AN ABSTRACT OF THE THESIS OF

Randolph Logan Chakerian for the degree of Master of Science in Genetics presented on May 29, 1985.

Title: Identification, Cloning and Nucleotide Sequencing of the Granulin Genes of Trichoplusia ni and Pieris brassicae Granulosis Viruses.

Redacted for Privacy

Abstract approved: _____

Peter S. Dawson

Certain insect baculoviruses are occluded in a proteinaceous, crystalline structure which serves to protect the virus outside its insect host. Two groups of occluded viruses have been defined: the nuclear polyhedrosis viruses (NPVs) (Baculovirus Subgroup A), having many virions occluded per crystal, and the granulosis viruses (GVs) (Baculovirus Subgroup B), having only one virion per crystal.

Using the cloned polyhedrin gene from the Orgyia pseudotsugata MNPV (OpMNPV) as a hybridization probe, a 2 kilobase Sal I fragment from the Trichoplusia ni granulosis virus (TnGV) was identified as the locus of the granulin gene. This fragment was cloned into pUC8 and mapped with restriction enzymes. Three fragments of about 400 base-pairs each were subcloned into reciprocal M13 vectors and their nucleotide sequences were determined by the dideoxy method.

A 1.3 kilobase Eco RI clone containing most of the granulin gene from TnGV was used as a hybridization probe to identify the granulin gene from Pieris brassicae granulosis virus (PbGV). This gene was also cloned, mapped, subcloned and dideoxy sequenced.

The amino acid sequences derived from both the PbGV and TnGV granulin gene sequences are about 70% conserved with respect to each other and about 50% conserved related to the Lepidopteran NPV polyhedrins. This suggests that the GVs form a distinct branch of

Baculoviruses which evolved before the extensive divergence of the Lepidopteran NPVs. Highly conserved sequences within these genes suggest regions of granulin and polyhedrin which might be determinants of higher-order structure. Comparison of the 5' flanking regions of both granulins, several NPVs and another hyperexpressed late gene, the "10k" protein of Autographa californica MNPV, reveals a highly conserved sequence which may be a regulatory element involved in governing the expression of these genes.

Identification, Cloning and Nucleotide Sequencing of the Granulin Genes of *Trichoplusia ni* and *Pieris brassicae* Granulosis Viruses

by

Randolph Logan Chakerian

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed May 29, 1985

Commencement June 1986

APPROVED:

Redacted for Privacy

Professor of Zoology in charge of major

Redacted for Privacy

Head of Genetics Program

Redacted for Privacy

Dean of Graduate School

Date thesis is presented May 29, 1985

Typed by himself for Randolph Logan Chakerian

ACKNOWLEDGEMENTS

Substantial contributions to the data included in this thesis were made by Donna Akiyoshi (TnGV subcloning and sequencing); George Rohrmann (cloning and subcloning of the -COOH termini of both genes, and reading of the sequencing gels); Mike Nesson (S sequencing of both genes); Doug Leisy (initial PbGV cloning and hybridization); and George Beaudreau (initial nick translations for PbGV hybridizations).

Beyond these material contributions, however, were the contributions to my understanding of biology, such as it is, made by a long list of friends and teachers: Denzel Ferguson, Joe Beatty, Pete Dawson, Ralph Quatrano, George Beaudreau, George Rohrmann, John Armstrong, Mike Nesson, George and Margot Pearson, Doug Leisy and last and still, Diane Henneberger.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

IDENTIFICATION, CLONING AND NUCLEOTIDE SEQUENCING OF THE GRANULIN GENES OF *TRICHOPLUSIA NI* AND *PIERIS BRASSICAE* GRANULOSIS VIRUS

INTRODUCTION

Pieris brassicae Granulosis Virus (*PbGV*) and *Trichoplusia ni* Granulosis Virus (*TnGV*) are members of the family Baculoviridae, a group of rod-shaped (baculum=rod), membrane-enveloped viruses restricted to arthropod hosts and containing supercoiled double-stranded DNA. Aside from two Baculoviruses which infect Crustacea (Couch, 1974; Lightner and Redman, 1981), the Baculoviridae appear to be confined exclusively to holometabolous insect hosts and particularly to the Lepidoptera (Rohrmann et al, 1981).

PbGV and *TnGV* are members of Baculovirus Subgroup B, the Granulosis Viruses (GVs), which along with the related Subgroup A, the Nuclear Polyhedrosis Viruses (NPVs), are characteristically occluded in a crystalline protein matrix called granulin and polyhedrin, respectively. The matrix serves to protect the virus from inactivation when it is not infecting its host: Douglas-fir Tussock Moth NPV recovered from forest soils 40 years after an epizootic was found to be still infectious (Thompson et al,1981). A third group of Baculoviruses, Subgroup C, or Non-Occluded

Viruses (NOVs), are not yet well characterized and will not be considered in this thesis.

GVs and NPVs, although relatively closely related in their morphology and in the biology of their infection process, differ in several important respects. GV nucleocapsids are enveloped singly in a lipoprotein membrane, are occluded one to a granule, and infect only lepidopteran hosts. NPVs are enveloped either singly (SNPVs) or in multiples (MNPVs). In addition, up to 100 bundles of MNPVs may be occluded in a single polyhedrin. Like the GVs, MNPVs are found only in the Lepidoptera. SNPVs, however, occur in the Hymenoptera, Diptera, and Trichoptera, as well as in the Lepidoptera (Table 1).

The usual route of occluded Baculovirus infection begins with the ingestion of an occlusion body by a host larva. The alkaline conditions of the larval midgut cause dissolution of the polyhedrin or granulin and release of the enveloped viral particle(s) into the midgut lumen. Fusion of the envelope with midgut epithelial cell membranes results in the passage of the viral particles into the cells, association of these particles with pores in the nuclear membrane, and insertion of viral DNA into the nucleus (Tweeten, Bulla & Consigli, 1981). Progeny viral particles appear in the nucleus after 6-8 hours (Summers, Volkman and Hsieh, 1978), but at this stage of infection they are not occluded. Rather, they initiate secondary infections in fat body, tracheolar, and other host cells.

TABLE 1

Baculovirus Occurrence in arthropods

Arthropoda	Virus	Morphotype	Number of Affected Species
Crustacea			
Decapoda	NPV	S	2
Insecta			
Trichoptera	NPV	S	1
Hymenoptera	NPV	S	26
Diptera	NPV	S	22
Lepidoptera	NPV	S,M	355
	GV	S	113
	NOBV	S	2
Coleoptera	NOBV	S	16

NPV: nuclear polyhedrosis virus
GV: granulosis virus
NOBV: non-occluded baculovirus
S: singly-enveloped nucleocapsid
M: multiply-enveloped nucleocapsid

(from Vlak & Rohrmann, 1984)

During this secondary infection GVs and NPVs follow divergent pathways. In both types of infection the size of the cell nucleus increases, but in GV infections the nuclear membrane disintegrates before occlusion of the viral particles occurs. In NPV infections occlusion occurs within the intact nucleus . There must therefore be a mechanism, absent in GVs, for the transportation of polyhedrin through the nuclear membrane during NPV infection.

A number of efforts have been made to clarify Baculovirus phylogeny for both applied and evolutionary studies. The restriction of GVs and MNPVs to Lepidoptera and of NOVs to Coleoptera and 2 Lepidoptera, and the wide distribution of SNPVs among the Decapoda, Trichoptera, Hymenoptera, Diptera and Lepidoptera (Table 1) are indications of a general phylogeny. More specific classification is by reference to the type host, although it is not clear that host specificity is always limited to a single insect species. PbGV and TnGV infect Pieris brassicae, the cabbage white butterfly, and Trichoplusia ni, the cabbage looper, respectively, but PbGV can infect other Pierid species (David, 1978) and Autographa californica NPV (AcNPV), whose type host is the alfalfa looper, infects at least 6 lepidopteran families (Kaya, 1977).

Studies attempting to define Baculovirus phylogeny with greater precision have used a variety of techniques to

analyze specific viral components - nucleocapsid proteins, envelope proteins, polyhedrin or granulin, total genomic DNA and, most recently, the polyhedrin gene itself - to establish relationships between and within viral groups. While some general relationships have been established, the data accumulated until recently have not always been consistent and in some cases have even been contradictory.

Immunological techniques have generally focussed on polyhedrin. A study of 17 baculoviruses (Krywienczyk and Bergold, 1960) defined 3 groups: hymenopteran NPVs, GVs, and a grouping of 11 lepidopteran NPVs and 2 other GVs. More recent immunological studies have shown no serological relatedness between lepidopteran NPVs and a dipteran NPV (Guelpa et al, 1977; Rohrmann et al, 1981), but did show some cross-reactivity between an MNPV and PbGV (Rohrmann et al, 1981).

Radial immunodiffusion studies of nucleocapsid proteins showed differences between SNPVs and MNPVs of Plusiinae baculoviruses (McCarthy and Lambiase, 1979), but the same study found granulin from SNPVs and MNPVs to be serologically similar. Western blots of nucleocapsid proteins from 17 baculovirus species, however, showed that MNPVs, SNPVs and GVs had similar antigenic determinants (Smith and Summers, 1981). Granulins again showed immunological similarity in the same study.

Monoclonal antibodies against polyhedrin from a number of Baculoviruses have also been prepared (Hohmann and

Faulkner, 1983; Quant et al, 1984), primarily for utilization in applied studies.

Tryptic peptide and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analyses of Baculovirus structural proteins (i.e., from both envelopes and nucleocapsids) and polyhedrins have not demonstrated clear patterns of relatedness. Rohrmann et al (1979) found that tryptic peptides of polyhedrin from an SNPV and an MNPV with the same host were very closely related. However, Summers and Smith (1978), using SDS-PAGE, found unique structural polypeptides among 8 Baculoviruses, including differences in 4 MNPVs between the normally-occurring multiply-enveloped forms and variant morphotypes of these having only singly-enveloped particles.

Amino-terminal residues and amino-terminal sequences of polyhedrins have shown patterns of similarities and differences among various baculoviruses. Maruniak and Summers (1978) found the amino-terminal residue was either proline or methionine for 6 NPVs and glycine for 2 GVs. Rohrmann et al (1979) showed that an MNPV and an SNPV with the same host differed at only 4 of the first 34 amino-terminal residues, and that the Bombyx mori SNPV amino-terminal sequence (Serebryani et al. 1977) was also different at only 4 or 5 residues respectively from either of these. A comparison of 10 amino-terminal polyhedrin amino acid sequences (Rohrmann, 1985) clearly shows close relatedness among 7 lepidopteran SNPVs and MNPVs; close

relatedness between 2 GVs; a somewhat more distant relatedness between the lepidopteran NPVs and the GVs; and a distinctly different amino-terminal sequence for a dipteran SNPV.

Many recent studies of Baculovirus relatedness have focussed on DNA, but until the application of direct DNA sequencing to the problem the data produced by these studies was even more contradictory than that of previous studies with other substrates.

Baculovirus genomes range in size from 90-165 kilobases. While closely related genomes can be compared on the basis of restriction enzyme profiles (Smith and Summers, 1979; Rohrmann, Martignoni and Beaudreau, 1982), this technique is not applicable to genomes whose sequences diverge by more than 15-25% (Upholt, 1977). A study of 18 Baculoviruses (Smith and Summers, 1982), for instance, showed a unique restriction pattern for every virus.

A very powerful technique, when applied cautiously, is hybridization between heterologous genomes or genomic fragments. Conditions under which these hybridizations are carried out must be carefully considered, however, and under even the least stringent conditions sequences which diverge by more than 33% will not be able to form stable heteroduplexes (Howley et al, 1979). Given the degeneracy of the genetic code, it would be possible for two proteins with identical peptide sequences to be coded for by nucleotide sequences diverging at every third base, or by 33%. In any

conditions other than those most permissive of heteroduplex formation - and quite possibly not even under those conditions - the two identical proteins in this hypothetical example might show little or even no homology between their coding sequences. Rohrmann (1985) has documented a case in which three different investigators, using DNA hybridization techniques, calculated the overall homology between the genomes of the same two NPVs as 1, 4 and 25%.

All of the techniques discussed above are based on inherently incomplete data sets, and are therefore subject to some degree of error. The elucidation of full-length amino-acid and nucleotide sequences, particularly the latter, are major steps towards generating complete and specific data with which related organisms can be studied. Such comparisons are most useful when they are made on the basis of sequences of functionally equivalent genes. Among Baculoviruses, the granulin and polyhedrin genes are obvious candidates for such a comparison.

The polyhedrin monomer is a protein of about 29,000 daltons coded for by a single-copy viral gene (van der Beek et al, 1980; Summers et al, 1980). Although polyhedrin message and polyhedrin monomers have been found as early as 12 hours post-infection (2-4 hours after the appearance of infectious non-occluded virus in midgut epithelial cells (Summers et al, 1978)), the message is not abundant until 18-24 hours post-infection (Smith, Vlak and Summers, 1983).

The formation of polyhedrin particles - crystalline aggregations of the monomers with embedded virions - is not apparent until some 48 hours post-infection (Smith, Vlak and Summers, 1983). Polyhedrin is thus a late gene in the cycle of virus infection, but in a somewhat special sense: its transcription and expression during the period of initial infection are at low levels, although complete virions, including envelopes, are synthesized at this stage. It is apparently only in secondarily infected cells that the gene is transcribed and expressed at its characteristically unusually high levels.

Polyhedrin can account for up to 17% of the total protein of an infected insect (Quant et al, 1984), and in cell culture polyhedrin from Autographa californica NPV (AcNPV) constituted about 25% of the total protein mass of infected cells of Spodoptera frugiperda, the fall armyworm (Smith, G.E., M.J. Fraser and M.D. Summers, 1983). Thus the single-copy polyhedrin gene appears to be expressed in greater abundance than any other protein known in a virus-infected eucaryotic cell.

The complete nucleotide sequences and some flanking sequences have previously been determined for 3 NPV polyhedrin genes. Complete amino acid sequences for 3 other NPV polyhedrins, as well as amino-terminal amino acid sequences for 2 other NPV polyhedrins, have also been elucidated (Table 2). Polyhedrin is thus the most thoroughly and precisely characterized of any Baculovirus gene product.

TABLE 2

Previously Reported Polyhedrin Sequences

Type Host		Common Name	Order:Family	Morphotype	Reference
			Complete Amino Acid Sequences		
Bombyx mori	(Bm)	Silkworm moth	Lepidoptera: Bombycidae	SNPV	Kozlov et al, 1981
Porthetria dispar (Lymantria dispar)	(Ld)	Gypsy moth	Lepidoptera: Lymantriidae	MNPV	Kozlov et al, 1981
Galleria mellonella	(Gm)	Greater wax moth	Lepidoptera: Pyralidae	MNPV	Kozlov et al, 1981
			Amino-Terminal Amino Acid Sequences		
Tipula paludosa	(Tp)	Crane fly	Diptera: Tipulidae	SNPV	Rohrmann et al, 1981
Neodiprion sertifer	(Ns)	European pine sawfly	Hymenoptera Diprionidae	SNPV	Rohrmann et al, 1981
			Complete Nucleotide Sequences		
Orgyia pseudotsugata	(Op)	Douglas-fir tussock moth	Lepidoptera: Lymantriidae	SNPV	Leisy et al, 1985
Orgyia pseudotsugata	(Op)	Douglas-fir tussock moth	Lepidoptera: Lymantriidae	MNPV	Leisy et al, 1985
Autographa californica	(Ac)	Alfalfa looper	Lepidoptera: Noctuidae	MNPV	Hooft van Iddekinge et al, 1983

From these 8 polyhedrin sequences, and the 2 GV granulin nucleotide sequences presented in this thesis and elsewhere (Akiyoshi et al, 1985; Chakerian et al, 1985), 3 broad sets of hypotheses can be elaborated: one characterizes and contrasts the granulin and polyhedrin proteins and, where the data permits, their genes; a second defines the phylogenetics and evolution of the Baculoviridae; and a third addresses the regulation of polyhedrin and granulin gene expression.

MATERIALS AND METHODS

Virus

InGV granule preparations were a gift of Dr. Gale Smith.

PbGV granule preparations were a gift of Dr. Norman Crook.

Isolation of Viral Genomic DNA

20-25 mg of granule preparations (viral particles and associated granulin matrix) were resuspended in 1 ml of water. The matrix was dissolved by the addition of 10 ul of 4 M NaOH. Virions were pelleted at 30,000 rpm for 30 min in a Sorvall AH 650 rotor at 5°C and resuspended in 0.8 ml of 0.1 M Tris (pH 7.0), 5 mM EDTA and 0.2 ml of 10% Sarcosyl. After 30 min at 37°C Proteinase K was added to a final concentration of 100 ug/ml and incubated for a further 30 min. Viral DNA was purified by extraction with an equal volume of phenol, followed by two back extractions of the phenol phase with equal volumes of water. The pooled aqueous phases were extracted once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1 vol:vol) and once with an equal volume of chloroform:isoamyl alcohol (24:1). The final aqueous phase was dialyzed against ice-cold TE buffer (10 mM Tris (pH 8.0), 1 mM EDTA) for 24 hours at 4°C with several changes of buffer.

Southern Blotting, Nick Translation and Hybridizations

Viral genomic and recombinant plasmid DNA for Southern blots was size fractionated in horizontal submarine agarose gels. Agarose concentrations ranged from 0.7% to 1.4% depending upon sizes of fragments to be separated (Maniatis et al, 1982). Both TAE (10X = 0.4 M Tris (pH 7.4), 0.4 M Acetate, 20 mM EDTA) and TBE (20X = 1 M Tris (pH 8.3), 1 M Borate, 20 mM EDTA) buffers were used. Ethidium bromide at 0.5 ug/ml was included in both the gels and the running buffer. All restriction enzymes were purchased from Bethesda Research Laboratories and used according to the manufacturer's specifications.

Bi-directional Southern blots were made according to Smith and Summers (1980).

Nick translations were performed according to Maniatis et al (1982) with the following modifications: 10X nick translation buffer was 0.4 M Tris (pH 7.6), 65 mM $MgCl$, and 10 mM B-Mercaptoethanol. Final concentration of cold dNTPs was 2 uM each. 20 uCi of labeled dNTP (either ^{32}P α-labeled dATP or ^{32}P α-labeled dCTP, Spec. Ac. = 800 Ci/mmol, purchased from New England Nuclear) was added giving a final concentration of 0.5 uM. 330 pg of DNase I were added for nicking, and 10-15 units of E. coli Polymerase 1 (Bethesda Research Laboratories) were used to translate the nicks.

Unincorporated nucleotides were removed by passage of the reaction mix over a Sephadex G50 column as outlined in

Maniatis et al (1982), except that a 3 ml disposable syringe was used in place of a 1 ml syringe. Incorporation of ~25% of total counts, with an associated specific activity of 1×10^7 - 1×10^8 cpm/ug, was considered near-optimum.

Hybridizations of Southern blots were performed according to Maniatis et al (1982) with the following modifications. Blots were presoaked in 4X SET (20X = 3 M NaCl, 0.4 M Tris (pH 7.8), 20 mM EDTA), 10X Denhardt's Solution for 1 hour at room temperature. Prehybridization was in 1 M NaCl, 10 mM Tris (pH 7.4), 1X Denhardt's Solution, 0.1% SDS and 40 ug/ml boiled salmon testis DNA at 42°C for 30 minutes in a volume of 50 ul/cm^2 of blot. The prehybridization solution was removed and replaced with an equal volume of the same solution with the addition of $1\text{-}5 \times 10^6$ cpm of boiled probe and either 20% or 50% formamide for low or high stringency, respectively. Hybridization was overnight at 42°C. Blots were washed once in 6X SSC, 1% SDS for 10 min; twice in 6X SSC for 10 min each; and once in 5mM Tris (pH 7.4) for 10 min, all at room temperature. Film was exposed at room temerature, without intensifying screens, from overnight to several days.

Fragment Isolation, Cloning and Transformation

Viral genomic fragments hybridizing to nick-translated granulin probes were isolated preparatively from agarose gels either in low melting temperature agarose (LMT agarose, Bethesda Research Laboratories) or on Whatman DE-81 paper.

In LMT agarose isolations a small well about 5 mm long

and slightly wider than the gel lane was cut below the target fragment, melted LMT agarose was poured in to slightly over-fill the well, and the gel was placed at 4°C for 10 minutes. When the fragment had fully migrated into the LMT agarose the well was cut out of the gel and placed in a 1.5 ml Eppendorf microfuge tube. About 5 volumes (about 1 ml) of TE buffer (20 mM Tris (pH 7.4), 1 mM EDTA) were added and the tube was placed at 65°C for 10 minutes. Fragment DNA was purified by phenol extraction as described above, and ethanol precipitated.

In DE-81 paper isolations the digests were run in the outside gel lanes so that a slit just below the target fragment could be cut from the edge of the gel and across the lane in which the digest was run. A piece of dry DE-81 paper slightly wider than the lane and slighty higher than the thickness of the gel was placed in the slit. When the fragment had migrated off the gel the paper was placed in a 0.5 ml Eppendorf microfuge tube with a small hole in the bottom. After 4 washes in TE buffer the DNA was eluted from the paper into a 1.5 ml Eppendorf tube with TE buffer/1 M NaCl. Washes and elutions were done in a table-top centrifuge at 1000 rpm at room temperature. The DNA was ethanol precipitated following elution.

Neither method proved fully satisfactory, and recovery in both methods was limited to 50% at best.

Genomic fragments were cloned into the Eco R1 site of pBR325 (PbGV) or pUC8 (TnGV) according to Maniatis et

al (1982). T4 DNA Ligase was purchased from Bethesda Research Laboratories.

Transformation of *E. coli* C600 was according to Maniatis et al (1982) using the $CaCl_2$ procedure. Small-scale plasmid isolations to assay for recombinants with the correct insert were performed according to Holmes and Quigley (1981). A variation of this method was used for large-scale (100 ml) plasmid isolations. The procedure is given in its entirety in Table 3.

Mapping, Subcloning and Nucleotide Sequencing

Recombinant plasmids containing desired viral genomic fragments were mapped initially using single and double restriction enzyme digests and hybridization to granulin probes. Sub-fragments showing hybridization were sub-cloned into the appropriate M13 sequencing vector (mps 8, 9 and 18 were used). Transfection of *E. coli* JM103, identification of recombinants, and sequencing protocols using ^{32}P α-labeled dATP were performed according to the Bethesda Research Laboratories sequencing manual (1980). Final sequencing was with ^{35}S α-labeled dATP and was performed according to the New England Nuclear sequencing manual (1983). 0.4 mm gels of either 6% or 8% acrylamide were used for both ^{32}P and ^{35}S sequencing.

TABLE 3

Large-Scale (100ml) plasmid isolations

1. Grow up 100 mls LB (with or without antibiotic) inoculated with a single recombinant colony overnight at 37° C.

2. Centrifuge 5k, 10 min, 5°C (GSA rotor).

3. Resuspend pellet in 70 ul/ml of original volume in STET buffer.

4. Add 5 ul/ml of original volume fresh lysozyme at 10 mgs/ml.

5. Boil 2 mins - at first over an open flame until the mixture begins to bubble (some caution is advised here) and then in a boiling water bath.

6. Centrifuge 15k, 15 min, 5° C.

7. Remove viscous pellet and add to remaining supernatant (7 ml) a half volume of 7.5 M NH OAC and 2-3 volumes of isopropanol. Place at -70°C for 30 min.

8. Spin 15k, 15 min, 5°C. Wash with 70% ethanol and dry down.

9. Resuspend pellet in 1 ml 0.1X E buffer.

10. Add DNase-free RNase to a final concentration of 50. Add Proteinase K to a final concentration of 100 ug/ml. Incubate 30 min at 37° C.

11. Extract with phenol, phenol-chloroform and chloroform, and ethanol precipitate.

This procedure yields from 0.5 - 1 mg of plasmid DNA from 100

ml of saturated overnight culture. Restriction digestion and subcloning of plasmid DNA prepared from this procedure were routinely accomplished.

RESULTS AND DISCUSSION

Identification and Cloning of the Granulin Genes of TnGV and PbGV

Because N-terminal amino acid sequences and immunological data showed some relatedness between OpMNPV polyhedrin and PbGV granulin (Rohrmann et al, 1981), the initial experimental strategy was to use cloned OpMNPV polyhedrin DNA to localize the PbGV granulin gene. Four experiments in which various restriction digests of PbGV genomic DNA were Southern blotted and probed with an OpMNPV polyhedrin clone (Rohrmann et al, 1982), either as a whole plasmid or as an isolated insert, showed no hybridization under non-stringent conditions (data not shown). In a fifth experiment, two lanes of heavily overloaded Eco RI-digested PbGV genomic DNA probed with the isolated insert of the OpMNPV polyhedrin gene both showed clear hybridization to a position containing 3 fragments of approximately 6 kb each (Fig. 1).

Since it was not possible to identify a discrete DNA fragment to which the probe hybridized, all 3 fragments were isolated preparatively from an agarose gel, ligated into Eco R1-cut pBR325, and transformed into E. coli C600. Fourteen of 1546 colonies assayed on antibiotic-containing agar plates showed chloramphenicol sensitivity. Plasmids isolated from these colonies which contained 6 kb inserts were

Figure 1. Hybridization of PbGV and OpSNPV with OpMNPV polyhedrin insert.

A: Ethidium bromide-stained agarose gel.

Lane 1 - Eco RI-digested PbGV genomic DNA
Lane 2 - Eco RI-digested OpSNPV genomic DNA
Lane 3 - Same as Lane 1

B: Autoradiogram of Southern blot of A, probed with OpMNPV insert

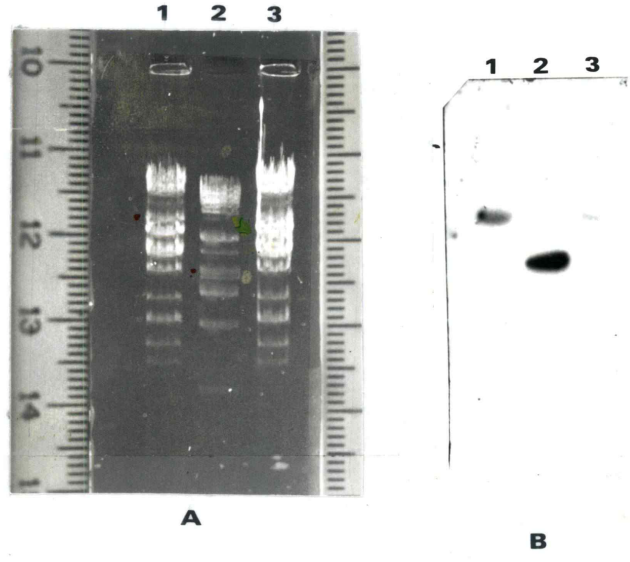


Figure 1

further assayed using double restriction digests. Three different clones were identified, one of which (GV14) had a total insert size of 12 kb and contained both a unique 6 kb fragment and the 6 kb fragment also contained in clone GV3. OpMNPV-probed Southern blots of digests of these 3 plasmids, however, showed no hybridization (data not shown).

Because of these anomalous results with PbGV, bi-directional Southern blots of Eco R1-digested TnGV genomic DNA were probed with nick-translated clones of the polyhedrin genes from OpSNPV and OpMNPV under non-stringent conditions (Fig. 2). A discrete fragment of approximately 1.3 kb hybridized with both probes and was cloned into pUC8. This cloned fragment was then used as a nick-translated probe and was hybridized under non-stringent conditions to a blot of restriction-digested genomic DNAs from TnGV, PbGV and OpMNPV. The homologous 1.3 kb Eco R1 TnGV fragment, the 6 kb region in Eco R1-digested PbGV genomic DNA, and the 5 kb Xho 1 fragment of OpMNPV containing the polyhedrin gene (Rohrmann et al, 1982) all showed hybridization. A reciprocal hybridization of the twin blot, using DNA from the cloned OpSNPV granulin gene as a probe, showed the same pattern of homology, including strong hybridization to the 6 kb region in Eco R1-digested PbGV genomic DNA (Fig. 3).

Blots of Eco R1/Bam H1 digests of the 3 PbGV plasmids probed with the 1.3 kb TnGV clone showed hybridization to common bands in the GV3 and GV14 digests (Fig. 4).

Figure 2. Hybridization of TnGV with OpMNPV and OpSNPV polyhedrin clones.

A: Ethidium bromide-stained agarose gel.

Lane 1 - Eco RI-digested TnGV genomic DNA
Lane 2 - Xho I-digested TnGV genomic DNA

B: Autoradiogram of Southern blot of A, probed with OpMNPV polyhedrin clone

C: Autoradiogram of duplicate Southern blot of A, probed with OpSNPV polyhedrin clone

Arrows in B and C indicate strong hybridization to a 1.3 kb Eco RI band, whose location is indicated but not visible (because its concentration is too low for ethidium bromide visualization) in A.

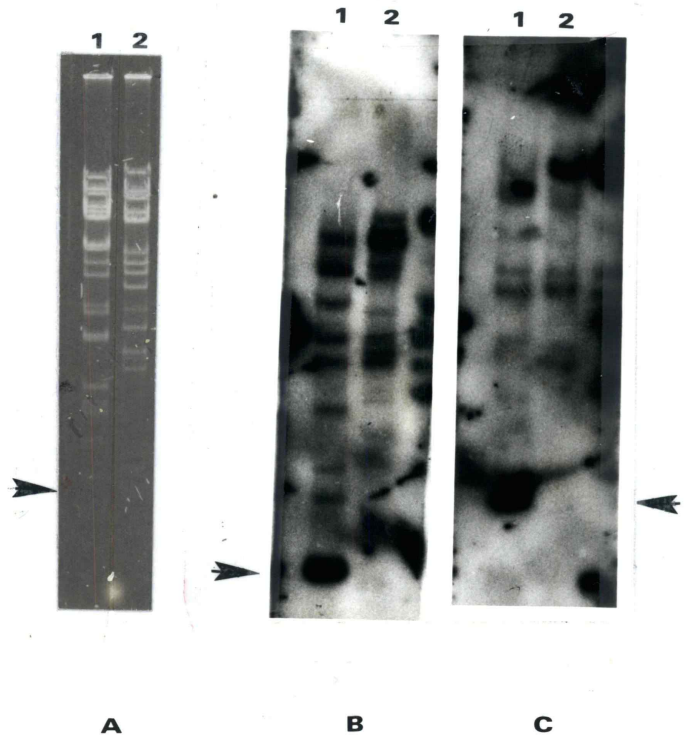


Figure 2

Figure 3. Hybridization of TnGV, PbGV and OpMNPV with OpSNPV polyhedrin and TnGV granulin clones

A: Ethidium bromide-stained agarose gel

Lane 1 - Hind III-digested Lambdaphage DNA size markers
Lane 2 - Eco RI-digested TnGV genomic DNA
Lane 3 - Eco RI-digested PbGV genomic DNA
Lane 4 - Xho I-digested OpMNPV genomic DNA
Lane 5 - same as lane 1

B: Autoradiogram of Southern blot of A, probed with OpSNPV polyhedrin clone

C: Autoradiogram of duplicate Southern blot of A, probed with TnGV granulin clone

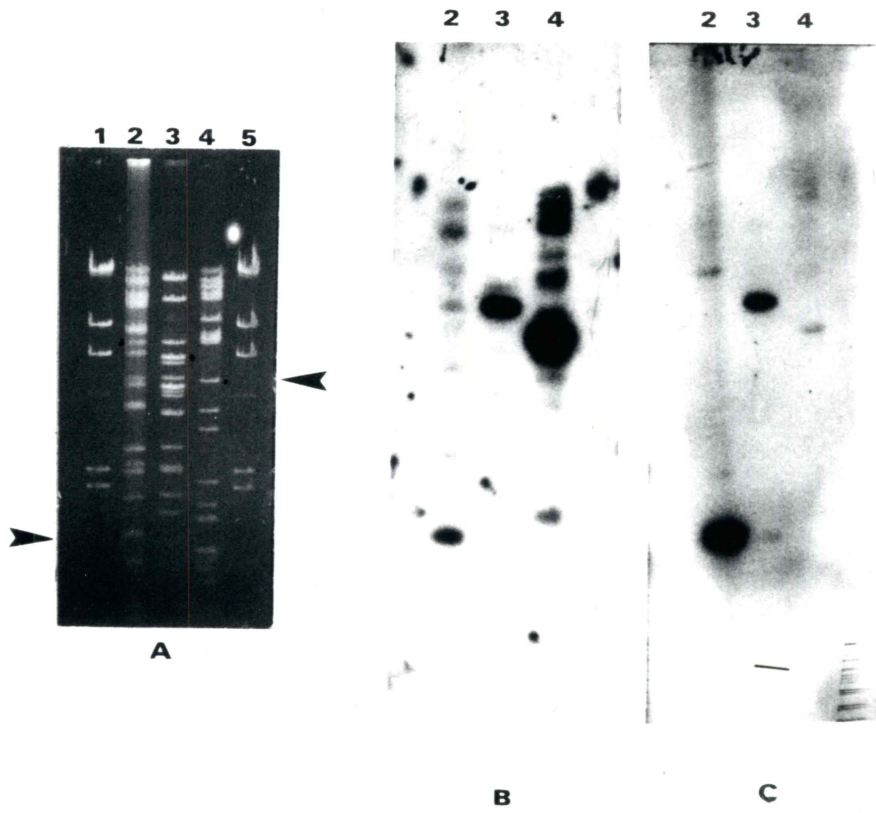


Figure 3

Figure 4. Hybridization of 3 PbGV clones with TnGV granulin clone

A: Ethidium bromide-stained agarose gel

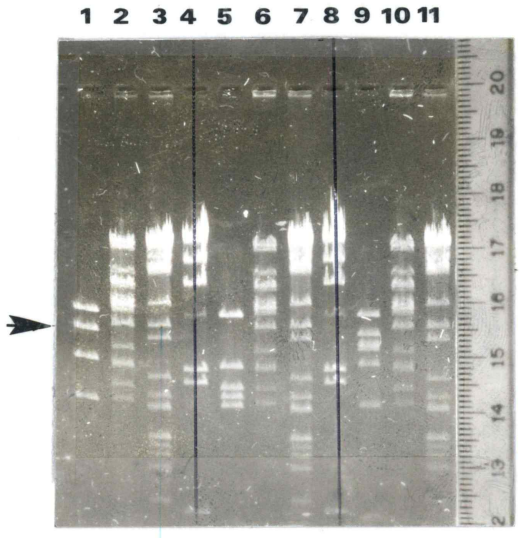
Lane 1 - Eco RI/Bam HI digest of clone GV3
Lane 2 - Eco RI digest of PbGV genomic DNA
Lane 3 - Xho I digest of OpMNPV genomic DNA
Lane 4 - Hind III digest of Lambdaphage DNA for size markers

Lane 5 - Eco RI/Bam HI digest of clone GV10
Lane 6 - same as Lane 2
Lane 7 - same as Lane 3
Lane 8 - same as Lane 4

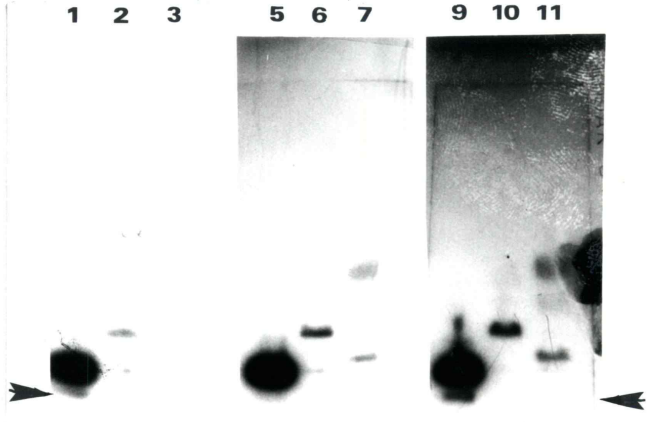
Lane 9 - Eco RI/Bam HI digest of clone GV14
Lane 10 - same as Lane 2
Lane 11 - same as Lane 3

B: Autoradiograms of Southern blot of A, probed with TnGV granulin gene

The strong signal in Lanes 1, 5 and 9 is the result of an extensive region of nucleotide identity between pBR325, the vector in which the inserts of clones GV3, GV10 and GV 14 were cloned, and pUC8, in which the TnGV granulin-containing fragment was cloned. Just below this signal, in Lanes 1 and 9 but not in 5, is a weaker signal resulting from hybridization between the TnGV probe insert and an insert fragment common to GV 3 and GV14. An arrow in A, B and C indicates the appropriate band.



A



B

Figure 4

Mapping and Sequencing of the Granulin Genes

Initial mapping of the granulin-containing fragments from both InGV and PbGV was with double restriction digests. Subsequent fine mapping to localize the gene on the fragment was done by hybridizing OpSNPV to blots of various digests. Figures 5 and 6 show restriction maps and indicate the sequencing strategies for the genomic clones of InGV and PbGV, respectively.

During the course of the sequencing it was realized that an Eco R1 site was conserved towards the 3' end of both genes, 12 amino acid residues (36 bases) from their -COOH termini. It was thus necessary to identify fragments containing the remainder of the two genes. Using the 1.3 kb Eco R1 InGV fragment as a probe a Sal 1 fragment of approximately 2 kb was identified as containing the entire InGV granulin gene (Fig. 5). An M13 0.2 kb Bam H1/Eco R1 subclone of the PbGV granulin gene, constructed for use in sequencing, was used as a probe to identify a 6 kb Bam H1/Sal 1 fragment which contained the remainder of the PbGV gene (Fig. 6). Both fragments were cloned into pUC vectors, from which appropriate M13 subclones were derived for carboxy-terminal and 3'-flanking region sequencing.

Figure 5. TnGV granulin restriction map and sequencing strategy.

The thick line indicates the location of the granulin gene.
N = 5' end of the gene
C = 3' end of the gene

Arrows below the gene indicate M13 clones which were used in sequencing.

E = Eco RI
S = Sal I
Hp = Hpa II
Hd = Hind III
Hc = Hinc II

from Akiyoshi et al, 1985

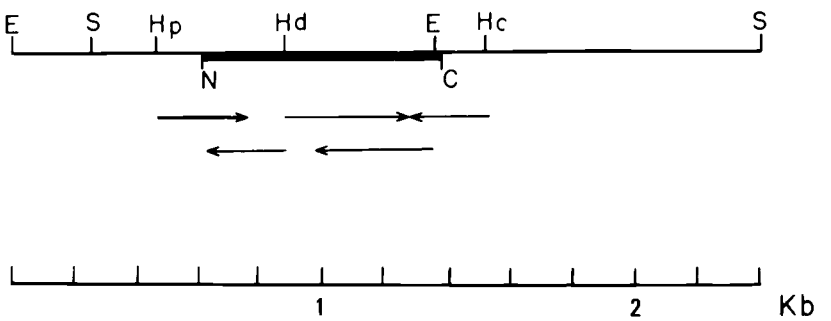


Figure 5

Figure 6. PbGV granulin restriction map and sequencing strategy.

The thick line indicates the location of the granulin gene.
N = 5' end of the gene
C = 3' end of the gene

Arrows below the gene indicate M13 clones which were used in sequencing.

E = Eco RI
Hc = Hinc II
Hp = Hpa II
B = Bam HI
K = Kpn I
S = Sal I

from Chakerian et al, in press

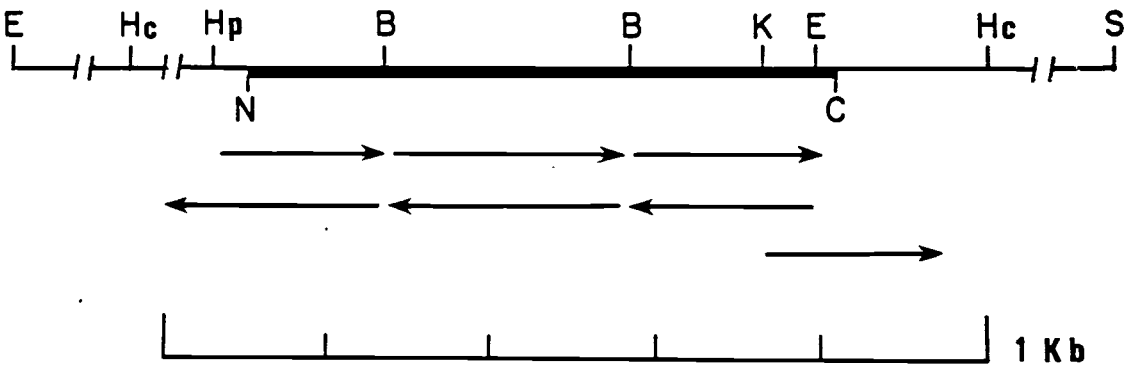


Figure 6

Characterization of Polyhedrin and Granulin: Genes and Proteins

A central question in Baculovirus studies has been the relation between GVs and NPVs. Aside from, and perhaps associated with, the differences in the biology of their infection has been the question about the differences between polyhedrin and granulin themselves. Since the major force of selection on these genes must be at the protein rather than the nucleic acid level, comparisons of amino acid rather than nucleotide sequences seem most apposite. Nevertheless, complete sequences are available for 3 polyhedrins (OpMNPV and OpSNPV, Leisy et al, in Ms.; AcMNPV, Hooft van Iddekinge et al, 1983) and the 2 granulins considered in this thesis, and some brief observations can usefully be made at this level. For convenience, and because the focus here is primarily on granulin, only one polyhedrin sequence, from AcMNPV, will be compared in detail with the granulins (Figures 7 and 8).

A base-for-base comparison of the NPV and GV nucleotide sequences shows virtually identical coding region lengths for both groups. An uninterrupted reading frame from an initial ATG codon to a UAA termination codon in all 3 sequences suggests that neither the granulin nor the polyhedrin genes contain introns. The molecular weights of

Figure 7. Nucleotide sequences of PbGV granulin, TnGV granulin and AcMNPV polyhedrin.

Dots indicate homology with PbGV sequence.
Dashes indicate missing nucleotides.

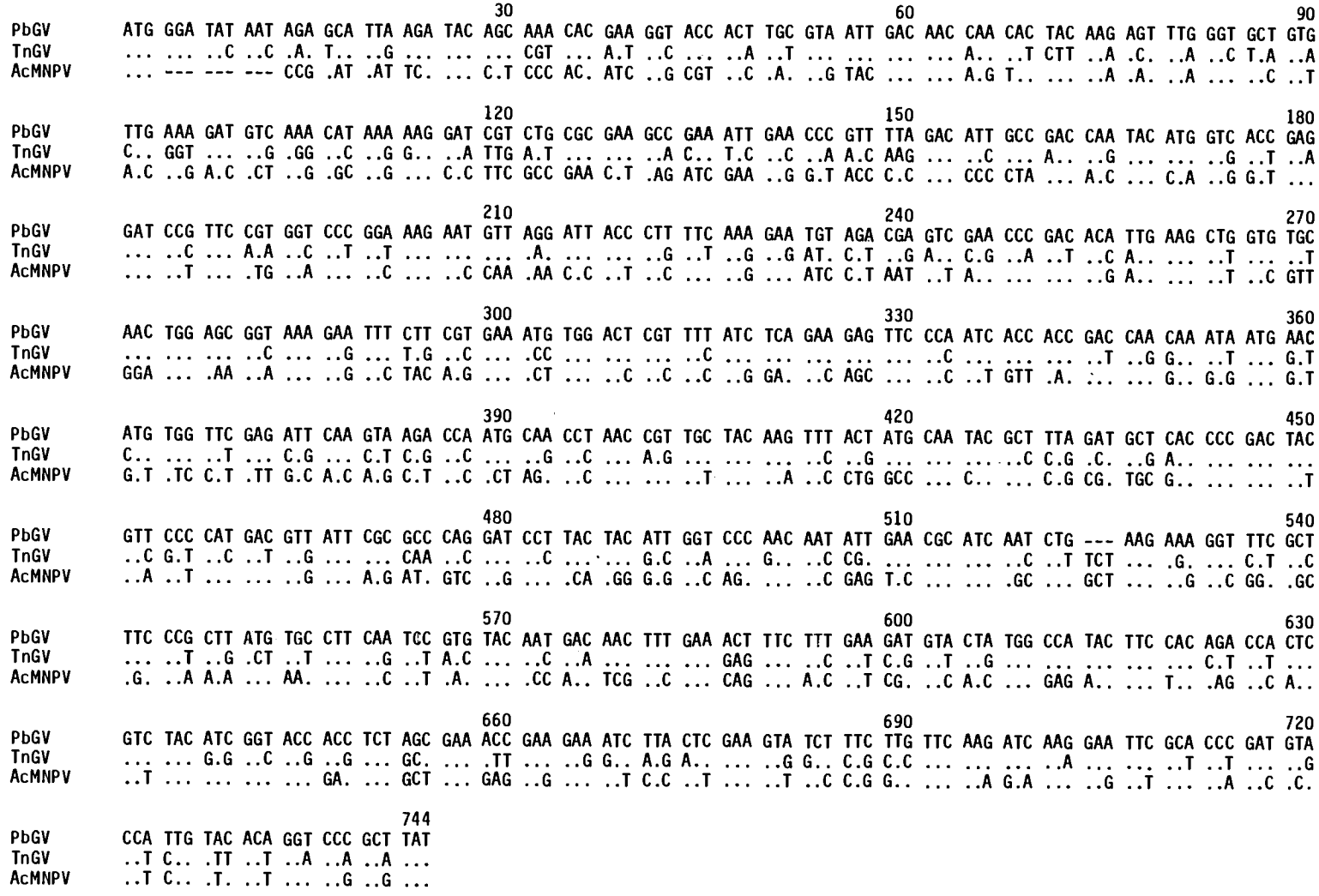

30
60
90
PbGV ATG GGA TAT AAT AGA GCA TTA AGA TAC AGC AAA CAC GAA GGT ACC ACT TGC GTA ATT GAC AAC CAA CAC TAC AAG AGT TTG GGT GCT GTG
TnGV ... .C ..C .A. T.. ..G ... ... CGT ... A.T ..C ... ..A ..T ... ... A.. ..T CTT ..A .C. ..A ..C T.A ..A
AcMNPV ... --- --- --- CCG .AT .AT TC. ... C.T CCC AC. ATC ..G CGT ..C .A. ..G TAC ... ... A.G T.. ... ..A .A. ..A ... ..C ..T
120
150
180
PbGV TTG AAA GAT GTC AAA CAT AAA AAG GAT CGT CTG CGC GAA GCC GAA ATT GAA CCC GTT TTA GAC ATT GCC GAC CAA TAC ATG GTC ACC GAG
TnGV C.. GGT ... ..G .GG ..C ..G G.. ..A TTG A.T ... .A C.. T.C ..C ..A A.C AAG ... ..C ... A.. ..G ... .G ..T ..A
AcMNPV A.C ..G A.C .CT ..G .GC ..G ... C.C TTC GCC GAA C.T .AG ATC GAA ..G G.T ACC C.C ... CCC CTA ... A.C ... C.A ..G G.T ...
210
240
270
PbGV GAT CCG TTC CGT GGT CCC GGA AAG AAT GTT AGG ATT ACC CTT TTC AAA GAA TGT AGA CGA GTC GAA CCC GAC ACA TTG AAG CTG GTG TGC
TnGV ... ..C ... A.A ..C ..T ..T ... ... .A. ... .G ..T ..G ..G AT. C.T ..G A.. C.G ..A ..T ..C A.. ... ..T ... ..T
AcMNPV ... ..T ... .TG ..A ... ..C ... ..C CAA .AA C.C ..T ..C ... ..G ... ATC C.T AAT ..T A.. ... .G A.. ... ..T ..C GTT
300
330
360
PbGV AAC TGG AGC GGT AAA GAA TTT CTT CGT GAA ATG TGG ACT CGT TTT ATC TCA GAA GAG TTC CCA ATC ACC ACC GAC CAA CAA ATA ATG AAC
TnGV ... ...C ... ..G ... T.G ..C ... .CC ... ...C ... ...C ... ...T ..G G.. ..T ... G.T
AcMNPV GGA ... .AA ..A ... ..G ..C TAC A.G ... .CT ... ..C ..C ..C ..G GA. ..C AGC ... ..C ..T GTT .A. ... ... G.. G.G ... G.T
390
420
450
PbGV ATG TGG TTC GAG ATT CAA GTA AGA CCA ATG CAA CCT AAC CGT TGC TAC AAG TTT ACT ATG CAA TAC GCT TTA GAT GCT CAC CCC GAC TAC
TnGV C.. ... ..T ... C.G ... C.T C.G ..C ... ..G ..C ... A.G ... ...C ..G ... ...C C.G .C. ..G A.. ... ...
AcMNPV G.T .TC C.T .TT G.C A.C A.G C.T ..C .CT AG. ..C ... .T ... ..A ..C CTG GCC ... C.. ... C.G CG. TGC G.. ... .T
480
510
540
PbGV GTT CCC CAT GAC GTT ATT CGC GCC CAG GAT CCT TAC TAC ATT GGT CCC AAC AAT ATT GAA CGC ATC AAT CTG --- AAG AAA GGT TTC GCT
TnGV ..C G.T ..C ..T ..G ... ... CAA ..C ... ..C ... ... G.C ..A ... G.. ..C CG. ... ...C ..T TCT ... .G. ... C.T ..C
AcMNPV ..A ..T ... .G ... A.G AT. GTC ..G ... .CA .GG G.G ..C AG. ... ..C GAG T.C ... ... .GC ... GCT ... ..G ..C GG. .GC
570
600
630
PbGV TTC CCG CTT ATG TGC CTT CAA TCC GTG TAC AAT GAC AAC TTT GAA ACT TTC TTT GAA GAT GTA CTA TGG CCA TAC TTC CAC AGA CCA CTC
TnGV ... ..T ..G .CT ..T ... ..G ..T A.C ... ..C ..A ... ... GAG ... ..C ..T C.G ..T ..G ... ... C.T ..T ...
AcMNPV .G. ..A A.A ... AA. ... ..C ..T .A. ... .CC A.. TCG ..C ... CAG ... A.C ..T CG. ..C A.C ... GAG A.. ... T.. .AG ..C A..
660
690
720
PbGV GTC TAC ATC GGT ACC ACC TCT AGC GAA ACC GAA GAA ATC TTA CTC GAA GTA TCT TTC TTG TTC AAG ATC AAG GAA TTC GCA CCC GAT GTA
TnGV ... ... G.G ..C ..G ..G ... GC. ... .TT ... ..G G.. A.G A.. ... ..G G.. C.G C.C ... ...A ... .T ..T ... ..G
AcMNPV ..T ... ... GA. ... GCT ... GAG ..G ... ..T C.C ..T ... ..T ..C C.G G.. ... ..A G.A ... ..G ..T ... ..A ..C .C.
744
PbGV CCA TTG TAC ACA GGT CCC GCT TAT
TnGV ..T C.. .TT ..T ..A ..A ..A ...
AcMNPV ..T C.. .T. ..T ... ..G ..G ...

Figure 7

Figure 8. Derived amino acid sequences of PbGV granulin, TnGV granulin and AcMNPV polyhedrin.

Translation of the nucleotide sequence shown in Figure 7. Dots indicate residues homologous with PbGV sequence. Dashes indicate missing residues.

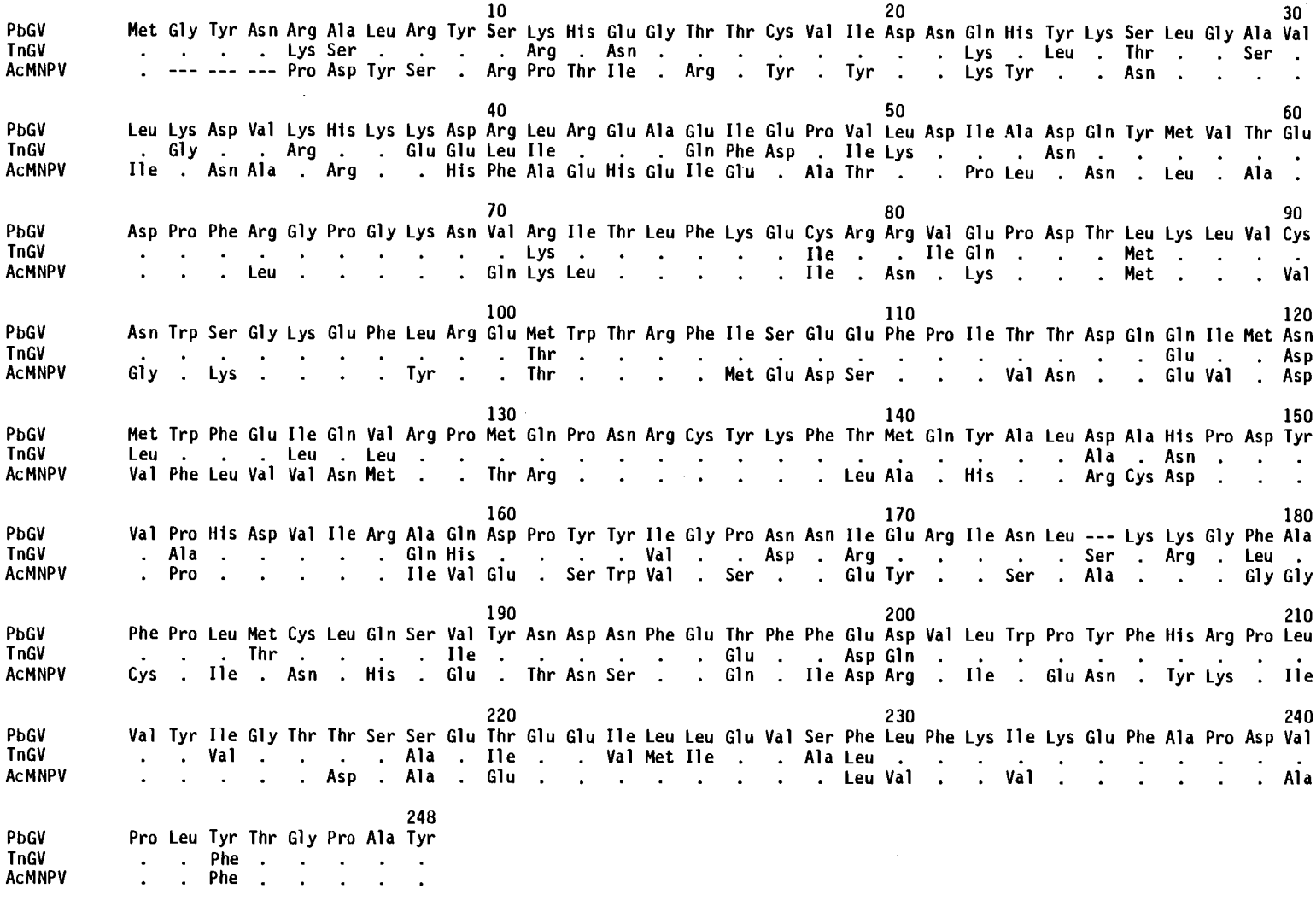

10 20 30
PbGV Met Gly Tyr Asn Arg Ala Leu Arg Tyr Ser Lys His Glu Gly Thr Thr Cys Val Ile Asp Asn Gln His Tyr Lys Ser Leu Gly Ala Val
TnGV . . . . Lys Ser . . . . Arg . Asn . . . . Lys . Leu . Thr . . Ser .
AcMNPV . --- --- --- Pro Asp Tyr Ser . Arg Pro Thr Ile . Arg . Tyr . Tyr . . Lys Tyr . . Asn . . . .
40 50 60
PbGV Leu Lys Asp Val Lys His Lys Lys Asp Arg Leu Arg Glu Ala Glu Ile Glu Pro Val Leu Asp Ile Ala Asp Gln Tyr Met Val Thr Glu
TnGV . Gly . . Arg . . Glu Glu Leu Ile . . . Gln Phe Asp . Ile Lys . . . Asn . . . . . .
AcMNPV Ile . Asn Ala . Arg . . His Phe Ala Glu His Glu Ile Glu . Ala Thr . . Pro Leu . Asn . Leu . Ala .
70 80 90
PbGV Asp Pro Phe Arg Gly Pro Gly Lys Asn Val Arg Ile Thr Leu Phe Lys Glu Cys Arg Arg Val Glu Pro Asp Thr Leu Lys Leu Val Cys
TnGV . . . . Lys . . . . . . Ile . . Ile Gln . . . Met . . . .
AcMNPV . . . Leu . . . . . Gln Lys Leu . . . . . Ile . Asn . Lys . . . Met . . . Val
100 110 120
PbGV Asn Trp Ser Gly Lys Glu Phe Leu Arg Glu Met Trp Thr Arg Phe Ile Ser Glu Glu Phe Pro Ile Thr Thr Asp Gln Gln Ile Met Asn
TnGV . . . . Thr . . . . Glu . . Asp
AcMNPV Gly . Lys . . . . Tyr . . Thr . . . . Met Glu Asp Ser . . . Val Asn . . Glu Val . Asp
130 140 150
PbGV Met Trp Phe Glu Ile Gln Val Arg Pro Met Gln Pro Asn Arg Cys Tyr Lys Phe Thr Met Gln Tyr Ala Leu Asp Ala His Pro Asp Tyr
TnGV Leu . . . Leu . Leu . . . . . Ala . Asn . . .
AcMNPV Val Phe Leu Val Val Asn Met . . Thr Arg . . . . . Leu Ala . His . . Arg Cys Asp . . .
160 170 180
PbGV Val Pro His Asp Val Ile Arg Ala Gln Asp Pro Tyr Tyr Ile Gly Pro Asn Asn Ile Glu Arg Ile Asn Leu --- Lys Lys Gly Phe Ala
TnGV . Ala . . . . . Gln His . . . . Val . . Asp . Arg . . . . . Ser . Arg . Leu .
AcMNPV . Pro . . . . . Ile Val Glu . Ser Trp Val . Ser . . Glu Tyr . . Ser . Ala . . . Gly Gly
190 200 210
PbGV Phe Pro Leu Met Cys Leu Gln Ser Val Tyr Asn Asp Asn Phe Glu Thr Phe Phe Glu Asp Val Leu Trp Pro Tyr Phe His Arg Pro Leu
TnGV . . . Thr . . . . . Ile . . . . . Glu . . Asp Gln . . . . . .
AcMNPV Cys . Ile . Asn . His . Glu . Thr Asn Ser . . Gln . Ile Asp Arg . Ile . Glu Asn . Tyr Lys . Ile
220 230 240
PbGV Val Tyr Ile Gly Thr Thr Ser Ser Glu Thr Glu Glu Ile Leu Leu Glu Val Ser Phe Leu Phe Lys Ile Lys Glu Phe Ala Pro Asp Val
TnGV . . Val . . . . Ala . Ile . . Val Met Ile . . Ala Leu . . . . . .
AcMNPV . . . . . Asp . Ala . Glu . . . . . Leu Val . . Val . . . . . Ala
248
PbGV Pro Leu Tyr Thr Gly Pro Ala Tyr
TnGV . . Phe . . . . .
AcMNPV . . Phe . . . . .

Figure 8

the derived amino acid sequences sum to roughly 29,000 daltons each, in close agreement to generally reported molecular weights for polyhedrin and granulin.

Nine nucleotides near the 5' end of the granulin sequences are absent in the NPVs, PbGV is missing nucleotides 523-525 of the TnGV sequence, and OpSNPV has 3 more nucleotides near its 5' end compared to OpMNPV and AcMNPV. The significance of these differences is difficult to assess solely on the basis of sequence data.

Of the nucleotides common to the 3 known polyhedrin sequences about 75% are conserved (Rohrmann, 1985), while about 73% of the common nucleotides are conserved between the two granulins. About half of the differences within each of the groups are third-base wobbles; the other half result in coding changes.

Using PbGV and AcMNPV sequences as representative of granulin and polyhedrin, respectively, only 59% of the common nucleotides are conserved. Of the nucleotides which differ between AcMNPV and PbGV, roughly 75% result in coding changes.

It is instructive to note, parenthetically, that the overall nucleotide sequence homology of 59% between PbGV and AcMNPV is well below the minimum 67% homology theoretically required for DNA-DNA hybridization under even the least stringent conditions. The final 150 bases of the carboxy terminal region, however, are roughly 67% homologous, and could thus account for the observed hybridizations.

It has been suggested that genes which are hyper-expressed, such as the polyhedrin and granulin genes, may exhibit "non-random selection" of codons (Grantham et al, 1981; McLachlan, Staden and Boswell, 1984). Analysis of the 3 polyhedrin sequences does not show a particularly skewed codon selection (Vlak & Rohrmann, in press), with about 50 of 61 possible codons being used. A similar situation is found in the 2 granulin sequences, with about 55 codons occurring in each gene (data not shown).

It is likely that tertiary structure is influential in the transportation of NPV polyhedrins across the nuclear membrane, as well as for the attachment of polyhedrin and granulin to the viral envelope and their subsequent quaternary structure. Cysteine and proline residues contribute significantly, although not exclusively, to higher-order protein structure, and although the overall number of cysteine and proline residues is similar in GVs and NPVs, the pattern of their distribution is somewhat - and perhaps importantly - different.

Of 20 total proline sites, 13 are found in all 3 sequences (Fig. 9). The GVs have 2 more in common between them, while AcMNPV has 2 prolines near its 5' terminus, one of which also occurs in the other 5 lepidopteran NPV polyhedrin sequences which have been determined (Table 4). A proline in both GVs at residue 48 may be functionally equivalent to an AcMNPV proline at residue 52. InGV seems to have undergone a transversion within its codon 152,

Figure 9. Location of Proline and Cysteine residues in PbGV granulin, TnGV granulin and AcMNPV polyhedrin

Same Figure as Fig. 8, except that Proline residues are highlighted in red and Cysteine residues are highlighted in green.

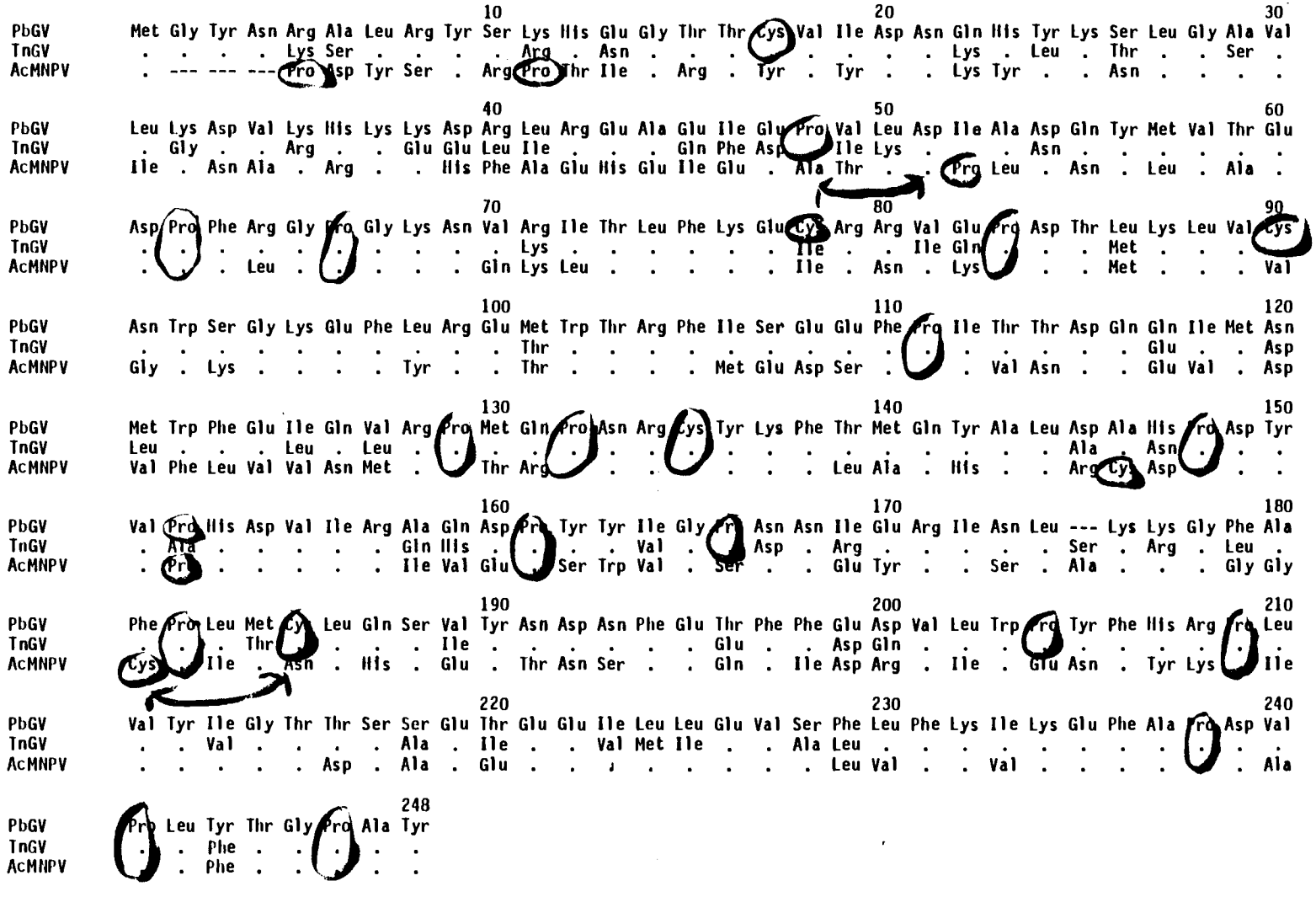

10 20 30
PbGV Met Gly Tyr Asn Arg Ala Leu Arg Tyr Ser Lys His Glu Gly Thr Thr Cys Val Ile Asp Asn Gln His Tyr Lys Ser Leu Gly Ala Val
TnGV . . . . Lys Ser . . . . . Arg . Asn . . . . . Lys . Leu . Thr . . Ser .
AcMNPV . --- --- --- Pro Asp Tyr Ser . Arg Pro Thr Ile . Arg . Tyr . Tyr . . Lys Tyr . . Asn . . . .
40 50 60
PbGV Leu Lys Asp Val Lys His Lys Lys Asp Arg Leu Arg Glu Ala Glu Ile Glu Pro Val Leu Asp Ile Ala Asp Gln Tyr Met Val Thr Glu
TnGV . Gly . . Arg . . Glu Glu Leu Ile . . . Gln Phe Asp Ile Lys . . . Asn . . . . . .
AcMNPV Ile . Asn Ala . Arg . . His Phe Ala Glu His Glu Ile Glu . Ala Thr . . Pro Leu . Asn . Leu . Ala .
70 80 90
PbGV Asp Pro Phe Arg Gly Pro Gly Lys Asn Val Arg Ile Thr Leu Phe Lys Glu Cys Arg Arg Val Glu Pro Asp Thr Leu Lys Leu Val Cys
TnGV . . . . Lys . . . . . Ile . . Ile Gln . . . Met . . . .
AcMNPV . . . Leu . . . . . Gln Lys Leu . . . . . Ile . Asn . Lys . . . Met . . . Val
100 110 120
PbGV Asn Trp Ser Gly Lys Glu Phe Leu Arg Glu Met Trp Thr Arg Phe Ile Ser Glu Glu Phe Pro Ile Thr Thr Asp Gln Gln Ile Met Asn
TnGV . . . . . Thr . . . . . Glu . . Asp
AcMNPV Gly . Lys . . . . Tyr . . Thr . . . . Met Glu Asp Ser . . . Val Asn . . Glu Val . Asp
130 140 150
PbGV Met Trp Phe Glu Ile Gln Val Arg Pro Met Gln Pro Asn Arg Lys Tyr Lys Phe Thr Met Gln Tyr Ala Leu Asp Ala His Pro Asp Tyr
TnGV Leu . . . Leu . Leu . . . . . Ala . Asn . . .
AcMNPV Val Phe Leu Val Val Asn Met . . Thr Arg . . . . . Leu Ala . His . . Arg Cys Asp . . .
160 170 180
PbGV Val Pro His Asp Val Ile Arg Ala Gln Asp Pro Tyr Tyr Ile Gly Pro Asn Asn Ile Glu Arg Ile Asn Leu --- Lys Lys Gly Phe Ala
TnGV . Ala . . . . . Gln His . . . . Val . . Asp . Arg . . . . . Ser . Arg . Leu .
AcMNPV . Pro . . . . . Ile Val Glu . Ser Trp Val . Ser . . Glu Tyr . . Ser . Ala . . . Gly Gly
190 200 210
PbGV Phe Pro Leu Met Cys Leu Gln Ser Val Tyr Asn Asp Asn Phe Glu Thr Phe Phe Glu Asp Val Leu Trp Pro Tyr Phe His Arg Pro Leu
TnGV . . . Thr . . . . . Ile . . . . . Glu . . Asp Gln . . . . . .
AcMNPV Cys . Ile . Asn . His . Glu . Thr Asn Ser . . Gln . Ile Asp Arg . Ile . Glu Asn . Tyr Lys . Ile
220 230 240
PbGV Val Tyr Ile Gly Thr Thr Ser Ser Glu Thr Glu Glu Ile Leu Leu Glu Val Ser Phe Leu Phe Lys Ile Lys Glu Phe Ala Pro Asp Val
TnGV . . Val . . . . Ala . Ile . . Val Met Ile . . Ala Leu . . . . . .
AcMNPV . . . . . Asp . Ala . Glu . . . . . Leu Val . . Val . . . . . Ala
248
PbGV Pro Leu Tyr Thr Gly Pro Ala Tyr
TnGV . . Phe . . . . .
AcMNPV . . Phe . . . . .

Figure 9

Table 4. Conservation of Cysteine and Proline Residues Among Eight Baculoviruses

derived from Rohrmann, 1985

TABLE 4

Conservation of Cysteine and Proline Residues Among 8 Baculoviruses

	10	20	30	40	50	60	70	80	90	100	110	120	130
AcMNPV	P P				P	P P		P			P		
GmMNPV	P P					P P		P			P		
OpMNPV	P P				P	P P		P			P		
LdMNPV	P					P P		P			P		
BmSNPV	P					P P		P			P		
OpSNPV	P				P	P P		P			P		
PbGV		C			P	P P		C P	C		P		
TnGV		C			P	P P		P	C		P		

	140	150	160	170	180	190	200	210	220	230	240	250
AcMNPV	P P C	C P P	P		CP			P			P P	P
GmMNPV	P P C	C P	P		CP			P			P P	P
OpMNPV	P P C	C P	P		CP			P			P P	P
LdMNPV	P P C	C P		P	CP			P			P P	P
BmSNPV	P P C	C P	P		CP			P			P P	P
OpSNPV	P P C	C P P	P		CP			P			P P	P
PbGV	P P C	P P	P	P	P C			P P			P P	P
TnGV	P P C	P	P	P	P C			P P			P P	P

changing the residue from the proline which occurs in PbGV and AcMNPV (and the other 5 lepidopteran NPVs) to alanine.

Of 7 cysteine sites, only one, at residue 135, is conserved among all 3 sequences as well as in the other 5 lepidopteran NPVs. Cysteine residues at 185 in the GVs are probably functionally equivalent to the cysteine residue at 181 found in all 6 lepidopteran NPVs. The GVs have 2 other cysteine residues in common, while the lepidopteran NPVs have 1 other. PbGV also has a unique cysteine residue at 78.

It is interesting to place these proline and cysteine residues against a background of "regional" amino acid sequence homology: prolines and cysteines are generally embedded in or associated with extended regions of sequence homology, indicating perhaps a strong selective pressure to maintain a specific conformation near or around these conserved sites (Fig. 9).

The region around the conserved cysteine at 135 is very highly conserved and includes within it conserved prolines at 129 and 132. The hypothesis that prolines and cysteines are important to the higher-order conformation of polyhedrin and granulin might be tested by deleting or mutating all of part of this region and assaying for polyhedrin and granulin higher-order structure formation. A similar experiment in the carboxy-terminus sequence, which contains 3 prolines in a very highly conserved region, might also be instructive.

Phylogenetics and Evolution of the Baculoviridae

The analysis of Baculovirus evolution using amino acid sequence data has already been elaborated in considerable detail elsewhere (Rohrmann, 1985; Vlak & Rohrmann, in press). Because the two GV sequences presented in this thesis formed a significant part of the data on which that phylogeny was based, the general conclusions of the study are presented here.

The broad outline of Baculovirus evolution may be inferred from what is known of their distribution (Table 1). It seems likely that the genus arose in the Arthropoda at some early point, perhaps before the divergence of the Insecta from the Crustacea, given the occurrence of two Crustacean Baculovirus hosts. Within the Insecta, Baculoviruses are confined to some, but not all, of the holometabolous orders and are not known to infect any of the hemimetabolous orders. This may result from some general aspect of hemimetabolous physiology which excludes the Baculoviridae from infecting those orders. An alernative phylogeny, however, given that there are only two Crustaceans which are reported Baculovirus hosts, is that the Baculovirdae arose initially in the holometabolous insects and horizontally transferred to the two Crustacean hosts.

Since SNPVs infect all insect orders in which Baculovirus infection occurs, they are most likely the

original Baculovirus type. Conversely, since MNPVs and GVs infect only the Lepidoptera (the most recently evolved insect order), it seems likely that these forms of Baculovirus are relatively recent evolutionary morphotypes.

Construction of the species-specific Baculovirus phylogeny was based on comparisons of amino acid sequences of polyhedrin and granulin from 10 insect species from 3 orders (Table 2). Polyhedrin and granulin are considered to be "orthologous", or functionally equivalent, proteins because of their highly conserved morphology and particularly because their resistance to environmental degradation and susceptibility to dissolution in alkaline conditions indicate functional conservation.

In a fashion analagous to phylogenies constructed using Cytochrome C sequences from a considerably more diverse group of species, polyhedrin and granulin sequences formed the basis for studies of the molecular evolution of the Baculoviridae. Phylogenetic relationships among the 8 NPV polyhedrins and the 2 GV granulins were quantified by comparing their protein sequences and scoring for divergences. A phylogenetic tree, without any reference to absolute time, was then drawn both for them and, by extension, for the viruses by which they are coded (Figure 10).

The groupings within the Lepidoptera of the GVs and the NPVs are not unexpected. The sequence-based phylogeny confirms previous immunological studies (Rohrmann et al,

Figure 10. Baculovirus phylogeny.

The numbers indicate the per cent amino acid difference between points on the tree.

from Rohrmann, 1985

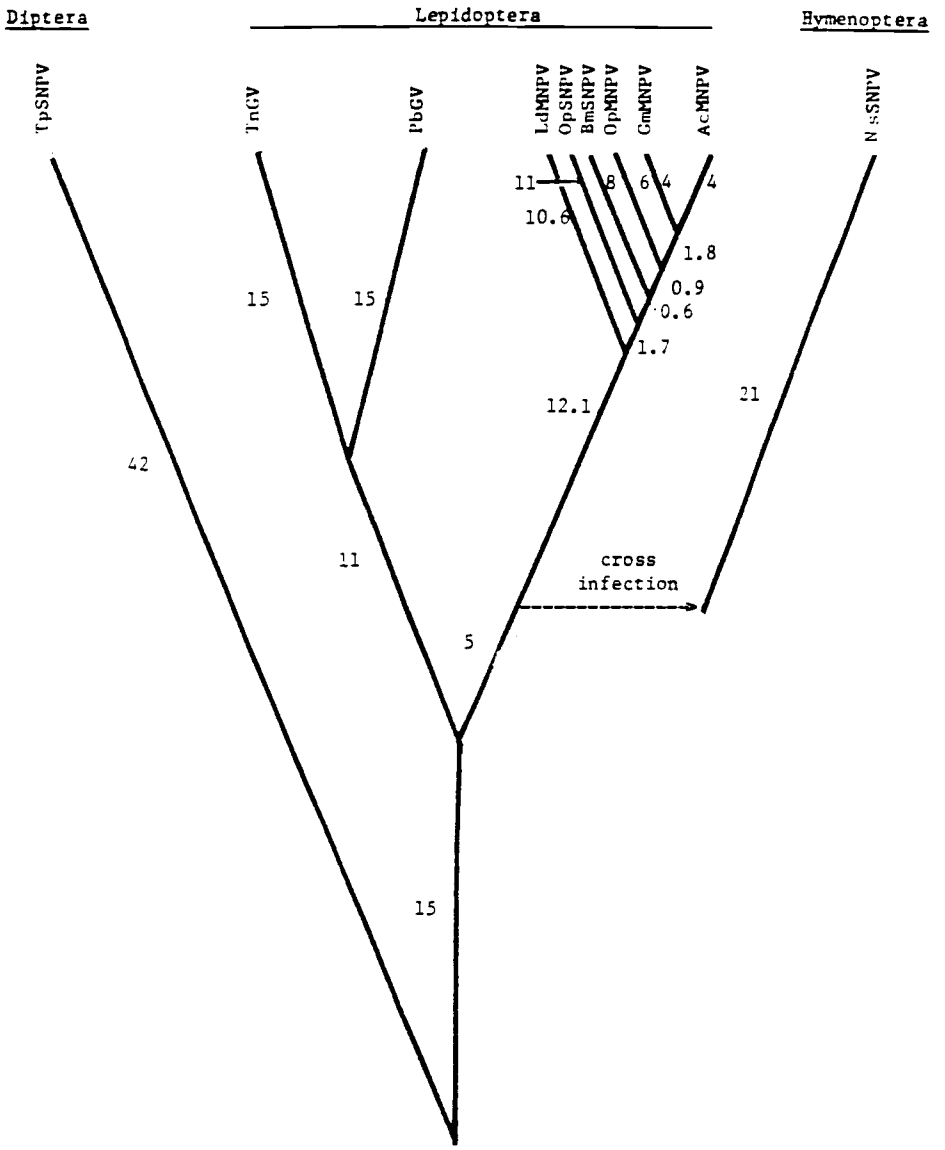

Diptera
Lepidoptera
Hymenoptera
TpSNPV
TnGV
PbGV
LdMNPV
OpSNPV
BmSNPV
OpMNPV
GmMNPV
AcMNPV
NsSNPV
11
8
6
4
4
10.6
1.8
0.9
0.6
1.7
15
15
12.1
21
42
11
cross
infection
5
15

Figure 10

1981) which had shown no antigenic relatedness between the dipteran SNPV and either a lepidopteran SNPV or MNPV. The same studies had shown some antigenic relation between the 2 lepidopteran NPVs and a hymenopteran SNPV: scoring of the amino acid sequence data confirmed the immunological data and suggested the construction of the cross infection from the lepidopteran NPV line to the hymenopteran SNPV.

It is interesting to compare the sequence-derived Baculovirus phylogeny of Figure 10 with a schematic, i.e., not quantified, phylogeny of Baculovirus hosts based on generally accepted taxonomic relationships (Figure 11). The Baculoviruses whose sequences were used to construct the phylogeny of Figure 10 have been placed on Figure 11 purely for illustrative purposes. It can readily be seen that the groupings based on polyhedrin and granulin sequence data are not consistent with classical phylogenetic constructs: for instance, while the sequence data for the two Gvs indicates close relationship between them, phylogenies based on morphological characteristics have placed the Pierids in the Papilionoidea and the Plusiinae in the Noctuoidea. Whether this inconsistency results from errors in the earlier phylogenies, or simply indicates that the Baculoviridae evolved independently of their hosts, can not be determined from the available data.

The fossil record for insects is inevitably limited. It may be expected, however, that at some future time enough fossil data will be available to enable the construction of

Figure 11. A schematic Baculovirus host phylogeny.

Unlike Figure 10, this phylogeny is merely schematic and does not attempt to quantify relationships. It is a host phylogeny, rather than a Baculovirus phylogeny, but Baculoviruses whose polyhedrin or granulin sequences have been determined in part or in whole have been placed in their host phylum for purposes of comparison with the phylogeny of Figure 10.

The Lepidopteran phylogeny is derived from Brock, 1971. The relationship of the Diptera, Hymenoptera and Lepidoptera is generally accepted as presented.

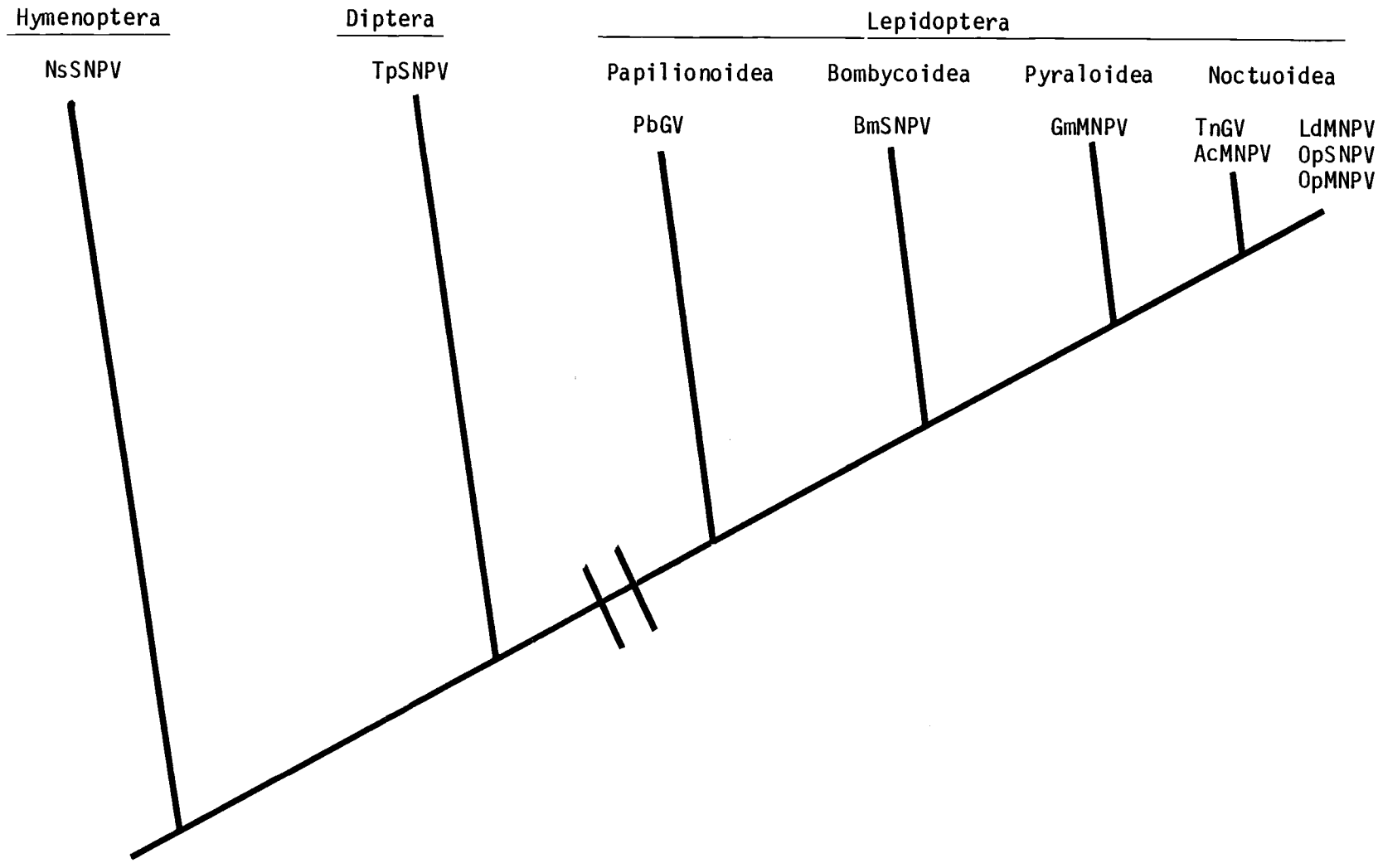

Hymenoptera
Diptera
Lepidoptera
NsSNPV
TpSNPV
Papilionoidea
Bombycoidea
Pyraloidea
Noctuoidea
PbGV
BmSNPV
GmMNPV
TnGV
AcMNPV
LdMNPV
OpSNPV
OpMNPV

Figure 11

insect phylogenetic lines with greater precision than is now possible. At that time Baculovirus phylogeny - perhaps augmented by further polyhedrin and granulin sequences, or by sequences of other orthologous proteins such as Cytochrome C - can be compared to the insect phylogeny, and the very interesting question of the possible co-evolution of the Baculoviridae with their hosts can be addressed. This has already been done for a group of Papovaviruses which infect mammalian hosts (Soeda et al, 1980).

Regulation of Polyhedrin and Granulin Gene Expression

The regulation of polyhedrin and granulin expression is of considerable interest not only because the proteins are produced in such extremely high concentrations from a single-copy gene but as well because they are expressed essentially only in the latter part of a "biphasic" infection cycle. Specific regulatory mechanisms are currently being investigated both for their potential application in eukaryotic expression vectors as well as for their intrinsic biological interest.

Fuchs et al (1983) identified an alpha-amanitin-resistant RNA polymerase activity in Spodoptera frugiperda cells infected with AcNPV. This activity increased over the course of viral infection until at 24 hours postinfection RNA synthesis was completely resistant. It seems probable either that there is a

viral-coded RNA polymerase "factor" which confers alpha-amanitin resistance to the host polymerase, or that the virus codes for an entirely novel polymerase II. In either case, the co-incidence of the timing of polyhedrin expression with the rise in resistant polymerase activity begs several questions: Is the resistant activity specific to late, hypertranscribed viral genes, and, if so, what is the mechanism which effects that specificity?

A different but perhaps complementary line of evidence can be adduced from an inspection of sequences which are likely to be involved in the regulation of expression of these genes. Comparison of the 5' flanking sequences from the 2 granulin genes considered here, together with those from the polyhedrin genes of AcMNPV (Hooft van Iddekinge et al, 1983), BmSNPV (Iatros, Ito and Witkiewicz, 1985), OpMNPV and OpSNPV (Leisy et al, in Ms.), and from another hypertranscribed late gene, the "10k protein" of AcMNPV (Lubbert and Doerfler, 1984), reveals a sequence of conserved nucleotides which might serve as a recognition site for a regulatory molecule (figure 12).

The region in which this sequence occurs is generally AT rich, a frequent characteristic of regulatory regions. Since an AT rich sequence has fewer hydrogen bonds than would be found in a sequence of random nucleotides, the strands of the double helix are easily dissociated and individual strands are thus more accessible to DNA-binding regulatory molecules (e.g., proteins or nucleic acids). In

Figure 12. 5' flanking nucleotide sequences of 7* hyperexpressed Baculovirus late genes.

The conserved sequence of 12 nucleotides is underlined in all 6 genes. Putative TATA boxes are underlined in AcMNPV polyhedrin and "10k" protein. Arrows above the sequences for AcMNPV polyhedrin and "10k" protein and OpMNPV indicate cap sites for mRNA. Numbers above sequences indicate distance from beginning of coding sequence.

*: A seventh 5' flanking region, for BmSNPV, was published during the final drafting of this thesis. It is identical to the AcMNPV 5' flanking sequence, except that an A replaces the G immediately following the AcMNPV putative TATA box, and the A at position -1 in AcMNPV is missing in BmMNPV. The first 6 bases of the AcMNPV putative TATA box are also missing.

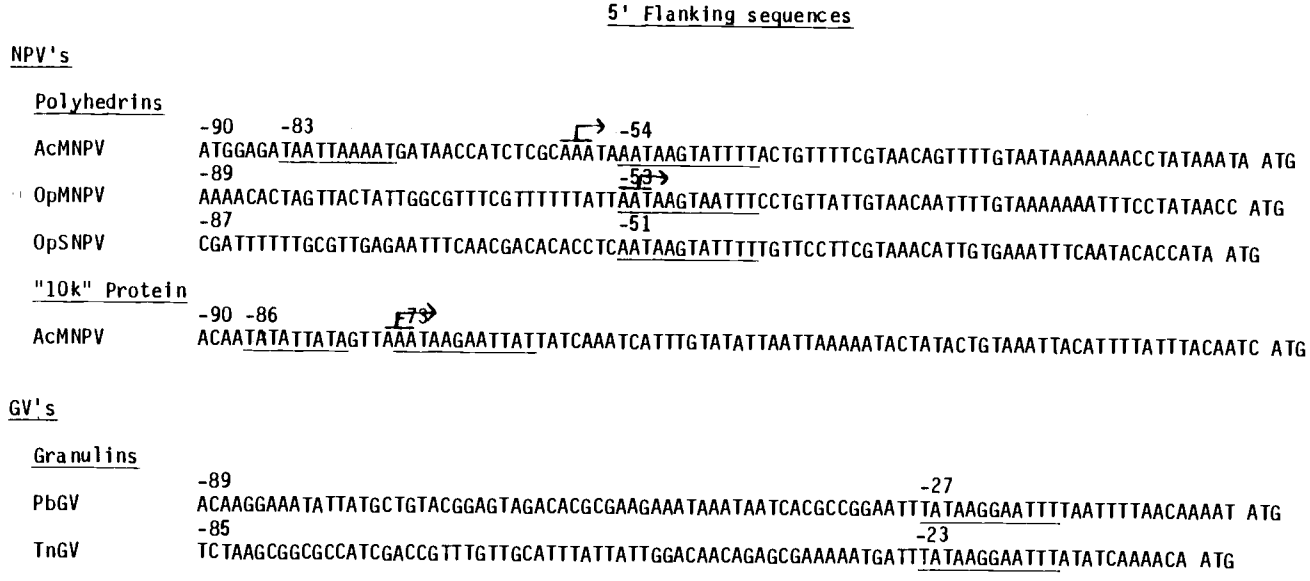

5' Flanking sequences
NPV's
Polyhedrins
-90 -83 -54
AcMNPV ATGGAGATAATTAAAATGATAACCATCTCGCAAATAAATAAGTATTTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAATA ATG
-89 -53
OpMNPV AAAACACTAGTTACTATTGGCGTTTCGTTTTTTATTAATAAGTAATTTCCTGTTATTGTAACAATTTTGTAAAAAAATTTCCTATAACC ATG
-87 -51
OpSNPV CGATTTTTTGCGTTGAGAATTTCAACGACACACCTCAATAAGTATTTTTGTTCCTTCGTAAACATTGTGAAATTTCAATACACCATA ATG
"10k" Protein
-90 -86 -79
AcMNPV ACAATATATTATAGTTAAATAAGAATTATTATCAAATCATTTGTATATTAATTAAAAATACTATACTGTAAATTACATTTTATTTACAATC ATG
GV's
Granulins
-89 -27
PbGV ACAAGGAAATATTATGCTGTACGGAGTAGACACGCGAAGAAATAAATAATCACGCCGGAATTTATAAGGAATTTTAATTTTAACAAAAT ATG
-85 -23
TnGV TCTAAGCGGCGCCATCGACCGTTTGTTGCATTTATTATTGGACAACAGAGCGAAAAATGATTTATAAGGAATTTATATCAAAACA ATG

Figure 12

point of fact, the AcMNPV polyhedrin and "10k protein" conserved sequences are located just downstream from TATA boxes.

Since no sequences are available for any viral genes other than hypertranscribed late genes, it is not possible to say that this conserved 5' flanking region is specific for that group of genes, nor, for that matter, that it is regulatory in nature. The biphasic infection cycle of occluded Baculoviruses implies the need, however, for late gene specific regulation, and this conserved sequence suggests itself as a potential regulatory element. Given as a working hypothesis that the sequence is in fact regulatory and is late gene-specific, two simple alternatives regarding its function come to mind. On the one hand it may be a positive regulatory element, e.g., a "promoter" or perhaps a sequence functioning in a manner similar to the SV40 enhancer sequence; alternatively it may be a negative regulatory element, analagous to the operator site in the E. coli lactose operon.

For the 3 genes for which the 5' end of the message has been mapped (AcMNPV polyhedrin and "10k protein", and OpMNPV polyhedrin) this conserved sequence begins at or within 1 to 5 bases 3' to the beginning of the message (Fig. 12). The possibility that the sequence functions as a positive regulatory element seems unlikely, then, since it occurs at the beginning of, or even a few bases within, the coding sequence of the message. Intuitively it seems that

placement of a regulatory element downstream from a promoter and precisely at the start of a message would result in a negative rather than positive effect. In fact, if the sequence were recognized by a repressor-type protein, an either/or situation would exist: either the repressor would be bound to the site, and would thus block transcription, or the repressor would not be bound, in which case polymerase activity would be unobstructed.

A third set of data which may be relevant is the repeated observation of the loss of polyhedrin expression during serial passage of NPVs in cell culture (Miller and Miller, 1982; Fraser and Hink, 1982; Potter et al, 1976). Mutations induced in the polyhedrin gene and resulting in non-occluded infectious virus have shown that the polyhedrin gene is not essential for functional, infectious virus (Smith, Fraser and Summers, 1983), and in vivo the viral particles which cause secondary infection in the host larvae are not occluded. Whether the "few polyhedrin" (FP) mutants which frequently appear in cell culture result from repression of polyhedin expression or from a more fundamental disruption of the gene itself is not clear and needs to be investigated further. (It is also possible that other functions, such as envelope membrane development, may be the root cause of FP mutants.)

If the lac operator repressor analogy is correct, however, absence of polyhedrin expression could be explained by the absence of an "inducer" molecule in cell culture. It

would be interesting to see if polyhedrin expression could be restored by the addition of "factors" from NPV-infected larvae. It would also be interesting to know whether an alpha amanitin resistant RNA Polymerase II is present in NPV-infected cell culture, whether this too disappears with the loss of polyhedrin expression, and whether this also could be restored by infected-larval "factors". Of course, none of these experiments would effectively resolve the question of positive or negative regulation.

Finally, it would be very interesting to know what function the "10k" protein serves in AcMNPV, and whether its homologues can be identified in other Baculoviruses. Although it is a hyperexpressed late protein like polyhedrin and granulin, the "10k" protein is also the most abundant protein being synthesized at about 20 hours post-infection, at a time when polyhedrin is not abundant (Smith, Vlak and Summers, 1982). By 48 hours post-infection, essentially the only proteins being produced are polyhedrin and the "10k" protein, both in great abundance. However, the "10k" protein is not a major component either of extruded, non-occluded AcMNPV, i.e., virus recovered from primary infection, or of AcMNPV virions purified from polyhedrons (Smith, Vlak and Summers, 1982). It is not clear whether this latter evidence rules out the possibility that the "10k" protein may be an envelope component intimately involved in the occlusion of virions in polyhedrons. An even more interesting possibilty, though one completely without evidence at the present time,

is that the "10k" protein might functon as an inducer molecule of some sort.

GVs and NPVs have recently been proposed both as insect pest management vectors, since they are limited to the Arthropoda, essentially to the holometabolous insects, and largely to the Lepidoptera (Miller, Lingg and Bulla, 1983), and as eucaryotic expression vectors (Miller, Miller and Adang, 1983; Smith, Summers and Fraser, 1983; Smith, Fraser and Summers, 1983). The polyhedrin gene has been an obvious cloning site for the latter application, since it does not appear necessary to viral replication in cell culture and since it is expressed at such a high intracellular concentration.

Several clones constructed by splicing various parts of the AcNPV polyhedrin coding and flanking sequences with the human beta-interferon gene produced and secreted more beta-interferon - by 2 orders of magnitude - than had been reported previously for expression systems using E. coli or mouse cells (Smith, Summers and Fraser, 1983). Interestingly, in the clone which had the highest interferon expression the gene was inserted precisely at the 5' startpoint of the polyhedrin message, at -59 (see Fig. 12), and the sequence from -58 to +175, including the putative conserved regulatory sequence, was deleted. This has implications with regard to the hypothesis that the sequence serves as an RNA Polymerase II recognition site, or, for

that matter, in any other positive regulatory function.

Before NPVs can be fully utilized as eucaryotic vectors their host ranges will have to be better understood and defined. A great deal of work also remains to be done before the regulation of polyhedrin and granulin expression can be understood with any confidence. It seems likely that the regulatory mechanism(s) will be fully as interesting as are those of the *lac* operon, bacteriophage *Lambda*, and the yeast mating type locus, and probably as different from them as they are from one another. That prospect marks near-term research in Baculovirus gene regulation as a most interesting undertaking.

BIBLIOGRAPHY

1. Akiyoshi, D., R. Chakerian, G.F. Rohrmann, M.H. Nesson and G.S. Beaudreau (1985), Cloning and sequencing of the granulin gene from the Trichoplusia ni granulosis virus. Virology 141:328-332.

2. Anon. (1980), M13 Cloning/"Dideoxy"Sequencing Manual Bethesda Research Laboratories, Gaithersburg, Md.

3. Anon. (1983), Step-by-Step Protocol for ^{35}S-Sanger Sequencing. New England Nuclear, Boston, Mass.

4. Brock, J.P. (1971), A contribution towards an understanding of the morphology and phylogeny of the Ditrysian Lepidoptera. J. Nat. Hist. 5:29-102.

5. Chakerian, R., G.F. Rohrmann, M.H. Nesson, D.J. Leisy, and G.S. Beaudreau (1985), The nucleotide sequence of the Pieris brassicae granulosis virus granulin gene. J. Gen. Virol.:in press.

6. Couch, J.A. (1974), An enzootic nuclear polyhedrosis virus of pink shrimp: ultrastructure, prevalence, and enhancement. J. Invert. Path. 24: 311-331.

7. David, W.A.L. (1978), The Granulosis Virus of Pieris brassicae (L.) and its relationship with its host. Adv. in Virus Res. 22:111-161.

8. Fraser, M.J. and W.F. Hink (1982), The isolation and characterization of the MP and FP plaque variants of Galleria mellonella nuclear polyhedrosis virus. Virology 117:366-378.

9. Fuchs, Y. L., M.S. Woods, and R.F. Weaver (1983), Viral transcription during Autographa californica nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected Spodoptera frugiperda cells. J. of Virology 48(3):641-646.

10. Grantham, R., C. Gautier, M. Gouy, M. Jacobzone and R. Mercier (1981), Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Research 9(1): r43-r74.

11. Guelpa, B., M. Bergoin and G. Croizier (1977), La proteine d'inclusion et les proteines du virion du baculovirus du diptere Tipula paludosa (Meigen). C.R. Acad. Sci. 284:779-782.

12. Hohmann, A.W. and P. Faulkner (1983), Monoclonal antibodies to baculovirus structural proteins:

determination of specificities by Western blot analysis. Virology 125:432-444.

13. Holmes, D.S. and M. Quigley (1981), A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114:193.

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

15. Howley, P.M., M.A. Israel, M. Law and M.A. Martin (1979), A rapid method for detecting and mapping homology between heterologous DNAs. J. Biol. Chem. 254 :4876-4883.

16. Iatrou, K., K. Ito and H. Witkiewicz (1985), Polyhedrin gene of Bombyx mori nuclear polyhedrosis virus. J. of Virology 54(2):436-445.

17. Kaya, H.K. (1977), Transmission of a nuclear polyhedrosis virus isolated from Autographa californica to Alsophila pometaria, Hyphantria cunea, and other forest defoliators. J. Econ. Entom. 70:9-12.

18. Kozlov, E.A., T.L. Levitina, N.M. Gusak and S.B. Serebryani (1981), Comprison of the amino acid sequences

of inclusion body proteins of nuclear polyhedrosis viruses of Bombyx mori , Porthetria dispar and Galleria mellonella. Bioorg. Khim. 7:1008.

19. Krywienczyk, J. and G.H. Bergold (1961), Serological studies of inclusion-body proteins by agar diffusion technique. J. Insect Pathol. 3:15-28.

20. Leisy, D.J., G.F. Rohrmann, M.H. Nesson, M.N. Pearson and G.S. Beaudreau (1985), The nucleotide sequence of polyhedrin genes from two Orgyia pseudotsugata baculoviruses. In Ms.

21. Lightner, D.V. and R.M. Redman (1981), A baculovirus-caused disease of the Penaeid Shrimp Penaeus monodon. J. Invertebrate Pathology 38:299-302.

22. Lubbert, H. and W. Doerfler (1984), Transcription of overlapping sets of RNAs from the genome of Autographa californica nuclear polyhedrosis virus: a novel method for mapping RNAs. J. of Virology 52(1):255-265.

23. Maniatis, T., E.F. Fritsch and J. Sambrook (1982), Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

24. Maruniak, J.E. and M.D. Summers (1978), Comparative

peptide mapping of baculovirus polyhedrins. J. Invert. Path. 32:196-201.

25. McCarthy, W.J. and J.T. Lambiase (1979), Serological relationships among Plusiinae bacloviruses. J. Invert. Path. 34:170-177.

26. McLachlan, A.D., R. Staden and D.R. Boswell (1984), A method for measuring the non-random bias of a codon usage table. Nucleic Acids Research 12(24):9567-9575.

27. Miller, L.K., A.J. Lingg and L.A. Bulla, Jr. (1983), Bacterial, viral, and fungal insecticides. Science 219 :715-721.

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

29. Miller, L.K., D.W. Miller and M.J. Adang (1983), "An Insect Virus for Genetic Engineering: Developing Baculovirus Polyhedrin Substitution Vectors," pp. 89-97, in Genetic Engineering in Eukaryotes, ed. by P.F. Lurquin and A. Kleinhofs. Plenum Press, New York.

30. Potter, K. N., P. Faulkner, E.A. MacKinnon (1976), Strain selection during serial passage of Trichoplusia

ni nuclear polyhedrosis virus. J. of Virology 18 (3):1040-1050.

31. Quant, R.L., M.N. Pearson, G.F. Rohrmann and G.S. Beaudreau (1984), Production of polyhedrin monoclonal antibodies for distinguishing two Orgyia pseudotsugata baculoviruses. Applied and Environmental Microbiology 48 (4):732-736.

32. Rohrmann, G.F. (1985), "Baculovirus relatedness and evolution," in The Biology of Baculoviruses , Vol. 1, ed. by R. Granados and B. Federici. CRC Press, N.Y.

33. Rohrmann, G.F., T.J. Bailey, B. Brimhall, R.R. Becker and G.S. Beaudreau (1979), Tryptic peptide analysis and NH_2-terminal amino acid sequences of polyhedrins of two baculoviruses from Orgyia pseudotsugata. Proc. Natl. Acad. Sci. USA 76 (10):4976-4980.

34. Rohrmann, G.F., D.J. Leisy, K-C. Chow, G.D. Pearson and G.S. Beaudreau (1982), Identification, cloning and R-loop mapping of the polyhedrin gene from the multicapsid nuclear polyhedrosis virus of Orgyia pseudotsugata. Virology 121:51-60.

35. Rohrmann, G.F., M.E. Martignoni and G.S. Beaudreau

(1982), DNA sequence homology between Autographa californica and Orgyia pseudotsugata nuclear polyhedrosis viruses. J. Gen. Virol. 62:137-143.

36. Rohrmann, G.F., M.N. Pearson, T.J. Bailey, R.R. Becker and G.S. Baudreau (1981), N-terminal polyhedrin sequences and occluded Baculoviurus evolution. J. Mol. Evol. 17:329-333.

37. Serebryani, S.B., T.L. Levitina, M.L. Kautsman, Y.L. Radavski, N.M. Gusak, M.N. Ovander, N.V. Sucharenko and E.A. Kozlov (1977), The primary structure of the polyhedral protein of nuclear polyhedrosis virus (NPV) of Bombyx mori. J. Invert. Pathol. 30:442-443.

38. Smith, G.E., M.J. Fraser and M.D. Summers (1983), Molecular engineering of the Autographa californica nuclear polyhedrosis virus genome: deletion mutations within the polyhedrin gene. J. of Virol. 46(2):584-593.

39. Smith, G.E. and M.D. Summers (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.

40. Smith, G.E. and M.D. Summers (1980), The

bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. Anal. Biochem. 109 :123-129.

41. Smith, G.E. and M.D. Summers (1981), Application of a novel radioimmunoassay to identify Baculovirus structural antigens that share interspecies antigenic determinants. J. Virol. 39:125-137.

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

43. Smith, G.E., M.D. Summers and M.J. Fraser (1983), Production of human beta interferon in insect cells infected with a Baculovirus expression vector. Mol. and Cell. Biol. 3(12):2156-2165.

44. Smith, G.E., J.M. Vlak and M.D. Summers (1982), In vitro translation of Autographa californica nuclear polyhedrosis virus early and late mRNAs. J.of Virology 44 (1):199-208.

45. Smith, G.E., J.M. Vlak and M.D. Summers (1983), Physical analysis of Autographa californica nuclear polyhedrosis virus transcripts for polyhedrin and 10,000-molecular-weight protein. J. Virol. 45:215-225.

46. Soeda, E., T. Maruyama, J.R. Arrand and B.E. Griffin (1980), Host-dependent evolution of three papova viruses. Nature 285:165-167.

47. Summers, M.D. and G.E. Smith (1978), Baculovirus structural polypeptides. Virology 84:390-402.

48. Summers, M.D., L.E. Volkman and C.-H. Hsieh (1978), Immunoperoxidase detection of Baculovirus antigens in insect cells. J. Gen. Virol. 40:545-557.

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

50. Thompson, C.G., D.W. Scott and B.E. Wickman (1981), Long-term persistance of the nuclear polyhedrosis virus of the Douglas-fir tussock moth, Orgyia pseudotsugata (Lepidoptera: Lymantriidae), in forest soil. Environ. Entom. 10:254-255.

51. Tweeten, K.A., L.A. Bulla, Jr. and R.A. Consigli (1981), Applied and molecular aspects of insect granulosis viruses. Microbiol. Reviews 45(3):379-408.

52. Upholt, W.B. (1977), Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nuc. Acids Res. 4(5):1257-1265.

53. van der Beek, C.P., J.D. Sayer-Riep and J.M. Vlak (1980), On the origin of the polyhedrin protein of Autographa californica nuclear polyhedrosis virus. Virology 100:326-333.

54. Vlak, J.M. and G.F. Rohrmann (in press), "The Nature of Polyhedrin", in Viral Insecticides for Biological Control , ed. by K. Maramorosch and K.E. Sherman. Academic Press, N.Y.