Characterization of the *Spodoptera exigua* baculovirus genome: structural and functional analysis of a 20 kb fragment

E.A. van Strien

proefschrift ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van de Landbouwuniversiteit Wageningen,
dr. C.M. Karssen,
in het openbaar te verdedigen
op vrijdag 5 september 1997
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen

The work presented in this thesis was carried out at the Department of Virology of the Wageningen Agricultural University, The Netherlands.

Strien, Elisabeth A. van.

Characterization of the Spodoptera exigua nucleopolyhedrovirus genome: Structural and functional analysis of a 20 kb segment - E.A. van Strien
Thesis Wageningen - with ref. - with summary in Dutch.
ISBN 90-5485-739-0
Subject headings: baculovirus

Characterization of the *Spodoptera exigua* baculovirus genome: structural and functional analysis of a 20 kb fragment

E.A. van Strien

promotor:	dr. R.W. Goldbach Hoogleraar in de Virologie
	dr. J.M. Vlak Persoonlijk Hoogleraar bij de vakgroep Virologie
co-promotor:	dr. D. Zuidema Universitair Docent vakgroep Virologie

1. De genvolgorde van baculovirussen is een bruikbaar kenmerk voor fylogenetische reconstructies.

   dit proefschrift

2. Het bezit van een gen voor ubiquitine is karakteristiek voor baculovirussen.

   dit proefschrift

3. De bewering dat herpesvirussen en baculovirussen een gemeenschappelijke voorouder hebben is niet bewijsbaar.

   M. Kool. 1994. Proefschrift, Landbouwuniversiteit Wageningen.
   A.J. Davison. 1992. Channel catfish virus: a new type of herpesvirus. Virology 186, 9-14.

4. De baculovirale genen *ie*2 en *ie*3 zijn ontstaan door duplicatie.

   R. Krappa & D. Knebel-Mörsdorf. 1991. Identification of the very early transcribed baculovirus gene PE-38. *Journal of Virology* 65, 805-812.
   D.A. Theilmann & S. Stewart. 1992. Molecular analysis of the trans-activating IE-2 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 187, 84-96.
   X. Wu *et al.* 1993. Alternative transcriptional initiation as a novel mechanism for regulating expression of a baculovirus *trans* activator. *Journal of Virology* 67, 5833-5842.

5. Baculovirussen zijn niet geschikt voor humane gen-therapie.

   D. Murges *et al.* 1997. Baculovirus transactivator IE1 is functional in mammalian cells. Journal of General Virology 78, 1507-1510.
   F.M. Boyce & N.L.R. Buchner. 1996. Baculovirus-mediated gene transfer into mammalian cells. Proceedings of the National Academy of Sciences USA 93, 2348-2352.
   C. Hofmann *et al.* 1995. Efficient gene transfer into human hepatocytes by baculovirus vectors. Proceedings of the National Academy of Sciences USA 92, 10099-10103.

6. Choi en Guarino concluderen ten onrechte dat IE1 geen structureel onderdeel van AcMNPV virions vormt.

   J. Choi & L.A. Guarino. 1995. Expression of the IE1 transactivator of *Autographa californica* nuclear polyhedrosis virus during viral infection. Virology 209, 99-107.

7. Goudsmit is te positief over het gebruik van een niet-pathogene HIV-stam voor de vaccinatie ter voorkoming van AIDS.

   J. Goudsmit. 1997. Vrijend virus, over aard en oorsprong van het aidsvirus. Uitg. Contact, Amsterdam.

8. Transgene planten met het *Bacillus thuringiensis* δ-endotoxine kunnen wel eens tegen meer plaaginsecten bestand zijn dan aanvankelijk werd beoogd.

   F.S. Walters & L.H. English. 1995. Toxicity of *Bacillus thuringiensis* δ-endotoxine toward the potato aphid in an artificial diet bioassay. Entomologia Experimentalis et Applicata 77, 211-216.

9. De conclusie dat een recombinant baculovirus met een antisense humaan *myc* construct werkzaam is dankzij een antisense mechanisme wordt niet ondersteund door de gepresenteerde resultaten.

S.Y. Lee *et al.* 1997. Insecticidal acitivity of a recombinant baculovirus containing an antisense c-*myc* fragment. Journal of General Virology 78, 273-281.

10. De fabrikanten van laboratorium-apparatuur beschouwen de aan- en uitknop kennelijk niet als noodzakelijk, gezien de doorgaans onbereikbare plaats van deze knop.

11. Het ontbreken van de Ascovirussen in de laatste officiële taxonomische indeling van virussen is symptomatisch voor het gebrek aan interesse in de virologie van invertebraten.

F.A. Murphy *et al.* (editors). 1995. Virus taxonomy. *Sixth Report of the International Committee on Taxonomy of Viruses*, pp. 586.

12. Het verdient aanbeveling om het effect van paracetamol-inname kort na besmetting met HIV, op verloop van infectie en op ontstaan van mutaties in het virale genoom, nader te onderzoeken.

13. Het ontbreken van BTW-heffing op vliegreizen suggereert dat de overheid vliegen nog boven de eerste levensbehoeften stelt.

14. Milieu-organisaties zouden zich moeten realiseren dat transgene gewassen voorzien van resistenties tegen ziekten en plagen een realistisch perspectief bieden op het terugdringen van het pesticidengebruik.

15. Het non-vaccinatiebeleid voor klassieke varkenspest is niet Normaal en een duidelijke misrekening.

16. De observatie dat in veel stedelijke gebieden de soortenrijkdom groter is dan in vele delen van het landelijk gebied zou voor natuurliefhebbers aanleiding moeten zijn verstedelijking van het Groene Hart te propageren.

Natuurverkenning '97. Uitg. Samsom H.D. Tjeenk Willink BV, Alphen aan de Rijn.

17. Een proefschrift is niet eetbaar.

Stellingen behorend bij het proefschrift:

Characterization of the *Spodoptera exigua* nucleopolyhedrovirus genome:
structural and functional analyis of a 20 kb fragment.

Wageningen, 5 september 1997 E.A. van Strien

VOORWOORD

In tegenstelling tot hetgeen die enkele naam op de kaft van een proefschrift suggereert, is dit boekje allerminst het resultaat van de inspanningen van één persoon. Het voorwoord van het proefschrift lijkt echter de enige plaats te zijn om alle mensen te bedanken die op enigerlei wijze hebben bijgedragen aan de totstandkoming van het geheel. Welnu, hier komt de lijst.

Als eerste wil ik Just Vlak noemen, en in niet minder mate, Douwe Zuidema. Niet alleen voor de concrete ondersteuning voorafgaand (omscholing van een verdwaalde entomoloog) en gedurende het onderzoek en tijdens het schrijven van het proefschrift, maar ook voor het vertrouwen en geduld dat zij hebben betoond. Rob Goldbach wil ik bedanken voor de kritische suggesties en opmerkingen die het proefschrift ten goede zijn gekomen.

Onmisbare ondersteuning was afkomstig van Els Klinge-Roode, Magda Usmany, Ine Derksen en Wout Rozeboom. Op deze plaats wil ik jullie bedanken voor de hoge kwaliteit en grote hoeveelheden primer extensies, insectencellen, rupsen, medium en chemicaliën. Dit alles binnen en soms ook buiten kantooruren.

Een promotie onderzoek op de vakgroep Virologie wordt ook gekenmerkt door de goede sfeer in en buiten het lab. In eerste instantie, in het oude gebouw, dankzij het gezelschap van Hans Flipsen, Marcel Kool, Ruud Mans, John Martens, Monique van Oers en Peter Roelvink. In het nieuwe gebouw kregen de wekelijkse "werkbesprekingen" met Jacco Heldens en Rene Broer een nieuwe impuls door de oprichting van de RBBS (Royal Baculovirus Beer Society) en het Se-promotie team. Ook dank zij Felix Bianchi, Lilian Colbers, Mariëlle van Hulten en Stan de Kort (met nuttige boekentips) en de schier eindeloos lijkende stromen studenten. Tevens wil ik mijn (ex)studenten Jeroen Bakker, Han Gerrits, Doeke Hoekstra, Fiona de Kaper, Jan-Willem de Kraker, Jessica Tersteege en in het bijzonder Bas Jansen bedanken voor hun bijdrage aan dit proefschrift.

De goede sfeer op Virologie is ook zeker een verdienste van de mensen van de EM afdeling en van de onderzoekers aan dat ándere insectenvirus: Joop, Erwin, Danielle, Marjolein, Richard, Jan, Cor, Frank, Marcel P., Marc, Fennet en Ineke.

Furthermore I would like to thank the foreign visitors of the baculovirus group for their stimulating interest and major contributions to a good atmosphere: Basil Arif, Primitivo Caballero, Johannes Jehle, Peter Krell, David Theilmann and Zhi-Hong Hu (Rose).

Buiten werktijd kwamen de broodnodige afleiding en morele ondersteuning van familie en vrienden: Marion, Wim, San, Theo, Roland en Anita, mijn moeder en Dick.

En tegen Erik wil ik hier alleen maar willen zeggen: het lijkt onderwaardering, zo'n vermelding als allerlaatste in het voorwoord, maar bedenk dat in de wetenschappelijke subcultuur van het namenrijtje hier de ere-positie zetelt.

Bep

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION AND SCOPE OF THE INVESTIGATION

Baculoviruses provoke a devastating disease in insects. A few days after infection the afflicted larvae become moribund and subsequently die and liquefy. The spectacular symptoms which are induced by these viruses were already noticed two millennia ago in China in the silkworm culture (Benz, 1986). Only much later the notion arose that baculoviruses could be exploited constructively as bio-insecticides to control insect pests in agriculture (Steinhaus, 1956). The ever more eminent drawbacks of chemical pest control (Carson, 1962) further fuelled the interest for and subsequent research of baculoviruses.

Nowadays, baculoviruses are used worldwide as microbial insecticides (Cunningham, 1995) and, due to some characteristic features of their biology, as expression vectors for the production of recombinant proteins of pharmaceutical interest. The latter application is a successful result of genetic engineering of the baculovirus genome. With this technology the properties of baculoviruses as biocontrol agents may be improved. Both types of application, *i.e.* engineering for improved insecticidal activity and foreign gene expression, require detailed knowledge of the genetic organization and expression of the baculovirus genome. Whereas a wealth of information was available for *Autographa californica* multinucleocapsid polyhedrosis virus (AcMNPV), the baculovirus type species, genetic information of other baculoviruses, such as *Spodoptera exigua* MNPV (SeMNPV), the subject of this thesis, was virtually absent at the start of this research.

The research described in this thesis was aimed at the determination of the molecular genetic properties of SeMNPV. Prior to the presentation of the experimental research, this introducing chapter will provide some background information concerning the biology and molecular genetics of baculoviruses in general and SeMNPV in particular.

BACULOVIRUS TAXONOMY, VIRUS MORPHOLOGY AND DNA STRUCTURE

Baculoviruses (family Baculoviridae) constitute a group of viruses of which the host is found exclusively among arthropods, distributed over seven different insect orders and one crustacean order. Most of the baculoviruses described are found in the order Lepidoptera (Murphy et al., 1995). The most apparent characteristic of this group of viruses is the production of large (0.1-10 µm) proteinaceous capsules, called occlusion bodies (OBs). These OBs contain rod-shaped virions, hence the name baculoviruses (baculum = rod), and they protect the virions from proteolytic attack during synthesis in the insect and from environmental decay.

The family Baculoviridae is divided into two genera, the genus Nucleopolyhedrovirus (NPV) and the genus Granulovirus (GV) (Murphy *et al.*, 1995), based on the type of occlusion body. In GVs a single virion is found in the occlusion body (granulum), which mainly consists of a protein named granulin. In NPVs several virions are found in one occlusion body (polyhedron), which mainly consists of the polyhedrin protein (Fig. 1.1, 1.2). Polyhedrin and granulin are structurally and functionally related proteins. Within the NPVs two morphotypes are recognized: single (SNPV) and multiple enveloped nucleocapsid (or in short: multicapsid) NPVs (MNPV). In the latter morphotype up to nine nucleocapsids are assembled in a single

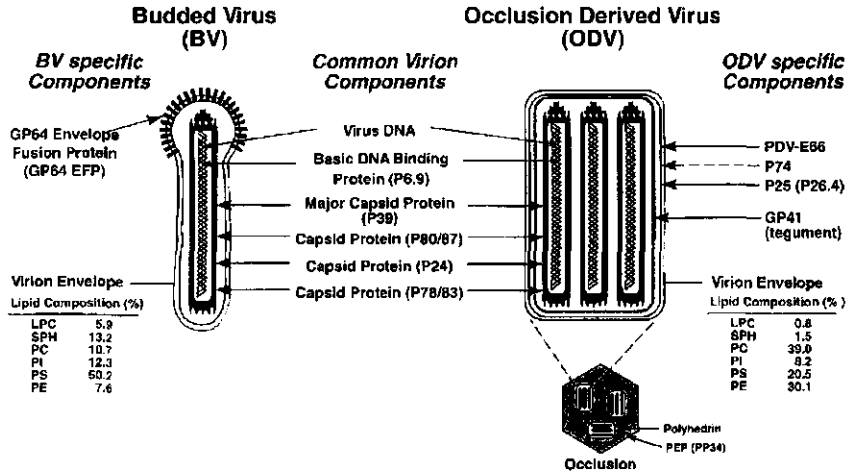


Figure 1.1. Structural composition of the two baculovirus phenotypes, the budded virion, BV, and the occlusion derived virion, ODV (from Blissard, 1996). The ODV structure represents the MNVP subgroup. Proteins common to both virion types are indicated in the middle of the figure. Proteins specific to either BV or ODV are indicated on the left and right, respectively. The polar nature of the baculovirus capsid is indicated in the diagram with the claw-like structure at the bottom and the ring-like nipple at the top of the capsid. The possible location of p74 is indicated by a dashed line. Lipid compostions of the BV and ODV envelopes derived from AcMNPV infected Sf9 cells (Braunagel & Summers, 1994) are indicated (LPC, lysophosphatidylcholine; SPH, sphingomyelin; PC, phosphatidylcholine; PI, phsophatidylinositol; PS, phosphatidylserine; PE, phosphatidylethonalamine).

virion envelop, before several of these packages are occluded into a polyhedron. The taxonomic significance of the NPV morphotypes is disputed and the latest taxonomic revision of the Baculoviridae does not treat the SNPVs and MNPVs as distinct groups any longer (Murphy *et al.*, 1995). The maturation of the occlusion bodies of NPVs takes places in the nucleus and that of the granuloviruses in the nuclear-cytoplasmic continuous mass of the infected cell.

The latest taxonomic overview lists 633 baculovirus species: 15 assigned and 483 putative NPV species and 4 assigned and 131 putative GV species (Murphy *et al.*, 1995). Although the cumulative host range of baculoviruses is very broad, each baculovirus can only infect the arthropod animals from a single or a few related species. SeMNPV only infects the larvae of a single insect species, whereas AcMNPV has a relatively broad host range, involving 73 species from 13 Lepidopteran families (Possee *et al.*, 1986, Hu & Vlak, 1997). An overview of baculoviruses and their hosts is available at the world wide web site of the Illinois Natural History Survey in the Ecological Database of the World's Insect Pathogens (EDWIP).

Both GVs and NPVs are often named after the host from which they were first isolated. The host range of a baculovirus often constitutes an important determinant for the ultimate designation of the taxonomic species status. The use of molecular data have been exploited only sparsely in baculovirus taxonomy, although the general awareness is increasing that at least restriction enzyme profile analysis is required to establish identity (Vlak, 1982, Cunningham, 1995).

The basic biology of NPVs is much better investigated than that of GVs. The replication of AcMNPV, the NPV type species and the best studied baculovirus, is introduced below in more detail.

The baculovirus genome consists of a circular, double stranded DNA molecule, ranging

in size between 80 and 160 kilobase pairs (kbp). On the AcMNPV genome, with a size of 133,894 base pairs, more than 150 potential open reading frames (ORFs) were identified (Ayres *et al.*, 1994). These ORFs are tightly packed and sometimes even overlap (Fig. 1.3). The baculovirus genome is further characterized by the presence of homologous repeat (*hr*) regions, which are involved in the enhancement of early gene expression and in DNA replication (see below). A growing number of baculovirus genes and genetic elements has been investigated (for recent reviews see Kool & Vlak, 1993, Blissard, 1996, Williams & Faulkner, 1996). Recently, the complete sequence of *Bombyx mori* (BmNPV) NPV (S. Maeda, accession number L33180) and *Orgyia pseudotsugata* (Op) MNPV (Ahrens *et al.*, 1997) have become available. However, the information on baculoviruses less closely related to AcMNPV is rather limited.

INFECTION CYCLE

Multiplication in vivo

A successful baculovirus infection (Fig. 1.2) starts when an insect larva of a susceptible species ingests a baculovirus in the form of OBs, usually concomitant with food intake. The midgut lumen of a lepidopteran larva constitutes a very alkaline environment (up to pH 12) in which the OBs dissolve and the so-called occlusion-derived virions (ODVs) are released. These virions fuse with the membrane of microvilli of the midgut epithelial cells whereafter they are transported to the nucleus, initiating the first of many infection cycles in the cells of the insect body.

Baculoviruses have a biphasic replication cycle, in which two genetically indentical but phenotypically distinct types of virions are formed (Fig. 1.1, 1.2). Newly formed virions are initially released by budding through the plasma membrane of the infected cell. They are designated by a variety of names: budded virions (BVs), extracellular virions (ECVs) or non-occluded virions (NOVs). The insect tracheal system and haemolymph have a major transport function of the BVs to other tissues of the insect body (Engelhard *et al.*, 1994, Flipsen *et al.*, 1995a).

BVs differ in several aspects from the ODVs which are formed later in infection (Fig. 1.1, 1.2). BVs are responsible for the systemic infection, throughout the body of the insect; ODVs, embedded within OBs, facilitate viral spread from one individual insect to an other, through the population and over time. The differences between BVs and ODVs encompass the timing of synthesis and the formation process itself, the protein and envelope composition up to the infection mechanism of the virions (Blissard & Rohrmann, 1990, Braunagel & Summers, 1994, Blissard, 1996, Rohrmann, 1992). BVs enter the cell by endocytosis, followed by membrane fusion of the viral envelope and the endosome membrane. The fusion process is mediated by a virus-encoded essential glycoprotein, gp64, which is exclusively found in BVs (Blissard, 1996, Monsma *et al.*, 1996).

The ODVs, the virions formed later in infection in the infected cell, are not released by budding but they acquire an envelope in the nucleus, followed by occlusion in polyhedra. Finally, the infected cell ruptures and the lysis of both the cellular and nuclear membrane allows the release of the newly formed, mature polyhedra. The polyhedra are surrounded by a carbohydrated polyhedral envelope (Zuidema *et al.*, 1989, Gombart *et al.*, 1989). The baculovirus p10 protein has been implicated in the formation of this envelope and in the rupture

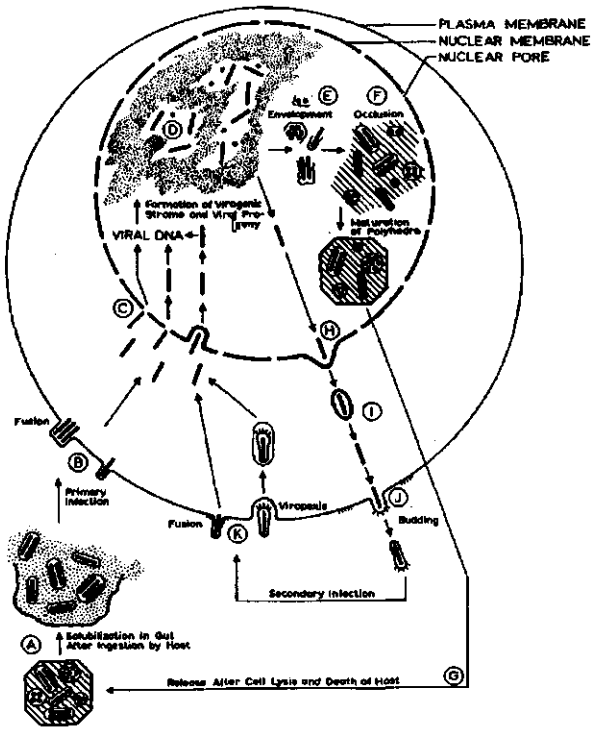


Figure 1.2. Schematic representation of the baculovirus infection cycle (from van der Beek, 1980). Ingested polyhedra are solubilized in the midgut and virions are released (A). The envelopes of the virions fuse with the plasma membrane of the insect cell (B). After traversing the cytoplasm virions enter the nucleus at a nuclear pore, uncoat and the viral DNA enters the nucleus (C). Progeny viral nucleocapsids are synthesized in the virogenic stroma (D). Following envelopment in the nucleus (E), progeny nucleocapsids are initially released by budding (H, I, J). Budded virions infect adjacent insect cells by endocytosis (K). Nucleocapsids produced in later stages of in infection become occluded in polyhedral protein (F). Finally, the occlusion bodies are released by lysis of the infected cell.

of the infected cell (van Oers, 1994a).

A baculovirus infection may be restricted to certain tissues of the insect, depending on the baculovirus and its host. In lepidopteran larvae the fat body is usually the major infected tissue and is therefore an important source of new polyhedra.

It is not until the final stages of infection that infected larvae become sluggish and cease feeding. Death occurs a week after infection under optimal conditions (temperature). The concomitant liquefaction of the insect body remnants by the viral proteins cathepsin (Ohkawa *et al.*, 1994) and chitinase (Hawtin *et al.*, 1995) facilitates the spread of the OBs in the environment.

Multiplication in cell culture (in vitro)

The processes that take place at the cellular level during baculovirus infection have been investigated in detail using cultured insect cells (Granados & Lawler, 1981, Granados & Federici, 1986, Blissard, 1996, Williams & Faulkner, 1996). Insect cells are relatively easy to cultivate and many are susceptible to baculovirus infection. Notably, studies of AcMNPV in *Spodoptera frugiperda* and other insect cells have provided fundamental knowledge of baculovirus DNA

replication, gene function, gene expression and the regulation thereof. Furthermore, the cell-virus system offers the possibility to investigate the molecular basis of host specificity (Friesen & Miller, 1986, Blissard & Rohrmann, 1990, Blissard, 1996, Williams & Faulkner, 1996). From a practical point of view the baculovirus-insect cell system has proven useful in generating, producing and maintaining engineered baculoviruses for the expression of foreign genes or for biological control purposes.

Gene expression and DNA replication

After the BV has entered the insect cell by absorptive endocytosis, the nucleocapsid is transported to the nucleus where the viral DNA is uncoated and gene expression and viral DNA replication take place. Baculovirus gene expression occurs in a cascade like fashion and is divided into several phases (Fig. 1.4). Whereas late genes are not expressed before the onset of viral DNA replication, the expression of early genes starts before this event (Friesen & Miller, 1986, Blissard & Rohrmann, 1990, Blissard, 1996). Of these, the immediate early (*ie*) genes do not require the presence of viral gene products for their transcription, but utilize the host RNA polymerase II complex. They are often involved in the transactivation of delayed early and late gene expression and some are essential, for viral DNA replication and for late gene expression (Kool *et al.*, 1995, 1994a, Friesen & Miller, 1986, Blissard & Rohrmann, 1990, Lu & Carstens, 1993, Blissard 1996). By definition, late gene expression starts after the onset of DNA replication. In AcMNPV-infected insect cells maintained at 27° C this event starts at about 6 h p.i., but this may be later in other baculovirus - cell systems. Host protein synthesis shut off is initiated around the same time. The expression of the polyhedrin and *p10*, designated very late genes, continues during the process of OB formation and ultimately ends at cell lysis.

Baculovirus gene expression is transcriptionally regulated. Early genes are transcribed by host RNA polymerase II (Fuchs, 1983, Glockner *et al.*, 1992, Huh & Weaver, 1990). The promoter sequences of some early baculovirus genes resemble insect RNA polymerase II promoters, notably the baculovirus gene promoter/initiator sequence CAGT preceded by the canonical TATA box (Blissard & Rohrmann, 1990, Blissard, 1996). In addition to these motifs other eukaryotic sequence motifs, GATA and CACGTG, were found in baculovirus genes to bind host factors and serve as components of a core promoter (Krappa *et al.*, 1992, Kogan & Blissard, 1994, Kogan *et al.*, 1995). Recently identified *Drosophila* transcription factors Myc and Mac recognized the same CACGTG sequence as their vertebrate homologs (Gallant *et al.*, 1996). It is thus likely that insect cell homologs perform the same function in transcription regulation of viral promoters with these sequences. As yet, the factors determining the transcription initiation of early baculovirus genes are still unclear. For instance, transcription initiated from different locations in the AcMNPV and BmNPV DNA polymerase genes, which was remarkable considering the high similarity of the promoter regions of these two viruses (Chaeychomsri *et al.*, 1995).

Transactivation by *hr*-sequences, which function as enhancers of transcription, plays a major role in transcription regulation (see below). Only for one baculovirus gene, AcMNPV *ie0*, the primary transcript is known to be spliced (Chisholm & Henner, 1988, Kovacs *et al.*, 1991). Generally, baculovirus genes do not contain introns.

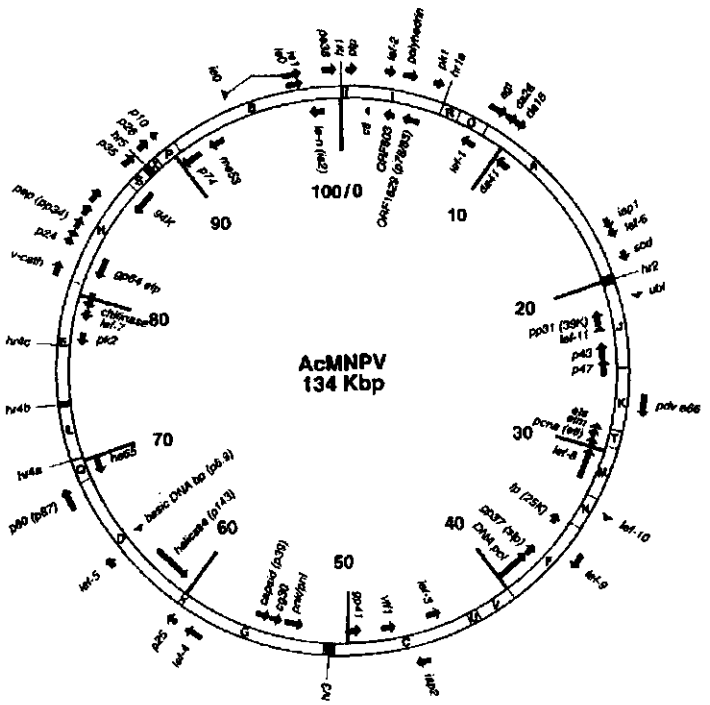


Figure 1.3. Representation of the circular AcMNPV genome with *Eco*RI fragments (A to X) . Units in map units (percentage of genome). Genes with assigned functions and *hr* elements are indicated (Blissard, 1996).

Late and very late genes utilize the canonical baculovirus late promoter annex initiator sequence (A/G/T)TAAG (Blissard & Rohrmann, 1990). Only limited sequence information beyond a few nucleotides surrounding the TAAG sequence is required for late promoter activity (Morris & Miller, 1994, Ooi *et al.*, 1989, Rankin *et al.*, 1988). The baculovirus late genes are transcribed by an α-aminitin resistant RNA-polymerase. The origin of this polymerase, either host-modified or, more likely, virus encoded, is still unclear (Grula *et al.*, 1981, Huh & Weaver, 1990, Glockner *et al.*, 1993, Beniya *et al.*, 1996, Xu *et al.*, 1995, Yang *et al.*, 1991), but host proteins may be involved in baculovirus late gene transcription as well (Jain & Hasnain, 1996).

Eighteen viral genes have been identified as being required for late gene expression (Todd *et al.*, 1995, Lu & Miller, 1995a). The predicted product of one of these genes, *lef-8*, showed homology to RNA polymerases (Passarelli *et al.*, 1994). In addition, a gene involved in very late gene expression, *vlf-1*, was identified (McLachlin & Miller, 1994). A subset of six (*ie-1, lef-1, lef-2, lef-3, helicase* and *DNA polymerase*) of these eighteen *lef* genes was essential for DNA replication (Kool *et al.*, 1994a). Baculovirus DNA replication and late gene expression seem to be intimately linked, but the nature of this link remains to be elucidated.

Baculovirus DNA contains so-called *hr* regions. These regions contain one or more repeated sequences and occur dispersed over the circular genome (Fig. 1.3). In AcMNPV and OpMNPV the *hr* regions function as strong enhancers of (early) gene transcription (Guarino *et al.*, 1986, Guarino & Summers, 1986, Choi & Guarino, 1995, Guarino & Dong, 1994, Rodems

& Friesen, 1995, Theilmann & Stewart, 1993). The enhancement takes place in close unison with the expression of immediate early 1 (*ie1*) gene (Choi & Guarino, 1995).

Hr sequences were also shown to function as origin of replication (*ori*) in an *in vitro* replication assay (Pearson *et al.*, 1992, 1993, Leisy & Rohrmann, 1993, Kool *et al.*, 1993). A second type of *ori*, with structural homology to eukaryotic origins, was later identified in both AcMNPV and OpMNPV (Kool *et al.*, 1994b, Ahrens et al., 1995).

SPECIFICITY AND VIRULENCE

Most baculoviruses are infectious for only a few species and the majority of insects is thus not susceptible to a given virus. Underlying mechanisms of specificity may be found at the cellular level, as observed in the sloughing off of infected midgut epithelial cells in the midgut lumen (Flipsen *et al.*, 1993, 1995a, Engelhard & Volkman, 1995) and in the encapsulation of infected tissue by haemocytes and subsequent clearing of infected cells (Washburn *et al.*, 1995, 1996). Insects do not possess an immune system against micro-organisms and viruses similar to the humoral response of vertebrates. Cellular immunity is an important defence system in arthropods. A potential but not very well investigated factor determining specificity may be differences in virus entry. However, beyond that stage the cellular defence reaction in the form of apoptosis seems to perform a major role and this reaction can already be observed in the midgut epithelial cells (Palli *et al.*, 1996). Apoptosis is a process of active cell death, initiated by cells as part of their programmed development or in response to infections or other triggers.

Most research on host specificity focused on the elucidation of the mechanisms at the molecular level. Several baculovirus genes (Clem *et al.*, 1991, Crook *et al.*, 1993, Birnbaum *et al.*, 1994, Clem & Miller, 1994) are involved in the prevention and counteraction of apoptosis. A recently elucidated apoptosis preventing mechanism is displayed by the AcMNPV p35, involving specific inhibition of caspases, a type of induced cellular protease involved in the apoptotic response (Bertin *et al.*, 1996, Bump *et al.*, 1995, Xue *et al.*, 1995). The *p35* gene is not required in all AcMNPV hosts or cell lines (Hershberger *et al.*, 1992, Clem & Miller, 1993, Clem *et al.*, 1991, *ibid.*, 1994).

Other viral gene products that were shown to be involved in host specificity are AcMNPV helicase (Croizier *et al.*, 1994, Maeda *et al.*, 1993) and lef-*7* (Chen & Thiem, 1997). The *hrf-1* gene from *Lymantria dispar* MNPV, but not the homologous gene from OpMNPV, extended the host range of AcMNPV (Thiem *et al.*, 1996). Different cell lines may impose different constraints on the set of genes essential for viral DNA replication: the AcMNPV *hcf-1* gene was required in TN-368 cells, but not in Sf21 cells in a plasmid replication assay (Lu & Miller, 1995b).

A viral gene that has been linked with virulence is the ecdysteroid UDP-glycosyltransferase (*egt*) gene (O'Reilly, 1995). This gene is required for the maintenance of functional Malpighian tubes after virus infection (Flipsen *et al.*, 1995b). Deletion of the *egt* gene resulted in a reduction of the time to kill the insect after infection and in a reduction of feeding damage.

Several structural proteins where shown to be essential in the baculovirus life cycle. The AcMNPV *p74* gene was required for the infectivity of the virus for insect larvae, but not for cultured cells (Kuzio, 1989).

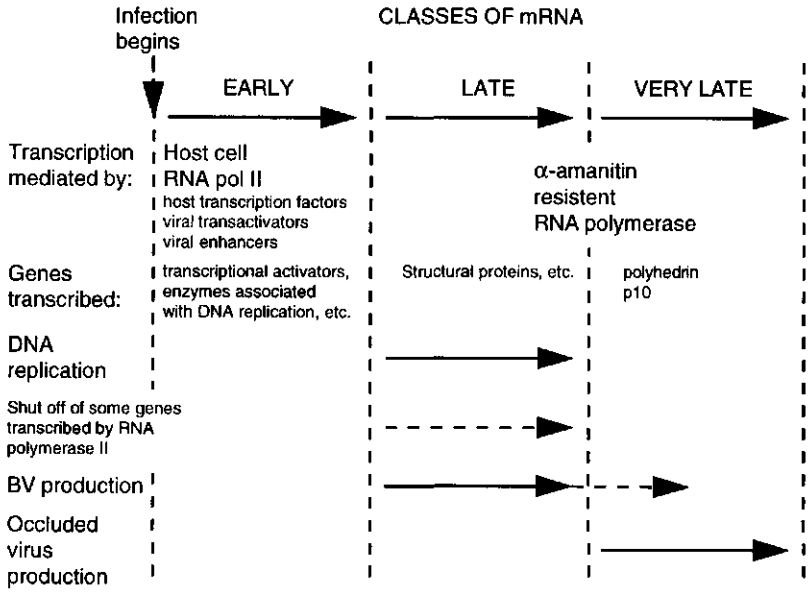


Figure 1.4. Schematic representation of the definitions and interrelationship between early, late and very late gene transcription and DNA replication during baculovirus infection (from Rohrmann, 1985).

GENETIC ENGINEERING OF BACULOVIRUSES: DEVELOPMENT OF VIRAL GENE VECTORS AND IMPROVEMENT AS BIO-INSECTICIDES.

The major occlusion body protein, polyhedrin or granulin, and the p10 protein are produced in large amounts in the baculovirus infected cell. The baculovirus occlusion body protein protects the virions from environmental decay. The exact function of the baculovirus p10 protein remains somewhat enigmatic (van Oers, 1994). It is a constituent of fibrillar structures found in both nucleus and cytoplasm of the infected cell (Russell *et al.*, 1991, van der Wilk *et al.*, 1987) and it is responsible for lysis of the nuclear membrane in the final stage of infection (van Oers, 1994). Furthermore, p10 is involved in polyhedron envelope morphogenesis (Russell *et al.*, 1991, Gross *et* al., 1994, Lee *et al.*, 1996). Neither polyhedrin nor p10 are essential for the infectivity of the virus in cultured insect cells. This finding has been exploited extensively in the construction of recombinant viruses, in which the very strong promoters of these genes drive the expression of a wide variety of foreign genes (for reviews see King & Possee, 1992, O'Reilly *et al.*, 1992, Luckow, 1991). Recombinants using promoters from other temporal classes of genes, for instance from early genes (Jarvis *et al.*, 1996b), are also under investigation (Blissard, 1996).

Baculoviruses have positive attributes as biological control agent such as a restricted host range, environmental safety and safety for vertebrates including humans, and compatibility with other agricultural practices. However, disadvantages are the relatively high registration costs, restricted market size due to their specificity, slow speed of action, low virulence for later larval stages and limited survival of the virus, both in the field (persistence) and as formulated product (shelf life). Recombinant baculoviruses, designed to alleviate some of the drawbacks mentioned above, have met varying success (for recent reviews see Vlak, 1993, Cunningham,

1995, Murhammer, 1996, Hu & Vlak, 1997). Successful engineered viruses with increased speed of action were obtained by deletion of the viral *egt* gene (O'Reilly & Miller, 1991) and by he introduction of a scorpion toxin gene (Stewart *et al.*, 1991, Maeda *et al.*, 1991), a mite toxin gene (Tomalski & Miller, 1991) or a modified JHE (juvenile hormone esterase) gene (Bonning *et al.*, 1995). Recombinant viruses expressing these genes under the control of the *ie1* promoter did not kill faster than the very late *p10* promoter recombinants (Jarvis *et al.*, 1996a). Failures and successes of baculovirus genetic engineering stress the importance of basic knowledge on baculovirus replication and gene expression as a prerequisite for rational design.

SeMNPV

The *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV), the subject of this thesis, is a baculovirus that has been officially registered in the Netherlands for the control of the beet army worm *Spodoptera exigua* on ornamental crops in glasshouses (Smits & Vlak, 1994). SeMNPV is also commercially used on the Asian and North-American continent in a variety of crops. During the 1970's an insecticide resistant population of the beet army worm, called "koppensneller" (head hunter) or "floridamot" by Dutch growers, was accidently introduced in the Netherlands. In Dutch glasshouses this subtropical, polyphagous insect rapidly developed into a serious pest of vegetable and ornamental crops like tomato and chrysanthum. In the research that was subsequently initiated (Smits, 1987) the SeMNPV stood out as a biological control agent with high potential (Smits *et al.*, 1988). The virus was characterized by establishing its identity. Genetic information of the SeMNPV genome was virtually absent at the start of the research presented in this thesis.

SCOPE OF THE INVESTIGATION

The baculovirus SeMNPV is biologically characterized as a monospecific, highly virulent baculovirus of the beet army worm *S. exigua*. These properties contrast with those of AcMNPV, which has a broad host range and is at least ten times less virulent for this pest insect than SeMNPV. The supposition that the biological differences between AcMNPV and SeMNPV must have a genetic basis formed the starting point of this thesis.

The nucleotide sequence of the AcMNPV genome is completely known (Ayres *et al.*, 1994) and a few candidate genes affecting host range and virulence have been identified. However, the available molecular data for SeMNPV were limited to restriction enzyme profiles of the viral DNA (Smith & Summers, 1982, Gelernter *et al.*, 1986; Smits, Ph.D. thesis, 1987). To ultimately achieve understanding of the mechanisms governing host range and virulence, a molecular and genetic analysis of the SeMNPV genome was initiated, with special emphasis on the genome organization and expression strategies of viral genes.

The first step was the identification and characterization of the SeMNPV polyhedrin gene. The location of this gene allows the determination of the zero point of the circular viral genome (Vlak & Smith, 1982). Analysis of polyhedrin gene transcription provided basic information on future exploitation of its promoter (Chapter 2). To set the coordinates on the circular SeMNPV genome the other major late gene, *p10*, was identified and transcriptionally characterized (Chapter 3).

From the observed distance between the polyhedrin and *p10* genes in the genome of AcMNPV (19 kb) and SeMNPV (11 kb) it was clear that the genetic organization of SeMNPV in this region would be different from AcMNPV and OpMNPV. The entire SeMNPV genome region (20 kbp) ranging from upstream of *p10* to downstream of polyhedrin was therefore sequenced and analyzed (Chapters 4, 5, 6 and 7). This investigation allowed the identification of a gene for ubiquitin, *v-ubi*, upstream of *p10* (Chapter 4). The highly conserved *v-ubi* gene occupied a different position on the genetic map of SeMNPV, as compared to its position in AcMNPV.

In AcMNPV the region between the two major late genes polyhedrin and *p10* is occupied by at least four immediate early genes, including *ie1*. The latter gene has been implicated in host specificity and encodes a multifunctional protein, involved in transactivation of viral gene expression as well as in DNA replication. The SeMNPV homolog of AcMNPV *ie1*, located in between polyhedrin and *p10*, was characterized molecularly with special reference to the potential occurrence of splicing (Chapter 5).

Chapter 6 describes the identification and characterization of a SeMNPV gene encoding a homolog of the large subunit of ribonucleotide reductase, the *rr1* gene. This gene is a newly-identified gene among the Baculoviridae and is absent in AcMNPV. The possible role of this gene in baculovirus virulence will be discussed (Chapter 6).

The genetic organization of the region spanning the polyhedrin-*p10* genomic region was further investigated and compared to the corresponding regions in the AcMNPV and OpMNPV genomes (Chapter 7). This region was also analyzed for the presence of *hr* sequences, genetic elements reported to function as enhancer of viral gene transcription and origin of baculovirus DNA replication. The high degree of homology among the polyhedrin genes was a remarkable exception to the pattern of divergence seen with other genes identified in this region. The suitability of the polyhedrin gene as a phylogenetic marker for baculovirus relationships is therefore evaluated. Finally, the results obtained from the molecular analysis, gene organization and gene expression studies are summarized and reflected against the present literature. The potential to use the gene organization in baculovirus genomes as a phylogenetic marker is discussed (Chapter 7).

CHAPTER 2

NUCLEOTIDE SEQUENCE AND TRANSCRIPTIONAL ANALYSIS OF THE POLYHEDRIN GENE OF *SPODOPTERA EXIGUA* NUCLEAR POLYHEDROSIS VIRUS.

SUMMARY

The nucleotide sequence of a 1.1 kbp fragment of the multiple nucleocapsid nuclear polyhedrosis virus (MNPV) of *Spodoptera exigua* (Se) containing the polyhedrin gene was determined. An open reading frame (ORF) of 738 nucleotides (nt) was detected. This ORF encoded a protein of 246 amino acids with a predicted M_r of 29 kDa. The nucleotide and amino acid sequences were compared with the sequences of eight other NPV polyhedrins. The SeMNPV polyhedrin protein was most closely related to *S. frugiperda* MNPV polyhedrin with differences in only five amino acids, and most distantly related to the *Lymantria dispar* MNPV polyhedrin. The size of the mRNA was approximately 1000 nt, as determined by Northern blot analysis. Using primer extension assays and S1 nuclease mapping the transcriptional start and stop sites of the polyhedrin mRNA were located. The 5' regulatory sequence appeared to be 44 nt in length with the mRNA start site predominantly at the first A of the TAAG consensus start sequence. Two degenerated polyadenylation signals were found immediately downstream of the translational stop signal. The transcriptional stop was located approximately 230 nt downstream from the translational stop signal, in an AT-rich sequence that appears to be common to all baculovirus polyhedrin genes. The SeMNPV polyhedrin mRNA does not appear to be polyadenylated.

This chapter has been published as: Strien, E.A. van, Zuidema, D., Goldbach, R.W., & Vlak, J.M. 1992. Nucleotide sequence and transcriptional analysis of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus. Journal of General Virology 73, 2813-2821.

INTRODUCTION

Spodoptera exigua multiple-nucleocapsid nuclear polyhedrosis virus (SeMNPV) is a member of the subfamily Eubaculovirinae (Family Baculoviridae) (Francki *et al.*, 1991). Baculoviruses are enveloped, dsDNA-containing viruses that cause acute disease in a wide range of insects. The rod-shaped virions are embedded in large protein capsules, called polyhedra, that are a characteristic phenotype of this group of viruses.

SeMNPV is specific for beet armyworm (*S. exigua*) and is being applied as a biological insecticide against this insect (Gelernter & Federici, 1986a; Smits *et al.*, 1988). *S. exigua* (Lepidoptera, Noctuidae) is a polyphagous pest insect of economically important crops in (sub)tropical areas of the northern hemisphere as well as in protected environments, such as glasshouses.

As for all baculoviruses, the speed of action of SeMNPV is relatively slow as it may take several days or even weeks before virus infection renders its effect on insects, e.g. in the form of reduced feeding (Smits *et al.*, 1988). Recently however, successful attempts have been reported on the improvement of the insecticidal activity of baculoviruses via the introduction by genetic engineering of toxin-encoding genes (Stewart *et al.* 1991; Tomalski & Miller, 1991; McCutchen *et al.*, 1991; Vlak, 1992). This advancement became possible from the detailed knowledge of the structure and function of the polyhedrin gene and the development of transfer systems to introduce foreign genes into baculoviruses (Luckow & Summers, 1988; Miller, 1988; Maeda, 1989). Most of these attempts have been undertaken with a baculovirus of *Autographa californica* (AcMNPV), this being the most thoroughly characterized baculovirus so far (Blissard & Rohrmann, 1990; Rohrmann, 1992). Other baculoviruses are not used as frequently because their molecular characterization lags far behind.

Polyhedrin is the most abundant protein of polyhedra and present in high amounts in infected cells. For many baculoviruses it has been the subject of detailed molecular and genetic studies. It is approximately 30 kDa in size and highly conserved among baculoviruses (Vlak & Rohrmann, 1985). Elucidation of the genome structure and expression of the SeMNPV polyhedrin gene is a starting point for the design and engineering of this virus to improve its insecticidal properties. In this paper the nucleotide sequence organization and transcriptional analysis of the polyhedrin gene of SeMNPV is reported. The nucleotide and amino acid sequences of SeMNPV polyhedrin are compared with those of polyhedrin genes of other MNPVs.

MATERIALS AND METHODS

Virus, insects and cells. The *S. exigua* MNPV field isolate (SeMNPV/US) (Gelernter & Federici, 1986b) was obtained from Dr B.A. Federici, Department of Entomology, University of California, Riverside, Ca., U.S.A. in the form of polyhedra. The virus was propagated in fourth instar larvae of *S. exigua* (Smits *et al.*, 1988).

Haemolymph from SeMNPV-infected insects was used as a source of extracellular virus (ECV) for infection of cultured *S. exigua* cells (Gelernter & Federici, 1986b). This cell line (UCR-SE-1), obtained from Dr B.A. Federici, was maintained in plastic tissue culture flasks in TNM-FH medium (Hink, 1970) supplemented with 10% foetal calf serum. Infectivity assays of the ECV

solutions were performed using the endpoint dilution method (Vlak, 1979) and the titres were expressed as $TCID_{50}$ units per ml.

SeMNPV DNA was obtained from alkali-liberated virions purified after alkaline treatment of polyhedra followed by sucrose gradient centrifugation (Caballero *et al.*, 1992). Alternatively, viral DNA was isolated from ECVs of a plaque-purified isolate of SeMNPV according to procedures described by Summers and Smith (1987).

Southern blot hybridization, cloning and sequencing. To locate polyhedrin gene-containing sequences, the SeMNPV DNA was digested with several restriction enzymes, separated in 0.7% agarose gels, transferred to Hybond N filters (Amersham) and hybridized with the [^{32}]P-labelled *Hin*dIII V fragment of *Autographa californica* MNPV DNA according to procedures described by Sambrook *et al.* (1989). This AcMNPV DNA fragment contained polyhedrin gene sequences (Smith *et al.*, 1983). SeMNPV DNA fragments hybridizing with the AcMNPV *Hin*dIII V probe were isolated from agarose gels by the freeze-squeeze method (Sambrook *et al.*, 1989) and (sub)cloned into the plasmids pTZ18R (Promega) and pUC. The SeMNPV DNA inserts were sequenced using T7 DNA polymerase (Promega) and [^{35}]S-dATP according to Sanger *et al.* (1977).

RNA isolation. For Northern blot hybridization, primer extension and S1 nuclease analysis, total RNA was isolated from infected *S. exigua* cells at 24 or 48 h post infection (p.i.). Cells were lysed in 4 M guanidine isothiocyanate, 25 mM sodium acetate, pH 6.0, and 0.83% 2-mercaptoethanol and stored frozen until centrifugation of the RNA through a CsCl cushion (5.7 M CsCl, 25 mM sodium acetate, pH 6.0) by centrifugation for 23 h at 32000 rpm (Beckmann ultracentrifuge). The pelleted RNA was dissolved in sterilized double-distelled water and ethanol-precipitated after addition of 0.1 volume of 1.5 M sodium acetate (pH 5.0). After washing the RNA pellet in 70% (v/v) ethanol the pellet was redissolved in sterilized water. The RNA concentration was determined by measuring the u.v. absorbance at 260 nm.

Northern blot hybridization. Total RNA was electrophoresed in a 1% (w/v) agarose gel in 10 mM phosphate buffer (5 mM NaH_2PO_4, 5 mM Na_2HPO_4, pH 6.5 to 7.0), with constant recirculation of the running buffer. Prior to electrophoresis the RNA was denatured in 50% (w/w) DMSO, 1 M glyoxal (deionized), 10 mM sodium phosphate buffer for 1 h at 50° C and a 0.2 volume of loading buffer (50% glycerol, 10 mM sodium phosphate, 0.4% bromophenol blue) was added.

The separated RNAs were transferred to GeneScreen Plus (Dupont) or Hybond N (Amersham) in 10 x SSC (1.5 M NaCl, 0.15 M Na-citrate, pH 7.0). The filter was prehybridized for 1 h at 48 °C in a volume of 0.2 ml hybridization buffer [6 x SSC, 1% SDS, 200 µg denatured herring sperm DNA, 1% Ficoll-400, 1% polyvinylpyrrolidone, 1% BSA (fraction V)] per cm^2 filter. The 5' end-labelled oligonucleotide used in the primer extension assay (see below) was used as a probe in the Northern blot hybridization and added to the prehybridization solution. After hybridization for 16 h at 48° C the filter was washed for 5 min in 4 x SSC, 1% SDS and then for 5 min in two changes 2 x SSC, 0.5% SDS at room temperature. The filter was autoradiographed for several days.

Primer extension. To identify the transcriptional start site of the SeMNPV polyhedrin gene, a 17-

mer oligonucleotide (5' TCGTACACGTAAGTGCG 3') complementary to the polyhedrin mRNA was synthesized and used in primer extension assays. The oligonucleotide was 5'-end labeled with [γ-^{32}P]dATP using T4 polynucleotide kinase (Gibco-BRL) in 50 mM Tris-HCl, pH 9.5, 10 mM $MgCl_2$, 5 mM DTT, 5% glycerol for 30° min at 37°C followed by heat denaturation of the kinase at 90° C for 10 min. The labeled oligonucleotide was purified on a 1 ml Sephadex G25 column and ethanol-precipitated after addition of a 0.1 volume of 1.5 M sodium acetate (pH 5.0). Fifteen ng labeled primer was added to 2 µg total RNA and this mixture was denatured for 10 min at room temperature in 25 mM methylmercury. After addition of diluted 2-mercaptoethanol, the reverse transcriptase reaction was carried out for 30 min at 37° C in a volume of 15 µl with 0.5 mM of each of the four dNTPs, 1 µl Moloney murine leukemia virus reverse transcriptase (Gibco-BRL) in the buffer supplied by the manufacturer. The reaction was stopped by addition of 5 µl of a solution containing 95% (v/v) formamide, 0.01% xyleen cyanol, 0.01% bromophenol blue. Five µl of this reaction mixture was analyzed in a 6% polyacrylamide sequencing gel. After drying the gel was subjected to autoradiography.

S1 nuclease analysis. S1 nuclease protection experiments were performed to identify the 3' end of the polyhedrin mRNA. A 3' end labeled dsDNA probe was prepared by filling in a suitable restriction enzyme site with a 5' overhang using Klenow large fragment of DNA polymerase I (Gibco-BRL) and [α-^{32}P]dATP. The probe was purified on a Sephadex G50 1 ml column, precipitated and dissolved in sterile water.

Five µg total RNA and 1000 c.p.m. labelled DNA were suspended in 80% formamide, 40 mM PIPES pH 6 to 7, 1 mM EDTA, 0.4 M NaCl and denatured at 85° C for 15 min. After hybridization for 4 h at 48° C, 300 µl ice-cold S1 nuclease digestion buffer [0.28 M NaCl, 50 mM sodium acetate, 20 µg/ml denatured herring sperm DNA, 0.1 or 0.5 units/µl S1 nuclease (Gibco-BRL)] was added and the mixture was incubated for 30 min at 37° C. The reaction was terminated by addition of 80 µl termination buffer (2.5 M ammonium acetate, 50 mM EDTA) and the nucleic acids were precipitated by addition of two volumes of ethanol. S1 nuclease-protected fragments were electrophoresed on a 6% polyacrylamide sequence gel. After drying the gel was subjected to autoradiography.

RESULTS

Localization, cloning and sequencing of the polyhedrin gene

The location of the polyhedrin gene on the SeMNPV genome was investigated by hybridization of various restriction enzyme digests of SeMNPV DNA with AcMNPV fragment *Hin*dII-V as a probe under non-stringent conditions (Fig. 2.1). This fragment contains approximately 500 nucleotides (nt) from the 3' end of the AcMNPV polyhedrin gene (Smith *et al.*, 1983). An *Xba*I fragment of 1.7 kb (lane 2) and a *Sph*I fragment of 11.4 kb (lane 3) of SeNPV DNA gave positive signalw with the AcNPV probe and were further analyzed. The probe also hybridized with *Eco*RI fragment I of AcMNPV confirming the specificity of the *Hin*dIII V (lane 1). SeMNPV *Xba*I fragment R (1.7 kb) and *Sph*I fragment D (11.4 kb) were further analysed for the presence of various restriction sites. On the basis of the physical map obtained (Fig. 2.2) subclones were prepared for DNA sequencing of the SeMNPV polyhedrin gene.

A contiguous segment of approximately 1120 nt was sequenced encompassing the

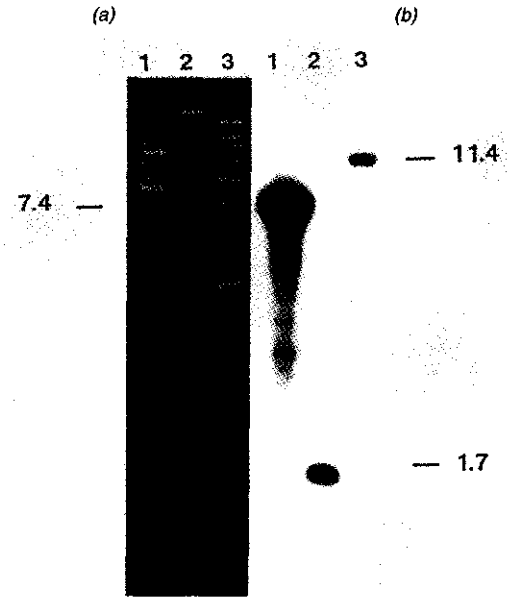


Figure 2.1. Localization of the SeMNPV polyhedrin gene. Restriction enzyme analysis (a) of AcMNPV DNA digested with *Eco*RI (lane 1) and SeMNPV DNA digested with *Xba*I and *Sph*I (lanes 2 and 3, respectively), and Southern blot hybridization (b) under non-stringent conditions with AcMNPV *Hin*dIII-V as a probe. The sizes of AcMNPV fragment *Eco*RI-I (7.4) and SeMNPV fragments *Xba*I-R (1.7) and *Sph*I-D are indicated (in kbp).

complete coding sequence of polyhedrin and some flanking sequences (Fig. 2.3). The open reading frame (ORF) of the SeMNPV polyhedrin gene contained 738 nt with the potential to encode a polypeptide of 246 amino acids. The predicted M_r (29 kDa) was similar to the apparent size of SeMNPV polyhedrin as determined by SDS-PAGE analysis (Caballero *et al.*, 1992).

In the 5' non-coding region two putative transcriptional start sites containing the canonical core TAAG (Vlak and Rohrmann, 1985) are found, one around residue -70 and another around residue -45 with respect to the translational start codon (Fig. 2.3). At residue -45 this core sequence is part of a dodecanucleotide sequence TGTAAGTAATTT, which is largely conserved among all polyhedrin genes sequenced so far and always contains the start signal for transcription (Rohrmann, 1992). No canonical TATA boxes are found in the region upstream from the two putative transcriptional start sites. Two short, imperfect tandem repeats with the sequence TATC(T)ATCGA were found at residues -85 and -112. Downstream from the putative transcription start site two other conserved motifs (Zanotto *et al.*, 1992) were located around residues -28 (TTCGTA) and -17 (TTGTGA), respectively.

In the 3' non-coding region of the polyhedrin gene the canonical polyadenylation signal sequence (AATAAA) was not found (Fig 2.3). However two signal-like sequences were observed, one (ATTAAA) as part of the translational stop signal and a second (AATTAAA) approximately 10 nt downstream from the first putative signal. Downstream from the polyhedrin-coding sequence the 3' end of another ORF was found on the opposite strand. At residue 782 a perfect putative polyadenylation signal for this gene was found.

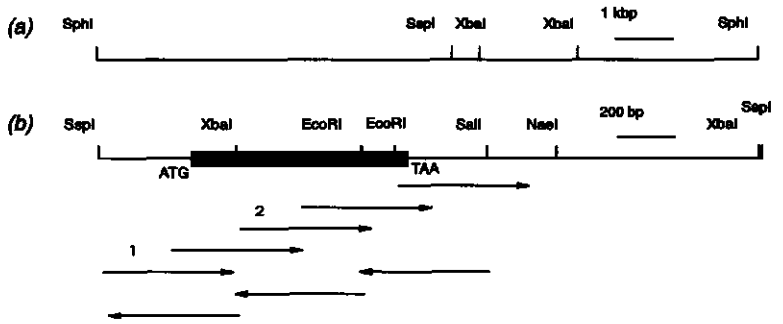


Figure 2.2. Physical map of an 11.4 kbp segment (a) of SeMNPV DNA and its 2.2 kbp subfragment (b) for restriction enzymes *Sph*I, *Ssp*I, *Xba*I, *Eco*RI and *Sal*I. The polyhedrin gene is indicated by the solid bar. The start (ATG) and stop (TAA) codons of the polyhedrin gene are indicated. The arrows indicate the direction in which the (sub)clones were sequenced, except 1 and 2 where specific primers were used.

Transcriptional analysis of the polyhedrin gene

To characterize the polyhedrin gene transcripts an attempt was made to isolate mRNA from *S. exigua* cells infected with SeMNPV. However, we were unable to isolate polyadenylated polyhedrin-specific mRNA by oligo(dT) column chromatography. In the same experiment large amounts of p10 mRNA were obtained (data not shown), suggesting that the polyhedrin mRNA is not polyadenylated or that its poly(A) tail is very short.

Northern blot analysis of total RNA from infected cells isolated at 24 and 48 h p.i. showed (Fig. 2.4a) an abundant mRNA of approximately 1000 nt when a polyhedrin-specific oligonucleotide (5'-TCGTACACGTAAGTGCG-3') was used as probe (lanes 2 and 3). Longer transcripts were not detected. The specificity of the probe was confirmed by the absence of a signal when RNA from uninfected cells was used (lane 1).

Primer extension was employed to determine the transcriptional start site of the polyhedrin mRNA. A 17-mer oligonucleotide, corresponding to the complement of the coding sequence from residue 36 to 53 (Fig. 2.3) was used in the primer extension assay of polyhedrin mRNA (Fig. 2.4b, lane 1). RNA from uninfected cells was used as a negative control (lane 2). An internal *Nru*I subclone of fragment *Sph*I-D served as a positive control for the assay and the sequence (other lanes). The reverse transcription extended as far as the A at residue -44 of the mRNA from the translational start site, whereas a minor signal suggested alternative starts at the T and A at positions -45 and -43, respectively (lane 1). The A at residue -44 is in the central region of the canonical transcriptional initiation motif (Rohrmann, 1986).

The 3' end of the SeMNPV polyhedrin mRNA was determined using S1 nuclease analysis using an *Eco*RI-*Sal*I probe of 320 nt (Fig. 2.4c, lanes 1-3) and an *Eco*RI-*Nae*I probe of 561 nt (Fig. 2.4c, lanes 4-6) overlapping the putative transcription termination sequence. Depending on the concentration of S1 nuclease (lanes 2-3 and 5-6) a fragment of 263 or 270 nt was protected. This positioned the 3' end of the polyhedrin gene transcript approximately 230 nt downstream of the translational stop signal in a highly AT-rich sequence. The same 3' end as well as the variation were found when the *Eco*RI-*Nae*I DNA probe of 561 nt was used (lanes 5-6).

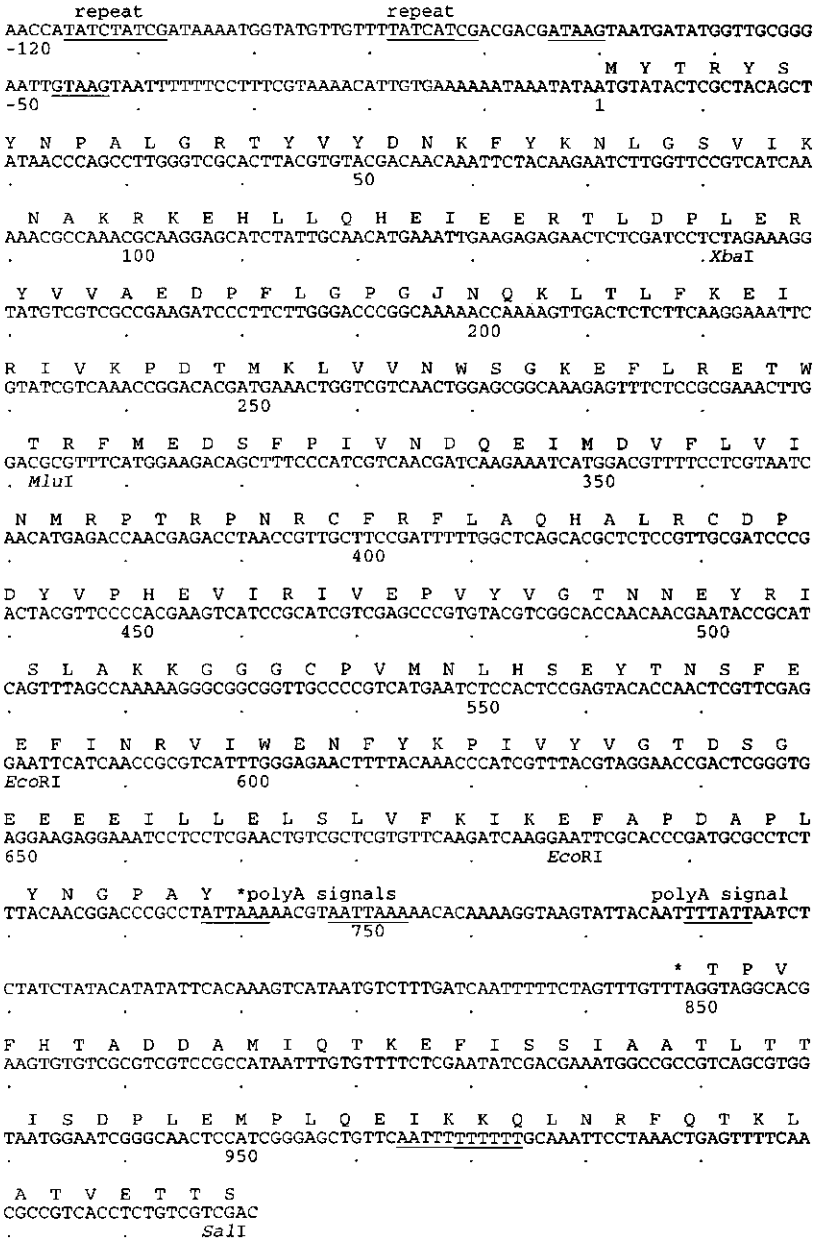


Figure 2.3. Nucleotide sequence of the SeMNPV polyhedrin gene and its flanking regions. The predicted amino acids are indicated with one-letter code designations. Putative transcription initiation and termination signals are underlined.

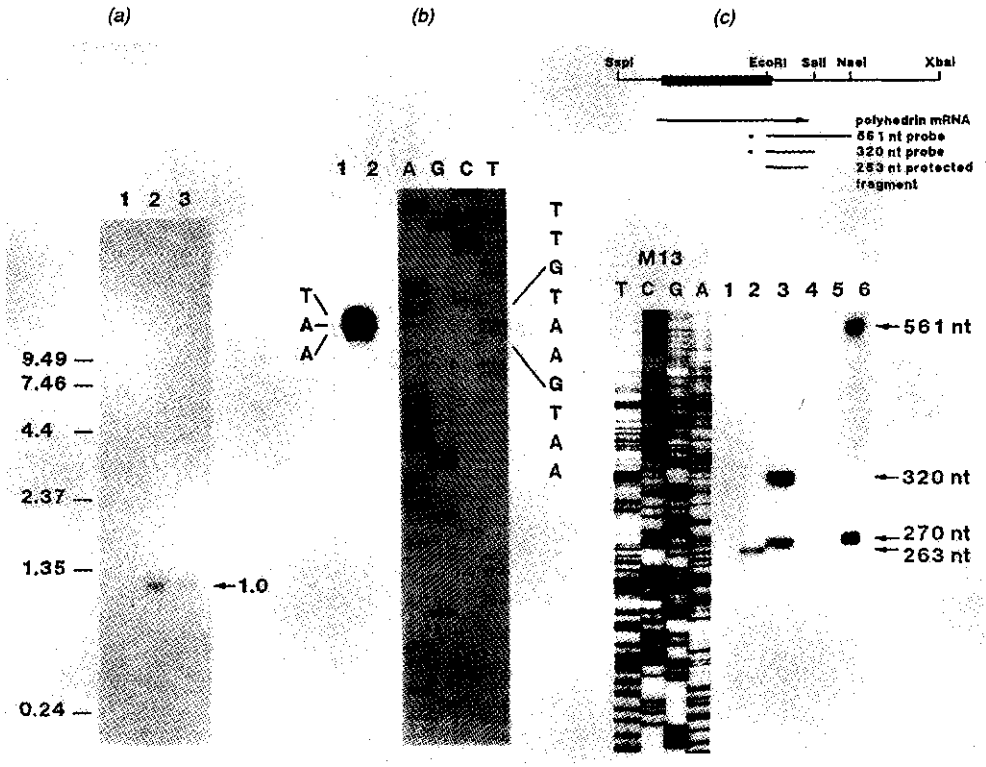


Figure 2.4. Transcriptional analysis of the SeMNPV polyhedrin gene. (a) Northern blot analysis of polyhedrin transcripts. Total RNA was extracted from uninfected *S. exigua* cells (lane 1) and SeMNPV-infected *S. exigua* cells 48 h p.i. (lane 2) and 24 h p.i. (lane 3) and analyzed in a 1% agarose gel (5 µg RNA per lane). The polyhedrin transcripts were detected with a ^{32}P-labeled SeMNPV polyhedrin gene-specific oligonucleotide probe (17-mer) complementary to the mRNA strand. The size of the mRNA is indicated in kb (right); a RNA ladder was used as molecular weight markers (left). (b) Size analysis of cDNA prepared by primer extension of an oligonucleotide complementary to the SeMNPV polyhedrin mRNA. The oligonucleotide was 5' end-labeled with [^{32}P]dATP, annealed to total RNA from SeMNPV-infected *S. exigua* cells 48 h p.i. (lane 1) and uninfected cells (lane 2) and elongated with reverse transcriptase. The sequence of the polyhedrin promoter (mRNA strand) is indicated by the lettering at the right. The sizes of the extension products were determined against a sequence ladder (AGCT) obtained from a SeMNPV plasmid clone using the same 17-mer extension oligonucleotide as in (a). (c) Nuclease S1 analysis of the 3' end of the SeMNPV polyhedrin mRNA. End-labeled probes were hybridized to total RNA (5 µg per lane) isolated from uninfected (lanes 1 and 4) or SeMNPV-infected *S. exigua* cells 48 h p.i. treated with 0.1 units (lanes 3 and 6) or 0.5 units (lanes 2 and 5) S1 nuclease, analyzed on denaturing sequencing gels. The sizes of the protected fragments were determined by comparison with a sequencing ladder of M13 ssDNA (M13).

Comparison of the SeMNPV polyhedrin gene with other baculoviruses

The amino acid sequence of the SeMNPV polyhedrin was compared to eight other MNPV polyhedrins (Fig. 2.5) and found to be most similar to SfMNPV polyhedrin with 98% amino acid identity and most distantly related to LdMNPV with only 81% identity (Table 2.1). The nucleotide sequences of the promoter region of SeMNPV and SfMNPV up to residue -58 were identical (Fig. 2.6). The sequences upstream from this region showed no similarity.

```
SeMNPV MYTRYSYNPA LGRTYVYDNK FYKNLGSVIK NAKRKEHLLQ HEIEERTLDP LERYVVAEDP
SfMNPV .........S .......... .......... ........AL .......... ..........
MbMNPV .........S .......... Y......... ..N..R.YIE ..L..K.... .D..L.....
PfMNPV .........S FP........ Y.....A... ..N..NDFIE .QL..K.... .D..L.....
BmMNPV .-PN.....T I......... Y.....GL.. .....K..IE ..K..KQW.L .DN.M.....
AcMNPV .-PD...R.T I......... Y.....A... .....K.FAE .....A.... .DN.L.....
OpMNPV .-PD...R.T I......... Y......... .....K...E ..ED.KH... .DH.M.....
AgMNPV .-PD...R.T I......... Y......... .....K...E .QEE.KS..G .DH.I.....
LdMNPV .HNF.N.S.. ..K....E.. Y.....T... Q...QK..E. ......S..H .D..L.....
                                                                        60

SeMNPV FLGPGKNQKL TLFKEIRIVK PDTMKLVVNW SGKEFLRETW TRFMEDSFPI VNDQEIMDVF
SfMNPV .......... .......... .......... .......... .......... ..........
MbMNPV .......... .......N.. .......... .......... .......... .....V....
PfMNPV .......... .......N.. .......... .......... .......... .....V....
BmMNPV .......... .....V.N.. ......I... .......... ...V...... .....V...Y
AcMNPV .......... .......N.. ........G. K....Y.... .......... .....V....
OpMNPV .......... .......N.. ......I... .......... ...V...... .....V....
AgMNPV .......... .......N.. ......I... .......... ...V...... .....V....
LdMNPV .Y........ .......N.. .......... .......... .......... .....V..IY
                                                                       120

SeMNPV LVINMRPTRP NRCFRFLAQH ALRCDPDYVP HEVIRIVEPV YVGTNNEYRI SLAKKGGGCP
SfMNPV .......... .......... .......... .......... ...N...... ..........
MbMNPV .......... ....K..... .......... .........S ...S.....V ....R.....
PfMNPV .......... ...YK..... ......E... .........S ...S.....V ....R.....
BmMNPV ..A.LK.... ...YK..... ...W.E.... ......M..S ...M...... ..........
AcMNPV ..V....... ...YK..... .......... .D.......S W..S...... ..........
OpMNPV ..V....... ...YK..... ...W.C.... .........S ...M...... ..........
AgMNPV ....L..... ...YK..... ...SHC.... .........S ...M...... ..........
LdMNPV .T..V..... ...YK.V... .....EG... .........S T.E-...... ....R.....
                                                                       180

SeMNPV VMNLHSEYTN SFEEFINRVI WENFYKPIVY VGTDSGEEEE ILLELSLVFK IKEFAPDAPL YNGPAY
SfMNPV .........H .......... .......... .......... .......... .......... ......
MbMNPV .......... .......... .......... .....A.... ....V..... .......... ......
PfMNPV .......... .......... .......... .....A.... ....V..... .......... ......
BmMNPV I..I...... ...S.V.... .......... I....A.... ..I.V..... .......... FT....
AcMNPV I......... ...Q..D... .......... I....A.... ....V..... V......... FT....
OpMNPV I..I.A.... ...S.V.... .......... I....S.... ..I.V..... V......... FT....
AgMNPV I..I...... ...S.V.... .......... I....S.... ..I.V..... V......... FT....
LdMNPV IR....A..T ...H.L.S.. .D........ ...T.A.... ....V..... .......... FQ....
                                                                       246
```

Figure 2.5. Comparison of the amino acid sequence of nine MNPV polyhedrins. The one-letter code designation is used. The hyphens denote identical amino acids. The abbreviation for the viruses are given in the text.

DISCUSSION

In this report the nucleotide sequence and transcriptional analysis of the SeMNPV polyhedrin gene are presented. Polyhedrin genes are highly conserved among baculoviruses (Vlak and Rohrmann, 1985) and the SeMNPV polyhedrin is no exception. The SeMNPV polyhedrin gene was most closely related to the polyhedrin gene of SfMNPV (Gonzales *et al.*, 1989), differing by only five amino acids (Fig. 2.5). DNA sequence comparison of SeMNPV and SfMNPV polyhedrin showed a difference in identity in 102 nt, of which only 9 were not in the wobble

Table 2.1. Amino acid and nucleotide sequence homology (identity) (%) of ten NPV polyhedrins*.

	*Sf*MNPV	*Mb*MNPV	*Pf*MNPV	*Ac*MNPV	*Op*MNPV	*Ag*MNPV	*Bm*MNPV	*Ld*MNPV
SeMNPV	**98**	**92**	**89**	**85**	**84**	**84**	**82**	**81**
	86	*82*	*81*	*76*	*77*	*75*	*74*	*73*
SfMNPV		**92**	**89**	**85**	**83**	**83**	**81**	**80**
		82	*81*	*79*	*77*	*74*	*75*	*76*
MbMNPV			**96**	**88**	**86**	**85**	**84**	**82**
			90	*80*	*77*	*74*	*75*	*74*
PfMNPV				**87**	**84**	**84**	**83**	**81**
				81	*77*	*76*	*75*	*75*
AcMNPV					**89**	**88**	**86**	**81**
					79	*76*	*75*	*75*
OpMNPV						**95**	**91**	**79**
						83	*81*	*75*
AgMNPV							**91**	**80**
							80	*75*
BmMNPV								**80**
								73

* Bold and italic lettering denote amino acid and nucleotide sequence identity, respectively.

position (data not shown). SeMNPV was most distantly related to LdMNPV (Smith *et al.*, 1988; Chang *et al.*, 1989) which were different with respect to 46 amino acids (Fig. 2.5).

The similarity between SeMNPV and SfMNPV polyhedrins is higher than with those of any of the other MNPVs (Table 2.1). The differences are located at sites known to be particularly variable among polyhedrins (Rohrmann, 1986) and are functionally conservative. This suggests that the polyhedrin genes of SeMNPV and SfMNPV have evolved from a common ancestor distinct from other MNPVs. Polyhedrins of SeMNPV and SfMNPV on the one hand are more related to *Mamestra brassicae* MNPV (MbMNPV, Cameron & Possee, 1989), *Panolis flammea* MNPV (PfMNPV, Oakey *et al.*, 1989) and *Orgyia pseudotsugata* SNPV (OpSNPV, Leisy *et al.*, 1986a), than to *Orgyia pseudotsugata* MNPV (OpMNPV, Leisy *et al.*, 1986b), *Anticarsia gemmatalis* MNPV (AgMNPV, Zanotto *et al.*, 1992), AcMNPV (Hooft van Iddekinge *et al.*, 1983), *Bombyx mori* NPV (BmMNPV, Iatrou *et al.*, 1985) or LdMNPV (Smith *et al.*, 1988; Chang *et al.*, 1989) (Table 2.1). Availability of polyhedrin protein sequences of other baculoviruses may aid in their classification and may help define baculovirus species.

By primer extension analysis the transcription initiation site for the polyhedrin gene mRNA of SeMNPV was mapped in the promoter within the TAAG motif at approximately residue -44 (Fig. 2.4c). This motif is conserved in all baculovirus late genes (Rohrmann, 1986; Blissard & Rohrmann, 1990). The major start site was at the A at -45, but the T at -46 and the A at -44 serve as minor start sites. This heterogeneity at the 5' end of the polyhedrin mRNA also exists in polyhedrin transcripts from other MNPVs as determined for AcMNPV (Hooft van Iddekinge *et al.*, 1983) and OpMNPV (Leisy *et al.*, 1986c) by S1 analysis and for AcMNPV (Howard *et al.*, 1986) and PfMNPV (Oakey *et al.*, 1989) by primer extension analysis. It is therefore unlikely that the upstream ATAAG motif at residue -72 is used for SeMNPV polyhedrin gene transcription. Whether this start site is used much earlier in infection, as has been demonstrated for other late genes, such as the gp64 gene of OpMNPV (Blissard & Rohrmann, 1989) requires further investigation.

The promoter regions of the SeMNPV and SfMNPV polyhedrin gene were identical as far

```
AcMNPV TGGAAATGTCTATCAATATATA---GTTGCTGATATCATGGAGATAATTAAAATGATAACCATCTCGCAA
          *  * ** ** **     **      *  *  ****** **       ***       *       *
SeMNPV CATATCTATCGATAAAATGGTA---TGTTGTTTTATCATCGACGACGATAAGTAATGATATGGTTGCGGG
       ****     *        ** ***    *  * ****  *            ***     *         ******
SfMNPV CATAATCCTATGCTGCATTGTAATCTAATCTTTTGCCGCATTTTTATCTAATCTTTTGCCGCACTGCGGG
       -120        .           .           .           .           .

AcMNPV ATAAATAAGTATTTTACTGTTTTCGTAACAGTTTTGTAATAAAAAAACCTATAAAT ATG
       *    ****** ***  *  ******** *  **    ** *** ***  ****     1
SeMNPV AATTGTAAGTAATTTTTTCCTTTCGTAAAACATTGTGAAAAAATAAA--TATA ATG
       ********************************************  **** 1
SfMNPV AATTGTAAGTAATTTTTTCCTTTCGTAAAACATTGTGAAAAAATAAA--TATA ATG
       -50          .           .           .           .          1
```

Figure 2.6. Computer alignment of the DNA sequences of the 5' sequence flanking the polyhedrin genes of SeMNPV, SfMNPV (Gonzales *et al.*, 1989) and AcMNPV (Smith *et al.*, 1983; Howard *et al.*, 1986). The tandem repeats and TAAG consensus sequence are underlined. Asteriks indicate nucleotide identity.

as residue -58 (with respect to the translational start). Upstream from the transcriptional start site the sequence was highly diverged (Fig. 2.6). A similar situation exists in the 5' flanking region of polyhedrins of other closely related MNPVs, such as MbMNPV (Cameron *et al.*, 1989) and PfMNPV (Oakey *et al.*, 1989). Two imperfect tandem repeats with the sequence TATC(T)ATCGA were found upstream from the transcriptional start site at residue -85 and -112. Similar repeats have been found in AcMNPV (Hooft van Iddekinge *et al.*, 1983) and SfMNPV (Gonzales *et al.*, 1989). There is, however, little evidence that these repeats contribute to the promoter activity of baculovirus polyhedrins as suggested from (deletion) mutagenesis experiments (Possee & Howard, 1987), although they may have an auxillary function (Ooi *et al.*, 1989).

Two conserved elements in polyhedrin promoters downstream from the TAAG motif have been found to share sequence homology and residual similarity with consensus regions for binding of transcriptional factors to 5S ribosomal RNA genes and to tRNA genes, respectively (Zanotto *et al.*, 1992). Such elements, TTCGTA around residue -28, and TTGTGA one helix turn downstream around residue -17, are also observed in the SeMNPV polyhedrin gene promoter. It has been suggested that the conserved positioning of these elements is important in transcriptional/translational regulation of polyhedrin gene expression (Zanotto *et al.*, 1992; Ohlendorf & Matthews, 1983; Ooi *et al.*, 1989).

The canonical polyadenylation signal AATAAA (Birnstiel *et al.*, 1985) is not present in the 3' end of the SeMNPV polyhedrin gene. However there are two similar motifs in proximity to the translational stop signal. One overlaps with the TAA (ATTAAA) of the translational stop and the other is located approximately 10 nt downstream. Both may serve as a signal for polyadenylation. The 3' non-coding sequence of the mRNA contained at least an additional 230 nt before polyadenylation could occur (Fig. 2.4a). This puts a transcriptional stop signal within an AT-rich region, which is almost invariably present in the transcription termination sequence of other baculovirus polyhedrin genes. The variation at the 3'-end of the message (Fig. 2.4c) may be due to microheterogeneity in the transcripts or to a slight 'breathing' of AT base pairs in this region during the S1 nuclease treatment.

In the process of mRNA isolation we were unable to isolate SeMNPV polyhedrin mRNA by oligo(dT)-cellulose chromatography. It is possible that the poly(A) tail, if present at all, is too

short to bind efficiently to the oligo(dT). This is supported by the observation that the mRNA size of about 1000 nt as observed in Northern analysis is close to the value of the primary transcript of 1020 nt as deduced from primer extension and S1 analysis (Fig. 2.4). If the SeMNPV polyhedrin is polyadenylated, then the poly(A) tail is either very small or rapidly degraded in the infected cell. The observation that the polyadenylation motif is not in the canonical format (Fig. 2.3) favours the hypothesis that the SeMNPV polyhedrin mRNA is not polyadenylated. This is in contrast to AcMNPV polyhedrin mRNA, which is polyadenylated (Smith *et al.*, 1983; Howard *et al.*, 1986).

There is a potential downstream ORF on the opposite strand terminating at residue 848 with a putative polyadenylation signal at residue 782 (Fig. 2.3). In AcMNPV an ORF of 1629 residues is present downstream of the polyhedrin gene which is running in the opposite orientation and which is essential for virus replication (Possee *et al.*, 1992). Comparison of our sequencing data indicate that the downstream sequences have no homology with the 1629 gene of AcMNPV. Transcripts from the polyhedrin gene and a putative downstream open reading frame have a considerable overlap in both AcMNPV and SeMNPV. Further sequencing, transcriptional mapping and deletion mutagenesis may elucidate whether the ORF in the downstream region is functionally similar to the AcMNPV 1629 gene.

The transcription of SeMNPV polyhedrin mRNA terminated in an AT-rich region within the sequence AATTTTTTTTT (AAT_9) at residue 970. Similar AT sequences are frequently found at the 3' end of other polyhedrin mRNAs. This AAT_9 motif is observed in SfMNPV (Gonzales *et al.*, 1989), MbMNPV (Cameron & Possee, 1989), PfMNPV (Oakey *et al.*, 1989) and AgMNPV (Zanotto *et al.*, 1992) in roughly the same position, but not in OpMNPV (Leisy *et al.*, 1986b) and AcMNPV polyhedrin mRNA (Possee *et al.*, 1992) although the 3' end of the polyhedrin mRNA in the latter two MNPVs is in an AT-rich region. Computer-assisted alignment of the 3' non-translated region of SeMNPV polyhedrin gene with other polyhedrins further showed limited sequence homology (data not shown). Some structural features, however, have been maintained, such as the location of (putative) polyadenylation signals and transcription termination sequences.

Information provided in this paper forms the basis for a future exploitation of the SeMNPV polyhedrin promoter for expression of foreign genes and for the engineering of SeMNPV to improve its insecticidal properties. Further sequencing and transcriptional analysis of the regions flanking the SeMNPV polyhedrin gene will give additional information on the organization and expression of the SeMNPV genome in comparison to AcMNPV and other baculoviruses.

Acknowledgements

We thank dr. B.A. Federici for *Spodoptera exigua* cell line UCR-SE-1 and for virus isolate SeMNPV (US). We appreciated the skilful technical assistance of Mrs M. Usmany during the course of these experiments.

CHAPTER 3

NUCLEOTIDE SEQUENCE AND TRANSCRIPTIONAL ANALYSIS OF THE P10 GENE OF *SPODOPTERA EXIGUA* NUCLEAR POLYHEDROSIS VIRUS

SUMMARY

The p10 gene of *Spodoptera exigua* multiple nuclear polyhedrosis virus (SeMNPV) was localized on the *Xba*I fragment H (5.1 kb) of the physical map of the viral genome. The coding sequence of the SeMNPV p10 gene appeared 264 nucleotides (nt) long corresponding with a predicted protein of 88 amino acids with a M_r of 9607 Da. The SeMNPV p10 protein showed only limited amino acid identity (39% and 26%, respectively) to the p10 proteins of *Orgyia pseudotsugata* MNPV (OpMNPV) and *Autographa californica* MNPV (AcMNPV) and thus appears less conserved than other viral proteins. The SeMNPV p10 gene was expressed by a transcript of approximately 450 nt in length, which started in the conserved baculovirus late gene promoter motif TAAG. The leader of the SeMNPV p10 transcript was AT rich (92%) and at 36 nt was the shortest leader of all baculovirus major late genes reported on so far. The SeMNPV p10 transcript terminated 6 nt downstream from a putative polyadenylation signal sequence (AATAAA); the latter was 61 nt downstream of the translational stop codon TAA. Upstream and downstream of the p10 gene, partial putative ORFs were found that showed significant amino acid sequence identity to the baculovirus p26 and p74 proteins. It is concluded that the region of SeMNPV DNA containing the p10 gene is collinear with the corresponding regions in the AcMNPV and OpMNPV genomes.

This chapter has been published as: Zuidema, D., Oers, M.M. van, Strien, E.A. van, Caballero, P.C., Klok, E.-J., Goldbach, R.W., & Vlak, J.M. 1993. Nucleotide sequence and transcriptional analysis of the p10 gene of *Spodoptera exigua* nuclear polyhedrosis virus. Journal of General Virology 74, 1017-1024.

INTRODUCTION

Nuclear polyhedrosis viruses (NPVs), a subgroup of the *Baculoviridae* (Francki *et al.*, 1991), produce two polypeptides, polyhedrin and p10, at very high levels in the late phase of infection. At 48 h post infection (p.i.) these two polypeptides constitute about half of the protein mass in insect cells (Smith *et al.*, 1982). Polyhedrin is the major constituent of the viral occlusion bodies which are found exclusively in the nuclei of baculovirus-infected insect cells (Vlak & Rohrmann, 1985). The p10 protein is associated with large fibrillar structures in both nuclei and cytoplasm of baculovirus-infected insect cells (Harrap, 1970; van der Wilk *et al.*, 1987; Williams *et al.*, 1989; Russell *et al.*, 1991).

In view of its abundance in infected cells and its particular structure it is surprising that no more specific functions have been attributed as yet to the p10 protein. Studies with p10 deletion mutants of *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) have shown that the p10 protein is not essential for viral replication both *in vivo* or *in vitro*, but that insect cells infected with p10 mutants of AcMNPV fail to liberate polyhedra (Williams *et al.*, 1989; Van Oers *et al.*, 1993).

Only two baculoviral p10 genes have been studied in detail. The p10 coding sequences of AcMNPV (Kuzio *et al.*, 1984) and *Orgyia pseudotsugata* MNPV (OpMNPV; Leisy *et al.*, 1986a) are similar in size, 282 and 276 nucleotides (nt), respectively. In contrast to the highly conserved polyhedrin genes with 80% nucleotide sequence and 90% amino acid sequence identity, the p10 genes have only 54% and 41% identity at the nucleotide and amino acid sequence levels, respectively (Leisy *et al.*, 1986b). Despite the high degree of similarity between the AcMNPV and OpMNPV genomes (Rohrmann, 1992) the p10 genes have diverged considerably. It is therefore of interest to investigate the p10 genes of more distantly related baculoviruses and to determine to which extent the structural homology in the p10 protein is preserved.

Spodoptera exigua MNPV (SeMNPV) is a baculovirus specific for a single insect species, the beet armyworm *Spodoptera exigua*, whereas AcMNPV for example can infect over 30 insect species. Information on the genetic organization of SeMNPV, in particular its nucleotide sequence and regulation of gene expression, may shed more light not only on the evolution and speciation of baculoviruses, but also on the molecular basis of host specificity. This report describes the identification, nucleotide sequence and transcriptional analysis of the SeMNPV p10 gene.

MATERIALS AND METHODS

Cells and virus. The *S. exigua* cell line UCR-Se-1, obtained from Dr B.A. Federici (Department of Entomology, University of California, Riverside, Ca., U.S.A.), was used and maintained in plastic tissue culture flasks (Nunc) in TNM-FH medium (Hink, 1970) supplemented with 10% foetal calf serum. A field isolate of SeMNPV (Gelernter & Federici, 1986a), kindly provided by Dr B.A. Federici, was used to infect fourth instar larvae of *S. exigua* (Smits *et al.*, 1988). The haemolymph of these infected larvae was used as to obtain budded virus for infection of cultured UCR-SE-1 cells (Gelernter & Federici, 1986b).

RNA isolation and cDNA synthesis. Total RNA was isolated from SeMNPV-infected cells at 24, 48 and 74 h p.i., essentially as described by van Strien *et al.* (1992). Polyadenylated RNA (74 h p.i.) was selected by oligo(dT)-cellulose column chromatography (Sambrook *et al.*, 1989) and served as template in a reverse transcription reaction. Poly(A)$^+$ RNA (1.5 µg) and oligo(dT)$_{12\text{-}18}$ (500 ng) were incubated in a volume of 5 µl for 5 min at 70 °C, followed by an incubation for 1 h at 42 °C in a reaction volume of 50 µl after addition of 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM DTT, 1 mM dGTP, 1 mM dTTP, 20 µCi [α-^{32}P]dATP, 20 µCi [α-^{32}P]dCTP, 80 units RNasin, 4 mM sodium pyrophosphate and 28 units of reverse transcriptase (Life-Technologies). After alkali degradation of the RNA, cDNA was purified over a Sephadex G50 column and directly used in Southern blot hybridization.

Viral DNA, Southern blot hybridization and cloning. SeMNPV DNA was isolated from virions, which were liberated from polyhedra upon alkali treatment and cleared by sucrose gradient centrifugation (Caballero *et al.*, 1992). SeMNPV DNA was digested with various restriction enzymes, electrophoretically separated in 0.7% agarose gels using standard techniques (Sambrook *et al.*, 1989) and transferred under alkali conditions to Hybond-N nylon membranes (Amersham) by capillary transfer according to the manufacturer's specifications. After prehybridization at 60 °C (3 h) in 6 x SSC, 0.5% (w/v) SDS, 1% Ficoll-400, 1% polyvinylpyrrolidone, 1% BSA (fraction V) and 100 µg/ml herring sperm DNA, blots were hybridized at 60 °C for 20 h to a ^{32}P-labeled cDNA probe (1 x 10^6 cpm/ml) synthesized from poly(A$^+$) RNA isolated from SeMNPV-infected cells 74 h p.i. After hybridization membranes were washed twice for 5 min at room temperature in 4 x SSC, 0.1% (w/v) SDS, followed by two 10 min washes in 2 x SSC, 0.1% (w/v) SDS at 45 °C. Membranes were exposed to Kodak XAR-5 X-ray film using intensifying screens. On hybridization with the cDNA probe, *Xba*I fragment H of SeMNPV DNA and accompanying subfragments were isolated from agarose gels by the freeze-squeeze method (Tautz & Renz, 1983) and (sub)cloned in plasmid pUC19 or Bluescript SK$^+$.

DNA sequencing and computer analysis. All sequencing reactions were performed with dsDNA plasmid templates using T7 DNA polymerase and the chain termination method of Sanger *et al.* (1977). Sequence analyses were performed using the UWGCG computer programs (Devereux *et al.*, 1984). The DNA sequence and the deduced amino acid sequence of the open reading frame (ORF) were compared with sequences in the updated Genbank/EMBL, SWISSPROT and NBRF databases, using the FASTA and TFASTA programs.

Northern blot hybridization. Total RNA was denatured, electrophoresed and blotted, as described by van Strien et al. (1992). To identify SeMNPV p10-specific transcripts, blots were hybridized for 16 h at 48 °C in the prehybridization buffer to a 5'-end labeled (^{32}P) oligonucleotide (5' GCTGCAAAGCATCGAC 3'), deduced from the p10 gene nucleotide sequence. After hybridization, blots were washed for 5 min at room temperature in 4 x SSC, 1% SDS and twice in 2 x SSC, 1% SDS. Autoradiography was performed as described above.

Primer extension and S1 nuclease assay. Primer extension analysis was performed, essentially as described by van Strien *et al.* (1992), with total RNA (2 µg) isolated at 48 h p.i. from mock-

and SeMNPV-infected UCR-SE-1 cells and the 16-base oligonucleotide (see above) complementary to the SeMNPV p10 mRNA. S1 nuclease protection assays were carried out, essentially as described by van Strien *et al.* (1992), using as a probe denatured DNA ^{32}P labelled at the 3' end, hybridized (4 h at 48 °C) to 5 µg total RNA isolated from mock- or SeMNPV-infected UCR-Se-1 cells.

RESULTS

Mapping, cloning and sequencing of the SeMNPV p10 gene

Southern blot hybridization at low stringency, using the AcMNPV p10 gene as heterologous probe [a 555 bp fragment (*Xho*I-*Ssp*I) of *Eco*RI fragment P of AcMNPV-DNA], did not result in the detection of a similar sequence on the SeMNPV genome. Assuming that p10 mRNA occurs abundantly late after infection (Smith *et al.*, 1983), poly(A)$^+$ RNA was isolated from mock- and SeMNPV-infected insect cells 74 h p.i. and used as template for cDNA synthesis. This radiolabeled cDNA was then hybridized to Southern blots containing SeMNPV DNA digested with *Xba*I and *Sph*I. A strong hybridization signal to an *Xba*I fragment of 5.1 kb (*Xba*I-H) and to a *Sph*I fragment of 2.4 kb (*Sph*I-R) was observed (Fig. 3.1, lanes 2 and 3, respectively). Since polyhedrin mRNA was not polyadenylated (Van Strien *et al.*, 1992), this hybridization signal was most likely to be due to the p10 sequences.

The *Xba*I fragment H was cloned and a detailed physical map for various restriction endonucleases was constructed (Fig. 3.2). After hybridization with radiolabeled cDNA as described above, subclones were identified and a 914 nt fragment (*Sph*I-*Sst*I) was sequenced (Fig. 3.3). It contained an ORF of 264 nt with the potential to encode a protein of 88 amino acids with a predicted M_r of 9607 Da. This ORF presumably represented the p10 gene of SeMNPV (see below).

The 5' non-coding sequence of the SeMNPV p10 gene contained a TAAG motif which is the consensus sequence of the transcriptional start site in baculovirus late and very late gene expression (Blissard & Rohrmann, 1990). The TAAG consensus sequence was located at position -37 to -34 with respect to the translational start codon of the SeMNPV p10 ORF (Fig. 3.3). A putative conserved motif ATTGTA was identified two DNA helix turns downstream of the putative late transcription start site TAAG of the SeMNPV p10 gene, and was suggested to be involved in ribosome binding (Zanotto *et al.*, 1992). A sequence motif consisting of six repeats of the triplet GAC was found upstream (position -63 to -80) and downstream (position 276 to 293) of the p10 polypeptide coding sequence. Two putative polyadenylation signal sequences (AATAAA) were found in the 3' non-coding region of the putative p10 gene, located approximately 60 nt and 210 nt downstream of the translational stop codon (Fig. 3.3).

Contiguous ORFs were identified upstream and downstream from the p10 gene, and these extended beyond the sequenced region (Fig. 3.3). The p10 gene and the upstream ORF had the same polarity, whereas the downstream ORF had an opposite polarity. The partially identified upstream ORF had a high amino acid sequence similarity (53% and 55%, respectively) to the p26 genes of AcMNPV and OpMNPV, whereas the determined sequences of the downstream ORF had a high amino acid similarity (77% and 71%, respectively) to the p74 genes of AcMNPV and OpMNPV. These partial ORFs thus were more highly conserved than the p10 gene. From these data it is likely that the 264 nt ORF encodes the SeMNPV p10 gene

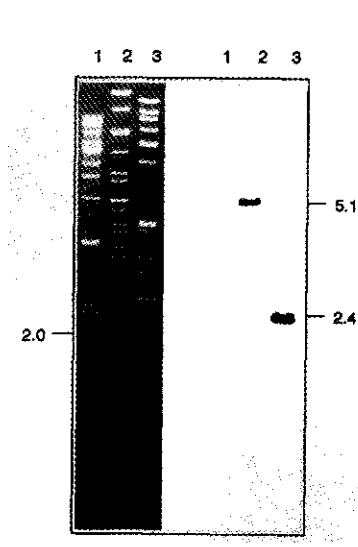


Figure 3.1. a) Identification and mapping of the SeMNPV p10 gene. Restriction enzyme analysis of AcMNPV DNA digested with *Eco*RI (lane 1) and SeMNPV DNA digested with *Xba*I and *Sph*I (lanes 2 and 3, repectively). *(b) Southern blot hybridization of (a) with cDNA obtained after reverse transcription of poly(A)* RNA extracted from SeMNPV-infected insect cells 74 h p.i. The sizes of AcMNPV fragment *Eco*RI-P (2.0) and SeMNPV fragments *Xba*I-H (5.1) and *Sph*I-R (2.4) are indicated.

and that the location of the SeMNPV p10 gene is collinear with the p10 gene and its flanking regions of AcMNPV and OpMNPV.

Transcriptional analysis of the SeMNPV p10 gene

SeMNPV p10-specific transcripts were identified by Northern blot hybridization (Fig. 3.4a). Total and poly(A)$^+$ RNA isolated from mock-infected and SeMNPV-infected cells at 24 and 48 h p.i. were hybridized with a synthetic 16-mer oligonucleotide probe (5' GCTGCAAAGCATCGAC 3'), complementary to the p10 coding sequence from residues 58 to 73 (Fig. 3.3). The oligonucleotide probe detected a single specific mRNA of approximately 450 nt (Fig. 3.4a, lanes 1, 2 and 3). No signal was observed with RNA isolated from mock-infected cells (lane 4).

To determine the p10 transcriptional start site, reverse transcription experiments were performed with RNA isolated 48 h p.i. using the synthetic p10-specific 16-mer oligonucleotide in a primer extension assay (Fig 3.4b, lane 1). RNA isolated from mock-infected cells served as a negative control (lane 2). The primer extension products were identified by running alongside a dideoxy-nucleotide sequencing ladder, obtained from a p10 plasmid clone and the p10-specific oligonucleotide, on a denaturing polyacrylamide gel (lanes A, G, C and T). Two major transcription initiation sites were located in the conserved canonical core sequence TAAG at residues -36 (A) and -37 (T) with respect to the p10 translational start site (lane 2). This implies that the leader sequence of the SeMNPV p10 transcripts is only 36 to 37 nt in length and has an A/T content of 92%. A minor reverse transcription product was detected starting at residue -208 (A). No obvious motif that would account for this minor transcription start could be found

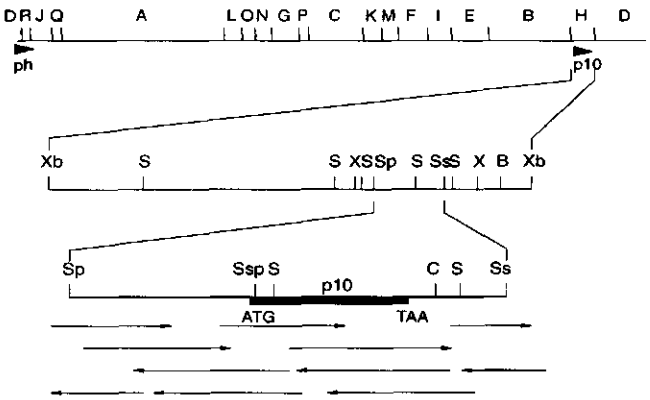


Figure 3.2. Location of the p10 gene on the physical map for *Xba*I restriction endonuclease of the SeMNPV genome and sequence strategy of the p10 gene region. The *Sph*I-*Sst*I region (914 bp) of *Xba*I fragment H was sequenced and the location of the SeMNPV p10 gene is indicated (thick line). The arrows below the lower map indicate the sequencing strategy of the 914-bp region. The location of the previously identified polyhedrin gene (ph) is shown as a reference (van Strien *et al.*, 1992). B, *Bam*HI; C, *Cla*I; S, *Sal*I; Sp, *Sph*I; Ss, *Sst*I; Ssp, *Ssp*I; X, *Xho*I; Xb, *Xba*I.

in the primary sequence (Fig. 3.3). Alternatively, this product may have originated from non-specific priming in the RNA pool. Computer-aided research did not reveal possible splice acceptor sites in the upstream sequences of the p10 gene.

The 3' end of the SeMNPV p10 transcripts was determined by S1 nuclease analysis using a *Sal*I-*Bam*HI fragment probe of 900 nt (Fig. 3.4c, lanes 1 and 2). This fragment was generated by digestion of *Xba*I fragment H with restriction endonuclease *Bam*HI and partially with *Sal*I. A major protected fragment of 292 nt was observed. This mapped the 3' terminus of the SeMNPV p10 transcript (Fig. 3.3, nt 339) within 10 nt downstream of the putative polyadenylation signal sequence AATAAA, which is located 61 nt downstream of the p10 ORF.

Comparison of the p10 genes of SeMNPV, AcMNPV, and OpMNPV

The SeMNPV p10 ORF (264 nt) was similar in size to the p10 ORFs of OpMNPV (276 nt) and AcMNPV (282 nt). The primary sequence of these genes has diverged considerably in these three viruses. The overall nucleotide and amino acid sequence identity was 26 to 39 % (Table 3.1 and Fig. 3.5). Despite this sequence divergence the hydrophilicity plots showed a reasonable similarity between the three p10 proteins (Fig. 3.6). A weakly hydrophobic region was present at the N terminus of the protein, whereas the C terminus was highly hydrophilic and contained several basic amino acid residues. Furthermore, the C-terminal half of all three p10 proteins shared a high proline content. The p10 proteins did not contain cysteine, tryptophan or histidine residues.

Dot plot comparisons between the three different p10 proteins showed two stretches of homology common to all three proteins (Fig. 3.7). The position of the first homology box - KV(D/T)(A/S)(L/V)Q - was conserved in the primary structure of the p10 proteins (Fig. 3.5), but the position of the second conserved homology box - (L/V/I)PD(V/L)P - was located at different positions in the primary structure.

The 5' untranslated leaders of the three p10 genes differed considerably in size (36 nt in

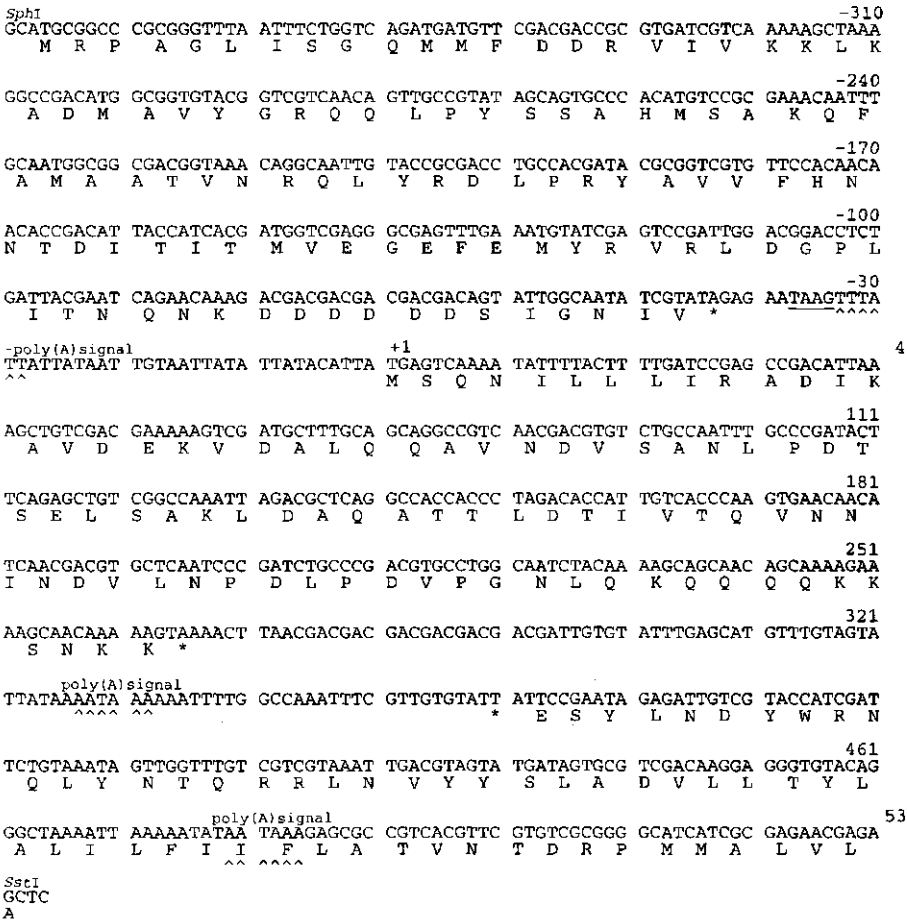


Figure 3.3. Nucleotide sequence of a 914-bp region of the SeMNPV genome containing the p10 gene. The p10 coding sequence starts at position +1 and terminates at position +267. The predicted amino acid sequence is indicated by the one-letter code. The transcription initiation signal (TAAG) is underlined. Putative transcription termination signals (AATAAA) are indicated by arrowheads.

SeMNPV, 52 nt in OpMNPV and 70 nt in AcMNPV), but were all A/T rich (92%, 77%, and 87%, respectively). The 3' untranslated sequences of the p10 mRNAs (85 nt in SeMNPV, 175 nt in OpMNPV, and 162 nt in AcMNPV) had no additional features in common, other than that all three p10 transcripts terminated at about 15 nt downstream of the eukaryotic consensus polyadenylation signal sequence AATAAA (Birnstiel *et al.*, 1985).

DISCUSSION

This study describes the identification and characterization of the putative p10 gene of SeMNPV. The latter was sequenced and its expression was analyzed by Northern blot, primer

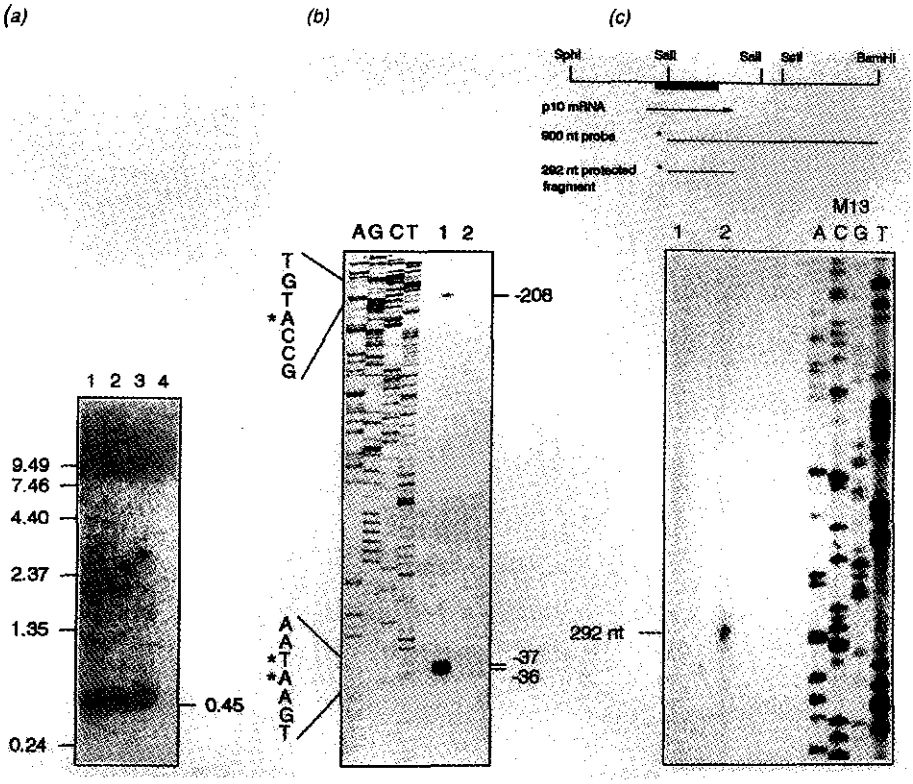


Figure 3.4. Transcriptional analysis of the SeMNPV p10 gene. (a) Northern blot analysis of total RNA extracted from uninfected (lane 4) and SeMNPV-infected insect cells 48 h (lane 1) and 24 h p.i. (lane 2), and poly A$^+$ RNA extracted from SeMNPV-infected insect cells 48 h p.i. (lane 3). The p10 transcripts were detected with a p10-specific oligonucleotide (16-mer) complementary to the p10 mRNA. An RNA ladder was used as size marker (left). The size of the p10 transcript is indicated on the right (in kb). (b) Primer extension analysis of the SeMNPV p10 transcripts performed with a 16-mer oligonucleotide complementary to the p10 mRNA, ^{32}P-labelled at the 5' end. The oligonucleotide was annealed to total RNA (5 µg) from uninfected (lane 2) and SeMNPV-infected insect cells (lane 1) 48 h p.i. and elongated by reverse transcription. The sizes of the extension products (right, see Fig. 3.3) were determined by comparison with a sequence ladder (lanes A, G, C and T) obtained from an SeMNPV p10 gene-containing plasmid clone and the same 16-mer oligonucleotide as a sequencing primer. Asterisks in the nucleotide sequence (left) indicate the position of the primer extension products. (c) S1 nuclease analysis of the 3' end of the SeMNPV p10 transcript. The 900 bp ^{32}P-labelled probe was hybridized to total RNA (5 µg) from uninfected (lane 1) and SeMNPV-infected insect cells 48 h p.i. (lane 2), treated with 0.2 units nuclease S1, and analyzed on a denaturing sequencing gel. The sizes of the protected fragments were determined by comparison with a sequencing ladder of M13 ssDNA (M13). The size and location of the probe used to map the 3' end, and the protected fragment are represented at the top.

extension and S1 analysis. It was mapped by Southern blot hybridization of digested SeMNPV DNA, using a cDNA probe which was derived from RNA isolated from SeMNPV-infected insect cells late after infection. The p10 gene was mapped to *Xba*I-H in the SeMNPV genome (Fig. 3.2) and is flanked by ORFs with homology to baculovirus p26 and p74 genes. This organization is collinear with the location of the p10 genes of AcMNPV and OpMNPV. The limited sequence

Table 3.1. Nucleotide sequence identity and amino acid sequence identity or similarity between p10 genes

of SeMNPV, OpMNPV and AcMNPV.

	Nucleotide level (%)	Amino acid level (%)	
	Identity	Identity	Similarity
SeMNPV-OpMNPV	36	39	55
SeMNPV-AcMNPV	38	26	44
AcMNPV-OpMNPV	53	42	62

similarity between the p10 genes accounted for the failure to detect the SeMNPV p10 gene with the AcMNPV p10 gene probe. The abundance of p10 mRNA in cells late after infection can be used to locate p10 genes of other baculoviruses via cDNA as probe. Alternatively, due to the conserved nature of p74 sequences and their apparent collinearity with p10 genes, p74 sequences can be used as probes to detect p10 genes.

Two major start sites of the SeMNPV p10 transcripts were mapped within the conserved canonical core TAAG at position -37 and -36 (Fig. 3.3, T and A, respectively). Transcription of most baculovirus late genes appear to involve the conserved TAAG motif (Blissard & Rohrmann, 1990). With a leader of 36 nt, a coding sequence of 264 nt and a 3' non-coding sequence of 75 nt the length of the SeMNPV p10 transcript, without a poly(A) tail, was calculated to be 375 nt. The Northern blot indicated a transcript size of about 450 nt (Fig. 3.4a). This implied that the SeMNPV p10 transcript contained a poly(A) tail of about 90 nt in length.

```
SeMNPV  M S Q - N I L L L I R A D I K A V D E K V D A L Q Q A V N D V S A N  33
OpMNPV  . . K P S . . T Q . L D A V R . . . S . . T . . . T Q . D Q L V E D  34
AcMNPV  . . K P N V . T Q . L D A V T E T N T . . D S V . T Q L N G L E E S  34
                                                  * * * * * * *

SeMNPV  L P D T S E L S A K L D A Q A T T L D T I V T Q V N N I N D V L N P  67
OpMNPV  S K T L E A . T D Q . G E L D N K V S D . Q S M L S V E E E L P E .  68
AcMNPV  F Q L L D G . P A Q . T D L N T K I S E . Q S I L T G D I V P D L .  68

SeMNPV  D L P D V P G N L Q K Q Q Q Q K K S N K K  88
OpMNPV  P A . A P E P E . P E I P D V P G L R R S R K Q  92
AcMNPV  D S L K P K L K S Q A F E L D S D A R R G K R S S K  94
```

Figure 3.5. Alignment of the deduced amino acid sequence of the p10 proteins of SeMNPV, OpMNPV (Leisy *et al.*, 1986a) and AcMNPV (Kuzio *et al.*, 1984). Homology box KV(D/T)(A/S)(L/V)Q-is indicated by asterisks and homology box -(L/I/V)PD(V/L)P has been underlined. Identical amino acids are indicated by dots. The hyphen represents a gap introduced to maximize matching of amino acid residues. Numbers indicate amino acid positions.

Recently Zanotto *et al.* (1992) identified two putative structural elements conserved in the 5' non-coding region of all baculovirus polyhedrin genes to date. These elements share sequence homology and positional similarity within the 5' regulatory consensus regions of 5S ribosomal and tRNA genes, which are involved in binding of class III transcription factors. One of these elements, ATTGTA, was found in the 5' non-coding region of the SeMNPV p10 gene, 13 to 18 nt downstream of the p10 transcription start site. A similar sequence, TTTGTA, existed in the promoter region of AcMNPV p10 gene (Kuzio *et al.*, 1984). However, neither of these motifs was present in the promoter region of OpMNPV p10 gene (Leisy *et al.*, 1986a). The relevance of

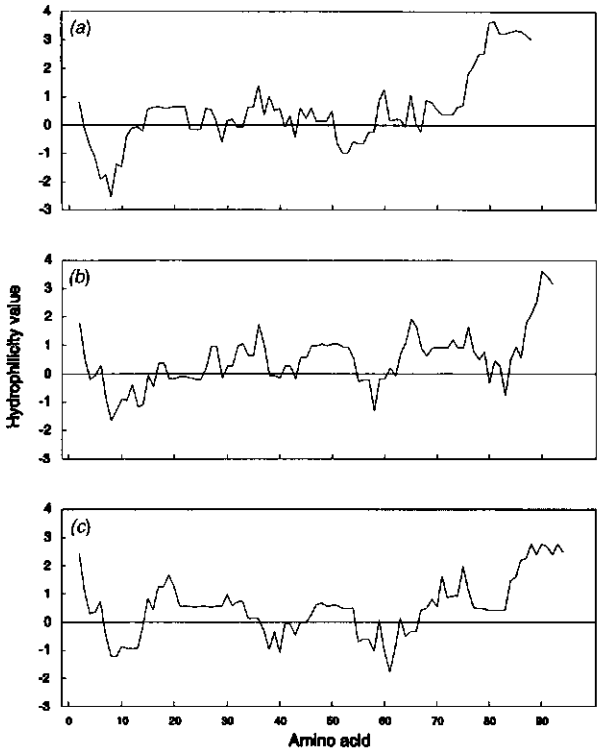


Figure 3.6. Hydrophilicity plots of the p10 proteins of SeMNPV, OpMNPV and AcMNPV. The methionine residue at the N-terminus of the p10 protein has been omitted. The plots were constructed using the algorithms of Kyte & Doolittle (1982), calculated with the PEPTIDESTRUCTURE programme of the GCG sequence analysis software package, and plotted with the LOTUS and FREELANCE programs (Lotus Development Corporation). The computer program plots the sums of hydrophilicity values for seven contiguous amino acids over the position of the middle amino acid in each sector. Above the zero line the plot indicates sectors with an overall hydrophilic character, whereas below the zero line it indicates sectors with an overall hydrophobic character.

these motifs for p10 gene expression therefore remains unclear.

Partial ORFs were detected upstream and downstream from the SeMNPV p10 gene, extending beyond the sequenced region. The putative upstream ORF, which terminated 5 nucleotides upstream of the ATAAG transcription initiation motif of the p10 gene, was transcribed in the same direction as the p10 gene (Fig. 3.3). This partial ORF shared 55% and 53% amino acid similarity with the OpMNPV and AcMNPV p26 ORFs, respectively. Both AcMNPV and OpMNPV p26 coding sequences were located directly upstream and were transcribed in the same direction as the their respective p10 genes. A perfect polyadenylation signal sequence (AATAAA) could not be detected between this putative p26 ORF and the p10 coding sequences, as was also the case in AcMNPV and OpMNPV. The SeMNPV p26 transcripts may coterminate with the p10 transcripts, as is the case for AcMNPV and OpMNPV (Rankin *et al.*, 1986; Bicknell *et al.*, 1987). However, Northern blot analysis of RNA isolated from SeMNPV-infected insect cells at 24 h p.i. and 48 h p.i., and probed with a p10-specific

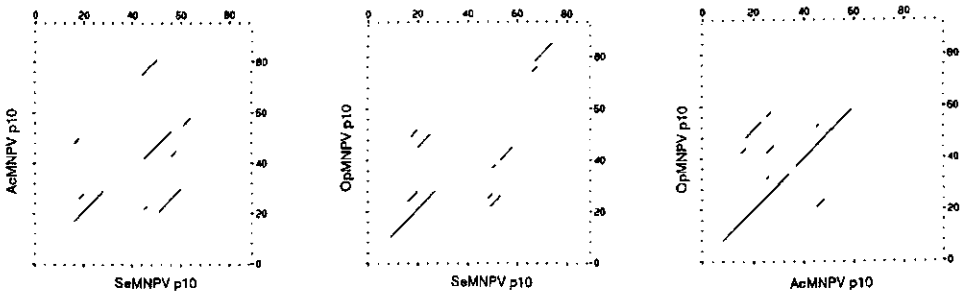


Figure 3.7. Dot plot comparisons of SeMNPV, OpMNPV and AcMNPV p10 proteins. The comparisons were performed with a window size of 21 and a stringency of 14 using the COMPARE and DOTPLOT programs of the GCG sequence analysis software package. The plots were generated with the FIGURE program of GCG.

oligonucleotide, detected only a single specific mRNA of approximately 450 nt in length (Fig. 3.4a), which encoded the p10 polypeptide.

The putative downstream ORF, which terminated at a TAA stop codon 26 nt downstream of the putative polyadenylation signal sequence (AATAAA) of the p10 transcripts, is transcribed in the opposite direction to that of p10 (Fig. 3.3). This partial ORF shared 77% and 71% amino acid similarity with the AcMNPV and OpMNPV p74 protein, respectively (Kuzio *et al.*, 1989; Leisy *et al.*, 1986a). Their was no putative polyadenylation signal sequence (AATAAA) detected between the putative p74 ORF and the p10 ORF in SeMNPV, as also observed in AcMNPV (Kuzio *et al.*, 1989). However, a putative polyadenylation signal sequence was detected in the opposite direction just downstream of the transcription start site (ATAAG) within the 5'-noncoding region of the SeMNPV p10 transcripts, a feature also shared by AcMNPV.

Although the p10 region seems to be conserved among different baculoviruses, the function of the p10 protein in the viral replication cycle is still an enigma, but a role in cell lysis has been suggested (Williams *et al.*, 1989). Mutational analysis of the AcMNPV p10 gene showed that the C terminus was essential for the formation of fibrillar structures, but not for liberation of polyhedra from the nuclei (van Oers *et al.*, 1993). Insect cells infected with AcMNPV mutants with deletions up to 15 amino acids from the carboxy terminus of the p10 protein showed a normal release of polyhedra from the cells, but no formation of fibrillar structures.

Recently the nucleotide sequence of the *Choristoneura fumiferana* MNPV (CfMNPV) p10 gene became available through the EMBL data library (accession no. M98513). Computer-aided research revealed that the putative CfMNPV p10 protein contained 81 amino acids, was devoid of cysteine, histidine and tryptophan residues and had, like other p10 proteins, a highly basic C terminus. The SeMNPV p10 protein shared an amino acid similarity of 37% with the CfMNPV p10 protein. The two conserved homology boxes (Fig. 3.5) were also preserved in the CfMNPV p10 coding sequences.

From the available data on the p10 genes of AcMNPV, OpMNPV, CfMNPV and SeMNPV it can be postulated that the size of the ORF, and the structure and function of the protein it encodes, are preserved among nuclear polyhedrosis viruses, whereas the nucleotide and amino acid sequence of the gene product are much more divergent than those of other baculovirus

genes encoding homologous functions (Rohrmann, 1992). This suggests that the p10 gene has evolved from an ancestor gene more rapidly than other baculovirus genes.

ACKNOWLEDGEMENTS

This work was financially supported in part by Solvay & Cie, Brussels, Belgium, through a fellowship to M.M. van Oers. D. Zuidema was a recipient of a fellowship of the Royal Netherlands Academy of Arts and Sciences. P. Caballero was the recipient of a grant from the Ministry of Education and Science of Spain. We thank Dr C. Santiago-Alvarez, University of Cordoba, Cordoba, Spain, for his continuous interest.

CHAPTER 4

SEQUENCE AND TRANSCRIPTIONAL ANALYSIS OF THE UBIQUITIN GENE CLUSTER IN THE GENOME OF *SPODOPTERA EXIGUA* NUCLEOPOLYHEDROVIRUS

SUMMARY

The nucleotide sequence of a 1200 bp DNA fragment of *Spodoptera exigua* nucleopolyhedrovirus (SeMNPV) was determined. This sequence contained a cluster of two open reading frames (ORFs), one coding for a viral ubiquitin (v-ubi) and another with homology to orf2 of *Autographa californica* (Ac) MNPV and *Orgyia pseudotsugata* (Op) MNPV. The v-ubi ORF is 240 nucleotides (nt) long, potentially encoding a protein of 80 amino acids with a predicted molecular mass of 9.4 kDa. The amino acid sequence of the v-ubi gene in SeMNPV has 75% and 81.6% identity with the v-ubi gene of AcMNPV and OpMNPV and approximately 84% with cellular ubiquitins. Northern blot analysis revealed three major small transcripts late in infection, of about 690, 550 and 400 nt long. Primer extension analysis showed that transcription started from within two consensus late promoter elements (TAAG), located at positions -6 and -30. The start site at position -4/-5 precedes the shortest leader reported to date for a baculovirus gene. The other ORF, xb187, was identified in the opposite orientation immediately upstream of the v-ubi gene. This ORF potentially encodes a 22 kDa protein with unknown function and has about 60% amino acid similarity to the products of the orf2 genes of AcMNPV and OpMNPV. The SeMNPV xb187 ORF is transcribed late in infection via two transcripts, 1.2 kb and 770 nt long. The v-ubi-xb187 gene cluster is located at map unit (m.u.) 89 on the genome of SeMNPV. This is different from the position of an identical cluster in the AcMNPV and OpMNPV genomes, located at relative m.u. 20.

This chapter has been published as: Strien, E.A. van, Jansen, B.J.H., Mans, R.M.W., Zuidema, D., & Vlak, J.M. 1996. Sequence and transcriptional analysis of the ubiquitin gene cluster in the genome of Spodoptera exigua nucleopolyhedrovirus. Journal of General Virology 77, 2311-2319.

INTRODUCTION

Ubiquitins are small proteins, abundantly present in eukaryotic cells and thought to be involved in an array of basic cellular processes, such as cell cycle control, stress response, ribosome biogenesis and cell differentiation (see Finley & Chau, 1991; Ciechanover & Schwartz, 1994 for review). Eukaryotic ubiquitins are 76 amino acids long and highly conserved (96% identity), suggesting an essential function in the life cycle of the cell. Cellular ubiquitins occur free or covalently linked to other proteins and are processed via the ubiquitin pathway. In this pathway ubiquitin monomers are covalently linked by their C-terminal Gly-76 to lysines residues of target proteins and to internal lysines of other ubiquitin molecules (Lys-48) to form multi-ubiquitin chains. Recently, it has been shown that Lys-63 and Lys-29 of ubiquitin can also be used as acceptor for ubiquitination (Arnason & Ellison, 1994; Johnson *et al.*, 1995). Ubiquitination has been implicated strongly in protein degradation (Doherty & Mayer, 1992).

Ubiquitins also play a role in the life cycle of a wide range of viruses. Host ubiqutins have been found covalently linked to some plant viruses (Dunigan *et al.*, 1988) and free ubiquitins were observed in avian leukosis virus particles (Putterman *et al.*, 1990). Herpes simplex virus 1 encodes a protein (ICP4) that induces polyubiquitin gene expression in the host (Kemp & Latchman, 1988). African swine fever virus contains ubiquinated virion proteins (Hingamp *et al.*, 1995) and codes for a protein that shows homology to enzymes involved in the ubiquitin pathway (Rodriguez *et al.*, 1992). Involvement of the ubiquitin pathway in the propagation of viruses has been suggested (Driscoll & Finley, 1992).

Ubiquitin-like genes have been found in the baculoviruses *Autographa californica* nucleopolyhedrovirus (AcMNPV) (Guarino, 1990) and *Orgyia pseudotsugata* (Op) MNPV (Russel & Rohrmann, 1993), showing 75% (AcMNPV) and 80% (OpMNPV) amino acid identity with eukaryotic ubiquitins. Failed attempts to delete v-ubi from AcMNPV suggest that ubiquitin is essential in the baculovirus life cycle (Guarino, 1990). Recently, Reilly and Guarino (1996) showed by generating a frame-shift mutation in the AcMNPV ubiquitin gene that v-ubi is not essential for virus replication in cell culture. However, production of budded virus by this mutant is much reduced. In AcMNPV both v-ubi as well as host ubiquitin are attached to the inner surface of budded virion membranes by a novel type of phospholipid anchor, referred to as phosphatidyl-ubiquitin (Guarino *et al.*, 1995). In this paper we report the presence of a ubiquitin homologue in the *Spodoptera exigua* (Se) MNPV genome.

SeMNPV is a member of the family Baculoviridae (Murphy *et al.*, 1995). In contrast to AcMNPV, SeMNPV is resticted to a single host, the beet army worm *S. exigua* (Lepidoptera: Noctuidae). SeMNPV is being applied as a biological control agent against this pest insect (Gelernter & Federici, 1986; Smits & Vlak, 1994). SeMNPV is distantly related to AcMNPV, the baculovirus type species, based on comparisons of polyhedrin (van Strien *et al.*, 1992) and p10 gene sequences (Zuidema *et al.*, 1993). The availability of the complete nucleotide sequence of the AcMNPV genome (Ayres *et al.*, 1994) allows comparison of gene sequences and gene organization among baculoviruses.

In this paper, we report the nucleotide sequence and transcriptional analysis of the v-ubi gene of SeMNPV and a flanking ORF (xb187). The amino acid sequence of v-ubi was compared to those of AcMNPV and OpMNPV and to several eukaryotic ubiquitins. In addition, we compared ORF xb187 with the putative homologous orf2 of AcMNPV (Guarino and Smith, 1990)

and OpMNPV (Russell & Rohrmann, 1993). Finally, we matched the relative location of the v-ubi and xb187 ORFs of SeMNPV with those of AcMNPV and OpMNPV.

MATERIALS & METHODS

Virus, insects and cells. The SeMNPV field isolate (SeMNPV/US) (Gelernter & Federici, 1986) was kindly provided as polyhedra by B.A. Federici (Department of Entomology, University of California, Riverside, U.S.A.). The polyhedra were propagated in fourth instar *S. exigua* larvae (Smits *et al.*, 1988). Extracellular virus (ECV), used in time course infection experiments, was obtained from the supernatant of IZD-Se-2109 cells infected with hemolymph obtained from SeMNPV-infected fourth instar larvae. The *S. exigua* cell line (obtained from B. Moeckel, Institute of Zoology, Technical University, Darmstadt, Germany) was maintained in plastic tissue culture flasks in TNM-FH medium (Hink, 1970) supplemented with 10% fetal bovine serum. ECV titers were determined by the end point dilution method (Vlak, 1979) and expressed as $TCID_{50}$ units per ml. Viral DNA was extracted from occlusion derived virions (ODVs) purified on sucrose gradients essentially as described by Caballero *et al.*, (1992).

Localization and DNA sequencing of the ubiquitin gene region. Colony filter hybridization of an SeMNPV *Eco*RI genomic library in pBluescriptKS+ (Stratagene), with an [α-^{32}P]dATP labelled probe was performed essentially as described by Sambrook *et al.* (1989). Subfragments of a selected fragment were isolated from agarose gels using the freeze-squeeze method (Sambrook *et al.*, 1989) and subcloned into the plasmids pTZ19R (Promega) or pBluescript KS(+) (Stratagene). Sequencing of SeMNPV inserts and cosmids with standard sequencing primers and custom designed primers (Eurogentec) was performed with *Taq* polymerase, using the chain termination method of Sanger *et al.* (1977) and an automatic sequencer (Applied Biosystems). Sequences were analyzed with the UWGCG computer programs (Devereux *et al.*, 1984) and DNA and deduced amino acid sequences were compared with the updated GenBank/EMBL, SWISSPROT and PIR databases using the BLAST and FASTA programs.

Isolation of total RNA and Northern blot hybridization. Total RNA for Northern blot analysis and primer extension was isolated from SeMNPV-infected IZD-Se-2109 cells at several time points post infection (p.i.), as described by Xie & Rothblum (1991). Total RNA was denatured, electrophoresed in agarose gels and blotted onto Hybond-N nylon membrane (Van Strien *et al.* 1992). To identify v-ubi and ORF xb187 transcripts, the blots were hybridized for 16 h at 65 °C with [α-^{32}P]CTP-labelled riboprobes. Riboprobes were generated by *in vitro* transcription (Sambrook *et al.* 1989) by using T7 or T3 RNA polymerase (Gibco-BRL). Templates used for *in vitro* transcription were generated by cloning of DNA fragments containing either v-ubi or ORF xb187 sequences into pBluescript KS(+) (Stratagene). After hybridization, the filters were washed for 5 min with 2 x SSC, 0.5% SDS at room temperature, 30 min with 2 x SSC, 0.1% SDS at 65 °C and 30 min with 0.1 x SSC, 0.1% SDS at 65 °C. The filters were exposed to Kodak XAR film.

Primer extension. To identify the transcriptional start site(s) of SeMNPV v-ubi, 15 ng of an oligonucleotide (5' GCGCGTATCCTCGAGC 3'), complementary to the v-ubi mRNA, was

labelled at the 5' end with [γ-^{32}P]dATP by using T4 polynucleotide kinase (Gibco-BRL) in 50 mM Tris-HCl, pH 9.5, 10 mM $MgCl_2$, 5 mM DTT, 5% glycerol for 45 min at 37 °C followed by heat denaturation at 90 °C for 10 min. The labelled oligonucleotide was purified on a 1 ml Sephadex G-25 column. Labelled primer was added to 3 µg of total infected-cell RNA and the mixture was denatured at 90 °C for 5 min and annealed at 54 °C for 15 min. Reverse transcription was carried out at 48 °C for 1 h in a volume of 15 µl, containing 5 mM of each of the dNTPs and 1 µl Superscript reverse transcriptase (Gibco-BRL) in a buffer supplied by the manufacturer. The reaction was stopped by addition of 5 µl 'stop' buffer containing 95% (v/v) formamide, 0.01% xylene cyanol and 0.01% bromophenol blue. Six µl of the reaction mixture was analyzed in a 6% polyacrylamide sequence gel, followed by drying and autoradiography.

RESULTS

Location and sequence of the ubiquitin region on the SeMNPV genome

Upon characterization of the region upstream of the p10 gene of SeMNPV (Zuidema *et al.*, 1993), a genomic *Eco*RI bank was screened with a probe, an *Eco*RI-*Xba*I fragment derived from the left-hand end of *Xba*I-H. This resulted in the detection of a 2.2 kb *Eco*RI fragment overlapping the *Xba*I site between the fragments *Xba*I-B and *Xba*I-H (Fig. 4.1). Sequences upstream of the 2.2 kb *Eco*RI fragment were obtained from cosmid COS22, encompassing fragments *Xba*I-B, -H, -D and -R (J.G.M. Heldens *et al.*, unpublished). A segment of approximately 1200 nucleotides (nt), located around map unit (m.u.) 89 on the right end of *Xba*I-B, was sequenced in both directions and contained two putative ORFs (Fig. 4.2). An ORF of 240 nt, potentially encoding a protein of 80 amino acids with a predicted molecular mass of 9.4 kDa, showed homology to ubiquitin genes when screened against databases. This gene will be further referred to as the SeMNPV v-ubi gene. Another ORF was found upstream of the v-ubi gene in the opposite orientation, with the potential to encode a protein of 187 amino acids and a predicted molecular mass of 22 kDa. This gene will be further referred to as ORF xb187.

The putative translational start codons of v-ubi and xb187 are separated by an intergenic region of 90 nucleotides. In this region, there are two late transcriptional consensus start sites (TAAG) (Blissard & Rohrmann, 1990) at nucleotide residues -6 and -30 with respect to translational start codon of v-ubi (Fig. 4.2). A potential late promoter TAAG sequence for xb187 ORF is located at position -18 relative to the translational start site. A canonical poly(A) signal (AATAAA) is located 258 nt downstream of the translational stop codon of the v-ubi gene of SeMNPV. In ORF xb187 a poly(A) motif constitutes part of the last codon and all of the stop signal (Fig. 4.2). An additional or alternative poly(A) signal is located 431 nt downstream of the stop codon of ORF xb187 (data not shown).

In the intergenic region a repeat of seven ACG triplets, at -28 and -42 (Fig. 4.2) relative to the translational start sites of ORF v-ubi and xb187, respectively, was observed. This repeat was located outside the putative transcriptional start sites of both genes. A similar motif has also been found in the 5' non-translated region of the SeMNPV p10 gene (Zuidema *et al.*, 1993), but its function is unknown.

Transcriptional analysis of the ubiquitin gene region

Transcriptional activity of the SeMNPV v-ubi gene in insect cells was ascertained by Northern

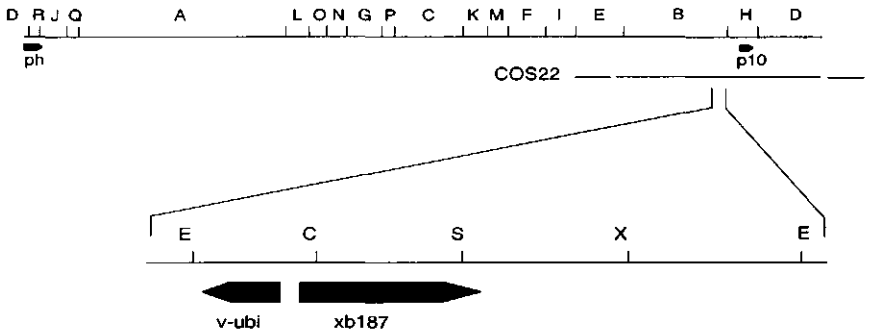


Figure 4.1. Location of the ubiquitin gene region on the genomic map of the SeMNPV genome. The *Xba*I restriction map of the SeMNPV genome is shown (top). The position of the cosmid (COS22) and the locations and orientation of the polyhedrin, p10, ubiquitin and xb187 genes are indicated (arrows). Detail physical map of the ubiquitin region (bottom). E, *Eco*RI; X, *Xba*I; C, *Cla*I; S, *Sal*I.

blot analysis of RNA isolated at various times after infection by using a strand-specific probe of the v-ubi gene. Three major transcripts, of approximately 400, 550 and 690 nt, were observed late in infection (Fig. 4.3a). Overexposure of the hybridized blot showed that the 690 nt transcript was first detectable at 8 h p.i., increased in intensity up to 24 h p.i. and appears to be less abundant very late in infection. The two smaller transcripts of 400 and 550 nt were present from 12 h p.i. and their concentration continued to increase up to 48 h p.i. The amount of the 550 nt transcript was considerably lower. Longer transcripts were also observed, which appear to originate further upstream of the v-ubi gene, as determined by an anti-sense riboprobe overlapping with ORF xb187 (data not shown).

RNA primer extension analysis was performed to determine the transcriptional start site(s) of the v-ubi gene (Fig. 4.3c). An oligonucleotide, complementary to nucleotides 147 to 162 with respect to the translational start site, was used (Fig. 4.2). The reverse transcription assay showed two major adjacent stops, at -4/-5 and at -28/-29 nt relative to the translational start codon (Fig. 4.3c), with a slight preference for the nucleotides at -4 and -28. This places the two transcriptional starts of SeMNPV v-ubi at the adenine residues of the consensus late promoter elements TAAG.

Nucleotide sequence analysis revealed the presence of a canonical poly(A)signal (AATAAA) (Birnstiel *et al.*, 1985) at 258 nt downstream of the translational stop codon of the SeMNPV v-ubi gene (Fig. 4.2). It is embedded in an A/T rich region. The predicted size of the v-ubi mRNA is therefore approximately 520 nt, indicating that the largest stretch of this mRNA is not translated into protein. This also suggests that the v-ubi mRNA has a tail of approximately 170 adenine residues to give a mRNA of 690 nt in size.

Northern analysis of the xb187 gene (Fig. 4.3b) shows a major transcript of approximately 770 nt and a minor transcript of 1.2 kb late in infection. Overexposure of the Northern blot revealed that both transcripts were first detectable at 6 h p.i., their intensity increased up to 24 h p.i. and decreased at 48 h p.i. This indicates that this gene is transcribed late in infection. The 770 nt transcript could start at -10 and terminate at the poly(A) signal just downstream of the TAA stop codon, assuming a poly(A) tail of about 150 adenine residues. A v-ubi antisense riboprobe was used to investigate if the longer transcript overlapped with the

v-ubi gene. Since no hybridization signal was detected, this implied that both xb187 transcripts were 5' coterminal (data not shown).

Comparison of the SeMNPV the v-ubi and xb187 genes

The amino acid sequence encoded by the SeMNPV v-ubi ORF was compared to those of AcMNPV (Guarino, 1990) and OpMNPV (Russell & Rohrmann, 1993), as well as to the ubiquitin monomer sequences of *Drosophila melanogaster* (Arribas *et al.*, 1986) and *S. frugiperda*

```
ATTAGTATTGTAATTTTTTTTATTGTAAAAATAAATGTATCGGACATACACACACATACA    60

ATGTTTTTCTTTATATCCCAAAAACTACTCTAGCAATATCAAGTCAATATGAGGTCCATC   120

GTGCCGAAGGTCGATCGCTTGTAATACCAGGCCGTTGTATATGAGACGACGGTGAGAATG   180
                                                        EcoRI.
AGCCGGCGGCGTAATCAATACCGCCGTACAATCTTTAAACTCTCCATCTAATACGGAATT   240

CCGTTGAGGCGTGTTTGAGTAATCATGATCATCACCACCACCATCTACCTCCGCGTAATC   300
                                    *  W  W  W  R  G  G  R  L

TCAACACCAAGTGGATTGTAGACTCTTTTTGAATGTTGTAATCGCTCATAGTGCGCGTAT   360
R  L  V  L  H  I  T  S  E  K  Q  I  N  Y  D  S  M  T  R  T

CCTCGAGCTGTTTTCCAGCGTAGATGAGACGCTGCTGATCGGGAGGTATGCCCTCCTTGT   420
D  E  L  Q  K  G  A  Y  I  L  R  Q  Q  D  P  P  I  G  E  K

CTGTAATTTTTTGTTTAACCTGTTCGACGGTGTCCGTGCTCTCGACCTCGACGGTCACCG   480
D  T  I  K  Q  K  V  Q  E  V  T  D  T  S  E  V  E  V  T  V

TTTTGCCGGTCAGAGTTTTTACAAAAATCTGCATGGCTTATTCTATTACTCGTTTACTCT   540
T  K  G  T  L  T  K  V  F  I  Q  M  <  v-ubi

CTTAAAAAAAGAAATCGTCGTCGTCGTCGTCGTATCGTCTATTAAGACGCTAGCGCA     600
                                                    xb187 >.
AAAATGTCAGCACAACAATTCACGGCCGATCTACAGGCTCTCATCGATCGGGTCAAGCGG   660
   M  S  A  Q  Q  F  T  A  D  L  Q  A  I  D  R  V  K  R  R

CGCCAAAATCCCGACGACAACGACGACGCCGCCGCCGTTAAACAACCATCGCGGTTGGGT   720
Q  N  P  D  D  N  D  D  A  A  A  V  K  Q  P  S  R  G  L  G

GACGTGATACAGCACCTCGGTCGCAACGGTCTGCTTTTGCAACGCAAAAAGGATGAAAAT   780
D  V  I  Q  H  L  G  R  N  G  L  L  L  Q  R  K  K  D  E  N

TTCGACATCAACGAAACCATAGACATTTCGGACGTGGCACGCGACTATCTCAATCTTTTG   840
F  D  I  N  E  T  I  D  I  S  D  V  A  R  D  Y  L  N  L  L

CAAACGGAAAAGTTGAGTTCGTGCCGATTGTGCTATCACAACGACGAAACGCTACGCTGC   900
Q  T  E  K  L  S  S  C  R  L  C  Y  H  N  D  E  T  L  R  C

GAGTTTCACAAAAAGTACATTTTCAACAAAAACCCTAAAGATCACTACGACGAATATGTG   960
E  F  H  K  K  Y  I  F  N  K  N  P  K  D  H  Y  D  E  Y  V

AATTTTCTCAACAGCGAGATGGGAGTAATCAGTTTCGTCGAACTCTATTACACGTATCTG  1020
N  F  L  N  S  E  M  G  V  I  S  F  V  E  L  Y  Y  T  Y  L

GGAGTGTCGTCTTGGAGAATTGTTTCGTTGATGATGATGCGCGACCTCACCGGATTCTCG  1080
G  V  S  S  W  R  I  V  S  L  M  M  M  R  D  L  T  G  F  S
                                                  SalI
TCCATTCGCGAACTATTGACGTACTACAACTACGAATGTAGCGACGACGTCGACACCGTA  1140
S  I  R  E  L  L  T  Y  Y  N  Y  E  C  S  D  D  V  D  T  V

CCTTACGAAACTATGGATTGTGAATAAATCATATG  1175
P  Y  E  T  M  D  C  E  *
```

Figure 4.2. Nucleotide sequence of an 1175 bp region of the SeMNPV genome containing the ubiquitin and xb197 genes. The ubiquitin coding sequence starts at nt 513 and terminates at nt 273; the xb187 gene starts at nt 604 and terminates at nt 1164. Late transcription initiation signals (TAAG) and putative transcription polyadenylation signals (AATAAA) are underlined.

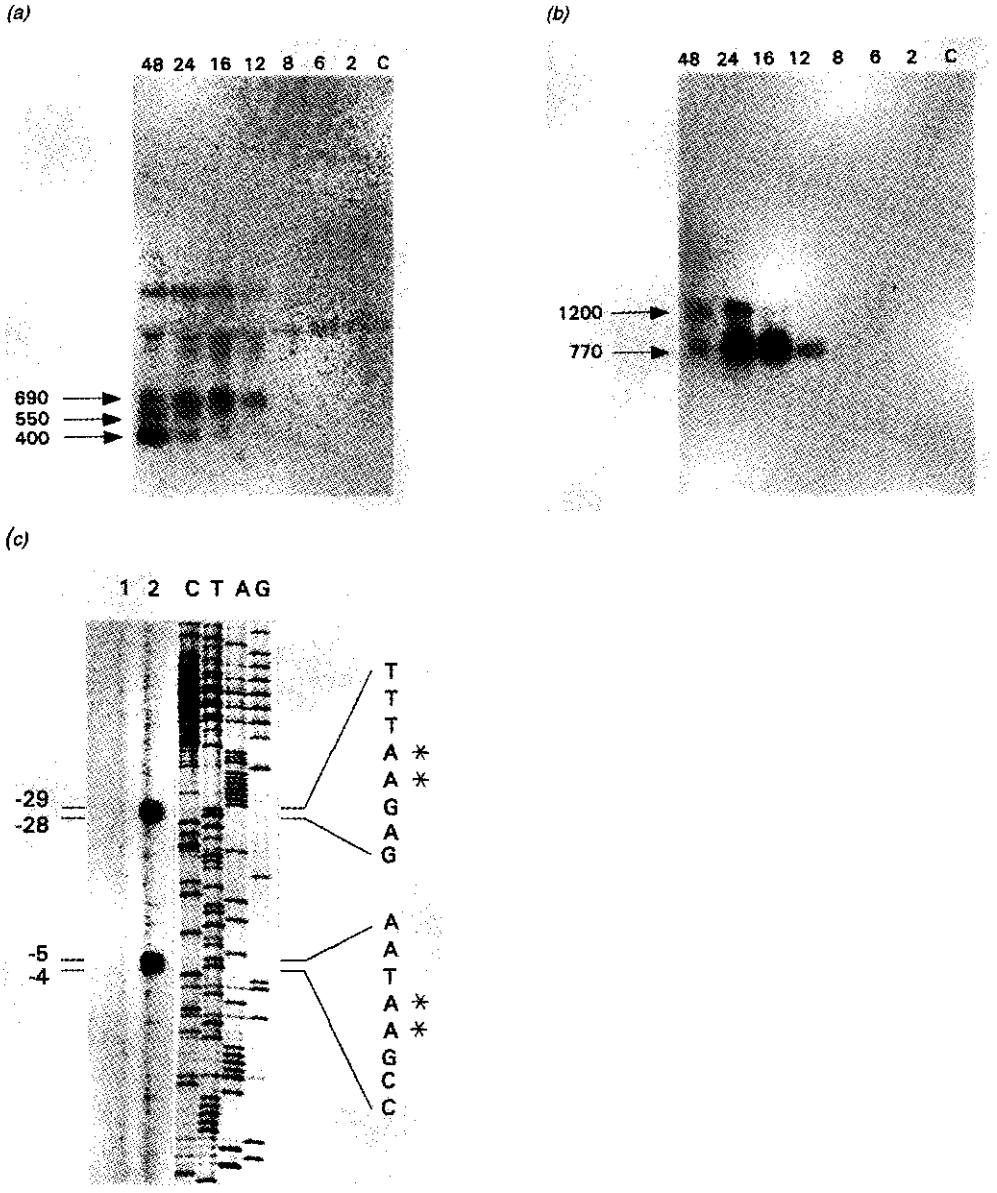


Figure 4.3. Transcriptional analysis of the SeMNPV v-ubi/xb187 gene region. Northern blot analysis of total RNA extracted from uninfected (lane C) and SeMNPV-infected IZD-Se2109 cells 2, 4, 6, 8, 12, 16, 24 and 48 h p.i. for v-ubi (a) and xb187 (b) transcripts. (c) Primer extension analysis of v-ubi transcripts performed with a 16-mer oligonucleotide complementary to the v-ubi RNA, ^{32}P-labelled at the 5 end. The oligonucleotide was annealed to total RNA from uninfected (lane 1) and SeMNPV-infected (lane 2) cells isolated 48 h p.i. and elongated by reverse transcription. The sizes of the extension products were determined by comparison with a sequence ladder run alongside (lanes *C*, *T*, *A* and *G*) obtained from an SeMNPV v-ubi containing plasmid clone and the 16-mer oligonucleotide as a sequence primer. Asterisks indicate the position of the 5'-terminal nucleotides.

(Guarino, 1990) (Fig. 4.4a). The baculoviral ubiquitins, including SeMNPV v-ubi, differ from the cellular ubiquitins in their C-terminal extension by one or more amino acids. Processed cellular ubiquitins have a fixed length of 76 amino acids. Therefore, only the first 76 amino acids of SeMNPV v-ubi were considered in calculating the extent of homology. The homology (identity) of SeMNPV ubiquitin with AcMNPV and OpMNPV ubiquitin is 75.0% and 81.6%, respectively, whereas the homology with *D. melanogaster* and *S. frugiperda* ubiquitin is 84.2% and 82.9%, respectively (Table 4.1). Notably, the ubiquitin protein of *Bombyx mori* (BmNPV) NPV is 100% identical to its AcMNPV homologue (complete BmNPV sequence is in Genbank/EMBL, accession no. L33180).

The residues that are known to be involved in the formation of ubiquitin-protein complexes in eukaryotes, Lys-29, Lys-48, Lys-63 and Gly-76 are all conserved (Fig. 4.4a; Johnson *et al.*, 1995; Chau *et al.*, 1989; Finley & Chau, 1991; Arnason & Ellison, 1994; Spence *et al.*, 1995). The altered amino acid residues, at positions 23 and 28, relative to the eukaryotic ubiquitins, are conserved among the baculoviral ubiquitins. The most heterogeneous parts of the SeMNPV ubiquitin are the areas from amino acid residues 15 to 31 and 53 to 57. In comparison to eukaryotic ubiquitins, SeMNPV ubiquitin has four additional amino acids at its C terminus. AcMNPV and OpMNPV ubiquitin have 1 and 18 additional amino acids at their C terminus, respectively. The C-termini of the other viral v-ubiquitins do not share homology to SeMNPV ubiquitin (Fig. 4.4a).

The xb187 gene shows homology (similarity) of 59%, 59% and 61% at the amino acid level with the orf2 genes of AcMNPV, BmNPV and OpMNPV (Fig. 4.4b). The amino acid sequence of the putative protein encoded by the xb187 gene is considerably shorter than that of the orf2 genes of AcMNPV, BmNPV and OpMNPV (187 amino acids versus 215,215 and 219, respectively), truncated by about 30 amino acids at the N terminus. The N-terminal region of the ORF2 and xb187 proteins is very heterogeneous among the three baculoviruses, whereas the C-terminal region displays a higher degree of homology. A search in the updated releases of the Genbank/EMBL database with FASTA and BLAST (Altschul *et al.*, 1990) of the GCG Software package did not show significant homologies to other genes.

DISCUSSION

In this report the nucleotide sequence and transcriptional analysis of two adjacent ORFs, v-ubi and xb187, on the *Xba*I-B fragment of the baculovirus SeMNPV is presented. The general features of the SeMNPV ubiquitin gene seem to be similar to those of other ubiquitins (Fig. 4.4 and Table 4.1). The amino acid residues involved in ubiquitination, Lys-29, Lys-48, Lys-63 and Gly-76, are conserved. The additional amino acids at the C termini of the baculoviral ubiquitins may be the remnant of the C-terminal extension of either an ancestral cellular ubiquitin-fusion gene or polyubiquitin gene. However, the residues known to be essential for proteolytic processing, Gly-Gly-X (Lopez-Otin *et al.*, 1989), are conserved and ubiquitin C-terminal hydrolase activity is present in AcMNPV-infected *S. frugiperda* cells (Moguilevksy *et al.*, 1994). Based on these findings it is likely that the C termini of the baculoviral ubiquitins are processed correctly to yield a ubiquitin of 76 amino acids.

Comparison of the predicted amino acid sequences suggest that SeMNPV ubiquitin is more closely related to OpMNPV than to AcMNPV ubiquitin. The identity with *D. melanogaster*

(a)

```
                                                    *                   *
SeMNPV ubi   MQIFVKTLTG KTVTVEVEST DTVEQVKQKI TDKEGIPPDQ QRLIYAGKQL 50
OpMNPV ubi   .......... ..I...T.PG ...G...... A....V.V.. .......... 50
AcMNPV ubi   ....I..... ..I.A.T.PA E..ADL.... A....V.V.. ....F..... 50
Dm ubi       .......... ..I.L...PS ..I.N..A.. Q......... ....F..... 50
Sf ubi       .......... ..I.L...PS ..I.N..A.. Q......... ....F..... 50


                         *
SeMNPV ubi   EDTRTMSDYN IQKESTIHLV LRLRGGRWWW 80
OpMNPV ubi   ..AQ.LA... ......L.M. ......NGLR KGKRRCLSLL QFI 93
AcMNPV ubi   ..SK..A... ......L.M. ......Y 77
Dm ubi       ..G..L.... ......L... ......  76
Sf ubi       G.G..L.... ......L... ......  76
```

(b)

```
SeMNPV xb187  ---------- ---------- -------MSA QQFTADLQAL IDRVKRRQNP  23
OpMNPV  orf2  ------MAAL VKPMLPLATF GGQQNGCLQH .LAKLVQARA RRGYEHDIGQ  44
AcMNPV  orf2  MTTVAVNAPL PPPLVELCNR RPIPTPRIIS L.RQLISTPV VKNYQADVQE  50
BmNPV   orf2  MTTVAVNAPL PPPMVELCNR RPIPTPRIIS L.RQLISTPV .KNYQADVQE  50

SeMNPV xb187  DDNDDAAAVK QPSRLGDVIQ HLGRNGLLLQ R--KKDENFD INETIDISDV  71
OpMNPV  orf2  LAEKLKKRQV ARGH....LE QM..QSE..P ELV.N..E.R .VQQR.LGHN  94
AcMNPV  orf2  AIDAFKRLNI T.GH..E..D TM.QQ.K..P EIIEA.DD.K V.Q.RNLSCK 100
BmNPV   orf2  AIDAFKRLNI T.GH..E..D AM.QQ.R..P EIIEA.DD.K V.Q.RNLSCK 100

SeMNPV xb187  ARDYLNLLQT EKLSSCRLCY HNDETLRCEF HKKYIFNKNP KDHYDEYVNF 121
OpMNPV  orf2  TIE......H D..F...... THAHW.W... ..THAYRGPR DISV.A..DH 144
AcMNPV  orf2  TVE...F.EN D..FR..... THADW.W.D. .RNHAYRGTR DITCNN..EH 150
BmNPV   orf2  TVE...F.EN D..FR..... THADW.W.D. .RNHAYRGTR DIACNN..EH 150

SeMNPV xb187  LNSEMGVISF VELYYTYLGV SSWRIVSLMM MRDLTGFSSI RELLTYYNYE 171
OpMNPV  orf2  ...D...VAL ..E..HW.SS CNDKAEAKRA LKT.ANVE.L GD..DS...- 193
AcMNPV  orf2  ...D...VML I.E.FYC.SS CNFKQDAKRA LQT..K.E.L SD.MAS..F- 199
BmNPV   orf2  ...D...VML ..E.FYC.SS CNFKQDAKLA LQT..N.E.L SD.MAS..F- 199

SeMNPV xb187  CSDDVDTVPY ETMDCE  187
OpMNPV  orf2  S.V.A..SS. .L..F.  209
AcMNPV  orf2  STP.L..NA. .L..F.  215
BmNPV   orf2  STP.L..NA. .L..F.  215
```

Figure 4.4. Amino acid sequence alignment of various baculoviral and cellular ubiquitins (a) and xb187 homologues (b). Dm, *D. melanogaster* (Arribas *et al.*, 1986); Sf, *S. frugiperda* (Guarino, 1990); Se, *S. exigua* MNPV; Ac, *A. californica* MNPV (Guarino, 1990); Op, *O. pseudotsugata* MNPV (Russell & Rohrmann, 1993); Bm, B.mori NPV (Genbank/EMBL, accession no. L33180). Identical amino acids are indicated by dots. Hyphens are introduced to maximize matching of amino acid residues. Amino acids involved in ubiquitination are indicated with asterisks. Numbers indicate amino acid positions.

and *S. frugiperda* ubiquitin is even greater (Table 4.1). However, since the variation among the baculoviral ubiquitins is much higher than that observed among cellular ubiquitins, these data show that baculoviral ubiquitins have diverged considerably from the highly conserved cellular ubiquitins and from each other. This may suggest that v-ubiquitins have a special function in baculovirus infection, as for example the association of ubiquitin with ECVs via a phospholipid anchor (Guarino *et al.*, 1995). With these data on v-ubi of SeMNPV, three nucleopolyhedrovirus

Table 4.1. Amino acid sequence identity (%) between the ubiquitin proteins of SeMNPV, AcMNPV, OpMNPV and *S. frugiperda* and *D. melanogaster*.

	AcNPV	OpNPV	Dm	Sf
SeNPV	75.0	81.6	84.2	82.9
AcNPV		85.5	76.3	75.0
OpNPV			80.3	78.9
Dm				98.7

ubiquitins have been sequenced. Moreover, the identification of a v-ubi gene in the genome of *Cydia pomonella* granulovirus (N.E. Crook, personal communication), suggests that this gene most likely is preserved in baculoviruses.

Analysis of SeMNPV v-ubi transcription revealed that transcripts started with equal frequency in the two canonical TAAG late promoter elements located at positions -6 and -30 relative to the translational start codon (Figs. 4.2 and 4.3c). This is unlike the situation in AcMNPV, where the two v-ubi transcripts are present in unequal amounts and have longer untranslated 5' leaders (Guarino, 1990). The 5' sequences of the baculovirus v-ubi genes differ considerably (Fig. 4.5). Only the TAAG consensus promoter sequences are conserved. The transcriptional start site of the SeMNPV v-ubi gene at position -4 is the shortest leader reported to date for a baculovirus gene and is reminiscent of the enhancin gene transcription of *Heliothis armigera* granulovirus (HaGV) (Roelvink *et al.*, 1995).

Three major v-ubi transcripts, of about 690, 550 and 400 nt in length were observed in infected cells late in infection. The size of the largest mRNA corresponds with the expected length of a polyadenylated transcript of v-ubi (Fig. 4.2), assuming a poly(A) tail of approximately 150-200 nt. The shorter transcripts of 550 and 400 nt, which appear later in infection (after 12 h p.i., Fig. 4.3a), may represent non-polyadenylated v-ubi mRNA. The 550 nt transcript might terminate near the poly(A) signal, whereas the 400 nt transcript might terminate in between the translational stop codon and the poly(A) signal, thereby precluding a poly(A) tail. There are a few previous reports on non-polyadenylated transcripts in baculoviruses. The SeMNPV polyhedrin gene transcripts are not polyadenylated (Van Strien *et al.*, 1992). Messenger RNA of the HaGV enhancin gene is also not polyadenylated and has an equally short leader (Roelvink *et al.*, 1995).

```
                                    -30                          -6    +1
SeMNPV ubi                  TTTTTTTAAGAGAGTAAACGAGTAATAGAATAAGCCATGCAGATTTTT
                                    **                           **
                                                  -13
OpMNPV ubi                          CATTTATAAG----AGATTCAAAATGCAAATATTT
                                                  -17
                  -195
AcMNPV ubi  GTTATATTAAGTCGTT--(163 nt)--CATTTATAAGTAATAGTGTAAAAATGCAAATATTC
                  *                             *
```

Figure 4.5. Alignment of the 5' noncoding leader sequence of the v-ubi genes of SeMNPV, AcMNPV and OpMNPV. Putative promoter elements are indicated in bold and transcriptional start sites are indicated by asterisks.

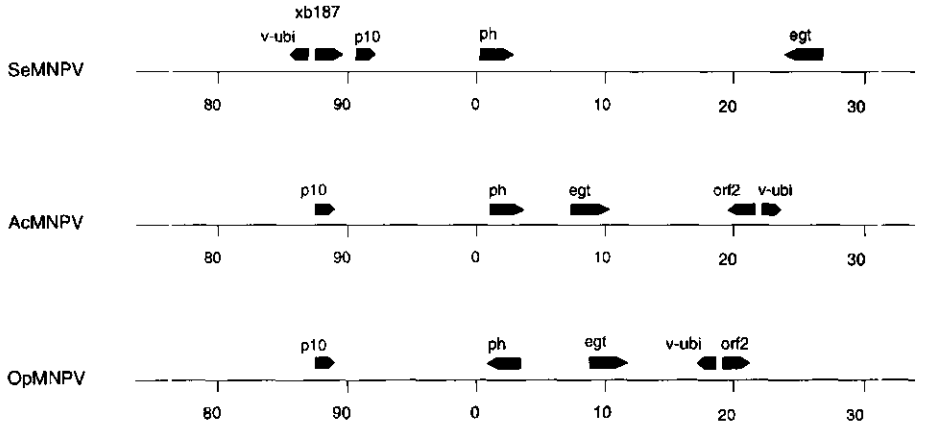


Figure 4.6. Alignment of genomes of AcMNPV, OpMNPV and SeMNPV, and location of polyhedrin, p10 and the v-ubi/xb187 gene cluster. The scale is given in map units (m.u.) of 1-100.

The xb187 gene of SeMNPV shows homology to the orf2 genes of AcMNPV (Guarino & Smith, 1990) and OpMNPV (Russell & Rohrmann, 1993) and is located next to the v-ubi gene in the opposite orientation in all three viruses. The highest homology is found at the C terminus, suggesting that functional domains are located in this part of the protein.
The xb187 gene is active as a late gene (Fig. 4.3b). The presence of two xb187 transcripts can be explained by the fact that there are two poly(A)signals, one overlapping with the 5 last nucleotides of the ORF and another 431 nt downstream of the translational stop codon (data not shown). The estimated sizes of the transcripts correspond well with the sequence data, including a presumed poly(A) tail of approximately 200 nt. The putative 5' ends of the mRNAs encoding v-ubi and xb187 do not overlap in SeMNPV and the intergenic region is much larger than in the genomes of AcMNPV and OpMNPV. In SeMNPV this region contains an ACG repeat of unknown function.

In the baculoviruses analyzed so far, the v-ubi and xb187 genes or its homologues are clustered, but located at different postitions in the viral genome. In SeMNPV the v-ubi/xb187 gene cluster is found around m.u. 89, whereas the same cluster in AcMNPV and OpMNPV is located around m.u. 22 and m.u. 19, respectively (Fig. 4.6). This supports the idea that the baculovirus genome organization has been rearranged during its evolutionary history (Gombart *et al.*, 1989). Sequence comparisons of the SeMNPV v-ubi/xb187 gene cluster with AcMNPV and OpMNPV, added to comparisons based on polyhedrin gene (Van Strien *et al.*, 1992; Zanotto *et al.*, 1993) and p10 gene (Zuidema *et al.*, 1993) sequences, confirmed the distant relationship of SeMNPV to these two viruses.

ACKNOWLEDGEMENTS

This work was financially supported in part by the Commission of the European Communities. We thank Magda Usmany and Els Klinge-Roode for technical support and Basil Arif for critical reviewing of the manuscript.

CHAPTER 5

UNUSUAL CHARACTERISTICS OF THE TRANSACTIVATOR GENE *ie1* OF *SPODOPTERA EXIGUA* NUCLEOPOLYHEDROVIRUS

SUMMARY

The genomic position and nucleotide sequence of the immediate early gene *ie1* of *Spodoptera exigua* nucleopolyhedrovirus (SeMNPV) were determined. The SeMNPV *ie1* gene had the potential to encode a protein of 714 amino acids with a predicted molecular mass of 82.0 kDa, representing the largest baculovirus IE1 known to date. The homology of SeMNPV IE1 with IE1 proteins from other baculoviruses was restricted to the basic C-terminal two-third of the protein, which is involved in DNA binding. Transcriptional analysis of the SeMNPV *ie1* gene by Northern blot hybridization showed the transient expression of a single transcript of 2.5 kb starting at 1 h p. i., peaking at 8 h p. i. and disappearing at 24 h p. i. From 8 h p.i. onwards a 1.7 kb transcript was detected which became more abundant at the late stage of infection. Primer extension analysis revealed the use of several start sites late in infection. One of these was located in the promoter motif CAGT located at position -15 relative to the translational start codon. This start site was used most intensively at 8 h p.i. A longer extension product, corresponding to a presumed non-translated 5' leader of 138 nt, was observed at late stages of infection. The different temporally expression of the *ie1* transcripts observed by Northern blotting and by primer extension analysis is as yet unexplained. Four kb upstream of the SeMNPV *ie1* gene, an open reading frame (SeMNPV ORF xd244) was identified, with homology to the left *ie0* exon of AcMNPV and OpMNPV. However, SeMNPV ORF xd244 did not contain the complete consensus splice donor sequences, which may explain the inability to identify a spliced *ie0* transcript early in SeMNPV infection. Phylogenetic analysis based on the baculovirus IE1 proteins indicated that SeMNPV shared the most recent ancestor with HzSNPV, although their IE1 proteins had diverged considerably from each other as well from those of other baculoviruses.

Manuscript in preparation: Strien, E.A. van, Gerrits, H., Zuidema, D., & Vlak, J.M. 1997. Unusual characteristics of the transactivator gene *ie1* of *Spodoptera exigua* nucleopolyhedrovirus.

INTRODUCTION

Baculoviruses are large, circular dsDNA viruses, replicating in the nucleus of infected cells of their arthropod hosts. Gene expression of baculoviruses is transcriptionally regulated and divided into several phases. Whereas late genes are not expressed before the onset of viral DNA replication, the expression of early genes starts before this event (Friesen & Miller, 1986, Blissard & Rohrmann, 1990). Of these, the immediate early genes utilize the host RNA polymerase II complex and do not require the presence of viral gene products for their transcription. They are often involved in the transactivation of delayed early and late gene expression (Friesen & Miller, 1986, Blissard & Rohrmann, 1990).

One of the most important baculovirus transactivator genes is the immediate early gene *ie1*, which encodes a multifuntional protein. The *ie1* genes of *Autographa californica* nucleopolyhedrovirus (AcMNPV) (Guarino & Summers, 1987), the baculovirus type species (Murphy *et al.*, 1995), of the closely related *Bombyx mori* (Bm) NPV (Huybrechts *et al.*, 1994) and of *Orgyia pseudotsugata* (Op) MNPV (Theilmann & Stewart, 1991) have been investigated in some detail. IE1 strongly transactivates the transcription of early genes, including its own transcription (Carson *et al.*, 1988, Guarino & Summers, 1986a). Ttransactivation by IE1 occurs in association with the so-called *hr* (homologous region) sequences (Choi & Guarino, 1995b), which are found dispersed around the viral genome (Cochran & Faulkner, 1983, Guarino *et al.*, 1986) and function as transcriptional enhancers (Guarino & Summers, 1986b, Kovacs *et al.*, 1991a, Theilmann & Stewart, 1993, Choi & Guarino, 1995a).

IE1 is also essential for viral DNA replication (Kool *et al.*, 1993, Kool *et al.*, 1994) and for late gene expression (Passarelli & Miller, 1993). IE1 has DNA binding properties as observed in gel retardation assays (Guarino & Dong, 1991) and is multiply phosphorylated (Theilmann & Stewart, 1993, Choi & Guarino, 1995b). The DNA binding properties of AcMNPV IE1 reside in the C-terminal part of the protein (Kovacs *et al.*, 1992). Furthermore, IE1 was associated with the budded virions (BVs) of OpMNPV (Theilmann & Stewart, 1993).

The coding sequences of AcMNPV and OpMNPV *ie1* are expressed in two ways, via an unspliced (*ie1*) and via a spliced transcript, *ie0*. In the latter case an exon, derived from a gene (AcMNPV ORF 141) located 4 kbp upstream of *ie1*, is spliced onto the 5' end of the *ie1* transcript. Compared with IE1, the *ie0* translation product (IE0) has several additional amino acids at the N terminus of IE1 (Chisholm & Henner, 1988, Kovacs *et al.*, 1991b, D.A. Theilmann, *pers.comm.*). The AcMNPV *ie0* and *ie1* translation products had distinct transactivation properties (Kovacs *et al.*, 1991a).

The baculovirus *Spodoptera exigua* nucleopolyhedrovirus (SeMNPV) is restricted to a single host, the beet army worm *S. exigua*. Several late genes of SeMNPV have been characterized, such as polyhedrin (van Strien *et al.*, 1992, Chapter 2), p10 (Zuidema *et al.*, 1993, Chapter 3) and v-ubi (van Strien *et al.*, 1996, Chapter 4). Comparison of these genes with their homologs in other baculoviruses showed that SeMNPV is distantly related to AcMNPV, BmNPV and OpMNPV. In view of the important role of IE1 in the baculovirus life cycle, the SeMNPV *ie1* gene has been mapped and analyzed. Its transcription pattern was different from that of other baculoviruses and did not involve splicing.

MATERIALS AND METHODS

Virus, insects and cells. The SeMNPV field isolate (SeMNPV/US) (Gelernter & Federici, 1986) was kindly provided as polyhedra by B.A. Federici (Department of Entomology, University of California, Riverside, USA). The virus was propagated in fourth instar *S. exigua* larvae (Smits *et al.*, 1988). Budded virus (BV), used in time course infection experiments, was obtained from the supernatant of IZD-Se-2109 cells (B. Moeckel, *pers.comm.*) and infected with haemolymph isolated from SeMNPV-infected fourth instar larvae. The *S.exigua* cell line IZD-Se-2109 was maintained at 27°C in plastic tissue culture flasks in TNM-FH medium (Hink, 1970) supplemented with 10% foetal bovine serum. BV titres were determined by the end point dilution method and expressed as $TCID_{50}$ units per ml. Viral DNA was extracted from SeMNPV occlusion body-derived virions (ODVs) purified on sucrose gradients essentially as described by Caballero *et al.* (1992).

Localization and DNA sequencing of the SeMNPV ie1 gene. A *Cla*I-*Xba*I DNA fragment (2.65 kb) of the AcMNPV *ie1* gene (Guarino & Summers, 1987) was [α-^{32}P]dATP labelled by nick-translation (Life Technologies Inc.) and hybridized under non-stringent conditions (Sambrook *et al.*, 1989) to a Southern blot of SeMNPV viral DNA on Hybond-N filters (Amersham). Hybridizing SeMNPV DNA fragments were purified from agarose gels by the freeze-squeeze method, (sub)cloned into pUC19 and sequenced according to the chain termination method with standard sequencing primers and custom designed oligonucleotide primers (Eurogentec) using the chain termination method (Sanger *et al.*, 1977).

Sequence analysis of the SeMNPV ie1 gene and ORF xd244. Sequences were analyzed with the UWGCG computer programs (Devereux *et al.*, 1984) and DNA and deduced amino acid sequences were compared with the updated GenBank/EMBL, SWISSPROT and PIR databases using the FASTA program and BLAST network service (Altschul *et al.*, 1990).
Phylogenetic analysis was performed with the PAUP 3.1 program (Swofford, 1993), using GCG pileup to produce input files of aligned protein sequences. Bootstrap analysis (Felsenstein, 1993), included in the PAUP package, was used to assess the integrity of the produced phylogeny.

Isolation of polyA$^+$ RNA and Northern blot hybridization. For Northern blot analysis polyA$^+$ RNA was prepared by oligo-dT column chromatography purification, according to the instructions of the manufacturer (Pharmacia), from SeMNPV-infected IZD-Se-2109 cells, harvested at various times post infection. An equivalent of polyA$^+$ RNA from 2 x 10^6 cells was denatured, electrophoresed in agarose gels and blotted onto Hybond-N nylon membrane (van Strien *et al.*, 1992). To identify *ie1* transcripts, the blot was hybridized for 16 h at 65 °C with an *ie1*-specific probe. An internal *Sal*I fragment of 530 nt (Fig. 5.1) was used as template to generate a single-stranded [α-^{32}P]dATP-labelled DNA probe using Klenow DNA polymerase (Life Technologies Inc.) and a standard sequencing primer. After hybridization, the filter was washed for 5 min with 2 x SSC, 0.5% SDS at room temperature, 30 min with 2 x SSC, 0.1% SDS at 65 °C and 30 min with 0.1 x SSC, 0.1% SDS at 65 °C. The filter was exposed to Kodak XAR film.

Isolation of total RNA and primer extension. Primer extension was performed with total RNA, isolated from SeMNPV-infected IZD-Se-2109 cells at various times post infection as described previously (van Strien *et al.*, 1992). To identify the transcriptional start site(s) of SeMNPV *ie1*, 15 ng of oligonucleotide A (5' GCGAAGCGGCGTGCTACCGGTGG 3'), or of oligonucleotide B (5' CTTCATGGCGCTCGAAACGCAC 3'), both complementary to the *ie1* mRNA, was labelled at the 5' end with [γ -^{32}P]dATP by using T4 polynucleotide kinase (Life Technologies Inc.) in 50 mM Tris/HCl, pH 9.5, 10 mM $MgCl_2$, 5 mM DTT, 5% glycerol for 45 min at 37 °C followed by heat denaturation at 90 °C for 10 min. The labelled oligonucleotide was purified on a 1 ml Sephadex G25 column. Labelled primer was added to 3 µg of total RNA and the mixture was denatured at 90 °C for 5 min and annealed at 54 °C for 15 min. Reverse transcription was carried out at 48 °C for 1 h in a volume of 15 µl, containing 5 mM of each of the dNTPs and 1 µl Superscript reverse transcriptase (Life Technologies Inc.) in a buffer supplied by the manufacturer. The reaction was stopped by addition of 5 µl 'stop' buffer containing 95% (v/v) formamide, 0.01% xylene cyanol and 0.01% bromophenol blue. Six µl of the reaction mixture was analyzed on a 6% denaturing polyacrylamide sequence gel, followed by drying and autoradiography.

RESULTS

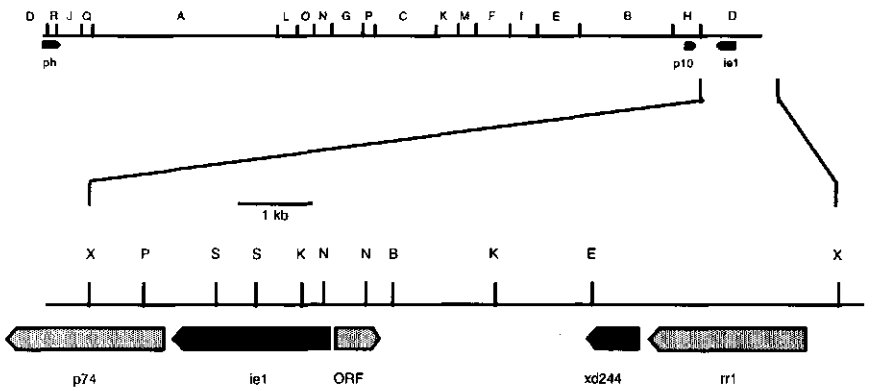
Location and sequence of the SeMNPV ie1 gene

In order to identify and localize the SeMNPV *ie1* gene on the genome, a 2.65 kb DNA fragment containing the AcMNPV *ie1* gene (Guarino & Summers, 1987) was used as a probe. Under low stringency hybridization and washing conditions the AcMNPV probe hybrizided to the SeMNPV XbaI-D fragment. Hybridizations with XbaI-D subclones showed that the smallest hybridizing fragment was a NruI-NruI fragment of 600 bp (data not shown). The SeMNPV XbaI-D fragment was (sub)cloned and partially sequenced (Fig. 5.1a, b).

An open reading frame (ORF) of 2145 nt was observed, which showed homology to the *ie1* genes from other baculoviruses (see below). No significant homologies to other proteins were found. The SeMNPV *ie1* gene had an anti-genomic orientation (Fig. 5.1a) and potentially encoded a protein of 714 amino acids (Fig. 5.1b) with a predicted molecular mass of 82.0 kDa.

A consensus baculovirus early promoter sequence, CAGT, (Blissard & Rohrmann 1990) was found 18 nt in front of the putative translational start codon. A TATA box was found 31 nt upstream of this CAGT motif, as part of the coding sequence of an ORF located upstream of SeMNPV *ie1*. This ORF displayed homology to the AcMNPV ORF 146, whose function is unknown (Ayres *et al.*, 1994). The putative translational start codons of SeMNPV *ie1* and the upstream ORF were separated by only 22 nt. The region encompassing the presumed SeMNPV *ie1* early start site, CAGT, and the preceeding TATA box, are highly conserved among the baculovirus *ie1* genes (Fig. 5.2a), but the nucleotide homology in the SeMNPV *ie1* 5' region is limited to these sequence motifs. It is possible that these sequences were responsible for the observed hybridization signal with the AcMNPV *ie1* DNA probe. Two consensus polyadenylation signals, AATAAA, were located 19 and 47 nt downstream of the translational stop codon of *ie1*. Downstream of SeMNPV *ie1*, sequences homologous to the AcMNPV p74 gene (Kuzio *et al.* 1989) were detected (Fig. 5.1). The SeMNPV *ie1* and p74 coding sequences were separated by 145 nt.

(a)



(b)

```
TATGGCGACGTTAGGAGGTGACTGCGATAGCCTAGTTTGAAATACAAAGATGCTCACAGA      60
I  A  V  N  P  P  S  Q  S  L  R  T  Q  F  V  F  I  S  V  S

GTTGATCGTGTTTAACATACTAAACGTTACAATATTAGTATTATCTTGCGCGTAGTATAA     120
N  I  T  N  L  M  S  F  T  V  I  N  T  N  D  Q  A  Y  Y  L

ATTGACGTTCATGTTTGTAATTGTTTCAGTTTGCTGAGACAATCATGCACACTCCGAGTC     180
  N  V  N  M  <  ORF                    IE1>  M  H  T  P  S  H    6
                                 primer A                 NruI
ACCACAATCAAATCAGTCCTGCCACTTACAAGACCACCGGTAGCACGCCGCTTCGCGACA     240
  H  N  Q  I  S  P  A  T  Y  K  T  T  G  S  T  P  L  R  D  S   26

GCCTCGGACCTTTTTTGAAATTTCAACATTCAATTCCGCAACCGCCACTCGACATGTTTA     300
  L  G  P  F  L  K  F  Q  H  S  I  P  Q  P  P  L  D  M  F  N   46

ACAACAACAACAACAACAATACCAATCAATTCGTCGACATTGACTACACGGACACCAATT     360
  N  N  N  N  N  N  T  N  Q  F  V  D  I  D  Y  T  D  T  N  Y   66

ACAATTTACTTATCAACAACGCCGAAATGATTAATCAAAATTACGACGCCGCCGCCGCCG     420
  N  L  L  I  N  N  A  E  M  I  N  Q  N  Y  D  A  A  A  A  A   86

CCGCCGCCGCCTCGCATCACCACGATCCCTTCAATGCTTCTTCGCCGCAAGACATTGATT     480
  A  A  A  S  H  H  H  D  P  F  N  A  S  S  P  Q  D  I  D  L  106
                                                      SalI
TGGATCTCCTATTTCCTCCTGAACAAGTGCCTTTCGTCGAAACTAAAAGCAATGTCGACG     540
  D  L  L  F  P  P  E  Q  V  P  F  V  E  T  K  S  N  V  D  A  126
     KpnI
CCGTGGTACCAACGATTAAGATTGACAAGAAAAATATGAACATCAAAGACAATAATAAAG     600
  V  V  P  T  I  K  I  D  K  K  N  M  N  I  K  D  N  N  K  E  146

AAAATGTGGCAGTCGTTTCTCCGTCTTTGACGTTGAAACATCACTCTTTGTCGATGTCGT     660
  N  V  A  V  V  S  P  S  L  T  L  K  H  H  S  L  S  M  S  S  166

CGCCCTCTTCATTTACATCGAAACCAATAATGTCGCCATTAAAAAAATCTACTATTAAAG     720
  P  S  S  F  T  S  K  P  I  M  S  P  L  K  K  S  T  I  K  E  186

AAAAGCGCGTCTCGATGAAGGGTATGGGCGCGGCGATGAAACGGATTCGCATGACCGACA     780
  K  R  V  S  M  K  G  M  G  A  A  M  K  R  I  R  M  T  D  T  206

CGTACGACGAAACGATGAGCAGCAGCAGCAGCAGCGACGACGACGACGAAAGCGGATCTA     840
  Y  D  E  T  M  S  S  S  S  S  S  D  D  D  D  E  S  G  S  S  226

GCGACGACGAAGCGAATAAGATTGAAAGAAATATTAAAAATTTTAATGACAACGACAAGA     900
  D  D  E  A  N  K  I  E  R  N  I  K  N  F  N  D  N  D  K  N  246

ATGACGATGATATCGGCGCCGTTGTCGAAGATCGAGAAAGAATTTTAATCAAACCAAAGA     960
  D  D  D  I  G  A  V  V  E  D  R  E  R  I  L  I  K  P  K  S  266

GCCGCGGCAGGTACGCCAAGAAAATGTGCGTTTCGAGCGCCATGAAGCCCGTTCACGTCG    1020
  R  G  R  Y  A  K  K  M  C  V  S  S  A  M  K  P  V  H  V  E  286

AGACACCGACTCCCAGCGATCCCGCAACCGATAGTCTTTTCCGCGAAATCATCACTAAAC    1080
  T  P  T  P  S  D  P  A  T  D  S  L  F  R  E  I  I  T  K  Q  306

AACTGCAACAGAACGACGATGTCGTCGTCGACAACAGCGGCATCTTTACCAGCCACATTC    1140
  L  Q  Q  N  D  D  V  V  V  D  N  S  G  I  F  T  S  H  I  L  326
                                                   SalI
TCGACACTAGTTACTACATGTTTATCATTAGCAAGTCGTCCAACGTCGACGAAGTGTATA    1200
  D  T  S  Y  Y  M  F  I  I  S  K  S  S  N  V  D  E  V  Y  S  346
```

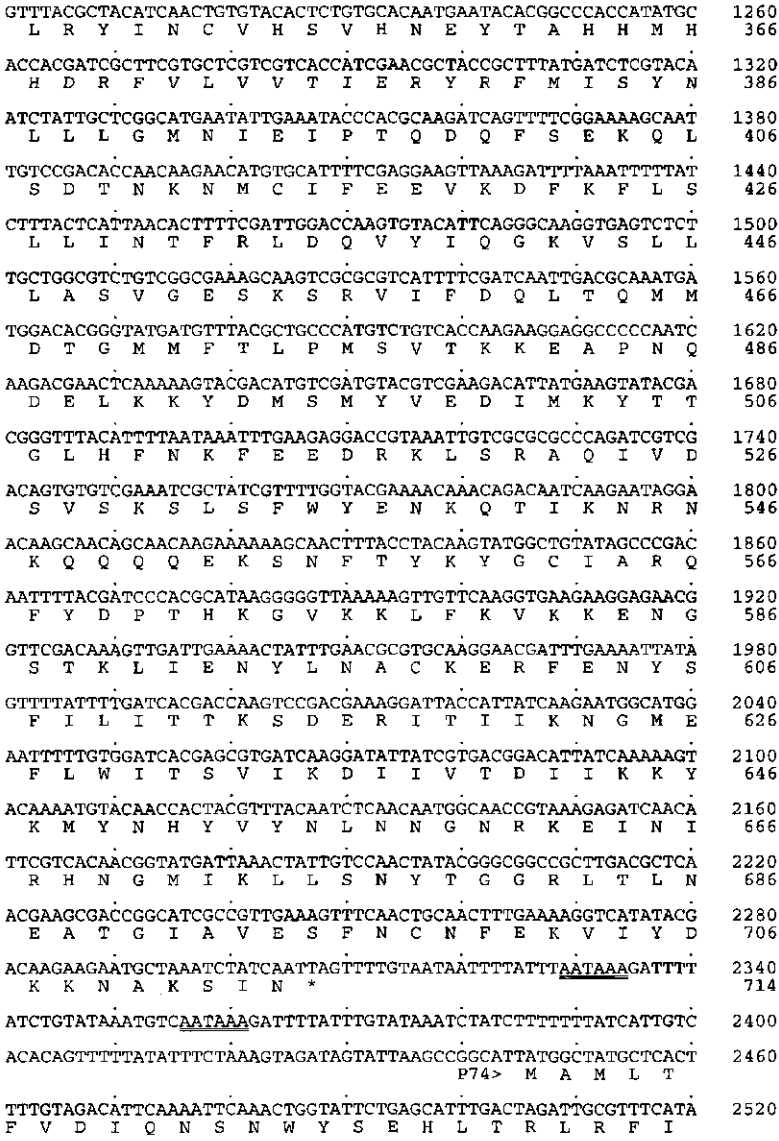


Figure 5.1. a) Location and restriction endonuclease recognition sites of the *ie1* gene region on the physical map of the SeMNPV genome (Heldens *et al.*, 1996). The location and orientation of the *ie1*, polyhedrin (ph), p74, xd200 and xd244 and p10 ORFs are indicated (arrows). E, *Eco*RI; X, *Xba*I; N, *Nru*I; Xh, *Xho*I; S, *Sal*I; B, *Bam*HI; K, *Kpn*I; P, *Pst*I. b) Nucleotide sequence and predicted amino acid translation of the SeMNPV *ie1* gene. Baculovirus consensus early and late transcriptional start sites are underlined. Putative polyadenylation signals are double underlined. Bold indicates putative transcriptional start sites calculated from the primer extension experiments. The locations of the oligonucleotide primers A and B are double underlined.

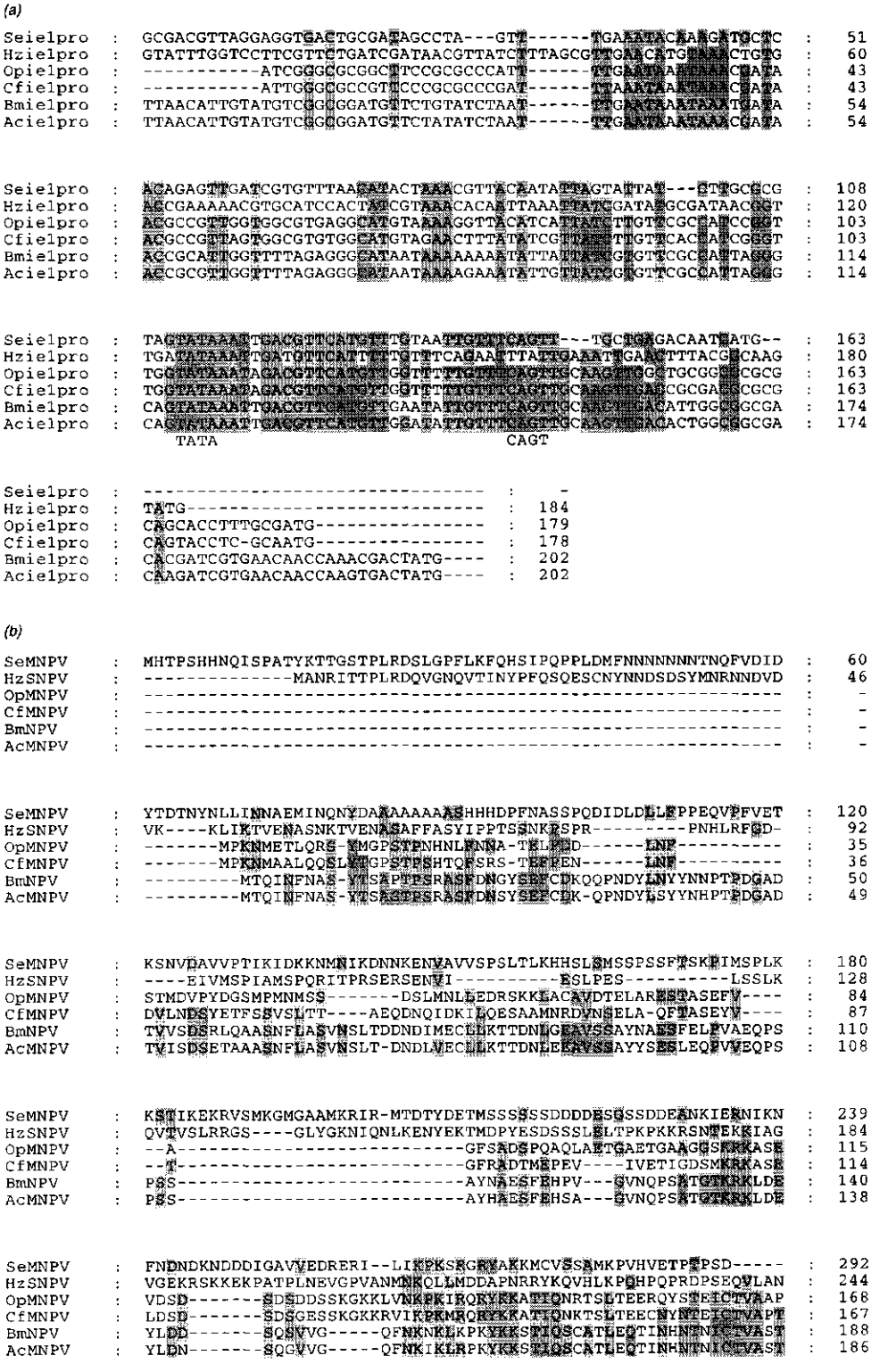
Transcriptional analysis of SeMNPV ie1

To investigate the transcriptional activity of the SeMNPV *ie1* gene, a specific probe was hybridized to a Northern blot with polyA$^+$ RNA isolated from SeMNPV-infected cells (Fig. 5.3a). At 1 h p.i. a single transcript of 2.5 kb in length was observed, which increased in intensity up to 8 h p.i. This transcript could not be detected anymore at 24 h p.i. Overexposure of the blot (Fig. 5.3b) showed only a single transcript at early times p.i., suggesting that the SeMNPV *ie1* gene did not produce spliced transcripts. From 8 h p.i. until 24 h p.i. (Fig. 5.3) and later (data not shown) multiple transcripts were visible. A transcript of 1.7 kb was first seen at 8 h p.i. and increased in intensity at 24 h.p.i. A less abundant transcript of 4.7 kb was also observed at this time point. The transcript was not detected with a probe specific for the SeMNPV p74 gene (data not shown), which suggests that the late 1.7 and 4.7 kb SeMNPV *ie1* transcripts were 3' coterminal.

The transcriptional start site of the SeMNPV *ie1* gene was determined by primer extension with primer A (Fig. 5.3.c). At late times p.i. several transcriptional start sites were used. An extension product of 87 nt, most abundant at 8 hp.i. and decreasing in intensity at later times p.i., corresponded with the T in the sequence TCAG**T**TTG (nt 150 in Fig. 5.1.b). The use of this start site would result in an untranslated 5' leader of 15 nt. An extension product of 210 nt was first observed at 8 h p.i. and increased in intensity at later times p.i. (Fig. 5.1.b and data not shown). This product corresponded with a presumed transcriptional start at the A in the sequence TGCG**A**TAG (nt 27 in Fig. 5.1.b). The use of this start site would result in an untranslated 5' leader of 138 nt. A minor extension product of 129 nt was observed at 24 h p.i. If this product was not due to a preliminary stop in the extension reaction, it correspondeded to the G in the sequence TCTT**G**CGC (nt 108 in Fig. 5.1.b). Overexposure of the autoradiogram and additional primer extension experiments including RNA isolated at 2 h p.i. showed no indication of primer extension products early in infection (data not shown).

Oligonucleotide B was used in primer extension experiments to investigate if the *ie1* transcript of 1.7 kb observed late in infection started internally in the *ie1* coding sequence. At late times p.i. extension products of 95, 82 and 80 nt were observed (Fig. 5.3.c). These products corresponded with the T in GATA**T**CGGC (nt 913 in Fig. 5.1.b) and with the C and A in the sequence TTGT**C**G**A**AGAT (nt 926 and 928 in Fig. 5.1.b). The first translational Met-codon following these presumed transcriptional start sites is observed at nt 984 (Fig. 5.1.b). This Met-codon could correspond with an ORF of 1653 nt, potentially encoding a protein of 551 amino acids. Overexposure of the autoradiogram showed no indication of the use of one of the TAAG sequences internal in the *ie1* coding sequence.

The length of the observed *ie1* transcript of 2.5 kb was in reasonable agreement with the length of the ORF (2142 nt), assuming a poly(A) tail of normal length (0.2 kb), a 5' leader of 15, 57 or 138 nt and transcription termination near one of the 3' polyadenylation signals. However, absence of extension products at 4 h p.i. was in disagreement with the early transcriptional activity observed in the Northern blot analysis (Fig. 5.3a, b). The length of the *ie1* transcript of 1.7 kb, observed in late stages of infection, was in reasonable agreement with a presumed start internal in the *ie1* ORF as determined with primer B (Fig. 5.3.c) (at nt 913, 926 or 928 in Fig. 5.1.b), assuming a poly(A) tail of normal length and transcription termination near one of the polyadenylation signals (at nt 2329 and 2356 in Fig. 5.1.b).


(a)
Seielpro
Hzielpro
Opielpro
Cfielpro
Bmielpro
Acielpro
TATA
CAGT
(b)
SeMNPV
HzSNPV
OpMNPV
CfMNPV
BmNPV
AcMNPV

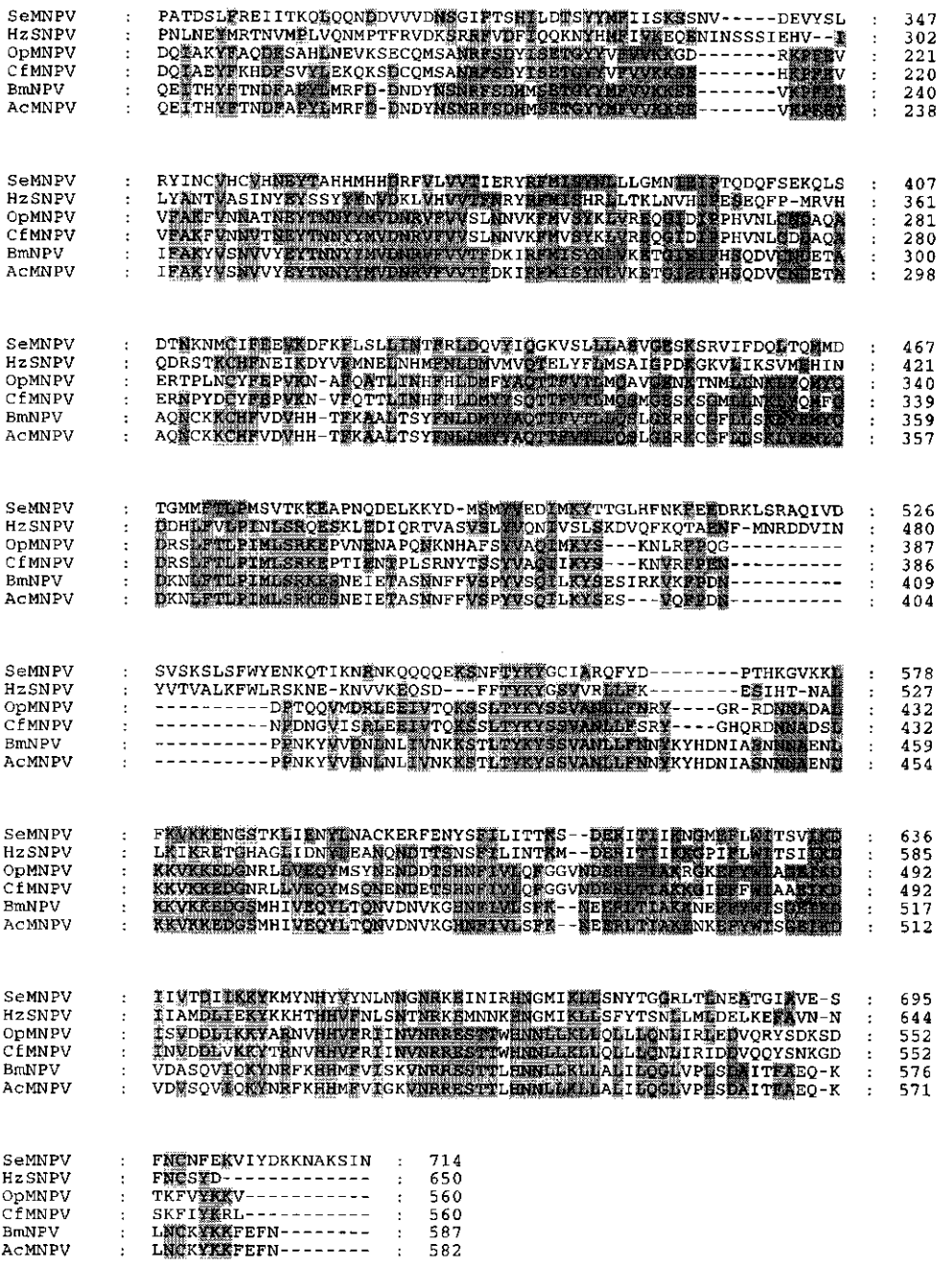


Figure 5.2. a) Alignment of the promoter region of the SeMNPV *ie1* gene with those of other baculovirus *ie1* genes. Gaps, introduced to optimize the alignment, are indicated by dashes. Grey shading is used to indicated the occurrence of identical nucleotides. b) Alignment of the predicted amino acid sequences of the IE1 proteins of SeMNPV, AcMNPV, BmNPV, OpMNPV, CfMNPV and HzSNPV. Gaps, introduced to optimize the alignment, are indicated by dashes. Shading is used to indicated the occurrence (at least 50%) of identical amino acids.

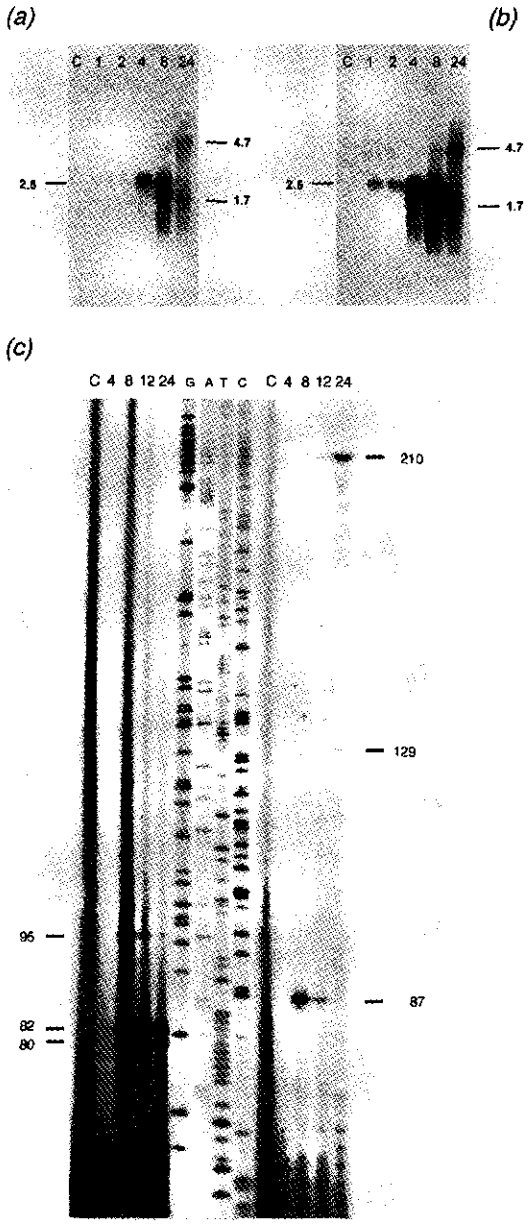


Figure 5.3. a) Northern blot analysis of polyA+ RNA extracted from uninfected (lane C) and SeMNPV-infected cells at 1, 2, 4, 8 and 24 h p.i. for *ie1* transcripts. The length of the transcripts is indicated in kb. b) Overexposure of the blot in a). c). Primer extension analysis of *ie1* transcripts performed with oligonucleotide primers A (right) and B (left), complementary to SeMNPV *ie1* mRNA, ^{32}P labelled at the 5' end. The oligonucleotides were annealled to total RNA from uninfected (lane C) and SeMNPV-infected cells isolated at 4, 8, 12 and 24 h p.i. Numbers indicated the sizes of the extension products, which were determined by comparison with a sequence ladder run alongside (lanes G, A, T and C). The ladder was obtained from an SeMNPV *v-ubi* containing plasmid clone and the *v-ubi* 16-mer oligonucleotide as sequence primer (van Strien *et al.*, 1996, Chapter 4).

Analysis of the SeMNPV "ie0 first exon" region

The SeMNPV *ie1* transcriptional analysis by Northern blot hybridization showed no indication of a larger, potentially spliced transcript early in infection. In AcMNPV-infected cells spliced transcripts (*ie0*) are produced throughout infection, of which the first exon originated approximately 4 kb upstream of the *ie1* gene (Chisholm & Henner, 1988, Kovacs *et al.*, 1991b). Similar spliced transcripts were observed in OpMNPV- (D.A. Theilmann, *pers.comm.*) and in *Lymantria dispar* (Ld) MNPV-infected cells (G.F. Rohrmann, *pers.comm.*).

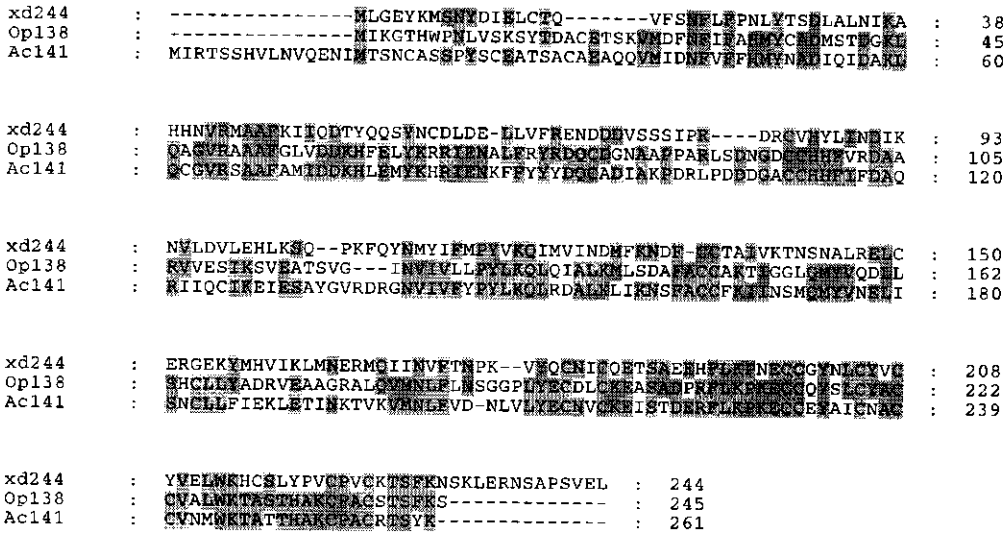
Assuming a similar gene arrangement in SeMNPV as in AcMNPV, the region located 4 to 4.5 kb upstream of the SeMNPV *ie1* gene was analyzed for the presence of sequences with homology to the AcMNPV *ie0* left exon. An ORF, denoted SeMNPV ORF xd244, was identified at the 3' end of the SeMNPV rr1 gene (van Strien *et al.*, 1997) in the same anti-genomic orientation as SeMNPV *ie1* (Fig. 5.1a). ORF xd244 potentially encoded a protein of 244 amino acids with a predicted molecular mass of 29 kDa, which showed 34% and 33% amino acid identity to the AcMNPV ORF 141 (Ayres *et al.*, 1994) and OpMNPV ORF 138 (Ahrens *et al.*, 1997), respectively. These are the ORFs from which the first *ie0* exon originates. The alignment of the predicted amino acid sequences (Fig. 5.4a) revealed the presence of several conserved amino acids, most notably several Cys residues reminiscent of a Zn-finger. However, no function has been assigned to these ORFs.

To investigate the presence of consensus splice donor sequences, the nucleotide sequence encompassing SeMNPV ORF xd244 was aligned with the sequence of the 5' exons and surrounding nucleotides of AcMNPV and OpMNPV *ie0* (Fig. 5.4b). This analysis showed that a consensus splice signal (AGGTRA) was not present in the sequences of SeMNPV ORF xd244, neither at the homologous location, nor at any other location in the SeMNPV ORF xd244 sequence. The early AcMNPV and OpMNPV *ie0* transcriptional start consensus sequence CAGT and preceeding TATA box were also absent from the region upstream of SeMNPV ORF xd244. Late in infection AcMNPV *ie0*/orf 141 was shown to produce spliced and unspliced transcripts, starting from several start sites (Kovacs *et al.*, 1991b). Only one of these, a consensus late promoter sequence TAAG, could be aligned with sequences 5' of the SeMNPV ORF xd244 (Fig. 5.4b).

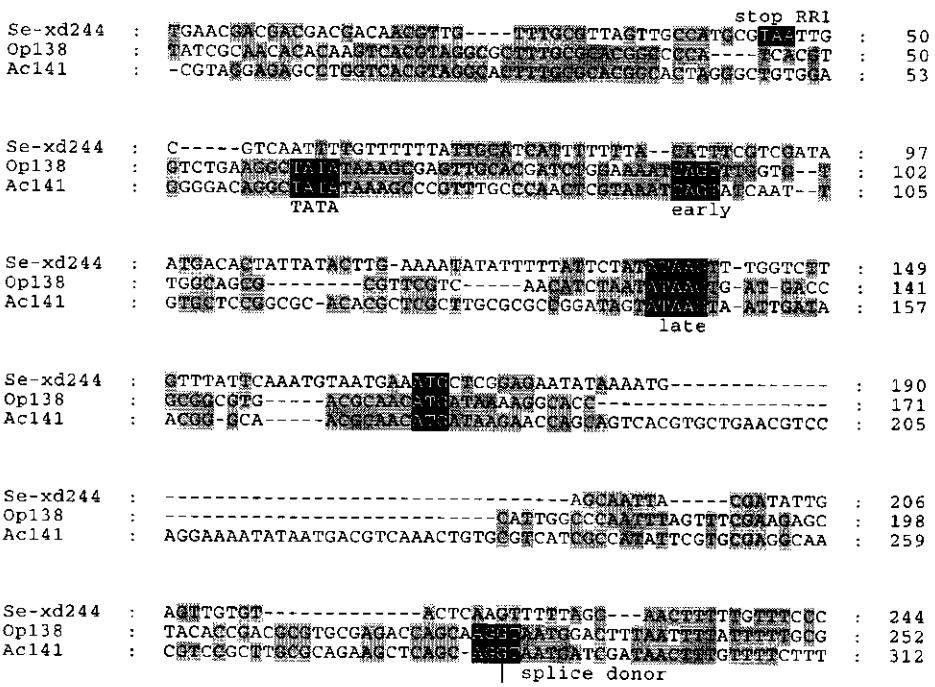
Comparison of IE1 proteins

SeMNPV IE1 showed a moderate similarity of around 50% to other IE1 proteins (Table 5.1). This value is much lower than the similarities of IE1 of AcMNPV, BmNPV, *Choristoneura fumiferana* (Cf) MNPV and OpMNPV, but equal in comparison to *Helicoverpa zea* (Hz) SNPV. This supports the view that SeMNPV is distinct from the other baculoviruses. The SeMNPV IE1 protein, encompassing 714 amino acids, is larger in size than any of the other baculovirus IE1 proteins (Fig. 5.2b). AcMNPV IE1 counted 582 amino acids and OpMNPV IE1 only 560 amino acids. The amino acid sequence alignment (Fig. 5.2b) showed that most of the extra length of the SeMNPV IE1 protein could be accounted for by additional amino acids at the N terminus. Most of the difference in amino acid number between the viruses OpMNPV and AcMNPV can be accounted for by the presence of a repeat of 18 amino acids (amino acids 91 to 109 and 110 to 128) in the latter. The same duplication is found in BmNPV IE1. There is no evidence that similar duplications are responsible for the additional length of the SeMNPV IE1 N terminus. The N termini of the IE1 proteins showed very little homology, although they were acidic and Ser-

(a)


xd244
Op138
Ac141

(b)


stop RR1
TATA
early
late
splice donor
Se-xd244
Op138
Ac141

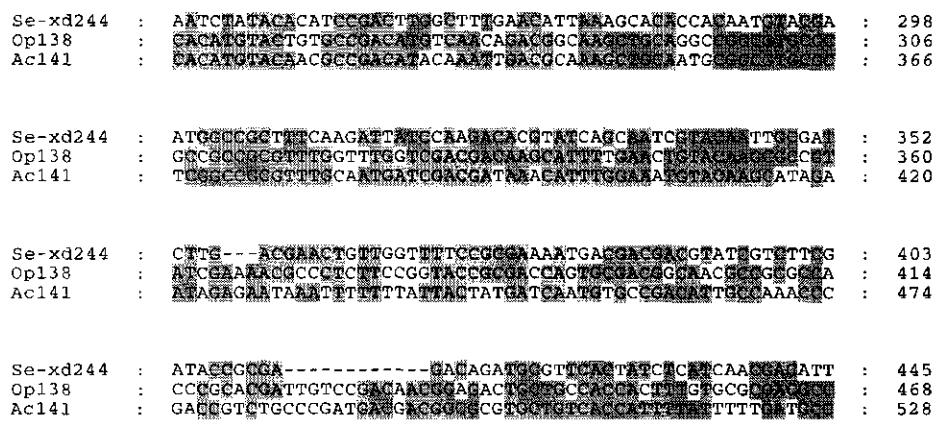


Figure 5.4. a) Alignment of the predicted amino acid sequences of the ORF xd244 protein of SeMNPV with AcMNPV ORF 141 and OpMNPV ORF 138. Gaps, introduced to optimize the alignment, are indicated by dashes. Grey shading is used to indicated the occurrence (at least 50%) of identical amino acids. b) Alignment of the nucleotide sequences of SeMNPV ORF xd244 with the *ie0* exon region of AcMNPV ORF 141 and OpMNPV ORF 138. Gaps, introduced to optimize the alignment, are indicated by dashes. Grey shading indicates the occurrence of identical nucleotides. The location of the AcMNPV and OpMNPV splice donor site, TATA box, early and late transcriptional start sequences are indicated.

and Thr-rich. The C-terminal part of the IE1 proteins (starting at position 235 in SeMNPV IE1) displayed a higher degree of amino acid homology and could be aligned with a few minor gaps. This segment of IE1 was rather basic with a predicted pI of 9.4. The predicted pI of the full length SeMNPV IE1 protein is 7.8.

The alignment of the C-terminal part of IE1 proteins was used to establish phylogenetic relationships between the six baculoviruses of which the *ie1* genes have been identified. An unrooted phylogenetic tree was constructed using parsimony (Swofford, 1993) and subjected to bootstrapping (Felsenstein, 1993) to assess the reliability of the deduced phylogeny. The phylogenetic analysis suggests that AcMNPV is most closely related to BmNPV and OpMNPV to CfMNPV, whereas the closest relative of SeMNPV is HzSNPV (Fig. 5.5). The bootstrap figures were high, indicative of a good reliability of the deduced phylogeny.

DISCUSSION

The complete nucleotide and predicted amino acid sequence of the SeMNPV *ie1* gene was determined. The homology with other baculovirus IE1 proteins was only moderate (Table 5.1) and resided predominantly in the C-terminal two-third of the protein (Fig. 5.4a). Two domains with different functions have been identified in the AcMNPV IE1 protein. The N-terminal part was responsible for transactivation and the C-terminal part for DNA binding and inhibitory activities (Kovacs *et al.*, 1992). In the latter part a consensus single-stranded DNA binding motif consisting of basic and aromatic amino acid residues (Wang & Hall, 1990) was found (Kool *et al.*, 1994). This motif in the SeMNPV IE1 sequence could extend beyond the consensus due to high content of basic and aromatic amino acids of the C-terminal region. A database search with the N-terminal amino acids of SeMNPV IE1 resulted in homology with an array of

Table 5.1. Pairwise amino acid homology of baculovirus IE1 proteins. The predicted sequence of the SeMNPV IE1 protein (length in brackets) was compared with IE1 of AcMNPV (Guarino & Summers, 1987), OpMNPV (Theilmann & Stewart, 1991), HzSNPV (Cowan *et al.*, 1994), BmNPV (Huybrechts *et al.*, 1994) and CfMNPV (Genbank acc.nr. L04945). Identity normal typesetting, similarity in bold.

	AcMNPV	BmNPV	OpMNPV	CfMNPV	HzSNPV
SeMNPV	29.2	30.0	28.7	30.3	32.3
AcMNPV		95.4	48.0	50.0	30.5
BmNPV			48.0	49.5	28.4
OpMNPV				73.9	28.8
CfMNPV					28.7

transcriptional activators, although only with low significance. The N-terminal part of the SeMNPV IE1 protein, as well as other baculovirus IE1s, had a high Ser and Thr content similar to the situation with other transcriptional activators (Ptashne, 1988). These Ser and Thr residues are potential phosphorylation sites. AcMNPV and OpMNPV are multiple phosphorylated (Choi & Guarino, 1995b, Theilmann & Stewart, 1993).

The conservation of the C-terminal DNA binding region in SeMNPV IE1, combined with the similarity with transactivators, suggests that SeMNPV IE1 has a similar role in transactivation of gene expression and/or viral DNA replication as AcMNPV and OpMNPV IE1. The diversity in the N-terminal part and a difference in phosphorylation status of SeMNPV IE1 may play a role in the spceficity of SeMNPV.

Phylogenetic analysis based on the baculovirus IE1 proteins indicates that SeMNPV shared the most recent common ancestor with HzSNPV. In spite of this, SeMNPV as well as HzSNPV appear to have diverged considerably from each other and from the other viruses. This is conspicuous from the branch lengths in the phylogenetic tree (Fig. 5.5) and from the low amino acid homology of the SeMNPV and HzSNPV IE1 proteins (Table 5.1). The high bootstrap figures suggest that the IE1 sequence may have a high reliabilty for baculovirus phylogenetic reconstructions.

The SeMNPV *ie1* gene was trancribed as an early (1 h p.i.) gene, as became apparent from the transcriptional analysis by Northern blotting. Expression of the SeMNPV *ie1* gene in the presence of cycloheximide would confirm the immediate early nature of this gene. Unlike the AcMNPV and OpMNPV *ie1* transcripts, the early SeMNPV 2.5 kb transcript disappeared at the later stages (24 h p.i.) of infection. Primer extension experiments revealed the use of the canonical early baculovirus promoter/initiator motif CAGT, similar to the transcriptional start of *ie1* of other baculoviruses. The use of this motif peaked at 8 h p.i. Unexpectedly, no extension products were observed at 4 h p.i. Later in infection a longer extension product was observed. The presence of two *ie1* primer extension products is reminiscent of the transcriptional activity in AcMNPV and in OpMNPV. However, in these viruses the longer product is the result of the expression of *ie0*, the spliced transcript. In contrast to SeMNPV, the longer extension product

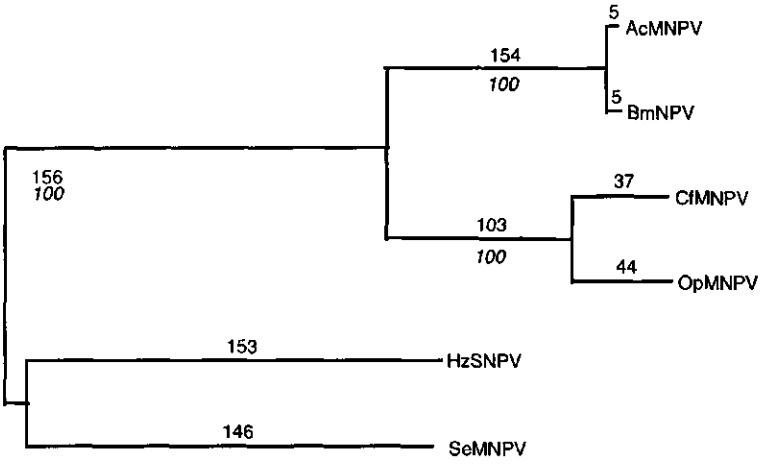


Figure 5.5. Bootstrap analysis (100 replicates, branch and bound search) of the most parsimonous unrooted phylogenetic tree (exhaustive search settings) based on baculovirus IE1 proteins, constructed with the PAUP 3.1 package. Numbers at the branches indicate branch length (normal typesetting) and frequency of cluster (italics).

of AcMNPV and OpMNPV decreased in later stages of infection.

The temporal expression pattern of *ie1* transcription as observed in primer extension experiments was in disagreement with the pattern observed by Northern blot analysis. Notwithstanding repeated experiments no extension products could be observed early in infection. Additional experiments are required to clearify SeMNPV *ie1* gene expression.

The origin of the late (24 h p.i.) 1.7 kb transcript was investigated by primer extension analysis with a primer complementary to sequences located internal in *ie1*. If the observed late extension products were the not due to preliminary stops in the primer extension reaction, the results indicated that late transcription initiated from sites located within the SeMNPV *ie1* coding sequences (Fig. 5.3.c, 5.1.b). A similar expression strategy is found in the OpMNPV *ie3* gene (Wu *et al.*, 1993). If the late transcript would start internally, the putative protein would only contain the conserved, basic C-terminal amino acids and would have a length more similar to other baculovirus IE1 proteins. It is tempting to speculate on the function of such a late protein, for instance as a structural part of the budded virions, as has been observed in OpMNPV (Theilmann & Stewart, 1993), or with a specific role in transactivation or viral DNA replication. Additional experiments are required to determine the origin of the late SeMNPV *ie1* transcripts and if a functional protein is made.

The conserved CAGT sequence is the consensus baculovirus early transcriptional start site and in the case of AcMNPV, OpMNPV and LdMNPV *ie0* also the splice acceptor site of the second exon. The baculovirus *ie0/ie1* gene is the only gene known to be spliced, and this transcription pattern factually results in the expression of two highly homologous proteins with different transactivating properties (Kovacs *et al.*, 1991a). The splice acceptor site, overlapping with the AcMNPV (Chisholm & Henner, 1988), OpMNPV (D.A. Theilmann, *pers.comm.*) and LdMNPV (G.F. Rohrmann, *pers.comm.*) *ie1* transcriptional start site CAGT, is conserved in SeMNPV *ie1* (Fig. 5.2a) but not in HzSNPV *ie1*. SeMNPV encoded an ORF, xd244, with amino

acid homology to the AcMNPV, OpMNPV and LdMNPV ORFs from which the first *ie0* exon originates. However, no consensus splice donor sequence was present in or near SeMNPV ORF xd244 (Fig. 5.4a). In addition, the early baculovirus *ie0* transcriptional start consensus sequence CAGT and preceeding TATA box were absent from SeMNPV. Only a late promoter consensus sequence TAAG could be recognized in front of SeMNPV ORF xd244.

It is of interest to determine the function of ORF xd244 in SeMNPV infection. The observed lack of two SeMNPV *ie1* transcripts early in infection, and the absence of splice donor and early start site consensus sequences in and upstream of ORF xd244, may suggest the absence of splicing in SeMNPV.

ACKNOWLEDGEMENTS

We appreciated the support of Ms. Jessica Tersteege and Mr. Jeroen Bakker in sequence analysis. David A. Theilmann and George F. Rohrmann are acknowledged for the supply of OpMNPV sequence data prior to publication. We would like to thank Rob W. Goldbach for critical reading of the manuscript. The fruitful discussions with R. Broer and J.G.M. Heldens were highly appreciated. This research was supported by DIARP Grant No. 93/20 from the Dutch-Israeli Agricultural Research Program.

CHAPTER 6

BACULOVIRUSES CONTAIN A GENE FOR THE LARGE SUBUNIT OF RIBONUCLEOTIDE REDUCTASE

SUMMARY

In the genomes of two baculoviruses, *Spodoptera exigua* and *S. littoralis* multicapsid nucleopolyhedroviruses (SeMNPV and SpliMNPV, respectively), an open reading frame (ORF) encoding the large subunit of ribonucleotide reductase (RR1) was identified. The predicted amino acid sequences of SeMNPV and SpliMNPV RR1 showed high homology to RR1 proteins from eukaryotes (ca. 70% and 80% similarity, respectively). The amino acid residues thought to be involved in catalytic function were conserved in the baculoviral RR1 ORFs. The RR1 ORFs in SeMNPV and SpliMNPV were located on different genomic positions. In SeMNPV, the RR1 ORF was located upstream of the polyhedrin gene, in an anti-genomic orientation. In SpliMNPV, the RR1 ORF preceded the *p74* gene. By searching databanks, sequences homologous to the N terminus of RR1 were also detected upstream of the polyhedrin gene of three other baculoviruses, Mamestra brassicae multicapsid NPV, Panolis flammea multicapsid NPV and Orgyia pseudotsugata single nucleocapsid NPV. The baculovirus type species, AcMNPV, however, does not encode RR (Ayres *et al.*, 1994). A 2.7 kb transcript could be detected throughout infection with SeMNPV, classifying the SeMNPV *rr1* as an early gene. Primer extension analysis revealed several early and late start sites. None of the major start sites showed similarity to previously characterized baculoviral transcriptional start motifs. Phylogenetic analysis of prokaryote, eukaryote and viral RR1 proteins suggested that SeMNPV and SpliMNPV acquired the gene for RR1 from a eukaryotic source, but independently from each other.

This chapter has been published as: Strien, E.A. van, Faktor, O., Hu, Z.H., Zuidema, D., Goldbach, R.W., & Vlak, J.M. 1997. Baculoviruses contain a gene for the large subunit of ribonucleotide reductase. Journal of General Virology 78, 000-000.

INTRODUCTION

Ribonucleotide reductase (RR) is a key enzyme in the biosynthesis of deoxyribonucleotides, catalyzing the reduction of ribonucleotides (for reviews see: Thelander & Reichard, 1979, Reichard, 1988, 1993, Elledge *et al.*, 1992). The RR (class I) enzyme is present in all eukaryotes and in some prokaryotes, such as *E.coli* grown under aerobic conditions (Reichard, 1993). The *E.coli* RR enzyme is considered to be the prototype and functions as a heterodimer consisting of two large (RR1) and two small (RR2) subunits encoded by different genes. The actual reduction of ribonucleotides takes place at the large subunit. The small subunit supplies the reducing capacity; it contains an active iron centre and provides electrons via a tyrosyl radical. The reduction of ribonucleotides into deoxyribonucleotides is subject to complex allosteric control exerted by binding of nucleotides to the large subunit. The overall activity level of the enzyme is regulated by the binding of ATP and dATP to the large subunit. Expression of the ribonucleotide reductase genes is tightly regulated during the cell cycle (Reichard, 1988, Elledge *et al.*, 1992).

Members from several groups of large DNA viruses, such as the poxviruses, herpesviruses, T-even bacteriophages and African swine fever virus (ASFV), encode their own RR. In the animal viruses RR also acts as a virulence factor. Deletion usually results in an attenuated phenotype with slightly impaired growth in cultured cells and increased restrictions in tissue specificity. RR-encoding viruses can apparently circumvent the strict regulation of RR activity by the cellular enzyme, as well as bypass the compromized host regulation machinery (Reichard , 1988, Conner *et al.*, 1994, Slabaugh *et al.*, 1984, Child *et al.*, 1990, Cunha and Costa, 1992, Howell *et al.*, 1993, Heineman & Cohen, 1994, Huszar & Bachetti, 1981, Goldstein & Weller, 1988). RR is an often sought target in the search for anti-viral, anti-microbial and cancer therapies, either suppressing enzyme activity exploited by the pathogen during infection, or by inhibiting the elevated RR enzyme activity of rapidly growing cancer cells, which are no longer subject to normal cell cycle control (Conner *et al.*, 1994, Lori *et al.*, 1994, Reichard, 1988).

RR has not yet been described for baculoviruses, another major group of large DNA viruses. Baculoviruses are pathogenic for arthropods and have a large (80 to 160 kbp) circular dsDNA genome (Murphy *et al.*, 1995). They replicate in the nucleus of the infected cell. Genes encoding RR have not been reported for the type species of the Baculoviridae, *Autographa californica* nucleopolyhedrovirus (AcMNPV)(Ayres *et al.*, 1994).

Here we present evidence that two other members of the baculovirus family, *i.e. Spodoptera exigua* and *S. littoralis* nucleopolyhedrovirus (SeMNPV and SpliMNPV, respectively), contain an open reading frame (ORF) with a high degree of homology to the large subunit of ribonucleotide reductase (RR1) of eukaryotic and viral origins. Expression of the SeMNPV *rr1* gene in infected insect cells was investigated by transcriptional analysis. To determine the ancestry of the two baculoviral RR1s a phylogenetic tree was constructed using parsimony.

MATERIALS AND METHODS

Virus, insects and cells. The SeMNPV field isolate (SeMNPV/US) (Gelernter & Federici, 1986b)

was kindly provided as polyhedra by B.A. Federici (Department of Entomology, University of California, Riverside, USA). The polyhedra were propagated in fourth instar *S. exigua* larvae (Smits *et al.*, 1988). Budded virus (BV), used in time-course infection experiments, was obtained from the supernatant of UCR-Se1 (Gelemter & Ferderici, 1986b) or IZD-Se-2109 cells (a gift from B. Moeckel, Institute of Zoology, Technical University, Darmstadt, Germany) which had been infected with haemolymph from SeMNPV-infected fourth instar larvae. The *S. exigua* cell lines were maintained at 27° C in plastic tissue-culture flasks in TNM-FH medium (Hink, 1970) supplemented with 10% foetal bovine serum. BV titres were determined by the end-point dilution method and expressed as $TCID_{50}$ units per ml. Cells were infected with a m.o.i. of 5 $TCID_{50}$ per cell. SpliMNPV type-B isolate E15 was plaque-purified from haemolymph of field-infected *S. littoralis* larvae in SPC-Sl-52 cells (Mialhe *et al.*, 1984), maintained at 27° C in TNM-FH medium (Sigma) supplemented with 10% heat-inactivated foetal calf serum.

DNA analysis. Viral DNA was extracted from SeMNPV occlusion body derived virions (ODVs) purified by sucrose gradient centrifugation essentially as described by Caballero *et al.* (1992). Viral DNA from SpliMNPV was purified from viral occlusion bodies propagated in third instar *S. littoralis* larvae.

Localization and DNA sequencing of the rr1 genes. SeMNPV fragment *Xba*I-D, containing N-terminal sequences of the polyhedrin gene (van Strien *et al.*, 1992) was cloned into pUC18. Subfragments upstream of the polyhedrin gene were isolated from agarose gels using the freeze-squeeze method (Sambrook *et al.*, 1989) and subcloned into the plasmids pTZ19R (Promega) or pBluescriptKS+ (Stratagene). Sequencing of SeMNPV inserts with standard sequencing and custom designed primers (Eurogentec) was performed with *Taq* polymerase using the chain termination method. SeMNPV *rr1* gene DNA fragments were [α^{32}P]dAT- labelled and hybridized under non-stringent conditions (Sambrook *et al.*, 1989) to a Southern blot of SpliMNPV viral DNA on Hybond-N filters (Amersham). Hybridizing SpliMNPV DNA fragments were purified from agarose gels with the Jetsorb kit (Genomad, USA), (sub)cloned into pUC19 and sequenced according to the chain termination method with standard sequencing primers using the Sequenase kit (United States Biochemical).

Sequence analysis of the rr1 genes. Sequences were analyzed with the UWGCG computer programs (Devereux *et al.*, 1984); DNA and deduced amino acid sequences were compared with the updated GenBank/EMBL, SWISSPROT and PIR databases using the FASTA program and BLAST network service (Altschul et al., 1990). Phylogenetic analysis was performed with the PAUP 3.1 program (Swofford, 1993), using GCG pileup to produce input files of aligned protein sequences. Bootstrap analysis (Felsenstein, 1993), included in the PAUP package, was used to assess the integrity of the produced phylogeny.

Isolation of total RNA and Northern blot hybridization. For Northern blot analysis and primer extension total RNA was isolated from SeMNPV-infected UCR-Se1 or IZD-Se-2109 cells at various times post infection (p.i.), as described by van Strien *et al.* (1992). Total RNA was denatured, electrophoresed in agarose gels and blotted onto Hybond-N nylon membrane (Van Strien *et al.* 1992). To identify *rr1* transcripts, the blot was hybridized for 16 h at 65 °C with an

[α^{32}P]ATP-labelled *rr1* specific probe. After hybridization, the filter was washed for 5 min with 2 x SSC, 0.5% SDS at room temperature, 30 min with 2 x SSC, 0.1% SDS at 65° C and 30 min with 0.1 x SSC, 0.1% SDS at 65° C. The filter was exposed to Kodak XAR film.

Primer extension. To identify the transcriptional start site(s) of SeMNPV *rr1*, 15 ng of an oligonucleotide (5' CAGACTATTCAAGAGCAGAG 3'), complementary to the *rr1* mRNA, was labelled at the 5' end with [γ ^{32}P]dATP by T4 polynucleotide kinase (Gibco-BRL) in 50 mM Tris/HCl, pH 9.5, 10 mM $MgCl_2$, 5 mM DTT, 5% glycerol for 45 min at 37°C followed by heat denaturation at 90° C for 10 min. The labelled oligonucleotide was purified on a 1 ml Sephadex G25 column. Labelled primer was added to 10 µg of total infected-cell RNA and the mixture was denatured at 90° C for 5 min and annealed at 54° C for 15 min. Reverse transcription was carried out at 48° C for 1 h in a volume of 15 µl, containing 5 mM of each of the dNTPs and 1 µl Superscript reverse transcriptase (Gibco-BRL) in a buffer supplied by the manufacturer. The reaction was stopped by addition of 5 µl 'stop' buffer containing 95% (v/v) formamide, 0.01% xylene cyanol and 0.01% bromophenol blue. The reaction mixture (6 µl) was analyzed on a 6% denaturing polyacrylamide gel.

RESULTS

Location and sequence of the rr1 genes of SeMNPV and SpliMNPV

In the course of sequencing the region upstream of the SeMNPV polyhedrin gene (van Strien *et al.*, 1992), a putative ORF of 2310 nt was found (Fig. 6.1a), potentially encoding a protein of 770 amino acids with a molecular mass of 87 kDa. This ORF, which was anti-genomic in orientation, had a high degree of homology to RR1 of eukaryotic and prokaryotic organisms and viruses and will be further referred to as the SeMNPV *rr1* gene.

Baculovirus early (CAGT) and late (TAAG) consensus transcriptional start sites (Blissard and Rohrmann, 1990) were observed 42 and 120 nt upstream of the translational start codon, respectively (Fig. 6.2c). No TATA box sequences could be identified. Immediately after the stop codon the nucleotide sequence was relatively A/T rich. A consensus polyadenylation signal (AATAAA) was first seen 639 nt downstream the translational stop codon (data not shown). The presumed translational start codon was in agreement with the Kozak consensus rule (Kozak, 1983). The putative start codons of the SeMNPV *rr1* and polyhedrin gene were separated by an intergenic region of 516 nt. The SeMNPV *rr1* stop codon and a putative downstream ORF (data not shown) were separated by 117 nt.

An *rr1* gene was also identified in SpliMNPV by hybridization with an SeMNPV *rr1* DNA probe (data not shown). The SpliMNPV *rr1* ORF was localized on the *Hind*III-E fragment of the SpliMNPV genome (Croizier *et al.*, 1989) and cloned and sequenced (Fig. 6.1b). The SpliMNPV *rr1* ORF consisted of 2345 nt with the potential to encode a protein of 781 amino acids with a predicted molecular mass of 87 kDa. Upstream of the presumed translational start site no consensus late baculoviral transcription start site (Blissard & Rohrmann, 1990) could be detected. A baculovirus consensus early CAGT sequence, without a preceding TATA box, was observed 79 nt in front of the translational start codon. Following the translational stop codon three consensus polyadenylation signals (AATAAA) were found. The ATG of the SpliMNPV *rr1* ORF was in favourable translational context (Kozak, 1983). SpliMNPV *rr1* preceded an ORF

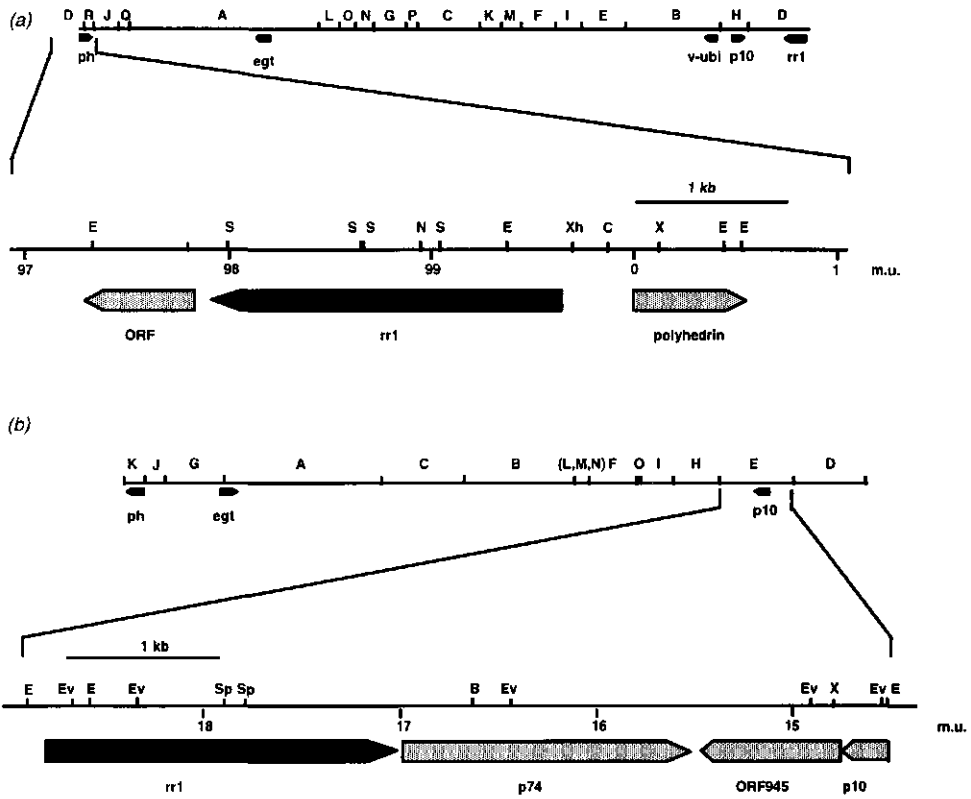
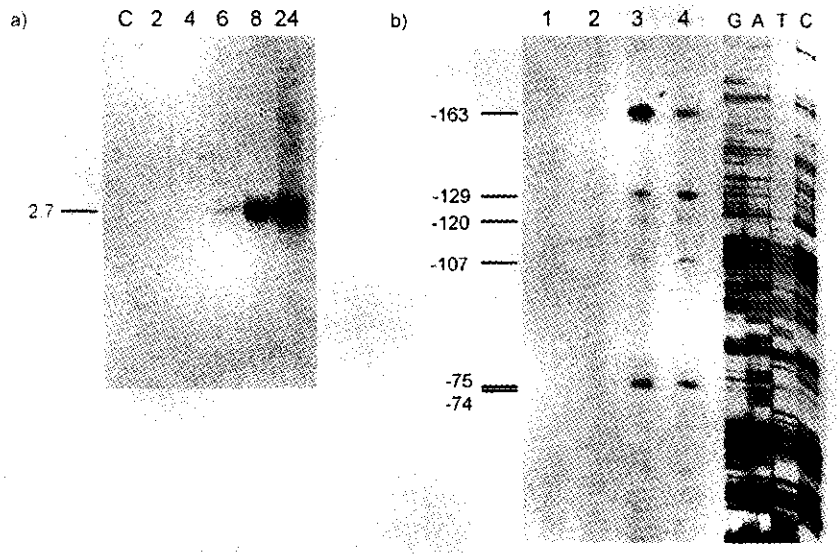


Figure 6.1. Location and restriction endonuclease recognition sites of the large subunit of the *rr1* gene region on the physical map of the SeMNPV (a) and SpliMNPV (b) genome. The location and orientation of the *rr1*, polyhedrin, p10 and *egt* genes are indicated (arrows). Map units (m.u.) are also given. a) Physical map of SeMNPV. E, *Eco*RI; X, *Xba*I; N, *Nru*I, Xh, *Xho*I, S, *Sal*I, C, *Cla*I. b) Physical map of SpliMNPV. E, *Eco*RI; Ev, *Eco*RV; Sp, *Sph*I; X, *Xba*I. The presentation of the SpliMNPV map is reversed compared to previous publications (Croizier *et al.*, 1989, Faktor *et al.*, 1997b).

with homology to the AcMNPV *p74* gene (Kuzio *et al.*, 1989). The SpliMNPV *rr1* stop codon and the presumed SpliMNPV *p74* start codon were separated by 101 nt.

Transcriptional activity of the SeMNPV rr1 gene

To investigate the transcription of the SeMNPV *rr1* gene, a Northern blot with total RNA isolated from SeMNPV-infected cells at several times p.i. was hybridized to a *rr1*-specific probe (Fig. 6.2a). The SeMNPV *rr1* gene transcript was approximately 2.7 kb in size and present throughout infection, as observed after overexposure of the autoradiogram (data not shown), thus classifying *rr1* as an early gene. The amount of transcript increased at later time points. The transcriptional start of the SeMNPV *rr1* gene was determined by primer extension (6. 2b). At early (4 h) and late (12, 24 h) times p.i. several transcriptional start sites were used. The major early start site was located at the A of the sequence TTT**A**TT at nt -163 with respect to the



(c)

```
CCCAAGGCTGGGTTATAGCTGTAGCGAGTATACATTATATTTATTTTTTCACAATGTTTT   60
 G  L  A  P  N  Y  S  Y  R  T  Y  M < polyhedrin

ACGAAAGGAAAAAATTACTTACAATTCCCGCAACCATATCATTACTTATCGTCGTCGATG   120

ATAAAACAACATACCATTTTATCGATAGATATGGTTATCGCGTCTAGTATATATATATAA   180

ACGTTTCGATGAAAATGTATCTAATAAACTCAAAGTTTCGACTTTGGTAAAATTGATATT   240

AACATGGAAGATTTATACTGTAACGAAGAATTGGGTTTGTCGCCGCCGTCGTCGCCGTTG   300

TCAATGTTGTCTTTTTCTGACGACGACGAACAACGACGATGGTTTGAAATAAATATTAGA   360

                  E,L -163                                L -129
CACTTGGATGAAAATTTATTATCGTTGAACGACGAGGACGAGGACGACGACGACGAGCAT   420

(L -120)          L -107                        L -75,-74
AAGAGGCAGCGGCAAATATTTGAAAAAATTAAATTGCAAGAGTTTATTTTAGAGCGAGCC   480

GCTCGAGTGGTTAAAAGACAGTTCGAAGATTTCGAAGAAAACAATCCTCGTTATTACATC   540

RR1 >
ATGTTCGATGAAAATAATTTTGAAATCGACAGTAAATTTAAAATTATACCATCAAACAAT   600
M  F  D  E  N  N  F  E  I  D  S  K  F  K  I  I  P  S  N  N

GTTTCGGAAAATTTAACTCTGCTCTTGAATAGTCTGTGCAAGAACATTGATACTCGTTAC   660
V  S  E  N  L  T  L  L  L  N  S  L  C  K  N  I  D  T  R  Y
```

Figure 6.2. Transcriptional analysis of the SeMNPV *rr1* gene region. a) Northern blot analysis of total RNA extracted from uninfected (lane C) and SeMNPV-infected UCR-Se cells at 2, 4, 6, 8, 12, and 24 h p.i. for *rr1* transcripts. b) Primer extension analysis of *rr1* transcripts performed with a 20-mer oligonucleotide complementary to the *rr1* RNA, ^{32}P-labelled at the 5' end. The oligonucleotide was annealed to total RNA from uninfected (lane 1) and SeMNPV-infected IZD-Se cells isolated 4, 12 and 24 h p.i. (lane 2, 3 and 4) and elongated by reverse transcription. Sizes of the extension products were determined by comparison with a sequence ladder run alongside (lanes G, A, T and C) obtained from an SeMNPV *rr1* containing plasmid clone and the 20-mer oligonucleotide as a sequence primer. Numbers indicate the position of the 5'-terminal nucleotides. c) Nucleotide sequence of the SeMNPV *rr1* promoter region. Transcriptional start sites as determined by primer extension analysis are indicated in bold. The location of the oligonucleotide primer is dubble underlined.

translational start codon. At late times p.i. this site was also used. In addition, major late transcriptial start sites were observed at GAC**G**AC (nt -129), at GGC**A**AA (nt -107) and at TTT**A**TTT (nt -74 and -75). The consensus late promoter sequence AT**A**AG (nt -120) was also used late in infection, although much less frequent than the other late start sites. Overexposure of the primer extension autoradiogram showed no indication of the use of this site at 4 h p.i., which also showed that the observed bands in the primer extension analysis were not due to preliminary stops. Comparable results were obtained with primer extension experiments at different annealing temperatures (data not shown). The length of the SeMNPV *rr1* transcript (2.7 kb) was in reasonable agreement with the length of the ORF (2.3 kb), assuming a polyA tail of normal length (0.2 kb), a 5' leader of 163 to 74 nt and transcription termination near the translational stop codon.

Comparison of baculoviral RR1 proteins with those from other sources

The amino acid sequences of SeMNPV and SpliMNPV RR1 were compared with those from other organisms and viruses. The overall homology (identity and similarity) is given in Table 1. The baculovirus RR1s have the highest homology to RR1s from eukaryotic organisms. The homology of SpliMNPV RR1 to these RR1s is higher than that of SeMNPV RR1. The homology between SeMNPV and SpliMNPV RR1, 69%, is essentially the same as that between SeMNPV and eukaryotic RR1s. The homology with RR1s from other DNA viruses is usually lower.

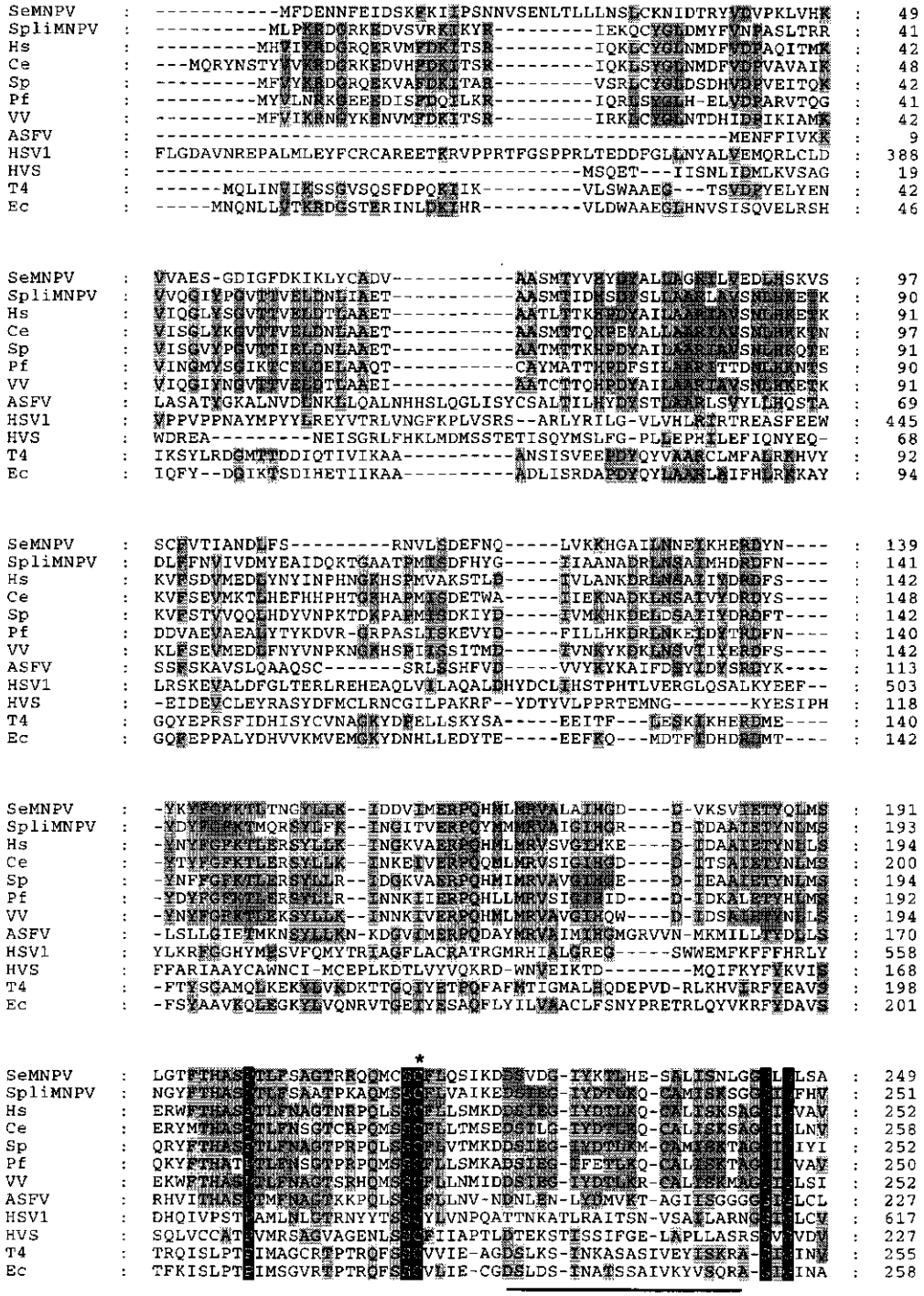
Alignment of baculoviral RR1 with an array of cellular and viral RR1s indicated that the overall length of SeMNPV and SpliMNPV RR1 fell within the range observed with other RR1s (Fig. 6.3). SeMNPV RR1 had diverged in several regions of the protein and showed a few gaps and amino acid insertions, in comparison to eukaryotic RR1. Several consensus amino acid residues (Mao *et al.*, 1992, Stubbe *et al.*, 1990, Uhlin & Eklund, 1994) could be recognized in the baculoviral RR1. These included the cysteine residues known to be involved in the formation of essential thiols in the *E.coli* RR1 protein, i.e. Cys225, Cys439, Cys462, Cys754 and Cys759 (*E.coli* numbering), as well as residues surrounding Cys225, Cys439 and Cys462 in the active site and two adjacent tyrosine residues (Tyr730 and 731) presumably involved in electron transfer from the small subunit.

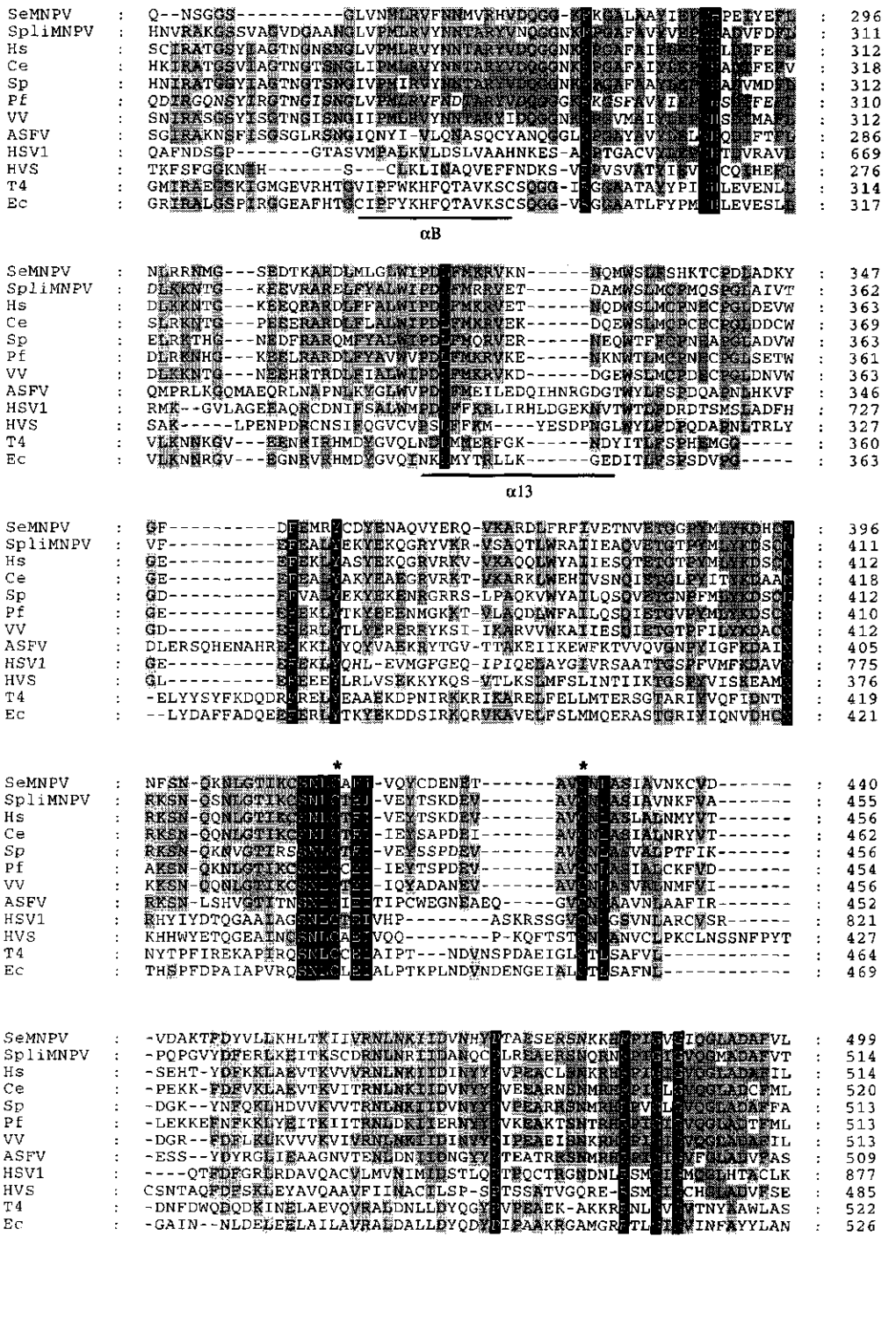
Structural analysis of the *E.coli* RR1 protein (Uhlin & Eklund, 1994) and previous research (Davis et al., 1994, Mao et al., 1992) revealed areas responsible for dimerization of RR1 and holoenzyme formation with RR2 subunits. The dimerization domains (two alpha helices, αA and αB) were well-conserved in SpliMNPV RR1 and slightly less in SeMNPV RR1. Holoenzyme formation involves the binding of the ultimate C-terminal residues of RR2 to two RR1 alpha helices (αI and α13). The baculoviral RR1s showed reasonable sequence homology to the corresponding eukaryotic regions.

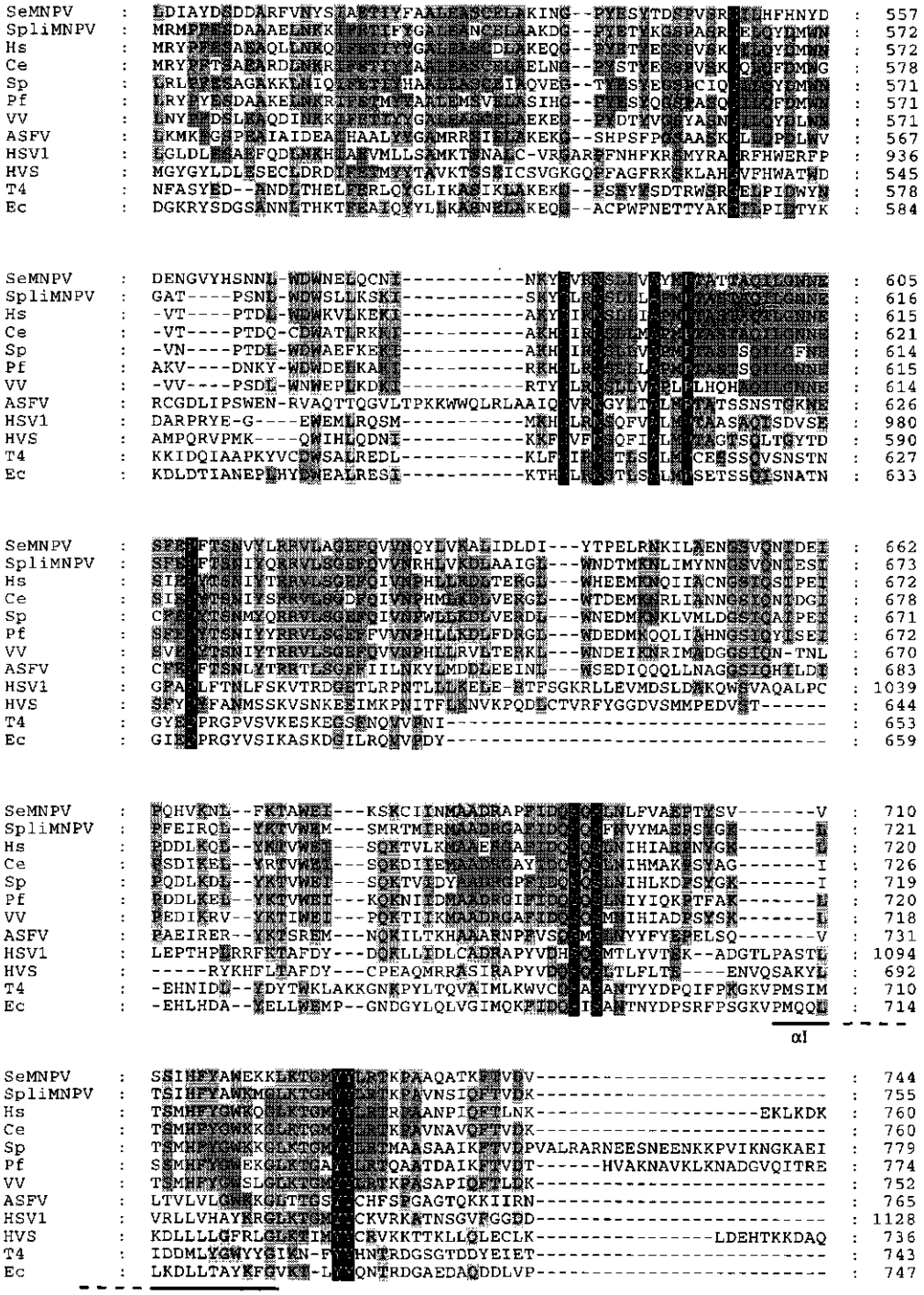
Databank searches revealed the occurrence of RR1-like sequences upstream of the

Table 1. Homology of SeMNPV and SpliMNPV RR1 with RR1 proteins from several organisms and viruses in percent identity and similarity (in bold). For abbreviations see legend Fig. 4.

	SpliMNP	Hs	Ce	Sp	Pf	VV	ASF	HS	HV	Ec
SeMNPV	50.1	51.3	51.2	48.9	48.9	51.3	38.3	29.8	27.0	26.
SpliMNPV		67.8	64.4	61.4	58.7	61.7	39.6	28.3	27.2	28.


SeMNPV
SpliMNPV
Hs
Ce
Sp
Pf
VV
ASFV
HSV1
HVS
T4
Ec
*
αA


αB
α13


αI
αI

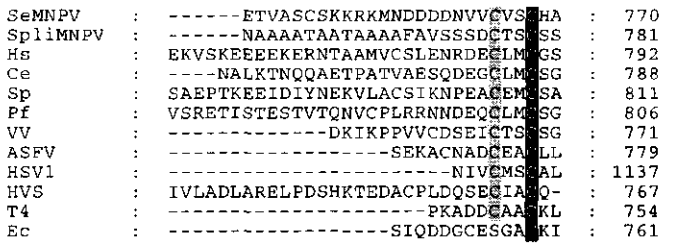


Figure 6.3. Alignment of the predicted amino acid sequences of the RR1 proteins of the baculoviruses SeMNPV and SpliMNPV with those of *Homo sapiens* (Hs, Parker *et al.*, 1991), *Caenorhabditis elegans* (Ce, Sulston *et al.*, 1992), *Plasmodium falciparum* (Pf, Ruben *et al.*, 1993), *Schizosaccharomyces pombe* (Sp, Fernandez *et al.*, 1993), vaccinia virus (VV, Tengelsen *et al.*, 1988), African swine fever virus (ASFV, Boursnell *et al.*, 1991), herpes simplex virus type 1 (HSV1, Nikas *et al.*, 1986), herpes virus saimiri (HVS, Nicholas *et al.*, 1992), bacteriophage T4 (Tseng *et al.*, 1988) and *E.coli* (Ec, Nillson *et al.*, 1988). Gaps, introduced to optimize the alignment, are indicated by dots. Shading is used to indicated the occurrence (black 100%, gray at least 60%) of identical amino acids. Asteriks indicate essential cysteine residues. Alpha helices involved in dimerization are underlined (see text).

polyhedrin gene in *Mamestra brassicae* NPV (MbMNPV, Cameron & Possee, 1989), *Pannolis flammea* NPV (PfMNPV, Oakey *et al.*, 1989) and *Orgyia pseudotsugata* NPV (OpSNPV, Leisy *et al.*, 1986b). The predicted amino acid sequence in these viruses showed homology to the N terminus of RR1 (Fig. 6.4). MbMNPV and PfMNPV RR1 showed the same gap around position 110 in amino acid alignment as SeMNPV RR1 in comparison with eukaryotic. However, the available sequences from MbMNPV and PfMNPV showed a somewhat higher homology to the eukaryote RR1 N terminus than SeMNPV RR1 did.

Phylogenetic reconstruction of the origin of the baculoviral rr1 genes

In order to investigate the origin of the *rr1* gene in SeMNPV and SpliMNPV, a phylogenetic tree was constructed with amino acid sequences of 30 RR1 proteins from prokaryotes, eukaryotes and viruses. An unrooted parsimonious tree was calculated with the PAUP heuristic search algorithm, followed by bootstrap analysis to assess the variability of the produced phylogeny (Fig. 6.5). The tree showed that prokaryotes and eukaryotes, as expected, were located on separate branches. Viral RR1 sequences were found on many different locations in the tree. The herpesvirus RR1s were all clustered, implying that they diverged from a common ancestor. Bootstrap analysis supported the phylogenetic separation of the herpesviruses into α- and γ-herpesviruses. Both SeMPNV and SpliMNPV grouped with the eukaryote organisms. The phylogeny of this branch was reasonably well supported by the bootstrap analysis. The tree suggested that SeMNPV and SpliMNPV RR1 do not share a recent common ancestor, but that these viruses acquired the gene for RR1 independent from each other as well as from other DNA viruses.

DISCUSSION

The complete nucleotide sequence of the *rr1* gene was determined in the baculoviruses SeMNPV and SpliMNPV. Transcriptional activity was investigated after RNA isolation from SeMNPV-infected cultured insect cells. Both Northern blotting and primer extension showed low

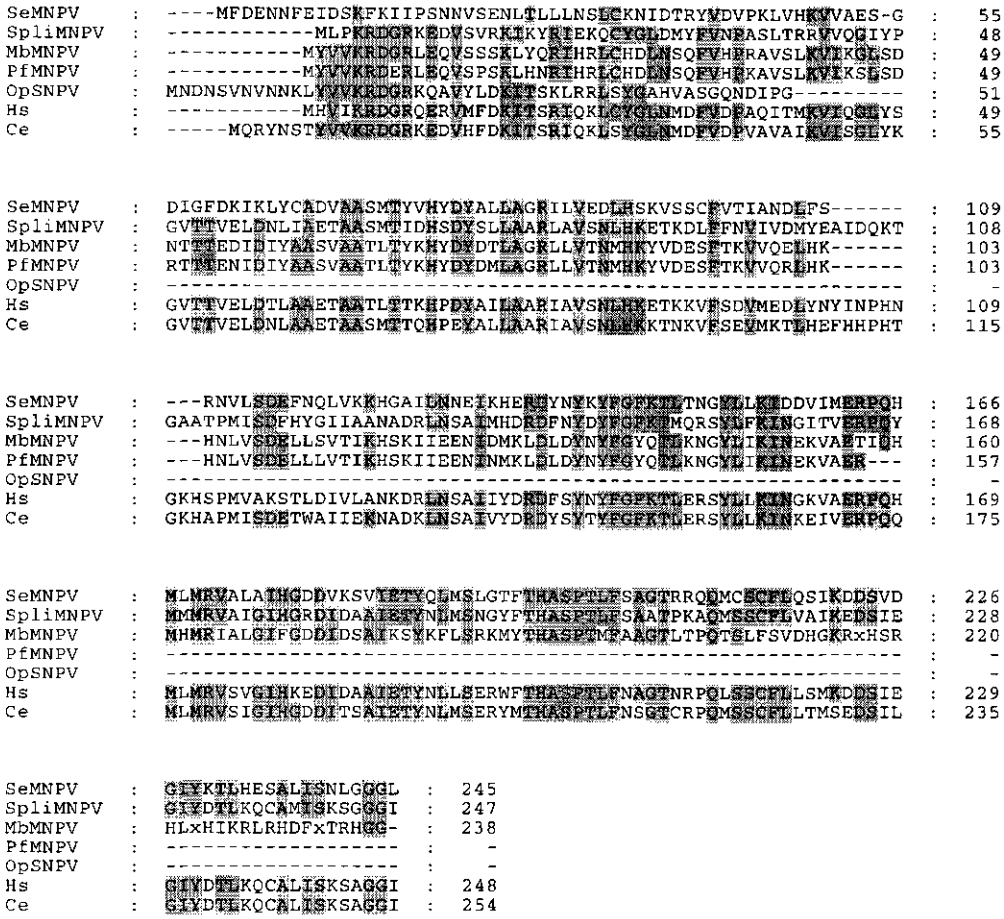


Figure 6.4. Alignment of the predicted amino acid sequences of the N terminus of the RR1 proteins of SeMNPV, SpliMNPV, *H.sapiens* and *C.elegans* with those of the baculoviruses MbMNPV, PfMNPV and OpSNPV. Shading indicates the occurrence of identical amino acids.

transcriptional activity early in infection, which increased at later stages. The early transcription of the SeMNPV *rr1* gene is in good agreement with its presumed role in deoxyribonucleotide synthesis required for DNA replication. Herpesvirus and vaccinia *rr* genes are also transcribed as early genes (Schmitt *et al.*, 1988, Swaink & Galloway, 1986).

The sequence surrounding the major transcriptional start sites showed no homology to previously characterized start sites of other baculovirus genes, for instance of those involved in DNA replication, such as helicase (Lu & Carstens, 1992) and DNA polymerase (Tomalski *et al.*, 1988). No homology between the sequences 5' of the translational start of SeMNPV, SpliMNPV and other baculoviruses *rr1* genes could be detected.

The alignment of SeMNPV and SpliMNPV RR1 with other RR1s (Fig. 6.3) showed that amino acid residues known to be involved in enzymatic activity were conserved. This suggests that the baculoviral *rr1* gene could code for a functional constituent of a RR enzyme. SeMNPV

RR1 had a lower homology to eukaryotic RR1 than SpliMNPV RR1. In comparison to the *E.coli* structural regions (Uhlin & Eklund, 1994), regions from eukaryotic and specifically viral RR1 which diverged from the *E.coli* sequence, often map to loops separating α-helices and β-barrels. One of the gaps in the SeMNPV RR1 alignment was located around amino acid 250. A similar situation was observed in several other viral RR1s. In *E.coli*, this region forms a loop (L1, amino acids 259-277) between the dimerizing domains of RR1. It has been suggested that this and other loops in the dimerization region may change from a flexible to a fixed position at subunit formation, hence stabilizing the holoenzyme (Uhlin & Eklund, 1994). The SeMNPV enzyme may thus show a stability different from the cellular enzyme.

Sequences with homology to the RR1 N terminus were found in the baculoviruses MbMNPV, PfMNPV and OpSNPV (Fig. 6.4). This suggests that RR1 is not only encoded by SeMNPV and SpliMNPV, but is widespread among baculoviruses. The *rr1* gene in MbMNPV, PfMNPV and OpSNPV had a similar genomic location to that from SeMNPV, *i.e.* positioned upstream of the polyhedrin gene. Phylogenetic trees based on polyhedrin gene sequences (Zanotto *et al.*, 1993b, Cowan *et al.*, 1994) placed SeMNPV, MbMNPV and PfMNPV in subgroup IIA. The position of OpSNPV was variant depending on which algorithms and of DNA or amino acid sequences were used. OpSNPV was either in subgroup IIB, together with SpliMNPV, or in IIA. OpSNPV might belong to subgroup IIA, based on the observed homology to the RR1 N terminus and its genomic location next to the polyhedrin gene.

Bootstrap analysis of the parsimonious phylogenetic tree (Fig. 6.5) confirmed the independent ancestry of the *rr1* gene in SeMNPV and SpliMNPV. This conclusion is in agreement with the different genomic location of the *rr1* gene in these two viruses, as well with the observed lower homology of SeMNPV RR1 with eukaryotic RR1s. Alternatively, if the common ancestor of SeMNPV and SpliMNPV already encoded RR1, the implication would be that the *rr1* gene evolved much faster in SeMNPV than in SpliMNPV. This assumption seems questionable and is not supported by determination of the evolutionary rates of the polyhedrin genes of SeMNPV and SpliMNPV (Zanotto *et al.*, 1993b, Cowan *et al.*, 1994), nor by comparison of other genes of SeMNPV and SpliMNPV (data not shown). However, the phylogenetic reconstruction may yield a different result when full length sequences of other baculovirus *rr1* genes become available.

The positioning of SeMNPV and SpliMNPV in the eukaryote branch of the tree suggests that each derived RR1 independently from an eukaryotic source, for example their host. The polyphyletic origin of baculovirus RR1 contrasted with the observed monophyletic grouping of the herpesviruses and of the two poxviruses. Only RR1 from SpliMNPV and poxviruses grouped inside a cellular clade, whereas RR1 from SeMNPV, ASFV, herpesviruses and bacteriophage T4 diverged more than their presumed ancestors did. This suggests that viral RR1 usually diverges faster than cellular RR1.

RR1 proteins need to associate with RR2 dimers in order to form a functional enzyme. In most large DNA viruses (with the exception of vaccinia virus) the genes for RR1 and RR2 are located adjacent to each other. A different situation exists in ß-herpesviruses such as HCMV, which only contain a RR1-like ORF (Chee *et al.*, 1990). However, unlike the baculoviral RR1 ORFs, the RR1-like ORFs in the ß-herpesviruses show only very limited homology to other RR1s and do not encode the amino acid residues known to be involved in ribonucleotide reduction. It has been suggested that these highly diverged genes may have acquired another

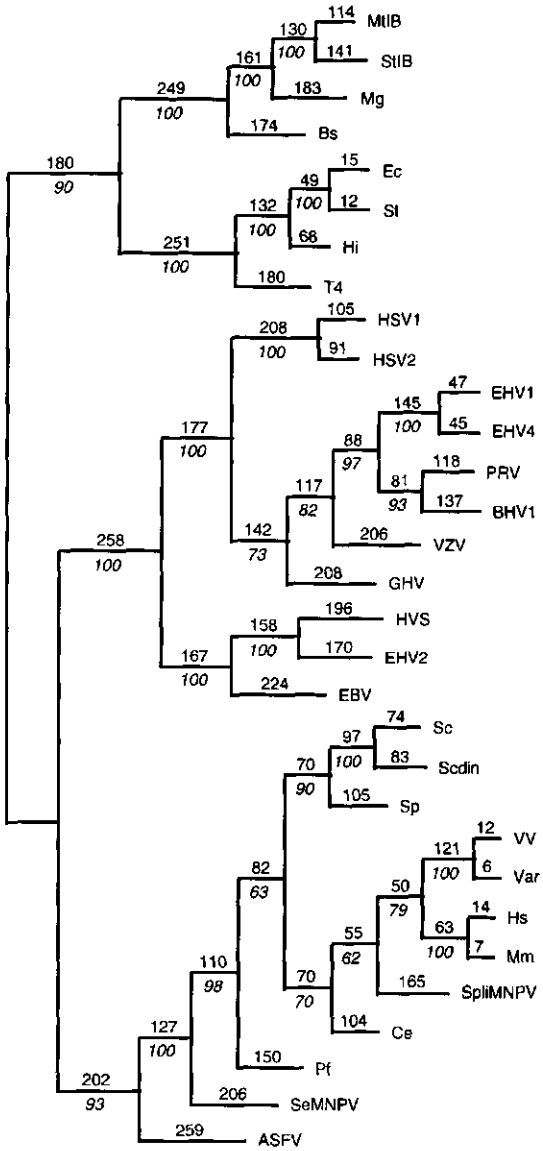


Figure 6.5. Bootstrap analysis (100 replicates) of an unrooted phylogenetic tree of RR1 proteins constructed with PAUP heuristic search algorithm. Numbers at the branches indicate branch length (normal typesetting) and frequency of cluster (italics).
Abbreviations: MtlB, *Mycobacterium tuberculosis* nrdE (Yang *et al.*, 1994); StlB, *Salmonella typhimurium* nrdE (Jordan *et al.*, 1994); Mg, *Mycoplasma genetialum* (Fraser *et al.*, 1995); Bs, *Bacillus subtillus* (acc.nr. Z68500); Ec, *Escherichia coli* nrdA (Nillson *et al.*, 1988); St, *Salmononella typhimurium* nrdA (Jordan *et al.*, 1994); Hi, *Haemophilus influenzae* (Fleischmann *et al.*, 1995); T4, bacteriophage T4 (Tseng *et al.*, 1988) HSV1, herpes simplex virus type 1 (Nikas *et al.*, 1986); HSV2, herpes simplex virus type 2 (Swain and Galloway, 1986);

EHV1, equine herpes virus type 1 (Telford *et al.*, 1992); EHV4, equine herpes virus type 4 (Riggio and Onions, 1994); PRV, pseudorabies virus (Kaliman *et al.*, 1994); BHV1, bovine herpes virus type 1 (Simard *et al.*, 1995); VZV, varicella zoster virus (Davison and Scott, 1986); GHV, gallid herpes virus 2 (Darteil *et al.*, 1995); HVS, herpes virus saimiri (Nicholas *et al.*, 1992); EHV2, equine herpes virus type 2 (acc.nr. U20824); EBV, Epstein-Barr virus (Baer *et al.*, 1984); Sc, *Saccharomyces cerevisae* (Elledge and Davis, 1990); Scdin, *Saccharomyces cerevisae* DNA damage inducible subunit (Yagle and McEntee, 1990); Sp, *Schizosaccharomyces pompe* (Fernandez *et al.*, 1993); VV, vaccinia virus (Tengelsen *et al.*, 1988); Var, variola virus (Shchelkunov *et al.*, 1993); Hs, *Homo sapiens* (Parker *et al.*, 1991); Mm, *Mus musculus* (Caras *et al.*, 1985); Ce, *Caenorhabditis elegans* (Sulston *et al.*, 1992); Pf, *Plasmodium falcifiparum* (Ruben *et al.*, 1993); ASFV, African swine fever virus (Boursnell *et al.*, 1991).

function during the ß-herpesvirus life cycle (Conner *et al.*, 1994). In SeMNPV, homology to RR2 could not be detected either in the 2,5 kb sequence upstream or in the 14 kb sequence downstream of the SeMNPV RR1 ORF (data not shown). However, an RR2 ORF might be located elsewhere on the genome. As yet, we do not know if SeMNPV and SpliMNPV encode their own RR2, or alternatively, combine with cellular RR2 proteins. Such a heterologous association is known in yeast, where two RR enzymes are recognized. One enzyme results from the combination of RR1 and RR2, whose normal expression is restricted to certain phases of the cell cycle. The other enzyme is induced by DNA damage and is a combination of the same RR2 and a different RR1 (Elledge *et al.*, 1992).

Virus-encoded RR enzymes from poxviruses, herpesviruses, ASFV and phage T4 are not essential and are distinct from the cellular enzyme in several aspects. It remains to be elucidated if similar characteristics hold for baculoviruses encoding or lacking RR.

Not all large DNA viruses encode RR. As can be concluded from sequence analysis of the complete genome, the baculovirus type species AcMNPV does not encode RR (Ayres *et al.*, 1994) and neither does the closely related *Bombyx mori* NPV (BmNPV, accession nr. L33180). This explains our failure in attempts to hybridize an SeMNPV *rr1* DNA probe to AcMNPV DNA (data not shown). The herpesvirus channel catfish virus (Davison, 1992) also does not encode a ribonucleotide reductase. The reasons for this are unknown. The lack of *rr1* in AcMNPV will allow the study of SeMNPV *rr1* in an AcMNPV recombinant expressing this gene.

During the course of our experiments Ahrens *et al.* (1997) published the entire sequence of the *Orgyia pseudotsugata* nucleopolyhedrovirus (OpMNPV) genome and also identified a *rr1* gene, which potentially encoded a relatively short protein of 593 amino acids. In juxtaposition of the OpMNPV *rr1* gene a gene with low sequence homology to *rr2* was found in this baculovirus. Preliminary phylogenetic analysis indicated that the OpMNPV *rr1* gene did not share a recent common ancestor with either SeMNPV or SpliMNPV, but was acquired independently in these three baculoviruses.

ACKNOWLEDGEMENTS

We thank E.C. Klinge-Roode for excellent technical assistance with the primer extension experiments and Basil M. Arif for reviewing earlier versions of this manuscript. This work was supported by Grant no. 93/20 from the Dutch-Israeli Agricultural Research Programme (DIARP).

CHAPTER 7

THE GENETIC ORGANIZATION OF A 20 KB DNA SEGMENT OF *SPODOPTERA EXIGUA* NUCLEOPOLYHEDROVIRUS: GENE ORDER AS A MARKER FOR BACULOVIRUS PHYLOGENY

SUMMARY

The organization of a contiguous genomic region of 20 kb (map unit 87 to 2) of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV) was analyzed and compared with the corresponding genomic regions of other baculoviruses. Twenty open reading frames (ORFs) and a potential homologous region (*hr*) were identified in this segment, which spanned the *p10* and the polyhedrin gene. All ORFs had homologs in *Autographa californica* (Ac) MNPV (except for a *rr1* gene, which encodes the large subunit of ribonucleotide reductase) and *Orgyia pseudotsugata* (Op) MNPV. A single SeMNPV ORF, xd379, had homology with two consecutive AcMNPV ORFs encoding ODV-E18 and ODV-EC27 (Braunagel *et al.*, 1996). The homology with the ORFs of AcMNPV and OpMNPV was generally low, except for *v-ubi* and polyhedrin. Multiple sequence alignments were created from which consensus amino acids could be assigned. Most of the genes were found to be preceded by the baculovirus consensus late promoter motif TAAG. Some had a consensus polyadenylation signal immediately downstream of the translational stop codon. An *hr* element of 900 nt was identified, consisting of four imperfect palindromic repeats of 72 nt in length, centered around a *Bgl*II site, and three direct repeats of 47 nt in length. In analogy to other baculoviruses this region might act as enhancer of transcription and origin of DNA replication. The very high homology of the polyhedrin gene, compared to the homology displayed by the other genes in the 20 kb region, questioned the value of this gene as a phylogenetic marker. Indeed the reliability of the baculovirus phylogeny based on this gene was found to be rather low. The high homology among the polyhedrin genes may suggest that this gene is either frequently exchanged between baculoviruses or alternatively, that is has been relatively recently acquired. Despite the conservation of the ORFs in the 20 kb SeMNPV DNA segment, the gene order has been extensively altered as compared to AcMNPV and OpMNPV. This up to now unnoticed gene shuffling in the baculoviral genome is compared with the situation in other large DNA viruses and may be used as a supplementary, independent character to reconstruct baculovirus phylogeny.

Manuscript in preparation: Strien, E.A. van, Zuidema, D., & Vlak, J.M. 1997. The genetic organization of a 20 kb DNA segment of *Spodoptera exigua* nucleopolyhedrovirus: gene order as a marker for baculovirus phylogeny.

INTRODUCTION

The baculovirus *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV) combines specificity for a single host species, the beet army worm *Spodoptera exigua*, with high virulence. In contrast, the baculovirus type species *Autographa californica* (Ac) MNPV can infect 73 different species, including *S. exigua* (Possee *et al.*, 1993). AcMNPV, however, has a tenfold lower virulence for this pest insect than SeMNPV (Smits *et al.*, 1988). These different characteristics raised the question whether the organization and expression of the SeMNPV genome would show major differences with AcMNPV, whose genome has been completely sequenced (Ayres *et al.*, 1994).

The major late genes polyhedrin and *p10* of SeMNPV have been previously characterized (van Strien *et al.*, 1992, Zuidema *et al.*, 1993). It was noted that the distance between the *p10* and polyhedrin gene was considerably smaller in SeMNPV (11 kb) than in AcMNPV (19 kb) and *Orgyia pseudotsugata* (Op) MNPV (22 kb). In AcMNPV and OpMNPV the region between polyhedrin and *p10* encompasses the major immediate early genes (*ie1*, *ie0*, *ie2* (*ien*) and *ie3* (*pe-38*)). The immediate early genes are the first genes to be transcribed in the infection process. Via their transactivation properties and involvement in viral DNA replication and late gene expression they are entailed in the first steps towards a successful infection. Recently, it was shown that the SeMNPV *ie1* gene (Chapter 5 and van Strien *et al.*, 1997b) had a comparable genetic location as in AcMNPV and OpMNPV, in between the *p10* and polyhedrin gene. To further investigate the genetic organization of SeMNPV, the complete nucleotide sequence of a continuous 20 kb segment, spanning the *p10* and polyhedrin genes, was determined.

The SeMNPV ORFs found in this segment, which represents 15% of the genome, appeared to have homologs in AcMNPV and/or other baculoviruses. The polyhedrin gene showed a remarkable high level of sequence conservation and therefore the significance of its use in baculovirus phylogenetic reconstructions was evaluated. Despite the structural conservation of the SeMNPV ORFs, their arrangement was strikingly different from other baculoviruses. The extent of baculovirus gene rearrangement is compared to the situation in other large DNA viruses and its potential as a phylogenetic marker is discussed.

MATERIALS AND METHODS

Virus. The SeMNPV field isolate (SeMNPV/US) (Gelernter & Federici, 1986) was kindly provided as polyhedra by B.A. Federici (Department of Entomology, University of California, Riverside, USA). The virus was propagated in fourth instar *S. exigua* larvae (Smits *et al.*, 1988). Viral DNA was extracted from SeMNPV occlusion derived virions (ODVs) purified on sucrose gradients essentially as described by Caballero *et al.* (1992).

Localization and DNA sequencing. SeMNPV DNA fragments were purified from agarose gels by the freeze-squeeze method (Sambrook et al., 1989), (sub)cloned into pUC19 and sequenced with standard sequencing primers and custom designed oligonucleotide (Eurogentec) primers using the chain termination method (Sanger, 1977).

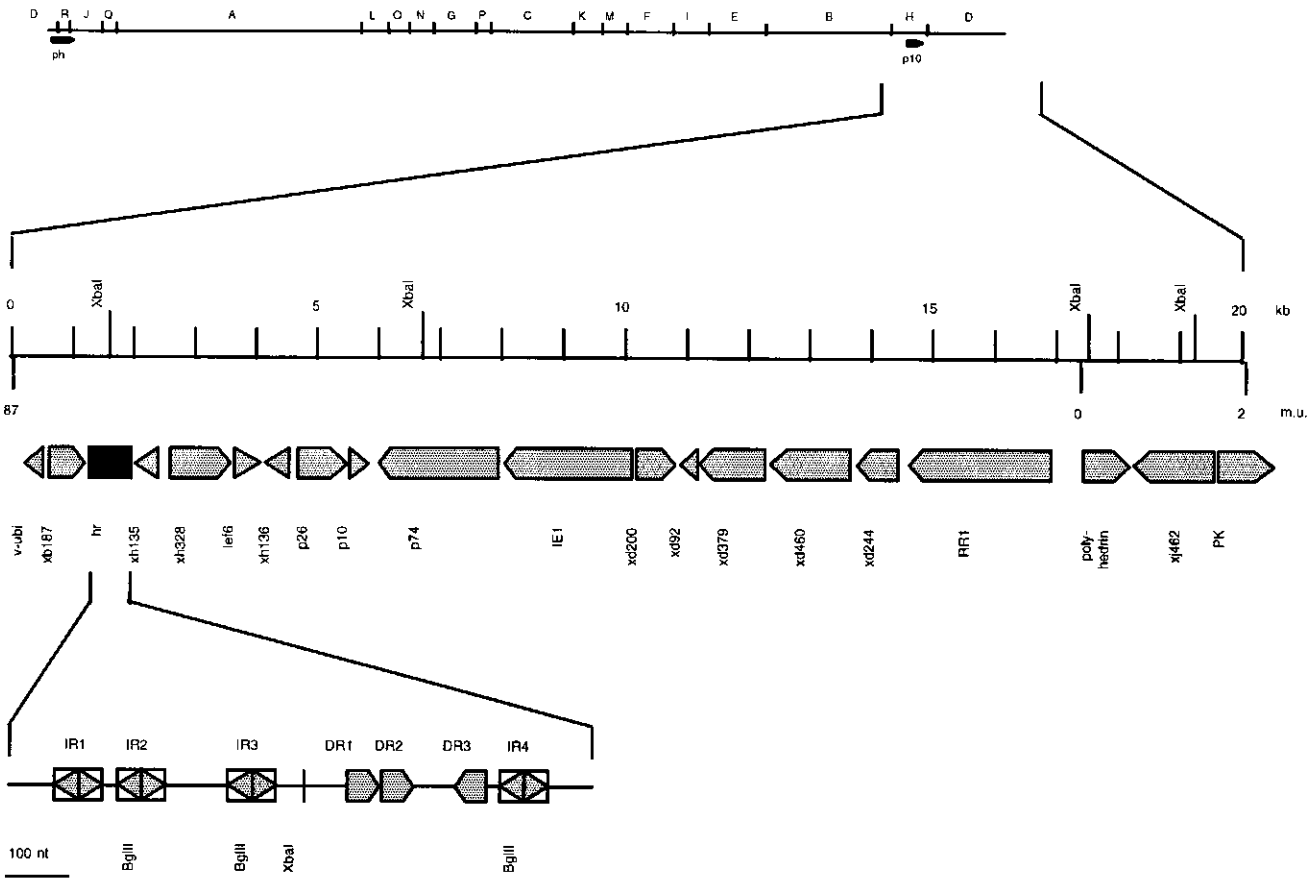
Analysis of the SeMNPV sequences. Sequences were analyzed with the computer programs UWGCG 8.0 (Devereux *et al.*, 1984) and Generunner 3.0 (Hastings Software Inc.). DNA and deduced amino acid sequences were compared with the GenBank/EMBL, SWISSPROT and PIR databases using the FASTA program and BLAST network service (Altschul *et al.*, 1990). Information as available in the daily updated databases was used until Jan. 1st, 1997. Phylogenetic analysis was performed with the PAUP 3.1 program (Swofford, 1993), using GCG pileup to produce input files of aligned protein sequences. Bootstrap analysis (Felsenstein, 1993), included in the PAUP package, was used to assess the integrity of the produced phylogeny.

For the polyhedrin gene alignments and phylogenetic analysis information was used from the following baculoviruses (names and abbreviations according to Murphy *et al.*, 1995):

Ac: AcMNPV, *Autographa californica* MNPV (Hooft van Iddekinge *et al.*, 1983);
Ace: ArceMNPV, *Archips cerasivoranus* MNPV (acc.nr. U40834);
Ag: AgMNPV, *Anticarsia gemmatalis* MNPV (Zanotto *et al.*, 1993b);
Ar: ArMNPV, *Attacus ricini* MNPV (Hu *et al.*, 1993b);
Bm: BmNPV, *Bombyx mori* NPV (Iatrou *et al.*, 1985);
Bs: BusuSNPV, *Buzura suppressaria* SNPV (Hu *et al.*, 1993a);
Cf: CfMNPV, *Choristoneura fumiferana* MNPV (B. Arif, *pers.comm.*);
Hc: HycuMNPV, *Hyphantria cunea* (acc.nr. D14573);
Hz: HzSNPV, *Helicoverpa zea* SNPV (Cowan *et al.*, 1994);
Ld: LdMNPV, *Lymantria dispar* MNPV (Smith *et al.*, 1988);
Ls: LsMNPV, *Leucania separata* (acc.nr. U30302);
Mb: MbMNPV, *Mamestra brassicae* MNPV(Cameron & Possee, 1989);
Mn: ManeNPV, *Malacosoma neustria* NPV (Kozlov *et al.*, 1994);
Op: OpMNPV, *Orgyia pseudotsugata* MNPV (Leisy *et al.*, 1986b);
OpS: OpSNPV , *Orgyia pseudotsugata* SNPV (Leisy *et al.*, 1986a);
Pf: PafIMNPV, *Panolis flammea* MNPV (Oakey *et al.*, 1989);
Pn: PenuMNPV, *Perina nuda* MNPV (Chou *et al.*, 1996);
Se: SeMNPV, *Spodoptera exigua* MNPV (van Strien *et al.*, 1992);
Sf: SpfrMNPV, *Spodoptera frugiperda* MNPV (Gonzalez *et al.*, 1989);
Sl: SpliMNPV, *Spodoptera littoralis* MNPV (Croizier & Croizier, 1994);
Slu: SpltMNPV, *Spodoptera litura* MNPV (acc.nr. X94437);
ClGV: CrleGV, *Cryptophlebia leutotreta* GV (Jehle & Backhaus, 1994b);
PbGV: *Pieris brassicae* GV (Chakerian *et al.*, 1985);
TnGV: *Trichoplusia ni* GV (Akiyoshi *et al.*, 1985).

Table 7.1. Genetic organization of a 20 kb segment of genomic SeMNPV DNA. Of each ORF the position of the translational start and stop codons, the number of amino acids and predicted molecular mass (in kDa) are indicated. Small alternative, overlapping ORFs showing no homology to baculoviral or cellular genes were disregarded. The homology to ORFs of AcMNPV and OpMNPV is given as percentage amino acid identity. ORFs from AcMNPV are numbered according to Ayres *et al.* (1994), from OpMNPV according to Ahrens *et al.* (1997). Unless an SeMNPV ORF showed clear homology to an ORF with an established name of known function, the name of the ORF was derived from the number of amino acids it potentially encoded and from the *Xba*I fragment on which the putative translational start codon was located.

SeMNPV							AcMNPV				OpMNPV			
ORF	left	dir	right	prom	amino acids	Mw	ORF	dir	amino acids	iden-tity	ORF	dir	amino acids	iden-tity
v-ubi	272	<<	514	L	80	9.4	**35**	>>	77	73 %	**25**	<<	93	77 %
xb187	604	>>	1167	L	187	22.0	**34**	<<	215	33 %	**26**	>>	209	40 %
HR														
xh135	2025	<<	2432	L	135	15.0	**26**	>>	135	33 %	**42**	<<	127	31 %
xh328	2578	>>	3564	L	328	38.0	**25**	<<	316	28 %	**43**	>>	300	31 %
lef6	3592	>>	4083	L	163	19.1	**28**	>>	173	26 %	**40**	<<	138	28 %
xh136	4123	<<	4533		136	16.3	**29**	<<	71	32 %	**39**	>>	75	34 %
p26	4638	>>	5474	L	278	31.1	**136**	>>	240	33 %	**132**	>>	230	33 %
p10	5516	>>	5782	L	88	9.6	**137**	>>	95	27 %	**133**	>>	92	37 %
p74	5875	<<	7833	E,L	652	74.3	**138**	<<	645	55 %	**134**	<<	644	55 %
ie1	7970	<<	10114	E	714	82.0	**147**	>>	582	29 %	**145**	>>	560	29 %
xd200	10147	>>	10749	L	200	22.4	**146**	<<	201	34 %	**144**	<<	197	33 %
xd92	10840	<<	11118	L	92	10.6	**145**	>>	77	56 %	**142**	>>	95	46 %
xd379	11133	<<	12272	L	379	43.1	**144**	>>	290	56 %	**141**	>>	297	51 %
							143	>>	90	41 %	**140**	>>	85	39 %
xd460	12310	<<	13692	L	460	54.1	**142**	>>	477	52 %	**139**	>>	484	50 %
xd244	13707	<<	14441	L	244	28.7	**141**	>>	261	30 %	**138**	>>	245	30 %
RR1	14564	<<	16873	L	770	87.4					**32**	<<	593	25 %
poly-hedrin	17382	>>	18122	L	246	28.9	**8**	>>	245	85 %	**3**	<<	245	84 %
xj462	18226	<<	19614	L	462	50.4	**9**	<<	543	30 %	**2**	>>	473	24 %
PK	19613	>>	20500	L	295	34.8	**10**	>>	272	41 %	**1**	<<	274	39 %


D R J Q A L O N G P C K M F I E B H D
ph
p10
0
XbaI
5
XbaI
10
15
XbaI
XbaI
20
kb
87
0
2
m.u.
v-ubi
xb187
hr
xh135
xh328
lef6
xh136
p26
p10
p74
IE1
xd200
xd92
xd379
xd460
xd244
RR1
poly-hedrin
xj462
PK
IR1
IR2
IR3
DR1
DR2
DR3
IR4
100 nt
BglII
BglII
XbaI
BglII

RESULTS

The complete DNA sequence of a 20 kb DNA segment of SeMNPV, spanning map unit (m.u.) 87 to m.u. 2, was analyzed for the presence of ORFs and other structural elements (Table 7.1, Fig. 7.1). Of the thirty-one partially overlapping ORFs larger than 75 amino acids, twenty ORFs in the 20 kb SeMNPV DNA segment had a homolog in AcMNPV (Ayres *et al.*, 1994) and/or OpMNPV (Ahrens *et al.*, 1997) and were therefore further investigated (Table 7.1, 7.1). The function of most ORFs is not yet known. Multiple sequence alignments are presented, including deduced conserved amino acid residues (Fig. 7.2 to 7.14). The 20 kb DNA segment was examined for the presence of baculovirus early and late promoter sequences at the 5' end (Blissard & Rohrmann, 1990) and consensus polyadenylation signals at the 3' end of the twenty ORFs. In total, the sequence (T/G/A)TAAG was recognized 54 times and the sequence CAGT 84 times. These sequences are only mentioned when they are located in front of the translational start codon. In addition, only when the early promoter/initator sequence CAGT is preceded by a TATA-box sequence (two times), the combination is marked as a putative early transcriptional start site. Examining the segment in the genomic, clockwise orientation, the following genes and genetic elements are encountered.

v-ubi and xb187

The two ORFs located at the 3' end of the 20 kb fragment, *v-ubi* and xb187, have been discussed elsewhere (Chapter 4, van Strien *et al.*, 1996). Cellular and viral ubiquitin are associated with the extra-cellulair virions (ECVs) via a phospholipid anchor (Guarino *et al.*, 1995). The viral ubiquitin is not essential, but mutation of the AcMNPV gene resulted in much reduced budded virus production (Reilly & Guarino, 1996). Recently is was shown that the rate of ATP-dependent degradation was lower for viral than for eukaryotic ubiquitin. These observations led to the suggestion that baculovirus v-ubi may function by blocking the normal protein destruction in the ubiquitin pathway (Haas *et al.*, 1996).

homologous repeat region

At the 3' end of ORF xb187 a region with a length of 900 nt was identified (Fig. 7.1) that consisted of four repeated imperfect palindromic sequences of 72 nt in length, centered around a *Bgl*II site (Fig. 7.2a). Furthermore, three direct repeats with a length of 47 nt were identified (Fig. 7.2b). This SeMNPV repeat region will be further referred to as the *hr* region.

In several other baculoviruses multiple copies of *hr*'s occur dispersed around the genome (Cochran & Faulkner, 1983, Maeda & Majima, 1990, Arif & Doerfler, 1984, Theilmann & Stewart, 1992a, McClintock & Dougherty, 1988, Xie *et al.*, 1995). To investigate if a similar situation existed in SeMNPV, the identified *hr* was used as a probe and hybridized to a blot with SeMNPV DNA under low stringency conditions. No clear hybridization signal other than from XbaI-H and -B was observed in these experiments (data not shown), suggesting that this region

Figure 7.1. Schematic representation of the genetic organization of the 20 kb DNA segment of SeMNPV. Top: physical map for *Xba*I of the SeMNPV genome. Middle: overview of the position and orientation of the identified ORFs and *hr* element. Distances in kb, map units and *Xba*I restriction sites are indicated. Bottom: schematic representation of the structural organization of the *hr* element. *Xba*I and *Bgl*II restriction sites are indicated. IR, inverted repeat/palindrome. DR, direct repeat.

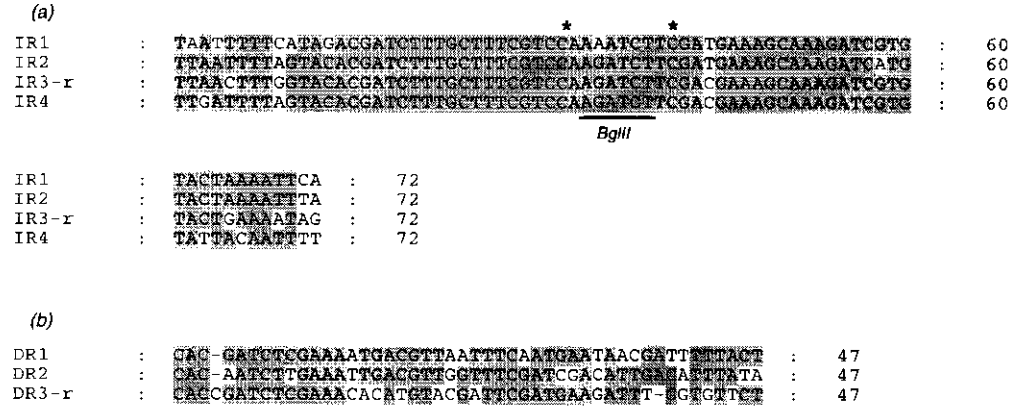


Figure7.2. a). Alignment of the four SeMNPV palindromes (IR1-4) from the *hr* region. The third palindrome is given in the reverse orientation (IR3-r). Asterisks indicate the location of the conserved mismatch. The *Bgl*II recognition site is overlined. b). Alignment of the three SeMNPV direct repeats (DR1-3) from the *hr* region. The third direct repeat is given in the reverse orientation (DR3-r). Dashes are introduced to optimize the alignment. Shading indicates the occurrence of identical nucleotides.

was present only once in the SeMNPV genome. However, the SeMNPV genome contained multiple, homologous *hr*'s with palindromic sequences, as became apparent from additional sequence information and from hybridization experiments with an oligonucleotide probe, based on a conserved sequence in the palindromic sequences (Broer *et al.*, in preparation).

In many baculoviruses the (locations of) *hr* regions are correlated with the occurrence of variable regions in baculovirus genomes, for instance as observed in different isolates (Garcia-Canedo *et al.*, 1996, Cochran & Faulkner *et al.*, 1983, Majima *et al.*, 1993, Arif & Doerfler, 1984, Brown *et al.*, 1984, Faktor *et al.*, 1997a). The differences in *Bgl*II restriction enzyme pattern that were observed in several SeMNPV field isolates (Caballero *et al.*, 1992) may therefore be correlated with polymorphism in the SeMNPV *hr*'s, since these are centered around a *Bgl*II restriction enzyme site.

The SeMNPV *hr* showed structural homology to the repeat regions of other baculoviruses, which also consist of a mixture of palindromic and direct repeats (Theilmann & Stewart, 1992a, Garcia-Canedo *et al.*, 1996, Guarino & Summers, 1986, Pearson & Rohrmann, 1995, Majima *et al.*, 1993). The conservation is most apparent in the palindromic sequence (Fig. 7.2a). Compared to other baculovirus *hr* palindromes, the SeMNPV palindrome differs in being longer (68 nt versus 28 nt in AcMNPV) but is similar in containing a conserved mismatch in the palindrome. The *hr* sequences of several baculoviruses were shown to be involved in enhancement of (early) gene transcription (Guarino & Summers, 1986, Theilman & Stewart, 1992, Faktor *et al.*, 1997a, Guarino *et al.*, 1986) and some functioned as origin of replication (ori) in an *in vitro* replication assay (Xie *et al.*, 1995, Kool *et al.*, 1993, Pearson *et al.*, 1992, 1993, Leisy & Rohrmann, 1993, Pearson & Rohrmann, 1995, Leisy *et al.*, 1995). As yet it is still unclear if the palindromic structure is important for these functions (Rodems & Friesen, 1995, Rasmussen *et al.*, 1996). The potential enhancer and ori properties of the SeMNPV *hr* are currently under investigation.

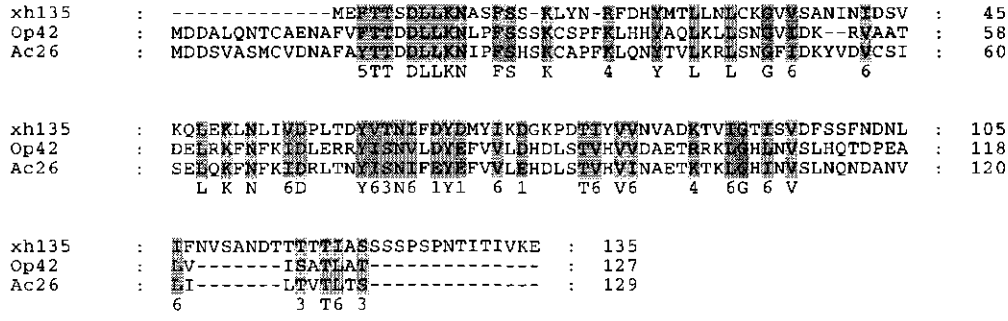


Figure 7.3. Alignment of SeMNPV ORF xh135 with AcMNPV ORF 26 (Ayres *et al.*, 1994, Braunagel *et al.*, 1992) and OpMNPV ORF 42 (Ahrens *et al.*, 1997). Dashes are introduced to optimize the alignment. Gray shading is used to indicate the occurrence of similar amino acids. Consensus amino acids are indicated below the alignment. Abbreviations: 1: D, E; 2: N,Q; 3: S, T; 4: K, R; 5: F, Y, W; 6: L, I, V, M.

ORF xh135

The *hr* region is flanked by an ORF of 135 amino acids in an antigenomic orientation, SeMNPV ORF xh135. This ORF potentially encodes a protein of 15.0 kDa. A baculovirus consensus late promoter motif GTAAG is found 58 nt upstream of the translational start codon. A consensus polyadenylation signal AATAAA is found 58 nt downstream of the translational stop codon. SeMNPV ORF xh135 displays a limited homology to AcMNPV ORF 26 and OpMNPV ORF 42, of 33% and 31%, respectively (Fig. 7.3). The function of this ORF is not yet known.

ORF xh328

SeMNPV ORF xh328 encodes, in a genomic orientation, a putative protein of 35 kDa with limited homology to AcMNPV ORF 25 and OpMNPV ORF 43 (Fig. 7.4). The nucleotide sequence of AcMNPV ORF 25 was published initially to encode two overlapping ORFs (Braunagel *et al.*, 1992), but was identified as a single ORF in a later publication by Ayres *et al.* (1994). The function of this ORF is not yet known. The C terminus of the ORF seems the most conserved part of this protein. SeMNPV ORF xh328 is preceded by a late consensus TTAAG at position -58 with respect to the translational start codon. A baculovirus consensus early promoter sequence CAGT, preceded by an AT rich region, is found at position -32. A consensus polyadenylation signal is found 4 nt downstream of the translational stop codon.

SeMNPV lef6

A SeMNPV homolog of AcMNPV *lef6* is found in a genomic orientation, potentially encoding a protein of 163 amino acids with a molecular mass of 19.1 kDa. SeMNPV lef6 shows a low degree of homology to the AcMNPV (26%) and OpMNPV (28%) lef6 proteins and only a few amino acids appear to be conserved (Fig. 7.5). The N-terminal part of the protein is rather basic and the C-terminal part acidic. SeMNPV *lef6* and ORF xh328, preceding SeMNPV *lef6*, are separated by 28 nt. An ATAAG sequence is present, 98 nt upstream of the putative *lef6* translational start codon, in the sequence encoding the C-terminal part of ORF xh328. No homology to the start sites of AcMNPV *lef6* can be recognized in the region 5' of SeMNPV lef6. A consensus polyadenylation signal is found 5 nt downstream of the translational stop codon.

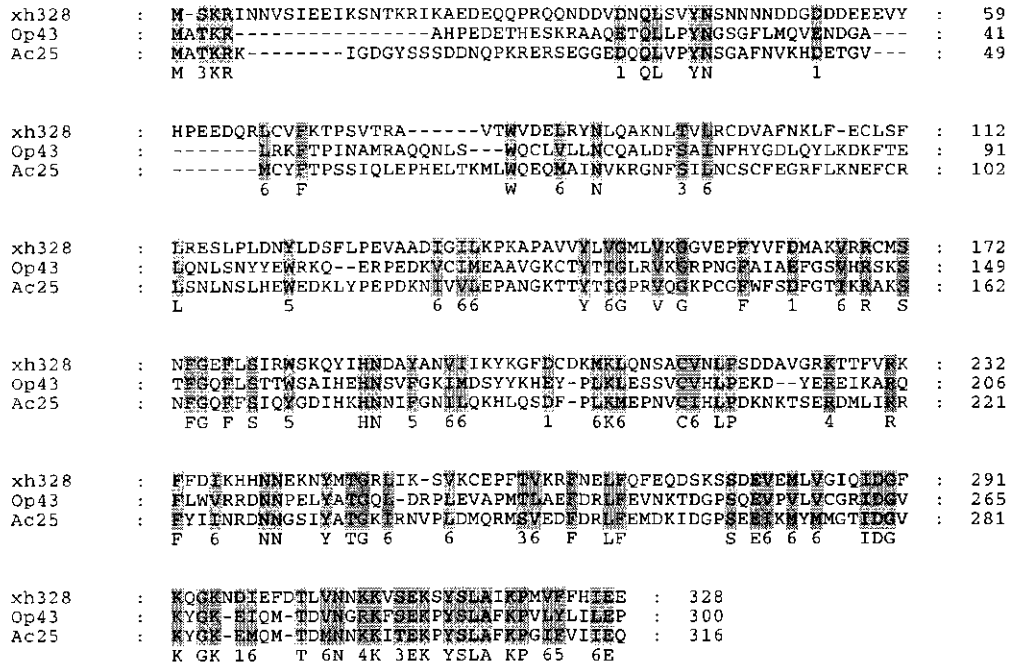


Figure 7.4. Alignment of SeMNPV ORF xh328 with the AcMNPV ORF 25 (Ayres *et al.*, 1994, Braunagel *et al.*, 1992) and OpMNPV ORF 43 (Ahrens *et al.*, 1997). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

The AcMNPV *lef6* gene is one of the 18 genes required for late gene expression and is transcribed early and late in infection (Passarelli & Miller, 1994, Lu & Miller, 1995). The promoter region of AcMNPV *lef6* is located in the sequences coding for the C-terminal part of AcMNPV IAP1 (ORF 27), the ORF that precedes AcMNPV *lef6* by 7 nt (Passarelli & Miller, 1994). In analogy to the situation in AcMNPV, the SeMNPV *lef6* promoter sequences may reside within the sequences encoding the C terminus of ORF xh328. The presence of a *lef6* homolog in SeMNPV suggests that at least some of the genes involved in late gene expression in SeMNPV will be similar as in AcMNPV.

ORF xh136

SeMNPV ORF xh136 is separated from *lef6* by an AT-rich region, which contains a potential polyadenylation site on both strands. This antigenomic ORF potentially encodes a basic protein (predicted pI 11.1) with a predicted molecular mass of 16.3 kDa and is the homolog of AcMNPV ORF 29 and OpMNPV ORF 39 (Fig. 7.6). The function of this ORF is not yet known. A homologous ORF was also identified in *Cydia pomonella* granulovirus (CpGV) (accession nr. U53466). The homology of the CpGV protein to the AcMNPV and OpMNPV ORFs is slightly higher than the homology displayed by SeMNPV xh136. SeMNPV ORF xh136 is longer than its homologs and the additional amino acids of SeMNPV xh136 are located at the N terminus (Fig. 7.6). However, in-frame Met codons are found in the SeMNPV xh136 ORF. The use of one of

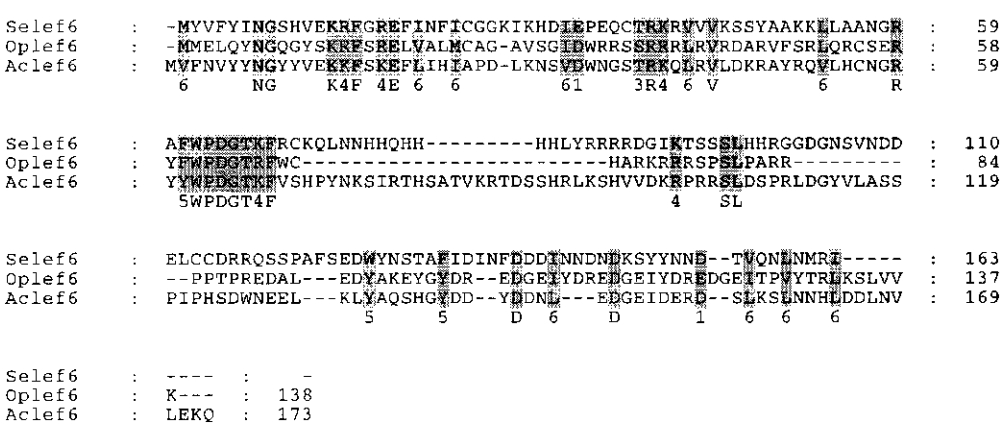


Figure 7.5. Alignment of SeMNPV lef6 with AcMNPV lef6 (Passarelli & Miller, 1994) (AcMNPV ORF 28, Ayres *et al.*, 1994) and OpMNPV ORF 40 (Ahrens *et al.*, 1997). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

these may result in an ORF with a length more similar to the homologous ORFs of other baculoviruses.

SeMNPV p26

A similar situation as with xh136 is observed for the SeMNPV p26, which also shows additional amino acids at the N terminus as compared to other baculovirus p26 proteins (Fig. 7.7). SeMNPV *p26* potentially encodes a protein of 278 amino acids with a predicted molecular mass of 31.1 kDa. Two late promoter sequences ATAAG are found in front of the SeMNPV *p26* ORF, at position -114 and -84 relative to the translational start codon. The role of *p26* in the baculovirus infection cycle is not known. The AcMNPV gene is transcribed early and late in infection (Liu *et al.*, 1986). A similar expression pattern is also observed with SeMNPV *p26* (data not shown). In AcMNPV the *p26* gene as well as the neighbouring homologous region *hr*5 are dispensable for infection of cultured insect cells (Rodems & Friesen, 1993).

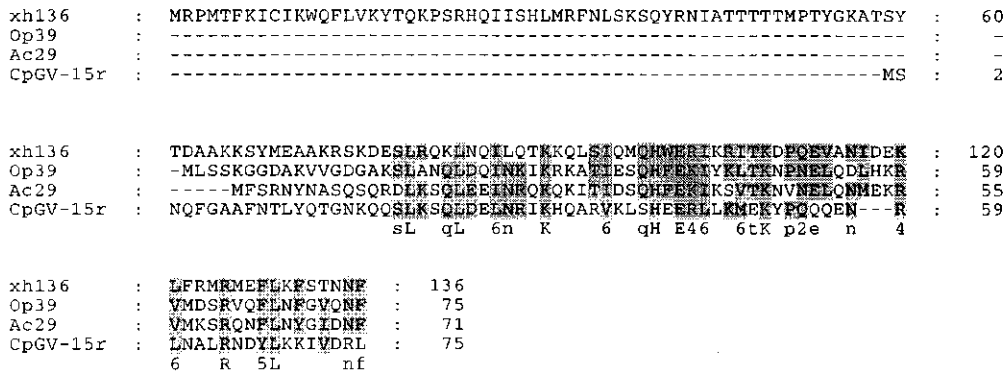


Figure 7.6. Alignment of SeMNPV ORF xh136 with AcMNPV ORF 29 (Ayres *et al.*, 1994), OpMNPV ORF 39 (Ahrens *et al.*, 1997) and *Cydia pommonella* (Cp) GV ORF 15R (acc.nr. U53466). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

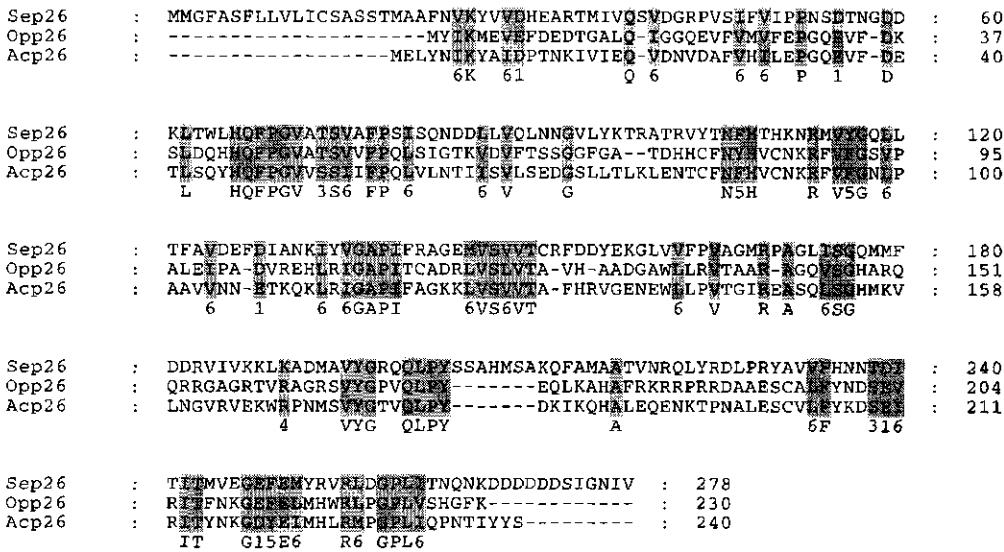


Figure 7.7. Alignment of SeMNPV p26 with AcMNPV p26 (Liu *et al.*, 1986) (AcMNPV ORF 136, Ayres *et al.*, 1994) and OpMNPV p26 (Bicknell *et al.*, 1987) (OpMNPV ORF 132, Ahrens *et al.*, 1997). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

Table 7.2. Pairwise amino acid identity and similarity (Blossum 35) percentages (bold) of ten baculovirus p10 proteins. For abbreviations see materials and methods section.

	LdMNPV		SpliMNPV		OpMNPV		AcMNPV		BmNPV-2
	BusuSNPV		PenuMNPV		CfMNPV		BmNPV-1		
SeMNPV	45	53	39	27	27	26	16	16	18
	62	**66**	**57**	**50**	**50**	**45**	**42**	**42**	**39**
LdMNPV		46	37	22	21	21	23	20	19
		67	**56**	**47**	**47**	**47**	**40**	**41**	42
BusuSNPV			47	29	30	34	20	19	20
			60	**50**	**51**	**47**	**40**	**40**	**37**
SpliMNPV				21	21	24	15	14	15
				39	**40**	**38**	**33**	**33**	**31**
PenuMNPV					94	37	34	32	33
					97	**56**	**57**	**57**	**56**
OpMNPV						40	35	32	32
						57	**58**	**58**	**55**
CfMNPV							40	37	35
							58	**60**	**58**
AcMNPV								87	65
								93	**74**
BmNPV-1									73
									78

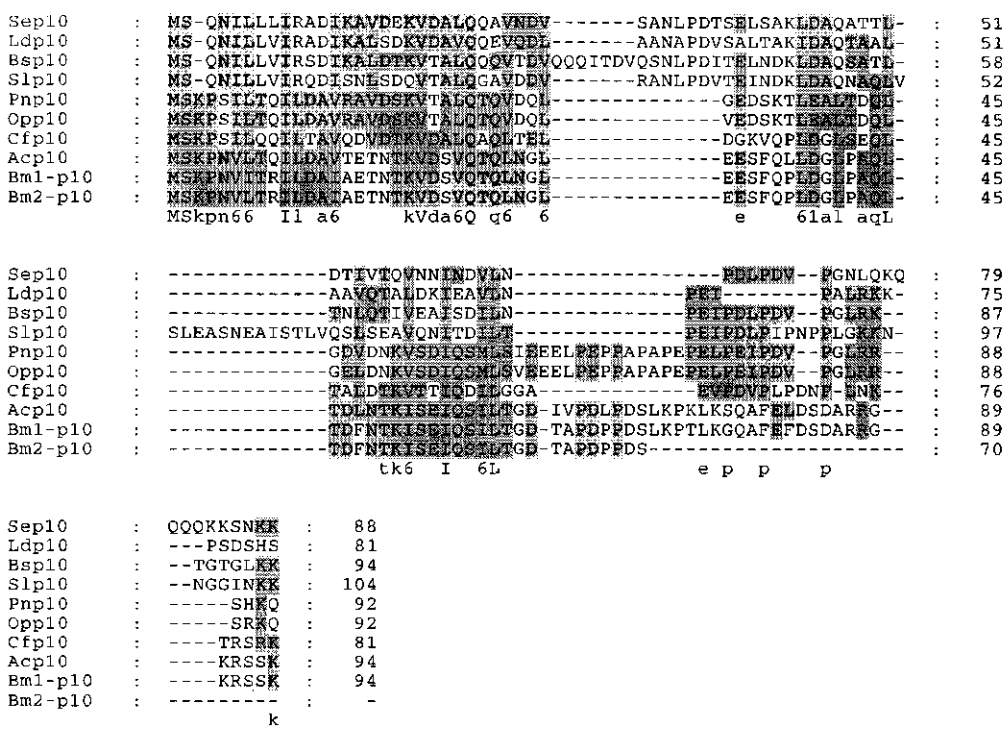


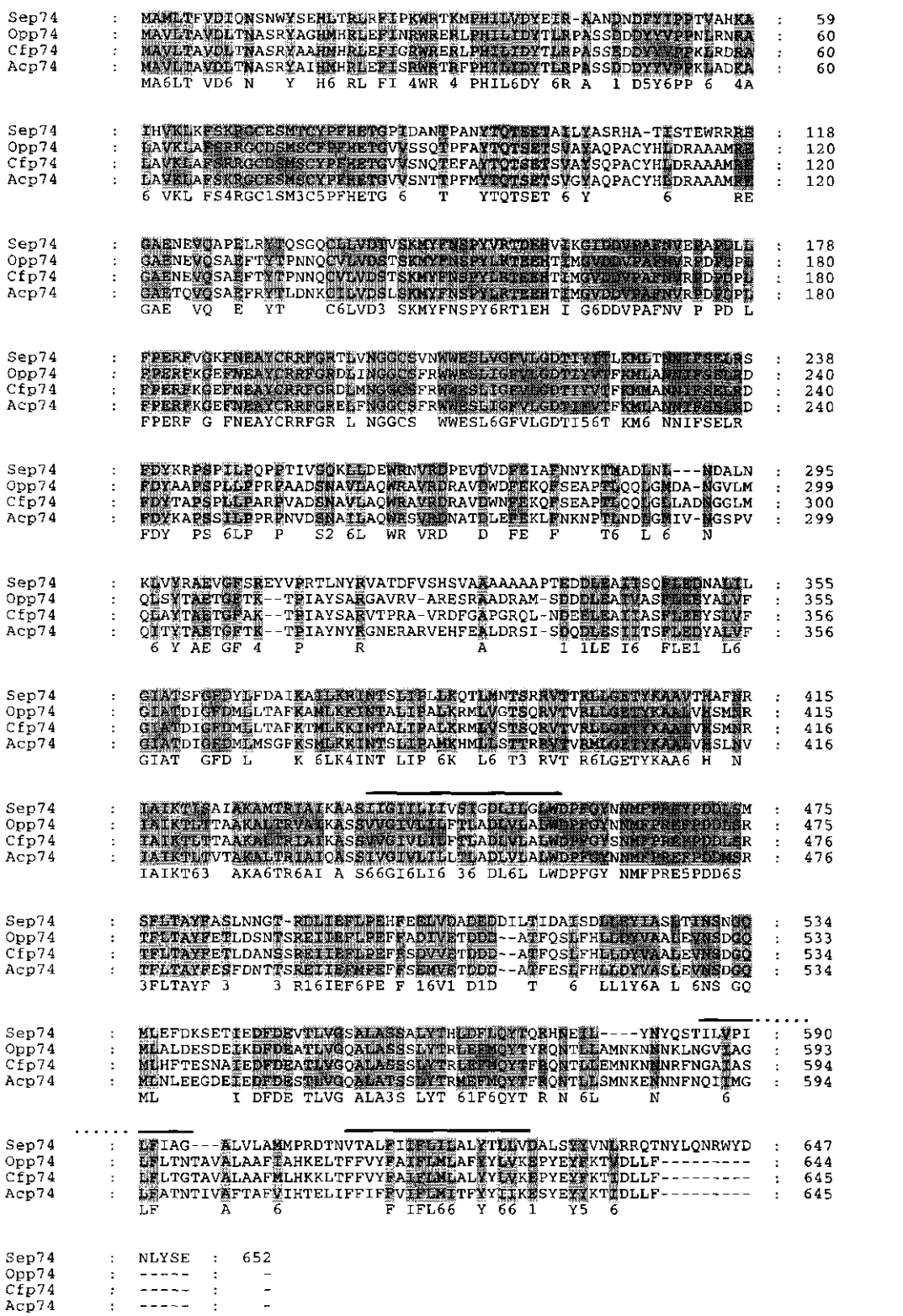
Figure 7.8. Alignment of SeMNPV p10 with the p10 proteins of AcMNPV (Kuzio *et al.*, 1984) (AcMNPV ORF 137, Ayres *et al.*, 1994)), OpMNPV (Leisy *et al.*, 1986c) (OpMNPV ORF 133, Ahrens *et al.*, 1997), CfMNPV (Wilson *et al.*, 1995), PenuMNPV (Chou *et al.*, 1992), SpliMNPV (Faktor, 1997b), LdMNPV (G.F. Rohrmann, *pers. comm.*) and BusuSNPV (Z.H. Hu, *pers.comm.*). BmNPV1 is a full length p10 of BmNPV (Zhang, 1992) and BmNPV2 is a truncated BmNPV p10 protein (Hu *et al.*, 1994). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

SeMNPV p10

The SeMNPV *p26* gene is flanked by the baculovirus major late *p10* gene, showing the same genomic orientation as *p26*. The baculovirus p10 protein, dispensable during infection, is involved in the formation of fibrillar structures, in polyhedron envelope morphogenesis and the desintegration of the nuclear membrane at the end of the infection (van der Wilk *et al.*, 1987, van Oers *et al.*, 1994, van Oers & Vlak, 1997, Russell *et al.*, 1991, Gross *et al.*, 1994, Lee *et al.*, 1996). The SeMNPV *p10* gene is discussed elsewhere (Chapter 3 and Zuidema *et al.*, 1993). Following this publication, several other *p10* genes were identified. Alignment of amino acid sequences (Fig. 7.8, Table 7.2) shows that SeMNPV p10 is most similar to p10 of *Buzuria suppressaria* SNPV (Z. Hu, *pers.comm.*), *Lymantria dispar* MNPV (G.F. Rohrmann, *pers.comm.*) and *Spodoptera littoralis* MNPV (Faktor *et al.*, 1997b). The p10 proteins of these four baculoviruses are more homologous to each other than to p10 proteins of AcMNPV, OpMNPV, *Bombyx mori* NPV, *Choristoneura fumiferana* (Cf) MNPV or *Perina nuda* MNPV.

SeMNPV p74

The SeMNPV *p74* gene potentially encodes a protein of 652 amino acids in length with a



predicted molecular mass of 74.3 kDa. The SeMNPV *p74* gene product has a high degree of homology to the *p74* gene products of AcMNPV, OpMNPV and CfMNPV(Fig. 7.9). The baculovirus consensus late promoter sequence TTAAG is found at position -13 with respect to the translational start codon. A baculovirus consensus early promoter CAGT sequence, preceded by a TATA box, is found at position -42.

Comparison of the SeMNPV p74 amino acid sequence with the p74 ORFs of AcMNPV, OpMNPV and CfMNPV reveals the presence of many conserved amino acids, including several Cys residues in the N-terminal half of the protein (Fig. 7.9). Multiple hydrophobic regions can be recognized in the C-terminal part of the protein, which are characteristic for transmembrane domains. SeMNPV p74 has a 15 amino acid C-terminal extension as compared to the other baculovirus p74 proteins. Several potential N-linked glycosylation sites can be recognized but only one, at amino acid position 485, is conserved in all p74 proteins. AcMNPV p74 is not glysolyated (T. Roberts in Rohrmann, 1992). The AcMNPV *p74* gene is dispensable during infection of cultured cells, but is essential for virus infectivity for insect larvae through the oral infection route. AcMNPV *p74* is expressed late in infection, but the consensus late baculovirus motif TAAG was not used to initiate the *p74* transcription (Kuzio, 1989). SeMNPV *p74* is also expressed late in infection (data not shown). The SeMNPV and AcMNPV *p74* 5' regions show no significant homology.

SeMNPV ie1

The SeMNPV *ie1* gene flanks the SeMNPV *p74* gene and shows a similar antigenomic orientation as *p74*. SeMNPV *ie1* is described extensively elsewhere (Chapter 5 and van Strien *et al.*, 1997b).

ORF xd200

SeMNPV ORF xd200, with a genomic orientation, potentially encodes a protein with a molecular mass of 22.4 kDa. A baculovirus consensus late promoter motif GTAAG is found 77 nt upstream of the translational start codon. This motif is thus located within the sequences encoding the SeMNPV IE1 N terminus. Homologs of the SeMNPV ORF xd200 are found immediately upstream of the *ie1* gene in all baculoviruses from which the *ie1* gene has been identified. SeMNPV ORF xd200 shows the highest homology to AcMNPV ORF 146. However, only a few conserved amino acids can be recognized (Fig. 7.10). The function of this ORF is not yet known.

ORF xd92

SeMNPV ORF xd92 potentially encodes a small Cys-rich ORF of 92 amino acids in an anti-genomic orientation. A baculovirus late promoter motif TTAAG is located 91 nt in front of the translational start codon and a consensus polyadenylation signal is found 9 nt downstream of the translational stop codon. ORF xd92 shows a relatively high degree of homology to AcMNPV ORF 145 and OpMNPV ORF 142. The ORFs of SeMNPV and OpMNPV have a hydrophobic

Figure 7.9. Alignment of SeMNPV p74 with the p74 proteins of AcMNPV (Kuzio *et al.*, 1989) (AcMNPV ORF 138, Ayres *et al.*, 1994), OpMNPV ORF 134 (Ahrens *et al.*, 1997) and CfMNPV (Hill *et al.*, 1993). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3. Putative transmembrane regions are overlined.

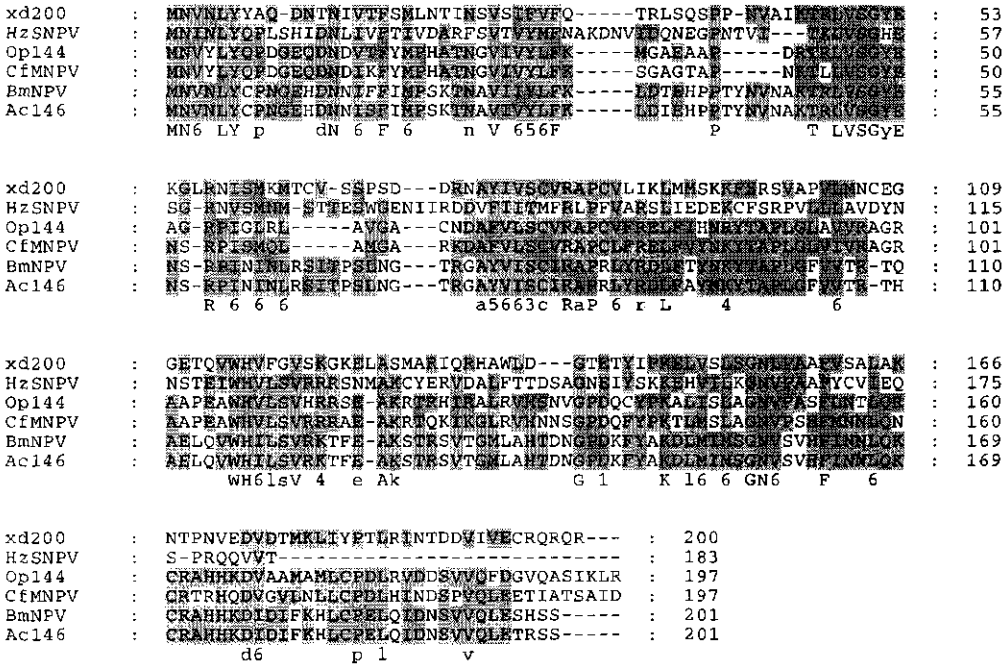


Figure 7.10. Alignment of SeMNPV ORF xd200 with the homologous ORFs from AcMNPV (ORF 146, Ayres *et al.*, 1994), OpMNPV (ORF 143, Ahrens *et al.*, 1997, Theilmann & Stewart, 1991), BmNPV (Huybrechts *et al.*, 1994), HzSNPV (Cowan *et al.*, 1994) and CfMNPV (acc.nr. L04945). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

N terminus which is absent from AcMNPV ORF 145 (Fig. 7.11). This ORF might be present twice in the AcMNPV genome, since it can also be aligned with AcMNPV ORF 150. Despite the presence of several conserved Cys residues, this ORF shows no significant homology to sequences in the data bank.

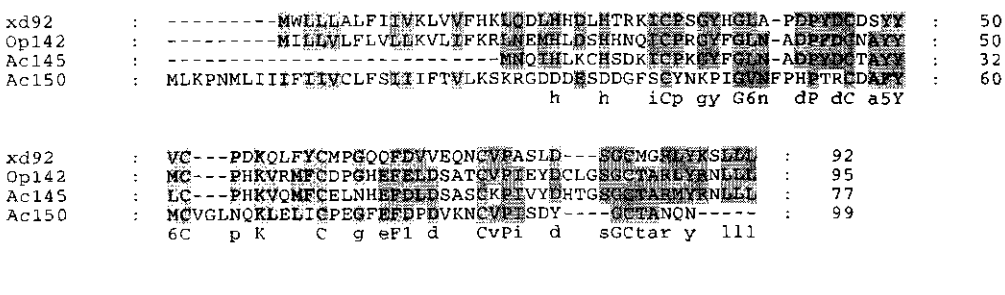


Figure 7.11. Alignment of SeMNPV ORF xd92 with the homologous ORFs from AcMNPV (ORF 145 and ORF 150, Ayres *et al.*, 1994), OpMNPV (ORF 141, Ahrens *et al.*, 1997). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

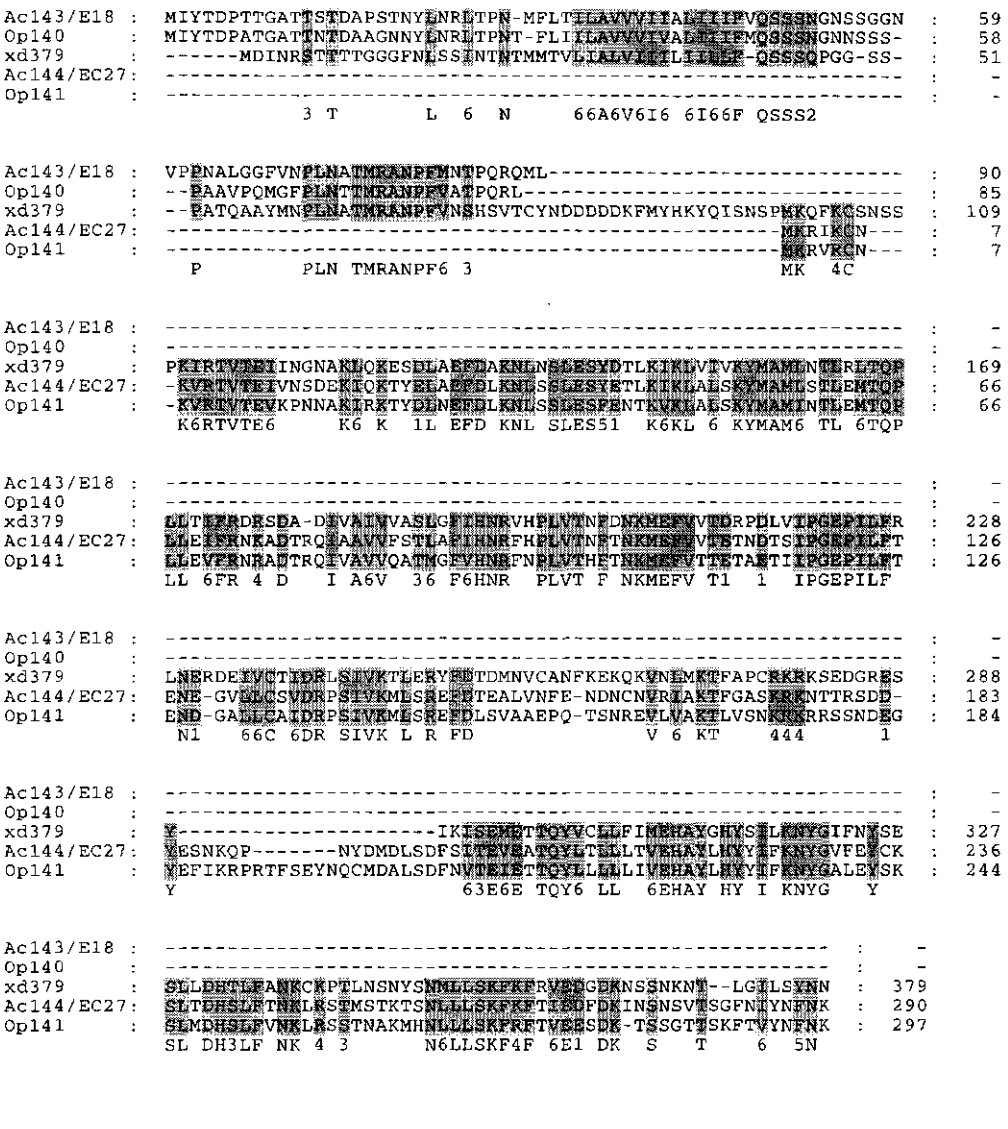


Figure 7.12. Alignment of SeMNPV ORF xd379 with the homologous ORFs from AcMNPV and OpMNPV. The N terminus of SeMNPV xd379 is aligned with AcMNPV ODV-E18 (Braunagel *et al.*, 1996) (AcMNPV ORF 143, Ayres *et al.*, 1994) and OpMNPV ORF 140 (Ahrens *et al.*, 1997). (ODV-E18, counting 90 amino acids, from Braunagel *et al.* (1996) is different from AcMNPV ORF 143, counting 62 amino acids. ODV-E18 has 28 additional amino acids at the N terminus as compared to AcMNPV ORF 143).
The C terminus of SeMNPV xd379 is aligned with AcMNPV ODV-EC27 (Braunagel *et al.*, 1996) (AcMNPV ORF 144, Ayres *et al.*, 1994) and OpMNPV ORF 145 (Ahrens *et al.*, 1997). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

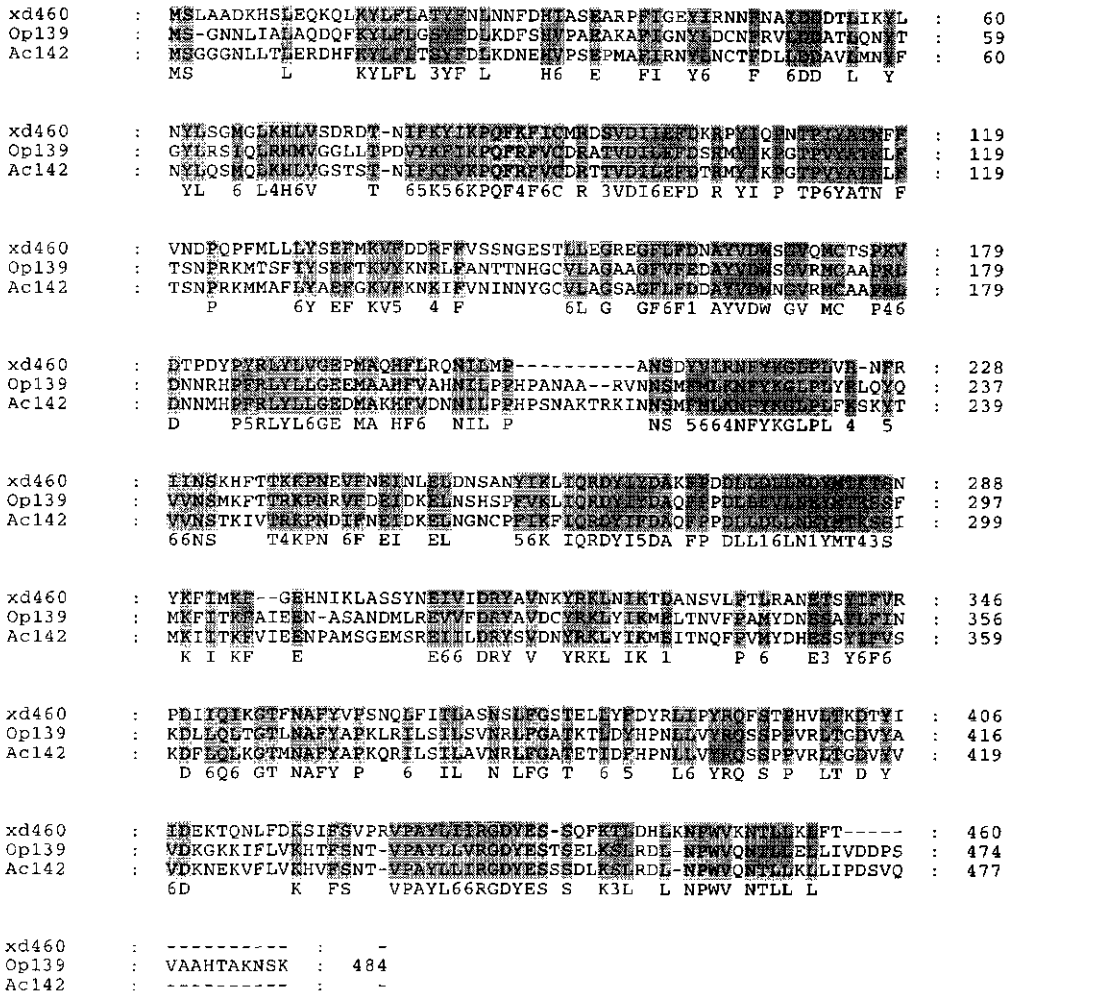


Figure 7.13. Alignment of SeMNPV ORF xd460 with the homologous ORFs from AcMNPV (ORF 142, Ayres *et al.*, 1994) and OpMNPV (ORF 139, Ahrens *et al.*, 1997). For shading, amino acid consensus and abbreviations see legend of Fig. 7.3.

ORF xd379

SeMNPV ORF xd379 potentially encodes a protein with a predicted molecular mass of 43.1 kDa. A baculovirus consensus late promoter ATAAG sequence is located 130 nt in front of the translational start codon. A polyadenylation signal is found 12 nt downstream of the translational stop codon.

SeMNPV ORF xd379, a single ORF in SeMNPV fragment XbaI-D, is homologous to two consecutive ORFs in AcMNPV and in OpMNPV. The N terminus of SeMNPV ORF xd379 shows homology to AcMNPV ORF 143 and to OpMNPV ORF 140; the C terminus shows homology to AcMNPV ORF 144 and to OpMNPV ORF 141 (Fig. 7.12).

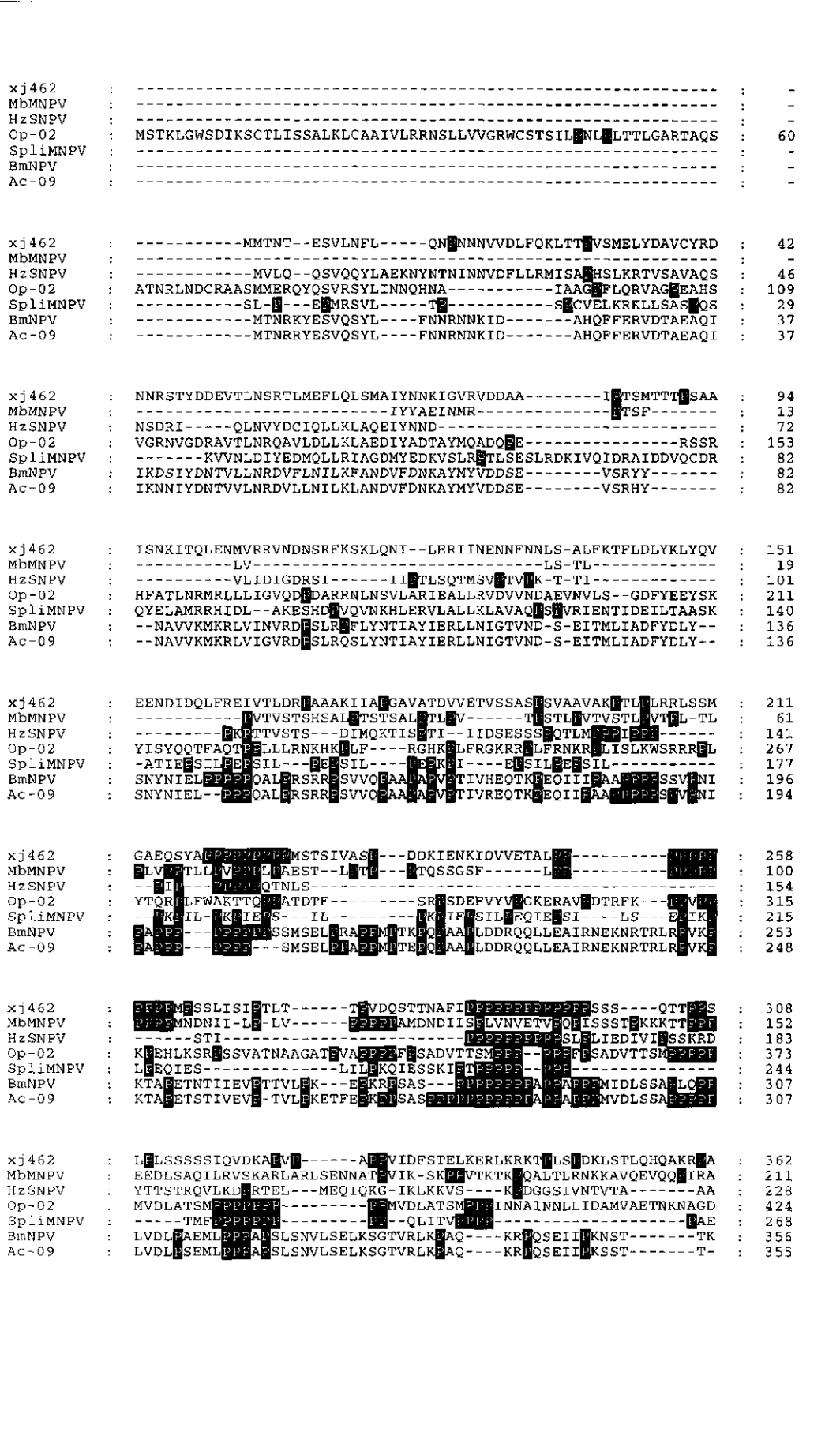
AcMNPV ORF 143 and ORF 144 were shown to encode occlusion derived viral envelope proteins of 18 kDa (ODV-E18) and 27 kDa (ODV-EC27), respectively. ODV-EC27 was also identified in the capsids of the occlusion derived virions (Braunagel *et al.*, 1996). In addition, this area produced a structural protein of 35 kDa, ODV-E35, which reacted with antibodies directed against ODV-E18 and with antibodies directed against ODV-EC27. The origin of ODV-E35 remained somewhat enigmatic and was therefore suggested to be the result of a ribosomal frameshift in the translation of a 2.0 kb transcript that encompassed AcMNPV ODV-E18 and ODV-EC27 (Braunagel *et al.*, 1996). In SeMNPV ORF xd379 no "in frame" stop codon is observed.

Alignment of SeMNPV ORF xd379 with AcMNPV ODV-EC27 and OpMNPV ORF 141 shows that the SeMNPV ORF xd379 sequence encodes an internal Met codon (Met-99) which coincides with the AcMNPV and OpMNPV translational start codons (Fig. 7.12). One of the late transcripts produced from this area originated from a consensus TAAG sequence located 15 nt in front of the translational start of AcMNPV ODV-EC27 (Braunagel *et al.*, 1996). SeMNPV xd379 has a TAAG sequence 25 nt in front of the homologous, corresponding Met-99 codon (data not shown). It is thus possible that SeMNPV uses this promoter motif to produce a late transcript, resulting in a protein with a size similar (280 amino acids, predicted molecular mass 32.3 kDa) to AcMNPV ODV-EC27.

A small protein, homologous to AcMNPV ODV-E18, would require processing of the full length translation product of SeMNPV ORF xd379. AcMNPV ODV-E18 transcripts were found to start from three late promoter sites, located 299, 175 and 110 nt upstream of the translational start codon of ODV-E18 (Braunagel *et al.*, 1996). These transcriptional start sites all reside within the coding sequence of the preceding AcMNPV ORF 142. The SeMNPV promoter sequence ATAAG at position -130, located in the coding sequence of SeMNPV ORF xd460 (see below), is homologous to the AcMNPV ATAAG at position -110 (data not shown). It is likely that SeMNPV will use this promoter motif for the late expression of ORF xd379. Transcriptional and translational analysis should clearify the actual coding capacity of this SeMNPV region.

ORF xd460

SeMNPV ORF xd460 is 460 amino acids long and potentially encodes a protein of 50 kDa in an antigenomic orientation. A late promoter sequence ATAAG is located at position -12 with respect to the translational start site and a GTAAG sequence is located at position -76. This ORF has a relatively high degree of homology to AcMNPV ORF 142 and OpMNPV ORF 139 (Fig. 7.13). The function of this ORF is not yet known.



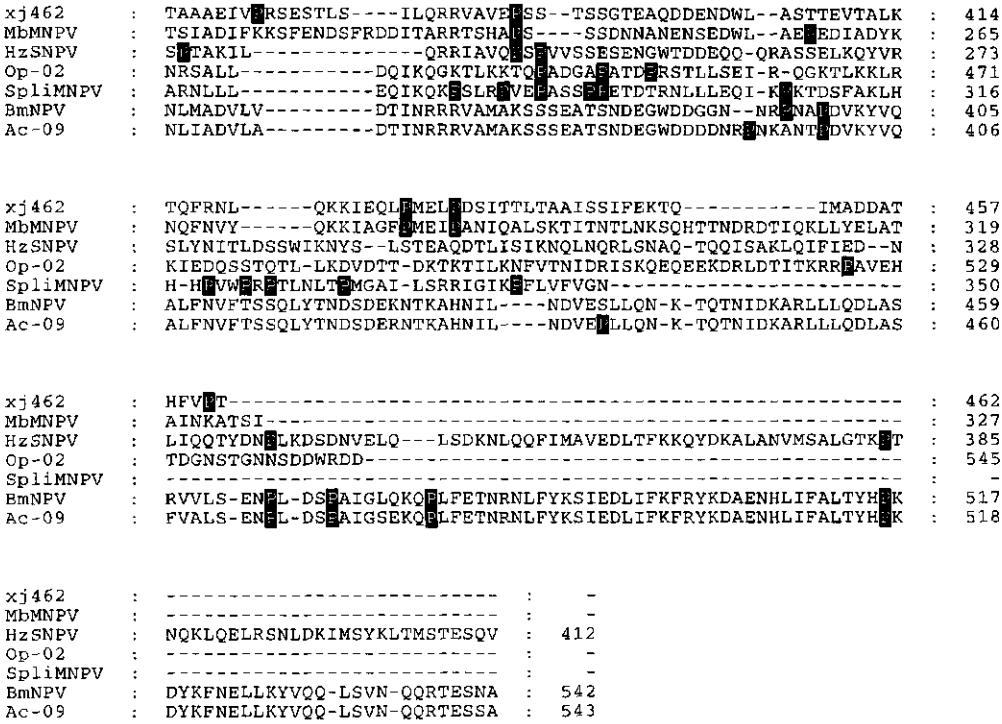


Figure 7.14. Alignment of SeMNPV ORF xj462 with ORFs from AcMNPV (Possee *et al.*, 1991) (AcMNPV ORF 9, Ayres *et al.*, 1994), OpMNPV (Russell *et al.*, 1996) (OpMNPV ORF 2, Ahrens *et al.*, 1997), BmNPV (acc.nr. L33180), HzSNPV (Cowan *et al.*, 1994), SpliMNPV (Faktor *et al.*, 1997b) and a partial sequence from MbMNPV (Cameron & Possee, 1989). Proline residues are indicated in black.

ORF xd244

SeMNPV ORF xd244 has homology to the region encoding the *ie0* left exon in AcMNPV and OpMNPV, but lacks the consensus splice donor sequence. SeMNPV ORF xd244 has been discussed elsewhere in detail (Chapter 5).

RR1

The SeMNPV ORF for the large subunit of ribonucleotide reductase (RR1) (Chapter 6, van Strien *et al.*, 1997a), located upstream of the polyhedrin gene, is encoded by many baculoviruses but not by AcMNPV (Ayres *et al.*, 1994). In OpMNPV, an abberant RR1 (ORF 32) is found at map unit 20. OpMNPV RR1 has a very low degree of homology with eukaryotic and baculoviral RR1's and its length is remarkably short. However, the amino acids thought to be involved in ribonucleotide reduction are conserved (Ahrens *et al.*, 1997). Phylogenetic analysis indicated that OpMNPV, SeMNPV and SpliMNPV acquired the RR1 gene independent from each other (Chapter 6, van Strien *et al.*, 1997a).

polyhedrin

The polyhedrin gene was the first characterized gene of SeMNPV and has been discussed elsewhere (Chapter 2 and van Strien *et al.*, 1992). The polyhedrin gene is the gene displaying the highest degree of homology with other baculoviruses, as compared to the generally much lower extent of homology displayed by other genes (Table 7.1).

ORF xj462

The SeMNPV polyhedrin gene is flanked by SeMNPV ORF xj462 in an antigenomic orientation. ORF xj462 is preceded by an ATAAG motif at position -39, located within the SeMNPV protein kinase (PK) coding region (see below). A consensus polyadenylation signal is located 61 nt downstream of the translational stop codon.

ORF xj462, potentially encoding a protein with a molecular mass of 50.4 kDa, shows very limited identity to baculovirus homologs of AcMNPV ORF 9 (ORF 1629). The multiple sequence alignment (Fig. 7.14) shows that a moderately conserved area is located in the N-terminal part of the protein (not indicated in the figure). Furthermore, the SeMNPV protein and its homologs from other baculoviruses are notable for the presence of multiple poly-proline tracts. SeMNPV xj462 displays the highest homology to an (in-complete) ORF identified downstream of the *Mamestra brassicae* (Mb) MNPV polyhedrin gene (Cameron & Possee, 1989).

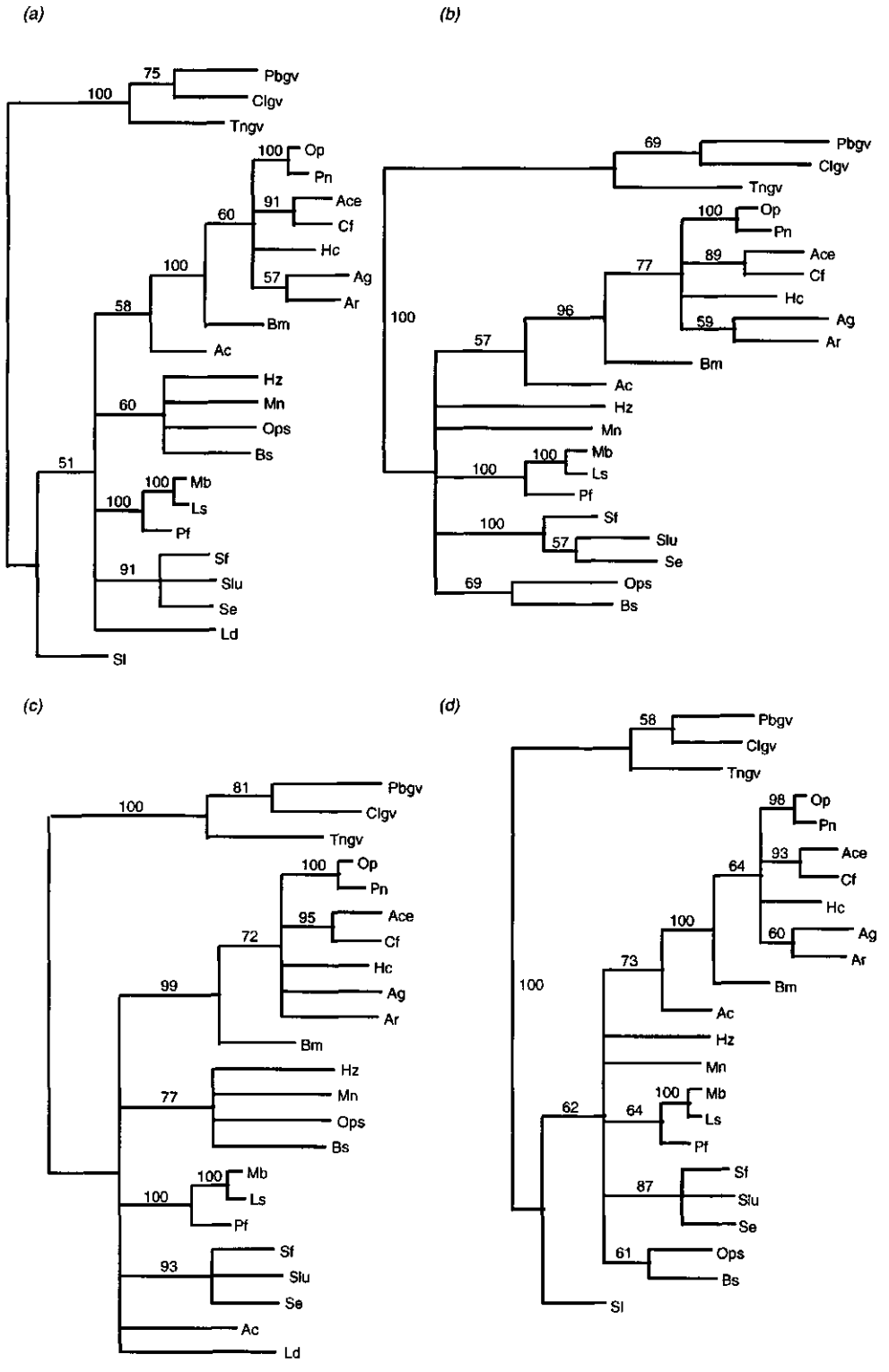
AcMNPV ORF 1629 encodes an essential phoshorylated protein, expressed late in infection. This protein was present in both budded and occluded virions of AcMNPV (Pham & Sivasubramanian, 1992, Pham *et al.*, 1993, Possee *et al.*, 1991, Kitts & Possee, 1993, Vialard & Richardson, 1993). The homologous protein from OpMNPV was shown to be specifically associated with one end of the viral capsid, the basal structure region (Russell *et al.*, 1997). Databank searches with SeMNPV ORF xj462 resulted in homology to a wide array of proline-rich cellular proteins with diverse functions.

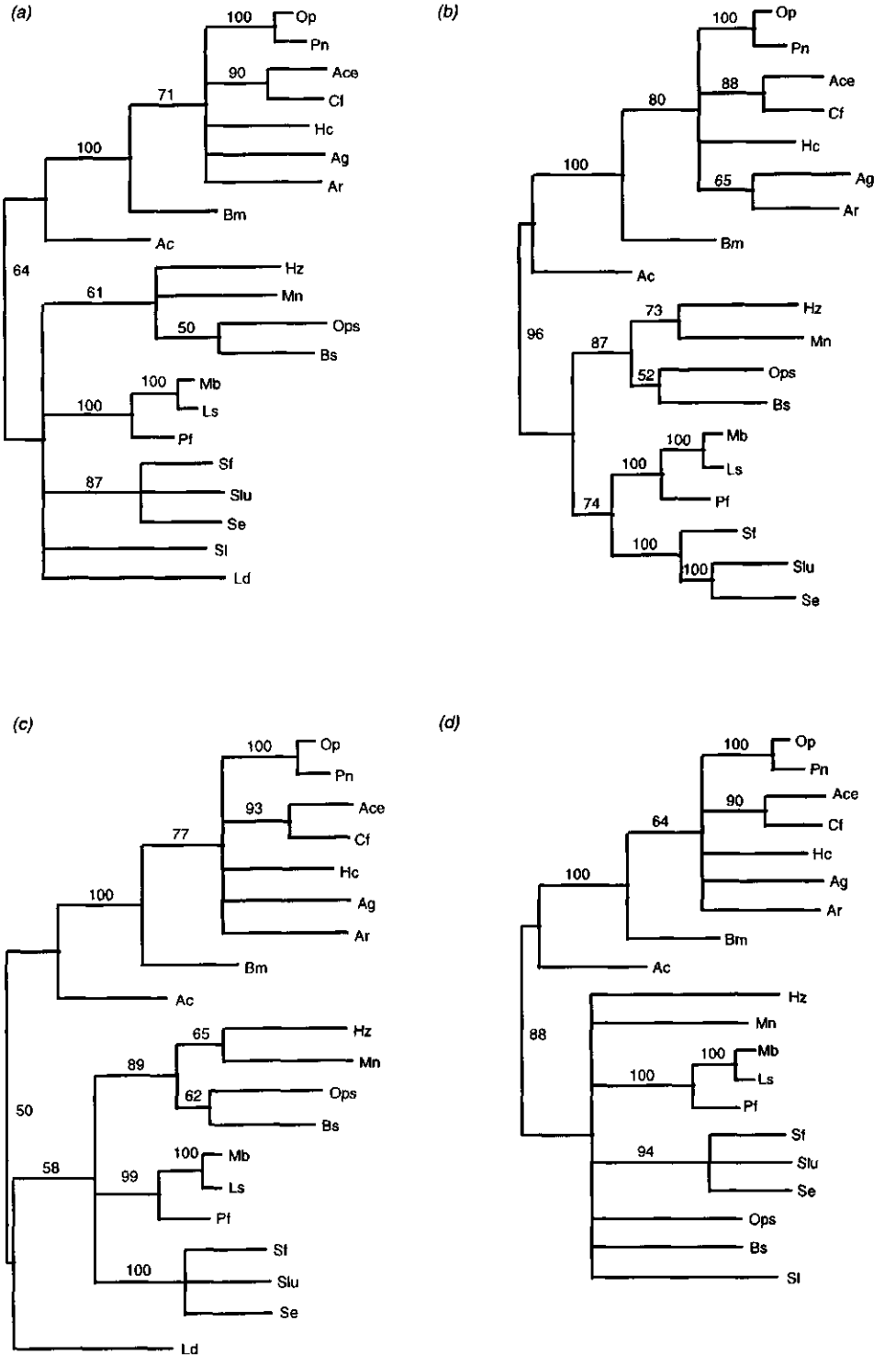
PK

The translational start codon of ORF xj462 overlaps with the translational start codon of the SeMNPV protein kinase (*pk*) gene, which has a genomic orientation. Only the first 103 amino acids encoded by the SeMNPV *pk* gene are determined. Two late promoter motifs are recognized, a GTAAG sequence located at position -188 and a ATAAG sequence located at position -142 with respect to the translational start codon. Both motifs reside within the ORF xj462 coding sequence.

The partially determined SeMNPV PK sequence shows 41% amino acid identity to AcMNPV protein kinase. Protein kinase genes have been identified in several baculoviruses on a similar genomic position relative to the polyhedrin gene as SeMNPV *pk*. The baculovirus PKs

Figure 7.15. a). Bootstrap analysis (100 replicates) of an unrooted phylogenetic tree of polyhedrin genes constructed with the heuristic search algorithm using parsimony (PAUP). Numbers at the branches indicate frequency of cluster. The nucleotide sequence alignment of occlusion body proteins produced by the GCG pileup program was adjusted by hand to be in accordance with the amino acid coding capacity (data not shown). The first 15 nt, which could not be aligned with certainty, were omitted from the analysis. To start as unbiased as possible no outgroup was defined. The three granulin genes were included to orient the tree, but the subgrouping of the granuloviruses was further ignored. For abbreviations of virus names see materials en methods section. b). Similar as in a), but without LdMNPV and SpliMNPV. c). Similar as in a), but without SpliMNPV. d). Similar as in a), but without LdMNPV. content (Smith *et al.*, 1988).


(a)
75
100
Pbgv
Clgv
Tngv
100
Op
Pn
91
Ace
Cf
60
Hc
100
57
Ag
Ar
58
Bm
Ac
Hz
60
Mn
Ops
Bs
51
100
Mb
100
Ls
Pf
Sf
91
Slu
Se
Ld
Sl
(b)
69
Pbgv
Clgv
Tngv
100
Op
Pn
89
Ace
77
Cf
Hc
96
59
Ag
100
Ar
57
Bm
Ac
Hz
Mn
100
Mb
100
Ls
Pf
100
Sf
57
Slu
Se
69
Ops
Bs
(c)
81
100
Pbgv
Clgv
Tngv
100
Op
Pn
95
Ace
72
Cf
Hc
99
Ag
Ar
Bm
Hz
77
Mn
Ops
Bs
100
Mb
100
Ls
Pf
Sf
93
Slu
Se
Ac
Ld
(d)
58
Pbgv
Clgv
Tngv
98
Op
Pn
93
Ace
64
Cf
Hc
100
60
Ag
Ar
73
100
Bm
Ac
Hz
Mn
100
Mb
62
64
Ls
Pf
Sf
87
Slu
Se
61
Ops
Bs
Sl


(a)
100
Op
Pn
90
Ace
71
Cf
Hc
100
Ag
Ar
Bm
Ac
64
Hz
61
Mn
50
Ops
Bs
100
Mb
100
Ls
Pf
Sf
87
Slu
Se
Sl
Ld
(b)
100
Op
Pn
88
Ace
80
Cf
Hc
100
65
Ag
Ar
Bm
Ac
73
Hz
96
Mn
87
52
Ops
Bs
100
Mb
100
Ls
74
Pf
Sf
100
100
Slu
Se
(c)
100
Op
Pn
93
Ace
77
Cf
Hc
100
Ag
Ar
Bm
Ac
65
Hz
89
Mn
62
Ops
50
Bs
100
Mb
58
99
Ls
Pf
Sf
100
Slu
Se
Ld
(d)
100
Op
Pn
90
Ace
64
Cf
Hc
100
Ag
Ar
Bm
Ac
Hz
88
Mn
100
Mb
100
Ls
Pf
Sf
94
Slu
Se
Ops
Bs
Sl

display similarity to serine-threonine protein kinases (Reilly & Guarino, 1994, Bischoff & Slavicek, 1994). Whereas the AcMNPV *pk* gene was expressed late in infection (Reilly & Guarino, 1994), the expression of the LdMNPV *pk* gene started early (Bischoff & Slavicek, 1994). The protein kinase of AcMNPV is involved in very late gene expression (Fan *et al.*, 1996).

Baculovirus polyhedrin gene phylogeny revisited

The polyhedrin gene, the most conserved gene among the baculoviruses, was the first gene used for the reconstruction of baculovirus phylogeny (Vlak & Rohrmann, 1985, Zanotto *et al.*, 1993a, Cowan *et al.*, 1994). Following these publications the number of determined polyhedrin sequences has roughly doubled. The available sequences were all used in parsimony analysis (Swofford, 1993) in an attempt to update the baculovirus phylogeny. However, the high homology of the polyhedrin genes prompted the need for a careful assessment of the deduced phylogeny. Just as sequences can be too diverged to be of use in phylogenetic reconstructions (Zanotto *et al.*, 1996), the homology can also be too high to offer sufficiently informative phylogenetic characters. The reliablity of the deduced phylogeny was determined by bootstrap resampling (Felsenstein, 1993).

A heuristic search with the amino acid alignment of the 28 available baculovirus occlusion body protein sequences resulted in 180 most parsimonous trees (data not shown). Bootstrapping with the amino acid alignment was therefore not further pursued and the nucleotide sequence alignment was used instead. The use of parsimony with the aligned nucleotide sequences of the 24 available species seemed well-justified, since the divergence among the polyhedrin genes had not fully reached the complete level of codon redundancy. The sequences of the granuloviruses were included to orient the polyhedrin tree. Heuristic searches with the aligned nucleotide sequences resulted in one to five most parsimonous trees, depending on the combination of species used (data not shown).

The result of the bootstrapping procedure with the aligned nucleotide sequences is depicted in Fig. 7.15a. Several polytomies are seen, which are due to unsupported relationships. A distinct subgroup was formed by AcMNPV, BmNPV, OpMNPV, CfMNPV, HycuMNPV, ArceMNPV, AgMNPV, PenuMNPV and ArMNPV (for abbreviations see materials and methods section). This group was previously recognized as NPV subgroup I by Zanotto *et al.* (1993a). New members are ArceMNPV, PenuMNPV and ArMNPV. Bootstrapping supported the admission of this group as distinct from the other NPVs, although the phylogeny within the subgroup was not completely resolved. The existence of subgroup II (A and B) (Zanotto *et al.*, 1993a, Cowan *et al.*, 1994), consisting of SeMNPV, SpfrMNPV, MbMNPV, PaflMNPV, BusuSNPV, HzSNPV, SpliMNPV, ManeNPV and OpSNPV, was not supported by the bootstrap

Figure 7.15. a). Bootstrap analysis (100 replicates) of an unrooted phylogenetic tree of polyhedrin genes constructed with the heuristic search algorithm using parsimony (PAUP). Numbers at the branches indicate frequency of cluster. The nucleotide sequence alignment of occlusion body proteins produced by the GCG pileup program was adjusted by hand to be in accordance with the amino acid coding capacity (data not shown). The first 15 nt, which could not be aligned with certainty, were omitted from the analysis. To start as unbiased as possible no outgroup was defined. The three granulin genes were included to orient the tree, but the subgrouping of the granuloviruses was further ignored. For abbreviations of virus names see materials en methods section. b). Similar as in a), but without LdMNPV and SpliMNPV. c). Similar as in a), but without SpliMNPV. d). Similar as in a), but without LdMNPV. content (Smith *et al.*, 1988).

analysis. The SNPVs formed a distinct subgroup (the phenotype of MnNPV is not known). They may thus constitute a monophyletic subgroup, although relationships within this subgroup and with the other NPVs are insecure.

The phylogeny and bootstrapping in Fig. 7.15a showed that the relationship between LdMNPV, SpliMNPV and the other NPVs is unclear, but LdMNPV and SpliMNPV do not reside within subgroup I. SpliMNPV as well as LdMNPV may be the single representatives of distinct subgroups, a view also supported by other observations. LdMNVP appeared to be rather different from the other NPVs, as judged from its larger genome which also has a relatively high (60%) GC content (Smith *et al.*, 1988). SpliMNPV appeared to be distantly related to other NPVs as seen in phylogenetic reconstructions based on the ecdysteroid UDP-glucosyltransferase (*egt*) gene (Clarke *et al.*, 1996, Hu *et al.*, 1997).

To further investigate the observed unclear relationships, bootstrapping was performed with the same data set excluding LdMNPV and/or SpliMNPV (Fig. 7.15b-d). The granulin genes were excluded from the data set to determine if the phylogeny might be better supported without them (Fig. 7.16a-d). Heuristic searches while in- or excluding species resulted in different parsimonous trees (data not shown) for which varying support was observed during bootstrapping (compare Fig. 7.15a-d, 7.16a-d). For instance, only when the granuloviruses, SpliMNPV and LdMNPV are excluded (Fig. 7.16d), subgroup II as defined by Zanotto *et al.* (1993a) could be recognized as a well supported branch. Data sets with other species did not support the presumed subgroup II (Fig. 7.15a-d, 7.16a-c). However, the viruses SeMNPV / SpfrMNPV /SpltMNPV and MbMNPV / PafiMNPV / LsMNPV are probably all quite closely related, based on comparison of sequences surrounding the polyhedrin gene (Chapter 2, 6, van Strien *et al.*, 1993, 1997a), but they were otherwise never observed to occur as a distinct, well supported subgroup. The SNPV classification, based on virion occlusion morphology, was sometimes supported by the bootstrapping, but the relationship with other NPVs remained unclear. Unresolved polytomies always existed within subgroup I. Exclusion of AcMNPV and BmNPV did not improve the supported phylogeny within this branch or in general (data not shown).

When the third nucleotide of each codon from the aligned nucleotide sequences was excluded, the resulting tree topology remained unstable and many branches were unsupported in the bootstrapping analysis (data not shown), similar to the analyses with all nucleotides included.

The available data considering the relationship between AcMNPV and BmNPV are in conflict with the phylogenetic reconstructions based on the polyhedrin gene, since these viruses were never subgrouped to share the most recent common ancestor (Cowan *et al.*, 1994, Zanotto *et al.*, 1993a, Fig. 7.15a-d, 7.16a-d). However, a close relationship between AcMNPV and BmNPV was observed in phylogenetic reconstructions based on the EGT (Clarke *et al.*, 1996, Hu *et al.*, 1997) and the immediate early 1 (IE1) sequence (Chapter 5). Furthermore, the complete genome sequence from AcMNPV (Ayres *et al.*, 1994), BmNPV (acc.nr. L33180) and OpMNPV (Ahrens *et al.*, 1997) is available and serves the estimation of the phylogenetic distance between these viruses. AcMNPV and BmNPV show a high overall homology and are thus very closely related. In this particular case the conflicting polyhedrin based phylogeny may be interpretated as the result of (partial) polyhedrin gene transfer into either AcMNPV (or BmNPV) from another baculovirus.

DISCUSSION

Gene sequences for phylogenetic reconstructions

In the 20 kb DNA segment, representing 15% of the SeMNPV genome, twenty major ORFs and an *hr* were identified. All these SeMNPV ORFs were also present in AcMNPV (except RR1) and in OpMNPV (Table 7.1). The ORFs of these viruses usually show a limited degree of homology (60% or less), suggesting a distant relationship between SeMNPV on the one hand and AcMNPV and OpMNPV on the other. Therefore, most of the conserved amino acids (Fig. 7.3-7.13) probably approach or represent the actual consensus of these proteins. For example, addition or omission of the CpGV sequence to the alignment with SeMNPV ORF xh136 (Fig. 7.6) hardly alters the consensus. This may indicate that granuloviruses are no more different from the NPVs than the NPVs are from each other. A similar conclusion can be made on the basis of the ODVP-6E gene homology (Theilmann *et al.*, 1996).

A few ORFs displayed a high degree of homology, most notably polyhedrin. SeMNPV v-ubi showed a high homology to baculoviral as well as to cellular ubiquitins. SeMNPV RR1 was highly homologous with eukaryotic RR1s (Chapter 6, van Strien *et al.*, 1997a), but not with OpMNPV RR1. Compared to the lower extent of homology among other baculovirus ORFs (Table 7.1 and Bjornson & Rohrmann, 1994), the high homology of the polyhedrin gene, apparent even at the nucleotide level, suggested that this gene may have been acquired only relatively recently by the baculoviruses.

From the new reconstructions of the baculovirus phylogeny based on either the polyhedrin amino acid or nucleotide sequences, it appeared that polyhedrin is not suited for this purpose. The deduced phylogeny was variant with the species used, not well supported by the bootstrap analyses (Fig. 7.15a-d, 7.16a-d) and often in conflict with baculovirus relationships determined by other criteria. The high degree of homology of the polyhedrin gene probably is the major cause of this. A high homology can interfere with the phylogenetic reconstruction in two ways: by reducing the number of informative characters and by increasing the chance to encounter a situation in which gene swapping has taken place. This is an unfortunate situation, because the polyhedrin gene is usually the first and only gene from a newly isolated baculovirus to be identified and sequenced, and to be used subsequently in phylogenetic reconstructions.

Other genes and proteins, preferably large and with non-structural functions (for instance IE1, helicase, etc.), may provide more consistent and reliable phylogenetic reconstructions than the polyhedrin gene does. Only with this type of information questions concerning the taxonomic position of the SNPVs, GVs, or single viruses like SpliMNPV and LdMNPV can be answered with greater confidence.

Genetic organization as phylogenetic marker

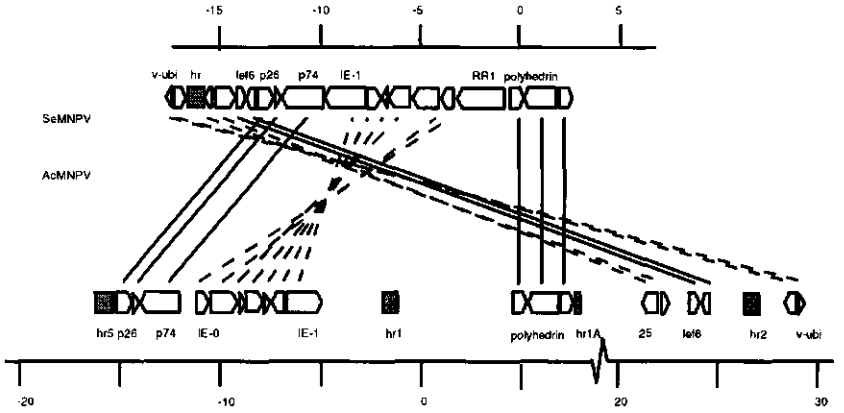
The most striking feature of the 20 kb segment of SeMNPV is not the homology of the ORFs to those of other baculoviruses or the presence of an *hr*, but the extensive difference in gene order, location and orientation between SeMNPV on the one hand and AcMNPV and OpMNPV on the other (Fig. 7.17a-c). Remarkable absentees from the SeMNPV 20 kb region are *ie2* (or *ien*, Carson *et al.*, 1991, Theilmann & Stewart, 1992a) and *ie3* (or *pe38*, Krappa & Knebel-Mörsdorf, 1991, Theilmann & Stewart, 1992b) which are located in between the *p10* and polyhedrin gene in both AcMNPV and OpMNPV. As immediate early genes they are involved

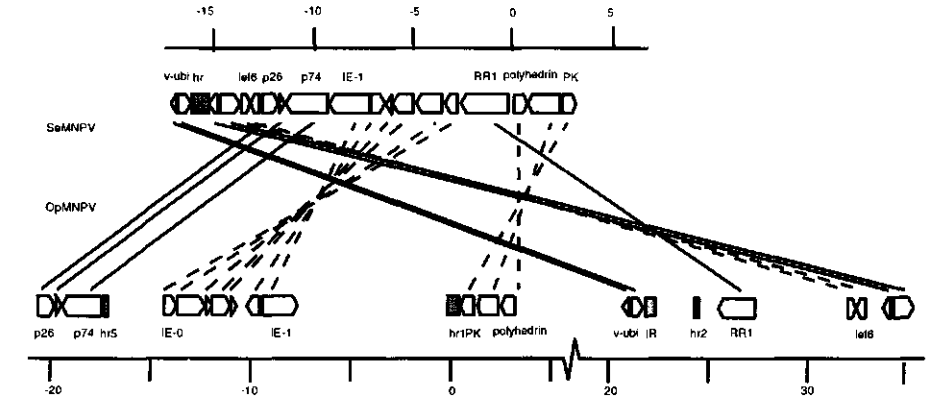
in transactivation, DNA replication and late gene expression (Theilmann & Stewart, 1992a, Wu et al., 1993, Yoo & Guarino, 1994, Passarelli & Miller, 1993, Todd et al., 1995, Lu & Miller, 1995, Kool et al., 1994, Ahrens & Rohrmann, 1995). It remains to be elucidated whether or not these genes are located somewhere else on the genome of SeMNPV. The only immediate early gene of SeMNPV that has been identified to date is the *ie1* gene. The ORF homologous to the AcMNPV and OpMNPV *ie0* left exon is present in SeMNPV, but the consensus splice donor sequence is absent. Therefore it seems that the clustering of the immediate early genes in between the very late major genes for p10 and polyhedrin, as observed in AcMNPV and OpMNPV, is not a conserved feature of baculoviruses.

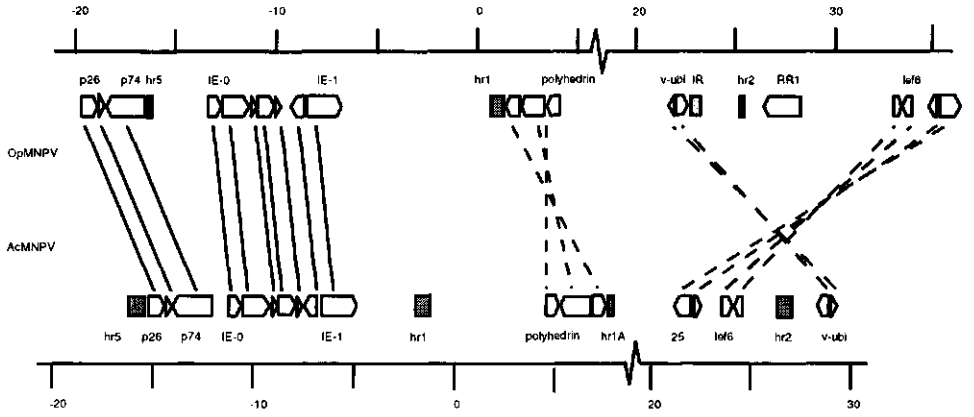
Comparing SeMNPV with AcMNPV (Fig. 7.17a) and OpMNPV (Fig. 7.17b) revealed that the largest conserved gene block encompassed SeMNPV ORFs xd92 until xd244. Genetic rearrangement occurred in the form of gene blocks as well as single genes, for instance in the case of ORF xh135 and ORF xh328. Rearrangements involved a large part of the genome. Genes found within the 20 kb segment of SeMNPV are dispersed over a 45 kb region in AcMNPV and even over 55 kb in OpMNPV (Fig 7.17a-c). The comparison of the complete genome sequences of AcMNPV and OpMNPV (Ayres *et al.*, 1994, Ahrens *et al.*, 1997) revealed that differences in genetic organization between these two are caused by gene capture and a few rearrangements in gene order. The rearrangements had occured in blocks or as single genes . However, the extent of the rearrangments between AcMNPV and OpMNPV is much smaller than that observed in comparisons with SeMNPV. A similar lack of conserved gene order was apparent in LdMNPV, in the region encoding the polyhedral envelope protein (Bjornson & Rohrmann, 1994). These data suggest that extensive genetic rearrangements are a common feature in baculoviruses. However, the gene order in the region around the histon-like basic protein *p6.9* gene was conserved between *Cryptophlebia leucotreta* GV and AcMNPV (Jehle & Backhaus, 1994). Preliminary sequence data indicated that a similar conservation could also be identified for about nine genes in a large contig encompassing the SeMNPV helicase gene (J.G.M. Heldens, *pers.comm.*).

The situation in the baculoviruses is to a certain extent comparable to that in other large eukaryotic DNA viruses, the poxviruses and herpesviruses. After the recent completion of the sequencing of the complete genome of the poxvirus Molluscum Contagiosum virus (MCV), comparison with the Vaccinia virus (VV) genome revealed that the genes which are conserved between the two viruses cluster in the central part of the genome, maintaining the same gene order (Senkevich *et al.*, 1996). Also in herpesviruses unique genes are found principally at the genome termini. Rearrangements occur frequently at the termini, where they involve rearrangement of a single or few genes. However, genetic rearrangements, predominantly in the form of large gene blocks, do occur throughout the herpesvirus genome (Gompels *et al.*,

Figure 7.17. Schematic overview of the genetic organization of the 20 kb DNA segment of SeMNPV and comparison with AcMNPV (Ayres *et al.*, 1994) and OpMNPV (Ahrens *et al.*, 1997). Units in kb. Continuous lines indicate homologous ORFs with similar genetic orientation, dashed lines indicate homologous ORFs with reversed orientation. Shading is used to indicated hr's and other repeated sequence elements.
a). Comparison between SeMNPV and AcMNPV. b). Comparison between SeMNPV and OpMNPV.
c). Comparison between AcMNPV and OpMNPV. Only the ORFs identified in the 20 kb SeMNPV fragment are indicated.


-15
-10
-5
0
5
v-ubi hr
lef6 p26
p74
IE-1
RR1 polyhedrin
SeMNPV
AcMNPV
hr5 p26
p74
IE-0
IE-1
hr1
polyhedrin hr1A
25
lef6
hr2
v-ubi
-20
-10
0
20
30


-15
-10
-5
0
5
v-ubi hr
lef6 p26
p74
IE-1
RR1 polyhedrin PK
SeMNPV
OpMNPV
p26
p74 hr5
IE-0
IE-1
hr1PK polyhedrin
v-ubi IR
hr2
RR1
lef6
-20
-10
0
20
30


-20
-10
0
20
30
p26
p74 hr5
IE-0
IE-1
hr1
polyhedrin
v-ubi IR
hr2
RR1
lef6
OpMNPV
AcMNPV
hr5
p26
p74
IE-0
IE-1
hr1
polyhedrin hr1A
25
lef6
hr2
v-ubi
-20
-10
0
20
30

1995, Davison & McGeoch, 1995). The mechanisms of herpesvirus divergence (nucleotide substitution, gene duplication, gene capture and rearrangement) have been reviewed by Davison & McGeoch (1995). Similar mechanisms seem to be responsible for baculovirus divergence. When the divergence among herpesviruses and poxviruses is compared with that among baculoviruses, it seems that in the two former virus groups extensive nucleotide substitution and gene capture have been relatively more important than in baculoviruses. Unfortunately, the low level of conservation of herpesvirus genes prohibited the estimation of the full magnitude of genetic rearrangement between the mammalian herpesviruses and Channel Catfish virus (CCV) (Davison, 1992).

In effect, the observed genetic rearrangement in some baculovirus genome regions is comparable to or greater than the situation in the herpesvirus termini. Additional sequence information is required to determine if a conserved central region, as is the case of poxviruses, is a distinctive feature of the baculoviruses. It remains to be elucidated if and how the baculovirus gene rearrangements are related to their mode of replication, for which the involvement of multiple origins, on dipersed locations on the genome, have been suggested. Furthermore, (the mechanism of) baculovirus genetic rearrangement might have unforseen implications for the assessment of the biosafety of recombinant baculoviruses with insecticidal genes.

The genetic rearrangements in the baculovirus genomes may provide new and supplementary phylogenetic information. Gene order has been used to determine herpesvirus phylogeny (Hannenhalli *et al.*, 1995). Boore *et al.* (1995) used the boundaries between the 37 mitochondrial genes to deduce a parsimonous arthropod phylogeny. The utilization of these kind of molecular data to determine phylogenies surpasses shortcomings in the alignment quality, compositional or substitutional bias and artificial associations of rapid evolving species. However, the method heavily relies on the ability to recognize homologous genes throughout several (virus) species. As seen in the comparison of CCV with other herpesviruses, recognition can constitute a real problem. Furthermore, gene transfer between species and gene duplication are possible interfering factors. Genes with different origins (for instance baculovirus *rr1* genes) must be excluded from the analysis. Nevertheless, when the number of contig or complete genome sequences from different baculoviruses will increase, baculovirus gene order may be useful as an independent phylogenetic character.

ACKNOWLEDGEMENTS

We thank R. Broer for the assistance in sequencing and G.F. Rohrmann for the access to the complete OpMNPV genome sequence prior to publication. Special thanks to R.W. Goldbach for reviewing earlier versions of this manuscript.

CHAPTER 8

SUMMARY AND CONCLUDING REMARKS

Baculoviruses are attractive, biological alternatives to chemical control agents of insect pests. These viruses are natural agents influencing the size of insect populations. They are often species-specific and some of them are highly efficacious, though their speed of action cannot meet that of chemical insecticides. In addition, chemical insecticides have a much wider range of target insects. With the advent of genetic engineering the host range of baculoviruses may be altered and their virulence possibly further improved. The principles underlying baculovirus virulence and specificity are not known. Therefore, the genetic organization and expression strategies of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV) have been investigated and described in this thesis. This virus is specific to the beet army worm *Spodoptera exigua* and highly virulent. The genetic information on the baculovirus type species *Autographa californica* (Ac) MNPV, which has a broad range of hosts and is less virulent, served as a reference point.

As a start the SeMNPV gene for the occlusion body protein, polyhedrin, was characterized. This protein is highly conserved among baculoviruses and comprises a major constituent of polyhedra. In contrast to other baculoviruses, the SeMNPV polyhedrin mRNAs were found to be non-polyadenylated (Chapter 2). The polyhedrin gene serves to set the origin of the physical map of the circular viral genome. By convention, the orientation of the map is set by the position of the other very late baculovirus gene, designated *p10*, relative to the polyhedrin gene. The presence of abundant amounts of mRNA allowed the identification of the SeMNPV *p10* gene via a cDNA probe. Both the SeMNPV polyhedrin and *p10* genes are transcribed late in infection and start in the canonical baculovirus late transcription initiation sequence, TAAG. In contrast to the polyhedrin mRNA the *p10* mRNA was polyadenylated (Chapters 2 and 3, respectively). The predicted amino acid sequence of SeMNPV p10 showed only low sequence homology with previously characterized baculoviral p10 proteins, but its size and domain structure were very similar (Chapter 3).

An equally small, but in contrast to *p10*, highly conserved gene encountered in the SeMNPV genome was the viral ubiquitin (*v-ubi*) gene. Ubiquitin is involved in protein degradation and is found associated with the non-occluded virions. The SeMNPV *v-ubi* gene was expressed late in infection and two baculovirus consensus late start sequences, TAAG, were used as transcriptional initiation sites. One was located only five nucleotides upstream of the translational start site, thereby providing the shortest untranslated leader reported to date for any baculovirus mRNA. The *v-ubi* gene had an unexpected genomic position in SeMNPV as compared to AcMNPV, suggesting differences in genome organization between baculoviruses (Chapter 4).

In constrast to *v-ubi*, the location of the SeMNPV immediate early gene *ie1* was more conserved as compared to AcMNPV. The *ie1* gene encodes a multifunctional protein in the baculovirus life cycle and is involved in transcriptional transactivation and in DNA replication. The SeMNPV *ie1* gene is larger than previously characterized baculovirus *ie1* genes and displayed an unusual transcription pattern and may be lacking splicing (Chapter 5).

Since the region between polyhedrin and p10 and encompassing *ie1* was much shorter

in SeMNPV (11 kb) than in AcMNPV (19 kb), it was of interest to characterize the entire region and compare its organization with that of AcMNPV. Upstream of the SeMNPV polyhedrin gene an open reading frame (ORF) was identified, encoding the large subunit of ribonucleotide reductase (RR1). This gene was not identified before in baculoviruses. Its product may be involved in nucleic acid synthesis. As expected from its putative function to generate DNA precursors, this gene was expressed as an early gene, prior to DNA replication. *S. littoralis* MNPV, another baculovirus infecting a *Spodoptera* species, also appeared to contain this gene. Phylogenetic analysis suggested that both viruses had acquired the *rr1* gene independently (Chapter 6). Ribonucleotide reductase is known to function as a virulence factor in other large DNA viruses. However, expression of the SeMNPV *rr1* gene in AcMNPV recombinants, using different promoter constructs, gave no indication of altered virulence (data not shown).

Finally, the genetic organization of a 20 kb segment of SeMNPV DNA representing 15% of the viral genome, encompassing the polyhedrin and *p10* gene, and potentially encoding 20 ORFs, was closely examined (Chapter 7). A non-coding sequence of 900 nucleotides in length, consisting of four imperfect palindromes centered around a *Bgl*II site and three direct repeats, was identified in this 20 kb segment. This sequence showed structural homology to *hr* regions found in other baculoviruses and which are well-established transcriptional enhancers and putative origins of DNA replication. All ORFs except for RR1 had a counterpart in AcMNPV, the baculovirus type species. The homology of SeMNPV with AcMNPV and OpMNPV genes was usually low and, except for polyhedrin and ie1, did not span promoter regions. A high level of homology was observed only with v-ubi and polyhedrin. A most remarkable aspect of the genetic organization of this genome segment of SeMNPV is the extensive rearrangement in gene order that must have taken place during baculovirus history.

The extremely high degree of conservation of the polyhedrin gene among baculoviruses questions its presumed value as a phylogenetic marker. Parsimony analysis of the large number of occlusion body protein sequences currently available, confirmed the notion that the polyhedrin gene is not suitable for this purpose (Chapter 7). The observation that the nucleotide sequence divergence of this gene had not reached the level of complete redundancy, suggested that it has been acquired by the baculoviruses rather recently, or alternatively, is frequently exchanged between different viruses. Several unresolved aspects of the baculovirus phylogeny, for instance the observed monophyletic grouping of the SNPVs, may better be investigated with other genes, e.g. *ie1* (Chapter 5). Finally, the baculovirus gene order itself might be used as an independent phylogenetic marker.

The initial characterization of SeMNPV DNA provides the basis for future genetic engineering of this highly specific baculovirus, in order to increase its efficacy as a bio-insecticide. The present data on the genetic structure and organization of the SeMNPV genome showed similarities to the better characterized AcMNPV. At the same time, major differences in gene order, coding capacity, gene sequence and expression were observed. The high divergence between SeMNPV and AcMNPV requires the complete sequencing of the SeMNPV genome to obtain a full view on the structural and genetic relatedness of baculoviruses. Ultimately, these analyses will be valuable in unravelling the issues related to the specificity and virulence in baculoviruses, two parameters of prime importance for their application as bio-control agents.

SAMENVATTING

SeMNPV, het kernpolyedervirus van de floridamot *Spodoptera exigua* (Lepidoptera, Noctuidae) en het onderwerp van dit proefschrift, behoort tot de familie van de baculovirussen. Dit zijn dubbelstrengs DNA virussen die alleen geleedpotigen (Arthropoda) infecteren. Het standaardvirus van deze familie is het kernpolyedervirus AcMNPV, *Autographa californica* multicapsid nucleopolyhedrovirus. De totale nucleotidenvolgorde van het circulaire, dubbelstrengs DNA genoom van AcMNPV is bepaald en van vele genen is de functie opgehelderd (Ayres *et al.*, 1994). AcMNPV heeft een uitgebreide gastheerreeks die rupsen van tientallen soorten (nacht)vlinders omvat, waaronder de larven van de floridamot. SeMNPV vertoont opvallende biologische kenmerken zoals specificiteit voor slechts één gastheer, de polyfage rupsen van de floridamot, gepaard aan een hoge virulentie voor dit plaaginsect. Deze eigenschappen, die SeMNPV zeer geschikt maken als middel om de floridamot biologisch te bestrijden, contrasteren met die van het minder virulente AcMNPV, het best onderzochte baculovirus.

Op basis van AcMNPV zijn genetisch gemodificeerde virussen ontworpen die, afhankelijk van de doelstelling, kunnen worden onderverdeeld in twee groepen. Er zijn recombinante virussen geconstrueerd die een heteroloog eiwit tot expressie brengen ten behoeve van farmaceutische of research doeleinden. De tweede groep recombinante virussen zijn ontworpen om het betreffende virus effectiever te maken als biologisch bestrijdingsmiddel. Beide doeleinden vereisen grondige moleculair-biologische kennis van de genetische organisatie en expressiestrategieën van het virale genoom, de eerste voorwaarde voor een rationeel ontwerp. Bij de start van het onderzoek was de moleculair-biologische kennis van het genoom van SeMNPV slechts zeer beperkt. Het hier beschreven onderzoek van het baculovirus SeMNPV dient als aanzet tot het verkrijgen van inzicht in de mechanismen die verantwoordelijk zijn voor gastheerspecificiteit en virulentie en tot het verbeteren van SeMNPV via genetische modificatie.

Als eerste werd het gen voor het "occlusion body"-eiwit, polyhedrine, gelokaliseerd en gekarakteriseerd. Dit eiwit is zeer geconserveerd bij baculovirussen en vormt het hoofdbestandeel van de polyeders, grote structuren die de virusdeeltjes omsluiten. In tegenstelling tot de polyhedrine transcripten van andere baculovirussen bleken de polyhedrine boodschapper RNA's van SeMNPV niet voorzien te zijn van een polyA-staart (Hoofdstuk 2). Het polyhedrine gen dient om het nulpunt van de fysische kaart van het circulaire virale genoom vast te leggen. Volgens afspraak wordt de oriëntatie van de fysische kaart bepaald door de positie van het andere zeer late baculovirus gen, *p10*, ten opzichte van het polyhedrine gen. Het *p10* gen van SeMNPV werd geïdentificeerd na de isolatie van complementair DNA, via grote hoeveelheden *p10* boodschapper RNA die laat in infectie nog werden aangetroffen in de cel. De "late" transcripten van zowel het polyhedrine als het *p10* gen bleken te starten in het transcriptie-initiatiemotief TAAG (Hoofdstuk 2 en 3). Deze sequentie is geconserveerd in (zeer) "late" baculovirus genen. De aminozuurvolgorde van het p10 eiwit van SeMNPV vertoonde slechts een geringe mate van overeenkomst met die van twee eerder gekarakteriseerde p10 eiwitten van AcMNPV en OpMNPV, het *Orgyia pseudotsugata* multicapsid nucleopolyhedrovirus. Grootte en domeinstructuur van de drie p10 eiwitten kwamen echter overeen (Hoofdstuk 3).

Een bijna even klein, maar in tegenstelling tot *p10* zeer geconserveerd gen, was het

virale ubiquitine (*v-ubi*) gen. Ubiquitine reguleert de eiwitafbraak en is een structureel eiwit van de "budded" virions. Het *v-ubi* gen van SeMNPV bleek laat in de infectie tot expressie te komen. Twee opeenvolgende consensus late baculovirus promoter-sequenties, TAAG, werden beiden gebruikt als transcriptiestartplaats. Eén van deze lag slechts 5 nucleotiden vóór de translatie-startplaats. Door deze beperkte afstand tussen transcriptie- en translatiestart is het de kortst bekende, niet-vertaalde "leader" van een baculovirus gen. Het *v-ubi* gen van SeMNPV lag op een onverwachte positie in vergelijking met de positie van het *v-ubi* gen in het genoom van AcMNPV. Deze observatie deed een verschil tussen de twee virussen in genomische structuur vermoeden (Hoofdstuk 4).

In tegenstelling tot de positie van het *v-ubi* gen is de positie van het "immediate early" gen *ie1* in het genoom van SeMNPV min of meer geconserveerd, vergeleken met de positie van *ie1* in het genoom van AcMNPV. Het *ie1* gen codeert voor een multifunctioneel eiwit in de levenscyclus van baculovirussen dat is betrokken bij de transactivatie van transcriptie en bij DNA replicatie. Het *ie1* gen van SeMNPV was groter dan eerder gekarakteriseerde baculovirus *ie1* genen en vertoonde een ongewoon transcriptiepatroon, waarbij "splicing" waarschijnlijk niet voorkwam (Hoofdstuk 5).

Aangezien het gebied tussen het polyhedrine en *p10* gen in het genoom van SeMNPV veel kleiner was dan in het genoom van AcMNPV, werd het van belang geacht om de gehele regio te karakteriseren en de organisatie met AcMNPV te vergelijken. Grenzend aan het polyhedrine gen van SeMNPV werd een open leesraam (open reading frame, ORF) voor de grote subunit van ribonucleotide reductase (RR1) geïdentificeerd. Dit gen was nog niet eerder aangetroffen in baculovirussen. Het gecodeerde eiwit is wellicht betrokken bij de synthese van nucleïnezuren. Zoals verwacht uit de vermoede functie, het produceren van DNA bouwstenen, begon de expressie van dit gen al in een vroeg stadium van de infectie, dus voorafgaand aan DNA replicatie (Hoofdstuk 6). In vele andere grote DNA virussen fungeert ribonucleotide reductase als virulentiefactor. Expressie van het *rr1* gen van SeMNPV in AcMNPV-recombinanten leverde echter geen aanwijzingen op dat dit ook bij baculovirussen het geval is.

Tenslotte werd de genetische organisatie van 15% van het genoom van SeMNPV, een DNA segment van 20 kbp dat zich uitstrekte tot voorbij de genen voor polyhedrine en *p10*, nauwgezet onderzocht (Hoofdstuk 7). Behalve twintig ORF's werd een niet-coderend segment van 900 nucleotiden gevonden. Dit segment was opgebouwd uit vier bijna-perfecte palindromen en drie andere herhaalde nucleotidensequenties. Het vertoonde structurele overeenkomst met baculovirus *hr* gebieden, die bekend staan als versterkers (enhancers) van transcriptie en als startpunt (origin) van DNA replicatie.

Alle geïdentificeerde ORF's, met uitzondering van RR1, hadden een tegenhanger in AcMNPV, het standaard baculovirus. De homologie van de genen van SeMNPV met de overeenkomstige genen van AcMNPV en OpMNPV was in het algemeen laag. De promotersequentie van de homologe virale genen vertoonde in het algemeen weinig overeenkomst, met uitzondering van de polyhedrine en *ie1* promotersequentie. Een hoge mate van overeenkomst in aminozuurvolgorde werd alleen waargenomen bij v-ubi en polyhedrine. Een zeer opmerkelijk aspect van de genetische organisatie van dit genoomsegment van SeMNPV is de uitgebreide herschikking in genvolgorde die moet hebben plaatsgevonden in de baculovirusgeschiedenis.

De uitzonderlijk hoge mate van overeenkomst in sequentie van het gen voor polyhedrine

tussen verscheidene baculovirussen riep vragen op over de waarde van dit gen als fylogenetisch kenmerk. De beschikbare sequenties van het gen voor het "occlusion body" eiwit zijn gebruikt in fylogenetische analyses, gebaseerd op parsimony. De evaluatie van deze analyses bevestigde de opvatting dat dit gen niet geschikt is voor fylogenetische reconstructies (Hoofdstuk 7). De waarneming dat de divergentie van de nucleotidensequentie van het gen voor polyhedrine het niveau van boventalligheid nog niet had bereikt, deed vermoeden dat baculovirussen dit gen recent hebben verworven of dat het regelmatig tussen verschillende virussen wordt uitgewisseld.

Verscheidene niet-opgeloste aspecten van de fylogenie van baculovirussen, zoals bijvoorbeeld de waargenomen monofyletische groepering van de SNPV's, kunnen wellicht beter worden onderzocht met behulp van andere genen en ORF's, bijvoorbeeld IE1 (Hoofdstuk 5). Tenslotte werd voor het eerst de mogelijkheid geopperd dat de genvolgorde in baculovirussen een bruikbaar fylogenetisch kenmerk kan zijn.

De hier beschreven moleculair-genetische karakterisering van SeMNPV legt de basis voor toekomstige genetische modificatie van dit zeer specifieke baculovirus. Op deze wijze kan de bruikbaarheid van SeMNPV als bio-insecticide wellicht nog worden vergroot. Uit de gepresenteerde gegevens met betrekking tot de genetische organisatie van het genoom van SeMNPV blijken overeenkomsten met het veel beter gekarakteriseerde genoom van AcMNPV, maar tegelijkertijd werden grote verschillen in de volgorde, samenstelling en expressie van genen gevonden. De waargenomen hoge mate van divergentie tussen het genoom van SeMNPV en het genoom van AcMNPV noodzaakt tot het bepalen van de nucleotidenvolgorde van het gehele genoom van SeMNPV, teneinde een volledig overzicht van de structurele en genetische verwantschap tussen baculovirussen te verkrijgen. Uiteindelijk zal dit soort analyses van nut blijken te zijn bij het ophelderen van vraagstukken betreffende de specificiteit en virulentie van baculovirussen, twee parameters die van groot belang zijn bij de toepassing van baculovirussen als biologisch bestrijdingsmiddel.

REFERENCES

Ahrens, C., & Rohrmann, G.F. (1995). Replication of *Orgyia pseudotsugata* baculovirus DNA: *ie-1* and *lef-2* are essential and *ie-2*, *p34* and *op-iap* are stimulatory genes. *Virology* 212, 650-662.

Ahrens, C.H., Russell, R.L.Q., Funk, C.J., Evans, J.T., Harwood, S.H., & Rohrmann, G.F. (1997). The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. *Virology* 229, 381-399.

Ahrens, C., Pearson, M.N., & Rohrmann, G.F. (1995). Identification and characterization of a second putative origin of DNA replication in a baculovirus of *Orgyia pseudotsugata*. *Virology* 207, 572-576.

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

Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.

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

Arnason, T.A., & Ellison, M.J. (1994). Stress resistance in Saccharomyces cerevisiae is strongly correlated with assemble of a novel type of multiubiquitin chain. *Molecular and Cellular Biology* 14, 7876-7883.

Arribas, C., Sampredo, J., & Izquierdo, M. (1986). The ubiquitin genes of *Drosophila melanogaster*: transcription and polymorphism. *Biochimica et Biophysica Acta* 868, 119-127.

Ayres, M.D., Howard, S.C., Kuzio, J., Ferber, M.L., & Possee, R.D. (1994). The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* 202, 586-605.

Baer, R.J., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G.F., Hudson, G.S., Satchwell, S.C., Seguin, C., Tuffnell, P.S., & Barrell, B.G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature* 310, 207-211.

Beek, C.P. van der. (1980). On the origin of the polyhedral protein of the nuclear polyhedrosis virus of *Autographa californica*. Mededelingen Landbouwhogeschool Wageningen 80-5.

Benz, G.A. (1986). Introduction: Historical perspectives. pp. 1-35 in: The biology of baculoviruses, volume 1: Biological properties and molecular biology. (eds. R.R. Granados & B.A. Federici). CRC-Press, Florida.

Beniya, J., Funk, C.J., Rohrmann, G.F., & Weaver, R.F. (1996). Purification of a virus-induced RNA polymerase form *Autographa californica* nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells that accurately initiates late and very late transcription in vitro. *Virology* 216, 12-19.

Bertin, J., Mendrysa, S.M., LaCount, D.J., Gaur, S., Krebs, J.F., Armstrong, R.C., Tomaselli, K.J., & Friesen, P.D. (1996). Apoptotic suppression by baculovirus p35 involves cleavage by and inhibition of a virus-induced DED-3/ICE-like protease. *Journal of Virology* 70, 6251-6259.

Bicknell, J.N., Leisy, D.J., Rohrmann, G.F., & Beaudreau, G.S. (1987). Comparison of the p26 gene region of two baculoviruses. *Virology* 161, 589-592.

Birnbaum, M.J., Clem, R.J., & Miller, L.K. (1994). An apoptosis-inhibiting gene from a nuclear polyhedrosis virus encoding a polypeptide with Cys/His motifs. *Journal of Virology* 68, 2521-2528.

Birnstiel, M.L., Busslinger, M., & Strub, K. (1985). Transcription termination and 3' processing: the end is in site! *Cell* 41, 349-359.

Bischoff, D.S., & Slavicek, J.M. (1994). Identification and characterization of a protein kinase gene in the *Lymantria dispar* multicapsid nuclear polyhedrosis virus. *Journal of Virology* 68, 1728-1736.

Bjornson, R.M., & Rohrmann, G.F. (1994). Nucleotide sequence of the polyhedron envelope protein gene region of the *Lymantria dispar* nuclear polyhedrosis virus. Corrigendum. *Journal of General Virology* 75, 3687.

Blissard, G.W. (1996). Baculovirus - insect cell interactions. pp. 73-93 in: Insect cell cultures. Fundamental and applied aspects. (eds. J.M. Vlak, C.D. de Gooijer, J. Tramper & H.G. Miltenburger) Kluwer, Dordrecht.

Blissard, G.W., & Rohrmann, G.F. (1990). Baculovirus diversity and molecular biology. *Annual Review of Entomology* 35, 127-155.

Blissard, G.W., & Rohrmann, G.F. (1989). Location, sequence, transcriptional mapping, and temporal expression of gp64 envelope glycoprotein gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 170, 537-555.

Bonning, B.D., Hoover, K., Booth, T.F, Duffey, S., & Hammock, B.D. (1995). Development of a recombinant baculovirus expressing a modified juvenile hormone esterase with potential for insect control. *Archives of Insect Biochemistry and Physiology* 30, 177-(194.

Boore, J.L., Collins, T.M., Stanton, D., Daehler, L.L., & Brown, W.M. (1995). Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. *Nature* 376, 163-165.

Boursnell, M., Shaw, K., Yáñez, R.J. Viñuela, E., & Dixon, L. (1991). The sequence of the ribonucleotide reductase genes from African swine fever virus show considerable homology with those of the Orthopox virus, vaccinia virus. *Virology* 184, 411-416.

Braunagel, S.C., He, H., Ramamurthy, P., & Summers, M.D. (1996). Transcription, translation, and cellular location of three *Autographa californica* nuclear polyhedrosis structural proteins: ODV-E18, ODV-E35, and ODV-EC27. *Virology* 222, 100-114.

Braunagel, S.C., & Summers, M.D. (1994). *Autographa californica* nuclear polyhedrosis virus, PDV, and ECV

viral envelopes and nucleocapsids: structural proteins, antigens, lipids and fatty acid profiles. *Virology* 202, 315-328.

Braunagel, S.C., Daniel, K.D., Reilly, L.M., Guarino, L.M., Hong, T., & Summers, M.D. (1992). Sequence, genomic organization of the *Eco*RI-A fragment of *Autographa californica* nuclear polyhedrosis virus, and identification of a viral-encoded protein resembling the outer capsid protein VP8 of Rotavirus. *Virology* 191, 1003-1008.

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

Bump, N.J., Hackett, M., Hugunin, M., Seshagiri, S., Brady, K., Chen, P., Ferenz, C., Franklin, S., Ghayur, T., Li, P., Licari, P., Mankovich, J., Shi, L., Greenberg, A.H., Miller, L.K., & Wong, W.W. (1995). Inhibition of ICE family proteases by baculovirus antiapoptosic protein p35. *Science* 269, 1885-1888.

Caballero, P., Zuidema, D., Santiago-Alvarez, C., & Vlak, J.M. (1992). Biochemical and biological characterization of four isolates of *Spodoptera exigua* nuclear polhedrosis virus. *Biocontrol Science and Technology* 2, 145-157.

Cameron, I.R., & Possee, R.D. (1989). Conservation of polyhedrin promoter function between *Autographa californica* and *Mamestra brassicae* nuclear polyhedrosis viruses. *Virus Research* 12, 183-200.

Caras, I.W., Levinson, B.B., Fabry, M., Williams, S.R., & Martin, D.W. (1985). Cloned mouse ribonucleotide reductase subunit M1 cDNA reveals amino acid sequence homology with *Escherichia coli* and herpesvirus ribonucleotide reductases. *Journal of Biological Chemistry* 260, 7015-7022.

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

Carson, D.D., Guarino, L.A., & Summers, M.D. (1988). Functional mapping of an AcMNPV immediately early gene which augments expression of the IE-1 trans-activated 39K gene. *Virology* 162, 444-451.

Carson, R.L. (1962). Silent spring. Penguin Press, Harmondsworth.

Chaeychomsri, S., Ikeda, M., & Kobayashi, M. (1995). Nucleotide sequence and transcriptional analysis of the DNA polymerase gene of *Bombyx mori* nuclear polyhedrosis virus. *Virology* 206, 435-447.

Chakerian, R., Rohrmann, G.F., Nesson, M.H., Leisy, D.J., & Beaudreau, G.S. (1985). The nucleotide sequence of the *Pieris brassicae* granulosis virus granulin gene. *Journal of General Virology* 66, 1263-1269.

Chang, M.T., Lanner-Herrera, C., & Fikes, M. (1989). Nucleotide sequence of *Lymantria dispar* nuclear polyhedrosis virus polyhedrin gene. *Journal of Invertebrate Pathology* 53, 241-246.

Chau, V., Tobias, J.W., Bachmair, A., Marriott, D., Ecker, D.J., Gonda, D.K., & Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in targeted short-lived protein. *Science* 243, 1576-1583.

Chee, M.S., Bankier, A.T., Beck, S., Bohni, R., Brown, C., Cerny, R., Horsnell, C.A., Hutchinson III, C.A., Kouzatides, T., Martignetti, J.A., Preddie, E., Satchwell, S.C., Tomlinson, P., Weston, K..M., & Barrell B.G. (1990). Analysis of the protein coding content of the sequence of human cytomegalovirus strain AD169. *Current Topics in Microbiology and Immunology* 154, 125-169.

Chen, C.-J.,& Thiem, S.M. (1997). Differential infectivity of two *Autographa californica* nucleopolyhedrovirus mutants on three permissive cell lines is the result of lef-7 deletion. *Virology* 227 88-95.

Child, S.J., Palumbo, G.J., Buller, R.M.L., & Hruby, D.E. (1990). Insertional inactivation of the large subunit of ribonucleotide reductase encoded by vaccinia virus is associated with reduced virulence *in vivo*. *Virology* 174, 625-629.

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

Choi, J., & Guarino, L.A. (1995a). The baculovirus transactivator IE1 binds to viral enhancer elements in the absence of insect cell factors. *Journal of Virology* 69, 4548-4551.

Choi, J., & Guarino, L.A. (1995b). Expression of the IE1 transactivator of *Autographa californica* nuclear polyhedrosis virus during viral infection. *Virology* 209, 99-107.

Chou, C.M., Huang, C.J., Lo, C.-F., Kou, G.H., & Wang, C.H. (1996). Characterization of *Perina nuda* nucleopolyhedrovirus (PenuNPV) polyhedrin gene. *Journal of Invertebrate Pathology* 67, 259-266.

Chou, J.M., Lo, C.F., Huang, C.J., & Wang, C.H. (1992). Isolation and nucleotide sequence of *Perina nuda* multicapsid nucleopolyhedrovirus (PenuMNPV) two late genes, polyhedrin and p10 gene. *Proc. XIX International Congress Entomology*, Bejing, China (Abstr. VIII-P #21, p280).

Ciechanover, A., & Schwartz, A.L. (1994). The ubiquitin-mediated proteolytic pathway: mechanisms of recognition of the proteolytic substrate and involvement in degradation of native cellular proteins. *The FASEB Journal* 8, 182-191.

Clarke, E.E., Tristem, M., Cory, J.S., & O'Reilly, D.R. (1996). Characterization of the ecdysteroid UDP-glucosyltransferase gene from *Mamestra brassicae* nucleopolyhedrovirus. *Journal of General Virology* 77, 2865-2871.

Clem, R.J., Robson, M., & Miller, L.K. (1994). Influence of infection route on the infectivity of baculovirus mutants lacking the apoptosis-inhibiting gene p35 and the adjacent gene p94. *Journal of Virology* 68, 6759-6762.

Clem, R.J., & Miller, L.K. (1994). Control of programmed cell death by the baculovirus genes p35 and *iap*. *Molecular and Cellular Biology* 14, 5212-5222.

Clem, R.J., & Miller, L.K. (1993). Apoptosis reduces both the in vitro replication and in vivo infectivity of a baculovirus. *Journal of Virology* 67, 3730-3738.
Clem, R.J., Fechheimer, M., & Miller, L.K. (1991). Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* 254, 1388-1390.
Cochran, M.A., & Faulkner, P. (1983). Location of homologous DNA sequences interspersed at five regions in the baculovirus AcMNPV genome. *Journal of Virology* 45, 961-970.
Conner, J., Marsden, H., & Clements, J.B. (1994). Ribonucleotide reductase of herpesviruses. *Reviews in Medical Virology* 4, 25-34.
Cowan, P., Bulach, D., Goodge, K., Robertson, A., & D.E. Tribe. (1994). Nucleotide sequence of the polyhedrin gene region of *Helicoverpa zea* single nucleocapsid nuclear polyhedrosis virus: placement of the virus in lepidopteran nuclear polyhedrosis group II. *Journal of General Virology* 75, 3211-3218.
Crook, N.E., Clem, R.J., & Miller, L.K. (1993). An apoptosis-inhibiting baculovirus gene with a zinc-finger like motif. *Journal of Virology* 67, 2168-2174.
Croizier, L., & Croizier, G. (1994). Nucleotide sequence of the polyhedrin gene of *Spodoptera littoralis* multiple nucleocapsid nuclear polyhedrosis virus. *Biochimica et Biophysica Acta* 1218, 457-459.
Croizier, G., Croizier, L., Argaud, O., & Poudevigne, D. (1994). Extension of *Autographa californica* nuclear polyhedrosis virus by interspecific replacement of a short DNA sequence in the p143 helicase gene. *Proceedings of the National Academy of Sciences* USA 91, 48-52.
Croizier, G., Boukhoudmi-Amiri, K., & Croizier, L. (1989). A physical map of *Spodoptera littoralis* B-type nuclear polyhedrosis virus genome. *Archives of Virology* 104, 145-151.
Cunha, C.V., & Costa, J.V. (1992). Induction of ribonucleotide reductase activity in cells infected with African swine fever virus. *Virology* 187, 73-83.
Cunningham, J.C. (1995). Baculoviruses as microbial insecticides. pp. 261-292 in: Novel approaches to integrated pest management (ed. R. Reuven), CRC Press, Boca Raton.
Darteil, R., Bublot, M., Laplace, E., Bouquet, J.-F., Audonnet, J.-C., & Riviere, M. (1995). Herpesvirus of turkey recombinant viruses expressing infectious bursal disease virus (IBDV) VP2 immunogen induce protection against an IBDV virulent challenge in chickens. *Virology* 211, 481-490.
Davis, R., Graslund, A., & Thelander, L. (1994). Purification, characterization and localization of subunit interaction of recombinant mouse ribonucleotide reductase R1 subunit. *Journal of Biological Chemistry* 269, 23171-23176.
Davison, A.J., & McGeoch, D.J. (1995). Herpesviridae. Chapter 20 in: Molecular basis of virus evolution (eds. A.J. Gibbs, C.H. Chalisher & F. Garcia-Arenal), Cambridge UP, Cambridge.
Davison, A.J. (1992). Channel catfish virus: a new type of herpesvirus. *Virology* 186, 9-14.
Davison, A.J., & Scott, J.E. (1986). The complete DNA sequence of varicella-zoster virus. *Journal of General Virology* 67, 1759-1816.
Devereux, J., Haeberli, P., & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* 12, 387-395.
Doherty, F.J., & Mayer, R.J. (1992). Intracellular proteins. IRL Press, London.
Driscoll, J., & Finley, D. (1992). A controlled breakdown: Antigen processing and the turnover of viral proteins. *Cell* 68, 823-825.
Dunigan, D.D., Dietzgen, R.G., Schoelz, J.E., & Zaitlin, M. (1988). Tobacco mosaic virus particles contain ubiquitinated coat protein subunits. *Virology* 165, 310-312.
Elledge, S.J., Zhou, Z., & Allen, J.B. (1992). Ribonucleotide reductase, regulation, regulation, regulation. *Trends in Biochemical Sciences* 17, 1(19-123.
Elledge, S.J., & Davis, R.W. (1990). Two genes differentially regulated in the cell cycle and by DNA damaging agents encode alternative regulatory subunits of ribonucleotide reductase. *Genes & Development* 4, 740-751.
Engelhard, E.K., & Volkman, L.E. (1995). Developmental resistance in fourth instar *Trichoplusia ni* orally inoculated with *Autographa californica* M nuclear polyhedrosis virus. *Virology* 209, 384-389.
Engelhard, E.K., Kam-Morgan, L.N.W., Washburn, J.O., & Volkman, L.E. (1994). The insect tracheal system: a conduit for the systemic spread of *Autographa californica* M nuclear polyhedrosis virus. *Proceedings of the National Academy of Sciences USA* 91, 3224-3227.
Faktor, O., Toister-Achituv, M., & Nachum, O. (1997a). Enhancer element, repetive sequences and gene organization in an 8-kbp region containing the polyhedrin gene of the *Spodoptera littoralis* nucleopolyhedrovirus. *Archives of Virology* 142, 1-15.
Faktor, O., Toister-Achituv, M., Nachum, O., & Kamensky, B. (1997b). The p10 gene of *Spodoptera littoralis* nucleopolyhedrovirus: nucleotide sequence, transcriptional analysis and unique organization in the p10 locus. *Journal of General Virology*, in press.
Fan, X., Thirunavukkarasu, K., & Weaver, R.F. (1996). Temperature-sensitive mutations in the protein kinase-1 (*pk-1*) gene of the *Autographa californica* nuclear polyhedrosis virus that block very late gene expression. *Virology* 224, 1-9.
Felsenstein, J. (1993). PHYLIP (Phylogeny Interference Package) Version 3.5. Distributed by the author. Department of Genetics, University of Washington, Seattle.
Fernandez Sarabia, M.J., McInery, C., Harris, P. , Gordon, C., & Fantes, P. (1993). The cell cycle genes cdc22+ and suc22+ of the fission yeast *Schizosaccharomyces pombe* encode the large and small

subunits of ribonucleotide reductase. *Molecular & General Genetics* 238, 241-251.
Finley, D., & Chau, V. (1991). Ubiquitination. *Annual Review of Cell Biology* 7, 25-69.
Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman,J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., & Venter, J.C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Science* 269, 496-512.
Flipsen, J.T.M., Martens, J.W.M., Oers, M.M. van, Vlak, J.M., & Lent, J.W.M. van. (1995a). Passage of *Autographa californica* nuclear polyhedrosis virus thrhough the midgut epithelium of *Spodoptera exigua* larvae. *Virology* 208, 328-335.
Flipsen, J.T.M., Mans, R.M.W., Kleefsman, A.W.F., Knebel-Mörsdorf, D., & Vlak, J.M. (1995b). Deletion of the baculovirus ecdysteroid UDP-glucosyltransferase gene induces early degeneration of Malpighian tubules in infected insects. *Journal of Virology* 69, 4549-4532.
Flipsen, J.T.M., Lent, J.W.M. van, Goldbach, R.W., & Vlak, J.M. (1993). Expression of polyhedrin and p10 in the midgut of AcMNPV infected *Spodoptera exigua* larvae: an immunoelectron microscopic investigation. *Journal of Invertebrate Pathology* 61, 17-23.
Francki, R.I.B., Fauquet, C.M., Knudson, D., & F. Brown (eds). (1991). Classification and nomenclature of viruses. *Fifth Report of the International Committee on Taxonomy of Viruses*, pp. 450.
Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton,G., Kelley, J.M., Fritchman, J.L., Weidman, J.F., Small, K.V., Sandusky,M., Fuhrmann, J.L., Nguyen, D.T., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.-F., Dougherty, B.A., Bott, K.F., Hu, P.-C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A.III & Venter, J.C. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science* 270, 397-403.
Friesen, P.D., & Miller, L.K. (1986). The regulation of baculovirus gene expression. *Current Topics in Microbiology & Immunology* 131, 31-49.
Fuchs, L.Y., Woods, M.S., & Weaver, R.F. (1983). Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *Journal of Virology* 43, 641-646.
Gallant, P., Shiio, Y., Cheng, P.F., Parkhurst, S.M., & Eisenman, R.N. (1996). Myc and Mac homologs in *Drosophila*. *Science* 274, 1523-1527.
Garcia-Canedo, A., Pavan, H.O., & Maruniak, J.E. (1996). A variable region of *Anticarsia gemmatalis* nuclear polyhedrosis virus contains tandemly repeated DNA sequences. *Virus Research* 41, 123-132.
Gelernter, W.D., & Federici, B.A. (1986a). Isolation, identification, and determination of virulence of a nuclear polyhedrosis virus from the beet armyworm, *Spodoptera exigua* (Lepidoptera: Noctuidae). *Environmental Entomology* 15, 240-245.
Gelernter, W.D., & Federici, B.A. (1986b). Continuous cell line from *Spodoptera exigua* (Lepidoptera: Noctuidae) that supports replication of nuclear polyhedrosis viruses from *Spodoptera exigua* and *Autographa californica*. *Journal of Invertebrate Pathology* 48, 199-207.
Glockner, B., Hoopes, R.R., Hodges, L., & Rohrmann, G.F. (1993). In vitro transcription from baculovirus late promoters: accurate mRNA initiation by nuclear extracts prepared from infected *Spodoptera frugiperda* cells. *Journal of Virology* 67, 3771-3776.
Glockner, B., Hoopes, R.R., & Rohrmann, G.F. (1992). *In vitro* transactivation of baculovirus early genes by nuclear extracts from *Autographa californica* nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells. *Journal of Virology* 66, 3476-3484.
Goldstein, D.J., & Weller, S.K. (1988). Herpes simplex virus type 1- induced ribonucleotide reductase activity is dispensable for virus growth and DNA synthesis, isolation and characterization of and ICP6 *lacZ* insertion mutant. *Journal of Virology* 62, 196-205.
Gombart, A.F., Pearson, M.N., Rohrmann, G.F., & Beaudreau, G.S. (1989a). A baculovirus PE-associated protein: genetic location, nucleotide sequence and immunocytochemical characterization. *Virology* 169, 182-(193.
Gombart, A.F., Blissard, G.W., & Rohrmann, G.F. (1989b). Characterization of the *Hin*dIII M region of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata* reveals major differences among baculoviruses. *Journal of General Virology* 80, 1815-1828.
Gompels, U.A., Nicholas, J., Lawrence, G., Jones, M., Thomson, B.J., Martin, M.E.D., Estathiou, S., Craxton, M., & Macaulay, H.A. (1995). The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. *Virology* 209, 29-51.
Gonzales, M.A., Smith, G.E., & Summers, M.D. (1989). Insertion of the SfMNPV polyhedrin gene into an AcMNPV polyhedrin deletion mutant during viral infection. *Virology* 170, 160-175.
Granados, R.R., & Lawler, K.A. (1981). *In vivo* pathway of *Autographa californica* baculovirus infection. *Virology* 108, 297-308.
Granados, R.R., & Federici, B.A. (1986). (Eds.) Biology of baculoviruses. Vol. I, CRC Press Inc., Boca Raton, Florida, USA.
Gross, C.H., Russell, R.L.Q., & Rohrmann, G.F. (1994). *Orgyia pseudotsugata* baculovirus p10 and polyhedron

envelope protein genes: analysis of their relative expression levels and polyhedron structure. *Journal of General Virology* 75, 1115-1123.
Grula, M.A., Buller, P.L., & Weaver, R.F. (1981). Alpha-amanitin resistant viral RNA synthesis in nuclei isolated form nuclear polyhedrosis virus-infected *Heliothis zea* larvae and *Spodoptera frugiperda* cells. *Journal of Virology* 38, 916-921.
Guarino, L.A., Smith, G., & Dong, W. (1995). Ubiquitin is attached to membranes of baculovirus particles by a novel type of phospholipid anchor. *Cell* 80, 301-309.
Guarino, L.A., & Dong, W. (1994). Functional dissection of the *Autographa californica* nuclear polyhedrosis virus enhancer element *hr5*. *Virology* 200, 328-335.
Guarino, L.A., & Dong, W. (1991). Expression of an enhancer-binding protein in insect cells transfected with the *Autographa californica* nuclear polyhedrosis virus IE1 gene. *Journal of Virology* 65, 3676-3680.
Guarino, L.A. (1990). Identification of a viral gene encoding a ubiquitin-like protein. *Proceedings of the National Academy of Sciences USA*. 87, 409-413.
Guarino, L.A., & Smith, M.W. (1990). Nucleotide sequence and characterization of the 39K gene region of *Autographa californica* nuclear polyhedrosis virus. *Virology* 179, 1-8.
Guarino, L.A., & Summers, M.D. (1987). Nucleotide sequence and temporal expression of a baculovirus regulatory gene. *Journal of Virology* 61, 2091-2099.
Guarino, L.A., Gonzales, M.A., & Summers, M.D. (1986). Complete sequence and enhancer function of the homologous DNA regions of *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 60, 224-229.
Guarino, L.A., & Summers, M.D. (1986a). Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene. *Journal of Virology* 60, 563-571.
Guarino, L.A., & Summers, M.D. (1986b). Interspersed homologous DNA of *Autographa californica* nuclear polyhedrosis virus enhances delayed-early gene expression. *Journal of Virology* 60, 215-223.
Haas, A.L., Katzung, D.J., Reback, P.M., & Guarino, L.A. (1996). Functional characterization of the ubiquitin variant encoded by the baculovirus *Autographa californica*. *Biochemistry* 35, 5385-5394.
Hannenhalli, S., Chappey, C., Koonin, E.V., & Pevzner, P.A. (1995). Genome sequence comparison and scenarios for gene rearrangements: a test case. *Genomics* 30, 299-311.
Harrap, K.A. (1970). Cell infection by a nuclear polyhedrosis virus. *Virology* 42, 311-318.
Hawtin, R.E., Arnold, K., Ayres, M.D., Zanotto, P.M. de A., Howard, S.C., Gooday, G.H., Chapell, L.H., Kitts, P.A., King, L.A., & Possee, R.D. (1995). Indentification and preliminary characterization of a chitinase gene in the *Autographa californica* nuclear polyhedrosis virus genome. *Virology* 212, 259-264.
Heineman, T.C., & Cohen, J.I. (1994). Deletion of the varicella-zoster virus large subunit of ribonucleotide reductase impairs growth of the virus *in vitro*. *Journal of Virology* 68, 3317-3323.
Heldens, J.G.M., Strien, E.A. van, Feldmann, A.M., Kulcsár, P., Munoz, D., Leisy, D.J., Zuidema, D., Golbach, R.W., & J.M. Vlak. (1996). *Spodoptera exigua* multicapsid nucleopolyhedrovirus deletion mutants generated in cell culture lack virulence *in vivo*. *Journal of General Virology* 77, 3127-3134.
Hershberger, P.A., Dickson, J.A., & Friesen, P.D. (1992). Site-specific mutagenesis of the 35-kilodalton protein gene encoded by *Autographa californica* nuclear polyhedrosis virus: cell-line specific effects on virus replication. *Journal of Virology* 66, 5525-5533.
Hill, J.E., Kuzio, J., Wilson, J.A., MacKinnon, E.A., & Faulkner, P. (1993). Nucleotide sequence of the p74 gene of a baculovirus pathogenic to the spruce budworm, *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus. *Biochimica et Biophysica Acta*, 1172, 187-189.
Hingamp, P.M., Leyland, M.L., Webb, J., Twigger, S., Mayer, R.J., & Dixon, L.K. (1995). Characterization of a ubiquitinated protein which is externally located in African swine fever virions. *Journal of Virology* 69, 1785-1793.
Hink, F. (1970) Established insect cell line from the cabbage looper, *Trichoplusia ni*. *Nature* 226, 466-467.
Hooft van Iddekinge, B.J.L., Smith, G.E., & Summers, M.D. (1983). Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus. *Virology* 131, 561-565.
Howard, S.C., Ayres, M.D., & Possee, R.D. (1986). Mapping of the 5' and 3' ends of *Autographa californica* nuclear polyhedrosis virus polyhedrin mRNA. *Virus Research* 5, 109-119.
Howell, M.L., Roseman, N.A., Slabaugh, M.B., & Mathews, C.K. (1993). Vaccinia virus ribonucleotide reductase, correlation between deoxyribonucleotide supply and demand. *Journal of Biological Chemistry* 268, 7155-7162.
Hu, J., Ding, H., & Wu, X. (1993). Cloning and sequencing of *Attacus ricini* nuclear polyhedrosis virus polyhedrin gene. *I Chuan Hsueh PAO* 20, 300-304.
Hu, N.-T., Lu, Y.-F., Hashimoto, Y., Maeda, S., & Hou, R.F. (1994). The p10 gene of natural isolates of *Bombyx mori* nuclear polyhedrosis virus encodes a truncated protein with an M_r of 7700. *Journal of General Virology* 75, 2085-2088.
Hu, Z.H., Broer, R., Westerlaken, J., Martens, J.W.M., Jin, F., Jehle, J.A., Wang, L.M., & Vlak, J.M. (1997). Characterization of the ecdysteroid UDP-glucosyltransferase gene of a single nucleocapsid nucleopolyhedrovirus of *Buzura suppressaria*. *Virus Research* 47, 91-97.
Hu, Z.-H., & Vlak, J.M. (1997). Engineering of biosafe baculoviruses with improved insecticidal properties: developments and prospects. *Virologica Sinica* 12, 14-25.
Hu, Z.H., Liu, M.F., Jin, F., Wang, Z.X., Liu, X.Y., Li, M.J., Liang, B.F., & Xie, T.E. (1993). Nucleotide sequence

of the *Buzura suppressaria* single nucleocapsid nuclear polyhedrosis virus polyhedrin gene. *Journal of General Virology* 74, 1617-1620.
Huh, N.E., & Weaver, R.F. (1990). Identifying the RNA polymerases that synthesize specific transcripts of *Autographa californica* nuclear polyhedrosis virus. *Journal of General Virology* 71, 195-202.
Huszar, D., & Bacchetti S. (1981). Partial purification and characterization of the ribonucleotide reductase induced by herpes simplex virus infection of mammalian cells. *Journal of Virology* 37, 580-588.
Huybrechts, R., Guarino, L.A., van Brussel, M., & Vulsteke, V. (1994). Nucleotide sequence of a transactivating *Bombyx mori* nuclear polyhedrosis virus immediate early gene. *Biochimica Biophysica Acta* 1129, 328-330.
Iatrou, K., Ito, K., & Witkiewicz, H. (1985). Polyhedrin gene of *Bombyx mori* nuclear polyhedrosis virus. *Journal of Virology* 54, 436-445.
Jain, A., & Hasnain, S.E. (1996). A 30 kDa host protein binds to two very late baculovirus promoters. *European Journal of Biochemistry* 239, 384-390.
Jarvis, D.L., Reilly, L.M., Hoover, K., Schultz, C., Hammock, B.D., & Guarino, L.A. (1996a). Construction and characterization of immediate early baculovirus pesticides. *Biological Control* 7, 228-235.
Jarvis, D.L., Weinkauf, C., & Guarino, L.A. (1996b). Immediate-early baculovirus vectors for foreign gene expression in transformed or infected insect cells. *Protein Expression and Purification* 8, 191-203.
Jehle, J.A., & Backhaus, H. (1994a). Genome organization of the DNA-binding protein gene region of *Cryptophlebia leucotreta* granulosis virus is closely related to that of nuclear polyhedrosis viruses. *Journal of General Virology* 75, 1815-1820.
Jehle, J.A., & Backhaus, H. (1994b). The granulin gene region of *Cryptophlebia leucotreta* granulosis virus: sequence analysis and phylogenetic considerations. *Journal of General Virology* 75, 3667-3671.
Johnson, E.S., Ma, P.C.M., Ota, I.M., & Varshavsky, A. (1995). A proteolytic pathway that recognizes ubiquitin as a degradation signal. *Journal of Biological Chemistry* 270, 17442-17456.
Jordan, A., Gibert, I., & Barbe, J. (1994). Cloning and sequencing of the genes of *Salmonella typhimurium* encoding a new bacterial ribonucleotide reductase. *Journal of Bacteriology* 176, 3420-3427.
Kaliman, A.V., Bolgdogkoi, Z., & Fodor, I. (1994). Large and small subunits of the Aujeszky's disease virus ribonucleotide reductase, nucleotide sequence and putative structure. *Biochimica et Biophysica Acta* 1219, 151-156.
Kemp, L.M., & Latchman, D.S. (1988). The herpes simplex virus type 1 immediate-early protein ICP4 specifically induces increased transcription of the human ubiquitin B gene without affecting the ubiquitin A and C genes. *Virology* 166, 258-261.
King, L.A., & Possee, R.D. (1990). A baculovirus expression vector derived from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus. *Journal of General Virology* 71, 971-976.
Kitts, P.A., & Possee, R.D. (1993). A method for producing recombinant baculovirus expression vectors at high frequency. *BioTechniques* 14, 810-817.
Kogan, P.H., Chen, X., & Blissard, G.W. (1995). Overlapping TATA-dependent and TATA-independent early promoter activities in the baculovirus gp64 envelope fusion protein gene. *Journal of Virology* 69, 1452-1461.
Kogan, P.H., & Blissard, G.W. (1994). A baculovirus *gp64* early promoter is activated by host transcription binding to CACGTG and GATA elements. *Journal of Virology* 68, 813-822.
Kool, M., Ahrens, C.H., Vlak, J.M., & Rohrmann, G.F. (1995). Replication of baculovirus DNA. *Journal of General Virology* 76, 2103-2118.
Kool, M., Ahrens, C.H., Goldbach, R.W., Rohrmann, G.F., & Vlak, J.M. (1994a). Identification of genes involved in DNA replication of the *Autographa californica* baculovirus. *Proceedings of the National Academy of Sciences USA* 91, 11212-11216.
Kool, M., Goldbach, R.W., & Vlak, J.M. (1994b). A putative non-hr origin of DNA replication in the *Hin*dIII-K fragment of *Autographa californica* nuclear polyhedrosis virus. *Journal of General Virology* 75, 3345-3352.
Kool, M., & Vlak, J.M. (1993). The structural and functional organization of the *Autographa californica* nuclear polyhedrosis virus genome: an overview. *Archives of Virology* 130, 1-16.
Kool, M., Voeten, J.T.M., Goldbach, R.W., Tramper, J., & Vlak, J.M. (1993). Identification of seven putative origins of *Autographa californica* multiple nuclear polyhedrosis virus DNA replication. *Journal of General Virology* 74, 2661-2668.
Kovacs, G.R., Choi, J., Guarino, L.A., & Summers, M.D. (1992). Functional dissection of the *Autographa californica* nuclear polyhedrosis virus immediate-early 1 transcriptional regulatory protein. *Journal of Virology* 66, 7429-7437.
Kovacs, G.R., Guarino, L.A., & Summers, M.D. (1991a). Novel regulatory properties of the IE1 and IE0 transactivators encoded by the baculovirus *Autographa californica* multicapsid nuclear polyhedrosis virus. *Journal of Virology* 65, 5281-5288.
Kovacs, G.R., Guarino, L.A., Graham, B.L., & Summers, M.D. (1991b). Identification of spliced baculovirus RNAs expressed late in infection. *Virology* 185, 633-643.
Kozak, M.J. (1983). Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNA. *Nucleic Acids Research* 12, 857-872.
Kozlov, E.A., Rodnin, N.V., Palchikivskaia, L.I., Levitina, T.L., Bobrovskaia, M.T., & Radomskii, N.F. (1994).

Primary structure of polyhedrin from the nuclear polyhedrosis virus of the silkworm *Malacosoma neustria*. *Bioorg. Khim.* 20, 543-545.
Krappa, R., Behn-Krappa, A., Jahnel, F., Doerfler, W., & Knebel-Mörsdorf, D. (1992). Differential factor binding at the promoter of early baculovirus gene PE38 during viral infection: GATA motif is recognized by an insect protein. *Journal of Virology* 66, 3494-3503.
Krappa, R., & Knebel-Mörsdorf, D. (1991). Identification of the very early transcribed baculovirus gene PE-38. *Journal of Virology* 65, 805-812.
Kuzio, J., Rohel, D.Z., Curry, C.J., Krebs, A., Carstens, E.B., & Faulkner, P. (1984). Nucleotide sequence of the p10 polypeptide gene of *Autographa californica* nuclear polyhedrosis virus. *Virology* 139, 414-418.
Kuzio, J., Jacques, R., & Faulkner, P. (1989). Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. *Virology* 173, 759-763.
Kyte, J., & Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157, 105-132.
Lee, S.-Y., Poloumienko, A., Belfry, S., Qu, X., Chen, W., MacAfee, N., Morin, B., Lucarotti, C., & Krause, M. (1996). A common pathway for p10 and calyx proteins in progressive stages of polyhedron envelope assembly in AcMNPV-infected *Spodoptera frugiperda* cells. *Archives of Virology* 141, 1247-1258.
Leisy, D.J., Rasmussen, C., Kim, H.T., & Rohrmann, G.F. (1995). The *Autographa californica* nuclear polyhedrosis virus homologous region 1A-identical sequences are essential for DNA replication activity and transcriptional enhancer function. *Virology* 208, 742-752.
Leisy, D.J.& Rohrmann, G.F. (1993). Characterization of the replication of plasmids containing *hr* sequences in baculoviruses-infected *Spodoptera frugiperda* cells. *Virology* 196, 722-730.
Leisy, D., Rohrmann, G., & Beaudreau, G. (1986a). The nucleotide sequence of the polyhedrin gene region from the multicapsid baculovirus of *Orgyia pseudostugata*. *Virology* 153, 280-288.
Leisy, D., Nesson, M., Pearson, M., Rohrmann, G., & Beaudreau, G. (1986b). Location and nucleotide sequence of the *Orgyia pseudotsugata* single nucleocapsid nuclear polyhedrosis virus polyhedrin gene. *Journal of General Virology* 67, 1073-1079.
Leisy, D.J., Rohrmann, G.F., Nesson, M., & Beaudreau, G.S. (1986c). 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., & Weaver, R.F. (1986). Nucleotide sequence of a portion of the *Autographa californica* nuclear polyhedrosis virus genome containing the *Eco*RI site-rich region (hr5) and an open reading frame just 5' of the p10 gene. *Journal of General Virology* 67, 2565-2570.
Lu, A., & Carstens, E.B. (1993). Immediate-early baculovirus genes trans-activate the p143 gene promoter of *Autographa californica* nuclear polyhedrosis virus. *Virology* 195, 710-718.
Lopez-Otin, C., Simon Mateo,c., Martinez, L., & Vinuela, E. (1989). Gly-Gly-X, a novel consensus sequence for the proteolytic processing of viral and cellular proteins. *Journal of Biological Chemistry* 262, 9107-9110.
Lori, F., Malykh, A., Cara, A., Sun, D., Weinstein, J.N., Lisziewicz, J., & Gallo, R.C. (1994). Hydroxyurea as an inhibitor of human immunodefiency virus-type 1 replication. *Science* 266, 801-805.
Lu, A., & Miller, L.K. (1995a). The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *Journal of Virology* 69, 975-982.
Lu, A., & Miller, L.K. (1995b).) Differential requirements for baculovirus late expression factor genes in two cell lines. *Journal of Virology* 69, 6265-6272.
Lu, A., & Carstens, E.B. (1992). Transcription analysis of the *Eco*RI D region of the baculovirus *Autographa californica* nuclear polyhedrosis virus identified an early 4-kilobase RNA encoding the essential p143 gene. *Journal of Virology* 195, 710-718.
Luckow, V.A. (1991). Cloning and expression of heterologous genes in insect cells with baculovirus vectors. pp. 97-152 in: *Recombinant DNA Technology and Applications* (eds. A. Prokop, R.K. Bjapai & C.S. Ho). McGraw-Hill, New York.
Luckow, V.L., & Summers, M.D. (1988). Trends in the development of baculovirus expression vectors. *Bio/Technology* 6, 47-55.
Maeda, S., Kamita, S.G., & Kondo, A. (1993). Host range expansion of *Autographa californica* nuclear polyhedrosis virus (NPV) following recombination of a 0.6-kilobase-pair DNA fragment originating from *Bombyx mori* NPV. *Journal of Virology* 67, 6234-6238.
Maeda, S., Volrath, S.L., Hanzlik, T.N., Harper, S.A., Majima, K., Maddox, D.W., Hammock, B.D., & Fowler, E. (1991). Insecticidal effects of an insect-specific neurotoxin expressed by a recombinant baculovirus. *Virology* 184, 777-780.
Maeda, S., & Majima, K. (1990). Molecular cloning and physical mapping of the genome of *Bombyx mori* nuclear polyhedrosis virus. *Journal of General Virology* 71, 1851-1855.
Maeda, S. (1989). Expression of foreign genes in insects using baculovirus vectors. *Annual Review of Entomology* 34, 351-372.
Majima, K., Kobara, R., & Maeda, S. (1993). Divergence and evolution of homologous regions of *Bombyx mori* nuclear polyhedrosis virus. *Journal of Virology* 67, 7513-7521.
Mao, S.S., Holler, T.P., Yu, G.X., Bollinger, J.M.Jr., Booker, S., Johnston, M.I., & Stubbe, J. (1992). A model for the role of multiple cysteine residues involved in ribonucleotide reduction, amazing and still confusing.

Biochemistry 31, 9733-9743.
McClintock, J.M., & Dougherty, E.M. (1988). Restriction mapping of *Lymantria dispar* nuclear polyhedrosis virus DNA: localization of the polyhedrin gene and identification of four homologous regions. *Journal of General Virology* 69, 2303-2312.
McCutchen, B.F., Choudary, P.V., Crenshaw, R., Maddox, D., Kamita, S.G., Volrath, S., Fowler, E., Hammock., B.D., & Maeda, S. (1991). Development of a recombinant baculovirus expressing an insect-selective neurotoxin - Potential for pest control. *Bio/Technology* 9, 848-852.
McLachlin, J.R., & Miller, L.K. (1994). Identification and characterization of *vlf-1*, a baculovirus gene involved in very late gene expression. *Journal of Virology*, 68, 7746-7756.
Mialhe, E., Quiot, J.M., & Paradis, S. (1984). Etablissement de deux lingnées celllulaires de *Spodoptera littoralis* [Lep.: Noctuidae] permissives pour des virus susceptibles d'être utilisés en lutte microbiologique. *Entomophaga* 29, 347-350.
Miller, L.K. (1988) Baculoviruses as gene expression vectors. *Annual Review of Microbiology* 42, 177-199.
Moguilevsky, N., Guillaume, J.-p.,Varsalona, F., Bulinckx, M., & Bollen, A. (1994). Correct processing of a chimeric ubiquitin-proapolipoprotein A-I fusion protein in baculovirus-infected insect cells. *Journal of Biotechnology*, 32, 39-43.
Monsma, S.A., Oomens, A.G.P., & Blissard, G.W. (1996). The gp64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. *Journal of Virology* 70, 4607-4616.
Morris, T.D, & Miller, L.K. (1994). Mutational analysis of a baculovirus major late promoter. *Gene* 140, 147-153.
Murhammer, D.W. (1996). Use of viral insecticides for pest control and production in cell culture. *Applied Biochemistry and Biotechnology* 59, 199-220.
Murphy, F.A., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., & Summers, M.D. (editors). (1995). Virus taxonomy. *Sixth Report of the International Committee on Taxonomy of Viruses*, pp. 586.
Nicolas, J., Cameron, K.R., Coleman, H., Newman, C., & Honess, R.W. (1992). Analysis of nucleotide sequence of the rightmost 43 kbp of herpesvirus saimiri (HVS) L-DNA, general conservation of genetic organization between HVS and Epstein-Barr virus. *Virology* 188, 296-310.
Nikas, I., McLauchlan, J., Davison, A.J., Taylor, W.R., & Clements, J.B. (1986). Structural features of ribonucleotide reductase. *Proteins* 1, 376-384.
Nillson, O., Arberg, A., Lundqvist, T., & Sjoberg, B.M. (1988). Nucleotide sequence of the gene coding for the large subunit of ribonucleotide reductase of *Escherichia coli*. Correction. *Nucleic Acids Research* 16, 4174.
Oakey, R., Cameron, I.R., Davis, B., Davis, E., & Possee, R.D. (1989). Analysis of transcription initiation in the *Panolis flammea* nuclear polyhedrosis virus polyhedrin gene. *Journal of General Virology* 70, 769-775.
Oers, M.M. van, & Vlak, J.M. (1997). The baculovirus 10-kDa protein. *Journal of Invertebrate Pathology* 70 , 1-17.
Oers, M.M. van. (1994). Functional analysis of the baculovirus 10 kilodalton protein. Ph.D. thesis, Agricultural University Wageningen.
Oers, M.M. van , Flipsen, J.T.M., Reusken, C.B.E.M., Sliwinsky, E.L., Goldbach, R.W., & Vlak, J.M. (1993). Functional domains of the p10 protein of *Autographa californica* nuclear polyhedrosis virus. *Journal of General Virology* 74, 563-574.
Ohkawa, T., Kajima, K., & Maeda, S. (1994). A cysteine protease encoded by the baculovirus *Bombyx mori* nuclear polyhedrosis virus. *Journal of Virology* 68, 6619-6625.
Ohlendorf, D.H., & Matthews, B.W. (1983). Structural studies of protein-nucleic acid interactions. *Annual Reviews of Biophysics and Bioengineering* 12, 259-284.
Ooi, B.G., Rankin, C., & Miller, L.K. (1989). Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. *Journal of Molecular Biology* 210, 721-736.
O'Reilly, D.R. (1995). Baculovirus-encoded ecdysteroid UDP-glycosyltransferases. *Insect Biochemistry and Molecular Biology* 25, 541-550.
O'Reilly, D.R., Miller, L.K, & Luckow, V.A. (1992). Baculovirus expression vectors: a laboratory manual. W.H. Freeman and Co., New York.
O'Reilly, D.R., & Miller, L.K. (1991). Improvement of a baculovirus pesticide by deletion of the *egt* gene. *Bio/Technology* 9, 1086-1089.
Palli, S.R., Caputo, G.F., Sohi, S.S., Brownright, A.J., Ladd, T.R., Cook, B.J., Primavera, M., Arif, B.M., & Retnakaran, A. (1996). CfMNPV blocks AcMNPV-induced apoptosis in a continuous midgut cell line. *Virology* 222, 201-213.
Parker, N.J., Begley, C.G., & Fox, R.M. (1991). Human M1 subunit of ribonucleotide reductase, cDNA sequence and expression in stimulated lymphocytes. *Nucleic Acids Research* 19, 3741.
Passarelli, A.L., & Miller, L.K. (1994). Identification and transcriptional regulation of the baculovirus *lef-6* gene. *Journal of Virology* 68, 4458-4467.
Passarelli, A.L., & Miller, L.K. (1993). Three baculovirus genes are involved in late and very late gene expression: *ie-1*, *ie-n* and *lef-2*. *Journal of Virology* 67, 2149-2158.
Passarelli, A.L., Todd, J.W., & Miller, L.K. (1994). A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases. *Journal of Virology* 68, 4673-4678.

Pearson, M., Bjornson, R., Pearson, G., & Rohrmann, G.F. (1992). The *Autographa californica* baculovirus genome: evidence for multiple replication origins. *Science* 257, 1382-1384.
Pearson, M.N., & Rohrmann, G.F. (1995). *Lymantria dispar* nuclear polyhedrosis virus homologous regions: characterization of their ability to function as replication origins. *Journal of Virology* 69, 213-221.
Pearson, M., Bjornson, R., Ahrends, C., & Rohrmann, G.F. (1993). Identification and characterization of a putative origin of DNA replication in the genome of a baculovirus pathogenic for *Orgyia pseudotsugata*. *Virology* 197, 715-725.
Pham, D.Q.-D., & Sivasubramanian, L.I. (1992). Sequence and in vitro translation analysis of a 1629-nucleotide ORF in *Autographa californica* nuclear polyhedrosis virus strain E2. *Gene* 122, 345-348.
Pham, D.Q.-D., Hice, R.H., Sivasubramanian, N., & Federici, B.A. (1993). The 1629-bp open reading frame of the *Autographa californica* multicapsid nuclear polyhedrosis virus encodes a virion structural protein. *Gene* 123, 275-280.
Possee, R.D., Hirst, M., Jones, L.D., & Bishop, D.H.L. (1993). Field tests of genetically engineered baculoviruses. Pp. 23-26 in: Opportunities for molecular biology in crop protection, (eds. D.J. Beadle, D.H.L. Bishop, L.G. Copping, G.K. Dixon, & D.W. Holleman *et al.*). BCPC Monograph No. 55, Farnham.
Possee, R.D., Sun, T-P, Howard, S.C., Ayres, M.D., Hill-Perkins, M., & Gearing K.L. (1992). Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis 9.4 kbp *Eco*RI-I and -R (polyhedrin gene) region. (1992). *Virology* 185, 229-241.
Possee, R.D., Sun, T.P., Howard, S.C., Ayres, M.D., Hill-Perkins, M., & Gearing, K.L. (1991). Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis virus 9.4 kbp EcoRI-I and -R (polyhedrin gene) region. *Virology* 185, 229-241.
Possee, R.D., & Howard, S.C. (1987). Analysis of the polyhedrin gene promoter of the *Autographa californica* nuclear polyhedrosis virus. *Nucleic Acids Research 15*, 10233-10248.
Ptashne, M. (1988). How eukaryotic transcriptional activators work. *Nature* 335, 683-689.
Putterman, D., Pepinsky, R.B., & Vogt, V.M. (1990). Ubiquitin in avian leukosis virus. *Virology* 176, 633-637.
Rankin, C., Ooi, B.G., & Miller, L.K. (1988). Eight base pairs encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin gene expression. *Gene* 70, 39-49.
Rankin, C., Ladin, B.F., & Weaver, R.F. (1986). Physical mapping of temporally regulated, overlapping transcripts in the region of the 10K protein gene in *Autographa californica* nuclear polyhedrosis virus. *Journal of Virology* 57, 18-27.
Rasmussen, C., Leisy, D.J., Ho, P.S., & Rohrmann, G.F. (1996). Structure-function analysis of the *Autographa californica* multinucleocapsid nuclear polyhedrosis virus homologous region palindromes. *Virology* 224, 235-245.
Reichard, P. (1993). From RNA to DNA, why so many ribonucleotide reductases? *Science* 260, 1773-1777.
Reichard, P. (1988). Interactions between deoxyribonucleotide and DNA synthesis. *Annual Review of Biochemistry* 57, 349-374.
Reilly, L.M., & Guarino, L.A. (1994). The pk-1 gene of *Autographa californica* multinucleocapsid nuclear polyhedrosis virus encodes a protein kinase. *Journal of General Virology* 75, 2999-3006.
Reilly, L.M., & Guarino, L.A. (1996). The viral ubiquitin gene of *Autographa californica* nuclear polyhedrosis virus is not essential for viral replication. *Virology* 218, 243-247.
Riggio, M.P., & Onions, D.E. (1994). Sequences of the ribonucleotide reductase-encoding genes of equine herpesvirus 4. *Gene* 143, 217-222.
Rodems, S.M., & Friesen, P.D. (1993). The *hr*5 transcriptional enhancer stimulates early expression from the *Autographa californica* nuclear polyhedrosis virus genome but is not required for virus replication. *Journal of Virology* 67, 5776-5785.
Rodems, S.M., & Friesen, P.D. (1995). Transcriptional enhancer activity of *hr*5 requires dual-palindrome half sites that mediate binding of a dimeric form of the baculovirus transregulator IE1. *Journal of Virology* 69, 5368-5375.
Rodriguez, J.M., Salas, M.L., & Vinuela, E. (1992). Genes homologous to ubiquitin-conjugating proteins and eukaryotic transcription factor SII in African swine fever virus. *Virology* 186, 40-52.
Roelvink, P.W., Corsaro, B.G., & Granados, R.R. (1995). Characterization of the *Helicoverpa armigera* and *Pseudaletia unipuncta* granulovirus enhancin genes. *Journal of General Virology* 76, 2693-2705.
Rohrmann, G.F. (1992). Baculovirus structural proteins. *Journal of General Virology* 73, 749-761.
Rohrmann, G.F. (1986). Polyhedrin structure. *Journal of General Virology* 67, 1499-1513.
Rubin, H., Salem, J.S., Li, L.-S., Yang, F.-D., Mama, S., Wang, Z.-M., Fisher, A., Hamann, C.S., & Cooperman, B.S. (1993). Cloning, sequence determination and regulation of the ribonucleotide reductase subunits from *Plasmodium falciparum*, a target for antimalarial therapy. *Proceedings of the National Academy of Sciences USA* 90, 9280-9284.
Russell, R.L.Q., Funk, C.J., & Rohrmann, G.F. (1997). Association of a baculovirus-encoded protein with the capsid basal region. *Virology* 227, 142-152.
Russell, R.L.Q., & Rohrmann, G.F. (1993). Nucleotide sequence of the ubiquitin-39K region from the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome. *Journal of General Virology* 74, 1191-1195.
Russell, R.L.Q., Pearson, M.N., & Rohrmann, G.F. (1991). Immunoelectron microscopic examination of *Orgyia*

pseudotsugata multicapsid nuclear polyhedrosis virus-infected *Lymantria dispar* cells: time course and localization of major polyhedron-associated proteins. *Journal of General Virology* 72, 275-283.
Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989). Molecular cloning, A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
Sanger, F., Nicklen, S., & Coulson, A.R. (1977). Nucleotide sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences USA* 74, 5463-5467.
Schmitt, J.F.C., & Stunnenberg, H.G. (1988). Sequence and transcriptional analysis of the vaccinia virus *Hin*dIII I fragment. *Journal of Virology* 62, 1889-1897.
Senkevich, T.G., Bugert, J.J., Sisler, J.R., Koonin, E.G., Darai, G., & Moss, B. (1996). Genome sequence of a human tumorigenic poxvirus: prediction of specific host response-evasion genes. *Science* 273, 813-816.
Shchelkunov, S.N., Blinov, V.M., Totmenin, A.V., Marennikova, S.S., Kolykhalov, A.A., Frolov, I.V., Chizhikov, V.E., Gytorov, V.V., Gashikov, P.V., & Belanov, E.F. (1993). Nucleotide sequence analysis of variola virus *Hin*dIII M, L, I genome fragments. *Virus Research* 27, 25-35.
Simard, C., Langlois, I., Styger, D., Vogt, B., Vlcek, C., Chalifour, A., Trudel, M., & Schwyzer, M. (1995). Sequence analysis of the UL39, UL38, and UL37 homologues of bovine herpesvirus 1 and expression studies of UL40 and UL39, the subunits of ribonucleotide reductase. *Virology* 212, 734-740.
Slabaugh, M.B., & Mathews, C.K. (1984). Vaccinia virus-induced ribonucleotide reductase can be distinguished from host cell activity. *Journal of Virology* 52, 501-506.
Smith, I.R.L., van Beek, N.A.M., Podgwaite, J.D., & Wood, H.A. (1988). Physical map and polyhedrin gene sequence of *Lymantria dispar* nuclear polyhedrosis virus. *Gene* 71, 97-105.
Smith, G.E., Vlak, J.M., & Summers, M.D. (1983). Physical analysis of *Autographa californica* nuclear polyhedrosis virus transcripts for polyhedrin and 10,000-molecular-weight protein. *Journal of Virology* 45, 215-225.
Smith, G.E., Vlak, J.M., & Summers, M.D. (1982). *In vitro* translation of *Autographa californica* nuclear polyhedrosis virus early and late mRNAs. *Journal of Virology* 44, 199-208.
Smith, G.E., & Summers, M.D. (1982). DNA homology among subgroup A, B and C baculoviruses. *Virology* 123, 393-406.
Smits, P.H., & Vlak, J.M. (1994). Registration of the first viral insecticide in the Netherlands: the development of Spod-X, based on *Spodoptera exigua* nuclear polyhedrosis virus. *Mededelingen Faculteit Landbouwwetenschappen Universiteit Gent* 59, 385-392.
Smits, P.H., van de Vrie, M., & Vlak, J.M. (1988). Nuclear polyhedrosis virus for control of *Spodoptera exigua* on glasshouse crops. *Entomologia Experimentalis et Applicata* 43, 73-80.
Smits, P.H. (1987). Nuclear polyhedrosis virus as biological control agent of *Spodoptera exigua*. Ph.D.thesis, Wageningen Agricultural University, Wageningen.
Spence, J., Sadis, S., Haas, A.L., & Finley, D. (1995). An ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Molecular and Cellular Biology* **15**, 1265-1273.
Steinhaus, E.A. (1956). Microbial control - the emergence of an idea. *Hilgardia* 26, 107-160.
Stewart, L.M.D., Hirst, M., López Ferber, M., Merryweather, A.T., Caylay, P.J., & Possee, R.D. (1991). Construction of an improved baculovirus insecticide containing an insect-specific toxin gene. *Nature* 352, 85-88.
Strien, E.A. van, Faktor, O., Hu, Z.H., Zuidema, D., Goldbach, R.W., & J.M. Vlak. (1997a). Baculoviruses contain a gene for the large subunit of ribonucleotide reductase. *Journal of General Virology* 78, 000-000.
Strien, E.A. van, Gerrits, H., Zuidema, D., Goldbach, R.W., & J.M. Vlak. (1997b). Unusual characteristics of the transactivator gene *ie1* of *Spodoptera exigua* nucleopolyhedrovirus. Manuscript in preparation.
Strien, E.A. van, Jansen, B.J.H., Mans, R.J., Zuidema, D., & J.M. Vlak. (1996). Sequence and transcriptional analysis of the ubiquitin gene cluster in the genome of *Spodoptera exigua* nucleopolyhedrovirus. *Journal of General Virology* 77, 2311-2319.
Strien, E.A. van, Zuidema, D., Goldbach, R.W., & J.M. Vlak. (1992). Nucleotide sequence and transcriptional analysis of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus. *Journal of General Virology* 73, 2813-2821.
Stubbe, J. (1990). Ribonucleotide reductases, amazing and still confusing. *Journal of Biological Chemistry* 256, 5329-5332.
Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., & Waterston, R. (1992). The *C. elegans* genome sequencing project, a beginning. *Nature* 356, 37-41.
Summers, M.D., & Smith, G.E. (1987). A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experiment Station Bulletin No. 1555.
Swain, M.A., & Galloway, D.A. (1986). Herpes simplex virus specifies two subunits of ribonucleotide reductase encoded by 3' coterminal transcripts. *Journal of Virology* 57, 802-808.
Swofford, D.L. (1993). PAUP: Phylogenetic analysis using parsimony. Version 3.1. Computer program distributed by the Illinois Natural History Survey, Champaign, IL.
Tautz, D., & Renz, M. (1983). An optimized freeze-squeeze method for the recovery of DNA fragments from

agarose gels. *Analytical Biochemistry* 132, 14-19.
Telford, E.A.R., Watson, M.S., McBride, K., & Davison, A.J. (1992). The DNA sequence of equine herpesvirus-1. *Virology* 189, 304-316.
Tengelsen, L.A., Slabaugh, M.M., Bibler, J.K., & Hruby, D.E. (1988). Nucleotide sequence and molecular genetic analysis of the large subunit of ribonucleotide reductase encoded by vaccinia virus. *Virology* 164, 121-131.
Thelander, L., & Reichard, P.(1979). Reduction of ribonucleotides. *Annual Review of Biochemistry* 48, 133-158.
Theilmann, D.A., Chantler, J.K., Stewart, S., Flipsen, H.T.M., Vlak, J.M., & Crook, N.E. (1996). Characterization of a highly conserved baculovirus structural protein that is specific for occlusion-derived virions. *Virology* 218, 148-158.
Theilmann, D.A., & Stewart, S. (1993). Analysis of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus trans-activators IE1 and IE2 using monoclonal antibodies. *Journal of General Virology* 74, 1819-1826.
Theilmann, D.A., & Stewart, S. (1992a). Molecular analysis of the trans-activating IE-2 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 187, 84-96.
Theilmann, D.A., & Stewart, S. (1992b). Tandemly repeated sequence at the 3' end of the IE-2 gene of the baculovirus *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus is an enhancer element. *Virology* 187, 97-106.
Theilmann, D.A., & Stewart, S. (1991). Identification and characterization of the IE-1 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 180, 492-508.
Thiem, S.M., Du, X., Quentin, M.E., & Berner, M.M. (1996). Identification of a baculovirus gene that promotes *Autographa californica* nuclear polyhedrosis virus replication in a non-permissive cell line. *Journal of Virology* 70, 2221-2229.
Todd, J.W., Passarelli, A.L., & Miller, L.K. (1995). Eighteen baculovirus genes, including lef-11, p35, 39K, and p47, support late gene expression. *Journal of Virology* 69, 968-974.
Tomalski, M.D., & Miller, L.K. (1991a). Insect paralysis by baculovirus-mediated expression of a mite neurotoxin gene. *Nature* 352, 82-85.
Tomalski, M.D., & Miller, L.K. (1991b). Insect paralysis by baculovirus-mediated expression of a insect-specific toxin gene. *Nature* 352, 85-88.
Tomalski, M.D., Wu, J., & Miller, L.K. (1988). The location, sequence, transcription and regulation of a baculovirus DNA polymerase gene. *Virology* 167, 591-600.
Tseng, M.-J., Hilfinger, J.M., Walsh, A., & Greenberg G.R. (1988). Total sequence, flanking regions and transcripts of bacteriophage T4 nrdA gene, coding for alpha chain of ribonucleoside diphosphate reductase. *Journal of Biological Chemistry* 263, 16242-16251.
Uhlin, A., & Eklund, H. (1994). Structure of ribonucleotide reductase protein R1. *Nature* 370, 533-539.
Vialard, J.E., & Richardson, C.D. (1993). The 1629-nucleotide open reading frame located downstream of the *Autographa californica* nuclear polyhedrosis virus polyhedrin gene encodes a nucleocapsid-associated phosphoprotein. *Journal of Virology* 67, 5859-5866.
Vlak, J.M. (1992). Genetic engineering of baculoviruses for insect control. pp.90-127 in: Molecular approaches to pure and applied entomology. (eds. J.G. Oakeshott & M.J. Whitten). Springer-Verlag, New York.
Vlak, J.M., & Rohrmann, G.F. (1985). The nature of polyhedrin. In *Viral Insecticides for Biological Control*, pp. 489-544. (eds. K. Maramorosch & K.E. Sherman). Orlando, Academic Press.
Vlak, J.M., & Smith, G.E. (1982). Orientation of the genome of *Autographa californica* nuclear polyhedrosis virus: a proposal. *Journal of Virology* 41, 1118-1121.
Vlak, J.M. (1982). Restriction endonucleases as tools in baculovirus identification. Proceedings of the 25th Annual Meeting of the Society for Invertebrate Pathology, pp. 218-225.
Vlak, J.M. (1979). The proteins of nonoccluded *Autographa californica* nuclear polyhedrosis virus produced in an established cell line of *Spodoptera frugiperda*. *Journal of Invertebrate Pathology* 34, 110-118.
Wang, Y., & Hall, J.D. (1990). Characterization of a major DNA-binding domain in the herpes simplex virus type 1 DNA-binding protein. *Journal of Virology* 64, 2082-2089.
Washburn, J.O., Kirkpatrick, B.A., & Volkman, L.E. (1996). Insect protection against viruses. *Nature* 383, 767.
Washburn, J.O., Kirkpatrick, B.A., & Volkman, L.E. (1995). Comparitive pathogenesis of *Autographa californica* M nuclear polyhedrosis virus in larvae of *Trichoplusia ni* and *Heliothis virescens*. *Virology* 209, 561-568.
Wilk, F. van der, Lent, J. van, & Vlak, J.M. (1987). Immunogold detection of polyhedrin, p10 and virion antigens in *Autographa californica* nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells. *Journal of General Virology* 68, 2615-2623.
Williams, G.V., & Faulkner, P. (1996). Replication patterns and cytopathology of cells infected with baculoviruses. pp. 95-110 in: Insect cell cultures. Fundamental and applied aspects. (eds. J.M. Vlak, C.D. de Gooijer, J. Tramper & H.G. Miltenburger) Kluwer, Dordrecht.
Williams, G.V., Rohel, D.Z., Kuzio, J., & Faulkner, P. (1989). A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion/deletion mutants. *Journal of General Virology* 70, 187-202.
Wilson, J.A., Hill, J.E., Kuzio, J., & Faulkner, P. (1995). Characterization of the baculovirus *Choristoneura*

fumiferana multicapsid nuclear polyhedrosis virus p10 gene that contains a coiled-coil domain. *Journal of General Virology* 76, 2923-2932.
Wu, X., Stewart, S., & Theilmann, D.A. (1993). Alternative transcriptional initiation as a novel mechanism for regulating expression of a baculovirus *trans*activator. *Journal of Virology* 67, 5833-5842.
Xie, W., & Rothblum, L.I. (1991). Rapid small-scale RNA isolation from tissue culture cells. *Bio/Techniques* 11, 323-327.
Xie, W.-D., Arif, B., Dobos, P., & Krell, P. (1995). Identification and analysis of a putative origin of DNA replication in the *Choristoneura fumiferana* multinucleocapsid nuclear polyhedrosis virus genome. *Virology* 209, 409-419.
Xu, B., Yoo, S., & Guarino, L.A. (1995). Differential transcription of baculovirus late and very late promoters: fractionation of nuclear extracts by phosphocellulose chromatography. *Journal of Virology* 69, 2912-2917.
Xue, D., & Horvitz, H.R. (1995). Inhibition of the *Caenorhabditis elegans* cell-death protease CED-3 by a CED-3 cleavage site in baculovirus p35 protein. *Nature* 377, 248-251.
Yagle, K., & McEntee, K. (1990). The DNA damage-inducible gene DIN1 of *Saccharomyces cerevisiae* encodes a regulatory subunit of ribonucleotide reductase and is identical to RNR3. *Molecular and Cellular Biology* 10, 5553-5557.
Yang, F., Lu, G., & Rubin, H. (1994). Isolation of ribonucleotide reductase from *Mycobacterium tuberculosis* and cloning, expression and purification of the large subunit. *Journal of Bacteriology* 176, 6738-6743.
Yang, C.L., Stetler, D.A., & Weaver, R.F. (1991). Structural comparison of the *Autographa californica* nuclear polyhedrosis virus induced RNA polymerase and three nuclear RNA polymerases from the host, *Spodoptera frugiperda*. *Virus Research* 20, 251-264.
Yoo, S., & Guarino, L.A. (1994). The *Autographa californica* nuclear polyhedrosis virus *ie2* gene encodes a transcriptional regulator. *Virology* 202, 746-753.
Zanotto, P.M. de A., Gibbs, M.J., Gould, E.A., & Holmes, E.C. (1996). A reevaluation of the higher taxonomy of viruses based on RNA polymerases. *Journal of Virology* 70, 6083-6096.
Zanotto, P.M. de A., Sampaio, M.J.A., Johnson, D.W., Roche, T.L., & Maruniak, J.E. (1993a). The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *Journal of General Virology* 73, 1049-1056.
Zanotto, P.M. de A., Kessing, B.D., & Maruniak, J.E. (1993b). Phylogenetic relationships among baculoviruses, evolutionary rates and host associations. *Journal of Invertebrate Pathology* 62, 147-164.
Zanotto, P.M. de A., Sampaio, M.J.A., Johnson, D.W., Rocha, T.L., &. Maruniak, J.E. (1992). The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *Journal of General Virology* 73, 1049-1056.
Zhang, Y,. Wu, X., & Li, Z. (1995). P10 genes of BmNPV and AcMNPV. *Sci. China B; Chemical Life Sciences & Earth Sciences* 38, 50-59.
Zuidema, D., Oers, M.M. van, Strien, E.A. van, Caballero, P.C., Klok, E.-J., Goldbach, R.W., & J.M. Vlak. (1993). Nucleotide sequence and transcriptional analysis of the p10 gene of *Spodoptera exigua* nuclear polyhedrosis virus. *Journal of General Virology* 74, 1017-1024.
Zuidema, D., Klinge-Roode, E.C., Lent, J.W.M. van, & Vlak, J.M. (1989). Construction and analysis of an *Autographa californica* nuclear polyhedrosis virus mutant lacking the polyhedral envelope. *Virology* 173, 98-108.

ACCOUNT

Strien, E.A. van, D. Zuidema and J.M. Vlak. 1991. Nucleotide sequence of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus. *IOBC/WPRS Bulletin* 14-7, 180-184.

Strien, E.A. van, D. Zuidema, R.W. Goldbach and J.M. Vlak. 1992. Nucleotide sequence and transcriptional analysis of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus. *Journal of General Virology* 73, 2813-2821.

Zuidema, D., M.M. van Oers, E.A. van Strien, P.C. Caballero, E.J. Klok, R.W. Goldbach and J.M. Vlak. 1993. Nucleotide sequence and transcriptional analysis of the p10 gene of *Spodoptera exigua* nuclear polyhedrosis virus. *Journal of General Virology* 74, 1017-1024.

Strien, E.A. van, B.J.H. Jansen, R.J. Mans, D. Zuidema and J.M. Vlak. 1996. Sequence and transcriptional analysis of the ubiquitin gene region in the genome of *Spodoptera exigua* nucleopolyhedrovirus. *Journal of General Virology* 77, 2321-2328.

Heldens, J.G.M., E.A. van Strien, D.J. Leisy, A.M. Feldmann, D. Muñoz, P. Kulcsar, D. Zuidema, R.W. Goldbach and J.M. Vlak. 1996. *Spodoptera exigua* multicapsid nucleopolyhedrovirus deletion mutants generated in cell culture lack virulence *in vivo*. *Journal of General Virology* 77, 3127-3134.

Strien, E.A. van, D. Zuidema, R.W. Goldbach and J.M. Vlak. 1997. Baculoviruses contain a large subunit of ribonucleotide reductase. *Journal of General Virology*, 78, 000-000

Strien, E.A. van, H. Gerrits, D. Zuidema, R.W. Goldbach and J.M. Vlak. 1997.Unusual characteristics of the transactivator gene *ie1* of *Spodoptera exigua* nucleopolyhedrovirus. submitted.

Broer, R., J.G.M. Heldens, E.A. van Strien, D. Zuidema and J.M. Vlak. 1997. Identification and characterization of multiple *hr*-type putative replication origins within the genome of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus. submitted.