrosis virus with two nonpermissive cell lines. *Intervirology* **21**:203-209.
- Harrap, K. A., and C. C. Payne. 1979. The structural properties and identification of insect viruses. *Adv. Virus Res.* **25**:273-355.
- Heckman, J. E., and U. L. Rajbhandary. 1979. Organization of tRNA and rRNA genes in *N. crassa* mitochondria: intervening sequences in the large rRNA gene and strand distribution of the RNA genes. *Cell* **17**:583-595.
- Himeno, H., F. Sakai, K. Onodera, H. Nakai, T. Fukada, and Y. Kawade. 1967. Formation of nuclear polyhedral bodies and nuclear polyhedrosis virus of silkworm in mammalian cells infected with viral DNA. *Virology* **33**:507-512.
- Hink, W. F., and E. Strauss. 1976. Replication and passage of alfalfa looper nuclear polyhedrosis virus plaque variants in cloned cell cultures and larval stages of four host species. *J. Invertebr. Pathol.* **27**:49-55.
- Hoof van Iddekinge, B. J. L., G. E. Smith, and M. D. Summers. 1983. Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus. *Virology* **131**:561-565.
- Iatrou, K., and S. G. Tsitolou. 1983. Coordinately expressed chorion genes of *Bombyx mori*: is developmental specificity determined by secondary structure recognition? *EMBO J.* **2**:1431-1440.
- Ignoffo, C. M. 1973. Effects of entomopathogens on vertebrates. *Ann. N.Y. Acad. Sci.* **217**:141-172.
- Kozlov, E. A., T. L. Levitina, N. M. Gusak, and S. B. Serebryani. 1981. Comparison of the amino acid sequence of inclusion body proteins of nuclear polyhedrosis viruses *Bombyx mori*, *Porthetria dispar* and *Galleria mellonella*. *Bioorg. Khim.* **7**:1008-1015.
- Lee, H. H., and L. K. Miller. 1979. Isolation, complementation, and initial characterization of temperature-sensitive mutants of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **31**:240-252.
- Lifson, S., and C. Sander. 1979. Antiparallel and parallel β -strands differ in amino acid residue preferences. *Nature (London)* **282**:109-111.
- Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74**:560-564.
- Messing, J., B. Gronenborn, B. Muller-Hill, and P. H. Hofschneider. 1977. Filamentous coliphage M13 as a cloning vehicle.

- cle: insertion of a *Hind*II fragment of the *lac* regulatory region in M13 replicative form *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **74**:3642–3646.
31. Miller, D. W., and L. K. Miller. 1982. A virus mutant with an insertion of a *copia*-like transposable element. Nature (London) **299**:562–564.
 32. Miller, L. K., and K. P. Dawes. 1979. Physical map of the DNA genome of *Autographa californica* nuclear polyhedrosis virus. J. Virol. **29**:1044–1055.
 33. Miller, L. K., A. J. Ling, and L. A. Bulla, Jr. 1983. Bacterial, viral and fungal insecticides. Science **219**:715–721.
 34. Pennock, G. D., C. Shoemaker, and L. K. Miller. 1984. Strong and regulated expression of *Escherichia coli* β -galactosidase in insect cells with a baculovirus vector. Mol. Cell. Biol. **4**:399–406.
 35. Proudfoot, N. J., and G. G. Brownlee. 1976. 3' non-coding sequences in eukaryotic messenger RNA. Nature (London) **263**: 211–214.
 36. Rohrmann, G. F., D. J. Leisy, K.-C. Chow, G. D. Pearson, and G. S. Beaudreau. 1982. Identification, cloning and R-loop mapping of the polyhedrin gene from the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*. Virology **121**:51–60.
 37. Rohrmann, G. F., M. N. Pearson, T. J. Bailey, R. R. Becker, and G. S. Beaudreau. 1981. N-terminal polyhedrin sequences and occluded baculovirus evolution. J. Mol. Evol. **17**:329–333.
 38. Ruther, U. 1982. pUR250 allows rapid chemical sequencing of both DNA strands of its inserts. Nucleic Acids Res. **10**: 5765–5772.
 39. Serebryani, S. B., T. L. Levitina, M. L. Kautzman, Y. L. Radavski, N. M. Gusak, M. N. Ovander, N. V. Sucharenko, and E. A. Kozlov. 1977. The primary structure of the polyhedral protein of nuclear polyhedrosis virus (NPV) of *Bombyx mori*. J. Invertebr. Pathol. **30**:442–443.
 40. Smith, G. E., M. J. Fraser, and M. D. Summers. 1983. Molecular engineering of the *Autographa californica* nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene. J. Virol. **46**:584–593.
 41. Smith, G. E., M. D. Summers, and M. J. Fraser. 1983. Production of human beta interferon in insect cells infected with a baculovirus expression vector. Mol. Cell. Biol. **3**:2156–2165.
 42. Smith, G. E., J. M. Vlak, and M. D. Summers. 1983. Physical analysis of *Autographa californica* nuclear polyhedrosis virus transcripts for polyhedrin and 10,000-molecular-weight protein. J. Virol. **45**:215–225.
 43. Smith, H. O., and M. L. Birnstiel. 1976. A simple method for DNA restriction site mapping. Nucleic Acids Res. **3**:2387–2398.
 44. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
 45. Tinsley, T. W. 1978. Use of insect pathogenic viruses as pesticidal agents, p. 199–210. In M. Pollard (ed.), Perspectives in Virology vol. 10. Raven Press, New York.
 46. Tjia, S. T., G. M. Altenschildesche, and W. Doerfler. 1983. *Autographa californica* nuclear polyhedrosis virus (AcNPV) DNA does not persist in mass cultures of mammalian cells. Virology **125**:107–117.
 47. Trager, W. 1935. Cultivation of the virus of grasserie in silkworm tissue cultures. J. Exp. Med. **61**:501–517.
 48. 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.
 49. Vlak, J. M., and K. G. Odink. 1979. Characterization of *Autographa californica* nuclear polyhedrosis virus deoxyribonucleic acid. J. Gen. Virol. **44**:333–347.
 50. Vlak, J. M., G. E. Smith, and M. D. Summers. 1981. Hybridization selection and *in vitro* translation of *Autographa californica* nuclear polyhedrosis virus mRNA. J. Virol. **40**:762–771.
 51. Wahl, G. M., M. Stern, and G. R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. U.S.A. **76**:3683–3687.
 52. Weiss, S. A., G. C. Smith, S. S. Kalter, and J. L. Vaughn. 1981. Improved method for the production of insect cell cultures in large volume. In Vitro **17**:495–502.
 53. Wood, H. A. 1980. Isolation and replication of an occlusion body-deficient mutant of the *Autographa californica* nuclear polyhedrosis virus. Virology **105**:338–344.