MOLECULAR BIOLOGICAL STUDIES OF THE GENOME OF NUCLEAR POLYHEDROSIS VIRUSES

核多角体病ウイルスの遺伝子構造に関する研究

KEI MAJIMA

1992

Molecular Biological Studies of the Genome of Nuclear Polyhedrosis Viruses

By

Kei Majima

B.A., Tottori University, Tottori, 1986
M.A., Tottori University, Tottori, 1988

DISSERTATION

DOCTOR OF PHILOSOPHY IN AGRICULTURE

in the

GRADUATE DIVISION

of the

FACULTY OF AGRICULTURE
TOTTORI UNIVERSITY, TOTTORI

DEDICATION

To my parents for their enduring love and support

Yoshiko Majima and Mototoshi Majima

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to Dr. Ryuzo Kobara, my major professor for his guidance during my work at Tottori University. My special thanks go to Dr. Hidenori Kai, Dr. Yasuo Maeta, and Dr. Masaaki Azuma for their advice and critical reviews for this dissertation.

I also wish to thank Dr. Susumu Maeda, Mr. Shizuo G. Kamita, and Dr. Atsushi Nakamura for their encouragement, guidance, and critical reading of this dissertation. I thank the students and staff of Dr. Kobara's lab for their friendship and help.

Finally, I wish to thank all my family who encouraged and supported me to complete my studies at Tottori University.

TABLE OF CONTENTS

I. Abstract

Bombyx mori nuclear polyhedrosis virus (BmNPV) and Spodoptera litura NPV (SlNPV) were analyzed using various molecular biological techniques to examine the structure and functions of baculoviral genomes. Gene libraries covering the entire genomes of the two viruses were initially constructed in plasmid vectors. Complete physical maps of the BmNPV and SlNPV genomes for several restriction endonucleases were constructed by hybridization and double digestion experiments.

In the BmNPV genome five major regions of homologously repeated (hr) sequences were found. The BmNPV hr sequences had very high conservation of sequence, location, and orientation compared to hr sequences previously found in Autographa californica NPV (AcNPV), however, an inversion and rearrangement of some sequences were observed. Excluding the hr sequences, the BmNPV genome was made up of unique sequences, which consisted of open reading frames (ORFs) and their flanking sequences. Most ORFs identified between 86.3 and 99.6 map units possessed late and/or early gene promoter motifs and poly $(A)^+$ signals, indicating that they were actively translated. Furthermore, the coding sequences and location of four immediate-early genes and one late gene (p74) were found to be well conserved between BmNPV and AcNPV.

Four different polyhedrin genes of BmNPV and SlNPV were isolated, sequenced, and characterized. An amino acid sequence substitution was found to be responsible for a change in the inclusion body shape from polyhedral to cuboidal. Sequence analysis of the polyhedrin genes of two SlNPV isolates, OT2 and CC5, showed that OT2 is an AcNPV variant and CC5 belongs to a previously unknown S. littoralis NPV group, which is related to Orygia pseudotsugata NPV. Based on the effects of the 5′ non-translated region of the polyhedrin gene, new transfer vectors with high expression characteristics and with multiple cloning sites were constructed for foreign gene expression.

The polyhedrin gene was used to study the mechanisms of protein localization in the nucleus. A specific consensus nuclear localization signal found in other eukaryotic cells was found in the BmNPV polyhedrin gene. When this region was deleted by treatment with Bal31 exonuclease and examined using a marker rescue technique, the resulting recombinant viruses expressed mutated polyhedrin which localized in the cell cytoplasm. These results showed that insect cells use the same protein trafficking mechanisms generally found in other eukaryotic cells. It was also found that nearly the entire amino acid sequence of polyhedrin is necessary for normal crystallization to occur. The deletion experiments showed that 1) the shape of polyhedra is controlled by C-terminal amino acids and 2) amino acid information controls polyhedral size.

II. Introduction.

Insect-pathogenic viruses are generally classified into seven families. Most insect-pathogenic viruses fall into the family, Baculoviridae, which consist of three subfamilies, nuclear polyhedrosis viruses (NPV), granulosis viruses (GV), and non-occluded viruses. Viruses in the family Baculoviridae are characterized by circular double-stranded DNA genomes 80-160 kilo bases (kb) in length and rod-shaped enveloped virions (Matthews, 1982).

NPVs are found in several orders of insects, mainly lepidopterans, and have the unique characteristic of producing proteinacious nuclear occlusion bodies, which can each incorporate a large number of progeny viral particles. GVs produce proteinacious inclusions, called granules, which can each incorporate one progeny viral particles. Other insect virus families, including cytoplasmic polyhedrosis virus (CPV) and insect poxvirus, are also able to produce large proteinacious inclusion bodies which contain many progeny viral particles. The production of viral inclusion bodies is a unique characteristic of insect-pathogenic viruses. These inclusion bodies play an important role for viral transmission in the field by protecting embedded viral particles from UV light, chemicals, etc. in the soil for prolonged periods, presumably more than several years.

Viral epidemics in the field are a major factor in the control of insect populations (Granados and Federici, 1986).

For this reason they have been considered for use as a natural insecticide. For example, Oryctes rhinoceros baculovirus has been heavily used for pest control of coconut trees in the South Pacific (Redford, 1986), and Anticarsia gemmatalis NPV has been used for pest control of soybean in Brazil (Johnson and Maruniak, 1989). These applications have been primarily based on the efficacy of baculoviruses for pest control (Granados and Federici, 1986) and safety factors (Summers et al., 1975). Spodoptera litura NPV (SlNPV) is also considered as an efficient controlling agent for S. litura (Okada, 1977), which is a major pest of vegetable crops in Southern and Western Japan, however, very few genetic studies have been performed on this virus (see Maeda et al., 1990).

Bombyx mori NPV (BmNPV) is a major disease of the silkworm, B. mori, in sericulture. Protection of reared silkworm colonies from viral infection is crucial for obtaining high quantities of high quality cocoon for silk production. Although many etiological and histopathological studies have been conducted, little has been reported on the genetics of BmNPV at the molecular level (see review of Horie & Watanabe, 1980). NPVs are generally characterized as either SNPV or MNPV depending on the number of nucleocapsids within a viral envelope embedded in the polyhedra. Although BmNPV is listed as representative of SNPV (Matthews, 1982), BmNPV also produces a MNPV form. In an established cell line, it has been reported that only the

SNPV form is observed (Inoue & Mitsuhashi, 1984). In larvae, most progeny viral particles (more than 90%) are the S- form and the ratio of S- and M- forms depends on the infected organ (Watanabe, 1975). Even after plaque purification, BmNPV still shows this heterogeneity and this ratio does not change significantly after passage in vivo or in vitro (see Maeda and Majima, 1990). The mechanisms behind the MNPV and SNPV phenotypes pose an interesting scientific question with regard to the processes of protein association within cells and viral assembly.

Baculoviruses have been considered as efficient vectors for expression of foreign genes in insect cells (Luckow and Summers, 1988; Miller, 1988; Maeda, 1989a; Luckow, 1991). Baculoviral expression systems exhibit numerous advantageous characteristics including high expression rates, correct post-translational modifications, and authentic antigenic properties. The very strong promoter of the polyhedrin gene is essential for high level production of foreign genes by recombinant baculoviruses, however, the polyhedrin gene or gene product is not essential for viral progeny production. The polyhedrin protein accounts for about 30% of total cellular proteins at a late stage of infection. Furthermore, since polyhedra are visible under the light microscope, a recombinant NPV carrying a foreign gene after the polyhedrin promoter by replacement of the polyhedrin gene can be isolated easily by screening for infected cells lacking polyhedral production.

Two NPVs, BmNPV and Autographa californica NPV (AcNPV), are commonly used for expression experiments, however, the BmNPV system has the advantage of having a well studied and easy to use in vivo host, the silkworm (Maeda et al., 1985). To date more than two hundred different genes have been expressed using baculovirus expression vectors (see review of Luckow, 1991) for basic research and specific applications. Of current interest in medicine in the United States is the use of a recombinant baculovirus-expressed envelope protein of the AIDS virus (HIV-1) (see Maeda, 1989a) for testing as a possible AIDS vaccine. Recently, the silkworm and the baculovirus expression vector system has also been used for the production of canine interferon alpha for veterinary use in Japan.

Until recently most studies of baculovirus replication have been conducted using in vivo systems (larvae) mainly by histological, histochemical and microscopic analysis. Initial infection of larvae by baculoviruses occurs after viral particles are released from ingested inclusion bodies by alkaline protease degradation in the digestive juice. Released viral particles attach to microvilli of midgut cells and replication initiates in the midgut tissue, however, only a limited number of midgut cells, columnar and regenerative cells, are initially infected. After the replication of baculoviruses in midgut cells, budded non-occluded viruses from the basement membrane spread into various tissues of the body cavity of larvae (see Keddie et

al., 1989). The initial target tissue in the body cavity is hemocyte. At a late stage of infection (3-4 days post infection), viruses can infect and replicate in almost all larval tissues. Fat body is the major target organ of viral replication and production of polyhedra is observed in most fat body cells. At a very late stage of infection (one or two days prior to larval death), fat body and other cells start to degenerate, resulting in the whitish appearance of hemolymph due to the large number (more than a hundred per cell) of polyhedral inclusion bodies released from lysed cells.

The establishment of an in vitro replication system (established cell lines) and molecular biological techniques (see Granados and Federici, 1986) has led to a greater understanding of the replication of baculoviruses. Most insect viruses do not replicate or replicate poorly in established cell lines. AcNPV, however, which possesses relatively wide host specificity replicates rapidly in vitro and a great wealth of knowledge of the mechanisms of viral replication has been accumulated using this baculovirus as a model.

Historically insect viruses including baculoviruses have been classified (Latin name plus subfamily name) based on their host specificity which is generally narrow. Recently, however, DNA restriction analysis has allowed more sensitive and accurate classification of baculoviruses. Using these techniques the genetic relatedness of many

insect viruses have been quantitatively compared (e.g., Smith and Summers, 1982). In addition, in vitro systems have spawned plaque isolation techniques for the recovery of pure viral clones including mutant clones. Using these techniques, wild stocks of baculoviruses have been found to be genetically heterogeneous (Lee and Miller, 1979) and are sometimes mixtures of completely different viruses (Maeda et al., 1990). Taxonomical distance among baculoviruses has also been studied using hybridization techniques (Smith and Summers, 1982). Although Kondo and Maeda (1991) have recently shown that host specificity of baculovirus can be changed (expanded) by recombination of two baculoviruses having different host specificity, the real mechanisms of host specificity are still unknown.

To date about 40% of the entire genome of AcNPV has been sequenced and more than ten genes have been identified, sequenced, and characterized. The polyhedrin gene of AcNPV was the first baculoviral gene to be identified and sequenced (van Iddekinge et al., 1983). Structural polypeptides of AcNPV that have been isolated and sequenced include: envelope protein (Whitford et al., 1989), capsid protein (Thiem and Miller, 1989), DNA binding protein (Wilson et al., 1987), and polyhedral envelope protein (Russell and Rohrmann, 1990). Non-structural proteins that have been isolated and sequenced include: DNA polymerase (Tomalski et al., 1988); p74 related to virulence (Kuzio et al., 1989); immediate early genes (IE-0 (Chisholm and Henner, 1988), IE-

Henner, 1988), IE-1 (Guarino and Summers, 1987), IE-N (Carson et al., 1991), PE-38 (Krappa and Knebel-Morsdorf, 1991)); PCNA (ETL and ETS) which accelerate late genes (Crawford and Miller, 1988); ubiquitine-like gene (Guarino, 1990); DNA helicase (Lu and Carstens, 1991); superoxide dismutase (Tomalski et al., 1991); ecdysteroid UDP-glucosyl transferase (O'Reilly and Miller, 1989); and apoptosis-preventing protein (Clem et al., 1991). Several structural and non-structural genes from Orygia pseudotsugata NPV (Blissard and Rohrmann, 1990) and BmNPV (Maeda et al., 1991a) have also been isolated.

Unique repeated sequences are characteristic of the baculoviral genome (Arif and Doerfler, 1984; Cochran and Faulkner, 1983). AcNPV has 6 repeated sequences in five discrete regions (Guarino et al., 1986). Since all baculoviruses examined have repeated sequences (see Blissard and Rohrmann, 1990), these regions are considered to be essential in the baculoviral genome construct. The gene arrangement of baculoviruses is also relatively conserved, although insertions, deletions, and inversions of genes have been observed (Blissard and Rohrmann, 1990).

Initial studies on the regulation of gene expression examined viral protein synthesis in established cell lines by SDS polyacrylamide gel electrophoresis and radio-labeling of protein synthesis (e.g., Dobos and Cochran, 1980). Thirty to forty viral polypeptides have been identified and their expression has been shown to be temporally controlled.

The control of protein synthesis is related directly to transcription levels. Baculovirus gene expression has been characterized by Northern blot analysis, cDNA cloning, primer extension, CAT (chloramphenicol acetyltransferase) assay, and S1 nuclease mapping.

Baculovirus gene expression is classified into four phases: immediate-early, delayed early, late, and very late (see Blissard and Rohrmann, 1990). Immediate-early genes do not require viral gene products for their expression, i.e., host factor(s) can activate the gene expression of immediate-early genes. Four major immediate-early genes, IE-0, IE-1, IE-N, and PE-38 have been isolated and characterized (see above references). The IE-1 and IE-N gene products can trans-activate delayed early genes (Carson et al., 1988), however, the real functions of these genes including their target sites are still unclear. The TATA box and a CAGT motif located about 25 bp downstream of the TATA box are believed to be essential for immediate-early gene expression (Blissard and Rohrmann, 1990). Most identified immdiate-early genes possess these sequences. The polyhedrin (Rohrmann, 1986) and p10 (Leisy et al., 1986) genes are expressed at a very late stage of infection. The upstream region of (very) late genes contain the consensus sequence ATAAG. Transcription of these genes starts from the second adenine of this consensus sequence. At a late stage of infection early gene and host gene expression are suppressed. Furthermore, splicing of host and viral genes

is also blocked (Chisholm and Henner, 1988). Ooi and Miller (1990) have hypothesized that this suppression is caused by the production of antisense RNA.

I have chosen two baculoviruses, BmNPV (T3 isolate) and SlNPV (OT2 isolate), for the study of baculoviral replication in insect cells. BmNPV is important as an expression vector and in the sericultural industries. SlNPV has potential for use as an effective controlling agent for a major pest of vegetable crops in Japan. Since published molecular biological studies of these two viruses are very limited, physical maps, which are essential for further experiments at the molecular level, of the viral genome, were initially constructed (Section III). A DNA fragment library of the viral genome, which can be directly used for gene analysis, was also constructed. In section IV, the characteristics of the polyhedrin genes of BmNPV T3, SlNPV OT2 and various mutants are reported. To test the relationship between nucleotide sequence and phenotypic characteristics such as the shape of the polyhedra, recombinant viruses were constructed. Four different polyhedrin genes were isolated from various mutants and the entire sequence of these genes was determined. Various mutants with appropriate deletions in the polyhedrin gene of BmNPV were examined to determine the mechanisms of 1) nuclear localization of polyhedrin, 2) crystallization of polyhedrin, and 3) shape and size determination of inclusion bodies. In section V, sequence analysis of the BmNPV genome

is reported. The structure of the repeated sequences of BmNPV were characterized after isolation and sequencing. A sequence of about 17 kb (13%) of the BmNPV genome, which contains four immediate-early genes was determined, and gene structure and expression in this region was examined.

III. Construction of DNA fragment libraries and physical maps of two nuclear polyhedrosis viruses.

A. Introduction

Recently, over 100 isolates from four different wild stocks of NPVs of Spodoptera litura (the same or closely related species to S. littoralis), which is a major agricultural pest in Africa, Asia, and Mediterranean regions, have been plaque-purified and characterized (Maeda et al., 1990). These S. litura NPV (SlNPV) isolates were classified into four distinct groups (SlNPV-A, SlNPV-B, SlNPV-C, and AcNPV (an AcNPV variant)) by in vitro host range and DNA restriction endonuclease patterns using EcoRI. Another NPV, the BmNPV T3 isolate, has been used extensively for basic molecular biological research (Maeda et al., 1991a) and its applications for foreign gene expression (Maeda et al., 1985; Maeda, 1989a), and as a model system for the construction of recombinant viral insecticides (Maeda et al., 1991b).

Plaque isolation and DNA analysis techniques have demonstrated the precise genetic relatedness of NPVs. AcNPV is the most well studied baculovirus at the molecular level and several AcNPV variants have been isolated and characterized (see Blissard and Rohrmann, 1990). Construction of a physical map of the viral genome using restriction endonucleases is essential for further experiments at the molecular level. Restriction

endonuclease maps have been constructed for several NPVs including: AcNPV and its variants (Miller and Dawes, 1979; Smith and Summers, 1979; Vlak, 1980; Cochran et al., 1982; Brown et al., 1984), Anticarsia gemmatalis NPV (Johnson and Maruniak, 1989), Panolis flammea NPV (Possee and Kelly, 1988), Mamestra brassicae NPV (Wiegers and Vlak, 1984; Possee and Kelly, 1988), Orgyia pseudotsugata NPV (Chen et al., 1988), Heliothis zea SNPV (Knell and Summers, 1984), Spodoptera littoralis NPV (Croizier et al., 1989) and Spodoptera frugiperda NPV (Loh et al., 1981; Maruniak, et al., 1984).

In this section, I describe the construction of restriction fragment libraries covering the entire genomes of BmNPV T3 and SlNPV OT2 (AcNPV variant) in plasmids, and construction of their physical maps for several endonucleases. In addition, five areas of EcoRI-rich repeated sequences were found and localized in the physical map (see section V for details).

B. Materials and Methods.

Chemicals, media, gene cloning techniques, and plasmid preparations: are described in Figs. III-1 to III-3.

Virus: BmNPV. A plaque purified isolate, T3, of BmNPV (Maeda, 1984) and a plaque purified isolate, OT2, of SlNPV (AcNPV variant) (Maeda et al., 1990) were used. For purification of BmNPV viral particles, polyhedral inclusion bodies propagated in the silkworm were used. Inclusion

A. Ligation: (using the Takara Ligation kit)

1. Take 2-5 ul of digested plasmid prefered for ligation, and add 4-8 times volume of A buffer and 1 times volume of B buffer. Mix by gentle tapping (do not vortex).

2. Incubate in a 16 C water bath for 30 min.

B. Transformation:

3. Mix the following in a microfuge tube:
   5 ul ligated plasmid (approx. 0.1 ug)
   40 ul competent cells

4. Incubate 10 min on ice.

5. (Invert once to mix) and incubate 40 sec in a 43.5 C water bath.

6. Quickly transfer tube to ice bath (0 C) for 2 min.

7. Add 70ul TUM without ampicillin and incubate at 37 C (air incubator) for 10 min. Mix by gentle tapping.

8. Spread on a plate (TA plates or Taxi plates) using sterile spreader and incubate at 37 C at least 7 hours.

Fig. III-1. Ligation and transformation with plasmid DNA

1. Add 1 ml of TUM/AMP* medium into sterile microfuge tubes using a sterile pipet.

2. Pick up a single colony using a sterile toothpick. Touch the toothpick to a replica plate, then insert it into a microfuge tube.

3. Incubate the microfuge cultures at least 7 hours at 37 C with rocking (175 rpm). Incubate the replica plates at 37 C without rocking.

4. Centrifuge cultures at 4000 rpm for 1 min.

5. Disgard supernatant and add 150 ul (300 ul for sequencing) of lysozyme-STET* (0.5mg/ml) solution into each tube and vortex until the pellet is completely disolved.

6. Boil tubes in a boiling water bath for 1 min.

7. Place each rack into an ice/water bath until cold (2 minutes).

8. Centrifuge at 12,000 rpm for 15 min.

9. Remove precipitate with sterile toothpicks.

10. Add 100 ul (200 ul for sequencing) of isopropanol. Mix well by inversion and shaking; incubate at -80 C for at least 10 min.

11. Centrifuge at 12,000 rpm for 10 min. Discard supernatant and dry in vacuum for 30 min or until dry.

***TUM/AMP Medium**

Bacto tryptone (Difco)		10 g
Bacto yeast extract (Difco)		5 g
NaCl		5 g
KCl		0.7 g
$MgSO_4.7H_2O$		2.5 g
H_2O	to	1 l

Autoclave for 15 minutes at 121 C. Stock at 5 C or room temperature. Add Ampicillin to 50 mg/liter.

***STET Buffer**

Sucrose		80 g
Triton X-100		50 g
0.5M EDTA		100 ml
1M Tris-Cl pH8		50 ml
H_2O	to	1 l

Fig. III-2 Minipreparation of plasmid DNA

1. Add 50 ml of TUM/Amp medium into a 150 ml Ehrenmyer flasks and innoculate media with a single bacterial colony using a sterile toothpick.

2. Incubate (140-160 rpm) at least 7 hours at 37 C.

3. Transfer culture medium into a 50 ml centrifuge tube and centrifuge at 3000 rpm for 10 min at 5 C.

4. Discard supernatant by decantation. Add 3ml of TE (20mM Tris-Cl, 1 mM EDTA pH 8.0). Completely suspend pellet by vortexing.

5. Add 1 ml of 2 mg/ml lysozyme in 25% sucrose-20mM Tris-Cl (pH 7.5). Mix well and incubate at room temperature for 5 min.

6. Add 4 ml of 1.5% SDS, 0.3 N NaOH. Shake vigorously. Incubate at room temp. for 5 min.

7. Add 4 ml of 5M Potassium Acetate. Shake vigorously. Centrifuge at 3,000 rpm for 10 min at 5 C.

8. Transfer supernatant into another 50 ml tube. Add 6 ml of isopropanol. Mix well. Incubate 15 min at room temp.

9. Centrifuge at 3,000 rpm for 10 min at 5 C. Discard supernatant by decantation. Turn upside-down on a Kimwipe.

10. Add 300 ul of RNase (10ug/ml) in TE (10:1), suspend pellet by vortexing and transfer to a microfuge tube. Centrifuge at 12,000 rpm for 2 min to remove precipitates.

11. Transfer into a new microfuge tube. Incubate for 30 min at 37 C.

12. Extract with 300 ul of phenol-chloroform (vortex well). Centrifuge 12,000 rpm for 3 min and transfer aqueous layer to a new tube.

13. Extract once with 300 ul of chloroform (vortex well). Centrifuge 12,000rpm for 3 min. Transfer aqueous layer to a new tube.

14. Transfer aqueous layer into another tube. Add 15ul 5M NaCl, 750ul EtOH and mix well. Incubate 30 min at -20 C.

15. Centrifuge 12,000 rpm, 10 min. Discard supernatant and dry under vacuum. Add 200-300 ul TE (10mM Tris-Cl,1mM EDTA) and suspend well.

Fig. III-3. Large-scale preparation plasmid DNA

bodies were dissolved in alkaline solution and the released viral particles were purified by ultracentrifugation (Kawarabata and Matsumoto, 1973). SlNPV OT2 was purified from the culture medium of infected Sf cells as described by Maeda et al. (1989).

Endonuclease analysis: Viral DNA was extracted from viral particles after treatment with proteinase K (Merk) in the presence of 1% SDS (Maeda, 1989b). DNA was cleaved by digestion with EcoRI, HindIII, PstI, BamHI, KpnI, or SmaI (New England Biolabs) under conditions recommended by the supplier. The cleaved fragments were separated on 0.7% agarose gels along with lambda phage DNA size markers using a Tris-acetate buffer system (Maniatis et al., 1982).

Southern blot analysis: Southern blot analysis was performed by a modification of the method of Maniatis (1982) as shown in Fig. III-4. Deionized formamide was prepared by adding 50-100 g of ion exchange resin (BioRad AG 501-X8) per liter of formamide and shaking gently for at least 1 hour. Salmon sperm DNA (Type-III sodium salt, Sigma) was dissolved in distilled water to a concentration of 2 mg/ml, stirred for 2 to 4 hours at room temperature, sheared by several passages through an 18 gauge needle, and boiled for 10 minutes. The DNA was stored at -20 C in small aliquots until used.

C. Results and Discussion.

1. BmNPV.

1. After electrophoresis, carefully transfer gel into a glass baking dish.

2. Denature the DNA by soaking the gel in about two volumes of 1.5 M NaCl and 0.5 M NaOH for 15 min at room temperature with constant shaking. Replace with 1.5 M NaCl and 0.5 M NaOH and soak for an additional 15 min.

3. Neutralize the gel by soaking in more than two volumes of 1 M Tris-Cl (pH 8.0) and 1.5M NaCl for 15 min at room temperature with constant shaking. Replace with 1 M Tris-Cl and 1.5M NaCl and shake for an additional 15 min.

4. Transfer DNA to a nitrocellulose filter using 10x SSC.

5. Wash the filter with 5x SSC* at room temperature.

6. Place filter onto Whatman 3MM paper and allow to dry.

7. Wrap filter in 3MM paper and bake for 1 hr. at 80 C under vacuum.

8. Wet filter in 5x SSC. Place filter in a hybridization incubator tube, add 10-15 ml of prehybridization solution*, and incubate 24 hours at 42 C. After prehybridization add 10 ul of denatured probe and incubate 6 to 24 hours at 42 C.

9. Wash twice with 2x SSC, 0.1% SDS (20 min per wash):

10. Wash twice with 0.1x SSC, 0.1% SDS (20 min per wash):

11. Air dry filter, and expose to X-ray film at -80 C.

***20x SSC**

NaCl		1753	g
Sodium Citrate		882	g
H_2O	to	10	l

***Prehybridization Solution**

Deionized Formamide (50%)	50 ml
50x Denhardt's solution* (5x)	10 ml
10% SDS (0.5%)	5 ml
2 mg/ml Salmon Sperm DNA (0.1 g/ml)	5 ml
20x SSC (5x)	25 ml
dd H_2O	5 ml

***50x Denhardt's Solution**

Ficoll		5	g
polyvinylpyrrolidone		5	g
BSA (Pentax Fraction V)		5	g
dd H_2O	to	1	l

Fig. III-4. Southern Transfer and Hybridization

Fig. III-5, shows BmNPV DNA digested with the indicated enzymes. No submolar bands, sometimes found in the preparation of other NPVs (Miller and Miller, 1982), were seen. Molecular weights of the fragments were estimated as shown in Table III-1 by comparison to the migration of the size markers. Molecular weights of some of the larger fragments were estimated by summing the sizes of the smaller fragments of which it was composed as described later. The molecular weight of the entire BmNPV genome was estimated to be about 130 kbp by summing the weights of the fragments generated by the restriction endonucleases. This estimated size is quite similar with that of AcNPV (Lee and Miller, 1978).

DNA fragments digested with EcoRI, HindIII, KpnI, PstI, or BamHI were cloned into pBR322, pUC9, or pUC19. Two large fragments of EcoRI, EcoRI A and B, were first cloned into Charon 4A (Maniatis, 1982), digested with restriction endonucleases, HindIII and/or BamHI, and then subcloned into pUC plasmids. Most of the cloned DNA was authenticated by comigration in agarose gels with digested viral DNA. All EcoRI fragments larger than 500 bp were successfully cloned into plasmids. These EcoRI clones and additional clones of the 22 DNA fragments listed in Table III-1 completely overlapped the entire viral genome without breaks at any restriction site.

Hybridization analysis was employed to construct physical maps of the BmNPV genome. Viral DNA was first

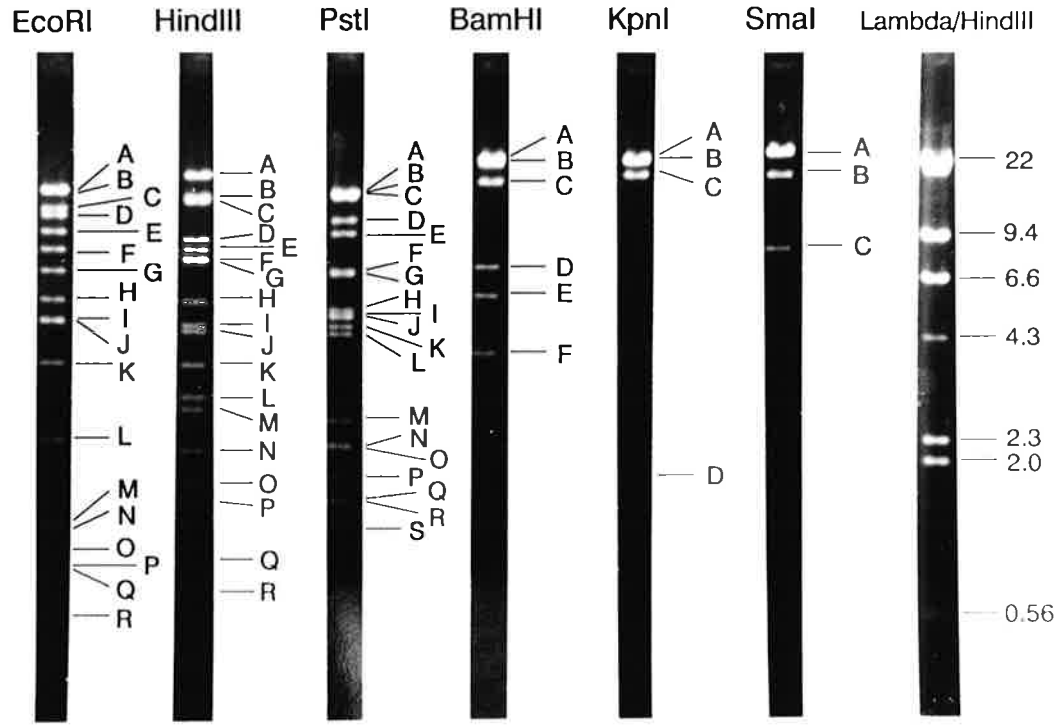


Fig. III-5. Cleavage patterns of BmNPV DNA by the restriction endonucleases EcoRI, HindIII, PstI, BamHI, KpnI, or SmaI. Viral DNA cleaved with these restriction endonucleases were separated on a 0.7 % agarose gel. Lane marked Lambda/HindIII shows molecular weight markers in kbp.

Table III-1. Sizes (kbp) of BmNPV restriction fragments

Fragment	EcoRI	HindIII	PstI	BamHI	KpnI	SmaI
A	20.4*	30.0	17.5*	54	52	97
B	20.1*	17.0	17.5*	36	48	25
C	14.5*	15.5*	17.0	22	28	8.0
D	13.9*	10.0	12.5*	7.3*	1.8*	
E	10.5*	8.9	10.8*	6.0*		
F	8.7*	7.8	7.2	3.9*		
G	7.5*	7.8	7.2*			
H	6.6*	5.8*	5.5*			
I	5.2*	5.1*	5.4*			
J	5.2*	4.8	5.4			
K	3.9*	3.8*	4.9*			
L	2.4*	3.1*	4.6			
M	1.3*	3.0	2.8*			
N	1.2*	2.2	2.3*			
O	1.0*	1.7	2.3			
P	0.9*	1.5	1.9*			
Q	0.8*	1.0*	1.5*			
R	0.5*	0.7	1.5			
S			1.3			

* cloned into plasmid

digested with the same restriction endonucleases used for cloning and electrophoresed on agarose gels. The separated DNA fragments were Southern-transferred onto a nitrocellulose filter, fixed at 80 C, pre-hybridized, then hybridized with cloned plasmid DNA probes labeled with [^{32}P]-dCTP (Maniatis, 1982). Hybridizations were carried out in 50% formamide at 42 C for 6-16 hours and the filters were washed with 0.1x SSC at room temperature. Analysis of the hybridization data showed that most of the sequences of the BmNPV genome were unique (see section V).

By comparing the overlapped areas of DNA for each restriction fragment a preliminary physical map was constructed. To obtain highly detailed maps, the larger cloned DNA fragments were digested with two or more restriction enzymes and physical maps of these fragments were constructed. To confirm the location of these smaller fragments, they were compared with the initially cloned fragments or with original viral DNA by electrophoresis in agarose gels. By combining the various data obtained above, detailed physical maps were constructed for six restriction endonucleases (Fig. III-6). The zero-point of the physical map was established as one end of the EcoRI E fragment following the proposal of Vlak and Smith (1982). The zero-point was chosen here because the polyhedrin gene (indicated by the arrowhead in Fig. III-6) has been mapped in this fragment (Maeda et al., 1985).

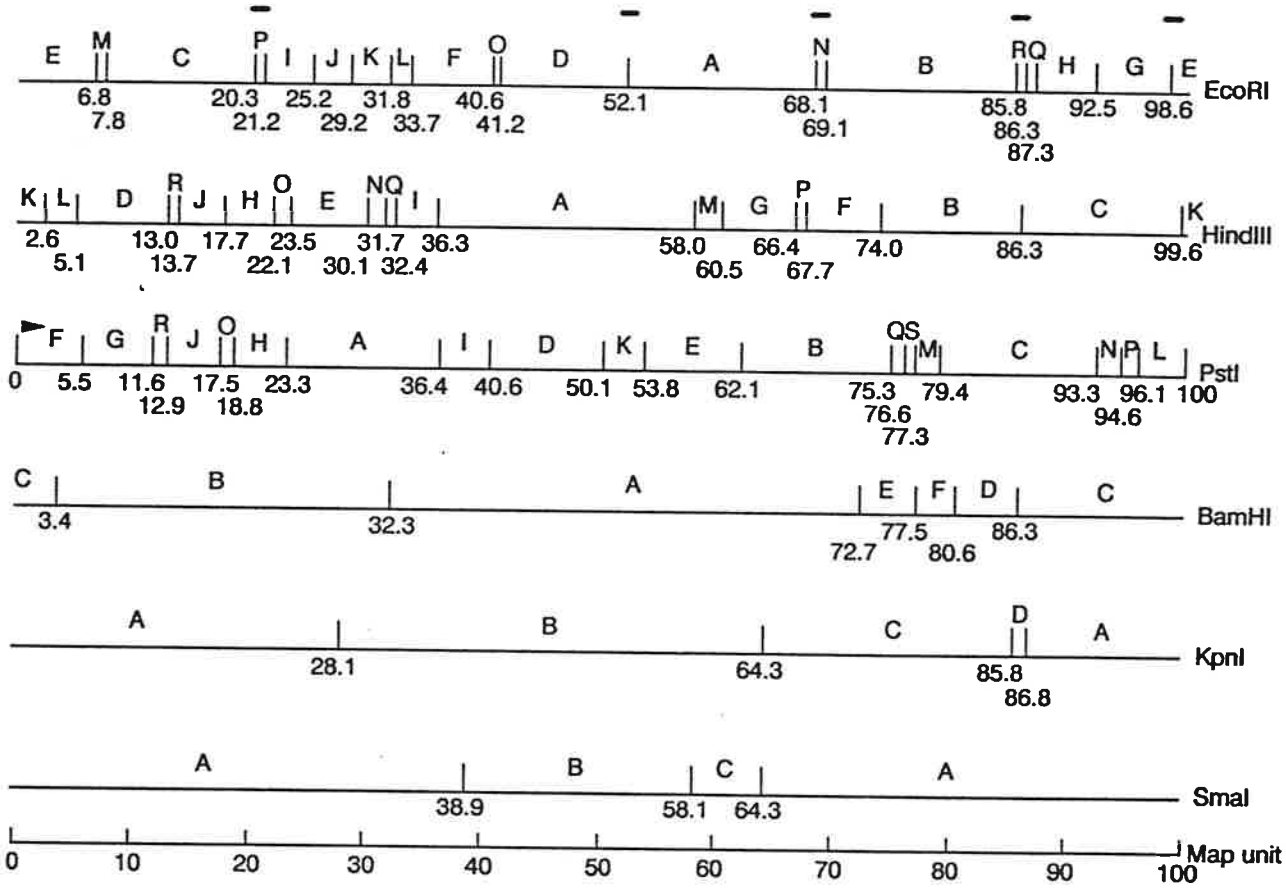


Fig. III-6. Physical map of BmNPV DNA for EcoRI, HindIII, PstI, BamHI, KpnI, and SmaI. The circular DNA is presented in linear form. Map units are calculated based on the total length of the genome being 100. The arrowhead indicates the position and direction of the polyhedrin gene. Bars represent areas of repeated sequences.

From the hybridization experiments, five regions containing homologously repeated sequences rich in EcoRI sites were found in the genome. To confirm the existence of the repeated sequences, cloned or subcloned plasmids (HindIII H, PstI K, HindIII-PstI (67.7-75.3 map unit) of PstI B, KpnI D, and PstI-HindIII (96.1-99.6 map unit) of HindIII), were digested with EcoRI and analyzed electrophoretically. As expected, several EcoRI fragments with molecular weights less than 400 bp were identified on an 1.5% agarose gel. These areas containing several smaller EcoRI fragments seemed to be the so-called repeated sequences found in AcNPV (Erlandson et al., 1984; Guarino et al., 1986) and other baculoviruses (Kuzio and Faulkner, 1984; Arif and Doerfler, 1984). Three repeated sequences could not be mapped exactly, however their locations were at either end or on both ends of the small EcoRI N, P, or R fragments, these areas were mapped and are indicated by the bars in Fig. III-6. All five repeated sequences of the BmNPV genome were located at similar positions as the repeated sequences mapped in the AcNPV genome (Summers and Smith, 1987). Furthermore, the positions of 7 (3.4, 28.1, 38.9, 64.4 (KpnI), 64.4 (SmaI), 77.5, and 80.6 map units) of 14 restriction endonuclease sites for BamHI, KpnI, and SmaI were located very closely to the analogous sites mapped in the AcNPV genome. This is consistent with data showing around 80% DNA homology of the polyhedrin gene areas between BmNPV and AcNPV (Maeda et al., 1985; Iatrou et al., 1985).

DNA hybridization analysis also indicated that the BmNPV genome was more than 50% homologous to the AcNPV genome by calculation from the intensity of film exposed to hybridized viral DNAs.

It has been reported that the genome organization of NPVs is relatively conserved even between viral DNAs with low DNA homology (Leisy et al., 1984). From our analysis AcNPV and BmNPV seem to be closely related viruses in terms of sequence homology. This is interesting because of the significant differences in the phenotypical and biological characteristics of the two viruses, such as host range. We are now analyzing the viral genome based on the physical map we constructed. The physical map and gene library will be useful for further genetic studies of BmNPV.

2. SlNPV.

SlNPV OT2 was originally isolated from a stock of SlNPV collected in Ogasawara, Japan and characterized as an AcNPV variant (Maeda et al., 1990; Kondo and Maeda, 1991). Isolates of the AcNPV group were obtained only from plaque assays on TN-368 cells and only from one of four SlNPV stocks obtained from various regions in Japan, indicating that this group is a minor part of SlNPV in Japan. The production rate of polyhedra per cell varied depending on the AcNPV isolate. A multiple polyhedra (MP) type isolate, OT2, was selected for the following experiments.

Restriction enzyme analysis is the most sensitive method for characterizing closely related viruses. To determine the genetic relatedness of SlNPV OT2 and AcNPV L1, purified OT2 and AcNPV L1 DNAs (Miller and Dawes, 1979) were digested with seven different restriction endonucleases EcoRI, HindIII, PstI, BamHI, KpnI, XhoI, and SmaI, electrophoresed in a 0.7% agarose gel and stained with ethidium bromide. As shown in Fig. III-7, the restriction patterns of cleaved OT2 DNA were similar to those of AcNPV L1. Four out of 20 EcoRI digested OT2 fragments migrated differently compared to the EcoRI digested AcNPV fragments. Four out of 25 HindIII fragments of OT2, one out of 15 PstI fragments, 3 out of 14 XhoI fragments, 4 out of 8 BamHI fragments, 0 out of 4 KpnI fragments, and 1 out of 4 SmaI fragments, migrated differently from the corresponding digested AcNPV fragments. In total about 81% of the OT2 fragments were indistinguishable from those of AcNPV L1, indicating that OT2 is genetically close to AcNPV. The size of each OT2 fragment was calculated by comparison with lambda DNA cleaved with HindIII (Table III-2). The sizes of larger fragments (> 20 kb) were calculated by summing the sizes of smaller fragments of which they were composed. The estimated genome size of OT2 was about 130 kb, which is similar to that of AcNPV (Lee and Miller, 1978).

To further analyze the genome structure, a DNA fragment library of OT2 was constructed in pUC19 plasmids using seven different restriction endonucleases EcoRI, HindIII, PstI,

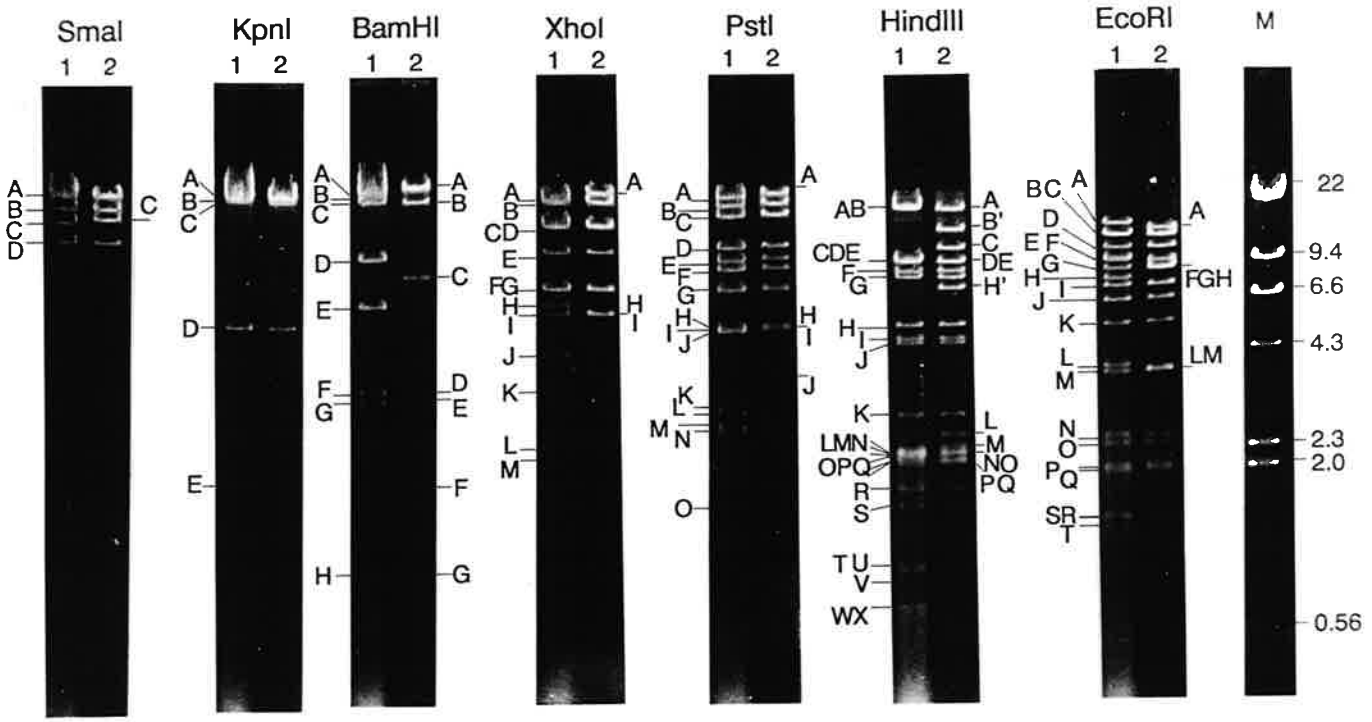


Fig. III-7. Restriction endonuclease analysis of the OT2 genome. Viral DNA (approximately 0.5 ug/lane) was digested with the indicated restriction endonucleases and electrophoresed on a 0.7% agarose gel. Lanes 1, OT2; lanes 2, L1. Lane M (size markers) is lambda DNA cleaved with HindIII. Sizes in kb are shown at the far right.

Table III-2. Sizes (kbp) of SlNPV OT2 restriction fragments

Fragment	EcoRI	HindIII	PstI	BamHI	KpnI	XhoI	XbaI
A	16.2*	20.3	24.5	49	47.8	25	18.0
B	13.0*	20.0	20.5	30.5	43.6	23.3	16.0
C	12.9*	10.5*	18.0*	23.4	29.9	14.4	15.5*
D	10.6	10.5*	11.0*	10.5	5.3*	14.4	12.9*
E	9.7*	10.4*	10.2*	6.3*	1.9*	10.4	11.5
F	9.5*	8.5*	9.3*	3.2*		7.4	11.5
G	8.8*	8.2*	7.5*	2.9*		7.3	10.8*
H	8.0*	5.5	5.5*	1.0*		6.2*	10.3*
I	7.4*	5.0*	5.5			5.7*	8.7*
J	6.7*	4.8	5.5*			4.1	7.9*
K	5.4*	2.8*	3.2*			3.0*	3.6*
L	3.9*	2.3	2.9*			2.25*	1.7*
M	3.7*	2.2*	2.7*			2.15	1.6*
N	2.5*	2.2	2.6			1.2*	
O	2.3*	2.1*	1.6*				
P	2.0*	2.1					
Q	1.9	2.05					
R	1.5*	1.8*					
S	1.5	1.6					
T	1.3	1.1					
U	1.0	1.0					
V		0.9					
W		0.8					
X		0.73					
Y		0.7					

* cloned into plasmid

BamHI, XbaI, XhoI, and KpnI. Plasmids in the library were extracted by the heat denaturation procedure (Maniatis et al., 1982), digested with the restriction endonucleases used for cloning, and electrophoresed in a 0.7% agarose gel with digested viral DNA in order to confirm the existence of an insert derived from viral DNA. Totally 57 different DNA fragments were successfully cloned into the plasmids (Table III-2).

Southern blot hybridization was performed to construct physical map of the OT2 genome using labeled fragments from the library as probes. Most probes hybridized specifically to a single fragment or a limited number of fragments on the nitrocellulose filter. However, the labeled HindIII Q fragment hybridized to several fragments, indicating the existence of EcoRI-rich repeated sequences (see section IV). To confirm the presence of repeated sequences, the corresponding HindIII fragments from the DNA fragment library were digested with EcoRI and analyzed by comigration with non-EcoRI digested fragments. HindIII A, B, F, L, and Q fragments generated several fragments smaller than 300 bp, indicating the existence of repeated sequences (data not shown). HindIII fragments G, M, N, O, and P did not generate small EcoRI fragments, indicating the lack of repeated sequences containing EcoRI. Rough physical maps were initially constructed by analysis of this preliminary data. More detailed physical maps were constructed by further hybridization analysis using plasmids in the

fragment library which were double digested with restriction endonucleases. The detailed physical maps of the OT2 genome for EcoRI, HindIII, PstI, BamHI, XhoI, KpnI, and SmaI are shown in Fig. III-8. The map units of OT2 were adjusted to that of HR3 as reported by Cochran et al. (1982) for ease of comparison, i.e., the two insertions and one deletion found only in the OT2 genome were not incorporated into the map, but rather shown separately at the top of the figure.

When the restriction patterns of OT2 and L1 (or HR3) (Cochran et al., 1982; AcNPV L1 and HR3 are nearly identical except for the existence of an additional HindIII site in the HindIII B (20.0 kb) fragment of L1) were compared for the HindIII, BamHI, PstI, EcoRI, XhoI, and KpnI physical maps, the cleavage patterns of BamHI showed the greatest disparity; 4 out of 8 BamHI digested OT2 fragments were different from those of AcNPV L1. This difference was caused by one deletion and two additions of BamHI sites in AcNPV L1. Deletion of a BamHI site of L1 between BamHI C (8.5 kb) and BamHI F (1.92 kb) at the 4.8 map units generated BamHI D (11.5 kb) of OT2. Insertion of two BamHI sites at 35.0 and 70.8 map units in the BamHI A (86.5 kb) fragment of L1 generated the three fragments, BamHI A (49.4 kb), BamHI B (30.8 kb), and BamHI E (6.3 kb) in OT2. The PstI patterns showed two detectable differences. PstI A (24.5 kb) of OT2 contained a 0.5 kb deletion causing it to be slightly smaller than PstI A (25.0 kb in our estimation; 25.7 kb by Cochran et al. (1982) (this value for L1 will be

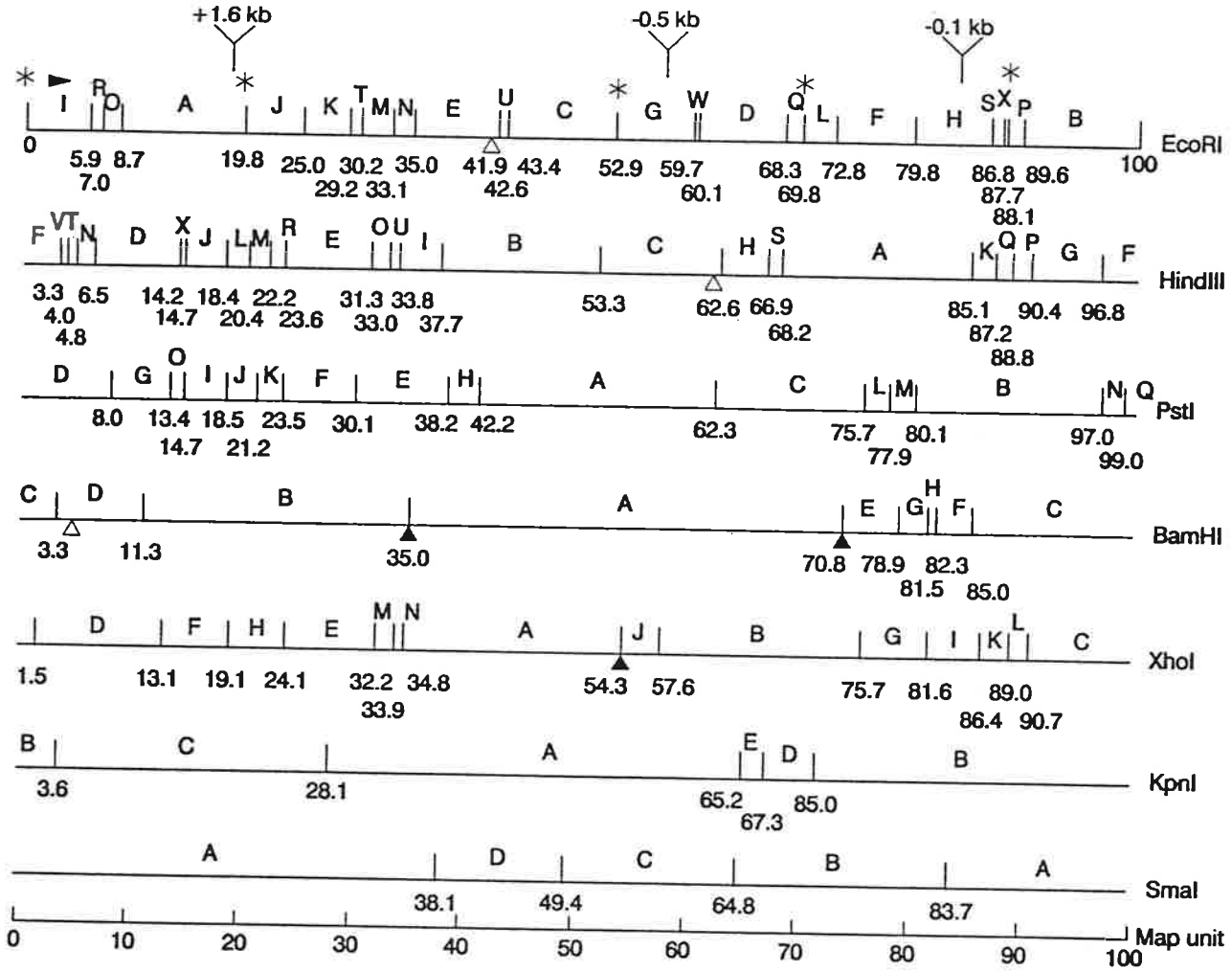


Fig. III-8. Physical maps of the OT2 genome for the indicated restriction endonucleases. The map position of OT2 is adjusted to the map units for HR3 (Cochran et al., 1982). Two insertions containing HindIII and XhoI sites and one deletion are shown at the top of the map. Asterisks indicate positions of repeated sequences. Differences in restriction enzyme sites between OT2 and HR3 are shown by triangles. △ indicate deletions of restriction endonuclease sites in the OT2 genome. ▲ indicate addition of restriction endonuclease sites in the OT2 genome. Detailed maps around the 1.6 kb insertion (19 map unit) are shown in Fig. III-9.

used hereafter)) of L1. PstI J (5.5 kb) of OT2 contained a 1.6 kb insertion compared to PstI J (3.45 kb) of L1. This insertion was also confirmed by the size difference of EcoRI A (14.2 kb) of L1 and EcoRI A (16.2 kb) of OT2. The EcoRI digestion patterns showed three detectable differences. One EcoRI site was deleted between EcoRI F (8.8 kb) and EcoRI V (0.94 kb) of L1 resulting in the 9.7 kb EcoRI E fragment of OT2. A 0.1 kb deletion in EcoRI H (8.7 kb) of L1 resulted in the 8.0 kb EcoRI H fragment of OT2. A 0.1 kb insertion in EcoRI L (3.8 kb) of L1 resulted in the 3.9 kb EcoRI L fragment of OT2. The XbaI digestion patterns showed no detectable difference. The addition of a XhoI site in XhoI A (29.2 kb) of L1 resulted in the 25 kb XhoI A and 4.5 kb XhoI J fragments of OT2. The SmaI digestions patterns produced minor detectable differences. SmaI C of OT2 was slightly smaller than that of L1 as a result of a deletion in the EcoRI G fragment (52.9-59.7 map position) of L1 as described earlier. No detectable differences in the KpnI digestions patterns were observed between OT2 and L1.

AcNPV variants have been reported from various lepidopteran insects including: Trichoplusia ni (Miller and Dawes, 1978; Smith and Summers, 1979), Galleria mellonella (Smith and Summers, 1979), Rachiplusia ou (Summers et al., 1980), and Spodoptera exigua (Brown et al., 1984). These AcNPV variants showed similar restriction patterns. Of these variants, GmNPV, showed the greatest difference in its restriction patterns compared to the other AcNPV variants

including Trichoplusia ni NPV (TnNPV). The BamHI digestion patterns showed the most differences between GmNPV and the other AcNPV variants. The BamHI pattern of OT2 seemed to be completely identical to that of GmNPV (Smith and Summers, 1979). Furthermore, the EcoRI pattern of OT2 was identical to that of GmNPV, and slightly different from five AcNPV isolates and TnNPV (Smith and Summers, 1979). Only OT2 and GmNPV possessed the larger (16.2 kb) EcoRI A fragment and the smaller (8.0 kb) EcoRI H fragment. When the XhoI digestion patterns were compared, only OT2 and GmNPV contained 25 kb (XhoI A) and 4.3 kb (XhoI J) fragments, while all of the other isolates possessed a 29.2 kb XhoI A fragment presumably corresponding to the XhoI A and J fragments of OT2 and GmNPV. OT2 and GmNPV also had smaller 10.5 kb HindIII C fragments, while AcNPV E2 (Fraser et al., 1983) and L1 have 11.1 kb HindIII C fragments. These results indicated that OT2 is genetically closer to GmNPV than to AcNPV.

The GmNPV EcoRI A fragment seemed to contain the 1.6 kb insertion found at 19 map units in OT2, since it was larger than the corresponding EcoRI A fragments of the other AcNPV variants. To further analyze this area, the PstI J fragment of OT2 in the constructed library was compared to the PstI J fragment of L1 cloned into pTZ18R. Since additional HindIII and XhoI sites were speculated in OT2 from previous restriction enzyme analysis, these two fragments were first digested with HindIII and/or XhoI. As shown in Fig. II-9,

an additional three XhoI and one HindIII fragments were observed in OT2. Two fragments from OT2 and L1 generated by HindIII and XhoI double digestion migrated to identical positions (0.7 kb and 0.5 kb) on an agarose gel (Fig. III-9), indicating that both ends of the PstI J fragment covered by these fragments were presumably identical (Fig. III-9). EcoRI digestion revealed different patterns between L1 and OT2. EcoRI cleavage of the L1 fragment generated 7 fragments (4 discrete bands in a 3% gel presumably due to repeated sequences) as previously reported (Cochran and Faulkner, 1983; Guarino et al., 1986). Although the OT2 fragment generated many EcoRI fragments 70-80 bp in length, fragments longer than 100 bp were not detected. Detailed physical map analysis showed approximately 20-25 of these 70-80 bp fragments. The positions of the four XhoI sites, two HindIII sites, and PstI J fragment are shown in Fig. III-9. Repeated sequences containing EcoRI sites extended through the second XhoI site as indicated in Fig. III-9.

Several reports on the insertion of DNA fragments into the baculovirus genome have been published. Hot spot(s) for DNA insertion have been found in the AcNPV genome between 8.4 and 9.6 map units (Kumar and Miller, 1987), 35.5-37.7 map units (Fraser et al., 1983; Beams and Summers, 1988; 1989), and 80-86 map units (Miller and Miller, 1982). It has also been reported that insertions which originated from the chromosomal DNA of the cell line used were often observed during serial passages in cell culture in these

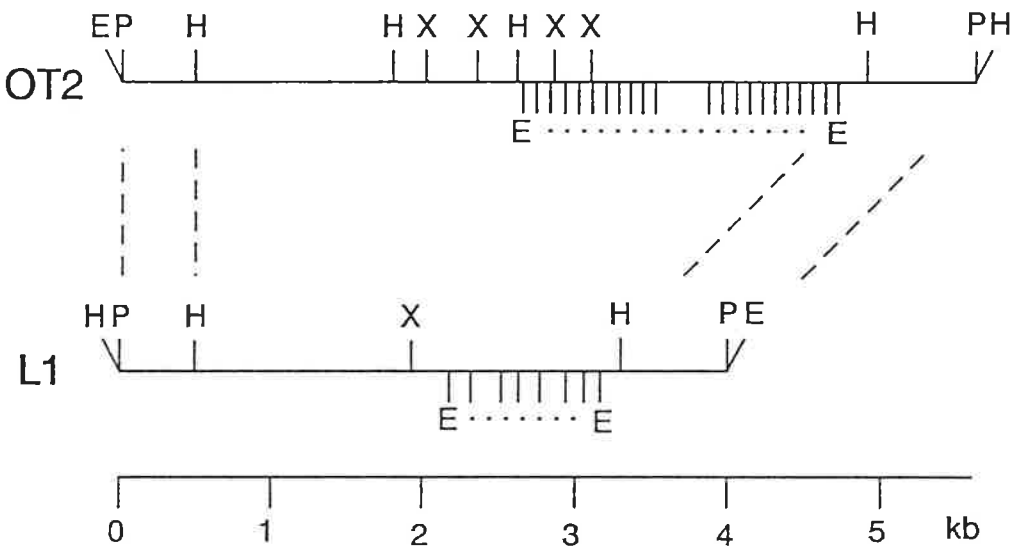


Fig. III-9. Physical maps of the PstI J fragments of OT2 and L1. The positions of several restriction endonucleases are shown. Symbols, P, PstI; X, XhoI; H, HindIII, and E, EcoRI.

areas. However, the insertion found at the 19 map unit position seemed not to be related to this type of insertion, since 1) the insertion site was different from other previously reported sties and 2) the physical mapping pattern of the inserted area (Fig. III-9) was unique.

Homology between OT2 and AcNPV was examined by comparing the nucleotide sequences of their polyhedrin genes. The polyhedrin gene of AcNPV has already been published (Iddekinge et al., 1983). The nucleotide sequence of the SlNPV OT2 polyhedrin gene was determined by dideoxy sequencing (see section IV). The nucleotide sequences of the polyhedrin gene was completely identical to that reported for the AcNPV L1, E2, and HR3 isolates, indicating that OT2 is closely related to previously characterized AcNPV isolates such as L1 and E2. Completely identical sequence homology of the polyhedrin genes was not expected from the restriction endonuclease patterns showing an average genomic difference of 19%. The perfectly conserved nucleotide sequence in the polyhedrin gene region may be explained by 1) the importance of the two genes for viral growth or replication, 2) specific insertions or deletions in other region, and/or 3) the existence of sequences (genes) essential for replication in Galleria mellonera other than the polyhedrin gene.

Restriction enzyme analysis showed that the OT2 isolate was very closely related to GmNPV, which is characterized as a variant of AcNPV (Smith and Summers, 1979). AcNPV

variants have also been isolated from NPV stocks from several lepidopteran insects by plaque purification, however, they have not been isolated in Japan until our finding. This is probably due to the fact that the original hosts of AcNPV such as Tricoplusia ni and Heliothis virescens do not exist or exist only as minor species in Japan, while G. mellonella is commonly found in Japan. OT2 was found as a very minor portion (less than 1%) of the population in only one of the four SlNPV stocks (Maeda et al., 1990) examined. These observations suggest that OT2 is a virus which originated in G. mellonella. The specific DNA pattern of GmNPV may be related to the specificity of this virus to G. mellonella, i.e., specific sequences (genes) may provide advantages to this virus.

IV. Cloning, sequence analysis, and expression of polyhedrin genes and their crystallization and nuclear localization mechanisms

A. Introduction

Baculoviruses produce many (presumably 100-150) structural and nonstructural polypeptides. The unique characteristic of producing two different types of progeny is considered to be controlled by mechanisms unique to baculoviruses. During an early stage of infection all viral components are transported to the cell surface from the cytoplasm (for gp64) and nucleus (for nucleocapsid containing genomic DNA, basic DNA binding protein, and capsid protein) where viral particles are assembled. At a late stage of infection, many structural polypeptides including polyhedrin, basic-DNA binding protein, capsid protein, and polyhedral envelope protein, as well as many nonstructural proteins (see section III; Blissard and Rohrmann, 1990) are transported into the nucleus where the viral envelope and polyhedral envelope are constructed. In general, polypeptides produced in eukaryotic cells are transported to target organs. These proteins have been shown to have or are considered to have specific signal sequences (e.g., specific amino acid sequences) for transportation (see review of Garoff, 1985). Baculoviral polypeptides which are transported into the nucleus are also expected to have specific sequences for transportation and

mechanisms similar to ones generally found in other eukaryotic cells. The nuclear membrane has pores which are considered wide enough to allow polypeptides up to 40 kDa to pass (see review of Silver, 1991). There are two major hypotheses for the translocation of proteins into the nucleus. One is the diffusion and trap theory, and the other is an energy dependant theory. In both cases, a recognition signal composed of specific amino acid sequences for nuclear localization seems to be important. There are many examples of signal sequences used for nuclear localization. A typical sequence which was originally found in the SV40 T antigen consists of five basic amino acids surrounded by hydrophobic amino acids (Garrof, 1985). In the BmNPV polyhedrin gene, a sequence of five basic amino acids surrounded alanine and leucine is found 31-35 amino acids from the N-terminus (Maeda et al., 1985).

Since polyhedra are visible under light microscopy, many mutants showing unique polyhedral shape and numbers, have been isolated. There are several mutants of CPV having interesting polyhedral characteristics. CPV produces polyhedra in the cytoplasm of midgut cells, however, mutants which form polyhedra in the nucleus, and mutants with cuboidal, or fibrous polyhedra have been isolated (Hukuhara and Yamaguchi, 1973). Polyhedra which localize in the nucleus are found to have four additional amino acids at the C-terminus caused by a single point mutation of the stop codon of the polyhedrin gene (Mori et al., 1989). These

mutations are speculated to be involved in the mutated polyhedral characteristics, however, this can not be proven due to the lack of an *in vitro* system for CPV replication. An *in vitro* system is necessary so that mutated viruses with insertions and deletions in the polyhedrin gene can be isolated.

In vitro systems are available for NPVs to directly study specific genes by insertion and deletion. Carstens *et al*. (1987) showed that a single point mutation causing a single amino acid substitution is responsible for the cuboidal shape of a mutant AcNPV polyhedra. The baculovirus expression vector system was used to introduce this mutation into the wild-type viral genome and confirm that the amino acid substitution is solely involved in the polyhedral structure.

To study the characteristics of polyhedra, polyhedrin genes of four mutant viruses of BmNPV and SlNPV were cloned and sequenced. The nucleotide and deduced amino acid sequences, phenotypic characteristics, and evolutionary relationships of the polyhedrin genes were examined. Gene expression of the polyhedrin gene in permissive and nonpermissive cells was analyzed. New transfer vectors with high expression characteristics were also constructed. To study the shape and nuclear localization of polyhedrin, recombinant viruses with various deletions in the polyhedrin gene were generated and isolated using direct deletions in the polyhedrin gene of BmNPV and the baculovirus expression

vector system. Amino acid sequences important for the nuclear localization of polyhedrin were identified. The relationship between amino acid sequence and shape of the polyhedra was also examined.

B. MATERIALS AND METHODS

DNA sequencing: Sequencing was performed mostly using double stranded plasmid DNA prepared on a mini-scale (Fig. III-2). The procedures employed are shown in Fig. IV-1 (denaturation and annealing of DNA), Fig. IV-2 (sequencing reaction), and in Fig. IV-3 (polyacrylamide gel preparation).

Northern blot analysis: The procedure employed for preparation of mRNA are shown in Fig. IV-4. Northern blot analysis was performed as shown in Fig. IV-5. Probes used for hybridization were labeled by random priming as shown in Fig. IV-6.

Fractionation of cell nucleus and cytoplasm: BmN cells were infected with BmNPV mutants at a moi of 5 and incubated for about 3 days at 27 C. BmN cells (about 10^7 cells) were washed twice, by centrifugation (4000 rpm, 5 min), and resuspension in PBS (pH 7.0, 0.14 M NaCl). After washing, the cell pellet was suspended gently in 110 ul of 1% Nonidet P-40, 30 mM Tris-Cl, pH 7.5, 10 mM $MgCl_2$ and incubated at room temperature for one min. Nuclei were precipitated by centrifugation (4000 rpm, 5 min) and washed with PBS. Polypeptides in the supernatant (cytoplasm) were

A. Denaturing:

1. Denaturing mix, mix the following for one sample (1 ug DNA in 8 ul H_2O):

dd H_2O	7 ul
Primer	1 ul
5N NaOH	2 ul

2. Add 10 ul of denaturing mix per microfuge tube.

3. Transfer 8 ul of DNA solution (from mini prep) into microfuge tube containing denaturing solution. Vortex to mix.

4. Denature by heating for 7 minutes at 85 C.

5. Precipitate by adding 10 ul of 5M Ammonium Acetate (3.85 g/10 ml) and 100 ul of 100% EtOH (-20 C).

6. Close lids and vortex.

8. Incubate for 10 min. at -80 C (-20 C for 60 min) (samples can be stocked at -80 C).

9. Centrifuge 12,000 rpm, 10 min. Remove supernatant.

10. Wash with 200 ul of 70% EtOH.

11. Centrifuge 12,000 rpm, 3 min. Remove supernatant.

12. Dry in vacuum. (tubes can be stocked at room temperature).

B. Annealing:

1. Add 10 ul of 1x annealing buffer to each.

2. Incubate at 37 C for 20 minutes.

3. Keep at room temperature for 10 minutes, (do not allow to cool under room temperature)

Fig. IV-1. Denaturing and annealing of plasmid DNA for sequencing

1. Add 1.3 ul each of A, C, G, and T solutions to a 96-well plate.

2. Mix the following just prior to use (on ice) and vortex gently:

DTT	21.0 ul
^{35}S-dATP	9.0 ul
Labeling mix	16.0 ul
Sequanase 2.0	2.5 ul

3. Add 2.2 ul of the above mix into the annealed DNA solutions and mix at room temperature. (add every 15 seconds for 20 samples).

4. Preheat the 96 well plate on a heat block at 37 C

5. Transfer 2.5 ul from each labeling reaction (microfuge tube from step 3) into each of the four wells containing A, C, G, and T solutions in the 96-well plate.

6. Add 3.6 ul of stop solution to each well.

7. Seal the plate with self adhesive plastic plate sealers. (Samples can be stored at -20 C.)

8. Incubate at 80 C, 2 min (air incubator) before loading onto the gel.

9. Load 1.1 ul of each sample onto a 5% acrylamide gel.

10. Run gels at 2400 V, 60 W for 6 and 12 hours.

11. Disassemble and wash gel twice with acidified-methanol solution (250 ml per wash for 10 minutes):

galcial acetic acid	50 ml
methanol	60 ml
H_2O	to 500 ml

Fig.IV-2. Sequencing of plasmid DNA using Sequanase

1. **5% acrylamide gel stock solution** for sequencing gel. Mix the following (stir vigorously) and stock at 5 C for less than a few months:

Total	
Acrylamide	48.3 g
Bis-acrylamide	1.68 g
Urea	420 g
10x TBE*	100 ml

2. Construct the necessary number of gel plates and set up.

3. Transfer the necessary amount of 5% acrylamide gel stock solution into a flask and degas for 10 min on ice. Add TEMED and 10 % ammonium persulfate as indicated below. Mix completely, but carefully not to introduce bubbles into the solution. Solution will not polymerize for about 45 min. if kept on ice.

5% stock	110 ml
TEMED	63 ul
10% ammonium-persulfate	630 ul

4. Cast gels.

5. Insert combs. Gel will polymerize in 30-60 min.

***10 x TBE buffer**

Tris	109 g
EDTA-4Na	8.3 g
Boric acid	53.4 g
H_2O	to 1 l

Fig.IV-3. Sequencing Gel

1. At the appropriate time after infection, discard the culture medium and wash attached cells gently once with ice-cold 1x PBS (5-10 ml/150 mm dish). Completely scrape-off cells with a rubber policeman, add fresh ice-cold 1x PBS, and transfer PBS and cells into a 50ml centrifuge tube. Centrifuge at 500 rpm for 5 minutes and discard supernatant. If necessary wash and centrifuge again.

2. Add Extraction Buffer, (200mM NaCl, 200mM Tris HCl, 1.5mM $MgCl_2$, 2% SDS) into the centrifuge tube (4 ml/150 mm dish) and mix well to disolve cells completely. (The solution should become viscous due to the extracted genomic DNA.) To extract mRNA from fat body of 5th instar larvae, add 5-10 ml of Extraction Buffer into a 15 ml centrifuge tube, add fat body collected from one larvae, and vortex immediately.

3. Shear genomic DNA by sonication (3-5 min) and passage (1 or 2 times) through an 18 gauge needle using a 50 ml syringe. Incubate (45 C water bath) for 1 to 2 hr.

4. Adjust the NaCl concentration of the lysate to 0.5M, by adding 0.95 mls of 5M NaCl for each 15 mls of Extraction Buffer (which is 0.2M NaCl to start).

5. Add the pre-equilibrated oligo-dT cellulose to the lysate. Rotate at room temperature for 40-60 min.

6. Centrifuge 3000 rpm at room temperature. Pour off the supernatant and resuspend the oligo-dT cellulose in 10-20 ml Binding Buffer (500mM NaCl, 10mM Tris HCl pH 7.5). Repeat this process 4 times (or until clear).

7. Add the washed oligo dT cellulose-mRNA to a disposable autoclaved column. Wash with 10 ml of binding buffer 2 or 3 times. (Allow the column to run dry or until no liquid remains on top of the resin layer.

8. Elute the poly A+ RNA with 0.5 ml of DEPC water (or 10 mM Tris-HCl) into a 1.5 ml microfuge tube containing 50 ul (0.1 vol.) of 2M sodium acetate. Add 1 ml (2 vol.) of 100% ethanol (-20 C). Precipitate at -20 C overnight.

9. Centrifuge at 14,000 rpm for 10 min. Dry briefly under vacuum, and add 100 ul of DEPC water. Take 1 ul into another microfuge tube and determine concentration by measuring absorbance at 260 - 280 nm. Stock at -80 C.

Fig. IV-4. Extraction of mRNA of infected cells

1. Soak electrophoresis tank and other equipment in contact with RNA in 3% hydrogen peroxide solution and rinse with DEPC H_2O or dd H_2O).

2. Preparation of the gel:
Microwave (or autoclave) the following:

Agarose	1 g
10x MOPS*	10 ml
DEPC H_2O	85 ml

After cooling to 60 C add:

Fromaldehyde (37%)	5.4 ml

Pour agarose into gel plate.

3. Sample Preparation:
Mix the following in microfuge tubes:

poly(A)+ RNA	1-2 ug
add DEPC H_2O up to	4.5 ul
10x MOPS*	2 ul
Formaldehyde	3.5 ul
Formamide* (deionized)	10 ul

Incubate: 65 C for 10 min

4. Add 2 ul RNA loading buffer (50% glycerol, 1mM EDTA, 0.4% bromophenol blue, 0.4% xylene cyanol) per tube and load samples onto gel. Electrophorese RNA at 150 V in 1000 ml of 1x MOPS buffer.

5. Soak the gel for 40 min in a 0.1-0.5 ug/ul ethidium bromide solution (2-10 ul of 10 mg/ml stock EtBr per 200 ml dd H_2O) with gentle rocking.

6. Mark the size standards a on clear plastic overlay using UV light box.

7. Denaturation: treat gel in 0.05 M NaOH - 0.15 M NaCl for 20 min. (0.4 g NaOH and 1.75 g NaCl in 200 ml dd H_2O)

8. Neutralization: pour off the denaturing solution and add 0.1 M Tris pH 7.5 - 0.15 M NaCl for 20 min. (1.753 g NaCl in 200 ml 0.1M Tris pH 7.5).

9. Transfer RNA to nitrocellulose or equivalent using PosiBlot Pressure Blotter.

10. Fix RNA to nitrocellulose using UV Stratalinker two times. Place filter in UV Stratalinker **RNA side up** push autocrosslink button.

11. Dry filter on 3MM filter paper (at 37 C for 10 min).

12. Bake filter at 70 C for 30 min under vacuum

Fig.IV-5. Northern Transfer and Hybridization (MOPS)

13. Wet filter in 5x SSC*. Place filter in hybridization incubator tube, add 10-15 ml of prehybridization solution*, and incubate 6 to 24 hours at 42 C. After prehybridization add 10 ul of denatured probe (see Random Primed DNA Labeling Protocol) and incubate 6 to 24 hours at 42 C.

14. Wash twice with 2x SSC*, 0.1% SDS (30 min per wash):

15. Wash once with 0.1x SSC*, 0.1% SDS (20 min per wash):

16. Air dry filter, wrap in Saran wrap (using 3MM filter paper as a backing), and expose to X-ray film overnight at -80 C.

Materials:

*__10x MOPS (pH 5.5-7.0)__

MOPS (0.2M)	41.9 g
sodium acetate (50 mM)	4.1 g
EDTA (pH 8.0) (10 mM)	20 ml 0.5M stock
add DEPC H_2O	to 1 l
mix well and autoclave	

*__Deionized Formamide__

Add approximately 50-100 g of ion exchange resin (BioRad AG 501-X8) per 1 liter of formamide.

*__20x SSC__

Dissolve 175.3 g NaCl and 88.2 g sodium citrate in 800 ml DEPC H_2O. Adjust pH to 7.0 with 3-4 drops of concentrated HCl. Adjust volume to 1 liter and autoclave.

*__Prehybridization Solution__

Deionized Formamide (50%)	50 ml
50x Denhardt's Solution* (5x)	10 ml
10% SDS (0.5%)	5 ml
2 mg/ml Salmon Sperm DNA (0.1 g/ml)	5 ml
20x SSC (5x)	25 ml
dd H_2O	5 ml

*__50x Denhardt's Solution__

Mix the following:

Ficoll	5 g
polyvinylpyrrolidone	5 g
BSA (Pentax Fraction V)	5 g
dd H_2O	to 1000 ml

Filter through disposable Nalgene filter. Store at -20 C.

Fig. IV-5. continued

1. Transfer 25-500 ng of DNA into a microfuge tube, add water up to 18 ul, and denature DNA by heating for 10 min. at 95 C and then chill on ice. Spin down. Adjust to 15 ul with distilled water.

2. Place the microfuge tube on ice and add the following (for ^{32}P dATP labeling):

2 ul, dNTP (1ul each of dCTP, dGTP, dTTP) (or mixture)
0.5 ul, random primer
2 ul, reaction mixture (10x buffer)
1 ul, ^{32}P dATP, 3000 Ci/mmole, aqueous solution
0.5 ul, Klenow enzyme

3. Incubate 30 min at 37 C

4. If necessary, terminate the reaction by adding 2 ul of 0.2M EDTA, pH 8 (or by heating 65 C for 10 min).

5. Store at -80 C before use. One ul of the labeled probe solution is sufficient per 10 ml of hybridization solution.

Before Use:

6. Denature probe by heating (in a water bath) at 95 C for 10 min. Chill on ice. Add 9 or 10 ul of the probe to the incubator tube containing the prehybridization solution and filter.

Fig. IV-6. Labeling of DNA by random priming

precipitated by addition of 20 ul of 50% TCA, incubated on ice for 30 min, and centrifuged at 12,000 rpm for 5 min. The precipitate was washed with 1/5 volume of 50% TCA. The cells and cytoplasm fractions were solubilized in 50 ul of 10 mM Tris-Cl, pH 8.0, 2% SDS, and 5% 20-mercaptoethanol, and boiled at 100 C for 3 min, and subjected to SDS-polyacrylamide gel electrophoresis as described by Laemmli (1982).

CAT assay: The plasmid, pA10CAT2, which contains the entire coding sequence of the CAT gene, a HindIII cleavage site at its 5' end and, a Bsp1286 cleavage site at the 3' end, was obtained from Dr. A. Fuse of Chiba University. The entire coding sequence the CAT gene in pA10CAT2 was cleaved with HindIII just upstream of the translational start, treated with the Klenow fragment of DNA polymerase I making blunt ends, and inserted into the AatI site of the BmNPV transfer vector, pBK283, as shown in Fig. IV-7A. The resultant recombinant transfer vector (BmCAT) was isolated by restriction endonuclease analysis from plasmids propagated on a mini-scale (Fig. III-2). BmCAT was propagated on a large-scale (Fig. III-3), and cotransfected into BmN cells with wild-type T3 viral DNA in the presence of calcium ion as described by Maeda et al. (1985). Recombinant virus carrying the CAT gene after the polyhedrin gene promoter was isolated by plaque assay by screening plaques for the absence of polyhedra as described by Maeda

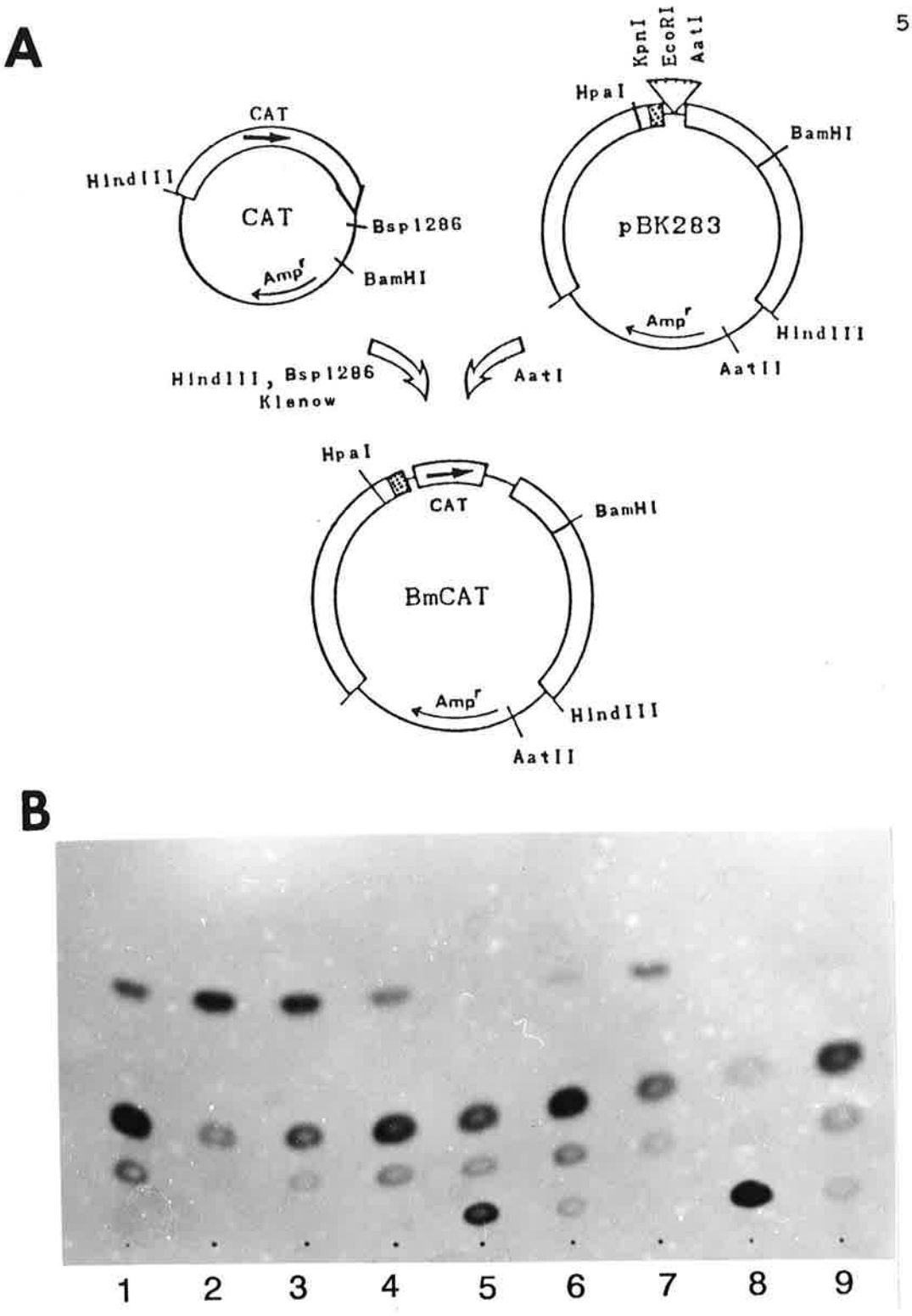


Fig. IV-7. (A) Construction of recombinant transfer vector carrying the chloramphenycol acetyltransferase gene (BmCAT). (B) Activation of the BmNPV polyhedrin promoter in non-permissive cells by coinfection with AcNPV (SlNPV OT2). Lane 1, CAT; lane 2, BmN infected with BmCAT; lane 3, BmN infected with BmCAT and SlNPV OT2; lane 4; TN-368 infected with BmCAT and SlNPV OT2; lane 5, TN-368 infected with BmCAT; lane 6, SF-21 infected with BmCAT; lane 7, SF-21 infected with BmCAT and SlNPV OT2; lane 8, CLS-79 infected with BmCAT; Lane 9, CLS-79 infected with BmCAT and SlNPV OT2.

(1989b). CAT assay was carried out as described by Guarino and Summers (1987).

C. Result and Discussion

1. Cloning and sequencing of the polyhedrin genes of the BmNPV and SlNPV isolates.

a. BmNPV BT31 with cuboidal polyhedra

BmNPV BT stock, which exhibits heat resistance and a cuboidal polyhedral shape, was originally collected in Tanegashima, Kyushu (Fig. IV-8). This virus stock was initially heterogeneous in terms of polyhedral shape, however, viruses producing only cuboidal polyhedra were selected from larvae infected *per os* and incubated at high temperature (35 C) for several passages (Watanabe, personal communication). The BmNPV BT31 isolate was plaque purified from this viral stock at high temperature.

To isolate the polyhedrin gene of the BmNPV BT 31 isolate, DNA restriction patterns of BmNPV BT31 were compared with BmNPV T3 isolate in 0.7% agarose gels as described in section III. As shown in Fig. IV-9, the restriction endonuclease patterns of BmNPV BT31 and T3 were generally similar, i.e, about 80% of the fragments had indistinguishable mobilities. However, several fragments of BT31 had no corresponding fragments. The estimated size, 130 kb, of the BT31 (by summing these digested fragments)

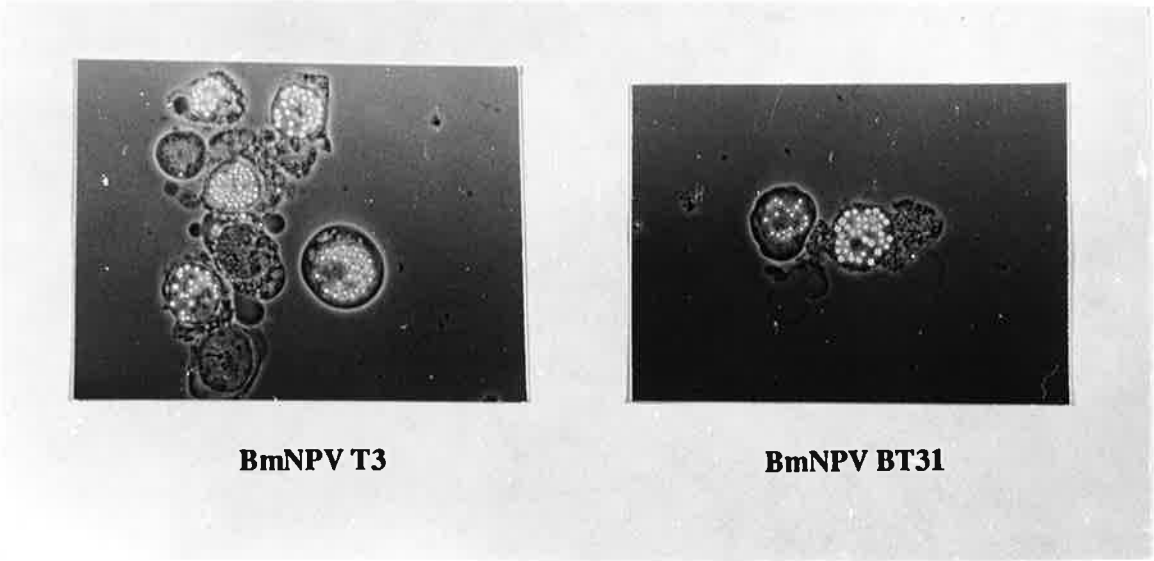


Fig IV-8. Polyhedra produced by the BmNPV T3 (polyhedral shape) and BmNPV BT31 (cuboidal shape) isolates.

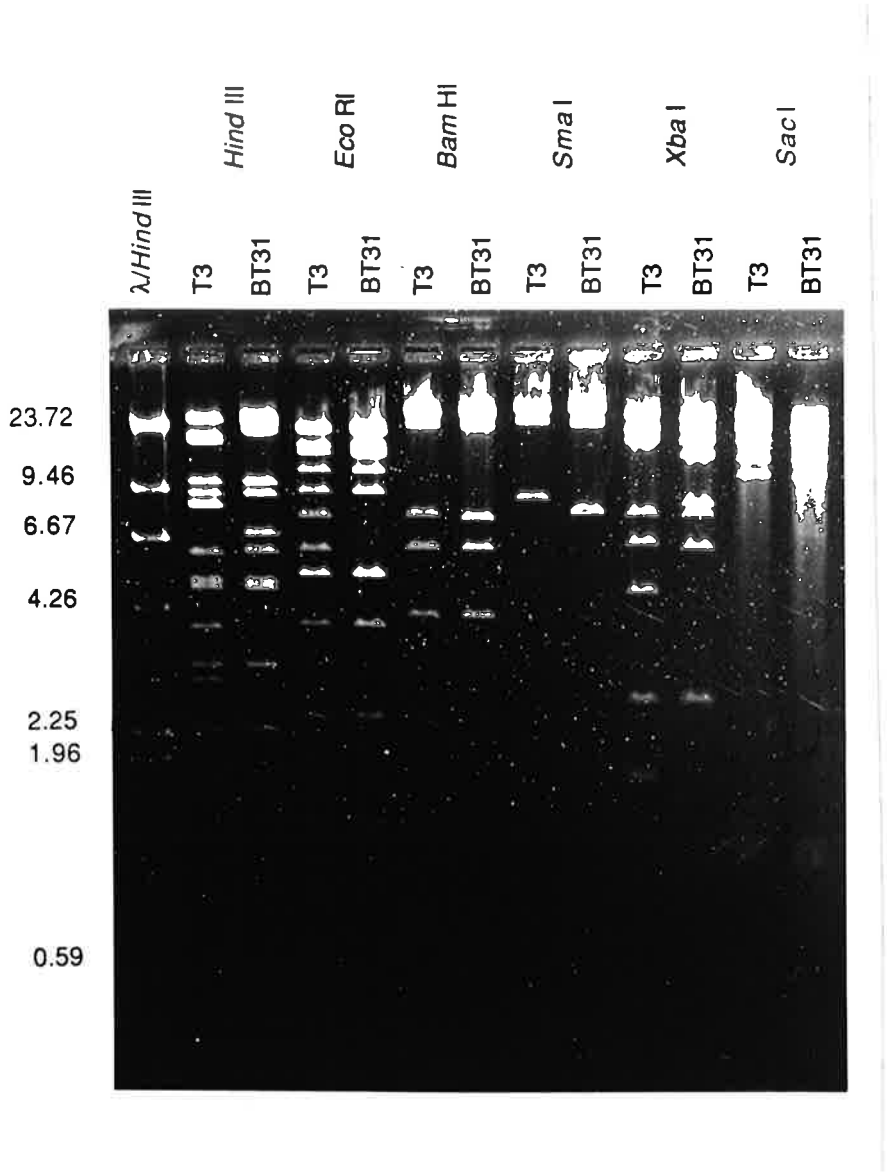


Fig. IV-9. Restriction endonuclease analysis of the BmNPV BT31 genome. Viral DNA (approximately 0.5 ug/lane) was digested with the indicated restriction endonucleases and electrophoresed on a 0.7% agarose gel. Lanes marked T3, BmNPV T3 DNA; lanes marked BT31, BmNPV BT31 DNA; lane marked Lambda/HindIII, lambda DNA cleaved with HindIII. Sizes in kb are shown at the far left.

and T3 genomes were similar. These results indicated that BT31 was a variant of T3, however, not a very closely related isolate.

To isolate the location the polyhedrin gene of the BT31 isolate, Southern blot analysis was carried out by the method described in Fig. III-4. BT31 Viral DNA was digested with EcoRI, HindIII, SmaI, XbaI, PstI, BamHI, SalI, and XhoI, and electrophoresed on a 0.7% agarose gel, then transferred and fixed to a nitrocellulose filter. When the EcoRI E fragment (10.5 kb) of the T3 isolate containing the polyhedrin gene was used as a probe for hybridization, the BT31 isolate showed hybridization patterns different from those produced by the T3 isolate (Fig. IV-10). The probe hybridized large (more than 7 kb) molecular weight fragments generated by EcoRI, SmaI, BamHI, XhoI. The hybridization pattern with HindIII digested DNA was especially different. The probe hybridized only to the 3.9 kb HindIII fragment of T3, however, it hybridized to three BT31 fragments, 0.9 kb strongly, and 3.5 and 10 kb weakly. These results indicated that the polyhedrin gene of BT31 was not located in the same HindIII fragment as T3.

Two additional probes were prepared to further characterize the polyhedrin gene location by Southern blot analysis. Probe #201 contained sequences 201 bp downstream and 2.7 kb upstream (HindIII site) of the translational start of the T3 polyhedrin gene, and probe #713 contained sequences 23 bp upstream and 4.3 kb (PstI site) downstream

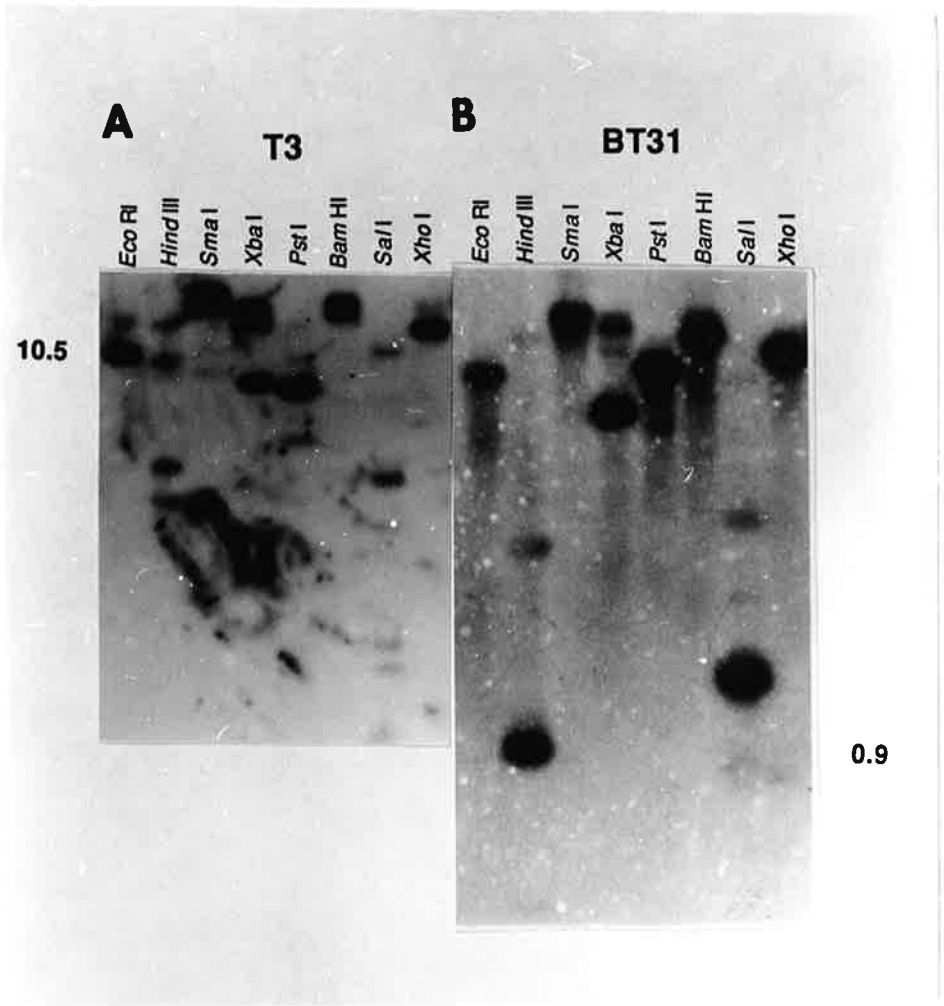


Fig. IV-10. Hybridization of BmNPV T3 viral DNA (A) and BmNPV BT31 viral DNA (B) digested with the indicated restriction endonucleases to the 10.5 kb EcoRI E fragment of BmNPV T3. Sizes in kb are indicated to the far left and right. Note that probe hybridized strongly to the 0.9 kb HindIII fragment of BT31.

of the polyhedrin stop signal of the T3 polyhedrin gene. Probe #201 hybridized to the 20 kb HindIII and 20 kb XbaI fragments of BT31. Probe #713 hybridized to the 0.9, 3.5, and 10 kb HindIII and 7.5 kb XbaI fragments of BT 31 (Fig. IV-11). These results indicated the existence of HindIII and XbaI sites in the polyhedrin gene. The XbaI site of BT31 seemed to be located at the same position as the polyhedrin gene of BmNPV T3 (Maeda et al., 1985), however, the HindIII site of BT31 seemed to be generated by nucleotide substitution.

Fragments of appropriate size for cloning and sequencing of BT31 were screened by double digestion and hybridization using the #210 and #713 probes. The enzymes used for the double digestions were EcoRI/XbaI, PstI/EcoRI, BamHI/HpaI, BamHI/PstI, XbaI/PstI, and HindiII/PstI. Hybridization patterns of the double digested BT31 DNA with the two probes are shown in Fig. IV-12. From these results, the EcoRI/XbaI 4.3 kb and the XbaI 7.5 kb fragments, which presumably contained sequences upstream and downstream of the polyhedrin gene, respectively, were cloned into pUC19 plasmid.

These two fragments were treated with Bal31 exonuclease to produce nested deletions, and cloned into pUC plasmids for sequencing as shown in Fig. IV-13. The nucleotide sequence of both directions of the coding and flanking regions were determined by dideoxy sequencing as shown in Figs. IV-14A and IV-15. When the polyhedrin gene sequence

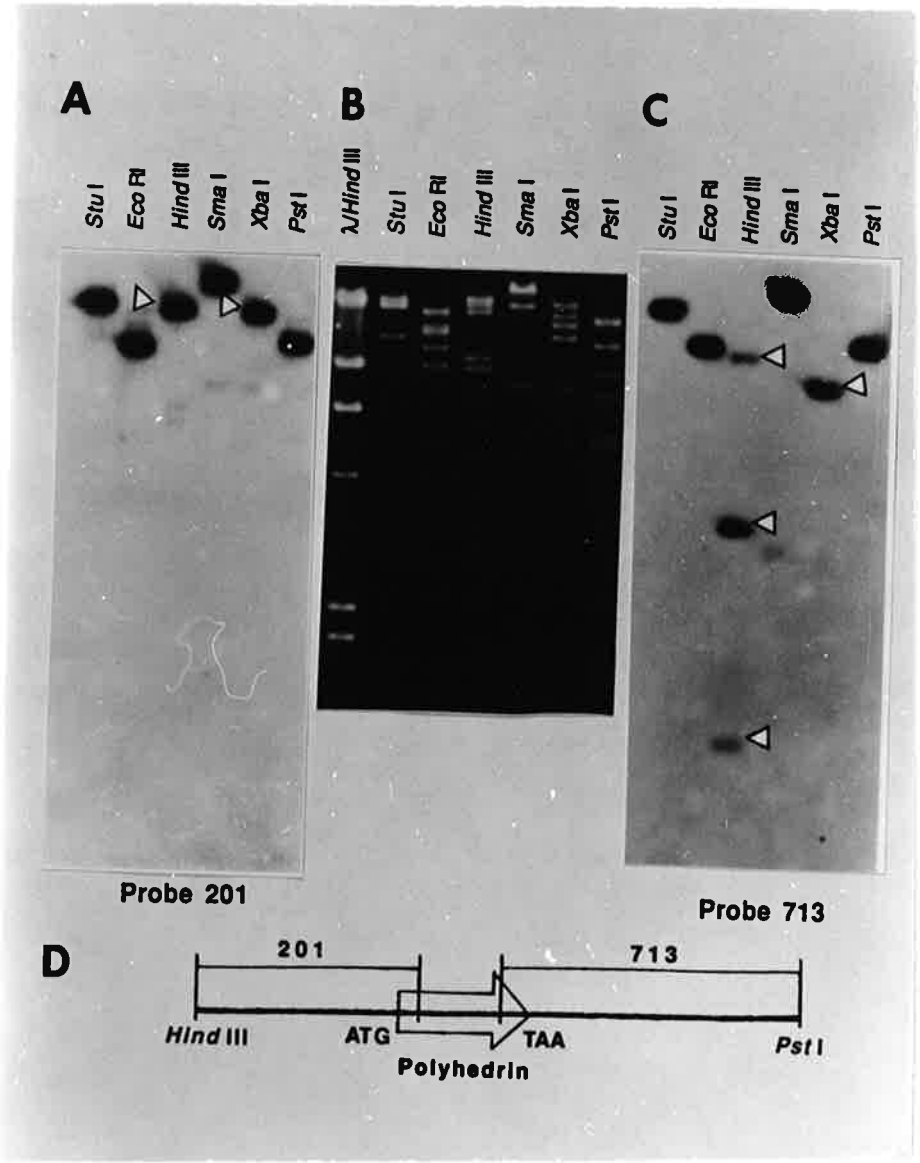


Fig. IV-11. BmNPV BT31 viral DNA digested with the indicated restriction endonucleases and electrophoresed on a 0.7% agarose gel (B), and transferred to a nitrocellulose filter and hybridized to probes #201 (A) and #713 (C). (A) Arrowheads indicate hybridization of probe #201 to the 20 kb HindIII and 20 kb XbaI fragments. (C) Arrowheads indicate hybridization of probe #713 to the 0.9, 3.5, and 10 kb HindIII and 7.5 kb XbaI fragments. (D) The polyhedrin gene and flanking regions of BmNPV T3 indicating the sequence regions of probes #201 and #713.

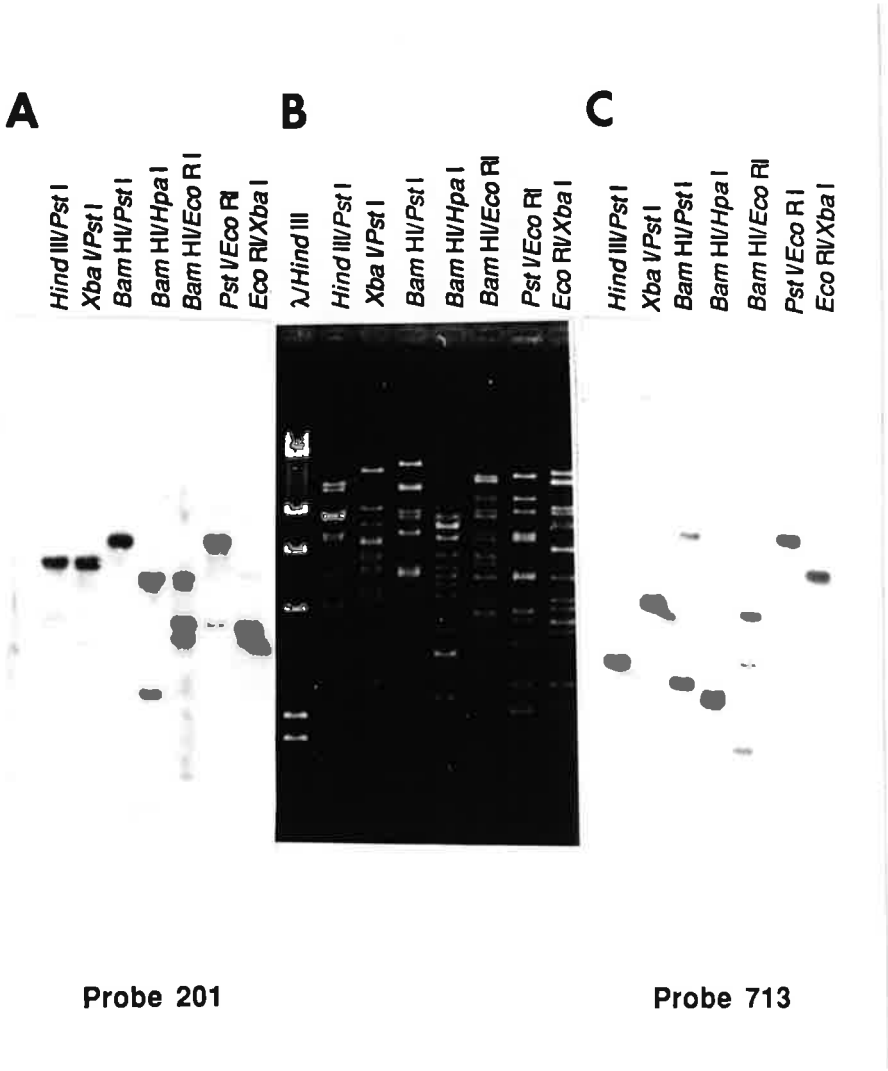


Fig. IV-12. BmNPV BT31 viral DNA double-digested with the indicated restriction endonucleases and electrophoresed on a 0.7% agarose gel (B), and transferred to a nitrocellulose filter and hybridized to probes #201 (A) and #713 (C).

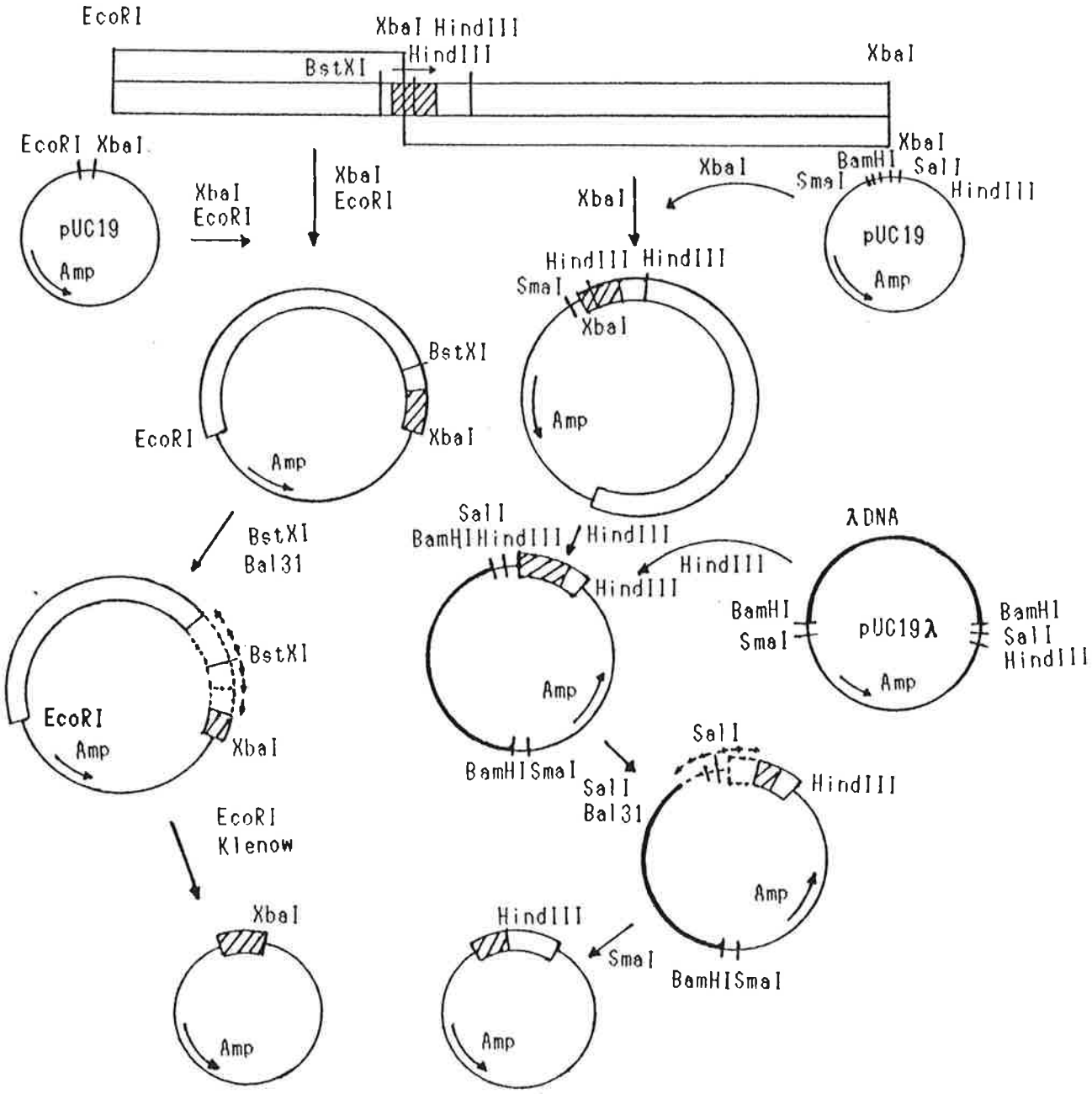


Fig. IV-13. Construction of plasmids with nested deletions of the BmNPV BT31 polyhedrin gene for sequencing.

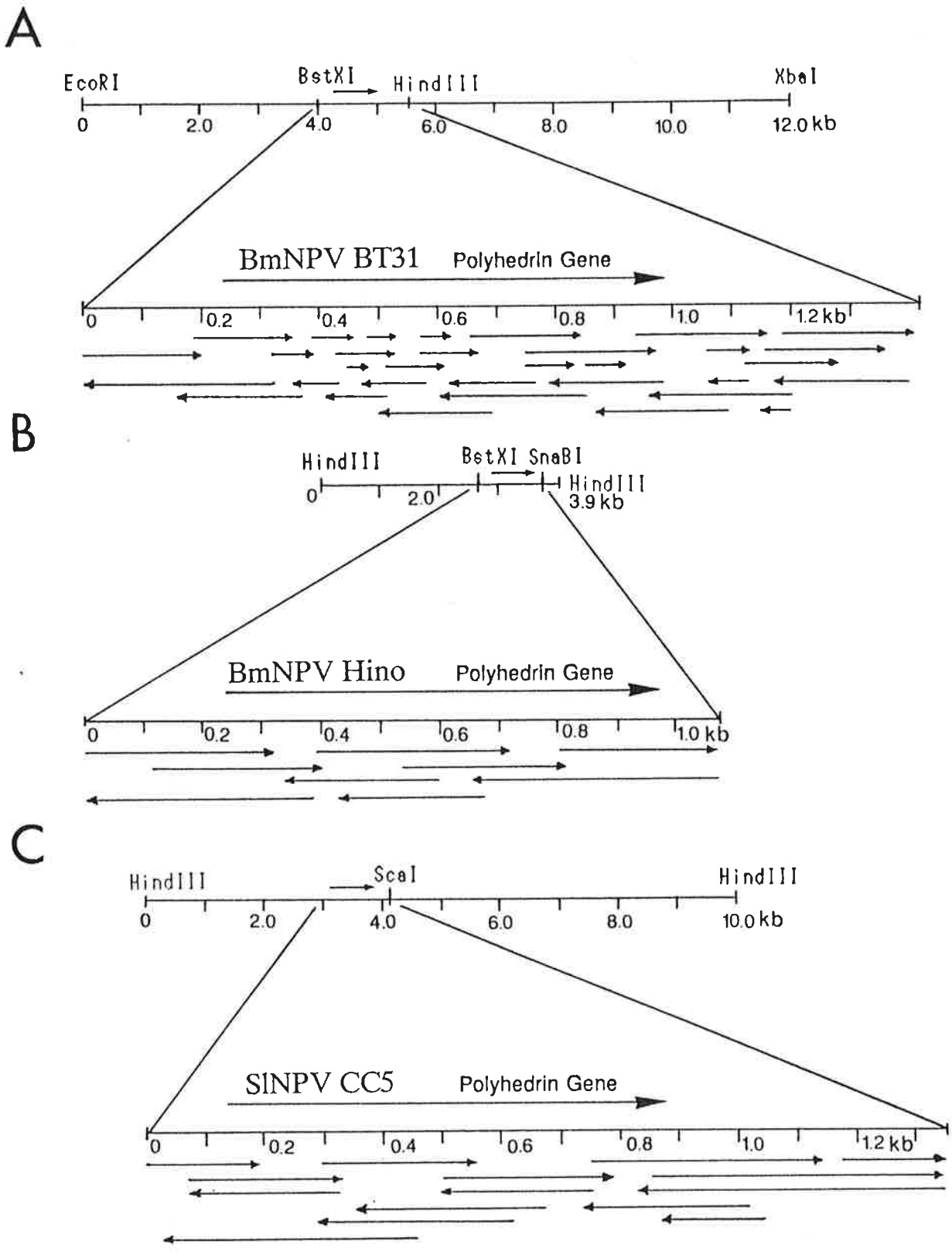


Fig. IV-14. Sequencing strategy of the polyhedrin genes of BmNPV BT31 (A), BmNPV Hino (B), and SlNPV CC5 (C).

```
Met Pro Asn Tyr Ser Tyr Thr Pro Thr Ile Gly Arg Thr Tyr Val Tyr      16
ATG CCG AAT TAT TCA TAC ACC CCC ACC ATC GGG CGT ACT TAC GTG TAC      48

Asp Asn Lys Tyr Tyr Lys Asn Leu Gly Cys Leu Ile Lys Asn Ala Lys      32
GAC AAT AAA TAT TAC AAA AAC TTG GGC TGT CTT ATC AAA AAC GCC AAG      96

Arg Lys Lys His Leu Val Glu His Glu Gln Glu Glu Lys Gln Trp Asp      48
CGC AAG AAG CAC CTA GTC GAA CAT GAA CAA GAG GAG AAG CAA TGG GAC     144
                                                              *

Leu Leu Asp Asn Tyr Met Val Ala Glu Asp Pro Phe Leu Gly Pro Gly      64
CTT CTA GAC AAC TAC ATG GTT GCC GAA GAT CCC TTT TTA GGA CCG GGC     192
                                          !

Lys Asn Gln Lys Leu Thr Leu Phe Lys Glu Ile Arg Asn Val Lys Pro      80
AAA AAC CAA AAG CTT ACC CTT TTT AAA GAA ATT CGC AAT GTG AAA CCC     240
              *                                 *

Asp Thr Met Lys Leu Ile Val Asn Trp Ser Gly Lys Glu Phe Leu Arg      96
GAT ACC ATG AAG TTA ATC GTC AAC TGG AGC GGC AAA GAG TTT TTG CGT     288

Glu Thr Trp Thr Arg Phe Val Glu Asp Ser Phe Pro Ile Val Asn Asp     112
GAA ACT TGG ACC CGT TTT GTT GAG GAC AGC TTC CCC ATT GTA AAC GAC     336

Gln Glu Val Met Asp Val Tyr Leu Val Ala Asn Leu Lys Pro Thr Arg     128
CAA GAG GTG ATG GAC GTG TAC CTC GTC GCC AAC CTC AAA CCC ACA CGC     384

Pro Asn Arg Cys Tyr Lys Phe Leu Ala Gln His Ala Leu Arg Trp Glu     144
CCC AAC AGG TGC TAC AAG TTC CTC GCT CAA CAC GCT CTT AGG TGG GAA     432

Glu Asp Tyr Val Pro His Glu Val Ile Arg Ile Val Glu Pro Ser Tyr     160
GAA GAC TAC GTG CCC CAC GAA GTA ATC AGA ATT GTG GAG CCA TCC TAC     480

Val Gly Met Asn Asn Glu Tyr Arg Ile Ser Leu Ala Lys Lys Gly Gly     176
GTG GGC ATG AAC AAC GAA TAC AGA ATT AGT CTG GCT AAA AAG GGC GGC     528

Gly Cys Pro Ile Met Asn Ile His Ser Glu Tyr Thr Asn Ser Phe Glu     192
GGC TGC CCA ATC ATG AAC ATC CAC AGC GAG TAC ACC AAC TCG TTC GAG     576

Ser Phe Val Asn Arg Val Ile Trp Glu Asn Phe Tyr Lys Pro Ile Val     208
TCG TTT GTG AAC CGC GTC ATA TGG GAG AAC TTC TAC AAA CCC ATC GTT     624

Tyr Ile Gly Thr Asp Ser Ala Glu Glu Glu Glu Ile Leu Ile Glu Val     224
TAC ATC GGC ACA GAC TCT GCC GAA GAA GAG GAA ATC CTA ATT GAG GTT     672

Ser Leu Val Phe Lys Ile Lys Glu Phe Ala Pro Asp Ala Pro Leu Phe     240
TCT CTC GTT TTC AAA ATA AAG GAG TTT GCA CCA GAC GCG CCT CTG TTC     720

Thr Gly Pro Ala Tyr ***                                             246
ACT GGT CCG GCG TAT TAA                                             738
```

Fig. IV-15. Sequence of the polyhedrin gene of BmNPV BT31. An asterisk (*) below a nucleotide sequence indicates a nucleotide substitution. An exclamation mark (!) above an amino acid sequence indicates an amino acid substitution. Three asterisks (***) indicate a stop codon.

of BT31 was compared to the polyhedrin gene of BmNPV T3 (Maeda et al., 1985; Iatrou et al., 1985), three nucleotide substitutions were observed: T at nucleotide 144 was changed to C; A at nucleotide 204 to G, and G at nucleotide 230 to A. One of these substitutions (nucleotide 230) resulted in an amino acid change from serine to aspartic acid. This substitution seemed to be responsible for the change in the shape of the inclusion body from polyhedral to cuboidal. It has been shown that an AcNPV mutant producing cuboidal polyhedra has a single amino substitution from proline to leucine at amino acid 58 (Carstens et al., 1986) of the AcNPV polyhedrin gene. Proline can cause "bends" in the amino acid chain due to unique characteristics of its side chain. However, the amino acid substitution found in the BT31 isolate was considered not to cause drastic changes in the structure of the polyhedra based on similar characteristics of the substituted amino acid.

Recombinant techniques were used to determine whether or not a single amino acid substitution was responsible for the change in polyhedral shape. The polyhedrin gene was exchanged between the T3 and BT31 isolates using a recombinant virus, BMIFN (carrying the human interferon alpha gene in place of the BmNPV T3 polyhedrin gene, Horiuchi et al., 1987). The complete polyhedrin gene sequence of BT31 was first constructed in a plasmid by ligation of the 4.3 kb EcoRI/XbaI and 7.5 kb XbaI fragments containing upstream and downstream sequences of the

polyhedrin gene and coding region, respectively. A plasmid containing the continuous complete polyhedrin gene was isolated by DNA fragment analysis of plasmid DNA propagated on a mini-scale. This plasmid was cotransfected with BmIFN DNA, and a polyhedron positive virus (BT-T3) was plaque purified.

To confirm that flanking sequences of the polyhedrin gene were not involved in the determination of polyhedral shape, the BT31 polyhedrin gene along with a fragment between the XbaI (at nucleotide 147 of the BT31 polyhedrin gene) and BamHI (1.3 kb downstream of the BT31 polyhedrin gene) sites was transferred to the transfer vector p89BX40 (Sekine et al., 1988) which contains about 3 kb of the flanking sequence derived from BmNPV T3. A recombinant virus (BTX40) was isolated after cotransfection of this plasmid and BmIFN DNA by the method described above.

When the BT-T3 and BTX40 viruses were infected into BmN cells, both produced cuboidal polyhedra (Fig. IV-16). These results indicate that the cuboidal shape of polyhedra is caused by substitution of one amino acid. When the polyhedrin gene of T3 which produces polyhedral shaped inclusion bodies, was transferred to the BT31 genome and examined, cuboidal polyhedra were not produced confirming that the BT31 polyhedrin gene is responsible for the cuboidal shape.

BmNPV BT31 DNA was extracted and cotransfected with recombinant plasmid having a deletion within the polyhedrin

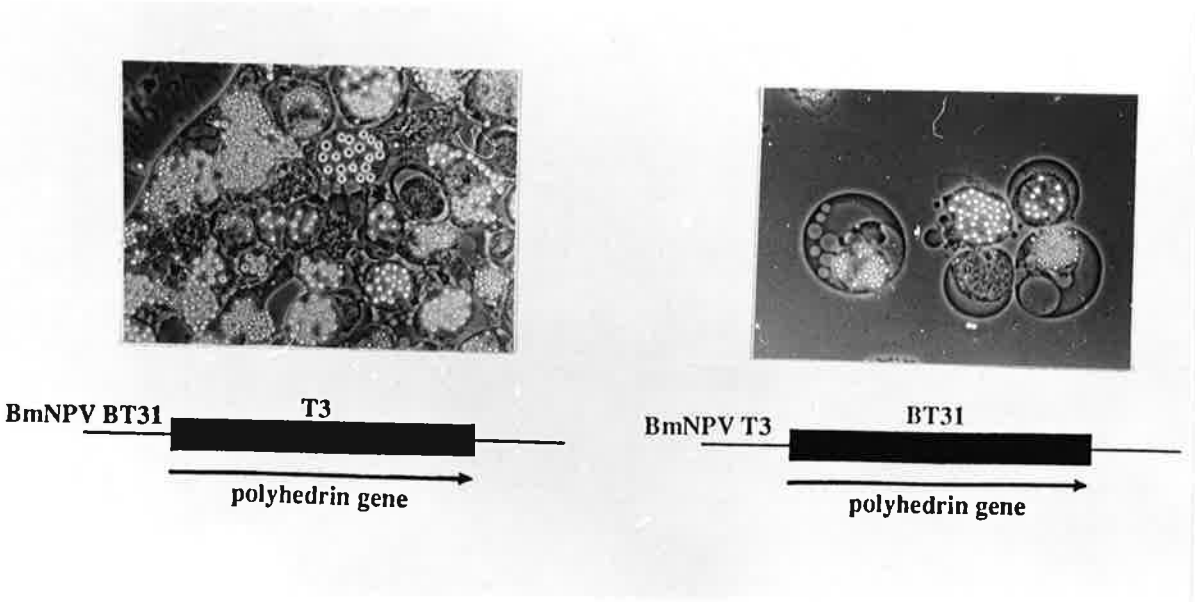


Fig. IV-16. Polyhedral production in BmN cells infected with a recombinant BmNPV BT31 carrying the polyhedrin gene of BmNPV T3 (left), and infected with a recombinant BmNPV T3 carrying the polyhedrin gene of BmNPV BT31 (right).

coding sequence (see Section V) into BmN cells. A recombinant virus of BT31 lacking polyhedral production was isolated after cotransfection of BmNPV BT31 DNA with a recombinant plasmid lacking the polyhedrin gene and plaque purification. Viral DNA of this recombinant BT31 was prepared and cotransfected with the EcoRI E (10.5 kb) (see Section III) fragment of BmNPV T3, which contains the polyhedrin gene. A polyhedra producing virus was plaque purified and reinfected into BmN. When these viruses were examined they produced polyhedral shaped inclusion bodies identical in appearance to those of T3. These results indicated that a single amino acid substitution was responsible for the cuboidal shape of the polyhedra.

b. The formaldehyde resistant BmNPV Hino isolate

The BmNPV Hino isolate was plaque purified and propagated in 5th instar silkworm larvae. Polyhedral inclusion bodies were purified from larvae by various centrifugation processes and viral particles were purified from inclusion bodies in a 10-40% sucrose gradient after alkaline treatment of polyhedra. Viral DNA was extracted from viral particles and analyzed by agarose gel electrophoresis after digestion with several endonucleases. Fig. IV-17 shows electrophoretic patterns of cleaved Hino DNA. Most cleaved fragments of the Hino isolate were indistinguishable from those of the T3 isolate. Southern blot analysis was conducted using the HindIII K fragment

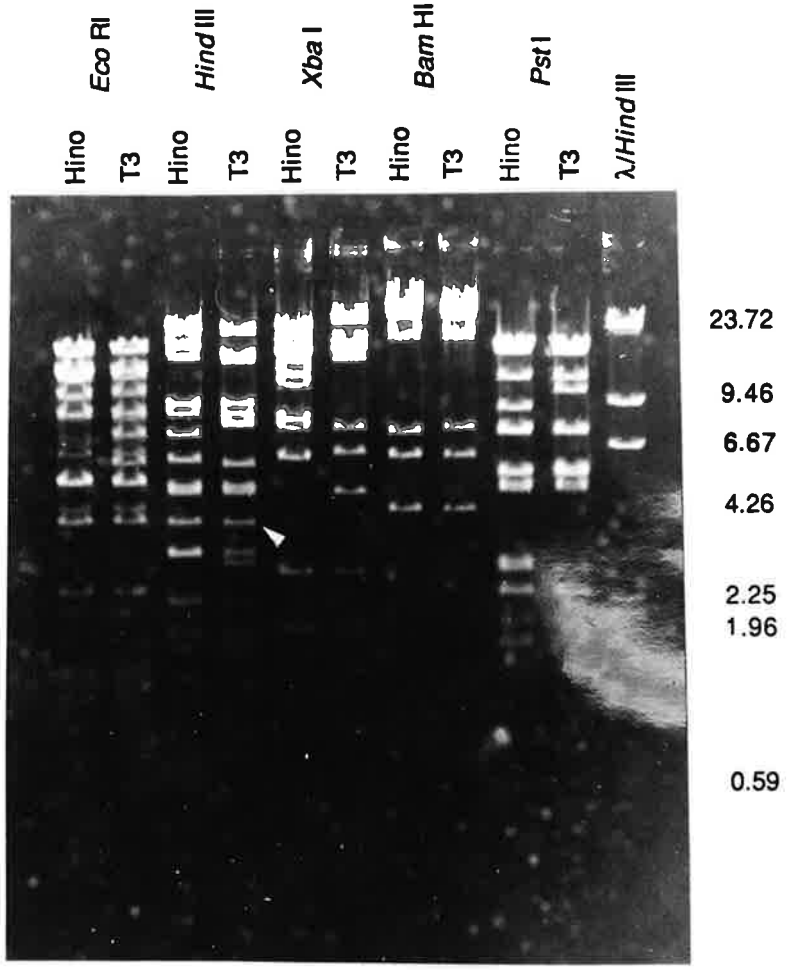


Fig. IV-17. Restriction endonuclease analysis of the BmNPV Hino isolate digested with the indicated restriction endonucleases and electrophoresed on a 0.7% agarose gel. Lanes marked Hino, BmNPV Hino DNA; lanes marked T3, BmNPV T3 DNA. Lane marked lambda/HindIII, lambda DNA cleaved with HindIII. Sizes in kb are shown at the far right. Arrowhead indicates the 3.9 kb HindIII fragment of BmNPV T3 containing the polyhedrin gene.

(see Section III) of BmNPV T3 containing the polyhedrin gene as a probe. The probe strongly hybridized to the 3.9 kb HindIII and the 10.5 kb EcoRI fragments of the Hino isolate (Fig. IV-18). The 3.9 kb HindIII fragment was assumed to have the polyhedrin gene and was cloned into the pTZ18R plasmid for treatment with Bal31 exonuclease to produce fragments with nested deletions. Fragments with nested deletions were recloned into pTZ18R (Figs. IV-14B, IV-19).

The entire nucleotide sequence of the polyhedrin gene region of the Hino isolate was determined by dideoxy sequencing of the nested fragments. No amino acid nor nucleotide substitutions were found in the coding sequence. Amino acid substitutions causing changes in structural conformation resulting in low fixation efficiency of polyhedra to formaldehyde have been hypothesized as the mechanism of resistance of polyhedral resistance to formaldehyde (Nitta and Watanabe, 1984). However, identical amino acid sequences of polyhedrin between the T3 and Hino isolate indicated that the polyhedrin gene alone was not involved in formaldehyde resistance. Since polyhedra are surrounded by an envelope (polyhedral envelope) and its associated protein (polyhedra envelope protein), it can be hypothesized that a change of the polyhedra envelope protein imparts increased resistance to formaldehyde in the Hino isolate. Experiments to examine this hypothesis will be performed by isolation of the polyhedron envelope protein

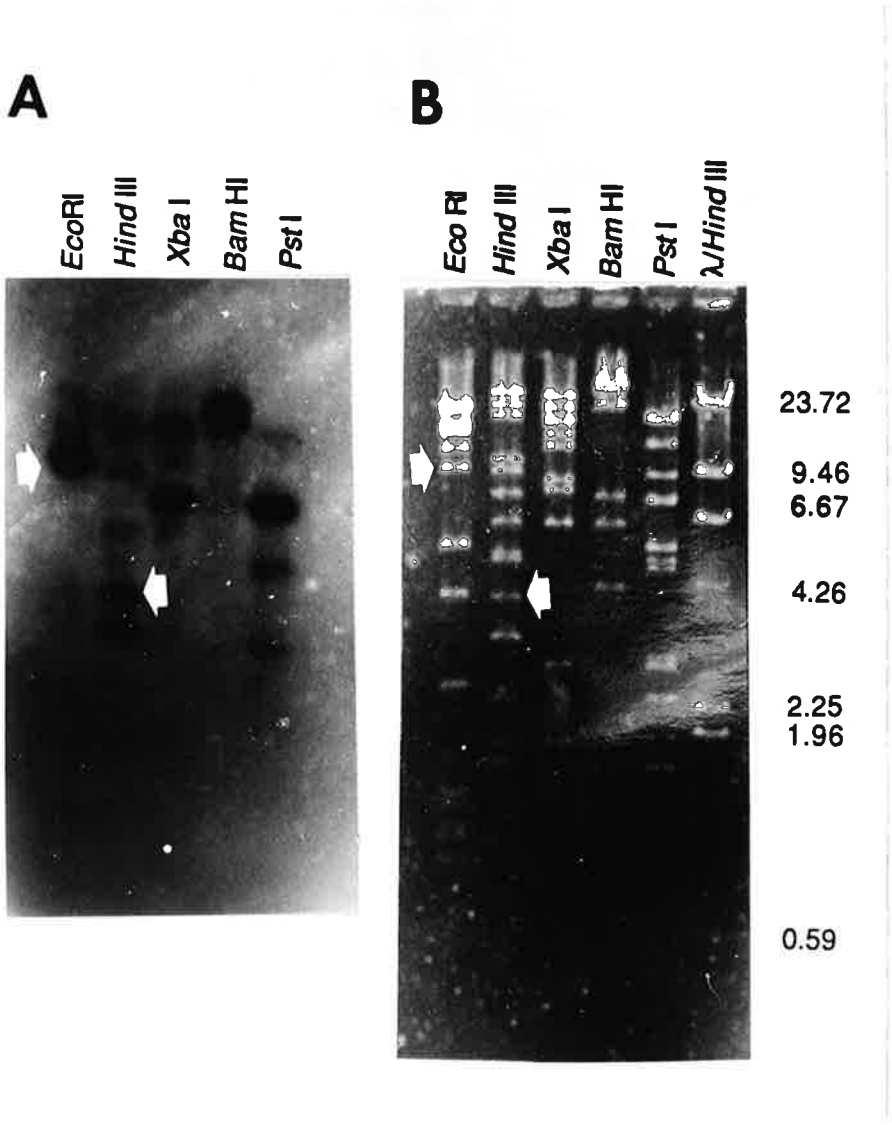


Fig. IV-18. Restriction endonuclease analysis of BmNPV Hino digested with the indicated restriction endonucleases and electrophoresed on a 0.7% agarose gel (B) and transferred to a nitrocellulose filter and hybridized with the 10.5 kb EcoRI E fragment of BmNPV T3 (A). Arrows indicate hybridization of the probe to the 10.5 kb EcoRI and 3.9 kb HindIII fragments of BmNPV Hino.

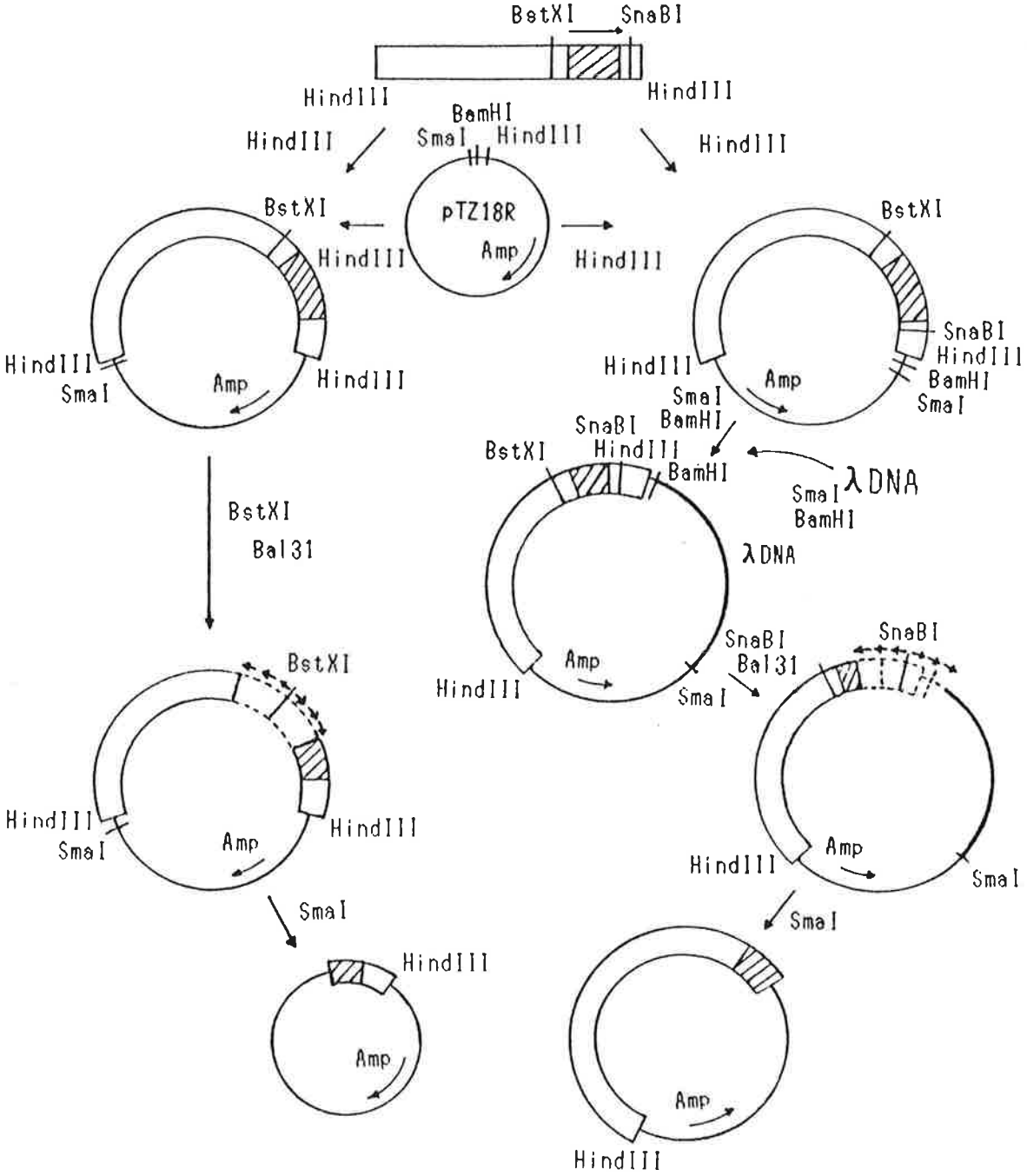


Fig. IV-19. Construction of plasmids with nested deletions of the BmNPV Hino polyhedrin gene for sequencing.

and exchange experiments using the baculovirus expression vector system.

c. SlNPV OT2

The SlNPV OT2 isolate is classified as an AcNPV variant (AcNPV OT2) as described in Section II. The nucleotide sequence of its polyhedrin gene was determined by dideoxy nucleotide sequencing and compared to the AcNPV polyhedrin gene to study the relatedness of OT2 and AcNPV. The polyhedrin sequence of the OT2 isolate was completely identical to that of AcNPV. In addition, other DNA characteristics of this virus have shown it to be an AcNPV variant and nearly identical to Galleria mellonera NPV.

d. SlNPV CC5

It has been shown that the SlNPV CC5 isolate belongs to the type C group of SlNPV by DNA hybridization and in vitro replication studies (Maeda et al., 1990). DNA hybridization experiments show low DNA homology of CC5 to OT2 (AcNPV variant) (Maeda et al., 1990), i.e. the 1.8 kb HpaI-HindIII DNA fragment of BmNPV T3 (Maeda et al., 1985) containing the polyhedrin gene only hybridizes to limited regions of CC5 (data not shown). To analyze the polyhedrin gene of SlNPV CC5, a 10 kb HindIII fragment, which presumably contained the polyhedrin gene, was cloned into the HindIII site of pUC19. Using this plasmid further hybridization analysis

narrowed the location of the polyhedrin gene to near the ScaI site located in the center of the HindIII fragment (see Fig. 14C). After cleavage at the SacI site, two fragments possessing the polyhedrin gene area at its ends were subcloned into pTZ18Rlambda and the polyhedrin gene area was sequenced as described in Fig. IV-20.

When sequence around the ScaI region was examined, one open reading frame 738 nucleotides in length (including the translational start) encoding a 245 amino acid long protein very similar to polyhedrin was found. Fig. IV-21 shows the entire sequence of the SlNPV CC5 polyhedrin gene and its deduced amino acid sequence along with amino acid sequences of the polyhedrin genes of other NPVs and the granulin gene of two GVs for comparison. The nucleotide sequence and amino acid sequences were generally similar to those of previously reported polyhedrin genes (Fig. IV-22; Rohrmann, 1986). When the polyhedrin amino acid sequence of CC5 was compared in detail to other polyhedrin amino acid sequences from AcNPV, G. mellonella NPV, O. pseudotsugata NPV, L. dispar NPV, and BmNPV (see section III), no identical sequences were found, and O. pseudotsugata NPV showed the highest homology (89%)(Rohrmann, 1986). These results suggests that there are some genetic relationships between OpNPV and the CC5 isolate of SlNPV group C. It is of interest that they are characterized as completely different viruses due to differences in their hosts.

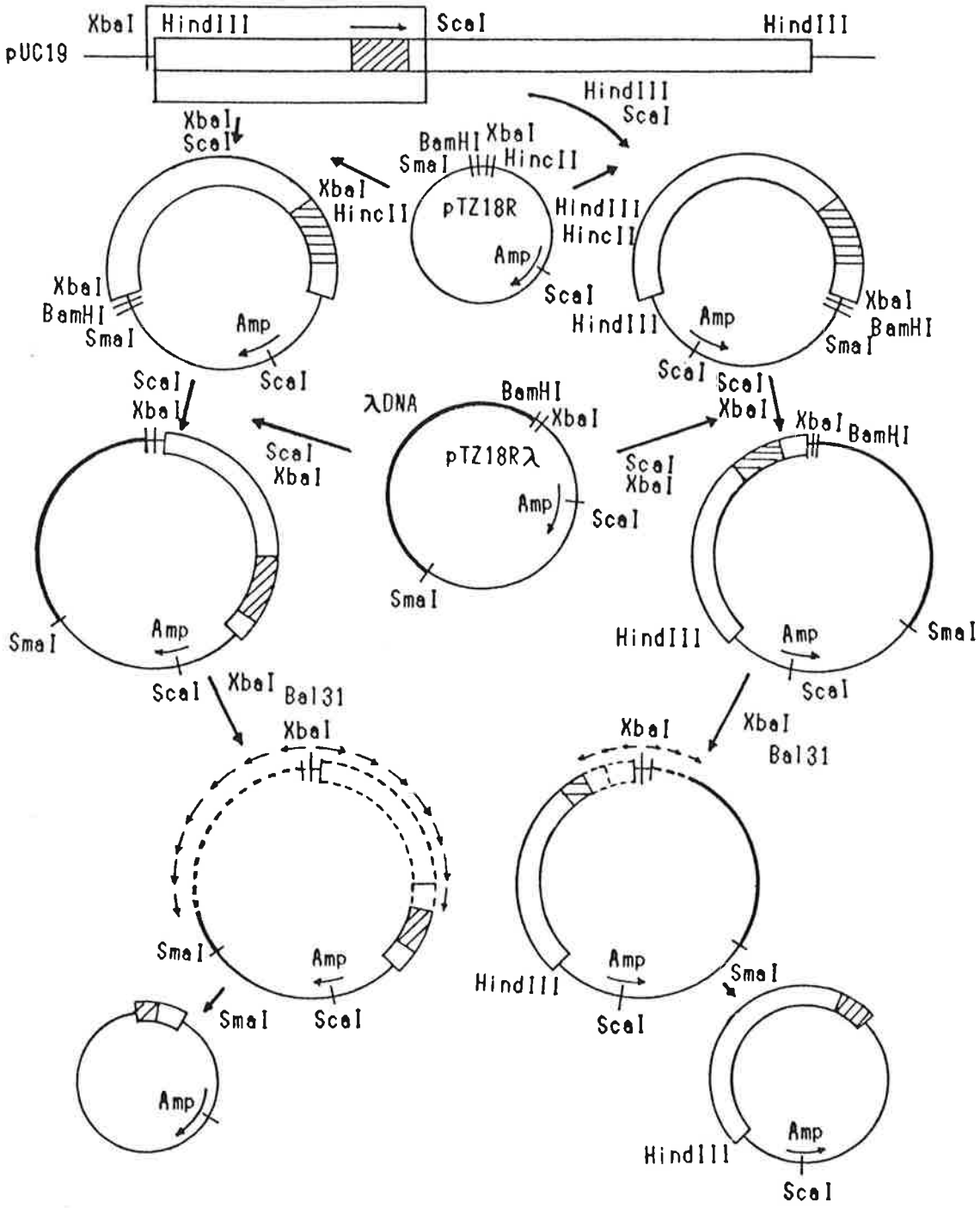


Fig. IV-20. Construction of plasmids with nested deletions of the SlNPV CC5 polyhedrin gene for sequencing.

```
Met Tyr Thr Arg Tyr Ser Tyr Asn Pro Ser Leu Gly Arg Thr Tyr Val      16
ATG TAT ACT CGC TAC AGC TAC AAC CCG TCT CTG GGT CGC ACC TAC GTG      48

Tyr Asp Asn Lys Phe Tyr Lys Asn Leu Gly Ser Val Ile Lys Asn Ala      32
TAC GAC AAC AAG TTT TAC AAA AAT CTC GGT TCA GTG ATC AAA AAC GCC      96

Lys Arg Lys Glu His Leu Val Gln His Glu Ile Glu Glu Arg Thr Leu      48
AAA CGG AAA GAA CAT TTG GTG CAA CAT GAA ATC GAA GAG AGG ACA TTG     144

Asp Pro Leu Glu Arg Tyr Val Val Ala Glu Asp Pro Phe Leu Gly Pro      64
GAT CCC CTG GAA AGG TAC GTC GTC GCC GAG GAC CCC TTC CTC GGA CCC     192

Gly Lys Asn Gln Lys Leu Thr Leu Phe Lys Glu Ile Arg Leu Val Lys      80
GGC AAG AAC CAA AAG TTG ACC CTC TTT AAA GAA ATC CGT CTC GTC AAG     240

Pro Asp Thr Met Lys Leu Val Val Asn Trp Ser Gly Lys Glu Phe Leu      96
CCC GAC ACG ATG AAA CTG GTC GTC AAC TGG AGC GGC AAA GAG TTT CTC     288

Arg Glu Thr Trp Thr Arg Phe Met Glu Asp Ser Phe Pro Ile Val Asn     112
AGG GAA ACT TGG ACC CGT TTC ATG GAA GAC AGC TTC CCC ATC GTC AAC     336

Asp Gln Glu Ile Met Asp Val Tyr Leu Val Ile Asn Met Arg Pro Thr     128
GAT CAA GAA ATA ATG GAC GTC TAT CTC GTC ATC AAC ATG AGA CCC ACT     384

Arg Pro Asn Arg Cys Tyr Arg Phe Leu Ala Gln His Ala Leu Arg Cys     144
AGA CCC AAC CGA TGC TAC AGA TTC TTG GCG CAA CAC GCT CTC CGT TGC     432

Asp Pro Asp Tyr Val Pro His Glu Val Ile Arg Ile Val Glu Pro Val     160
GAT CCT GAC TAC GTT CCT CAC GAA GTG ATC CGC ATC GTC GAG CCC GTA     480

Tyr Val Gly Ser Asn Asn Glu Tyr Arg Ile Ser Leu Ala Lys Lys Gly     176
TAC GTC GGC TCC AAC AAC GAG TAC CGC ATC AGC TTG GCC AAA AAG GGC     528

Gly Gly Cys Pro Val Met Asn Leu His Ser Glu Tyr Thr His Ser Phe     192
GGC GGC TGC CCG GTC ATG AAT CTA CAC TCT GAG TAC ACG CAC TCT TTC     576

Glu Glu Phe Ile Asn Arg Val Ile Trp Glu Asn Phe Tyr Lys Pro Ile     208
GAA GAA TTC ATC AAC CGC GTC ATC TGG GAA AAC TTC TAC AAA CCC ATC     624

Val Tyr Val Gly Thr Asp Ser Gly Glu Glu Glu Glu Ile Leu Leu Glu     224
GTG TAC GTA GGA ACC GAC TCC GGT GAA GAG GAA GAA ATC CTC CTC GAA     672

Val Ala Leu Val Phe Lys Ile Lys Glu Phe Ala Pro Asp Ala Pro Leu     240
GTG GCC CTC GTT TTC AAG ATC AAA GAG TTT GCG CCC GAC GCG CCT CTC     720

Tyr Asn Gly Pro Ala Tyr ***                                        247
TAC AAC GGA CCC GCA TAT TGA                                        741
```

Fig. IV-21. Sequence of the polyhedrin gene of SlNPV CC5.

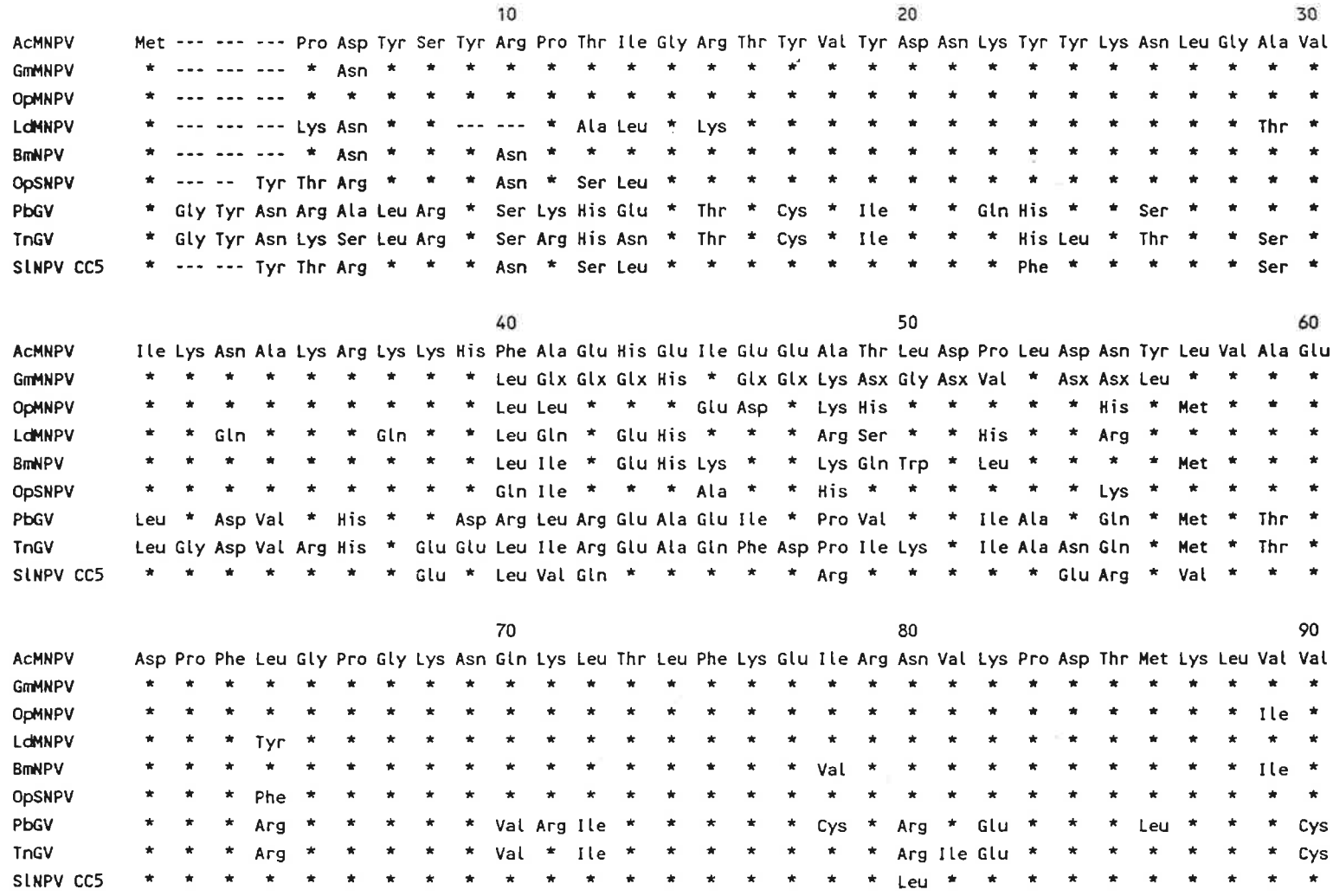
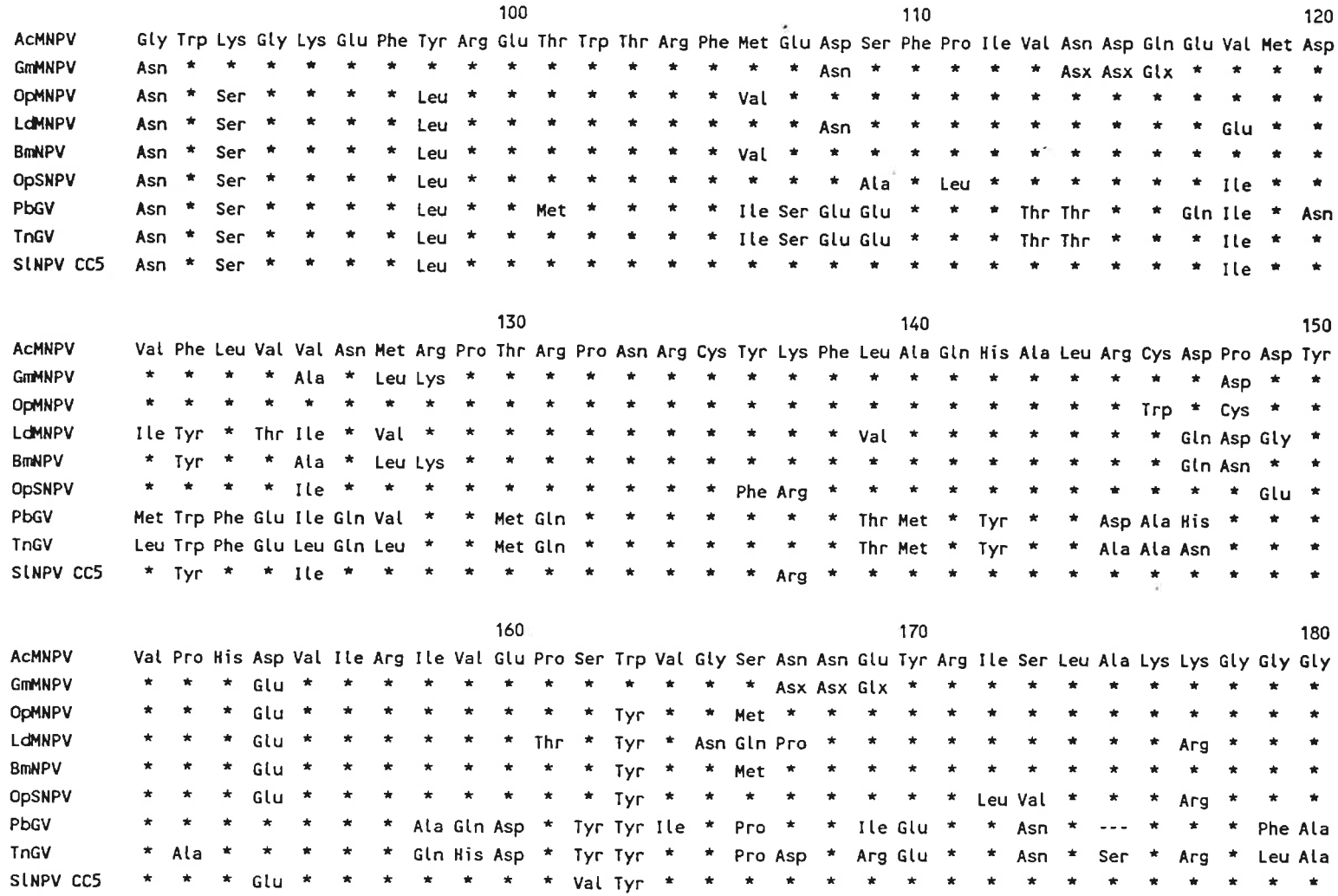


Fig. IV-22. Comparison of the amino acid sequence of the AcMNPV polyhedrin gene with the polyhedrin and granulin amino acid sequences of other NPVs and GVs. Dashes (---) indicate deletions. An asterisks (*) indicates an identical amino acid.



	100																110												120
AcMNPV	Gly	Trp	Lys	Gly	Lys	Glu	Phe	Tyr	Arg	Glu	Thr	Trp	Thr	Arg	Phe	Met	Glu	Asp	Ser	Phe	Pro	Ile	Val	Asn	Asp	Gln	Glu	Val	Met	Asp
GmMNPV	Asn	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Asn	*	*	*	*	*	Asx	Asx	Glx	*	*	*	*
OpMNPV	Asn	*	Ser	*	*	*	*	Leu	*	*	*	*	*	*	*	Val	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LdMNPV	Asn	*	Ser	*	*	*	*	Leu	*	*	*	*	*	*	*	*	*	Asn	*	*	*	*	*	*	*	*	*	Glu	*	*
BmNPV	Asn	*	Ser	*	*	*	*	Leu	*	*	*	*	*	*	*	Val	*	*	*	*	*	*	*	*	*	*	*	*	*	*
OpSNPV	Asn	*	Ser	*	*	*	*	Leu	*	*	*	*	*	*	*	*	*	*	Ala	*	Leu	*	*	*	*	*	*	Ile	*	*
PbGV	Asn	*	Ser	*	*	*	*	Leu	*	*	Met	*	*	*	*	Ile	Ser	Glu	Glu	*	*	*	Thr	Thr	*	*	Gln	Ile	*	Asn
TnGV	Asn	*	Ser	*	*	*	*	Leu	*	*	*	*	*	*	*	Ile	Ser	Glu	Glu	*	*	*	Thr	Thr	*	*	*	Ile	*	*
SlNPV CC5	Asn	*	Ser	*	*	*	*	Leu	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Ile	*	*

	130																140												150
AcMNPV	Val	Phe	Leu	Val	Val	Asn	Met	Arg	Pro	Thr	Arg	Pro	Asn	Arg	Cys	Tyr	Lys	Phe	Leu	Ala	Gln	His	Ala	Leu	Arg	Cys	Asp	Pro	Asp	Tyr
GmMNPV	*	*	*	*	Ala	*	Leu	Lys	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Asp	*	*
OpMNPV	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Trp	*	Cys	*	*
LdMNPV	Ile	Tyr	*	Thr	Ile	*	Val	*	*	*	*	*	*	*	*	*	*	*	Val	*	*	*	*	*	*	*	Gln	Asp	Gly	*
BmNPV	*	Tyr	*	*	Ala	*	Leu	Lys	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Gln	Asn	*	*
OpSNPV	*	*	*	*	Ile	*	*	*	*	*	*	*	*	*	*	Phe	Arg	*	*	*	*	*	*	*	*	*	*	*	Glu	*
PbGV	Met	Trp	Phe	Glu	Ile	Gln	Val	*	*	Met	Gln	*	*	*	*	*	*	*	Thr	Met	*	Tyr	*	*	Asp	Ala	His	*	*	*
TnGV	Leu	Trp	Phe	Glu	Leu	Gln	Leu	*	*	Met	Gln	*	*	*	*	*	*	*	Thr	Met	*	Tyr	*	*	Ala	Ala	Asn	*	*	*
SlNPV CC5	*	Tyr	*	*	Ile	*	*	*	*	*	*	*	*	*	*	*	Arg	*	*	*	*	*	*	*	*	*	*	*	*	*

	160																170												180
AcMNPV	Val	Pro	His	Asp	Val	Ile	Arg	Ile	Val	Glu	Pro	Ser	Trp	Val	Gly	Ser	Asn	Asn	Glu	Tyr	Arg	Ile	Ser	Leu	Ala	Lys	Lys	Gly	Gly	Gly
GmMNPV	*	*	*	Glu	*	*	*	*	*	*	*	*	*	*	*	*	Asx	Asx	Glx	*	*	*	*	*	*	*	*	*	*	*
OpMNPV	*	*	*	Glu	*	*	*	*	*	*	*	*	Tyr	*	*	Met	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LdMNPV	*	*	*	Glu	*	*	*	*	*	*	Thr	*	Tyr	*	Asn	Gln	Pro	*	*	*	*	*	*	*	*	*	Arg	*	*	*
BmNPV	*	*	*	Glu	*	*	*	*	*	*	*	*	Tyr	*	*	Met	*	*	*	*	*	*	*	*	*	*	*	*	*	*
OpSNPV	*	*	*	Glu	*	*	*	*	*	*	*	*	Tyr	*	*	*	*	*	*	*	*	Leu	Val	*	*	*	Arg	*	*	*
PbGV	*	*	*	*	*	*	*	Ala	Gln	Asp	*	Tyr	Tyr	Ile	*	Pro	*	*	Ile	Glu	*	*	Asn	*	---	*	*	*	Phe	Ala
TnGV	*	Ala	*	*	*	*	*	Gln	His	Asp	*	Tyr	Tyr	*	*	Pro	Asp	*	Arg	Glu	*	*	Asn	*	Ser	*	Arg	*	Leu	Ala
SlNPV CC5	*	*	*	Glu	*	*	*	*	*	*	*	Val	Tyr	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Fig. V-22. continued

	190																200										210
AcMNPV	Cys	Pro	Ile	Met	Asn	Leu	His	Ser	Glu	Tyr	Thr	Asn	Ser	Phe	Glu	Gln	Phe	Ile	Asp	Arg	Val	Ile	Trp	Glu	Asn	Phe	Tyr	Lys	Pro	Ile
GmMNPV	*	*	*	*	*	*	---	---	---	---	*	Asx	*	*	Glx	Glx	*	*	*	*	*	*	*	Glx	Asx	*	*	*	*	*
OpMNPV	*	*	*	*	*	Ile	*	Ala	*	*	*	*	*	*	*	Ser	*	Val	Asn	*	*	*	*	*	*	*	*	*	*	*
LdMNPV	*	*	*	Arg	*	*	*	*	Ala	*	*	His	*	*	*	Thr	*	---	---	---	---	---	---	Asp	*	*	*	*	*	*
BmNPV	*	*	*	*	*	Ile	*	*	*	*	*	*	*	*	*	Ser	*	Val	Asn	*	*	*	*	*	*	*	*	*	*	*
OpSNPV	*	*	Val	*	*	*	Gln	Ala	*	*	*	*	*	*	*	Glu	*	*	Glu	*	*	His	*	*	*	*	*	*	*	*
PbGV	Phe	*	Leu	*	Cys	*	Gln	*	Val	*	Asn	Asp	Asn	*	*	Thr	*	Phe	Glu	Asp	*	Leu	*	Pro	Tyr	*	His	Arg	*	Leu
TnGV	Phe	*	Leu	Thr	Cys	*	Gln	*	Ile	*	Asn	Asp	Asn	*	*	Glu	*	Phe	*	Gln	*	Leu	*	Pro	Tyr	*	His	Arg	*	Leu
SlNPV CC5	*	*	Val	*	*	*	*	*	*	*	*	His	*	*	*	Glu	*	*	Asn	*	*	*	*	*	*	*	*	*	*	*

	220																230										240
AcMNPV	Val	Tyr	Ile	Gly	Thr	Asp	Ser	Ala	Glu	Glu	Glu	Glu	Ile	Leu	Leu	Glu	Val	Ser	Leu	Val	Phe	Lys	Val	Lys	Glu	Phe	Ala	Pro	Asp	Ala
GmMNPV	*	*	*	*	*	*	Thr	Ser	*	*	*	Glx	Leu	Ile	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
OpMNPV	*	*	*	*	*	*	*	Ser	*	*	*	*	*	*	Ile	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
LdMNPV	*	*	Val	*	*	Thr	Ala	Ser	*	*	*	Gln	*	*	*	*	*	*	*	*	*	*	Ile	*	*	*	*	*	*	*
BmNPV	*	*	*	*	*	*	Ala	Ser	*	*	*	Gln	*	*	Ile	*	*	*	*	*	*	*	Ile	*	*	*	*	*	*	*
OpSNPV	*	*	Val	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	Ile	*	*	*	*	*	*	*
PbGV	*	*	*	*	*	Thr	*	Ser	*	Thr	*	*	*	*	*	*	*	*	Phe	Leu	*	*	Ile	*	*	*	*	*	*	Val
TnGV	*	*	Val	*	*	Thr	*	*	*	Ile	*	*	Val	Met	Ile	*	*	Ala	*	Leu	*	*	Ile	*	*	*	*	*	*	Val
SlNPV CC5	*	*	Val	*	*	*	*	*	Gly	*	*	*	*	*	*	*	*	Ala	*	*	*	*	Ile	*	*	*	*	*	*	*

	248
AcMNPV	Pro Leu Phe Thr Gly Pro Ala Tyr
GmMNPV	* * * * * * * *
OpMNPV	* * * * * * * *
LdMNPV	* * * Gln * * * *
BmNPV	* * * * * * * *
OpSNPV	* * Tyr Ser * * * *
PbGV	* * Tyr * * * * *
TnGV	* * * * * * * *
SlNPV CC5	* * Tyr Asn * * * *

Fig. V-22. continued

2. Expression of the polyhedrin gene.

a. Expression of the polyhedrin gene in vivo

The polyhedrin genes of baculoviruses are expressed at a late stage of infection, however, few studies have been conducted on the expression of the polyhedrin and other baculoviral genes. To study gene expression of the polyhedrin gene of BmNPV, especially in vivo, the recombinant virus, BmDH5 (Maeda, 1989c) was used. After BmDH5 was injected into the body cavity, fat body was collected at appropriate times after infection and poly $(A)^+$ RNA was extracted by the method described in Fig. IV-4. Poly $(A)^+$ RNA was electrophoresed on a 1.5% agarose gel, and transferred to a nitrocellulose filter. The filter was subjected to Northern blot analysis using a plasmid probe containing a synthetic diuretic hormone adjacent to the polyhedrin promoter.

Three days post infection, an 800 nt band was detected which intensified at 4 days post infection (Fig. IV-23. Four days post infection, two additional minor bands were identified with mobilities of 1800 and 2400 nt. No detectable band was observed in fat body collected from uninfected control larvae. Transcription of the polyhedrin gene initiates 53 bp upstream of the polyhedrin translational start (Horiuchi et al., 1987) and the poly (A)+ signal is located 435 bp downstream of the polyhedrin stop codon (Iatrou et al., 1985). Since the size of the inserted gene (DH) was 95 bp, the expected size of the

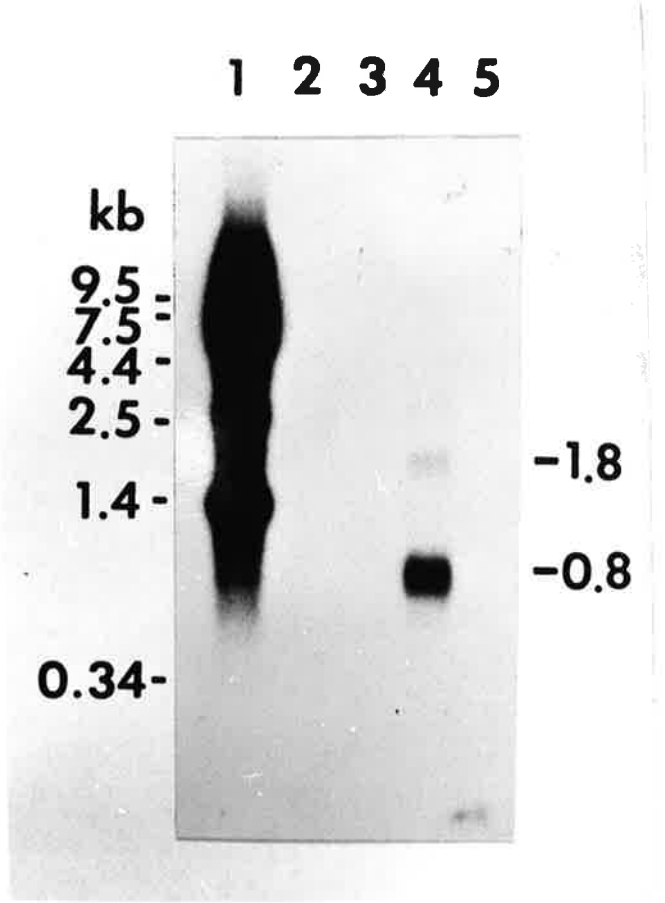


Fig. IV-23. Northern blot analysis of the expression of the synthetic diuretic hormone gene in the fat bodies of infected silkworm larvae using a pTZ plasmid containing the entire DH5 gene as a probe. Lane 1, RNA marker ladder (BRL); lane 2, mock-infected fat body; lane 3, BmDH5-infected fat body (2 days p.i.); lane 4, BmDH5-infected fat body (3 days p.i.); lane 5, control wildtype infected fat body (3 days p.i.).

polyhedrin promoter derived mRNA was about 700 nt, which was almost identical to the 800 nt observed in the Northern blot analysis. These results indicated that the major band originated from a polyhedrin promoter derived mRNA. The polyhedrin promoter was not activated at an early stage of infection and strongly activated at a late stage of infection *in vivo*, in fat body cells. The additional minor bands 1800 and 2400 nt in size were considered as extended mRNAs which were read through the poly $(A)^+$ signals, since the polyhedrin mRNA of other baculoviruses show similar read through products in infected cells (Friessen and Miller, 1985; Lubbert and Doerfler, 1984). This type of read through was not found to occur in fat bodies infected with BmNPV.

b. Activation of the polyhedrin gene promoter by heterologous viral infection.

During our experiments, a recombinant virus carrying two different polyhedrin gene promoters was isolated as follows. BmIFN viral DNA was cotransfected with plasmid DNA containing the 7.2 kb *Eco*RI fragment containing the polyhedrin gene of the SlNPV (AcNPV) OT2 isolate, and a recombinant virus producing polyhedra was isolated by plaque purification. Since the upstream region of the polyhedrin gene of the OT2 isolate had no apparent homology between AcNPV and BmNPV (Maeda *et al.*, 1985), the location of the polyhedrin gene of this recombinant virus was analyzed by

DNA restriction enzyme analysis. The polyhedrin gene unexpectedly inserted at a position, which had relatively high (80-90%) DNA homology, but was different from the original polyhedrin gene location. The precise region determined by double digestions of several enzymes is shown in Fig. IV-24. These results indicated that 1) a different polyhedrin gene promoter of SlNPV (AcNPV) OT2 is activated in BmN cells, 2) the polyhedral production of AcNPV occurred in BmN cells by activation of BmNPV gene products, and 3) the location of the polyhedrin gene promoter in the viral genome is not important for expression.

Cat assay was used to further study the activation of various polyhedrin promoters in different cells. CAT assays are commonly employed to quantitate gene expression levels due to its high sensitivity and lack of endogenous activity in eukaryotic cells. BmNPV carrying the CAT gene was initially inserted after the polyhedrin gene promoter (designated BmCAT). BmCAT was able to express CAT at a high levels so that expression was detected with a high degree of sensitivity. BmCAT was propagated and infected into permissive cells with and without the SlNPV (OT2) isolate.

c. Construction of new transfer vectors of BmNPV

Since the 5' non-translated region of the polyhedrin gene is important for high level expression of foreign genes, new transfer vectors were constructed containing these sequences. Plasmids containing the polyhedrin gene

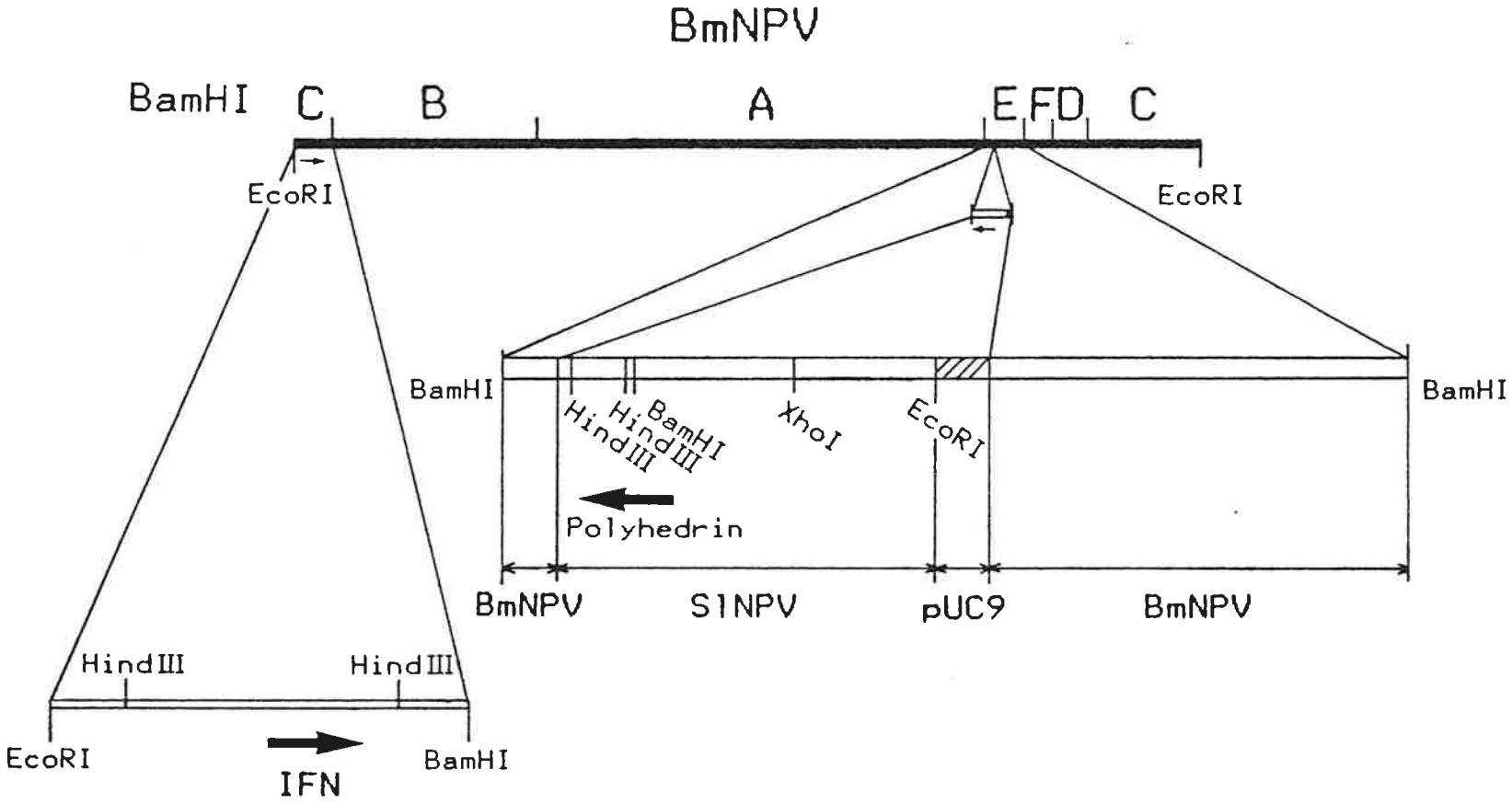


Fig. IV-24. Insertion by homologous recombination of the SlNPV OT2 polyhedrin gene and flanking sequences into the BamHI E fragment of a recombinant BmNPV carrying the human alpha-interferon (IFN) gene. The expected site of insertion was the BamHI C fragment. The recombinant virus from this construct expressed both IFN and SlNPV polyhedrin (albeit at lower levels) indicating that both BmNPV and SlNPV promoters were active.

and the 5′ flanking sequence were used as a base for the construction of new vectors. After Bal31 digestion, a plasmid containing nearly the entire 5′ flanking sequence (3 bp from the translational start of the polyhedrin gene) was isolated (pol-left) as shown in Fig. IV-25 (details of the construction are described in section IV-3).

Using a similar strategy, a DNA fragment containing the 3′ region including 5 bp upstream (in the coding sequence) from the translational stop of the polyhedrin gene was isolated (pol-right) (Fig. IV-25).

The HindIII-EcoRI fragment (HindIII is located to the left end in Fig. IV-25 and EcoRI is connected to a pUC linker) containing the 5′ region of the polyhedrin gene of pol-left and the EcoRI-PstI fragment (PstI is located to the right end in Fig. IV-25) of pol-right were ligated using the EcoRI and ScaI sites. When the resulting plasmid, pBE274, was tested for foreign gene expression levels using the human interferon alpha gene, a 5-fold increase in expression activity was obtained in infected culture medium and in the hemolymph of infected larvae. Expression of a synthetic insect-specific scorpion toxin gene also exhibited high level expression in cell culture and in larvae. Similar vectors, pBE284, pBK273, pBK283 were constructed by ligation with the 3′ region and a part of the linker region of the plasmid pBM030 plasmid (Horiuchi et al., 1987) (Fig. IV-26).

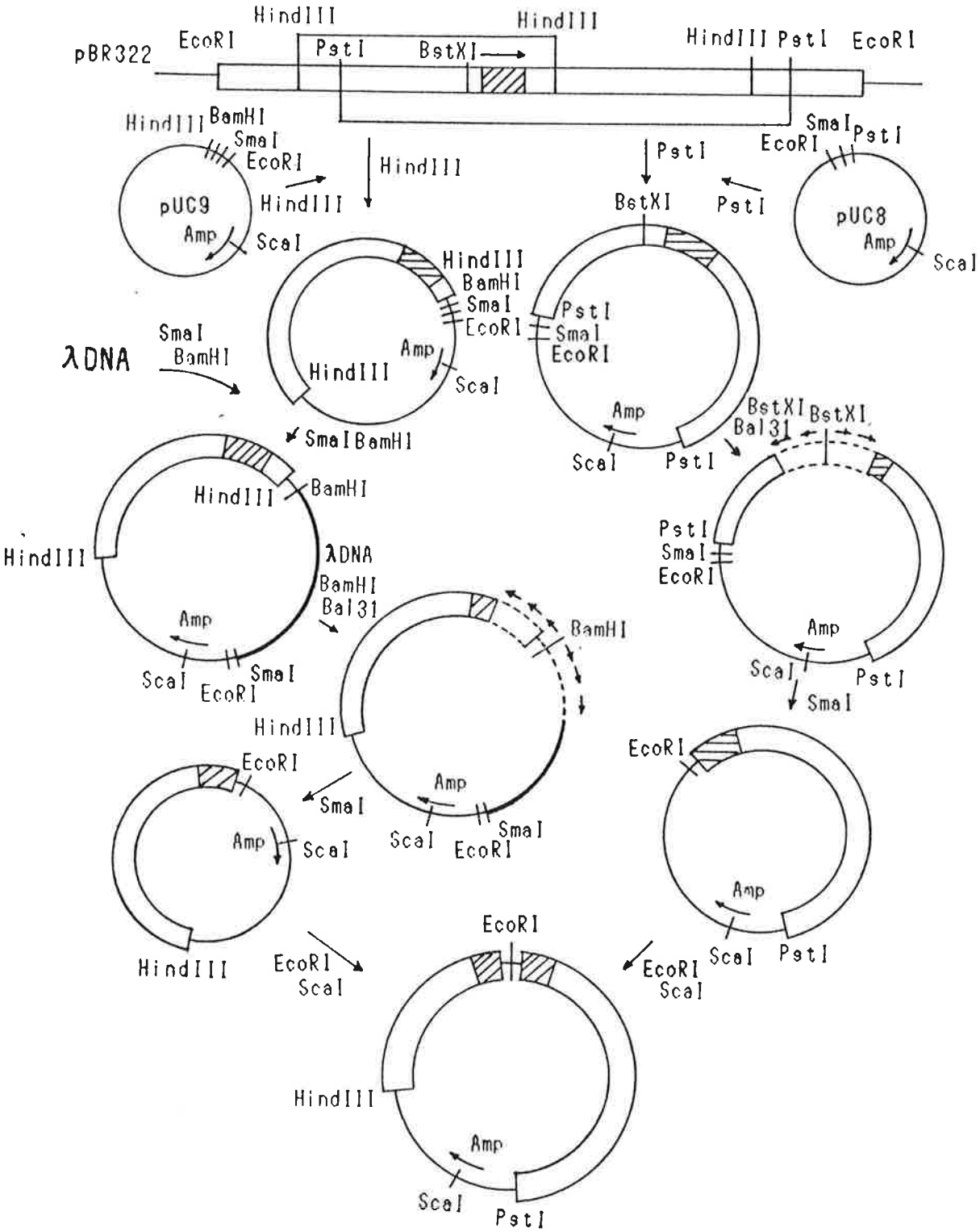
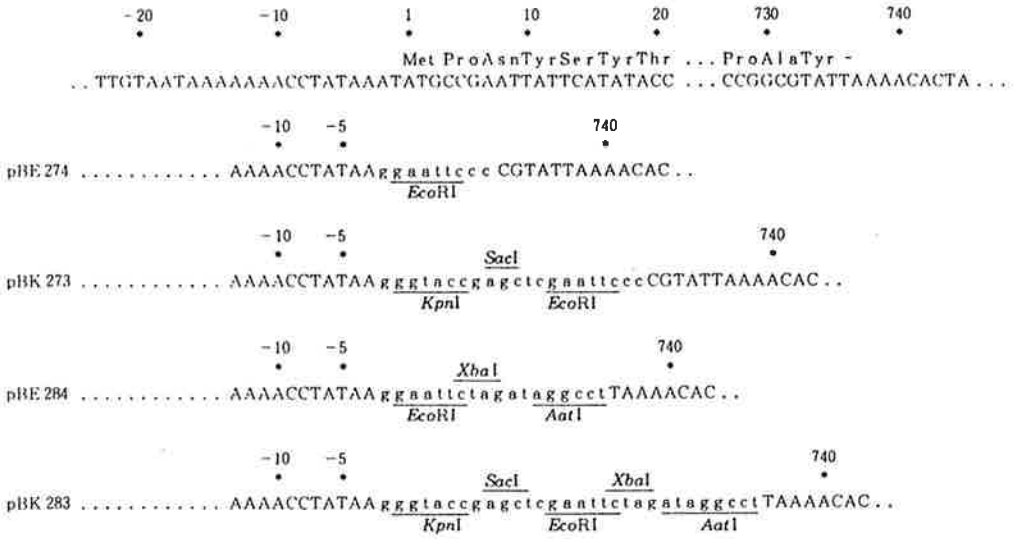


Fig. IV-25. Construction of plasmids containing BmNPV T3 polyhedrin genes with deletions in the coding region.



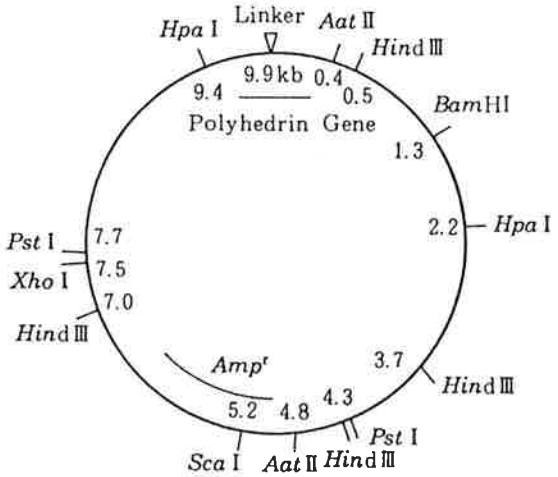


Fig. IV-26. Physical map of the pBE and pBK series transfer vectors (bottom). Multiple cloning sites of the pBE and pBK transfer as indicated (top).

3. Structure and localization of the polyhedron crystal

Deletions were made in the polyhedrin gene by Bal31 exonuclease digestion for both directions (Fig. IV-25). These deleted fragments were sequence by dideoxy nucleotide sequencing and the exact sites where deletions occurred were determined. Totally, mutant with 115 and 42 different deletions from the 5′ and 3′ ends of the coding sequence, respectively, were isolated. The deleted ends were ligated to the SmaI site adjacent to EcoRI, which was derived from the pUC8/9 plasmid. The two fragments were connected at the EcoRI sites after calculation and adjustment in the reading frame, so as to generate a continuous amino acid sequence with the addition of a SmaI-EcoRI (5′-GGGAATCCC-3′) linker sequence. By this procedure, all mutants had an additional 9 bases which expressed 3 amino acids.

The plasmid constructs with various deletions were transfected with wild-type BmNPV DNA into BmN cells and the cotransfection cell culture medium was subjected to plaque assay using BmN. When the plaques were screened, about 1% lacked polyhedral production, which was similar to the percentage of recombinant plaques found when a foreign gene is inserted into the polyhedrin gene of BmNPV (Maeda, 1989b). Since no plaques lacking polyhedral production were observed in a control transfection of the wild-type BmNPV DNA only, plaques without polyhedra were considered to be mutants with the expected deletions in the polyhedrin gene.

Several isolates without polyhedral production were plaque purified and propagated in BmN cells. BmN cells infected with these isolates did not shown any polyhedral production at a high (50) moi of infection. When the infected cells were collected and subjected to SDS-polyacrylamide gel electrophoresis, a major band was detected by Coomassie blue staining at a position expected for polyhedrin with the induced deletions (Fig. IV-27).

When the #79-105 recombinant plasmid (containing a deletion between 79 and 105 nt from the translational start) was cotransfected with BmNPV T3 DNA, recombinant virus with cuboidal shaped polyhedra were produced and isolated by plaque assay. When the plasmid #79-105 was cotransfected with BmIFN DNA, recombinant viruses producing cuboidal shaped polyhedra were also produced and isolated. Plaque purified #79-105 recombinants showing the cuboidal phenotype also produced polyhedra 3-5 times larger than wild-type BmNPV (Fig. IV-28). Furthermore, only a few polyhedra per nucleus were produced, probably because of the large size of each polyhedra. #124-141 and #99-112 recombinants also produced a similar type of polyhedra in the nucleus (Fig. IV-28).

When a deletion at the C-terminus (26 bp) was introduced into the polyhedrin gene, polyhedra which lacked straight edges and non-uniform in shape were produced (Fig. IV-28). The number of polyhedra per cell nucleus was similar to those of wild-type BmNPV T3.

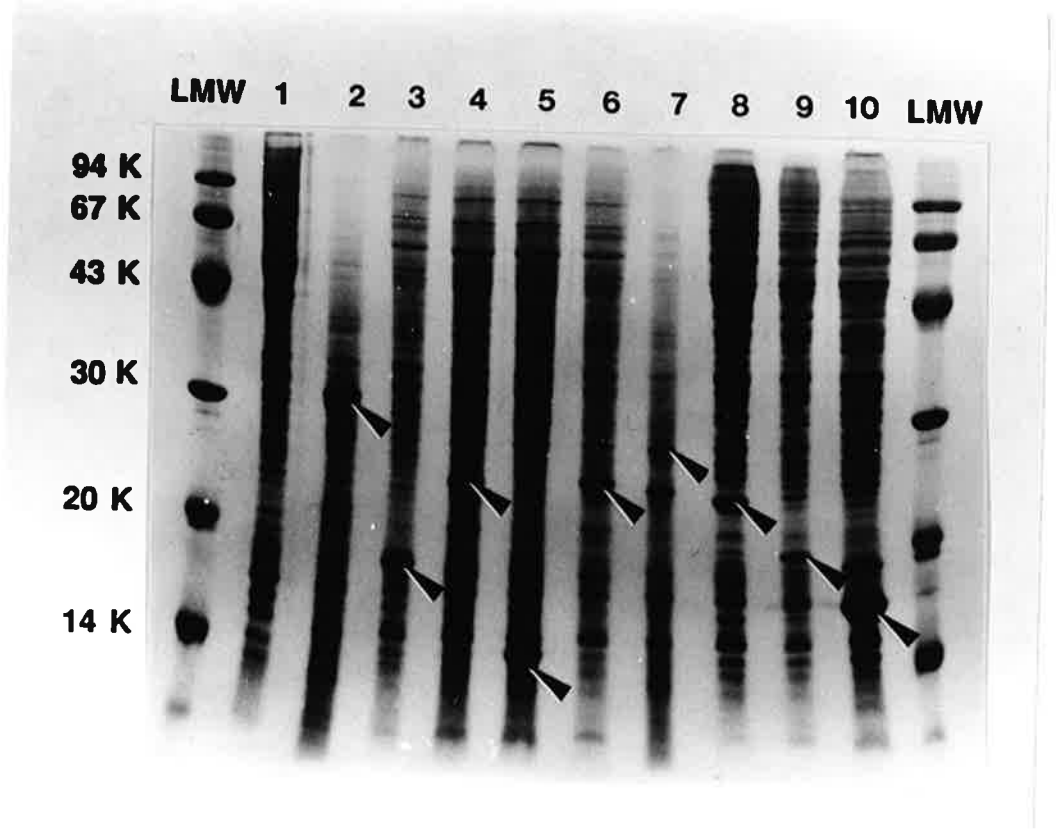


Fig. IV-27. SDS-PAGE (12.5% acrylamide) of proteins stained with Coomassie brilliant blue. Lane LMW, low molecular weight markers (BRL); lane 1, mock infected BmN; lane 2, BmNPV T3 infected BmN; lane 3, BmN infected with a recombinant BmNPV containing a deletion between nucleotides 5 and 310 of the polyhedrin gene; lane 4, deletion between nucleotides 5 and 172; lane 5, 5 and 499; lane 6, 185 and 410; lane 7, 352 and 477; lane 8, 386 and 574; lane 9, 360 and 632; and lane 10, 386 and 715. Arrowheads indicate polyhedrin. Sizes in kilo-Daltons are indicated to the far left.

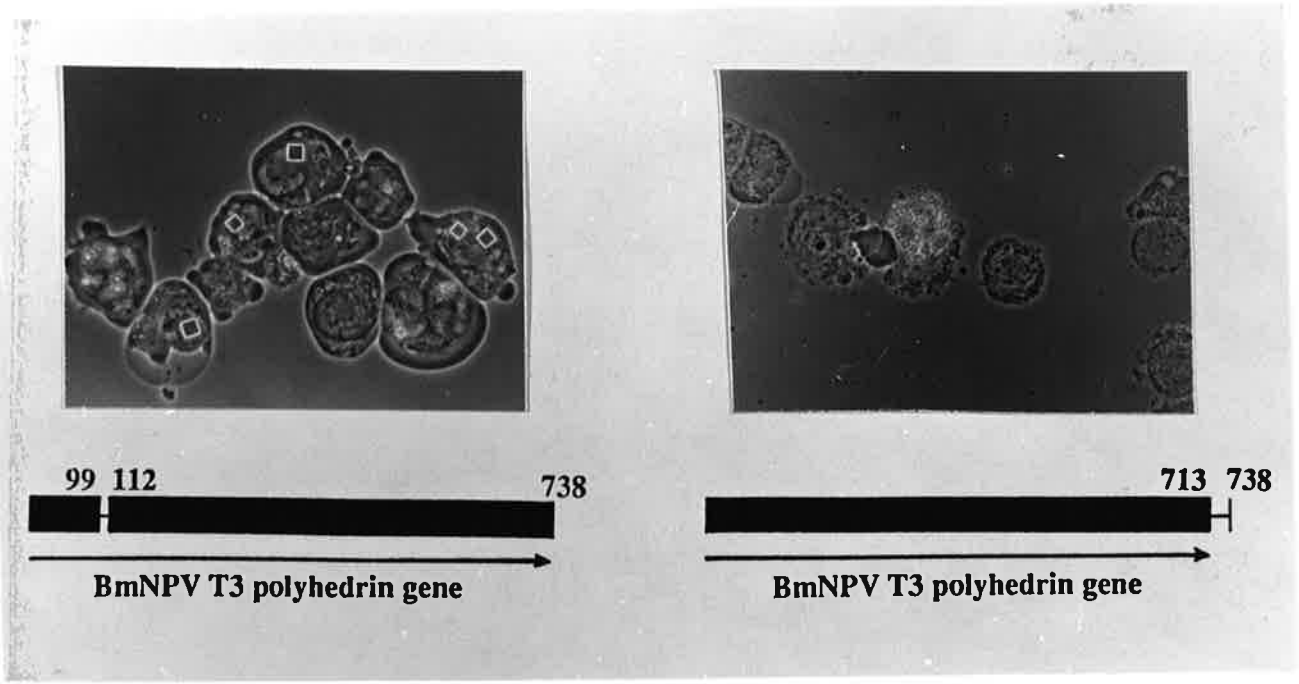


Fig. IV-28. Polyhedral production of BmN cells infected with recombinant BmNPV carrying the BmNPV T3 polyhedrin gene with deletions of nuclotides 100 to 111 (left), and 714 to 738 (right).

These results indicated that 1) proper polyhedral inclusion body assembly required most amino acid sequences of the polyhedrin gene, 2) a minor deletion or substitution within the polyhedrin gene caused the production of cuboidal polyhedra of abnormally large size, and 3) the C-terminus amino acid sequence is responsible for the straight edge phenotype of polyhedra. These results are reasonable in light of the fact that sequence analysis of mutants with different polyhedra shape have shown that they possess only a single amino acid substitution (Carstens et al., 1987; Mori et al., 1989). The amino acid sequence of the polyhedrin gene in various NPVs are highly conserved (more than 80% in the polyhedrin genes of seven NPVs) (see Fig. IV-20). Furthermore, the number of amino acids are very highly conserved (Rohrmann, 1986), indicating that the size of the amino acid chain is important for polyhedra. This study also supports the hypothesis that the size of the amino acid chains is important for the normal formation of polyhedra.

All cuboidal shaped polyhedra were found to be located in the nucleus. To analyze the mechanisms of the nuclear localization of polyhedrin, the location of the deleted polyhedrins were analyzed. Cells infected with various polyhedrin deletion mutants were initially observed using FCTC-conjugated antibody against polyhedral proteins. However, the staining patterns with antibodies were not clear, probably due to the expansion of the cell nucleus and

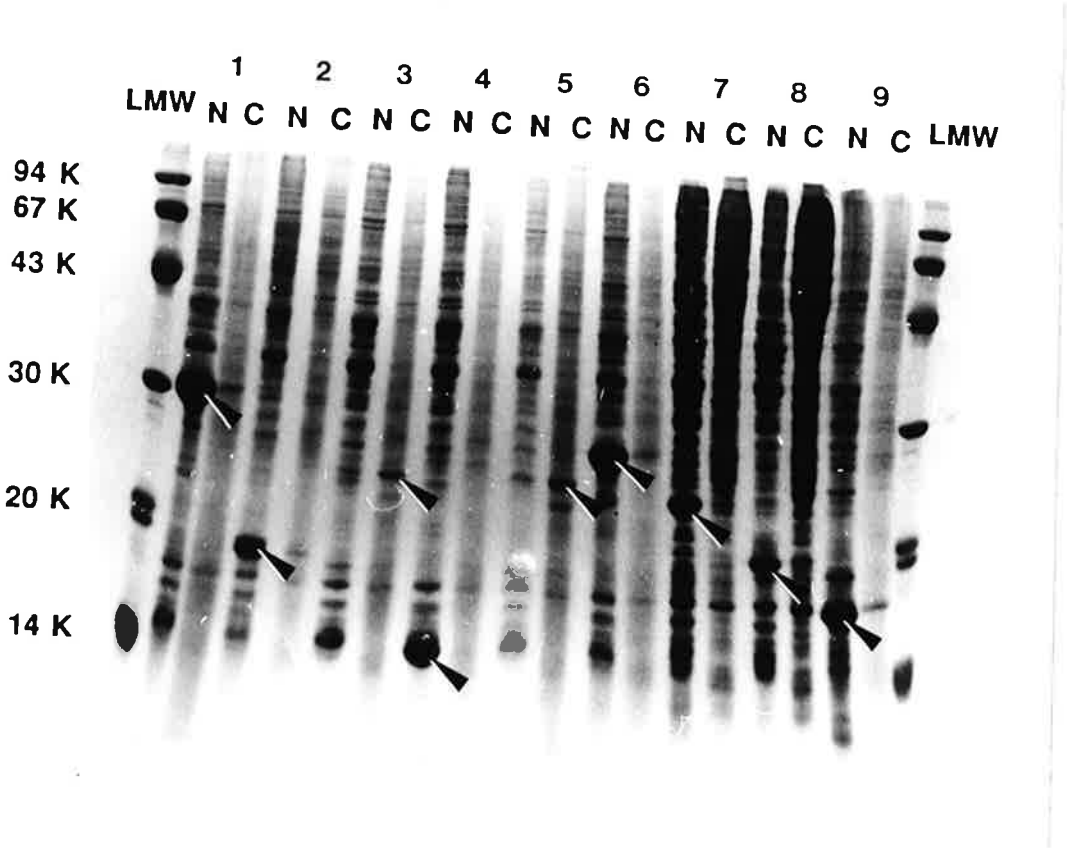


Fig. IV-29. SDS-PAGE (12.5% acrylamide) of proteins from the nucleus (lanes marked N) and cytoplasm (lanes marked C) stained with Coomassie brilliant blue. Lane LMW, low molecular weight markers (BRL); lane 1, BmNPV T3 infected BmN; lane 2, BmN infected with a recombinant BmNPV containing a deletion between nucleotides 5 and 310 of the polyhedrin gene; lane 3, deletion between nucleotides 5 and 172; lane 4, 5 and 499; lane 5, 185 and 410; lane 6, 352 and 477; lane 7, 386 and 574; lane 8, 360 and 632; and lane 9, 386 and 715. Arrowheads indicate polyhedrin. Sizes in kilo-Daltons are indicated to the far left.

low affinity of the antibody to polyhedrin under the non-denaturation condition without SDS treatment. Due to the poor results of the immunoassays, the nucleus was separated from the cytoplasm and polyhedrin localization was analyzed by SDS-polyacrylamide gel electrophoresis (Fig. IV-29). When various samples were stained with Coomassie blue, most polyhedrins with deletions were found in the nucleus fraction. Polyhedrin with deletions between its 2nd and 56th amino acids (recombinant #5-172), localized in the cytoplasm but not in the nucleus. When the deleted region of the polyhedrin gene was narrowed from nucleotides 5-172 to 5-112 (recombinant #5-112), the polyhedrin was also found only in the cytoplasm. Polyhedrin of recombinant #5-61 which lacked nucleotides 5-61 of the polyhedrin gene localized in the nucleus. These results indicate that a signal sequence for nuclear localization is found, between nucleotides 61 and 112. This region corresponded to amino acids 19 to 36. In this region polyhedrin has a sequence of Ala(30)-Lys-Arg-Lys-Lys-His-Leu(36). This sequence of basic amino acids surrounded by nonpolar amino acids is similar to the nuclear localization signal of other peptides such as the large SV40 T antigen (Fig. IV-30) which is localized in the nucleus. However, the #5-139 mutant without this sequence containing a similar sized deletion, localized both in the cytoplasm and nucleus, and mutants with large deletions including this nuclear localization signal produced polyhedrins which localized in the nucleus. These

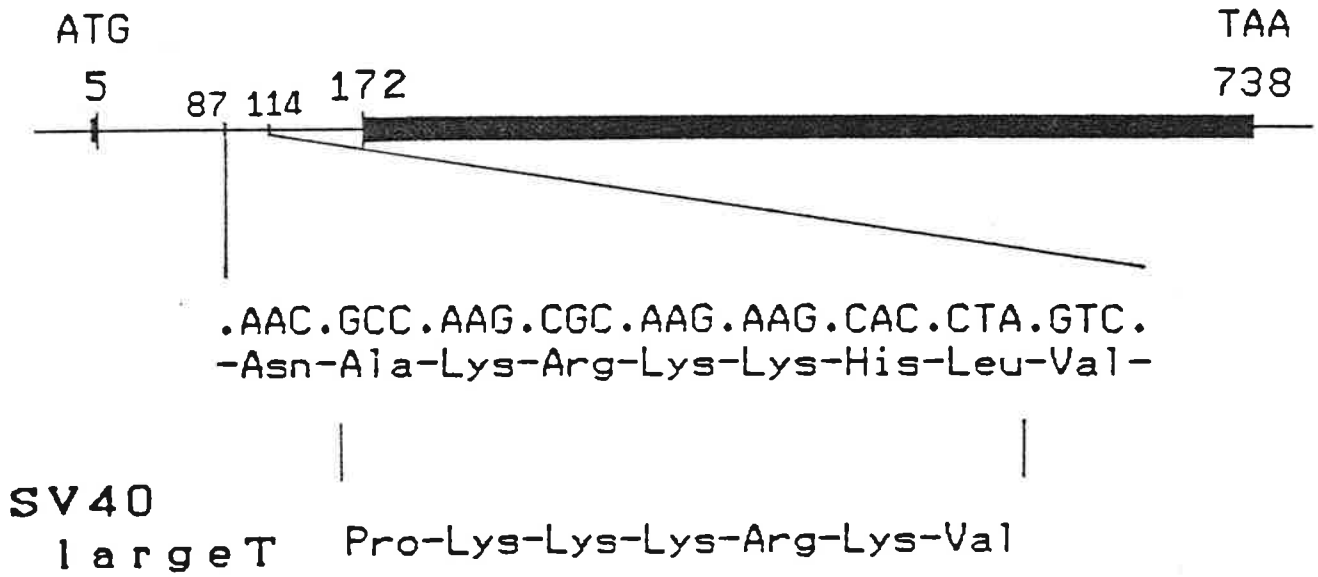


Fig. IV-30. A schematic drawing of the polyhedrin gene of BmNPV T3 with deletions between nucleotides 5 and 172. Polyhedrin of recombinant BmNPVs expressing this mutated gene localized in the cytoplasm. The deleted nucleotides 87 to 114 encode the amino acid sequence indicated (5 basic amino acids flanked by non-polar amino acids). This motif is similar to the putative nuclear localization sequence of the SV40 large T antigen.

results indicated that other mechanisms besides the amino acid localization sequence mentioned above of polyhedrin might be responsible for nuclear localization.

V. Gene structure and expression of BmNPV.

A. Introduction

Nucleotide sequencing of target DNA fragments and computer analysis is one of the most effective strategies for the study of specific genome characteristics. As discussed in previous sections, BmNPV was used for this analysis. BmNPV has not yet been studied at the DNA level in spite of its importance in agriculture and as an expression vector. AcNPV, on the other hand, has been widely studied at the DNA level. BmNPV and AcNPV have high DNA homology yet have different host specificities.

In this section, homologously repeated sequences in the BmNPV genome were isolated, sequenced, and analyzed by comparison to AcNPV. The HindIII C fragment (86.4-99.6 map position) of BmNPV, which supposedly contains four immediate early genes, was also subcloned, and the entire sequence of this region was sequenced. Genome structure and gene expression was examined by computer analysis.

B. Materials and Methods

DNA sequencing: The dideoxy sequencing method as described in section III was employed to determine the nucleotide sequence of genomic fragments inserted and propagated in plasmid vectors. DNA fragment libraries for sequence analysis were constructed from 1) restriction endonuclease digested genomic DNA (section III; Maeda and

Majima, 1990) and 2) sonicated genomic DNA (Maeda et al., 1991a). The ends of all DNA fragments in the two libraries were initially sequenced. To determine internal sequences, nested deletions of DNA fragments from the "restriction endonuclease" library were generated with Bal31 exonuclease treatment. Plasmids with nested deletions were subjected to mini-prep propagation, and plasmids with appropriate deletions were directly sequenced using a pUC primer.

Sequence analysis by personal computer: Sequence information was read and transferred to an IBM PC using a program written in Basic and analyzed by various subprograms of DNASIS. Homology search was carried out using Gene Bank R68.

C. Results and Discussion

1. Homologously repeated sequence

Physical mapping and hybridization experiments revealed the existence of five distinct areas containing homologously repeated (hr) sequences in the genome (designated hr1 to hr5) (see Section III) (Figs. III-6, V-1A). The map positions of these repeated sequences in the BmNPV genome were similar to those found in the AcNPV genome (Cochran and Faulkner, 1983; Guarino et al., 1986). Due to the repetitive nature of the repeated regions, it was not possible to obtain accurate sequence information by computer analysis of information accumulated by random sequencing of

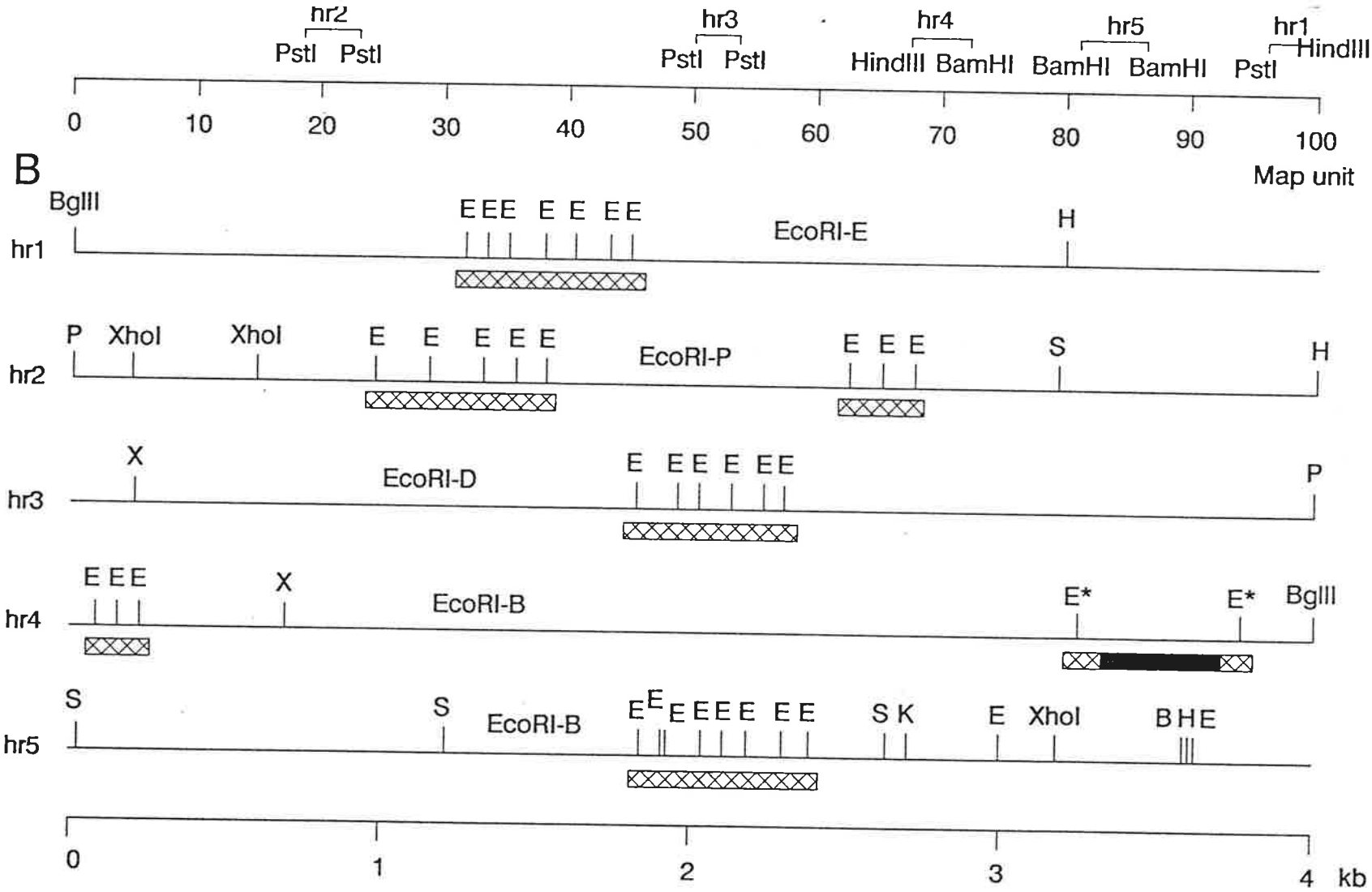


Fig. V-1. (A) Physical map of BmNPV T3 for EcoRI with map units indicated. Five major homologously repeated (hr) sequences (hr1, hr2, hr3, hr4, and hr5) are indicated. (B) Detailed physical maps of the hr sequences with length in kilobase pairs indicated. Cross-hatched bars indicate EcoRI-rich repeated sequences. The black bar indicates previously unknown non-EcoRI-rich repeated sequences. E indicates a core EcoRI sequence and E* indicates a core EcoRI motif.

the genome library from sonicated viral DNA fragments (Maeda et al., 1991a). To obtain accurate sequence information of the hr regions, DNA fragments each containing a complete hr region were prepared from the DNA restriction fragment library (see section III; Maeda and Majima, 1990). The library fragments used for sequencing containing the hr regions were the 4.2 kb PstI-HindIII (96.1-99.6 map units) fragment, the 4.1 kb PstI H (18.8-23.3 map units) fragment, the 4.9 kb PstI K (50.1-53.8 map units) fragment, the 6.1 kb HindIII-BamHI (67.7-72.7 map units) fragment, and the 7.6 kb BamHI D (80.6-86.3 map units) fragment, which contained hr1, hr2, hr3, hr4, and hr5, respectively, as determined by comparison of the physical maps of BmNPV T3 (section III) and AcNPV (Cochran et al., 1982). (Fig. V-1A). The fragments containing the hr regions were treated with Bal31 exonuclease to produce nested deletions for sequencing as described in section III. The sequence of the fragments with nested deletions was determined by the dideoxy sequencing method in two directions. All sequence data was melded and analyzed by the DNASIS program using an IBM compatible personal computer.

When the sequence of the five DNA fragments were analyzed by the Harr plot analysis subprogram of DNASIS, seven major hr regions were identified. The original hr2 and hr4 regions were further subdivided into hr2-left and hr2-right, and hr4-left and hr4-right, respectively. Six of the seven major hr regions were located in regions high in

EcoRI sites as shown in Fig. V-1B. When the number and position of the BmNPV hr regions were compared with the six hr regions (Guarino et al., 1986) of AcNPV; hr2-right of BmNPV seemed not to have a counterpart in AcNPV, as discussed below. Although hr4-right (71.5 map unit) of BmNPV had no EcoRI sites, it corresponded to hr4-right of AcNPV (Guarino et al., 1986). The other hr regions of BmNPV corresponded very closely to their respective counterpart regions in AcNPV. In conclusion, hr1, hr2-left, hr3, hr4-left, hr4-right, and hr5 of BmNPV seemed to correspond to hr1, hr2, hr3, hr4-left, hr4-right, and hr5 of AcNPV, respectively.

Sequence information of the seven BmNPV hr regions is shown in Figs. V-2 to V-6. The BmNPV hr regions, hr1, hr2-left, hr2-right, hr3, hr4-left, hr4-right, and hr5 contained 7, 5, 3, 6, 3, 0, and 8 EcoRI sites, respectively. The number of EcoRI sites in each of the BmNPV hr regions did not correspond to those of AcNPV. In AcNPV, hr1, hr2, hr3, hr4-left, hr4-right, and hr5 contain 5, 8, 8, 2, 4, and 6 EcoRI sites. The number of nucleotides between the EcoRI sites in the BmNPV hr regions mostly varied between 72 and 109 bp and averaged 83 bp. However, hr2-left had two 180 bp separations and hr5 had an 18 bp separation (Fig. V-7). When the repeated sequences of the hr regions of BmNPV were aligned as shown in Fig. V-7, core sequences possessing an EcoRI site at its center were recognized. More than 95% of the nucleotides in the core sequence of the repeated

:1

```
AGATAAAAA CGTTGTAAAT GCGTATGACG CTATCGATGT TGACCCCAAC AAAAAATTTA    60
AATTAATCA TAATCACGAA CAAGTCGATG AAACAAACAA ACAAGAAGTC GTCGATAAAA   120
TGACGCAAC AACATACAAT TCTTGCATCA TAAAAATTTA AATGATATTA TAATTTAAAA   180
AACAATGAC ATCATCGTTT GATTGTGTTT TACACGTAGA ATTCTACTCG TAAAGCCAGT   240
CAGTTTTGA AAAACAAATG ACATCACCTC TTGATTATGT TTTACACGTA GAATTCTACT   300
GTAAAGCCA GTTCAGTTTT GAAAAACAAA TGACGTCATT TCTTGATCAT GTTTTACACG   360
AGAATTCTA CTCGTAAAGC GAGTTCAGTT TTGAAAAAAC AAATGACATC AATTTTTATT   420
TTATAATAA ATGACATCAT TTCTTGATCA TGTTTTACAC GTAGAATTCT ACTCGTAAAG   480
GGGTTCAGT TTTGAAAAAC AAGTGACATC ATTTCTTAAA TTAAGTTTTG AAAAACAAAT   540
ACATCACCT TTCGATCATG TTTTACACGT AGAATTCTAC TCGTAAAGCG AGTTCAGTTT   600
GAAAAACAA GTGACATCAT TTCTTAAATT AAGTTTTGAA AAACAAATGA CGTCATTTCT   660
GATCATGTT TTACACGTAG AATTCTACTC GTAAAGCGAG TTCAGTTTTG AAAAACAAAT   720
ACATCATTT TTTTGATTAT GTTTTACAAG TAGAATTCTA CTCGTAAAAC GAGTTCGGTT   780
TGAGCCGTG TGCAAAAAAT GACATCAGCT TATGACATCA CCCACTGATC GTGCGTTACA   840
GTATAATTC TACTCGTAAA GCGAGTACAT ATTTAGTTAC GTTTCTGAGA TAAGATTGAA   900
GCACGTGTA AAATGTTTCC CGCGCGTTGG CACAACTATT TACAGTGCGG CCAAGTTATA   960
AAGATTCTA ATCTAATATG TTTTAAAACA CCTTTGCAAC CCGAGTTGTT TGCGTACGTG  1020
CTAGCGAAG AAGATGTGTG GACCACAGAA CAGATAGTAA AACAAAACCC CAGTATTGGA  1080
CAATAATCG ATTTAACCAA CACGTCTAAA TATTATGATG GTGTGCACTT TTTGCGGGCG  1140
```

ig. V-2. Nucleotide sequence of hr1 of BmNPV. EcoRI core sequences are underlined.

hr2-left

CAAAGACGAT CTGGGCCTTA CCGAGCATCC GTTGAGCAAA ACAACCGGCA ATTCTGACGG 60

CCGTTTGGGA TGCGGAATAA TTGCCATATG TAAATGATGT CATCGCTTTA ACTCGCTTTA 120

CGAGTAGAAT TCTACGTGTA AAACACAATC AAGAGATGAT GTCATTTGTT TTTCAAAACT 180

GAATTTAAGA AATGATGTCA TTTGTTTTTC AAAACTGAAC TGGCTTTACG AGTATAATTC 240

TACTTGTAAA ACATAATCAA GGGATGATGT CATTTGTTTT TCAAAACTGA ACTCGCTTTA 300

CGAGTAGAAT TCTACGTGTA AAACACAATC AAGAGATGAT GTCATTTGTT TTTCAAAACT 360

GAATTTAAGA AATGATGTCA TTTGTTTTTC AAAACTGAAC TGGCTTTACG AGTATAATTC 420

TACTTGTAAA ACATAATCAA GGGATGATGT CATTTGTTTT TCAAAACTGA ACTGGCTTTA 480

CGAGTAGAAT TCTACGTGTA AAACACAATC AAAAGATGAT GTCATTTGTT TTTCAAAACT 540

GAATGATGTC ATTTGTTTTT CAAAACTGAA CTGGCTTTAC GAGTAGAATT CTACGTGTAA 600

AACACAATCA AAAAATGATG TCATTTGTAG AATGATGTCA TTTGTTTTTC AAAATTAAAC 660

TCGCTTTACG AGTAGAATTC TACGTGTAAC GCATGATCAA AAAATGATGT CATCTTTTAC 720

ACATGATTAT AAACGTGTTT ATGTATGATT CATTTGTTTT TCAAAACTGA ACTCGCTTTA 780

CGAGTAAAAT TCTACTTGTA ACGCAAGATC GGTGGATGAT GTCATTTATT TGTGCAAAGC 840

hr2-right

ATGTGTAATC AATGTGTATT TTACAATAAA AACATTTTTA TTTAAATAAT TATTTTATTT 60

TAGTAAATGA CATCATCCCT TGATTGTGTT TTACACGTAG AATTCTACTC GTAAAGCGAG 120

TTCAGTTTTG AAAAACAAAT GACATCATTT CTTAAATTTA GTTTTGAAAA ACAAATGACA 160

TCATCCCTTG ATTGTGTTTT ACACGTAGAA TTCTACTCGT AAAGCGAGTT TAAACAAATG 240

ACATCATTTC TTAAATTTAG TTTTGAAAAA CAAATGACAT CATCCCTTGA TTGTGTTTTA 300

CACGTAGAAT TCTACTCGTA AAGCGAGTTT AGTTTTGAAA AACAAATGAC ATCATTTCTT 360

AAATTGAGAC TAGTTAATAA TTTGTGTGAT CAGATAACTA TTAACGTCCA CATGGTATTT 420

Fig. V-2. Nucleotide sequence of hr1 of BmNPV. EcoRI core sequences are underlined.

ır3

:GTAGAATCG GATAAATTAT ATTATTTGGA TAATTTACAA GAAGATTCCA TTGTATAAAC 60

\TTTTATGTC GAAAACAAAT GACATCAGCT TATGATTCAT ACTTAATCGT GCGTTACAAG 120

:AGAATTCTA CTTGTAAAGC GAGTTTAATT TGAAAAACAA ATTAGTCATT ATTAAACATG 180

:TAACAATCG TGTATAAAAA TGACATCAGT TTAATGATGA CATCATCTCT TGATTATGTT 240

:TACACGTAG AATTCTACTC GTAAAGCCGG TTCAGTTTTG AAAAACAAAT GACATCATCT 300

:TCGATTGTG TTTTACACGT AGAATTCTAC TCGTAAAGCC AGTTCAGTTT TGAAAAACAA 360

\TGACATCAT TTTTTTAAAT TCAGTTTTGA AAAACAAATG ACATCATCTC TTGATCATGT 420

:TTACACGTA GAATTCTACT CGTAAAGCGA GTTCAGTTTT GAAAAACAAA TGACATCATT 480

:AGTTTTGAA AAACAAATGA CATCATCTTT CGATTGTGTT TTACACGTAG AATTCTACTC 540

;TAAAGCCAG TTCAGTTTTG AAAAACAAAT GACATCATCT TTTGATTGTG TTTTACACGT 600

\GAATTCTAC TCGTAAAGCC AGTTCAATTT TGAAAAACAA ATGACATCAT CTTAGAGGTA 660

;ACCCGTCGC CAAGACGGGT CTGCTCATAT GTCGTTTTGT ATTTGTCATT GCCTCTTTTC 720

'ig. V-3. Nucleotide sequence of hr3 of BmNPV. EcoRI core sequences are underlined.

hr4-left

ATTGTGTACA AAATATGACT CATCGATCGA TCGTGCGTTA CACGTAGAAT TTTACTCGTA 60

AAGCGAGGTT TTAATTAGAG CCGTGTGCAA AACATGACAT CATAACAAAT CATGTTTATA 120

ATCATGTGCA AAATATGACA TCATCCGACA ATTGTGTTTT ACAAGTAGAA TTCTACTCGT 180

AAAGCCAGTT CAGTTTTGAA AAAACAAATG ACATCATCTC TTGATTGTGT TTTACACGTA 240

GAATTCTACT CGTAAAGCCA GTTCAGTTTT GAAAAACAAA TGACATCATC TCTCGATTAT 300

GTTTTACAAG TAGAATTCTA CTTGTAAAGC GAGTTTAAAA ATTTTGTGAC GTCAATGAAA 360

CAACGTGTAG TATTTTTTAC AATATTTAAG TGAAACATTA TGACTTCCAA CAATTTTGTG 420

hr4-right

TTAATAAACG TACATAATAA CACGTTTATC GATACTCACA TTTGATTTTA TCATTCGTTT 60

TTCAAAATTG AACTCGCTTT ACGAGTATAA TTCTACTTGT AAAACACAAT CGAAAGATGA 120

TGTCATTTTT GTACGTTATT ATAAACATGT TTAAACATAG TGCATTGAAC TTAATTTTTG 180

CAAGTTGATA AACATGATTA ATGTATGACT CATTTGTTTG TGCAAGTTAA TAAACGTGTG 240

ATTATAAACA TGTTTAAACA TAGTTCATTG AACTTAATTT TTGCAAGTTG ATAAACATGA 300

TTAATGTATG ACTCATTTGT TTGTGCGAGT TAATAAACAT GTTTAAACAT AGTTCATTGA 360

ACTTAATTTT TGCAAGTTGA TAAACATGAT TAGTGTATGA CTCATTGTTT GTGCAAGTTA 420

ATAAACGTGA TTAATAACAT AGTTCATTGA ACTTGGTTTT TGCAAGTTGA TAAACACGAT 480

TAATGTATGA CTCATTTGTT TGTGCAAGTT AATAAACGTG ATTAATATAT GATTCATATG 540

TTTGTGTAAA AATGATGTCA TTGTACAAAC TCGCTTTACG AGTAGAATTT TACGTGTAAA 600

ACACGATCAC AGCACTTCGT AGTTGTATCG AAAATTGTTC AAGGGCTCTT TGTTAATGTC 660

Fig. V-4. Nucleotide sequence of hr4-left (top) and hr4-right (bottom) of BmNPV. EcoRI core sequences are underlined.

hr5

ACACTCAACA AAAATGTAAT ATTAGACATG ATTAAATAAA TGTTAAAATT TATTGCCTAA 60

TATTATTCTT TGTCATTGCT TGATTGTGTT TTACGAATAT AATTCTACTC GTAACGCAAG 120

ATTTGAGATG ATGTCATTTA TGAAATGATG TCATTTGTTT TTCAAATTGA ACTGGCTTTA 180

CGAGTAGAAT TCTACTTGTA ACGCACGATC AGGTATGAAT CATAAGCTGA TGTCATGGTT 240

TGCACACGGC TCATAACCGA ATTCGTTTTA CGAGTAGAAT TCTACTTGTA ACGCACAATC 300

AAGAAATGAT GTCATTTGTT TTGCAATATG ATATCATACA ATATGACTCA TTTGTTTTTC 360

AAAATCGAAC TTGATTTACG GGCAGAATTC TACTTGTAAA GCACAATCAA AAAAATGATG 400

TCATTTGTTT TTCAAAACTG AACTGGCTTT ACGAGTAGAA TTCTACGTGT AAAACATGAT 480

CAAGAAATGA TGTCATGAGG TGATGTCATT TGTTTTTCGA AACTGAACTG GCTTTACGAG 540

TAGAATTCTA CGTGTAAAAC GCGATCAAGA AATGATGTCA TTTGTTTTTC AAAACTGAAT 600

TTAAGAAATG ATGTCATTTG TTTTTCAAAA CTGAACTGGC TTTACGAGTA GAATTCTACG 660

TGTAAAACAC AATCGAAAGG TGATGTCATT TGTAGAATGA TGTCACTTGT TTTTCAAAAC 720

TGAACTGGCT TTACGAGTAG AATTCTACGT GTAAAACACA ATCAAAATAT GATGTCATTT 780

GTTATAAAAA TAAAATTGAT GTCATGTTTT GCACACGGCT CATAACCAAC TCGCTTTACG 840

AGTAAAATTC TACGCGTAAA ACCCGATTGA TAATTAAATA ATTTATTTGC AAGCTATACG 900

TTGAATCGAA CGGACGTTAT GGAATTGTAT AATATTAAAT ATGCAATCGA TCCAACAAAT 960

Fig. V-5. Nucleotide sequence of hr5 of BmNPV. EcoRI core sequences are underlined.

hr1

ACATCATCGTTT-GATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAAAA-CAAATG
ACATCACCTCTT-GATTATGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAAAA-CAAATG
ACGTCATTTCTT-GATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCAGTTTTGAAAAAACAAATG (28)
ACATCATTTCTT-GATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGGGTTCAGTTTTGAAAAA-CAAGTG (36)
ACATCACCTTTC-GATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCAGTTTTGAAAAA-CAAGTG (36)
ACGTCATTTCTT-GATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCAGTTTTGAAAAA-CAAATG
ACATCATTTTTTTGATTATGTTTTACAAGTAGAATTCTACTCGTAAAACGAGTTCGGTT
acatcatttctt-gatcatgttttacacgtagaattctactcgtaaagcgggttcagttttgaaaaa-caagtg con

hr2-left

AACTCGCTTTACGAGTAGAATTCTACGTGTAAAACACAATCAAGAGATGATGTCAT (98)
ATGATGTCATTTGTTTTTCAAAACTGAACTCGCTTTACGAGTAGAATTCTACGTGTAAAACACAATCAAGAGATGATGTCAT (98)
ATGATGTCATTTGTTTTTCAAAACTGAACTGGCTTTACGAGTAGAATTCTACGTGTAAAACACAATCAAAAGATGATGTCAT (17)
ATGATGTCATTTGTTTTTCAAAACTGAACTGGCTTTACGAGTAGAATTCTACGTGTAAAACACAATCAAAAAATGATGTCAT (7)
ATGATGTCATTTGTTTTTCAAAATTAAACTCGCTTTACGAGTAGAATTCTACGTGTAACGCATGATCAAAAAATGATGTCAT
atgatgtcatttgtttttcaaaactgaactggctttacgagtagaattctacgtgtaaaacacaatcaaaagatgatgtcat con

hr2-right

AAATGACATCATCCCTTGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCA (48)
AAATGACATCATCCCTTGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTTA (39)
AAATGACATCATCCCTTGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTTA
aaatgacatcatcccttgattgtgttttacacgtagaattctactcgtaaagcgagttta con

hr3

ATGA--TTCATACTTAATCGTGCGTTACAAGTAGAATTCTACTTGTAAAGCGAGTTTAATTTGAAAAACAAA (55)
ATGACATCATCTCTTGATTATGTTTTACACGTAGAATTCTACTCGTAAAGCCGGTTCAGTTTTGAAAAACAA
ATGACATCATCTTTCGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAAAACAA (37)
ATGACATCATCTCTTGATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCAGTTTTGAAAAACAA (27)
ATGACATCATCTTTCGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAAAACAA
ATGACATCATCTTTTGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAATTTTGAAAAACAA
atgacatcatcttttgattgtgttttacacgtagaattctactcgtaaagccagttcaattttgaaaaacaa con

hr4-left

AAAATATGACATCATCCGACAATTGTGTTTTACAAGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAAA
AACAAATGACATCATCTCTTGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAA
AACAAATGACATCATCTCTCGATTATGTTTTACAAGTAGAATTCTACTTGTAAAGCGAGTTTAAAAATTTTG
aacaaatgacatcatctctcgattgtgttttacaagtagaattctactcgtaaagccagtttagttttgaaa con

Fig. V-7. Determination of a consensus sequence (con) of the sequence repeats of the BmNPV hr regions (hr1-hr5) indicated. The EcoRI core sequence is underlined. Dashes (-) indicate missing nucleotides. Dot (.) indicates mismatched nucleotides. Numbers in parenthesis indicate the number of unique nucleotide sequences in between the sequence repeats. Summary shows the determination of the consensus (CON) sequence of all hr region consensus sequences. The core EcoRI sequence motif is shown in bold characters. Ac indicates the AcNPV hr region consensus sequences. Differences between BmNPV and AcNPV consensus sequences are underlined.

hr4-right1

GAACTCGCTTTACGAGTATAATTCTACTTGTAAAACACAATCGAAAGATGATGTCATTTTTGTACGTTATTATAAACATGTTT (415)
AAACTCGCTTTACGAGTAGAATTTTACGTGTAAAACACGATCACAGCACTTCGTAGTTGTATCGAAAATTGTTCAAGGGCTCTTTGTTAATGTC
.aactcgctttacgagta.aatt.tac.tgtaaaacac.atc.. con

hr4-right2

AAACATAGTGCATTGAACTTAATTTTTGCAAGTTGATAAACATGATTAATGTATGACTCATTTGTTTGTGCAAGTTAATAAACGTGTGATT (12)
AAACATAGTTCATTGAACTTAATTTTTGCAAGTTGATAAACATGATTAATGTATGACTCATTTGTTTGTGCGAGTTAATAAACATGTTT
AAACATAGTTCATTGAACTTAATTTTTGCAAGTTGATAAACATGATTAGTGTATGACTCATT-GTTTGTGCAAGTTAATAAACGTGATTAA
TAACATAGTTCATTGAACTTGGTTTTTGCAAGTTGATAAACACGATTAATGTATGACTCATTTGTTTGTGCAAGTTAATAAACGTGATTA
aaacatagttcattgaacttaatttttgcaagttgataaacatgattaatgtatgactcatttgtttgtgcgagttaataaacatgttt con

hr5

AAATTGAACTGGCTTTACGAGTAGAATTCTACTTGTAACGCACGATCA (24)
TGGTTTGCACACGGCTCATAACCGAATTCGT
TTTACGAGTAGAATTCTACTTGTAACGCACAATCAAGAAATGATGTCAT (46)
AAATCGAACTTGATTTACGGGCAGAATTCTACTTGTAAAGCACAATCAAAAAAATGATGTCAT (10)
AAACTGAACTGGCTTTACGAGTAGAATTCTACGTGTAAAACATGATCAAGAAATGATGTCAT (23)
AAACTGAACTGGCTTTACGAGTAGAATTCTACGTGTAAAACGCGATCAAGAAATGATGTCAT (46)
AAACTGAACTGGCTTTACGAGTAGAATTCTACGTGTAAAACACAATCGAAAGGTGATGTCAT (27)
AAACTGAACTGGCTTTACGAGTAGAATTCTACGTGTAAAACACAATCAAAATATGATGTCAT
aaactgaactggctttacgagtagaattctacgtgtaaaacacaatcaaaatatgatgtcat con

Summary

1. ACATCATTTCTTGATCATGTTTTACACGTA**GAATTC**TACTCGTAAAGCGGGTTCAGTTTTGAAAAACAAGTG
2L (-) ATGACATCATCTTTTGATTGTGTTTTACACGTA**GAATTC**TACTCGTAAAGCCAGTTCAGTTTTGAAAAACAAATG
2R AAATGACATCATCCCTTGATTGTGTTTTACACGTA**GAATTC**TACTCGTAAAGCGAGTTTA
3 ATGACATCATCTTTTGATTGTGTTTTACACGTA**GAATTC**TACTCGTAAAGCCAGTTCAATTTTGAAAAACAA
4L AAATGACATCATCTCTCGATTGTGTTTTACAAGTA**GAATTC**TACTCGTAAAGCCAGTTTAGTTTTGAAA
4R1(-) GAT.GTGTTTTACA.GTA.**AATT.**TACTCGTAAAGCGAGTT
5 (-) ATGACATCATATTTTGATTGTGTTTTACACGTA**GAATTC**TACTCGTAAAGCCAGTTCAGTTT
con aaatgacatcatcttttgattgtgttttacacgtag**aattc**tactcgtaaagccagttcagttttgaaaaacaaatg

Ac aatgacatcatcccttgattgtgttttacaagtag**aattc**tactcgtaaagcgagtttagttttgaaaaacaaagc

Fig. V-7. continued

sequences of each hr region were conserved (Fig. V-7). Sequences (28 bp) flanking the core EcoRI site formed a palindrome in all conserved consensus regions (Fig. V-8). This palindrome motif was highly conserved with a maximum of only one or two nucleotide mismatches. Furthermore, the mismatch(s) was always found at the same position(s) of the palindromes, i.e., 7 and/or 8 bp to the left and/or right of the core EcoRI (Fig. V-8). This characteristic was the same as that observed in AcNPV (Guarino et al., 1986). However, one-third of the mismatched bases in the palindrome were not the same as that found in AcNPV. Furthermore, the palindrome of the BmNPV hr4-left region was perfect (without mismatch); perfect palindromes were not found in any AcNPV hr region. As hypothesized by Guarino et al. (1986), mismatches in the palindrome motif seemed to offer some functional advantages. Recently, this palindrome motif of AcNPV was shown to have the ability to bind a specific protein by gel retardation assay. The binding protein has been speculated to be IE-1 by Carson et al. (1991) or a host protein by Y. Hashimoto (personal communication). The existence of mismatches at different nucleotide positions in the palindrome will change binding affinity slightly so that the function of each hr region can be differentiated. Also, a functional direction (e.g. cis-activation of genes located only one side of the palindrome) can be generated. Repeated sequences have an enhancing function that can cis-activate delayed-early genes, e.g. the p39 gene, however, the

(A)

hr1	Bm	TTTAC**AC**GTAGAATTCTAC**TC**GTAAA
	Ac	TTTACA<u>A</u>GTAGAATTCTACTCGTAAA
hr2l	Bm	TTTAC**GA**GTAGAATTCTAC**GT**GTAAA
	Ac	TTTACGAGTAGAATTCTAC<u>T</u>TGTAAA
hr2r	Bm	TTTAC**AC**GTAGAATTCTAC**TC**GTAAA
	Ac	TTTAC<u>GA</u>GTAGAATTCTACT<u>T</u>GTAAA
hr3	Bm	TTTAC**AC**GTAGAATTCTAC**TC**GTAAA
	Ac	TTTACA<u>A</u>GTAGAATTCTACTCGTAAA
hr4l	Bm	TTTACAAGTAGAATTCTACTTGTAAA
	Ac1	TTTACAAGTAGAATTCTACT<u>C</u>GTAAA
hr4r1	Bm	TTTACGAGTAgAATTcTACt**T**GTAAA
	Ac2	TTTACGAGTAGAATTCTACTTGTAAA
hr5	Bm	TTTAC**GA**GTAGAATTCTAC**GT**GTAAA
	Ac	TTTACGAGTAGAATTCTACGTGTAAA

(B)

hr4r2	Bm	AAA**CAT**AGTTCATTGAACT**TAA**TTT
	Ac	AAACATGGT<u>A</u>CATTGAACTTAATTT

Fig. V-8. (A) Comparison of the palindrome motifs of the indicated hr regions for BmNPV (Bm) and AcNPV (Ac). Bold characters indicate mismatches in the BmNPV palindrome motif. Underlining indicates mismatch between the BmNPV and AcNPV palindrome motifs. Small characters indicate one of two nucleotides possible. All mismatches were found either 7 and/or 8 nucleotides to the left and/or right of the core <u>Eco</u>RI motif. (B) Comparison of newly identified palindrome motifs in the hr4-right regions of BmNPV and AcNPV. Mismatches were found 7 to 9 nucleotides to the left and right of the center, "T", of the BmNPV palindrome. See (A) for bold and underlined characters.

location of delayed early genes does not alter its enhancing activity (Guarino et al., 1986). This enhancing activity was activated in the expression of the IE-1 gene product (Guarino et al., 1986).

Guarino et al. (1986) have shown that homologous repeats have polarity. When the polarity of the BmNPV hr regions was examined by the Harr plot analysis, five of the six hr regions of BmNPV had the same polarity to the corresponding AcNPV hr regions (Guarino et al., 1986). When the consensus sequences of the seven BmNPV hr regions were compared to each other after orienting them to the same polarity, 97% sequence conservation was observed. The consensus sequence for all hr regions is shown in Fig. V-7. When this BmNPV hr consensus sequence was compared to the AcNPV hr consensus sequence, after orienting them to the same polarity, they showed high sequence conservation (92%) in the region 80 bp flanking the core EcoRI site (Fig. V-7).

Harr plot analysis showed 12 of 15 nucleotides were identical between the BmNPV and AcNPV hr1 core sequences (Fig. V-9A). Harr plot analysis also showed the BmNPV hr1 and AcNPV hr1 had 7 and 5 repeats, respectively, each showing high homology. The two extra repeats in BmNPV hr1, were consistent with the 180 bp longer length of BmNPV hr1 and the two extra EcoRI sites found in BmNPV hr1. Sequences flanking the BmNPV and AcNPV hr1 regions contained no homologous repeats, however, they showed high homology as indicated by arrows in Fig. V-9A. When the flanking regions

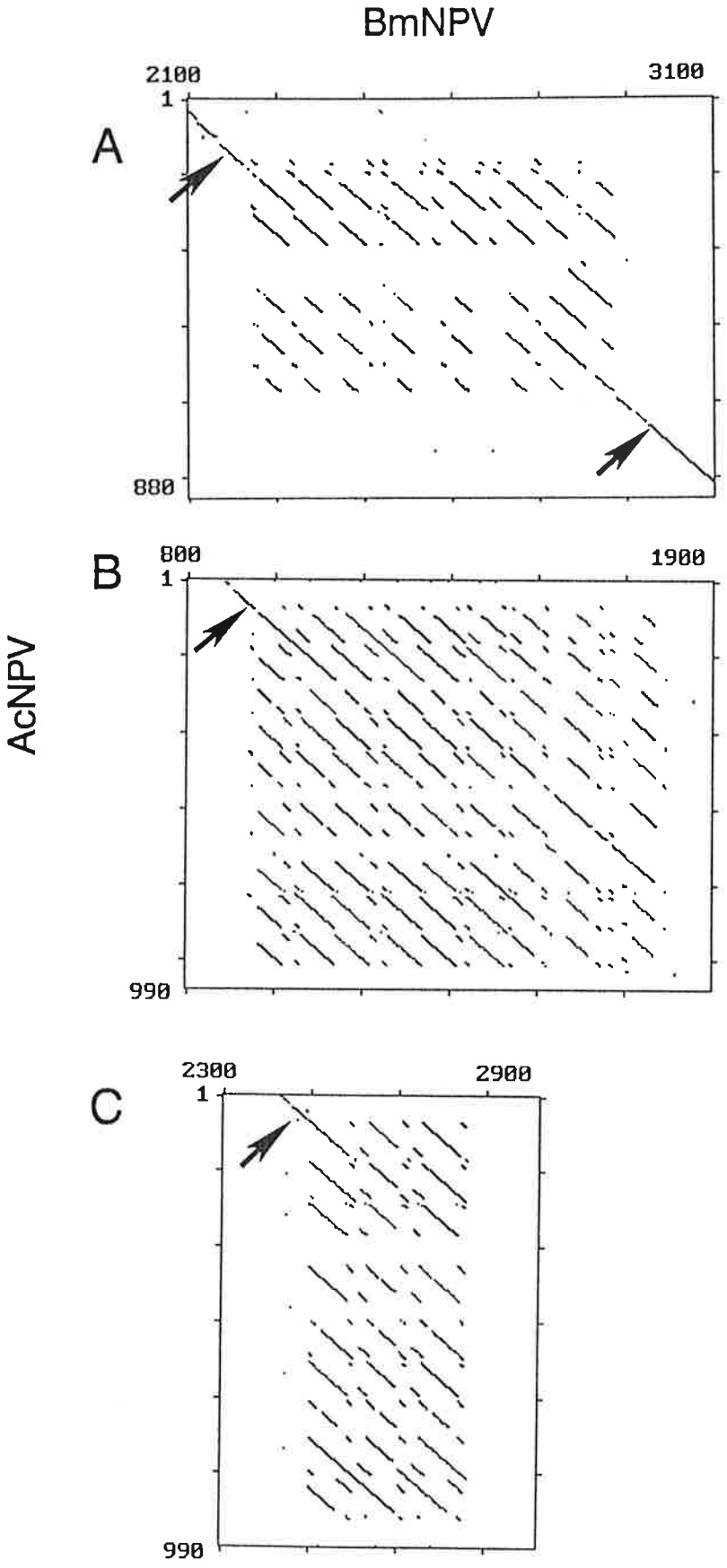


Fig. V-9. (A) Harr plot analysis of the hr1 regions of AcNPV and BmNPV. (B) Harr plot analysis of the hr2-left regions of AcNPV and BmNPV. (C) Harr plot analysis of the hr2-right regions of AcNPV and BmNPV.

of hr1 of BmNPV and AcNPV were compared sequences 246 bp to the left (upstream in the physical map) of hr1 (containing hr and flanking sequences) were 88.4% identical when deletions and insertions (14 bp) were excluded, and 83.4% identical when included. Sequences 353 bp to the right (downstream) of hr1 (containing hr and flanking sequences) were 95.4% and 90.4% identical excluding and including deletions and insertions, respectively. Both flanking and hr sequences possessed similar mismatch ratios, indicating the mutation rate of these areas was similar. Mismatches were presumably accumulated during evolution by random mutations, i.e. small (a few bases) deletions or insertions. The two additional repeats in BmNPV hr1 indicated that the insertion of repeated sequences could occur more frequently in hr regions than in non-hr regions. It is of interest to note that AcNPV hr1 had a 50 bp sequence near its center having no apparent homology to other hr regions of BmNPV and AcNPV as indicated by the discrete lines in Fig. V-9A.

Hr2-left and hr2-right of BmNPV, are separated by the 0.9 kb EcoRI P fragment. AcNPV seems not to have the same sequence corresponding to hr2-right of BmNPV as speculated by the lack of EcoRI sites. When the number of repeated sequences of BmNPV hr2 and AcNPV hr2 were compared, AcNPV hr2 has 8, while, BmNPV hr2-left has 5 and BmNPV hr2-right has 3 (see Fig. V-1B). Flanking sequences to the left side of AcNPV hr2 and BmNPV hr2-left showed homology as indicated

sequences of each hr region were conserved (Fig. V-7). Sequences (28 bp) flanking the core EcoRI site formed a palindrome in all conserved consensus regions (Fig. V-8). This palindrome motif was highly conserved with a maximum of only one or two nucleotide mismatches. Furthermore, the mismatch(s) was always found at the same position(s) of the palindromes, i.e., 7 and/or 8 bp to the left and/or right of the core EcoRI (Fig. V-8). This characteristic was the same as that observed in AcNPV (Guarino et al., 1986). However, one-third of the mismatched bases in the palindrome were not the same as that found in AcNPV. Furthermore, the palindrome of the BmNPV hr4-left region was perfect (without mismatch); perfect palindromes were not found in any AcNPV hr region. As hypothesized by Guarino et al. (1986), mismatches in the palindrome motif seemed to offer some functional advantages. Recently, this palindrome motif of AcNPV was shown to have the ability to bind a specific protein by gel retardation assay. The binding protein has been speculated to be IE-1 by Carson et al. (1991) or a host protein by Y. Hashimoto (personal communication). The existence of mismatches at different nucleotide positions in the palindrome will change binding affinity slightly so that the function of each hr region can be differentiated. Also, a functional direction (e.g. cis-activation of genes located only one side of the palindrome) can be generated. Repeated sequences have an enhancing function that can cis-activate delayed-early genes, e.g. the p39 gene, however, the

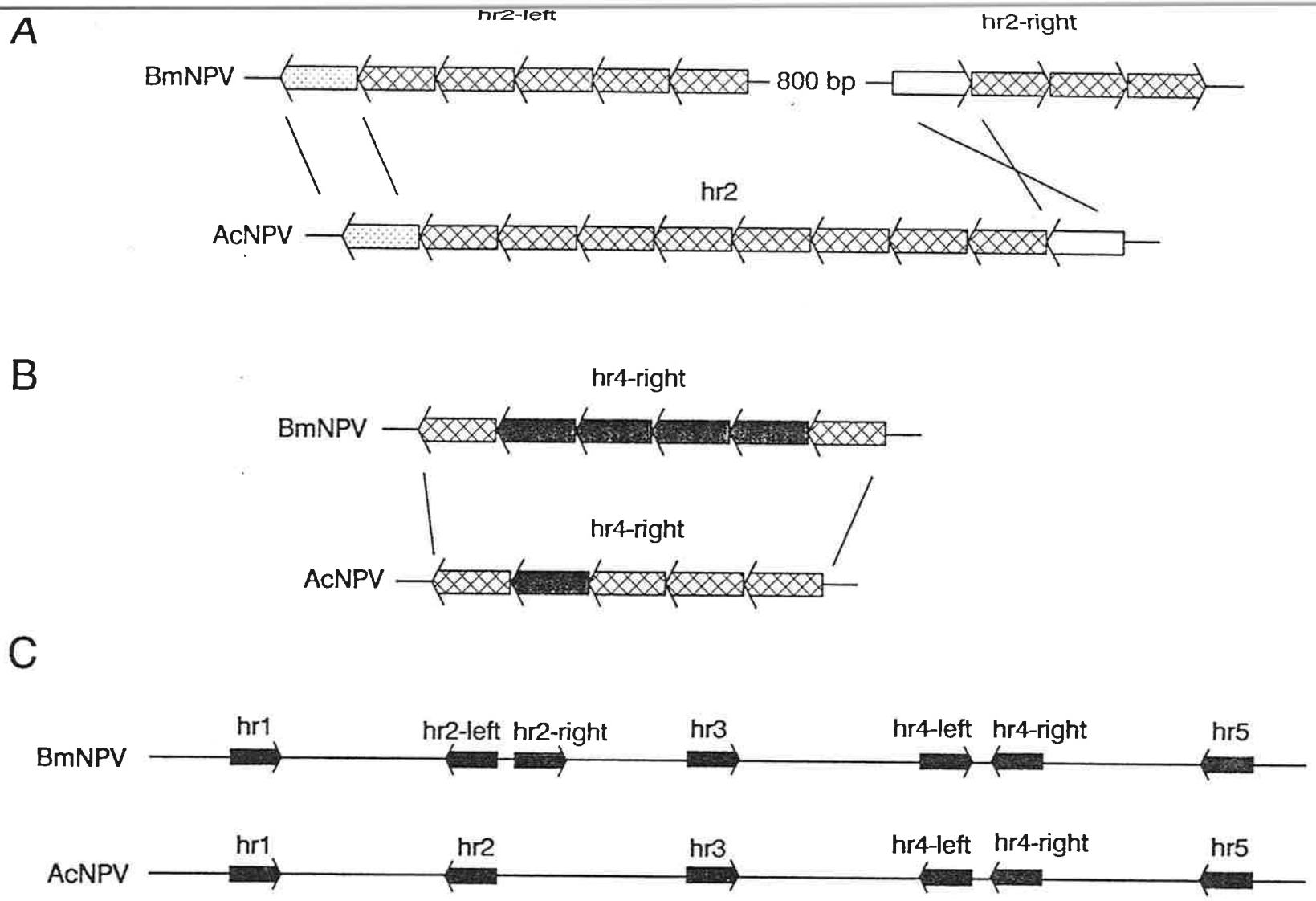


Fig. V-10. (A) Comparison of the hr2 regions of BmNPV and AcNPV. Cross-hatched arrows indicate homologous repeats and their orientation. Dotted and plain arrows indicate non-repeated, homologous flanking sequences and their orientation. (B) Comparison of the hr4-right regions of BmNPV and AcNPV. Cross-hatched arrows indicate homologous repeats and their orientation. Black arrows indicate previously unknown homologous repeats and their orientation. (C) Comparison of the location and orientation of the indicated hr regions of BmNPV and AcNPV.

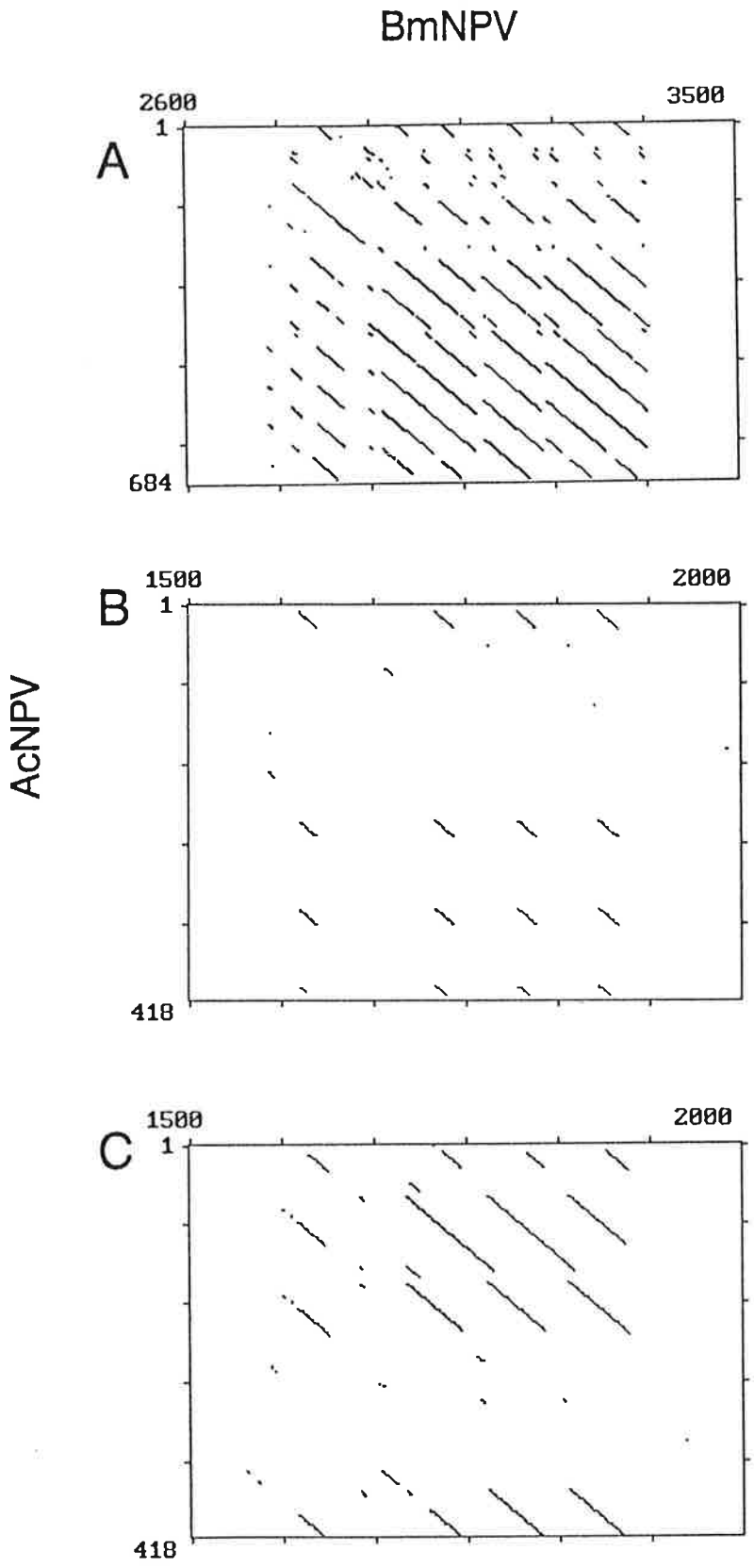


Fig. V-11. (A) Harr plot analysis of the hr3 regions of AcNPV and BmNPV. (B) Harr plot analysis of the hr4-left regions of AcNPV and BmNPV oriented in the same direction. (C) Harr plot analysis of the hr4-left regions of AcNPV and BmNPV oriented in the opposite direction.

AcNPV had 3 and 2 repeated sequences, respectively. When both hr regions were analyzed by the Harr Plot analysis program, they showed relatively low homology (Fig. V-11B). However, they showed high homology when they were compared in the same orientation (Fig. V-11C), indicating that this region was inverted between BmNPV and AcNPV. Since information of the flanking sequence of the hr4-left region AcNPV is not available, analysis of the flanking sequences of hr4-left of BmNPV and AcNPV was not possible.

As described above, hr4-right of BmNPV had no EcoRI sites, however, two consensus core hr sequences were found by homology sequence search. Two single nucleotide substitutions in the core EcoRI motif from GAATTC to TAATTC and GAATTT (indicated by E* in Fig. V-1B) were found in both BmNPV core sequences. The two core sequences were found at the ends of hr4-right in BmNPV (indicated by the cross-hatched bar in Fig. V-1B). Between these two core sequences, four repeated sequences each about 90 bp long were identified (indicated by the solid black bar in Fig. V-1B). These repeated sequences had no apparent homology to other hr sequences, however, the center of these repeats possessed inverted palindrome motifs (Fig. V-8). Fig. V-10B shows a schematic drawing of hr4-right (four repeats surrounded by typical hr sequences) of BmNPV. AcNPV hr4-right had a single repeat between the flanking repeated sequences having high homology to the middle four repeats of BmNPV hr4-right (Fig. V-12A). This region in AcNPV was

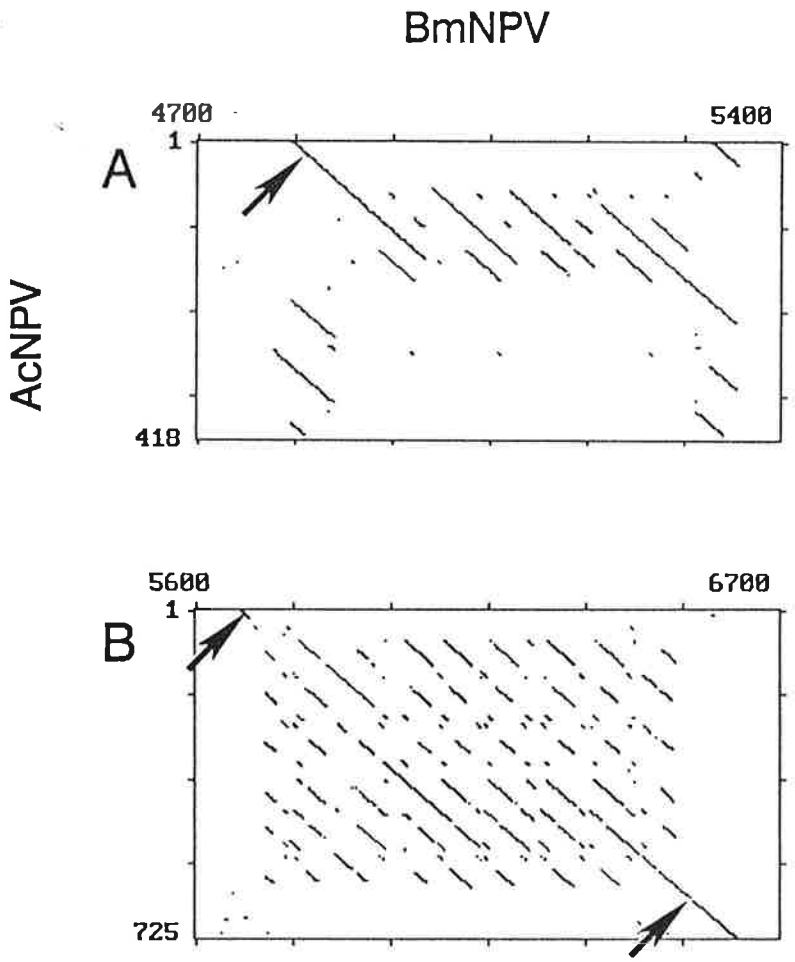


Fig. V-12. (A) Harr plot analysis of the hr4-right regions of AcNPV and BmNPV. (B) Harr plot analysis of the hr5 regions of AcNPV and BmNPV.

characterized as a nonhomologous insertion between hr sequences and contributed to the low sequence homology (64%) of hr4-right when compared to the other hr regions (80-87%) (Guarino et al., 1986). The four repeats of BmNPV hr4-right were presumably generated tandemly from the repeat unit found in AcNPV hr4-right during its evolution. This newly identified repeated sequence motiff was not found in a search of other areas of the BmNPV genome by homology search.

Harr plot analysis showed relatively high overall sequence homology between the hr5 regions of BmNPV and AcNPV (Fig. V-12B). BmNPV hr5 is the largest hr region of BmNPV and possessed two more repeats than the AcNPV hr5 region. Flanking sequences both to the left and right of hr5 were also conserved, indicating the importance of hr5 and its flanking regions. This region contains the p35 and p26 genes to the left and right sides of the hr, respectively, in both BmNPV (unpublished observations) and AcNPV (Friesen et al., 1986; Liu et al., 1986).

When sequences (approx. 80 bp) flanking the core palindrome motif were compared, high sequence conservation (around 92%) was observed. The conservation rate was similar to that calculated for regions flanking the hr regions. These values were also very similar to rates found in other regions of the baculovirus genome such as the DNA-binding protein (96%, Maeda et al., 1991a). On the other hand, differences in the number of homologous repeats

between the corresponding hr regions, indicated that sequence duplication or deletions occurred much more frequently in hr regions. These types of mutations were probably generated by the formation of loops and homologous recombination between looped DNA regions as was probably the case in hr2 of BmNPV. Simple duplication was also identified in several regions of the baculovirus genome and seemed to occur relatively frequently; e.g. repeat of a 30 bp fragment in the basic DNA binding protein gene of BmNPV (Maeda et al., 1991a), 16 repeats of 48 bp in the gene encoding the polyhedron envelope protein of AcNPV (Gombart et al, 1989).

In general, the corresponding hr regions between BmNPV and AcNPV were relatively conserved in their sequence, polarity, and location. Only hr4-left of BmNPV and AcNPV showed different polarities. From detailed sequence analysis, BmNPV hr2 seemed to have evolved from an ancestor having a sequence similar to AcNPV hr2 by partial inversion and recombination. Gene inversion was also reported in the polyhedrin gene of O. pseudotsugata. However, since the overall gene arrangement of baculovirus genomes is conserved, gene rearrangement seems to occur infrequently and only in limited areas. The difficulty of gene rearrangement is also supported by the fact that there is no evidence of gene rearrangement in two regions that were examined between the hr regions, e.g., between hr1 and hr5 and between hr2 and hr3. Baculovirus may also have

suppressive mechanisms to block large inversion by recombination which could cause gene rearrangements.

A new type of repeated sequence was found in the hr4-right region of BmNPV. The newly identified repeated sequence had no apparent homology to previously identified hr sequences, however, it possessed a core palindrome motif. The structure of these homologous repeats, four similar repeats surrounded by two different repeats, is very interesting, however, its function is yet to be identified.

The real function of the hr regions in the baculovirus genome are still unclear. Originally hr regions were speculated as sites for the initiation of replication (Kuzio and Falkner, 1984). Recently, Guarino et al. (1986) have shown them to be regions enhancing cis-activation of delayed early genes by binding the transactivating product of IE-1 onto its core palindrome. We are currently attempting to inactivate one of the hr regions by insertion of the beta-galactosidase gene and examining its activity. Detailed comparison of this mutant lacking an hr region with wild-type virus will help reveal the true function of this hr region.

2. Nucleotide sequence analysis of the 86.3 and 99.6 map unit fragment of the BmNPV genome.

Sequence information at the nucleotide level is essential for molecular biological study of baculoviruses. The entire sequence of the 15.5 kb HindIII C fragment (86.3-

99.6 map position) of the BmNPV genome was sequenced by the dideoxy sequencing procedure as described in Section III and IV. Initially, the 15.5 kb HindIII C fragment was digested into four fragments, two HindIII-PstI fragments, and two PstI fragments (see Fig. III-6, V13A) and subcloned into plasmids. After obtaining sequence data for both directions, the entire sequence was melded by the combining program of DNASIS. The nucleotide sequence of HindIII C is shown in Figs. V-14. This HindIII fragment was found to be 16,945 bp long by direct sequencing rather than 15.5 kb which was initially speculated by calculation from its mobility in an agarose gel (section III). The 1.4 kb difference was partially due to the poor separation of fragments over 10 kb in length in agarose gels. Furthermore, when the initial HindIII C fragment map was constructed by summing the sizes of the smaller fragments of which it is composed, several repeated sequences containing EcoRI sites were not properly accounted for especially in the 1 kb long hr1 region (see section IV), which is found in this fragment.

When the entire sequence of the HindIII C fragment was analyzed by the Harr plot analysis program, one homologously repeated sequence was found as described previously in this section. The remaining sequence was mostly unique and contained 12 open reading frames (ORFs) (of more than 500 bp in length) and their flanking sequences containing putative promoters and poly $(A)^+$ signals. To confirm the translation

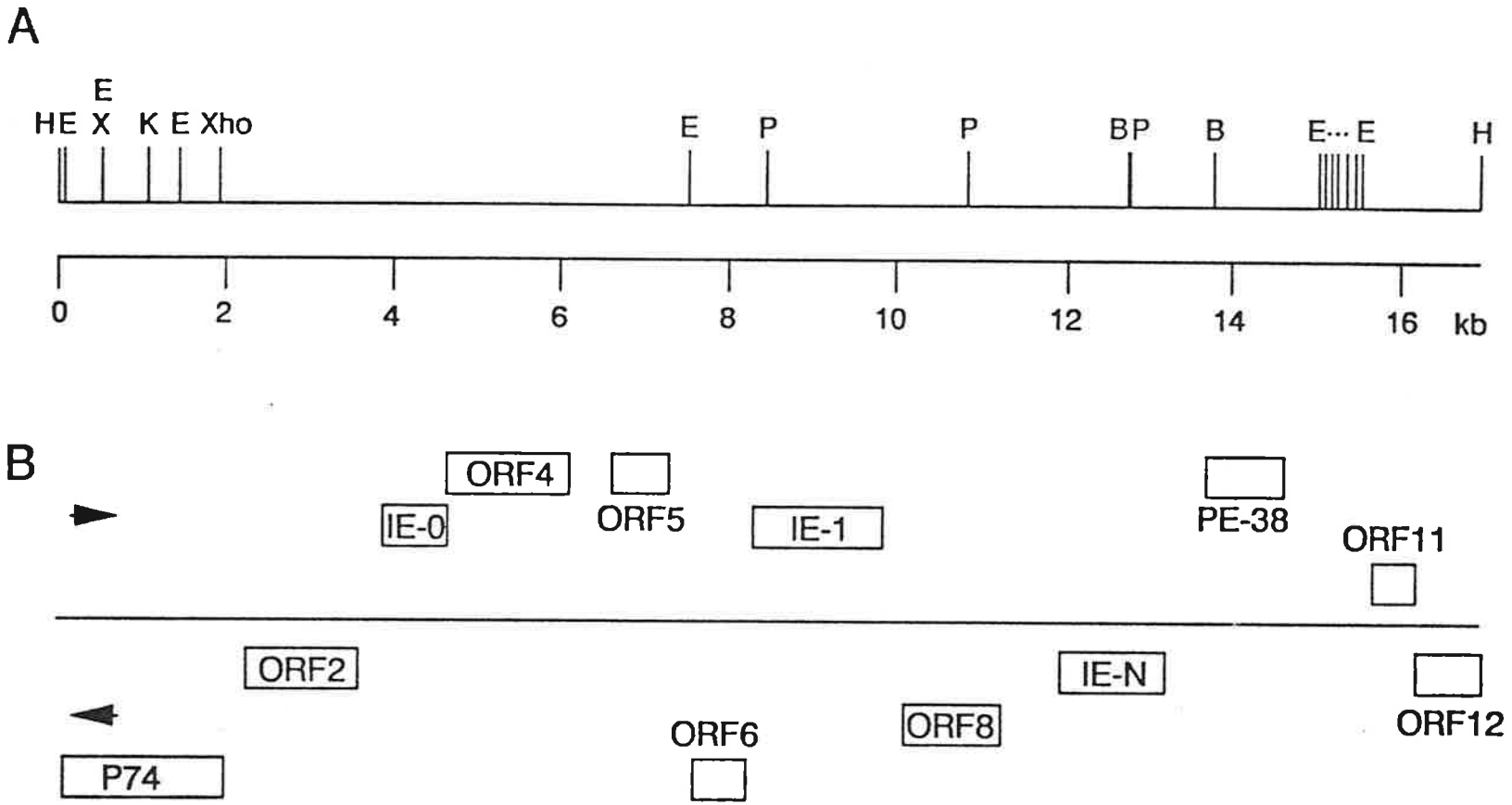


Fig. V-13. (A) Physical map of the HindIII C fragment of BmNPV (86.3-99.6 map units). Symbols: H, HindIII; E, EcoRI; X, XbaI; K, KpnI; Xho, XhoI; P, PstI; B, BamHI. Sizes in kilo-base pairs are indicated below the map. (B) Open reading frames found in the HindIII C fragment of BmNPV. Arrowheads indicate direction of transcription.

of these open reading frames, direct detection of translational products is necessary. However, the probability that these regions are transcribed and translated can be speculated using information accumulated from previous studies of baculovirus gene expression (see Blissard and Rohrmann, 1990).

The 12 ORFs found in HindIII C were initially compared with genes found in a part of the corresponding region in the AcNPV genome that has been sequenced previously (Fig. V-13B). Highly conserved open reading frames corresponding to p74, IE-0, IE-1, IE-N, PE-38 of AcNPV and an open reading frame designated ORF12 in BmNPV were identified. The position and directions of the corresponding genes between the BmNPV and AcNPV genomes were the same (Fig. V-13B). All sequences were generally conserved between BmNPV and AcNPV. However, the outside open reading frames were less conserved and deletions and insertions were found in these regions. In the coding regions of IE-N and PE-38, there were 96 and 51 bp deletions and insertions, respectively. Amino acid sequence homology between the, p74, IE-0, IE-1, IE-N, and PE-38 genes showed 91.5%, 96.8%, 94.5%, 72.8%, and 86.1% identity, respectively. These results indicate high conservation of gene structure between the two viruses. These six BmNPV genes were considered to be actively transcribed and translated, since transcription and translation of the corresponding genes in the AcNPV genome have been confirmed. When the 5′ flanking sequences of the

proposed translational start sites were analyzed, 8 open reading frames had a typical late gene consensus sequence (ATAAG) for transcription within 200 bp of the proposed translational signal, (information for the 5′ flanking sequence of ORF12 was not available). ORF6, IE-1, and IE-N, did not contain a late gene transcriptional consensus sequence, however, IE-1 and IE-N possessed consensus early gene promoters of baculoviruses, i.e., the TATA box located 21-25 upstream of the CAGT motif (marked by #### in Fig. V-14). In addition, IE-0 and PE-38 also had this consensus sequence for early gene expression. Only ORF6 had neither early nor late gene consensus transcription sequences of baculoviruses. When the 3′ regions of the 12 open reading frames were examined, p74, IE-0, ORF4, ORF6, and ORF11 had no poly $(A)^+$ signal (AATAAA) within 300 bp downstream of the translational stop, however, IE-0 was shown to have a spliced signal (identified by Chisholm and Henner, 1988) and ORF4 possessed many AT rich sequences in the 3′ flanking sequence, which might have a similar function as the poly $(A)^+$ signal. From these results, ORF6 seemed not to be translated since it had no early or late gene consensus transcriptional start sequences, and it was relatively small (605 bp). ORF11 possessed a short open reading frame (506 bp) and no poly $(A)^+$ signal. The remaining ten ORFs were most likely transcribed and translated.

Yih and Maeda have recently found that IE-0, IE-1, PE-38, and IE-N of BmNPV are transcribed at an early (2 hours

post infection) stage of infection (unpublished observations). These early genes of BmNPV all had an early gene transcription consensus sequence and were likely transcribed as early genes. The remaining open reading frames, ORF2, ORF3, ORF4, ORF5, ORF8, and ORF12 were considered to be transcribed at a late stage of infection. When the putative amino acid sequence of these regions were compared to amino acid sequences in GeneBank R68 (1991), amino acid sequences with high homology were not found.

When the nucleotide and amino acid sequences of the six gene areas between BmNPV and AcNPV having sequencing information were compared, they showed relatively low differences indicating the importance of the amino acid sequence of p74, IE-0, and IE-1. The numerous insertions and deletions observed outside these genes suggest that most areas outside limited small regions, such as promoters and poly $(A)^+$ signals, are not important in the viral genome construct.

```
                                                      ......
AAGCTTTTGAATTCGATTCAGACGCTCGTCGTGGTAAACGCAGTTCCAAGTAAATAAATCGTTTTTTTTTAAAATAACAAATCAATTGTT   90
                                                                        *  F  L  L  D  I  T

                                  ......
TTATAATATCCGTACGATTCTTTGACTATGTAATAAAATGTAATCATTAGGAAGATTACGAAAAATATAAAAAATATGAGTTGTTTGTGT  180
K  Y  Y  G  Y  S  E  K  V  I  Y  Y  F  T  I  M  L  F  I  V  F  F  I  F  F  I  L  Q  K  H

ATAACAAATGCTGTAAACGCCACAATTGTGTTTGTTGCAAACAAACCCATTATTATTTGATTAAAATTGTTGTTTTTTTTGTTCATAGAC  270
I  V  F  A  T  F  A  V  I  T  N  T  A  F  L  G  M  I  I  Q  N  F  N  N  N  K  K  N  M  S

AATATTGTGTTTTGCCTAAACGTGTACTGCATAAACTCCATGCGAGTGTATAGCGAGCTAGTGGCTAACGCTTTCCCCACCAAAGTAGAT  360
L  I  T  N  Q  R  F  T  Y  Q  M  F  E  M  R  T  Y  L  S  S  T  A  L  A  K  G  V  L  T  S

                                                                                      ......
TCGTCAAAATCCTCAATTTTATCACCCTCATCCAATTTTAACATTTGGCCGTCGGAATTAACTTCTAAAGATGTGACATAATTTAATAAA  450
E  D  F  D  E  I  K  D  G  E  D  L  K  L  M  Q  G  D  S  N  V  E  L  S  T  V  Y  N  L  L

TGAAATAGAGATTCAAACGTAGCTTTATCGTCCGTTTCGACCATATCCGAAAAGAATTCGGGCATAAACTCTATGATTTCTCTAGACGTG  540
H  F  L  S  E  F  T  A  K  D  D  T  E  V  M  D  S  F  F  E  P  M  F  E  I  I  E  R  S  T

GTGTTGTCTAAACTTTCAAAGTATGCAGTCAGGAACGTGCGCGACATGTCGTCGGGAAACTCGCGCGGAAACATGTTGTTGTAACCGAAC  630
T  N  D  L  S  E  F  Y  A  T  L  F  T  R  S  M  D  D  P  F  E  R  P  F  M  N  N  Y  G  F

                                  ......
GGGTCCCATAGCGCCAAAACCAAATCTGCCAGCGTCAATAAAATGAGCACAATGCCGATGACGGAGCTGGCTTTGATAGCAATTCGAGTT  720
P  D  W  L  A  L  V  L  D  A  L  T  L  L  I  L  V  I  G  I  V  S  S  A  K  I  A  I  R  T

AACGCTTTGGCAGTCGTGGTCAGCGTTTTGATGGCAATCGCGTTGAACGAGTGCACTATCGCGGCTTTGTAAGTCTCTCCCAAGATGCGC  810
L  A  K  A  T  T  T  L  T  K  I  A  I  A  N  F  S  H  V  I  A  A  K  Y  T  E  G  L  I  R

GCGGTCACGCGCCGAGTCGTGCTAACCAACATGCGTTTCAAGGCCGGAATGAGAGTAGTGTTAATTTTTTTCAACATGCTTTTAAACCCG  900
A  T  V  R  R  T  T  S  V  L  M  R  K  L  A  P  I  L  T  T  N  I  K  K  L  M  S  K  F  G

GACATTAGCATCTCAAAGCCAGCGTCCGTAGCAATACCGAAAATGATCGCGTAATCTTCCAAAAACTTTGTTATAATTGACTCCAAATCT  990
S  M  L  M  E  F  G  A  D  T  A  I  G  F  I  I  A  Y  D  E  L  F  K  T  I  I  S  E  L  D

TGGTCGCTGATTGAACGGTCGAGCGCCTCAAAATGTTCGACATGTACACGTTCGGTACCGCGGTAATTGTATGTGATCGGAGTTTTAGTA 1080
Q  D  S  I  S  R  D  L  A  E  F  H  E  V  H  V  R  E  T  G  R  Y  N  Y  T  I  P  T  K  T

AAGCCGGTTTCGGCCGTGTACGTGATCTGGACGGGCGACCCGTCGACAATCATGCCCAAATCGTTTAGTGTTGGCCTTTTGGTAAAAAGT 1170
F  G  T  E  A  T  Y  T  I  Q  V  P  S  G  D  V  I  M  G  L  D  N  L  T  P  R  K  T  F  L

TTTTCAAATTCCAAGTCTGTAGTGGTATCGCGCACGCTGCGCCATTGCGCTAATATTGCGTTGGAGTCCACGTTGGGTCGTGGCGGTAGT 1260
K  E  F  E  L  D  T  T  T  D  R  V  S  R  W  Q  A  L  I  A  N  S  D  V  N  P  R  P  P  L

ATGCTGGAAGGCGCTCTGTAATCAAAATTGCGCAATTCGCTAAAAACGTTGTTGGCCAGCATTTTGAAAGTGACAAGAATCGTGTCGCCC 1350
I  S  S  P  A  R  Y  D  F  N  R  L  E  S  F  V  N  N  A  L  M  K  F  T  V  L  I  T  D  G

AGCACGAATCCGACGAGCGATTCCCACCATCTAAACGAACAGCCGCCGTTGAATAGATTTCTGCCGAATCGTCGACAGTAGGCTTCGTTG 1440
L  V  F  G  V  L  S  E  W  W  R  F  S  C  G  G  N  F  L  N  R  G  F  R  R  C  Y  A  E  N
```

Fig. V-14. Nucleotide sequence of the 16.92 kb HindIII C fragment of BmNPV (86.3-99.6 map units). The deduced amino acid sequences of the open reading frames are indicated below the nucleotide sequences. The twelve open reading frames are named in the parenthesis. Asterisk (*) indicates stop codon. The arrow indicates the direction of transcription. Dots (.....) indicate poly $(A)^+$ signal. Asterisks (*****) indicate late gene consensus sequence. The number symbols (####) indicate early gene consensus sequence.

```
AATTCGCCTTTAAAGCGTTCGGGAAACAAGGGATCGGGATCGGGCCGAACGTTAAAAGCCGGCACATCGTCCACGCCCATGATCGTGTGT 1530
F  E  G  K  F  R  E  P  F  L  P  D  P  D  P  R  V  N  F  A  P  V  D  D  V  G  M  I  T  H

TCTTCGGTGCGCAAGTATGGGCTGTTAAAGTACATTTTGGACAGCGAGTCCACCAAGATGCATCTGTTGTCGGGCGTGTATCTAAACTCG 1620
E  E  T  R  L  Y  P  S  N  F  Y  M  K  S  L  S  D  V  L  I  C  R  N  D  P  T  Y  R  F  E
         ......
GCAGACTGAACTTTATTTTCGGCGCCTTCACGCATGGCCGCCGCCCTGTCCAAATGGTAGCACGCTGGCTGCGCGTAACCCACTCTAGTC 1710
A  S  Q  V  K  N  E  A  G  E  R  M  A  A  A  R  D  L  H  Y  C  A  P  Q  A  Y  G  V  R  T

TCTGAGGTCTGCGTGTACATGAACGGTGTCGTGTTGGACACGACTCCGGTTTCGTGAAACGGATAGCAGCTCATGCTTTCGCACCCGCGC 1800
E  S  T  Q  T  Y  M  F  P  T  T  N  S  V  V  G  T  E  H  F  P  Y  C  S  M  S  E  C  G  R

TTGCTGAAAGTCAGTTTGACGGCCAGCGCTCTGTCGGCCAATTTCGGCGGCACATAATAATCGTCGTCACTTGACGCGGGACGCAACGTG 1890
S  A  P  R  L  T  Y  D  I  L  I  H  P  F  R  T  R  W  R  S  I  F  E  L  R  H  M  H  I  A
                                      ......
TAGTCGATCAGTATATGCGGAAACCTAGTGCGCCATCTCGAAATAAACTCGAGACGATGCATATGTATGGCATATCTACTGGCATTAGTT 1980
K  S  F  T  L  K  V  A  L  A  R  D  A  L  K  P  P  V  Y  Y  D  D  D  S  Y  R  S  A  N  T
                                                                  ......
AAATCGACGGCTGTTAAGACCGCCATGTTATATATGACTTAAATTAAACAACGATAATAATGAAATATTTATTAGATTATACTATAACAA 2070
L  D  V  A  T  L  V  A  M  <--(P74)
         ......                  ......
TACATTTGCATTTATTATGATTATATTATATTTGCATTTATTATAACAATTACATATGAGCAACCCCGTCGTCAGGACGGGTCTACCTCT 2160
                                 *****                                  ......
AACAGCGTCGTAAAAAGATAGCCGCATACAGTGCTTATTACCAAATCAGGCCTGATTCAATAACGATACATTTTATTTAGACATTGTTGT 2250
                                                                        *  V  N  N  N

TTACAATATTAATTAACTTTTTATACATTTTTAAATCATAATATATAATCATTTCGTTGTGCATTTCAAAGTTTTTGATCGCTTTAGAAT 2340
  V  I  N  I  L  K  K  Y  M  K  L  D  Y  Y  I  I  M  E  N  H  M  E  F  N  K  I  A  K  S  Y

AGTACACAAGTTTAGAGTATTCAGGAAAATGATAAACGTTGGTAAACCCGCATTTGGTACAATATAACACGGGATTTTTATAATACAGTT 2430
  Y  V  L  K  S  Y  E  P  F  H  Y  V  N  T  F  G  C  K  T  C  Y  L  V  P  N  K  Y  Y  L  K

TAGTTTTTTTACACAATTTGCAATAATTGTTAATTGTAGGTCTCGAAGGAAACGTGATTGCGCCGCCGTCCAATACCTCGGTAAATTTTT 2520
  T  K  K  C  L  K  C  Y  N  N  I  T  P  R  S  P  F  T  I  A  G  G  D  L  V  E  T  F  K  K

TGACTTTAACAGTGGCAAACACGGTTCCTTTGATACCCGAAAATCGGTTGTCTTGCAGAGCGGCCATCATTTCGCTTGGCTCTTGAAGTA 2610
  V  K  V  T  A  F  V  T  G  K  I  G  S  F  R  N  D  Q  L  A  A  M  M  E  S  P  E  Q  L  I

TAAAACAATTAACGTCGTCCGCCGCGTCGGGTCTGGTGCACATGCTTCGGTAGCGCTGCAACACTATATTGGTGTACGTTTCTCTGAGAA 2700
  F  C  N  V  D  D  A  A  D  P  R  T  C  M  S  R  Y  R  Q  L  V  I  N  T  Y  T  E  R  L  V

CGAGACCGCCGGTGGTGCTAAGATCGATCGTTTGAATGCGCTCGTTGGGCTCTTTGTGATTTCGAATTATGCGCCTAATTATTTCAAACA 2790
  L  G  G  T  T  S  L  D  I  T  Q  I  R  E  N  P  E  K  H  N  R  I  I  R  R  I  I  E  F  V

CTTTGCAATTGTGATCGTCAATTTTCAATTCTTTAACTTCCGTCGTGTGCTCCAAACTTACAGGGAAAATGTATTGGTAAAAAAACCTCT 2880
  K  C  N  H  D  D  I  K  L  E  K  V  E  T  T  H  E  L  S  V  P  F  I  Y  Q  Y  F  F  R  E

CTCTGGCTAAATAGCTGAGGTCGACCAAATTGATAGAAGGATATATTTCGTACGAGGTCTTTGGAACGTTGTGATATAGATAGCATTTTT 2970
  R  A  L  Y  S  L  D  V  L  N  I  S  P  Y  I  E  Y  S  T  K  P  V  N  H  Y  L  Y  C  K  Q

GACAGCATATGTCTATGCGGTCAGAATCGTCCAACGGCTTTTCGATGTGAACCACAACATACAAAAACCATGCGCGCGTGTTGTCTTTGA 3060
  C  C  I  D  I  R  D  S  D  D  L  P  K  E  I  H  V  V  V  Y  L  F  W  A  R  T  N  D  K  F

ATCTATAATTGCAAGTGGTGCATCGCGAATCGCTCATGTGCTGCTCCGTAGTCTTTTTGTATTTTACAGAACTGCTTGCAAATTTGCCCG 3150
  R  Y  N  C  T  T  C  R  S  D  S  M  H  Q  E  T  T  K  K  Y  K  V  S  S  S  A  F  K  G  T
```

Fig. V-14. continued.

```
TCATGCGCATATCTTTACTGTTTATATAACCCATAATGTAATTGGTGGAAAATTTTAGCGTGGCTTTCATGATGTTGCGTTCTAAATCGC 3240
  M  R  M  D  K  S  N  I  Y  G  M  I  Y  N  T  S  F  K  L  T  A  K  M  I  N  R  E  L  D  S

TCATAAAATGCATACGTAGATCACGCTCTTGTTTGAAATCCAGTTTGTCGCTGTACGCGGGCAAACTTTCCAATTTGTTACCAAACTCGG 3330
  M  F  H  M  R  L  D  R  E  Q  K  F  D  L  K  D  S  Y  A  P  L  S  E  L  K  N  G  F  E  P

GCGGCACAAAATATCCATCTTTTCTGTTGACGACTGGTTTTTTACTTACAATGCTGCTGCTGATGCTGCGCTCCAACGGATTGGCCGAAA 3420
  P  V  F  Y  G  D  K  R  N  V  V  P  K  K  S  V  I  S  S  S  I  S  R  E  L  P  N  A  S  F

AAGTGCGCGTAGGCAAAGAGATGCGCGTAGGTGGTTTGATGTTAGATTTTGGCGGCTGACGAACAGGCGACAGCGGCGAGTTGGCGACAG 3510
  T  R  T  P  L  S  I  R  T  P  P  K  I  N  S  K  P  P  Q  R  V  P  S  L  P  S  N  A  V  P

GCGTTGGCAAAGATTTGACACGGCCCTTGCCACCGGTCTTTGGCGCGTCAAAAATGTTATTCTCTCGAAAAAAACGGTTCATTGTAACTG 3600
  T  P  L  S  K  V  R  G  K  G  G  T  K  P  A  D  F  I  N  N  E  R  F  F  R  N  M  <--(ORF2)
                                                                  *****       ......

TTGGTTAGCACTCAAAAATCAACACAATACTGTGCACGTTCAGCCATCGAGAGGCTTTATATACGAAAACCTTATCTATAGAAATAAAAT 3690
                                                                  ########

TGTATATGCGTAGGAGAGCCTGGTCACAAGGCACTTTGCGCGCGACACTAGGGCTGTGGAGGGGACAGGCTATATAAAGCCCGTTTGTCC 3780
          ####                                                    *****

AACTCGTAAATCAGTATCGATTGTGTCACGGCGCACACGCTTGCTTGCGCGCTGGATAGTATAAGTGATTGATAACGAGCAACGCAACAT 3870
                                                                                (IE-0)-->  M

GATAAGAACCAGCAGTCACGTGCTGAACGTCCAGGAAAATATAATGACGTCAAACTGTGCGTCATCGCCATATTCGTGCGAGGCAACGTC 3960
  I  R  T  S  S  H  V  L  N  V  Q  E  N  I  M  T  S  N  C  A  S  S  P  Y  S  C  E  A  T  S

CGCTTGCGCAGAAGCTCAGCAGGTAATGATTGATAACTTTGTTTTCTTTCACATGTACACCGCCGACATACAAATCGACGCAAAGGTGCA 4050
  A  C  A  E  A  Q  Q  V  M  I  D  N  F  V  F  F  H  M  Y  T  A  D  I  Q  I  D  A  K  V  Q
                                                                  ......    ......

ATGCGGCGTGCGTTCGGCCGCGTTTGCAATAATTGACGATAAACATTTGGAAATGTACAAGTATAGAATAGAGAATAAATTTTTTTATTA 4140
  C  G  V  R  S  A  A  F  A  I  I  D  D  K  H  L  E  M  Y  K  Y  R  I  E  N  K  F  F  Y  Y
                                                                  ......

CTATGATCAATGTGCCGACATTGCCAAACCCGACCGTCTTCCCGATGACGACGGTGCGTGCTGTCACCATTTTATTTTCGATGCCCAACG 4230
  Y  D  Q  C  A  D  I  A  K  P  D  R  L  P  D  D  D  G  A  C  C  H  H  F  I  F  D  A  Q  R

TATTATTCAATGTATTAAAGAAATTGAAGGCGCGTACGGCGTGCGTGATCGCGGCAATGTAATAGTGTTTTATCCGTACTTGAAACAGTT 4320
  I  I  Q  C  I  K  E  I  E  G  A  Y  G  V  R  D  R  G  N  V  I  V  F  Y  P  Y  L  K  Q  L

GCGAGACGCGTTGAAGCTAATCAAAAACTCTTTTGCGTGTTGTTTTAAAAATATAAATTCTATGCAAATGTACGTGAACGAGTTAATATC 4410
  R  D  A  L  K  L  I  K  N  S  F  A  C  C  F  K  N  I  N  S  M  Q  M  Y  V  N  E  L  I  S
            ......                  ......

AAATTGCCTGTTGTTTATTGAAAAGCTGGAAACTATTAATAAAACTGTTAAAGTTATGAATTTGTTCGTAGACAATTCGGTTTTGTACGA 4500
  N  C  L  L  F  I  E  K  L  E  T  I  N  K  T  V  K  V  M  N  L  F  V  D  N  S  V  L  Y  E

ATGCAATGTTTGTAAAGAAATATCTACGGATGAAAGATTTTTAAAGCCAAAAGAACGTCGCGAATACGCTATATGCAACGCGTGCTGCGT 4590
  C  N  V  C  K  E  I  S  T  D  E  R  F  L  K  P  K  E  R  R  E  Y  A  I  C  N  A  C  C  V
                                                                  *****

TACCATGTGGAAAACGGCCACCACGCACGCAAAGTGTCCAGCGTGCAGGACATCGTATAAATAAGCACGCAACGCAAAATGAGTGGCGGC 4680
  T  M  W  K  T  A  T  T  H  A  K  C  P  A  C  R  T  S  Y  K  *     (ORF4)-->  M  S  G  G
                                      ......

GGCAACTTGTTGACTCTGGAAAGAGATCATTTTAAATATTTATTTTTGACCAGCTATTTTGATTTAAAAGATAATGAACATGTTCCTTCA 4770
  G  N  L  L  T  L  E  R  D  H  F  K  Y  L  F  L  T  S  Y  F  D  L  K  D  N  E  H  V  P  S
            ......

GAGCCGATGCGATTTATTCGCAACTATTTGAATTGCACGTTTGATTTGCTAGACGATGCCGTGCTCATGAACTATTTCAATTACTTGCAA 4860
  E  P  M  R  F  I  R  N  Y  L  N  C  T  F  D  L  L  D  D  A  V  L  M  N  Y  F  N  Y  L  Q

AGCATGCAGTTGAAACATTTGGTGGGCAGCACGTCGACAAACATTTTCAAGTTTGTAAAGCCACAATTTCGATTTGTGTGCAATCGCACA 4950
  S  M  Q  L  K  H  L  V  G  S  T  S  T  N  I  F  K  F  V  K  P  Q  F  R  F  V  C  N  R  T
```

Fig. V-14. continued.

```
ACTGTGGACATTTTAGAATTTGACACGCGCATGTACATAAAGCCCGGCACGCACGTGTACGCCACGAACCTGTTTACGTCCAATCCTCGT 5040
 T  V  D  I  L  E  F  D  T  R  M  Y  I  K  P  G  T  H  V  Y  A  T  N  L  F  T  S  N  P  R
                                                  .....

AAAATGATGGCTTTCCTGTACGCTGAATTTGGCAAGGTGTTTAAGAATAAAATATTCGTAAACATTAACAACTACGGCTGCGTGTTAGCG 5130
 K  M  M  A  F  L  Y  A  E  F  G  K  V  F  K  N  K  I  F  V  N  I  N  N  Y  G  C  V  L  A

GGCAGTGCCGGTTTTTTGTTCGACGACGCGTACGTGGATTGGAATGGTGTGCGAATGTGTGCGGCGCCGCGATTAGATAACAACATGCAT 5220
 G  S  A  G  F  L  F  D  D  A  Y  V  D  W  N  G  V  R  M  C  A  A  P  R  L  D  N  N  M  H

CCGTTCCGATTGTATCTACTGGGCGAGGACATGGCTAAGCACTTTGTCGATAATAATATACTACCGCCGCACCCTTCTAACGCAAAGACT 5310
 P  F  R  L  Y  L  L  G  E  D  M  A  K  H  F  V  D  N  N  I  L  P  P  H  P  S  N  A  K  T

CGCAAAATCAACAATTCAATGTTTATGCTAAAAAACTTTTACAAAGGTCTGCCGCTGTTTAAATTAAAGTACACGGTGGTGAACAGCACT 5400
 R  K  I  N  N  S  M  F  M  L  K  N  F  Y  K  G  L  P  L  F  K  L  K  Y  T  V  V  N  S  T
                                                                                 .....

AAAATCGTGACCCGAAAACCCAACGATATATTTAATGAGATAGATAAAGAATTGAATGGCAACTGTCCGTTTATCAAGTTTATTCAGCGC 5490
 K  I  V  T  R  K  P  N  D  I  F  N  E  I  D  K  E  L  N  G  N  C  P  F  I  K  F  I  Q  R

GACTACATATTCGACGCTCAGTTTCCGCCAGATTTGCTTGATCTGCTAAACGAATACATGACCAAAAGCTCGATCATGAAAATAATTACC 5580
 D  Y  I  F  D  A  Q  F  P  P  D  L  L  D  L  L  N  E  Y  M  T  K  S  S  I  M  K  I  I  T

AAGTTTGTGATTGAAGAAAACCCCGCTATGAACGGTGAAATGTCTCGCGAGATTATTCTTGATCGCTACTCGGTAGACAATTATCGCAAG 5670
 K  F  V  I  E  E  N  P  A  M  N  G  E  M  S  R  E  I  I  L  D  R  Y  S  V  D  N  Y  R  K

CTGTACATAAAAATGGAAATAACCAACCAGTTTCCTGTCATGTATGATCATGAATCGTCGTACATTTTTGTGAGCAAAGACATTTTGCAA 5760
 L  Y  I  K  M  E  I  T  N  Q  F  P  V  M  Y  D  H  E  S  S  Y  I  F  V  S  K  D  I  L  Q

TTGAAAGGCACTATGAACGCGTTCTACGCGCCCAAGCAGCGTATATTAAGTATTTTGGCTATAAATCGTTTGTTTGGCGCCACGGAAACT 5850
 L  K  G  T  M  N  A  F  Y  A  P  K  Q  R  I  L  S  I  L  A  I  N  R  L  F  G  A  T  E  T
                                                                              *****

ATCGACTTTCATCCCAACCTGCTCGTGTACCGGCAGAGTTCGCCGCCGGTCCGTTTGACGGGCGACGTGTATGTTGTTGATAAGAACGAA 5940
 I  D  F  H  P  N  L  L  V  Y  R  Q  S  S  P  P  V  R  L  T  G  D  V  Y  V  V  D  K  N  E
                                                        *****

AAAGTTTTTTTAGTCAAACACGTGTTCTCAAACACGGTGCCTGCATATCTTTTAATAAGAGGTGATTACGAAAGTTCGTCTGAGTTGAAA 6030
 K  V  F  L  V  K  H  V  F  S  N  T  V  P  A  Y  L  L  I  R  G  D  Y  E  S  S  S  E  L  K

TCCCTTCGCGATTTAAATCCGTGGGTTCAGAACACGCTTCTCAAATTATTGATCCCCGACTCGGTATAATAATAATATGATTTACACAGA 6120
 S  L  R  D  L  N  P  W  V  Q  N  T  L  L  K  L  L  I  P  D  S  V  *

TCCCACTACTGGCGCTACGACTAGCACAGATGTGCAGTCCGCAAACTATTTAAACAGGCTAACTCCAAACATGTTTTTGACCATTTTGGC 6210

TGTAGTAGTAATTATTGCTTTAATAATTATGTTTGTTCAATCTAACAGTAATAATGGAAACAGCTCGGGTGCTGGCTCTGGTGGTAATGG 6300

TGGCGGTGGTGGTGCTGCTACACCTCCAAGCGCGACAGGTTTTATGAATCCTTTAAACGCTACCATGCGAGCTAATCCCTTTATGAACAC 6390
                  *****
GCCTCAAAGGCAAATGTTGTAGATAAGTGTATAAAAATGAAACGTGTCAAATGCAACAAAGTTCGAACGGTCACCGAAATTGTAAACAGC 6480
                        (ORF5)--> M  K  R  V  K  C  N  K  V  R  T  V  T  E  I  V  N  S

GATGAAAAGATCCAAAAAACCTACGAATTGGCCGAATTTGATTTAAAAAATCTAAGCAGTTTAGAAAGCTATGAAACTCTGAAAATTAAA 6570
 D  E  K  I  Q  K  T  Y  E  L  A  E  F  D  L  K  N  L  S  S  L  E  S  Y  E  T  L  K  I  K

TTGGCGCTCAGCAAATACATGGCTATGCTCAGCACCCTGGAAATGACCCAGCCGCTGTTGGAAATATTTAGAAACAAAGCAGACACTCGG 6660
 L  A  L  S  K  Y  M  A  M  L  S  T  L  E  M  T  Q  P  L  L  E  I  F  R  N  K  A  D  T  R
```

Fig. V-14. continued.

```
CAAATTGCCGCCGTGGTGTTTAGCACATTAGCTTTTATACACAATAGATTCCATCCACTTGTTACTAATTTTACTAACAAAATGGAGTTT 6750
 Q  I  A  A  V  V  F  S  T  L  A  F  I  H  N  R  F  H  P  L  V  T  N  F  T  N  K  M  E  F

GTGGTCACTGAAACTAATGACACAAGCATTCCCGGAGAACCCATTTTGTTTACAGAAAACGAGGGGGTGCTGTTGTGCTCCGTCGACAGA 6840
 V  V  T  E  T  N  D  T  S  I  P  G  E  P  I  L  F  T  E  N  E  G  V  L  L  C  S  V  D  R

CCGTCTATCGTTAAAATGCTAAGCCGCGAGTTTGACACCGAGGCTTTAGTAAACTTTGAAAACGATAACTGCAACGTGCGGATAGCCAAG 6930
 P  S  I  V  K  M  L  S  R  E  F  D  T  E  A  L  V  N  F  E  N  D  N  C  N  V  R  I  A  K
                                                        ......

ACGTTTGGCGCCTCTAAGCGCAAAAACACGACGCGCAGCGATGATTACGAGTCAAATAAACAACCAGATTACGATATGGATTTGAGCGAT 7020
 T  F  G  A  S  K  R  K  N  T  T  R  S  D  D  Y  E  S  N  K  Q  P  D  Y  D  M  D  L  S  D

TTTAGCATAACTGAGGTTGAAGCCACTCAATATCTAACTCTGTTGCTGATCGTCGAACATGCCTATTTGCATTATTATATTTTTAAAAAT 7110
 F  S  I  T  E  V  E  A  T  Q  Y  L  T  L  L  L  I  V  E  H  A  Y  L  H  Y  Y  I  F  K  N

TACGGGGTGTTTGAATATTGCAAGTCGCTAACGGACCATTCGCTTTTTACCAACAAATTGCGATCGACAATGAGCACAAAAACGTCTAAT 7200
 Y  G  V  F  E  Y  C  K  S  L  T  D  H  S  L  F  T  N  K  L  R  S  T  M  S  T  K  T  S  N
                                                  ......

TTACTGTTAAGCAAATTCAAATTTACCATTGAAGATTTTGACGAAATAAATTCAAATTCTGTAACATCAGGCTTTAATATATATAATTTT 7290
 L  L  L  S  K  F  K  F  T  I  E  D  F  D  E  I  N  S  N  S  V  T  S  G  F  N  I  Y  N  F
......

AATAAATAATTAAATAATATACAATGCTGTTATTAATTATATTTTTAATATTAATTAAAGTATTAATATTTAAAAAAATGAATCAAACTC 7380
 N  K  *

ATCTAAAGTGTCACAGCGATAAAATTTGTCCTAAAGGGTATTTTGGCCTCAACGCCGACCCCTATGATTGCACTGCGTATTACATCTGTC 7470

CGCATAAAGTGCAAATGTTTTGTGAATCAAATCACGAATTCGACTTGGACTCGGCCAGCTGCAAGCCTATCGTGTACGATCGCACGGGCA 7560

GCGGGTGTGCGGCTCGCATGTATAGAAACTTGTTATTATGAAGAGTGGCTTTCTAGTTGCACAACACTATTATCGATTTGCAGTTCGGGA 7650
                                  *  S  S  H  S  E  L  Q  V  V  S  N  D  I  Q  L  E  P
                                                        ......

CATAAATGTTTAAATATATCAATGTCTTTGTGATGCGCGCGACATTTTTGTAAGTTATTAATAAAATGCACTGACACGTTGCCCGACATT 7740
 C  L  H  K  F  I  D  I  D  K  H  H  A  R  C  K  Q  L  N  N  I  F  H  V  S  V  N  G  S  M

ATCATTAAATCCTTGGCGTAGAATTTGTCGGGTCCGTTGTCCGTGTGCGCTAGCATGCCCGTAACGGACCTTGTGCTTTTGGCTTCAAAG 7830
 I  M  L  D  K  A  Y  F  K  D  P  G  N  D  T  H  A  L  M  G  T  V  S  R  T  S  K  A  E  F

GTTTTGCGCACAGACAAAATGTGCCACACTTGCAGCTCTGCTTGTGTGCGCGTTACCACAAATCCCAACGGCGCAGTGTACTTGTTGTAT 7920
 T  K  R  V  S  L  I  H  W  V  Q  L  E  A  Q  T  R  T  V  V  F  G  L  P  A  T  Y  K  N  Y
   ......

GTAAATAAATCTCGATAAAGGCGCGGCGCGCGAATGCAGCTGATCACGTACGCTCCTCGTGTCCCGTTCAAGGACGGTGTTATCGACCTC 8010
 T  F  L  D  R  Y  L  R  P  A  R  I  C  S  I  V  Y  A  G  R  T  G  N  L  S  P  T  I  S  R
AGATTAATATTTATCGGCCGACTGTTTTCGTATCCGCTCACCAAACGTGTTTTTGCATTAACATTGTATGTCGGCGGATGTTCTGTATCT 8100
 L  N  I  N  I  P  R  S  N  E  Y  G  S  V  L  R  T  K  A  N  V  N  Y  T  P  P  H  E  T  D
      ..........                                ......                                  ######

AATTTGAATAAATAAATGATAACCGCATTGGTTTTAGAGGGCATAATAAAAAAAAATATTATTATCGTGTTCGCCATTAGGGCAGTATAAA 8190
 L  K  F  L  Y  I  I  V  A  N  T  K  S  P  M  I  F  F  I  N  N  D  H  E  G  N  P  C  Y  L
 ##                    ####
TTGACGTTCATGTTGAATATTGTTTCAGTTGCAAGTTGACATTGGCGGCGACACGATCGTGAACAACCAAACGACTATGACGCAAATTAA 8280
 N  V  N  M  <--(ORF6)                                                  (IE-1)-->  M  T  Q  I  N

TTTTAACGCGTCGTACACCAGTGCTCCGACGCCGTCCCGAGCGTCGTTCGACAACGGCTATTCAGAGTTTTGTGATAAACAACAGCCCAA 8370
   F  N  A  S  Y  T  S  A  P  T  P  S  R  A  S  F  D  N  G  Y  S  E  F  C  D  K  Q  Q  P  N
```

Fig. V-14. continued.

```
CGACTATTTGAATTATTATAACAATCCCACGCCGGATGGAGCCGACACGGTAGTATCTGACAGCGAGACTGCAGCAGCTTCAAACTTTTT 8460
  D  Y  L  N  Y  Y  N  N  P  T  P  D  G  A  D  T  V  V  S  D  S  E  T  A  A  A  S  N  F  L
                                                                                          *

GGCAAGCGTCAATTCGTTAACTGATGATAACGATATAATGGAATGTTTGCTCAAGACCACTGATAATCTCGGAGAAGCAGTTAGTTCTGC 8550
  A  S  V  N  S  L  T  D  D  N  D  I  M  E  C  L  L  K  T  T  D  N  L  G  E  A  V  S  S  A
****                                                  *****

TTATTATTCGGAATCCCTTGAGCTGCCTGTTGCGGAGCAACCATCGCCCAGTTCTGCTTATAATGCGGAATCTTTTGAGCAGTCTGTTGG 8640
  Y  Y  S  E  S  L  E  L  P  V  A  E  Q  P  S  P  S  S  A  Y  N  A  E  S  F  E  Q  S  V  G
                                                                                       ..

TGTGAACCAACCATCGGCAGCTGGAACTAAACGGAAGCTGGACGAATACTTGGACGATTCACAAAGTGTGGTGGGCCAATTTAACAAGAA 8730
  V  N  Q  P  S  A  A  G  T  K  R  K  L  D  E  Y  L  D  D  S  Q  S  V  V  G  Q  F  N  K  N
....

TAAATTGAAGCCTAAATACAAGAAAAGCACAATTCAAAGCTGTGCAACCCTTGAGCAGACAATTAATCACAACACGAACATTTGCACGGT 8820
  K  L  K  P  K  Y  K  K  S  T  I  Q  S  C  A  T  L  E  Q  T  I  N  H  N  T  N  I  C  T  V

CGCTTCAACTCAAGAAATTACGCATTATTTTACTAATGATTTTGCGCCGTATTTGATGCGTTTCGACGACAACGACTACAATTCCAACAG 8910
  A  S  T  Q  E  I  T  H  Y  F  T  N  D  F  A  P  Y  L  M  R  F  D  D  N  D  Y  N  S  N  R

GTTCTCCGACCATATGTCCGAGACTGGTTATTACATGTTTGTGGTTAAAAAAAGTGAAGTAAAGCCGTTTGAAATTATATTTGCCAAGTA 9000
  F  S  D  H  M  S  E  T  G  Y  Y  M  F  V  V  K  K  S  E  V  K  P  F  E  I  I  F  A  K  Y

CGTGAGCAATGTGGTGTACGAATATACAAACAACTATTACATGGTAGATAATCGTGTGTTTGTGGTAACGTTTGATAAAATTAGATTTAT 9090
  V  S  N  V  V  Y  E  Y  T  N  N  Y  Y  M  V  D  N  R  V  F  V  V  T  F  D  K  I  R  F  M

GATTTCGTACAATTTGGTTAAAGAAACCGGCATAGAAATTCCTCATTCTCAAGATGTGTGCAACGACGAGACGGCTGCACAAAATTGTAA 9180
  I  S  Y  N  L  V  K  E  T  G  I  E  I  P  H  S  Q  D  V  C  N  D  E  T  A  A  Q  N  C  K

AAAATGCCATTTTGTCGATGTGCATCACACGTTTAAAGCTGCTCTGACTTCATATTTTAATTTAGATATGTATTACGCGCAAACCACATT 9270
  K  C  H  F  V  D  V  H  H  T  F  K  A  A  L  T  S  Y  F  N  L  D  M  Y  Y  A  Q  T  T  F
                                                                                      ......

TGTGACTTTGTTACAATCGTTGGGCGAAAGAAAGTGTGGGTTTCTTTTGGGCAAGTTGTACGAAATGTATCAAGATAAAAATTTATTTAC 9360
  V  T  L  L  Q  S  L  G  E  R  K  C  G  F  L  L  G  K  L  Y  E  M  Y  Q  D  K  N  L  F  T

TTTGCCTATTATGCTTAGTCGTAAAGAGAGTAATGAAATTGAGACTGCATCTAATAATTTTTTTGTATCGCCGTATGTGAGTCAAATATT 9450
  L  P  I  M  L  S  R  K  E  S  N  E  I  E  T  A  S  N  N  F  F  V  S  P  Y  V  S  Q  I  L

AAAGTATTCGGAAAGCGTAAAGTTTCCCGACAATCCCCCAAACAAATATGTGGTGGACAATTTAAATTTAATTGTTAACAAAAAAAGTAC 9540
  K  Y  S  E  S  V  K  F  P  D  N  P  P  N  K  Y  V  V  D  N  L  N  L  I  V  N  K  K  S  T

GCTCACGTACAAATACAGTAGTGTCGCTAATCTTTTGTTTAATAATTATAAATATCATGACAATATTGCGAGTAATAATAACGCGGAAAA 9630
  L  T  Y  K  Y  S  S  V  A  N  L  L  F  N  N  Y  K  Y  H  D  N  I  A  S  N  N  N  A  E  N

TTTAAAAAAGGTTAAGAAGGAGGACGGCAGCATGCACATTGTCGAACAGTATTTGACTCAGAATGTGGATAATGTAAAAGGTCACAATTT 9720
  L  K  K  V  K  K  E  D  G  S  M  H  I  V  E  Q  Y  L  T  Q  N  V  D  N  V  K  G  H  N  F
                                                                  ......

TATAGTATTGTCTTTCAAAAACGAAGAGCGGTTGACTATAGCTAAGAAAAACGAAGAGTTTTATTGGATTTCTGGCGAGATTAAAGATGT 9810
  I  V  L  S  F  K  N  E  E  R  L  T  I  A  K  K  N  E  E  F  Y  W  I  S  G  E  I  K  D  V

AGACGCTAGTGCAAGTAATTCAAAAATATAATAGATTTAAGCATCACATGTTTGTAATCAGTAAAGTGAACCGAAGAGAGAGCACTACAT 9900
  D  A  S  A  S  N  S  K  I  *

TGCACAATAATTTGTTAAAATTGTTAGCTTTAATATTACAGGGTCTGGTTCCGTTGTCCGACGCTATAACGTTTGCGGAACAAAAACTAA 9990

ATTGTAAATATAAAAAAATTGAATTTAATTAATTATACATATATTTTGAATTTAATTAATTATACATATCTTTTGAATTTAATTAATTAT 10080
```

Fig. V-14. continued.

```
                                .....
ACATATATTTTATATTATTTTTGTCTTTTATTATCGACGAGGGGCCGTTGTTGATGCGGGGTGTTGCATAATAACAATGGGAGTTGGTGC 10170
                                * R R P A T T S A P H Q M I V I P T P A

                                                                .....
GCCACCGCTTCCTCCTCCTCCTCCTCCTCCTTTTGTCATGTATCTGTAGATAAAATAAAGTATTAAACCTAAAAACAAGACCGCGCCTAT 10260
 G G S G G G G G G G K T M Y R Y I F Y L I L G L F L V A G I

CATCATAATGATGGGCATTATTTTGTTGCGGATGCCGTCACTACTGTTGGACGATTTGCCGACTAAACCTTCTCTTCCCAGTAACCAATC 10350
 M M I I P M I K N R I G D S S N S S K G V L G E R G L L W D

TAAACCCAAGTCGCCAACTAAATCACTAAACGAGTAAGGTTCGATGCACATGATTGTTTGGCCCGCAGGAAGATCGCTGATATCTACGTA 10440
 L G L D G V L D S F S Y P E I C M I T Q G A P L D S I D V Y

TTGAGGCGAATTTGGGTTAGCGGCCGGATTGCTGCCGCGACAAACTGTTTTTTCTGTTTCATAGTTAAATCCTTGGCACATGTTGGTTAG 10530
 Q P S N P N A A P N S G R C V T K E T E Y N F G Q C M N T L

TAGGGGCGAATCGTTAGCCAACAAGGGGTCTCTTGAGCAAATGTTAACATCCGACTGAGCTAGATTGCGGTCTTGACGACAAGTGCGCTG 10620
 L P S D N A L L P D R S C I N V D S Q A L N R D Q R C T R Q

CAATAACAAACAGGACTCGACGTTTTCTCCGGCGTTTCTACCTTGCACATAATAACTTCCGCCGGTTCTATTGATGGCGTTGATTATATC 10710
 L L L C S E V N E G A N R G Q V Y Y S G G T R N I A N I I D

TTGTACTAATGTGGAGGCGGTAAACAAAAGATAACCGCCGCCGGCCAAGAGTATGCCCACTCCTGCTACTTTCAAGGTTCTCATGTGATT 10800
 Q V L T S A T F L L Y G G G A L L I G V G A V K L T R M H N

ATGTAAACGGGGGTTTTGCTGCAGTGCGTTTTGAACACCTTCCGGCGTGCGCACGTTGGTCTCTGGGAAGTTTTGTCTGACTGCATTGGA 10890
 H L R P N Q Q L A N Q V G E P T R V N T E P F N Q R V A N S

                                                        .....
TCGCGTCTGTTTGGTGTGGTAATGAAAGTCTGGCACGTTGTCCATGCGCCGCAATTGGCGCAATGAGTTTATTTGAGGGTCTGAAATGCC 10980
 R T Q K T H Y H F D P V N D M R R L Q R L S N I Q P D S I G

CTGAAATACATTGCGTATGTTGGGGACATCGTTGTTACGAGTGATTCTGTTTATGTCTGAAGTGCTCACAAACCGGTTGTTAGATAATTG 11070
 Q F V N R I N P V D N N R T I R N I D S T S V F R N N S L Q

ATAGCCCGGCTGATATCTGTTGTTTCCAAGGTTGCGTACACTGGGCGCGTTGAGCACATTTGTGAAACCGGCGGGAGTGCTTGTTAAAAG 11160
 Y G P Q Y R N N G L N R V S P A N L V N T F G A P T S T L L

                                        .....                    *
ACGCGTATTATCAGCAAGAAAACTGGCCTGATTAGGATACAATTTATTGACTCTACGAAGATTTGTAAAAAAACTCATTTTAAAGCAAAC 11250
 R T N D A L F S A Q N P Y L K N V R R L N T F F S M <--(ORF8)
****  .....
TTATTTAATAAATATATCACAGTAAAGGTTTTGCAAAATTGCCGTCGTCAATACAACACGGCTGCGGCGCCATGTTGGTAAAATCTAATC 11340

TTCTCCTTGCTTTAGATTTTGGGCGAGAGGGCGCATTTGTTGTGTCAGTCATTTCGACGTCTGCATTATTTGTTGTGTAAGGTACTTCAA 11430

TGTATGAAGCGACTTTAACATTATTATAATTTTTTTTAAATATTGATGCGTTCCACGGCGCTCGTTGATACGGATGATATCTCTCCATTG 11520

                  .....                                 .....
TATGATCGCTAAAGTTATATACCGTTTCAATAAATATGTTCAAACCCAACATGATTAGAATATACATAATAGTTTATTTGTTTTTTATAA 11610
     .....
TTATTTTTATTGTTTTGAAATCTAAAAGAGGTAACGATAACGATAACGATAACGATAACGATAACGATAACGATTCAGACCACGAGTTCA 11700

GTTGCGTTAACAGACCAATTGGAGTACATTTTCCGCATCCTACTAAATGTAATGCTTTTTACATGTGTGTCGGTATAAATCATAGATTAG 11790

AGTTACTTTGTTCTGAAGGATTTGAATTTGATCCAAATGTTAAAGATTGTGTTCCTATATCAGATTATGGATGTACTGCTAACCAAAACA 11880
```

Fig. V-14. continued.

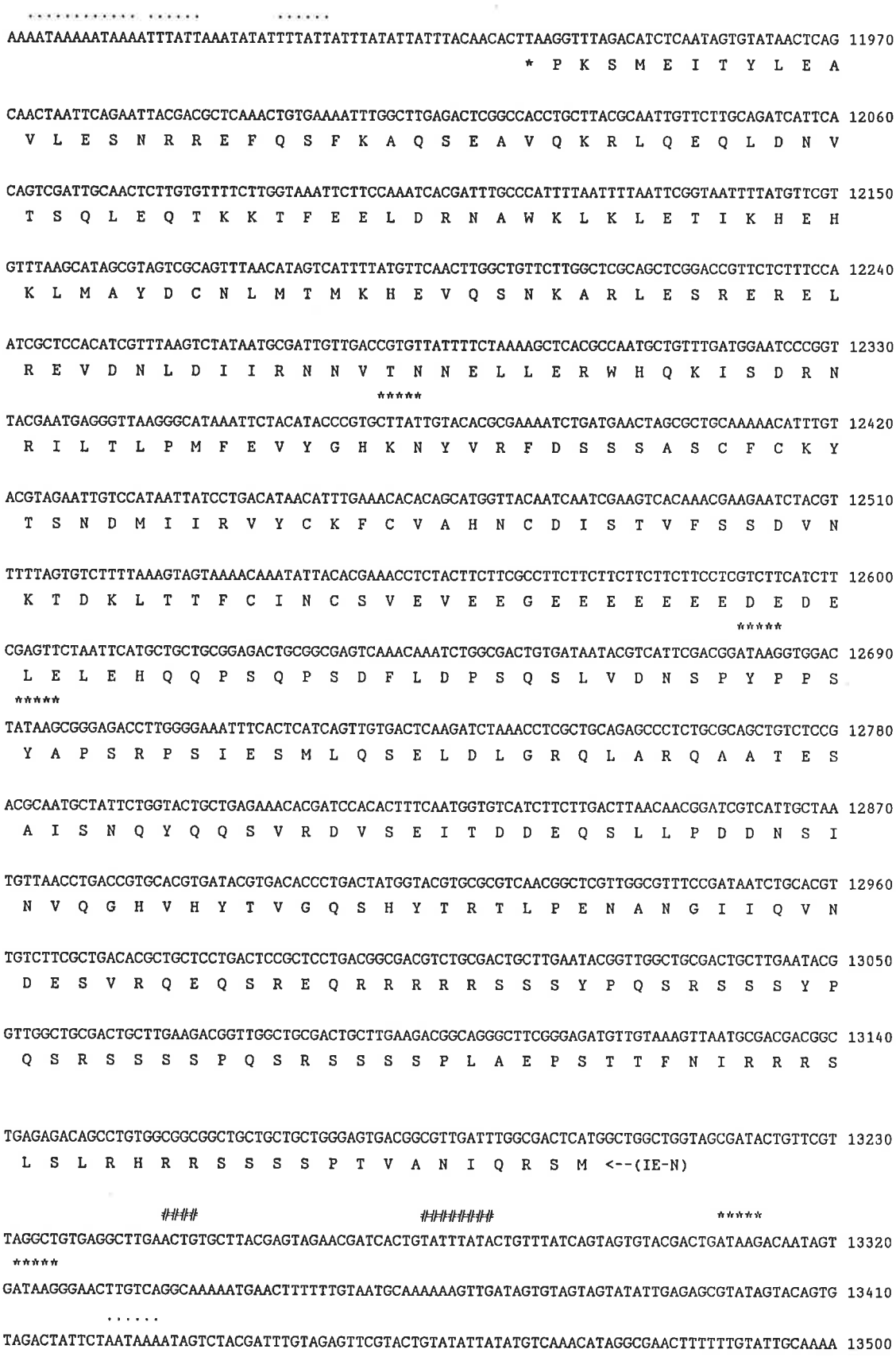


Fig. V-14. continued.

```
AAATTCATTTTTTAATTTATCATATCACAGGCTTCAGTACTGTTATCTGTCCCCACTCAGGCGTACAGCTATAACATATCACAGGCTTCA 13590
                                  #######                         ####
GTTTCTGTTATCTGTCCCCACTCAGGCGTACAGCTATAAAAGCAGACACTTGTCAACTCGTAAGCACAGTTTGTTGCGGAGAGCCTGCCA 13680
*****
ATAAGCAAAATGCCAAGGGACACCAACAATCGCCGCCGCCACCGGTCTACGCCATATGAACGTCCTACGCTTGAAGATCTCCACAGACAG 13770
(PE-38)--> M  P  R  D  T  N  N  R  R  R  H  R  S  T  P  Y  E  R  P  T  L  E  D  L  H  R  Q
                                                                                  ......
TTGGAAGACGCTTTGGAGAGGCGCTATCAAATGCGTAGAAGGCAGCGTCAAAACCGGCTCCGCACCATACAAATAAAACAGCAGCGAATG 13860
 L  E  D  A  L  E  R  R  Y  Q  M  R  R  R  Q  R  Q  N  R  L  R  T  I  Q  I  K  Q  Q  R  M

ATGGCGGAATTAAAGAAAGAGCCGGTAATTAATTTTAAATTTGAGTGCAGTGTGTGTTTCGAAACATATTCTCAACAATCTAACGATACT 13950
 M  A  E  L  K  K  E  P  V  I  N  F  K  F  E  C  S  V  C  F  E  T  Y  S  Q  Q  S  N  D  T

TGTCCTTTTTTGATTCCGACTACGTGCGACCACGGTTTTTGTTTCAAATGCGTCATCGATCTGCAAAGCAACGCGATGAATATTCCACAT 14040
 C  P  F  L  I  P  T  T  C  D  H  G  F  C  F  K  C  V  I  D  L  Q  S  N  A  M  N  I  P  H

TCAATTGTGTGCTGTCCATTGTGCAATACCCAGGTAAAAATGTGGCGTTCTTTAAAACCTAACGCTGTTGTGACGTGTAAGTTTTACAAG 14130
 S  I  V  C  C  P  L  C  N  T  Q  V  K  M  W  R  S  L  K  P  N  A  V  V  T  C  K  F  Y  K

AAAACTCAAGAAAGAGTTCCGGCCGTGCAGCAATATAAAAACATTATCAAAGTGCTACAAGAACGGAGCGTGATTAGTGTCGAAAACAGC 14220
 K  T  Q  E  R  V  P  A  V  Q  Q  Y  K  N  I  I  K  V  L  Q  E  R  S  V  I  S  V  E  N  S

GACAACAATTGTGACATAAATATGGAGAATCAGGCAAAAATAGTTGCTTTGGAAGCTGAATTAAAGAACGAAAAAAATCACAGTGATCAA 14310
 D  N  N  C  D  I  N  M  E  N  Q  A  K  I  V  A  L  E  A  E  L  K  N  E  K  N  H  S  D  Q

GTAACTTCTGAAAACCGACAGCTAATAGAGGAAAATACTCGTCTCAACGAACAGGTTCAAGAATTGCAGCGTCAGGTGAGGACATTGGCG 14400
 V  T  S  E  N  R  Q  L  I  E  E  N  T  R  L  N  E  Q  V  Q  E  L  Q  R  Q  V  R  T  L  A

CCGCAACGTGGCATTACGGTTAATCCGCAAATAGGCCGTGACGACCGTGCGCCAGCCGAGCTGAACGAGCGTTTTCGCTCACTCGTCTAT 14490
 P  Q  R  G  I  T  V  N  P  Q  I  G  R  D  D  R  A  P  A  E  L  N  E  R  F  R  S  L  V  Y
                  ......
TCGACTATTTCAGAACTGTTTATTGAAAATCGCGTTCATAGTATTCAAAATTATGTTTATGCCGGAACTTCTGGTGCTGCTAGTTCATGT 14580
 S  T  I  S  E  L  F  I  E  N  R  V  H  S  I  Q  N  Y  V  Y  A  G  T  S  G  A  A  S  S  C
                                                                                  .....
GATGTAAATGTTACTGTTAATTTTGGGTTTGAAAATTAATGTGATATGTATGTATAAATGTATATATAAAAATGCTGAAACAACTAATAA 14670
 D  V  N  V  T  V  N  F  G  F  E  N  *
.    ......      ......
ACGTTTTTTATTTGTTTTTTTTATTTTATGTGATTAAGAAACTTTTAACATGGATAGTAGTAATTGTATTAAAATAGATGTAAAATACCA 14760

TATGCCGTTACATTATCAATGTGACATTAACGCAGATAAAAACGTTGTAAATGCGTATGACGCTATCGATGTTGACCCCAACAAAAAATT 14850

TATAATTAATCATAATCACGAACAAGTCGATGAAACAAACAAACAAGAAGTCGTCGATAAAACTGACGCAACAACATACAATTCTTGCAT 14940

CATAAAAATTTAAATGATATTATAATTTAAAATAACAATGACATCATCGTTTGATTGTGTTTTACACGTAGAATTCTACTCGTAAAGCCA 15030

GTTCAGTTTTGAAAAACAAATGACATCACCTCTTGATTATGTTTTACACGTAGAATTCTACTCGTAAAGCCAGTTCAGTTTTGAAAAACA 15120

                                                                                  ....
AATGACGTCATTTCTTGATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCAGTTTTGAAAAAACAAATGACATCAATTTTTA 15210
..    ......
TTTTTATAATAAATGACATCATTTCTTGATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGGGTTCAGTTTTGAAAAACAAGTGACA 15300

TCATTTCTTAAATTAAGTTTTGAAAAACAAATGACATCACCTTTCGATCATGTTTTACACGTAGAATTCTACTCGTAAAGCGAGTTCAGT 15390

TTTGAAAAACAAGTGACATCATTTCTTAAATTAAGTTTTGAAAAACAAATGACGTCATTTCTTGATCATGTTTTACACGTAGAATTCTAC 15480
```

Fig. V-14. continued.

```
TCGTAAAGCGAGTTCAGTTTTGAAAAACAAATGACATCATTTTTTTGATTATGTTTTACAAGTAGAATTCTACTCGTAAAACGAGTTCGG 15570

                                *****
TTATGAGCCGTGTGCAAAAAATGACATCAGCTTATGACATCACCCACTGATCGTGCGTTACAAGTATAATTCTACTCGTAAAGCGAGTAC 15660
                        *****
ATATTTAGTTACGTTTCTGAGATAAGATTGAAAGCACGTGTAAAATGTTTCCCGCGCGTTGGCACAACTATTTACAGTGCGGCCAAGTTA 15750
                                  (ORF11)--> M  F  P  A  R  W  H  N  Y  L  Q  C  G  Q  V

TAAAAGATTCTAATCTAATATGTTTTAAAACACCTTTGCAACCCGAGTTGTTTGCGTACGTGACTAGCGAAGAAGATGTGTGGACCACAG 15840
I  K  D  S  N  L  I  C  F  K  T  P  L  Q  P  E  L  F  A  Y  V  T  S  E  E  D  V  W  T  T

AACAGATAGTAAAACAAAACCCCAGTATTGGAGCAATAATCGATTTAACCAACACGTCTAAATATTATGATGGTGTGCACTTTTTGCGGG 15930
E  Q  I  V  K  Q  N  P  S  I  G  A  I  I  D  L  T  N  T  S  K  Y  Y  D  G  V  H  F  L  R
                                                                      ......
CGGGCCTGTTATACAAAAAAATTCAAGTACCTGGCCAGACTTTGCCGTCTGAAAGCATCGTTCAAGAATTTATTGACACGGTAGAAGAAT 16020
A  G  L  L  Y  K  K  I  Q  V  P  G  Q  T  L  P  S  E  S  I  V  Q  E  F  I  D  T  V  E  E

TTACAGAAAAGTGTCCCGGCATGTTGGTGGGCGTGCATTGCACACACGGCATTAATCGCACCGGTTACATGGTGTGCAGATATTTAATGC 16110
F  T  E  K  C  P  G  M  L  V  G  V  H  C  T  H  G  I  N  R  T  G  Y  M  V  C  R  Y  L  M
                                                                                        ..
ACACCCTGGGTATTGCGCCGCAGGAAGCCATAAATAGATTTGAAAAAGCCAGAGGTCACAAAATTGAAAGACAAAATTACGTTCAAGATT 16200
H  T  L  G  I  A  P  Q  E  A  I  N  R  F  E  K  A  R  G  H  K  I  E  R  Q  N  Y  V  Q  D
....                                    ......
TATTAATTTAATTAATATTATTTGCATTTTTTAACAAATACTTTATTCTATTTTCAAATTGTTGCGCTTCTTCTAGCGAATCAAAACTAA 16290
L  L  I  *
      *  N  I  N  N  A  N  K  L  L  Y  K  I  R  N  E  F  Q  Q  A  E  E  L  S  D  F  S  I

TGCTTCGCTTGCTACGTTTACTTTGTAGCCGATCAGTGGCGTTGTTCCAATCGACGGTAGGATTAGGCCGGATATTCTCCACCACAATGT 16380
 S  R  K  S  R  K  S  Q  L  R  D  T  A  N  N  W  D  V  T  P  N  P  R  I  N  E  V  V  I  N

TGGCAGCGTCGATGTTACGTTTGTGCTTTTGGCTTTCCACGTACGTCTTTTGGCCGGTAATAGCCGTAAACGTAGTGCCGTCGCGCGTCA 16470
 A  A  D  I  N  R  K  H  K  Q  S  E  V  Y  T  K  Q  G  T  I  A  T  F  T  T  G  D  R  T  V

CGCACAGCACCGGATGTTTGCGCTTGTCCGCGGGGTATTGAACCGCGCGATCCGACAAATCTATCATTTTGGCAATTAAATTGGAGACCT 16560
 C  L  V  P  H  K  R  K  D  A  P  Y  Q  V  A  R  D  S  L  D  I  M  K  A  I  L  N  S  V  Q

GCGTGTCTTTAACCTGCAACAACTCGTCTTTTTTGTGCATCATTTCATCTTTTCTCTGCATAGTTTCCTGAAACCCGGTGTACATGCGGT 16650
 T  D  K  V  Q  L  L  E  D  K  K  H  M  M  E  D  K  R  Q  M  T  E  Q  F  G  T  Y  M  R  N

TGAGGTCGGTCATCACGCGCGTCACTTGCAAGTCTTTGGCCTCAATCTGCTTGTCCTTAATTGCGATGATGCGTTCGATAAACTCTTGCT 16740
 L  D  T  M  V  R  T  V  Q  L  D  K  A  E  I  Q  K  D  K  I  A  I  I  R  E  I  F  E  Q  K

TTTTAAAAAGTTCCTCGGTTTTTTGCGCCACCACCGCTTGCAGCGCGTTTGTGTGCTCGGTAAATGTCGCAATCAGTTTAGTCACCAACT 16830
 K  F  L  E  E  T  K  Q  A  V  V  A  Q  L  A  N  T  H  E  T  F  T  A  I  L  K  T  V  L  Q

GTTTGTTCTTTTCCTCCTGTTGTTTAATTGCGGGGTCGTACTTGCCCGTGCATAGCACTTGAGGAATTACTTCTTCTAAAAGCCATTCTT 16920
 K  N  K  E  E  Q  Q  K  I  A  P  D  Y  K  G  T  C  L  V  Q  P  I  V  E  E  L  L  W  E  Q

GTAATTCTATGGCGTACGGAAGCTT 3' 16945
 L  E  I  A  Y  P  L  K
```

Fig. V-14. continued.

VI General Discussion and Conclusion

Expression systems using baculoviruses especially NPVs have recently become very popular for basic research and its applications to agriculture and medicine. Recent advances in molecular biological techniques have also allowed the genome of viruses and living organisms to be manipulated. Several NPVs including BmNPV and AcNPV possess advantageous characteristics which allow these types of manipulations. BmNPV and AcNPV have established cell lines which support sufficient viral replication and allow the isolation of pure clones by plaque assay. They also possess a viral genome composed of circular double-stranded DNA, which can be easily handled by modifying enzymes. Furthermore, NPV possess a unique gene, the polyhedrin gene, whose gene product can aggregate and form proteinacious crystals which encapsulate many viral progenies. The polyhedrin gene is not essential for production of viral particles, and has a very strong promoter activated at a late stage of infection. Crystals of this gene product, polyhedra, are visible under the light microscope, and can be used as a marker for insertion of foreign genes. These characteristics have made baculoviruses extremely useful as gene expression vectors.

In contrast to other DNA viruses with similar or larger genomes, such as the vaccinia virus (180 kb genome), molecular biological knowledge of NPVs is limited. Nucleotide sequencing is one of the most powerful techniques

to study viruses at the DNA level. Baculoviruses possess genomes of about 130 kb in length, however, sequence information of only 40% of the AcNPV, 10% of O. pseudotsugata NPV, and 1-5% of several other NPVs including BmNPV is available. My thesis problem has focused on the clarification of the genome structure of baculoviruses. BmNPV was chosen because it is a major disease in sericulture and is currently used as a highly efficient vector for the expression of foreign genes. The BmNPV expression system in particular has an additional advantageous characteristic of having a useful larval host, the silkworm, for expression in vivo. Another advantage of using BmNPV for basic research is the availability of AcNPV sequence information for comparison. SlNPV was also analyzed as part of my thesis problem. This virus has been extensively studied for application in the control of a major pest of vegetable crops in Japan. Little information of the molecular biology of SlNPV is available (see Maeda et al., 1990).

From this starting point, gene libraries covering the entire genomes of the two viruses were initially constructed using plasmid vectors. By hybridization and double-digestion experiments, complete physical maps for several restriction enzymes of the two viral genomes were constructed. This basic information has contributed greatly to the molecular biological studies of this thesis. This information will also be very useful for basic research and

its applications in the future by other scientists. Using the gene library and associated information, the characteristics of the polyhedrin gene and gene structure of NPVs were studied extensively.

In section IV, the polyhedrin gene was analyzed by various procedures. Studies of the sorting mechanisms of proteins localized in nucleus and the crystallization of polyhedra and amino acid sequence relationships were conducted. Four different polyhedrin genes of BmNPV and SlNPV were isolated, sequenced, and characterized. Amino acid sequence substitutions were found to be responsible for changes in the polyhedral structure, e.g. polyhedral to cuboidal. By analysis of sequence information of the polyhedrin genes of two isolates (OT2 and CC5) of SlNPV, it was speculated that OT2 belongs to AcNPV and CC5 belongs to a new virus group of SlNPV, which has some relatedness to O. pseudotsugata NPV. These studies indicate that sequence analyses offer an efficient approach to viral systematics and evolutionary studies. Based on the studies of the effects of the 5′ non-translating region of the polyhedrin gene for foreign gene expression (Horiuchi et al., 1987), new transfer vectors with high expression characteristics and with multiple cloning sites were constructed. This type of transfer vector is currently used for high expression of foreign genes all over the world.

To utilize the polyhedrin gene, the basic scientific question of, "why does the nucleus contain specific

proteins?", was studied. Protein localization in specific organelles in eukaryotic cells is very important. Recent molecular biological studies have revealed that proteins have their own internal signals for localization. Two major hypothesis have been proposed for the mechanisms of nuclear localization of proteins. However, the true mechanism is still unclear. I chose the polyhedrin gene for this study because the polyhedrin gene is highly expressed, modified easily, and non-essential for viral production. A specific consensus nuclear localization signal found in other eukaryotic cells was also found in the BmNPV polyhedrin gene. These results showed that insect cells follow the same protein trafficking rules within insect cells as found in other eukaryotic cells. Furthermore, crystallization of polyhedra was found to be controlled by most of the amino acid sequence of polyhedrin. This result is consistent with data showing high conservation of amino acid sequences of polyhedrins (Rohrmann, 1986). It was also revealed through deletion experiments that 1) the shape of polyhedra is controlled by C-terminus amino acids and 2) polyhedral size is controlled by amino acid sequence information. These polyhedrin deletion experiments may contribute to study of the mechanisms of size determination of organelles in eukaryotic cells.

In section V, the gene structure of BmNPV was discussed. Homologously repeated sequences (hr) were found in all baculoviruses examined. It has been shown that hr

has an enhancing activity, however, its true mechanisms are still unknown. Detailed comparison of sequence information between AcNPV and BmNPV, hr regions showed high conservation of sequence and location, indicating the importance of these regions in the baculovirus genome. In addition, one inversion of an hr sequence was speculated in the BmNPV genome by detailed homology search. About 1 kb was inverted at the 21.2 map unit position of BmNPV and AcNPV. Since inversion of the polyhedrin gene of OpNPV has also been reported (Blissard and Rohrmann, 1990), inversions associated with hr regions were speculated to occur commonly.

The entire sequence of a fragment (about 13% of the genome) containing four important immediate-early genes in the AcNPV genome was determined. Detailed analysis of this region, showed that baculoviruses were covered by relatively unique sequences except in the hr regions. These unique regions contained open reading frames, which were presumably transcribed and translated, and separated with up to 500 bp of non-translated regions containing promoters and termination signals. Early gene promoters or/and late gene promoters (see Blissard and Rohrmann, 1990) were found. Most ORFs had a poly $(A)^+$ signal close to its transcriptional stop. When the corresponding ORFs between BmNPV and AcNPV were compared, the amino acid sequences were generally highly conserved, indicating the importance of the gene products. It was also shown that non-translated

regions had relatively high mutation rates including deletions and insertions. In the near future, the entire BmNPV genome sequence will be determined. This information will show a clearer picture of how the baculovirus genome operates.

In conclusion, I devised a basic system to study two important NPVs using viral genomic libraries. Based on the unique characteristics of the polyhedrin gene, the following were performed: analysis of viral systematics, structural analysis at the DNA level, anlysis of the nuclear localization mechanisms of polyhedrin, and construction of improved expression vectors. Finally, the genomic structures of NPVs were characterized by analyzing homologously repeated and unique sequences. This information has shed light on the organization and mechanisms of the baculoviral genome and will be helpful for the application of these viruses to agriculture and medicine.

VII. References

Arif,B.M. and Doerfler,W. (1984) Identification and location of reiterated sequences in the Choristoneura fumiferana MNPV genome. EMBO J., 3, 525-529.

Beames,B. and Summers,M.D. (1988) Comparisons of host cell DNA insertions and altered transcription at the site of insertions in few polyhedra baculovirus mutants. Virology, 162, 206-220.

Blissard,G.W. and Rohrmann,G.F. (1990) Baculovirus diversity and molecular biology. Ann. Rev. Entomol., 35, 127-155.

Brown,S.E., Maruniak,J.E., and Knudson,D.L. (1984) Physical map of SeMNPV baculovirus DNA: an AcMNPV genomic variant. Virology, 136, 235-240.

Carson,D.D., Summers,M.D., and Guarino,L.A. (1991) Molecular analysis of a baculovirus regulatory gene. Virology, 182, 279-286.,

Carstens,E.B., Lin-bai,Y., and Faulkner,P. (1987) A point mutation in the polyhedrin gene of a baculovirus, Autographa californica MNPV, prevents crystallization of occlusion bodies. J. Gen. Virol., 68, 901-905.

Chen,D.D., Nesson,M.H., Rohrmann,G.F., and Beaudreau,G.S. (1988) The genome of the multicapsid baculovirus of Orgyia pseudotsugata restriction map and analysis of two sets of GC-rich repeated sequences. J. gen. Virol., 69, 1375-1381.

Chisholm,G.E. and Henner,D.J. (1988) Multiple early transcripts and splicing of the Autographa californica nuclear polyhedrosis virus IE-1 gene. J. Virol., 62, 3193-3200.

Clem,R.J., Fechheimer,M., and Miller,L.K. (1991) Prevention of apoptosis by a baculovirus gene during infection of insect cells. Science, in press.

Cochran,M.A. and Faulkner,P. (1983) Location of homologous DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. J. Virol., 45, 961-970.

Cochran,M.A., Carstens,E.B., Eaton,B.T., and Faulkner,P. (1982) Molecular cloning and physical mapping of restriction endonuclease fragments of Autographa californica nuclear polyhedrosis virus DNA. J. Virol., 41, 940-946.

Crawford,A.M. and Miller,L.K. (1988) Characterization of an early gene accelerating expression of late genes of the baculovirus Autographa californica nuclear polyhedrosis virus. J. Virol., 62, 2773-2781.

Croizier,G., Boukhoudmi-Amiri,K., and Croizier,L. (1989). A physical map of Spodoptera littoralis B-type nuclear polyhedrosis virus genome. Arch. Virol., 104, 145-151.

Dobos,P. and Cochran,M.A. (1980) Protein synthesis in cells infected by Autographa californica nuclear polyhedrosis virus (Ac-NPV): the effect of cytosine arabionoside. Virology, 103, 446-464.

Erlandson,M., Skepasts,P., Kuzio,J., and Carstens,E.B. (1984) Genomic variants of a temperature-sensitive mutant of Autographa californica nuclear polyhedrosis virus containing specific reiterations of viral DNA. Virus Research, 1, 565-584.

Fraser,M.J., Smith,G.E., and Summers,M.D.(1983) Acquisition of host cell DNA sequences by baculovirus: relationship between host DNA insertion and FP mutants of Autographa californica and Galleria mellonella nuclear polyhedrosis viruses. J. Virol., 47, 287-300.

Friesen,P.D. and Miller,L.K. (1985) Temporal regulation of baculovirus RNA: overlapping early and late transcripts. J. Virol., 54, 392-400.

Friesen,P.D., Rice,W.C., Miller,D.W., and Miller,L.K. (1986) Bidirectional transcription from a solo long terminal repeat of the retrotransposon TED: symmetrical RNA start sites. Mol. Cell. Biol., 6, 1599-1607.

Garoff, H. (1985) Using recombinant DNA techniques to study protein targeting in the eucarytotic cell. Ann. Rev. Cell Biol., 1, 403-445.

Gombart,A.F., Pearson,M.N., Rohrmann,G.F., and Beaudreau,G.S. (1989) A baculovirus polyhedral envelope-associated protein: genetic location, nucleotide sequence, and immunocytochemical characterization. Virology, 169, 182-193.

Granados,R.R. and Federici,B.A. (1986) The Biology of Baculoviruses, Vol. I, Biological Properties and

Molecular Biology, p. 275; Vol. II, Practical Application for Insect Control, p 276, CRC Press, Boca Raton, FL.

Guarino,L.A. (1990) Identification of a viral gene encoding a ubiquitin-like protein. Proc. Natl. Acad. Sci. USA, 87, 409-413.

Guarino,L.A., Gonzalez,M.A., and Summers,M.D. (1986) Complete sequence and enhancer function of the homologous DNA regions of Autographa californica nuclear polyhedrosis virus. J. Virol., 60, 224-229.

Guarino,L.A. and Summers,M.D. (1987) Nucleotide sequence and temporal expression of a baculovirus regulatory gene. J. Virol., 61, 2091-2099.

Horie,Y. and Watanabe,H. (1980) Recent advances in sericulture. Ann. Rev. Entomol. 25, 49-71.

Horiuchi,T., Marumoto,Y., Saeki,Y., Sato,Y., Furusawa,M., Kondo,A., and Maeda,S. (1987) High-level expression of the human-alpha-interferon gene through the use of an improved baculovirus vector in the silkworm, Bombyx mori. Agric. Biol. Chem., 51, 1573-1580.

Hukuhara,T. and Yamaguchi,K. (1973) Ultrastructural investigation on a strain of a cytoplasmic-polyhedrosis virus with nuclear inclusion. J. Invertebr. Pathol., 22, 6-13.

Iatrou,K., Ito,K., and Witkiewicz,H. (1985) Polyhedrin gene of Bombyx mori nuclear polyhedrosis virus. J. Virol., 54, 436-445.

Inoue,H. and Mitsuhashi,J. (1984) A Bombyx mori cell line susceptible to a nuclear polyhedrosis virus. J. Seric. Sci. Jpn., 53, 108-113.

Johnson,D.W. and Maruniak,J.E. (1989) Physical map of Anticarsia gemmatalis nuclear polyhedrosis virus (AgMNPV-2) DNA. J. Gen. Virol., 70, 1877-1883.

Kawarabata,T. and Matsumoto,K. (1973) Isolation and structure of a nuclear polyhedrosis virus from polyhedra of the silkworm, Bombyx mori. Appl. Ent. Zool., 8, 227-233.

Keddie,B.A., Aponte,G.W., and Volkman,L.E. (1989) The pathway of infection of Autographa californica nuclear polyhedrosis virus in an insect host. Science, 243, 1728-1730.

Knell,J.D. and Summers,M.D. (1984) A physical map for the Heliothis zea SNPV genome. J. Gen. Virol., 65, 445-450.

Kondo,A. and Maeda,S. (1991) Host range expansion by recombination of the baculoviruses Bombyx mori nuclear polyhedrosis virus and Autographa californica nuclear polyhedrosis virus. J. Virol., 65, 3625-3632.

Krappa,R. and Knebel-Morsdorf,D. (1991) Identification of the very early transcribed baculovirus gene PE-38. J. Virol., 65, 805-812.

Kumar,S. and Miller,L.K. (1987) Effects of serial passage of Autographa californica nuclear polyhedrosis virus in cell culture. Virus Research, 7, 335-349.

Kuzio,J. and Faulkner,P. (1984) Regions of repeated DNA in the genome of Choristoneura fumiferana nuclear polyhedrosis virus. Virology, 139, 185-188.

Kuzio,J., Jaques,R., and Faulkner,P. (1989) Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. Virology, 173, 759-763.

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227, 680-685.

Lee,H.H. and Miller,L.K. (1978) Isolation of genotypic variants of Autographa californica nuclear polyhedrosis virus. J. Virol. 27., 754-767.

Lee,H.H. and Miller,L.K. (1979) Isolation, complementation, and initial characterization of temperature-sensitive mutants of the baculovirus Autographa californica nuclear polyhedrosis virus. J. Virol., 31, 240-252.

Leisy,D.J., Rohrmann,G.F., and Beaudreau,G.S. (1984) Conservation of genome organization in two multicapsid nuclear polyhedrosis viruses. J. Virol., 52, 699-702.

Leisy,D.J., Rohrmann,G.F., Nesson,M., and Beaudreau,G.S. (1986) Nucleotide sequencing and transcriptional mapping of the Orgyia pseudotsugata multicapsid nuclear polyhedrosis virus p10 gene. Virology, 153, 157-167.

Liu,A., Qin,J., Rankin,C., Hardin,S.E., and Weaver,R.F. (1986) Nucleotide sequence of a portion of the Autographa californica nuclear polyhedrosis virus genome containing the EcoRI site-rich region (hr5) and

an open reading frame just 5' of the p10 gene. J. Gen. Virol., 67, 2565-2570.

Loh,L.C., Hamm,J.J., and Huang,E. (1981) Spodoptera frugiperda nuclear polyhedrosis virus genome: physical maps for restriction endonucleases BamHI and HindIII. J. Virol., 38, 922-931.

Lu,A. and Carstens,E.B. (1991) Nucleotide sequence of a gene essential for viral DNA replication in the baculovirus Autographa californica nuclear polyhedrosis virus. Virology., 181, 336-347.

Lubbert,H. and Doerfler,W. (1984) Transcription of overlapping sets of RNAs from the genome of Autographa californica nuclear polyhedrosis virus: a novel method for mapping RNAs. J. Virol., 52, 255-265.

Luckow,V.A. (1991) Cloning and expression of heterologous genes in insect cells with baculovirus vectors. In Recombinant DNA technology and application (Eds, Prokop,A., Bajpai,R.K., Ho,C.S.), pp. 97-152. McGrow-Hill, Inc., New York.

Luckow,V.A. and Summers,M.D. (1988) Trends in the development of baculovirus expression vectors. Biotechnology, 6, 47-55.

Maeda,S. (1984) A plaque assay and cloning of Bombyx mori nuclear polyhedrosis virus. J. Seric. Sci. Jpn., 53, 547-548.

Maeda,S. (1989a) Expression of foreign genes in insects using baculovirus vectors. Ann. Rev. Entomol., 34, 351-372.

Maeda,S. (1989b) Gene transfer vectors of a baculovirus, Bombyx mori nuclear polyhedrosis virus, and their use for expression of foreign genes in insect cells. In Invertebrate Cell System Applications, vol. I, J. Mitsuhashi, Ed. (CRC Press, Boca Raton, FL, 1989), p. 167-181.

Maeda, S. (1989c) Increased insecticidal effect by a recombinant baculovirus carrying a synthetic diuretic hormone gene. Biochem. Biophys. Res. Commun., 165, 1177-1183.

Maeda,S. and Majima,K. (1990) Molecular cloning and physical mapping of the genome of Bombyx mori nuclear polyhedrosis virus. J. Gen. Virol., 71, 1851-1855.

Maeda,S., Kawai,T., Obinata,M., Fujiwara,H., Horiuchi,T., Saeki,Y., Sato,Y., and Furusawa,M. (1985) Production of human alpha-interferon in silkworm using a baculovirus vector. Nature, 315, 592-594.

Maeda,S., Mukohara,Y., and Kondo,A. (1990) Characteristically distinct isolates of the nuclear polyhedrosis virus from Spodoptera litura. J. Gen. Virol., 71, 2631-2639.

Maeda,S., Kamita,S.G., Kataoka,H. (1991a) the basic DNA-binding protein of Bombyx mori nuclear polyhedrosis

virus: the existence of an additional arginine repeat. Virology, 180, 807-810.

Maeda,S., Volrath,S.L., T.N.,Hanzlik, Harper,S.A., Majima,K. Maddox,D.W., Hammock,B.D., and Fowler,E. (1991b) Insecticidal effects of an insect-specific neurotoxin expressed by a recombinant baculovirus. Virology, 184, 777-782.

Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982) Molecular Cloning; A Laboratory Manual. p544. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.

Maruniak,J.E., Brown,S.E., and Knudson,D.L. (1984) Physical maps of SfMNPV baculovirus DNA and its genomic variants. Virology, 136, 221-234.

Matthews,R.F.F. (1982) Classification and nomenclature of viruses. Intervirology, 17, 1-199.

Miller,L.K. (1988) Baculoviruses as gene expression vectors. Ann. Rev. Microbiol., 42, 177-199.

Miller,D.W. and Miller,L.K. (1982) A virus mutant with an insertion of a copia-like transposable element. Nature, 299, 562-564.

Miller,L.K. and Dawes,K.P. (1979) Physical map of the DNA genome of Autographa californica nuclear polyhedrosis virus. J. Virol., 29, 1044-1055.

Mori,H., Minobe,Y, Sasaki,T., and Kawase,S. (1989) Nucleotide sequence of the polyhedrin gene of Bombyx mori cytoplasmic polyhedrosis virus A strain with

nuclear localization of polyhedra. J. Gen. Virol., 70, 1885-1888.

Nitta, M. and Watanabe, H. (1984) Resistance of nuclear polyhedrosis virus polyhedra Bombyx mori to formaldehyde treatment. J. Seric. Sci. Japan, 53, 146-150.

Okada,T. (1977) Studies on the utilization and mass production of Spodoptera litura nuclear polyhedrosis virus for control of the tobacco cutworm, Spodoptera litura Fabricius. Rev. Plant Prot. Res., 10, 102-120.

Ooi,B.G. and Miller,L.K. (1990) Transcription of the baculovirus polyhedrin gene reduces the levels of an antisencse transcript initiated downstream. J. Virol., 64, 3126-3129.

O'Reilly,D.R. and Miller,L.K. (1989) A baculovirus blocks insect molting by producing ecdysteroid UDP-glycosyl transferase. Science, 245, 1110-1112.

Possee,R.D. and Kelly,D.C. (1988) Physical maps and comparative DNA hybridization of Mamestra brassicae and Panolis flammea nuclear polyhedrosis virus genomes. J. gen. Virol., 69, 1285-1298.

Redford,G.O. (1986) Biological control of the rhinoceros (Oryctes rhinoceros) in the South Pacific by baculovirus. Agric. Exosystems Environ., 15, 141-147.

Rohrmann,G.F. (1986) Polyhedrin structure. J. Gen. Virol., 67, 1499-1513.

Russell,R.L.Q. and Rohrmann,G.F. (1990) A baculovirus polyhedron envelope protein: immunogold localization in infected cells and mature polyhedra. Virology, 174, 177-184.

Sekine,H., Fuse,A., Tada,A., Maeda,S., and Simizu,B. (1988) Expression of human papillomavirus type 6b E2 gene product with DNA-binding activity in insect (Bombyx mori) cells using a baculovirus expression vector. Gene, 65, 187-193.

Silver,P.A. (1991) How proteins enter the nucleus. Cell 64, 489-497.

Smith,G.E. and Summers,M.D. (1979) Restriction maps of five Autographa californica MNPV variants, Trichoplusia ni MNPV, and Galleria mellonella MNPV DNAs with endonucleases SmaI, KpnI, BamHI, SacI, XhoI, and EcoRI. J. Virol., 30, 828-838.

Smith,G.E. and Summers,M.D. (1982) DNA homology among subgroup A, B, and C baculoviruses. Virology, 123, 393-406.

Summers,M., Engler,R., Falcon,L.A., and Vail,P. (1975) Baculoviruses for insect pest control: safety consideration. American Society for Microbiology, Washington

Summers,M.D. and Smith,G.E. (1987) Laboratory manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experiment Station Bulletin No. 1555.

Summers,M.D., Smith,G.E., Knell,J.D., and Burand,J.P. (1980) Physical maps of Autographa californica and Rachiplusia ou nuclear polyhedrosis virus recombinants. J. Virol., 34, 693-703.

Thiem,S.M. and Miller,L.K. (1989) Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus Autographa californica nuclear polyhedrosis virus. J. Virol., 63, 2008-2018.

Tomalski, M.D., Eldridge, R., and Miller,L.K. (1991) A baculovirus homolog of a Cu/Zn superoxide dismutase gene. Virology, 184, 149-161.

van Iddekinge,B.J.L., Smith,G.E., and Summers,M.D. (1983) Nucleotide sequence of the polyhedrin gene of Autographa californica nuclear polyhedrosis virus. Virology, 131, 561-565.

Vlak,J.M. (1980) Mapping of BamHI and SmaI DNA restriction sites on the genome of the nuclear polyhedrosis virus of the alfalfa looper, Autographa californica. J. Invertebr. Pathol., 36, 409-414.

Vlak,J.M. and Smith,G.E. (1982) Orientation of the genome of Autographa californica nuclear polyhedrosis virus: a proposal. J. Virol., 41, 1118-1121.

Watanabe,H. (1975) Variation in the numbers of nucleocapsid within the envelope of nuclear polyhedrosis virus multiplied in tissues of the silkworm, Bombyx mori. J. Seric. Sci. Jpn., 43, 29-34.

Whitford,M., Stewart,S., Kuzio,J., and Faulkner,P. (1989) Identification and sequence analysis of a gene encoding gp67, an abundant envelope glycoprotein of the baculovirus Autographa californica nuclear polyhedrosis virus. J. Virol., 63, 1393-1399.

Wiegers,F.P. and Vlak,J.M. (1984) Physical map of the DNA of a Mamestra brassicae nuclear polyhedrosis virus variant isolated from Spodoptera exigua. J. Gen. Virol., 65, 2011-2019.

Wilson,M.E., Mainprize,T.H., Friesen,P.D., and Miller,L.K. (1987) Location, transcription, and sequence of a baculovirus gene encoding a small arginine-rich polypeptide. J. Virol., 61, 661-666.