The genome of *Spodoptera exigua* multicapsid nucleopolyhedrovirus: a study on unique features

W.F.J. IJkel

Promotoren: prof. dr. R.W. Goldbach
                Hoogleraar in de Virologie

                prof. dr. J.M. Vlak
                Persoonlijk Hoogleraar bij de Leerstoelgroep Virologie

Co-promotor: dr. D. Zuidema
                Universitair Docent bij de Leerstoelgroep Virologie

Samenstelling promotiecommissie:
                dr. J.A. Jehle (Zentrum Grüne Gentechnik, Germany)
                prof. dr. W.M. de Vos (Wageningen Universiteit)
                prof. dr. W.J. Stiekema (Plant Research International, Wageningen)
                dr. F.A.M. Rijsewijk (Dept. Mammalian Virology, ID-Lelystad)

The genome of *Spodoptera exigua* multicapsid nucleopolyhedrovirus: a study on unique features

W.F.J. IJkel

Stellingen

1. Kennis van de volledige genensamenstelling van SeMNPV geeft nog geen inzicht in de specifieke biologische eigenschappen van dit virus.
   Dit proefschrift.

2. Baculovirussen van het genus Nucleopolyhedrovirus, subtype groep II, gebruiken geen GP64 homoloog maar een nieuw type fusie-eiwit om insectencellen te infecteren.
   Dit proefschrift.

3. De voortdurende wijzigingen in de schrijfwijze van virusnamen zijn eerder een reflectie van de wisselende samenstelling van de ICTV dan een gevolg van voortschrijdend wetenschappelijk inzicht.
   Murphy et al. (Eds.) (1995). In "Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses", pp. 1-586. Springer-Verlag, Wien New York.
   Fauquet and Pringle (1999) Archives of Virology 144, 2265-2271.
   van Regenmortel et al. (Eds.) (2000). In "Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses", pp. 1-1162. Academic press, New York.

4. Alhoewel de term 'fylogenetische analyse' anders doet vermoeden, geeft deze geen uitsluitsel over de genetische oorsprong van genen.
   Ronneberg et al. (2000). PNAS USA 97, 13690-13695.

5. Mensen die 'lobbyen' tegen genetische modificatie ontkennen hun eigen identiteit.
   International human genome sequencing consortium (2001). Nature 409, 860-921.
   Venter et al. (2001). Science 291, 1304-1351.

6. De grootte van een ramp bepaalt de mate van verandering in beleid.

7. De aanduiding "rondje om de kerk" voor het slepende conflict tussen NS-directie en rijdend personeel wijst aan hoe dit conflict opgelost gaat worden: men moet erin of eraan geloven.

Stellingen behorende bij het proefschrift
'The genome of *Spodoptera exigua*
multicapsid nucleopolyhedrovirus:
a study on unique features'
W.F.J. IJkel, Wageningen, 1 juni 2001

Aan mijn ouders J.N. IJkel en D.C. IJkel-van Apeldoorn

Voorwoord

Het voorwoord van dit proefschrift wil ik gebruiken om een aantal mensen te bedanken, die ieder op hun eigen wijze hebben geholpen bij het voltooien van mijn promotie onderzoek. Allereerst wil ik mijn co-promoter Douwe Zuidema en beide promotoren Just Vlak en Rob Goldbach bedanken voor hun zeer positieve inzet en stimulerende begeleiding. De altijd kritische blik waarmee de manuscripten werden bekeken, maar ook de snelheid waarmee ze geretourneerd werden, hebben zeker bijgedragen aan de prettige sfeer waarin dit proefschrift tot stand is gekomen.

Dankbaar ben ik ook voor de hulp, die ik heb gehad van Angela Vermeesch-Markslag bij de uitvoering van de weefselkweek experimenten. De vele gezellige uren, die we samen werkend in één flowkast hebben doorgemaakt, zal ik niet snel vergeten. Ook Els Roode, Magda Usmany en Jan van Lent wil ik bedanken voor alle assistentie. Wout Rozeboom verdient een pluim voor bewezen diensten binnen maar zeker ook buiten kantooruren.

Natuurlijk wil ik ook Antoine Beerens, Jeroen Witteveldt, Joke van Vugt, Marjolein Op den Brouw en Robert-Jan Lebbink niet onvermeld laten, die allen als student, vele proeven hebben uitgevoerd in het kader van dit proefschrift en waarvan sommigen een plaatsje hebben veroverd in dit proefschrift. Bedankt voor alle inzet en gezelligheid.

Naast alle andere kamergenoten, wil ik in het bijzonder Mariëlle van Hulten en Marcel Westenberg bedanken voor hun nauwe betrokkenheid en prettige sfeer op onze kamer. Uiteraard is de prettige werksfeer ook te danken aan alle andere baculo's, whispo's, tospo's en de inwoners van 'small chinatown'. De internationale en informele omgang is verrijkend en heeft me goed gedaan.

Tenslotte wil ik mijn ouders bedanken voor hun voortdurende interesse in en relativering van mijn promotie onderzoek. Dit heb ik mede tot uiting gebracht door dit proefschrift aan jullie op te dragen.

Heleen, je bent het levende bewijs dat er veel meer dan 'exacte' wetenschap bestaat. Je hulp, steun en aanmoediging zijn hartverwarmend en ik hoop dat we in de komende tijd met heel veel plezier van het (nieuwe) leven zullen genieten.

Contents

Chapter 1

General Introduction

Introduction to the baculoviruses

There are over a million different species of insects, of which many are infected by a diverse spectrum of viruses. Although most of these insect viruses have little direct effect on human health, they nevertheless represent important viruses from an environmental point of view. Viral diseases in beneficial insects can lead to ecological imbalance, while diseases of insect pests are often welcomed as a means of reducing agricultural and medical problems. Baculoviruses are beneficial viruses by their potential to control insect pests of mainly moth and butterfly larvae. These larvae often cause severe feeding damage to agriculturally important crops, ornamental plants or to forest trees. Along with the research to improve their efficacy and safety as bio-pesticides came the use of these viruses and their genes in both basic and applied biomedical as well as veterinary research because of their ability to produce large quantities of proteins from foreign genes. This thesis focuses on several fundamental aspects of the baculovirus *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV), such as genome organization, the viral infection route and characterization of genes unique to this virus. In this chapter the baculovirus family will be introduced and the outline of the thesis be given.

Host range and persistence in nature

Baculoviruses exclusively infect arthropods, predominantly holometabolous insects. Nowadays baculovirus diseases have been described in over 800 different species of insects, reflecting the enormous diversity in baculoviruses. Most baculoviruses have been isolated from insect species of the lepidopteran order (butterflies and moths). In addition, some baculoviruses are also pathogenic for members of the order *Hymenoptera* (sawflies), *Diptera* (flies, mosquitoes), *Coleoptera* (beetles) and the crustacean order *Decapoda* (shrimp) (Adams and McClintock, 1991; Federici and Maddox, 1996; Federici, 1997; Summers, 1977). The majority of baculoviruses has a very limited host range and infects only closely related insects within a single order (Table 1.1). Although baculoviruses are commonly isolated from insects, they are also designed to survive outside their host. They can reside in soil, water or in the crevices

of plants for years before infecting arthropods that inhabit terrestrial or marine ecosystems (Adams and McClintock, 1991).

**Table 1.1** Host range of selected baculovirus in insects.

Virus species	Insect family	Insect species	Reference
Autographa californica MNPV	13	73	Adams & McClintock, 1991
Anagrapha falcifera MNPV	11	31	EDWIP
Bombyx mori NPV	7	9	EDWIP
Mamestra brassicae MNPV	5	44	EDWIP
Spodoptera litoralis MNPV	2	4	EDWIP
Orgyia pseudotsugata MNPV	2	3	EDWIP
Xestia c-nigrum GV	1	6	Goto *et al.*, 1992
Cydia pomonella GV	1	4	EDWIP
***Spodoptera exigua* MNPV**	**1**	1	EDWIP
Buzura suppressaria SNPV	1	1	EDWIP
Spodoptera litura MNPV	1	1	EDWIP
Lymantria dispar MNPV	1	1	EDWIP
Plutella xylostella GV	1	1	EDWIP

Virus host ranges are derived from the Ecological Database of the World's Insect Pathogens (EDWIP) on http://insectweb.inhs.uiuc.edu/Pathogens/EDWIP/index.html.

Taxonomy and structure

The baculoviruses are a family (*Baculoviridae*) of large, enveloped double-stranded DNA viruses that are characterized by their ability to form proteinaceous occlusion bodies (OBs) within infected cells. The family is taxonomically subdivided into two genera, *Nucleopolyhedrovirus* (NPV) and *Granulovirus* (GV), based upon distinct OB morphology (Blissard *et al.*, 2000). The NPVs produce large (ranging in size from 0.15-15 µm) polyhedron-shaped OBs called polyhedra that contain many virions, whereas the GVs have smaller (about 0.3 x 0.5 µm) ovicylindrical-shaped OBs called granules that normally contain a single virion. A baculovirus virion consists of one or more rod-shaped nucleocapsids (30-60 nm in diameter and 250-300 nm in length) that have a distinct structural polarity and are enclosed within an envelope. The name baculovirus (from *baculum* meaning stick) is derived from these rod-shaped nucleocapsids. The latter are composed of a single molecule of circular viral DNA of approximately 80-180 kbp in size packaged in a rod-shaped protein structure. The

NPVs are designated as single (S) or multiple (M) depending on the potential number of nucleocapsids (NC) packaged in an occluded virion. Although this differentiation is commonly used in describing NPVs, the biological significance of these morphotypes is unclear. Besides the virions present in OBs, baculoviruses usually produce a second type of virion known as the budded virus (BV) (Fig. 1.1).
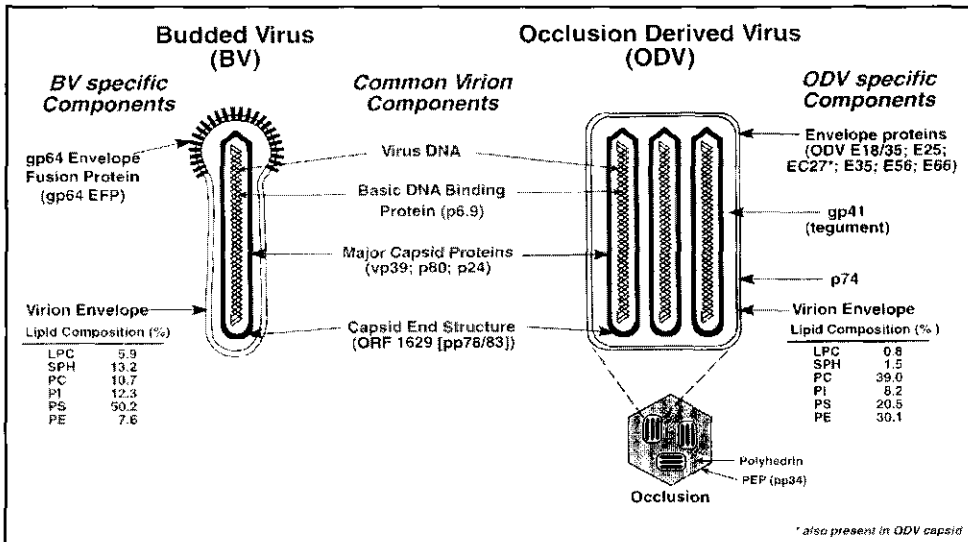
Fig. 1.1 Structural composition of the budded virus and the occlusion derived virus. Figure adapted from Funk *et al.* (1997). Proteins common to both phenotypes are indicated in the middle of the figure. Proteins specific to BV or ODV are indicated on the left and right, respectively. Lipid compositions of the BV and ODV envelopes are derived from AcMNPV infected Sf-9 cells (Braunagel and Summers, 1994) and indicated (LPC, LysoPhosphaditylCholine; SPH, sphingomyelin; PI, P-inositol; PS, P-serine; PE, P-ethanolamine).

To date, twelve different baculovirus species have been definitely assigned to the NPV genus and five to the GV genus (Blissard *et al.*, 2000). Among these are the NPVs *Autographa californica* (Ac), *Bombyx mori* (Bm), *Orgyia pseudotsugata* (Op), *Lymantria dispar* (Ld) and *Spodoptera exigua* (Se) as well as the GVs *Xestia c-nigrum* (Xc) and *Trichoplusia ni* (Tn). All baculoviruses are named after the host from which they were first isolated. From a taxonomic viewpoint this is rather confusing because the same virus may infect several insect species and be called by a variety of names. Therefore, it is necessary to establish the identity of baculovirus species for example by genome sequencing to allow a definite taxonomic status. In recent years, the increasing amount of sequence data available enabled preliminary studies of the phylogeny of this virus family. Such studies have now been carried out using the

sequences of several different genes. These studies confirmed that NPVs and GVs represent two well-separated clades. In addition, based upon the sequences of the *polyhedrin/granulin* (Zanotto *et al.*, 1993), *ecdysteroid UDP-glucosyltransferase* (Chen *et al.*, 1997), *late essential factor-2* (Chen *et al.*, 1999) and *DNA polymerase* (Bulach *et al.*, 1999) genes, a further subdivision has been proposed distinguishing lepidopteran NPVs into two distinct groups, named group I and II. The baculoviruses AcMNPV (Ayres *et al.*, 1994), BmNPV (Gomi *et al.*, 1999) and OpMNPV (Ahrens *et al.*, 1997) are members of group I, whereas SeMNPV (IJkel *et al.*, 1999) and LdMNPV (Kuzio *et al.*, 1999) belong to group II. For the further purposes of this thesis, we focus mainly on the NPVs, which have the widest ranges of hosts and have been studied most intensively.

Infection cycle

In the environment, NPVs are commonly found on plant surfaces and in the soil as stable polyhedra. A NPV infection starts with the uptake of polyhedra by the insect larvae. Upon ingestion the OBs dissolve in the alkaline environment of the larval midgut liberating numerous enveloped virions, which are termed 'occlusion derived virus' (ODV). After direct membrane fusion of the ODV virion envelope with the microvilli of midgut columnar epithelial cells (Granados, 1978; Granados and Lawler, 1981; Horton and Burand, 1993), the virions are uncoated and transported to the nucleus, where gene expression, DNA replication and assembly of progeny NCs occur (Fig. 1.2). Progeny NCs can be observed as early as 8 h post infection (p.i.) assembling within and around a dense virogenic stroma that develops in the enlarged nucleus upon infection. The newly assembled NCs then migrate from the virogenic stroma towards the plasma membrane on the basal side of the epithelial cell.

During infection of group I NPVs, a viral encoded major envelope glycoprotein, GP64, is synthesized and transported to the plasma membrane (Volkman and Goldsmith, 1984; Volkman *et al.*, 1984; Blissard and Rohrmann, 1989; Oomens *et al.*, 1995). When a single NC buds from the basal side of the epithelial cells into the hemocoel or tracheoblast, it acquires a loosely adhering plasma membrane envelope containing the GP64 protein. This protein is required for efficient budding (Monsma *et al.*, 1996; Oomens and Blissard, 1999). This second virion phenotype is termed 'budded virus' (BV) and believed to be essential for systemic infection, mediating movement of the virus from midgut to other tissues (fat body, muscle, trachea, hemocytes, epithelial cells) and propagating the infection from cell to cell within the infected animal (Flipsen, 1995; Keddie *et al.*, 1989; Granados and Lawler, 1981).
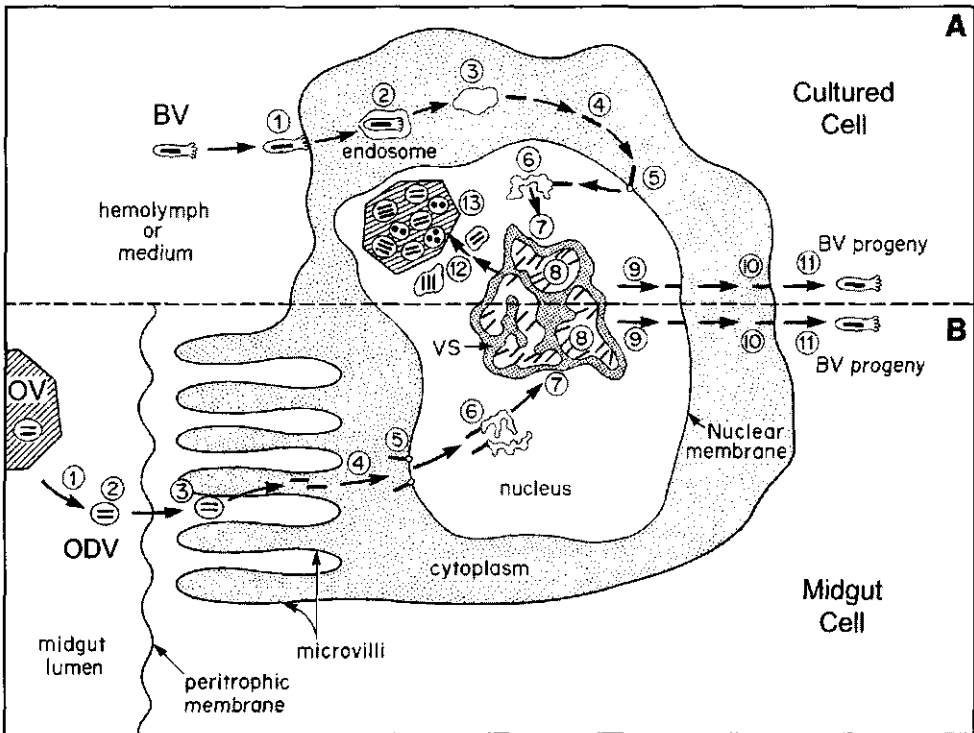
Fig. 1.2 The replication cycle of a baculovirus, specifically a MNPV.

Figure adapted from Miller (1996). (A): The process of budded virus (BV) infection of a cultured cell. A BV attaches to the cell surface (1) and enters by endocytosis (2). As the endosome acidifies, virus and endosomal envelopes fuse (3), releasing nucleocapsids (NCs) into the cytoplasm. NCs move to the nucleus (4) where they interact with a nuclear pore (5). Upon entering the nucleus, the core is released (6) and the viral DNA transcribed (7), replicated and packaged into NCs (8) in association with the virogenic stroma (VS). During the late stage of infection, NCs leave the nucleus (9) and travel to the plasmamembrane where they bud (10) to produce BV (11). During the very late phase, NCs are enveloped within the nucleus (12) and then embedded in a polyhedrin matrix to form occlusion bodies (13). (B): The process of infection of a midgut epithelial cell. The polyhedrin matrix of an ingested occluded virus (OV) is dissolved in the midgut lumen, releasing (1) occlusion derived virions (ODV) which cross the peritrophic membrane (2). The membranes of the ODVs fuse with the membranes of the microvilli of the cell, releasing NCs into the cytoplasm. The remaining events appear to be similar to those in cultured cells, except that little polyhedra formation occurs in columnar cells of the midgut epithelium.

BVs of group I NPVs infect other tissues through the endocytotic pathway (Volkman and Goldsmith, 1985). After BV binding to the cell membrane and uptake into an endosome (Hefferon *et al.*, 1999) the acidification of the endosome triggers GP64-mediated fusion of the viral and endosomal membrane (Blissard and Wenz, 1992; Kingsley *et al.*, 1999; Plonsky *et al.*, 1999). Then the nucleocapsids are released

into the cytoplasm and are transported to the nucleus, where viral transcription and a second round of DNA replication occur. This results in the assembly of progeny NCs that may have two fates. They may move out of the nucleus into the cytoplasm and bud through the plasma membrane (BV phenotype) or they may be enveloped *de novo* in the nucleus and later be occluded into polyhedra (ODV phenotype). So, the two NPV virion phenotypes differ in virion morphology, protein composition (Fig. 1.1), tissue specificity and the mode of viral entry into host cells. Typically, OBs are released on the death of the infected cell.

A NPV infection in the insect initially results in few symptoms and infected larvae continue feeding. As the viral infection progresses, molting of the larval instars is blocked by the production of a virus encoded enzym, UDP-glucosyltransferase, which inactivates insect ecdysteroids. The late stages of infection are characterized by an enormous production of polyhedra within the infected cells. At the end of the infection the insect becomes sluggish and ceases feeding. Two virally encoded enzymes, a cathepsin-like protease and a chitinase enhance the disintegration of the larvae, by degrading connective and epidermal tissue of the larva. Finally, the larvae die within four days to three weeks p.i. depending on the virulence of the virus for the insect and abiotic factors, such as temperature.

Gene expression and DNA replication

Baculovirus gene expression is temporally regulated in a cascaded manner. Two main classes of genes are recognized: early and late (Blissard and Rohrmann, 1990). Early genes may be further subdivided as immediate and delayed early, while late genes are distinghuished as late and very late. The gene classes occur dispersed on the baculovirus genome, and both strands of the genome are involved in coding functions. Baculovirus genes are frequently clustered into transcription units that produce overlapping RNAs both in the same and in the opposite orientation. The differential expression of baculovirus genes may be influenced by these overlapping transcripts.

Early gene transcription starts before the initiation of replication of the genome and utilizes the host RNA polymerase II complex (Friesen, 1997). The transcription of many early genes starts with the binding of the polymerase to a 5'-TATAA-3'-recognition motif and the transcript is initiated within a conserved 5'-ATCA(G/T)T(C/T)-3' promoter sequence. The transcription of the early genes can be enhanced by so-called upstream or downstream activating region (UAR or DAR) elements. One of the first genes transcribed is the major immediate early gene, *ie-1*, which is essential for the stimulation of expression of other early genes necessary for

DNA replication. The mRNA of this gene appears to be spliced in some baculoviruses resulting in several differentially expressed forms of the IE1 protein. Although the IE1 protein plays an essential role in DNA replication and gene transactivation, splicing does not appear to be a general phenomenon in baculovirus transcription regulation.

Once DNA replication starts there is a switch in transcription patterns although many of the early gene products are involved in the stimulation of late gene transcription. The replication of the viral genome starts at putative replication origins consisting of repeated sequences found at multiple locations within the genome. These sequences, termed *homologous repeat* (*hr*) sequences do not appear to be highly conserved between different baculovirus species. *Hr* regions are also enhancers of transcription (Kool *et al.*, 1995). Single copy, non-*hr* putative replication origins have also been detected. Following initiation of viral DNA replication, late gene expression is initiated and the transcription of host genes is repressed. So far, nineteen virus-encoded proteins known as Late Expression Factors (LEFs) are identified, which are necessary for very late gene expression. The transcription of late genes initiates within or near a highly conserved 5'-(A/T/G)TAAG-3' motif, which functions as a transcriptional start site. In contrast to early genes, late genes are transcribed by a viral RNA polymerase. This polymerase is α-amanitin resistent and composed of at least four virus encoded proteins, namely LEF4, LEF8, LEF9 and P47 (Guarino *et al.*, 1998). The majority of the late genes encode viral structural proteins or proteins involved in virion morphogenesis. Some of the very late genes encoding occlusion-specific proteins, such as polyhedrin, are transcribed at very high levels at the end of the infection process.

Genomics of baculoviruses

Baculoviruses have circular, double stranded, supercoiled DNA genomes that vary in size from 80 kbp to 180 kbp. The nucleotide sequences of two MNPVs, AcMNPV (Ayres *et al.*, 1994) and OpMNPV (Ahrens *et al.*, 1997) were determined at the onset of this thesis. The genome of AcMNPV is composed of 133,894 bp, potentially encoding 154 proteins (Ayres *et al.*, 1994). Twenty-eight ORFs are unique to AcMNPV, whereas the other ORFs have homologues in OpMNPV. Eight *hr* regions, implicated as transcriptional activators and as putative origins of replication, are present in the AcMNPV genome. The OpMNPV genome contains of 131,990 bp and potentially encodes 152 proteins (Ahrens *et al.*, 1997). Twenty-six genes are unique to OpMNPV. Only five *hr* sequences are present, dispersed throughout the OpMNPV genome. The genomes of these group I NPVs (Zanotto *et al.*, 1993) are basically

7

similar in size, gene content and arrangement, except for some small gene inversions and insertions (or deletions).

History on *Spodoptera exigua* and SeMNPV

The beet armyworm (Fig. 1.3), *Spodoptera exigua* (Hübner), originates in Southeast Asia but has unintentionally been introduced in other regions. It was first discovered in North America in Oregon (approx. 1876), reached Florida in 1924 and since then has invaded the southern half of the United States. In 1976 the insect was accidently introduced in the Netherlands from Florida with chrysanthemum cuttings and became rapidly a serious pest of vegetable and flower crops in greenhouses. The beet army worms causes damage to e.g. sweet pepper, tomato and cabbage lettuce but also to chrysanthemum, rose and gerbera plants. The larvae destroyes not only the seedlings and flowers but also consumes large portions of the leaves and induced stunting by feeding on buds.

Fig. 1.3 The beet army worm (*Spodoptera exigua*).
Beet armyworms vary in color but are usually a shade of olive green with many fine, wavy, light colored stripes down the back and sides. The body surface is smooth and almost hairless. Mature may be up to 4 cm. Moths lay their eggs in scale-covered on leaf surfaces. When eggs first hatch, the feed in groups near the egg mass, skeletonizing or completely consuming leaves. As they grow older, larvae disperse and move toward the center of the plant. Beet armyworms build up as weather warms and are most common on late summer and fall crops.

Since this insect shows broad insecticide resistance (Smits, 1987) alternative strategies, such as biological control using baculoviruses, have been explored. From Californian populations of beet armyworm the baculovirus SeMNPV had been isolated (Hunter and Hall, 1968; Smith and Summers, 1978; Gelernter and Federici, 1986b), which was later called SeMNPV-US1 (Muñoz *et al.*, 1998). Comparison of SeMNPV to other NPVs showed that it was five times more virulent than other SeMNPV isolates and also kills the larvae faster (Smits and Vlak, 1988b). For these reasons and the fact that SeMNPV is monospecific (only infectious for *S. exigua*) it was considered a most suitable candidate as biological control agent of the beet armyworm. Nowadays, preparations of SeMNPV, called Spod-X, are registered for use in greenhouses on ornamental plants. In 1999 Spod-X is also allowed as biological pesticide on vegetable crops.

Scope of the thesis

Along with the development of SeMNPV as biological insecticide, research was initiated on fundamental molecular aspects of this virus, such as function and regulation of genes, genome organization, mode of entry, DNA replication and virus factors that determine host range and virulence. This thesis aims at the further molecular characterization of the baculovirus SeMNPV to gain insight in its genetic make-up in comparison to those of other baculoviruses. This enables evalution of the preliminary taxonomic status of SeMNPV as a group II NPV and may reveal the assumed molecular basis of its specific biological properties.

At the onset of this thesis research, the complete nucleotide sequence of two group I NPVs, namely AcMNPV and OpMNPV, was known. However, sequence information on group II NPVs was sketchy and limited to some individually characterized genes. For SeMNPV the sequences of a number of genes, such as *polyhedrin* (van Strien *et al.*, 1992), *p10* (Zuidema *et al.*, 1993), *ubiquitin* (van Strien *et al.*, 1996), *ribonucleotide reductase* large subunit (van Strien *et al.*, 1997) and *p143* (Heldens *et al.*, 1997b), had been elucidated and characterized. The location of these characterized genes on the SeMNPV genome, based on a physical map constructed for the American-isolate (US1), differed considerably from that of AcMNPV and OpMNPV, suggesting that the genetic organization of SeMNPV is markedly different. It was noted, however, that among baculoviruses the genomic region located in the 'centre' of the linearized genome was highly conserved (Heldens *et al.*, 1997b). To be able to investigate the genetic difference of SeMNPV with group I NPVs and the molecular background of its biological properties, the complete sequence of the SeMNPV genome is elucidated (**Chapter 2**). The coding potential of SeMNPV is determined and the individual genes are compared to other baculovirus genes. Also the relative gene order is analysed as a second independent means of measuring the relatedness between baculoviruses.

The availability of the full DNA sequence of the SeMNPV genome allowed us to identify genes for which no homologs are found as yet among the *Baculoviridae*. As these genes may play key roles in biological properties specific to SeMNPV, such as host range and virulence, a selected set of genes has been studied into further detail (**Chapter 3**). These genes (Se116 and Se117) are expressed and their possible functions during the viral infection process are discussed. The temporal expression patterns of these unique genes are analysed and their possible appearance in SeMNPV budded or occlusion-derived virus is investigated. The research on unique genes is extended by the characterization of Se17/18, so far unique among NPVs, but strikingly, with a homolog in a distantly related granulovirus (**Chapter 4**). The

9

Se17/18 transcript is characterized, the subcellular location of its product determined and its importance discussed.

The comparison of AcMNPV to SeMNPV in Chapter 2 reveals also that the latter lacks a homolog of the *gp64* BV envelope fusion protein (EFP) gene. In AcMNPV and OpMNPV, this protein plays a pivotal role in BV entry, spread of the infection in the insect and is required for efficient virus budding. In **Chapter 5** the mechanism of entry into target cells is therefore examined and the identification of the SeMNPV EFP described. This investigation indicates that SeMNPV uses a similar entry mechanism as group I NPVs but other proteins.

Due to the inability of *S. exigua* cell lines to properly replicate and maintain the SeMNPV genome stably, conventional techniques do not result in viable recombinants with biological activity (Heldens *et al.*, 1996). This technical difficulty hampers the research on the function of specific SeMNPV genes. The potential of a novel strategy to generate SeMNPV recombinants is explained in **Chapter 6**. With the molecular genetic information of SeMNPV available, the possibility whether recombinants with improved insecticidal properties can be generated is investigated. Furthermore, experiments to obtain the first SeMNPV recombinant are described.

In **Chapter 7** the experimental data obtained during this PhD research are discussed in the context of recent literature data and current insights on baculovirus taxonomy and diversity. Also, a comparison is drawn between the genomes of baculoviruses and those of herpesviruses to extract common features in the genetic make-up of large DNA viruses.

Chapter 2

Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome

The nucleotide sequence of the DNA genome of *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV), a group II NPV, was determined and analysed. The genome contains 135611 bp and has a G+C content of 44 mol%. Computer-assisted analysis revealed 139 ORFs of 150 nucleotides or larger; 103 have homologues in *Autographa californica* MNPV (AcMNPV) and a further 16 have homologues in other baculoviruses. Twenty ORFs are unique to SeMNPV. Major differences in SeMNPV gene content and arrangement were found compared with the group I NPVs AcMNPV, *Bombyx mori* (Bm) NPV and *Orgyia pseudotsugata* (Op) MNPV and the group II NPV *Lymantria dispar* (Ld) MNPV. Eighty-five ORFs were conserved among all five baculoviruses and are considered as candidate core baculovirus genes. Two putative *p26* and *odv-e66* homologues were identified in SeMNPV, each of which appeared to have been acquired independently and not by gene duplication. The SeMNPV genome lacks homologues of the major budded virus glycoprotein gene *gp64*, the immediate-early transactivator *ie-2* and *bro* (baculovirus repeat ORF) genes that are found in AcMNPV, BmNPV, OpMNPV and LdMNPV. Gene parity analysis of baculovirus genomes suggests that SeMNPV and LdMNPV have a recent common ancestor and that they are more distantly related to the group I baculoviruses AcMNPV, BmNPV and OpMNPV. The orientation of the SeMNPV genome is reversed compared with the genomes of AcMNPV, BmNPV, OpMNPV and LdMNPV. However, the gene order in the 'central' part of baculovirus genomes is highly conserved and appears to be a key feature in the alignment of baculovirus genomes.

Introduction

The *Baculoviridae* are a family of rod-shaped viruses with large circular, covalently closed, double-stranded DNA genomes. The family is subdivided into two genera, *Nucleopolyhedrovirus* (NPV) and *Granulovirus* (GV), distinguished by occlusion body morphology (Volkman *et al.*, 1995). The NPVs produce large, polyhedron-shaped occlusion bodies called polyhedra that contain many virions, whereas the GVs have smaller occlusion bodies called granules that normally contain a single virion. The NPVs are designated as single (S) or multiple (M) depending on the potential number of nucleocapsids packaged in a virion.

Baculoviruses almost exclusively infect insects, belonging mainly to the orders *Lepidoptera, Hymenoptera* and *Diptera* (Adams and McClintock, 1991). SeMNPV infects only a single insect species, the beet army worm *Spodoptera exigua* (Lepidoptera: Noctuidae), a worldwide insect pest of agricultural importance. SeMNPV differs from many other baculoviruses in that it is monospecific and highly virulent for *S. exigua* larvae (Smits, 1987). However, the molecular mechanism associated with these properties is unknown. Therefore, it is important at this point to study the genetic information available for the virus and the expression of its genes.

The best-characterized baculoviruses are *Autographa californica* (Ac) MNPV (Ayres *et al.*, 1994), *Bombyx mori* (Bm) NPV (Gomi *et al.*, 1999), *Orgyia pseudotsugata* (Op) MNPV (Ahrens *et al.*, 1997) and *Lymantria dispar* (Ld) MNPV (Kuzio *et al.*, 1999). The genome of AcMNPV is composed of 133894 bp, potentially encoding 154 proteins (Ayres *et al.*, 1994). Fourteen ORFs are unique to AcMNPV, whereas most of the other ORFs have baculovirus homologues. Eight homologous regions (*hr*), implicated as transcriptional activators and as putative origins of replication, are present in the AcMNPV genome. BmNPV is 128413 bp in size and contains 136 putative genes (Gomi *et al.*, 1999). Only four ORFs are unique to BmNPV, all other ORFs possessing a baculovirus homologue. Five copies of an AcMNPV ORF2 homologue, named *bro*, and seven *hr* sequences are present, dispersed along the BmNPV genome. The OpMNPV genome contains 131990 bp and potentially encodes 152 proteins (Ahrens *et al.*, 1997). Twenty-six genes are unique to OpMNPV. OpMNPV contains one complete *bro* gene and two truncated ORFs that show homology to the *bro* genes. Only five *hr* sequences are present, dispersed throughout the OpMNPV genome. AcMNPV, BmNPV and OpMNPV belong to the group I NPVs (Zanotto *et al.*, 1993). The genome of LdMNPV (group II) is composed of 161046 bp and contains 163 ORFs (Kuzio *et al.*, 1999). Forty-seven genes are unique to LdMNPV. The large size of the LdMNPV genome is largely due to the presence of 13 *hr* sequences and 16 *bro* gene homologues.

A detailed physical map was recently constructed for an American isolate of SeMNPV (US1) in order to map a mutant SeMNPV. This mutant was obtained within the first passage in insect cell culture. It has a single deletion of approximately 25 kb and is unable to infect *S. exigua* larvae (Heldens *et al.*, 1996). The sequences of a number of SeMNPV genes, including *polyhedrin* (van Strien *et al.*, 1992), *p10* (Zuidema *et al.*, 1993), *ubiquitin* (van Strien *et al.*, 1996), *ribonucleotide reductase* large subunit (van Strien *et al.*, 1997) and *p143* (Heldens *et al.*, 1997b), have been elucidated and characterized. The locations of these genes on the SeMNPV genome differ considerably from those of other baculoviruses, such as AcMNPV, BmNPV, OpMNPV and LdMNPV, suggesting that the genetic organization is markedly different. It was noted, however, that among baculoviruses the genomic region located in the 'centre' of the linearized genome was highly conserved (Heldens *et al.*, 1997b). Six *hr* sequences have been identified on the SeMNPV genome, which are similar in structure to those of other baculoviruses (Broer *et al.*, 1998). Here, we present the complete sequence and gene organization of the SeMNPV genome and compare them to other baculoviruses by genomic and phylogeny analysis.

Results and discussion

Nucleotide sequence analysis of the SeMNPV genome

The SeMNPV genome was assembled into a contiguous sequence of 135611 bp, in good agreement with a previous estimate of 134.1 kb based on restriction enzyme analysis and physical mapping (Heldens *et al.*, 1996). The adenine residue at the translational initiation codon of the *polyhedrin* gene was designated previously as the zero point of the physical map of SeMNPV (Heldens *et al.*, 1996). The orientation of the physical map was, by convention, set by the location of the *p10* gene (Vlak and Smith, 1982).

One hundred and thirty-nine ORFs, defined as methionine-initiated ORFs encoding more than 50 amino acids and with minimal overlap with other ORFs, were present on the SeMNPV genome (Fig. 2.1). The SeMNPV ORFs were, in general, tightly packed with minimal intergenic distances and their orientation was distributed almost evenly along the genome (55% clockwise, 45% anticlockwise; Fig. 2.1). The locations, orientations and sizes of the predicted ORFs are shown in detail in Table 2.1. The distribution of the ATG, TAG and TGA codons in the SeMNPV sequence was not random, while the TAA frequency (1.58%) was not significantly different from the expected random distribution (1.56%). The ATG codon (1.77%) and TGA
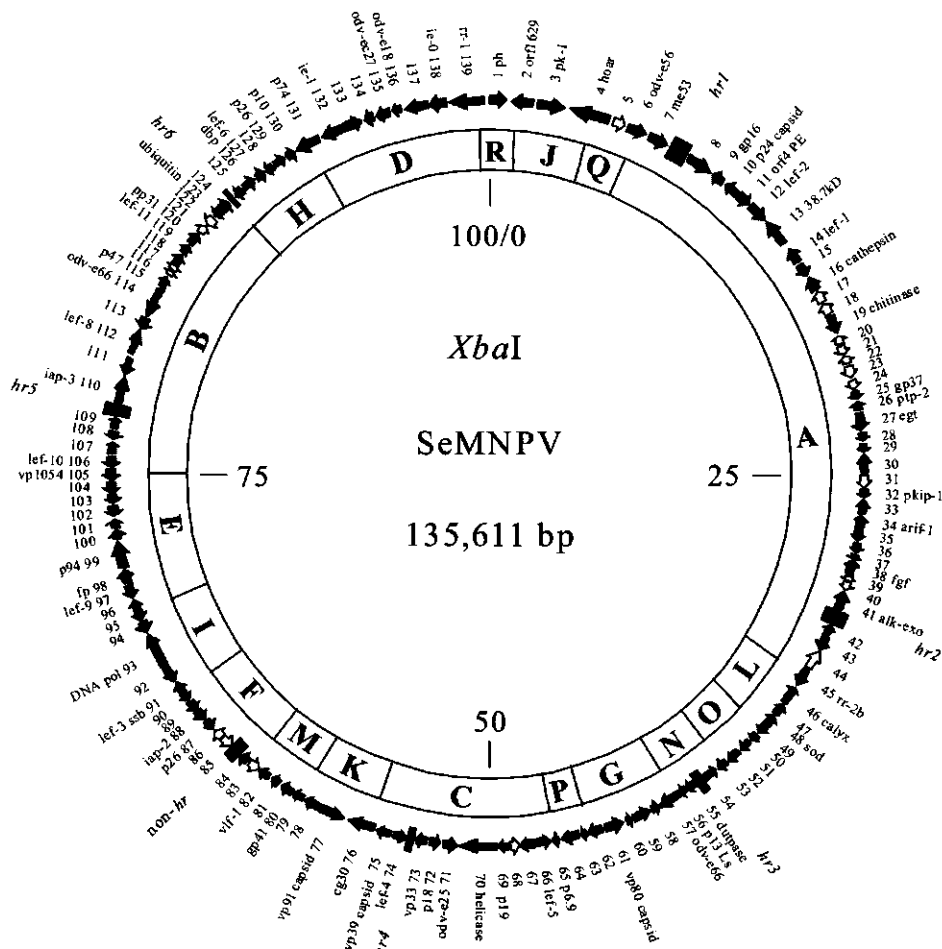
Fig. 2.1 Circular map of the SeMNPV DNA genome showing genomic organization. Sites for restriction enzyme *Xba*I are shown; fragments are indicated A to R according to size from the largest to the smallest (Heldens *et al.*, 1996). The positions of the 139 ORFs identified are indicated by arrows that also represent the direction of transcription. Shaded arrows indicate ORFs with homologues in other baculoviruses in the protein sequence databases. Open arrows represent ORFs unique to SeMNPV. The numbers alongside the ORFs represent the SeMNPV ORF number (see Table 2.1). The positions of the *hr* sequences are indicated by black boxes. The scale on the inner circle is in map units.

stop codon (1.78%) occurred more frequently in the SeMNPV sequence, while there was paucity of TAG stop codons (0.83%), as is the case for AcMNPV (Ranjan and Hasnain, 1995). Predicted ORFs represented 90% coding density, with a mean ORF length of 875 nucleotides. Twenty ORFs had small (<25 aa) overlaps with adjacent ORFs. One hundred and nineteen (86%) of the 139 SeMNPV ORFs had an assigned

function or had homologues among other baculovirus genes (Table 2.1). Twenty ORFs are so far unique to SeMNPV. These ORFs accounted for 7% (9.3 kb) of the genome. Six *hr* sequences similar in structure to those of other baculovirus *hr* sequences have been identified previously (Broer *et al.*, 1998) and no further *hr* sequences were detected in the complete sequence. Furthermore, one non-*hr*, putative origin of replication, is present on the *Xba*I-F fragment (Heldens *et al.*, 1997a). The positions of the *hr* sequences are presented in Fig. 2.1.

Table 2.1 Potentially expressed ORFs in SeMNPV strain US1.

The positions and orientations of 139 potentially expressed ORFs in the SeMNPV genome are shown. The presence of baculovirus early (E, E2) and late (L) promoter elements, located within 120 nt of the ATG, is indicated. E and E2 indicate a TATA sequence with CAKT (E) or CGTGC (E2) start site sequence 20-40 nt downstream. L indicates the presence of a DTAAG motif. Transcription elements were identified by the FINDPATTERNS program of GCG. Homologues ORFs in the genomes of AcMNPV (Ayres *et al.*, 1994), BmNPV (Gomi *et al.*, 1999), OpMNPV (Ahrens *et al.*, 1997) and LdMNPV (Kuzio *et al.*, 1999) are shown with the percentage amino acid sequence identity to the homologous ORF. The different clusters identified by GeneParityPlot analysis (Fig. 2.2) are shown.

ORF	Name	Position	Length (aa)	Pred. M$_r$	Prom.	Homologous ORFs Ac	Bm	Op	Ld	Identity to homologues (%) Ac	Bm	Op	Ld	Other	Cluster
1	*polyhedrin*	1 > 741	246	28950	L	8	1	3	1	85	82	84	82		1
2	*orf1629*	845 < 2233	462	50358	L	9	2	2	2	34	26	26	27		1
3	*pk-1*	2232 > 3119	295	34819		10	3	1	3	40	40	39	48		1
4	*hoar* *	3182 < 5356	712	82301	E2									27 *	
5		6190 > 7713	507	58571	E										
6	*odv-e56*	7848 > 8963	371	40970	L	148	124	146	14	47	45	47	51		
7	*me53*	9261 > 10433	390	45399		139	116	137	23	30	29	27	36		
	hr1	10515 11645													
8		12498 > 14495	656	76323	E, L	23	14	21	130	22	22	23	41		
9	*gp16*	14632 < 14916	94	10892	L	130	107	128		31	31	26			
10	*p24capsid*	14953 < 15699	248	27379	L	129	106	127		44	44	41			
11	*orf4 PE*	15784 > 16101	105	11936	L				137a				28		15
12	*lef-2*	16064 > 16693	209	24364		6	135	6	137	44	47	39	41		15
13	*38.7kd*	16738 < 17829	363	42338		13	5	12	122	34	33	28	33		2
14	*lef-1*	17829 < 18479	216	25369		14	6	13	123	44	43	45	51		2
15		18520 > 18984	154	17572	L				124				47		[2]
16	*cathepsin*	18990 < 20003	337	38387	L	127	104	125	78	55	56	55	56		3
17		20121 < 20456	111	12963											
18		20471 < 20989	172	20186											
19	*chitinase*	21122 > 22840	572	63509	L	126	103	124	70	65	65	65	63		3
20		22907 < 23344	145	17723											
21		23524 > 23820	98	11408											
22		24059 > 24370	103	12019											
23		24594 > 25229	211	24059	L										
24		25229 > 25447	72	8243											

ORF	Name	Position	Length (aa)	Pred. M_r	Prom.	Homologous ORFs Ac	Bm	Op	Ld	Identity to homologues (%) Ac	Bm	Op	Ld	Other	Cluster
25	gp37	25512 > 26315	267	30393	L	64	52	69	68	56	55	57	58		
26	ptp-2	26330 < 26827	165	19174	L		9						30		
27	egt	26928 > 28499	523	59751	E	15	7	14	125	50	49	50	54		
28		28720 > 29292	190	21818	E				127				26		16
29		29332 > 29973	213	24432	L	17	9	16	128	25	27	28	32		16
30		30009 < 32669	886	103481	E				129				29		16
31		32708 > 33433	241	28960											
32	pkip-1	33513 > 34007	164	19650	L	24	15	44	110	26	28	25	30		17
33		34039 < 34377	112	12762	E				111				23		17
34	arif-1	34389 < 35234	281	31454		21	12	19	118	25	27	22	32		4
35		35146 > 36387	413	47379		22	13	22	119	57	57	56	62		4
36		36415 > 37995	526	60472		119	97	119	155	45	45	46	43		5
37		37992 > 38234	80	9480		120	98	120	155a	28	28	20	30		5
38	fgf	38262 < 39476	404	45040		32	24	27	156	24	24	28	26		[5]
39		39448 > 39630	60	6720	E2										
40		39806 > 40531	241	27986	E										
41	alk-exo	40540 < 41781	413	48138	L	133	110	131	157	42	41	41	44		[5]
hr2		**41812**				**42353**									
42		42361 < 42606	81	9445	E	19	11	18	159	27	26	28	30		6
43		42696 > 43856	280	32472	E, L	18	10	17	158	25	25	23	24		6
44		43890 < 44312	140	16306											
45	rr2b	44408 > 45349	313	36608	E, L			34	120			20	55		
46	calyx / pep	45417 < 46424	335	37901	L	131	108	129	136	35	38	37	47		
47		46525 < 46836	103	11655		117	96	117		32	34	24			
48	sod	46971 < 47426	151	16105	L	31	23	29	145	67	67	66	66		18
49		47514 > 47906	130	15032	E				144				23		18
50		47932 > 48576	214	23587		115	95	115	143	45	46	42	51		18
51		48581 > 49009	142	16700					142				17		18
52		49019 > 50608	529	61477					141				21		18
53		50667 > 51335	222	25585	L	106/107	90	107	140	58/32	52	50	55		18
54		51404 < 52498	364	42915	E	33			32	32					
hr3		**52710**				**53237**									
55	dutpase	53308 > 53739	143	15856			31		116			27	50		
56	p13 ‡	53935 > 54786	283	33398	L									72/40 ‡	
57	odv-e66	54855 > 57026	723	83842	E, L	46	37	50	131	29	30	30	44		
58		57023 < 57367	114	12644	L	108	91	108	108	40	40	34	45		7
59		57376 < 58446	356	41278		109	92	109	107	49	49	48	49		7
60		58430 < 58609	59	7039		110			106	33			43		
61	vp80capsid	58606 < 60276	556	63804	E	104	88	105	105	22	21	19	21		8
62		60336 > 61463	375	44162	L	103	87	104	104	54	52	48	55		8
63		61453 > 61773	106	11780	L	102	86	103	103	25	27	26	40		8
64		61806 > 62972	388	43655	E, L	101	85	102	102	44	43	39	47		8
65	p6.9	63037 > 63264	75	9184	L	100	84	101	101	65	55	62	66		8
66	lef-5	63261 < 64100	279	32847		99	83	100	100	55	54	47	51		8
67		63996 > 64898	300	35809	L	98	82	99	99	48	47	43	51		8

ORF	Name	Position	Length (aa)	Pred. M_r	Prom.	Homologous ORFs				Identity to homologues (%)					Cluster
						Ac	Bm	Op	Ld	Ac	Bm	Op	Ld	Other	
68		64928 > 65413	161	18509	L										
69		65502 < 66014	170	19532	L	96	79	97	98	51	51	47	63		8
70	helicase	65980 > 69648	1222	143327	L	95	78	96	97	44	44	39	50		8
71	odv-e25	69732 < 70382	216	24096	L	94	77	95	96	50	48	45	69		8
72		70379 < 70852	157	18104		93	76	94	95	46	46	44	66		8
73		70864 > 71622	252	30597	E2	92	75	93	94	55	55	54	55		8
	hr4	71712 71779													
74	lef-4	71844 < 73244	466	54354		90	73	91	93	51	51	44	51		8
75	vp39capsid	73243 < 74223	326	36963	E, L, L	89	72	90	92	44	42	46	54		8
76	cg30	74454 > 75839	461	50821	E	88	71	89		24	22	22			8
77	vp91capsid	75935 < 78376	813	93285	L	83	69	86	91	43	42	41	44		8
78		78345 < 78935	196	21741	L	82	68	85	90	33	32	29	49		8
79		78742 > 79464	240	27736	L	81	67	84	89	50	50	48	51		8
80	gp41	79448 > 80443	331	37131	L	80	66	83	88	54	54	46	56		8
81		80465 > 80848	127	14521	E, L	78	64	81	87	45	42	41	46		8
82	vlf-1	80850 > 81968	372	43934	E, L	77	63	80	86	65	65	63	65		8
83		82262 < 82621	119	14134											
84		82519 > 83154	211	23959		44	35	49		25	24	24			
	non-hr														
85		83336 < 83593	85	9854	L										
86		83527 > 83703	58	7016											
87	p26	83928 < 84680	248	28268	E	136	113	132	40	24	24	24	26		
88	iap-2	84751 < 85704	317	36821	L	71	58	74	79	33	34	37	44		9
89		85499 < 86398	299	34210		69	57			44	44				9
90		86388 < 86789	133	15335		68	56	73	80	41	41	41	43		9
91	lef-3	86788 > 88056	422	48747	L	67	55	72	81	30	28	30	33		9
92		88110 < 90224	704	80851		66	54	71	82	23	23	24	27		9
93	DNA pol	90226 > 93417	1063	123109	E	65	53	70	83	47	47	44	57		9
94		93456 < 93845	129	15130	L	75	61	78	84	26	27	28	50		10
95		93856 < 94113	85	9972	E, L	76	62	79	85	45	44	42	71		10
96		94236 > 94577	113	12673	L	150	126		30	34	36		26		
97	lef-9	94614 > 96101	495	56629		62	50	65	64	63	63	62	69		11
98	fp[†]	96189 > 96776	195	22694	E2, L	61	49	64	63	59	59	57	28/38[†]		11
99	p94	96968 > 99127	719	85640	L	134				25					
100		99254 > 99523	89	10325	E, L	60	48	63	62	45	45	38	52		11
101		99537 > 100124	195	23027	L	59		62	61	44		35	45		11
102		100117 < 100653	178	20603		57	46	61	60	45	47	41	49		11
103		100813 < 101094	93	10747		56	45	60		27	28	29			11
104		101036 < 101239	67	7884	L	55	44	59	58	42	40	37	42		11
105	vp1054	101370 < 102410	346	40457	L	54	43	58	57	45	45	39	55		11
106	lef-10	102265 < 102498	77	8262	L	53a	42a	57	56	49	49	35	49		11
107		102694 > 103728	344	38949	L				55				24		11
108		103820 < 104233	137	16359	L	53	42	56	54	49	51	51	60		11
109		104296 < 104784	162	19632		52	41		53	24	24		34		11
	hr5	104879 105679													

ORF	Name	Position	Length (aa)	Pred. M_r	Prom.	Homologous ORFs Ac	Bm	Op	Ld	Identity to homologues (%) Ac	Bm	Op	Ld	Other	Cluster
110	*iap-3*	105825 > 106766	313	36082	L		35		139			48	26		
111		106826 < 108073	415	47226		51	40	55		24	23	23			11
112	*lef-8*	108094 > 110814	906	105197		50	39	54	51	64	64	59	65		11
113		110859 < 111032	57	7228	L	43	34	48		34	33	33			
114	*odv-e66*	111066 < 113123	685	78650	E2, L	46	37	50	131	29	29	30	32		
115	*p47*	113169 > 114371	400	46935		40	31	45	48	56	55	53	66		
116		114465 > 115142	223	26020	E										
117		115249 > 115824	191	22537	E										
118		115874 > 116659	261	31360	E, L	38	29	22	46	58	57	54	59		12
119	*lef-11*	116650 > 116964	103	12212	L	37	28	23	45	36	35	34	42		12
120	*39K / pp31*	116927 > 117880	317	35754		36	27	24	44	34	38	35	40		12
121		117908 > 118198	96	11525											
122		118299 > 118508	69	7693											
123	*ubiquitin*	118496 < 118738	80	9364	L	35	26	25	43	73	73	77	74		12
124		118828 > 119391	187	21960	L	34	25	26	42	35	36	41	46		12
	hr6	119442 120222													
125		120249 < 120656	135	14996	L	26	17	42	36	34	34	31	33		
126	*dbp*	120802 > 121788	328	37991		25	16	43	47	31	31	32	44		
127	*lef-6*	121816 > 122307	163	19098	L	28	19	40		28	28	25	33		
128		122347 < 122757	136	16347		29	20	39	38	32	36	35			
129	*p26*	122862 > 123698	278	31121	L	136	113	132	40	32	30	33	39		13
130	*p10*	123740 > 124006	88	9607	E, L	137	114	133	41	33	32	32	53		13
131	*p74*	124099 < 126060	653	74214	E, L	138	115	134	27	56	56	57	59		
132	*ie-1*	126197 < 128341	714	82059	E	147	123	145	15	29	28	29	33		14
133		128374 > 128976	200	22358	E2, L	146	122	144	16	34	34	35	38		14
134		129067 < 129345	92	10576	L	145	121	142	17	46	45	46	53		14
135	*odv-ec27*	129360 < 130205	281	32610	L	144	120	141	18	58	58	51	58		14
136	*odv-e18*	130260 < 130502	80	8546		143	119	140	19	56	50	46	58		14
137		130540 < 131922	460	54122	L	142	118	139	20	52	52	50	52		14
138	*ie-0*	131937 < 132671	244	28719	L	141	117	138	21	31	30	31	36		14
139	*rr-1*	132794 < 135106	869	99542	E, L		32		148			26	28		

* HzSNPV ORF name taken from Le *et al.* (1997). Percentage amino acid identity to HzSNPV *hoar* ORF is shown.

‡ LsNPV ORF name taken from Wang *et al.* (1995). Percentage amino acid identity is shown to LsNPV *p13* and the translated sequence form the *p13* to the *xe* ORF (Wang *et al.*, 1995; GenBank U30303 and AB000383).

† In the sequenced LdMNPV strain, there is a frame-shift in the *fp* gene. The two resulting ORFs were indicated by the same number (ORF63) by Kuzio *et al.* (1999). However, the amino acid identity of the SeMNPV *fp* gene to the resulting LdMNPV homologues is not the same. The values refer to the smaller and larger LdMNPV homologues, respectively.

Comparison of the SeMNPV gene content with that of other baculoviruses

The SeMNPV genome was compared with those of AcMNPV, BmNPV, OpMNPV and LdMNPV. These five genomes have a cumulative total of 245 different ORFs, of which 86 are unique to individual baculovirus genomes and 85 are conserved among the above baculoviruses. Putative functions have been assigned to approximately half of the common baculovirus genes, suggesting that these genes are required for basic baculovirus features, such as virus structure, transcription, DNA replication and auxiliary functions on the cellular or organism level (Table 2.1).

The overall characteristics of the different baculovirus genomes are shown in Table 2.2. The G+C content of the SeMNPV genome was 43.8 mol%, which is similar to that of AcMNPV (Ayres *et al.*, 1994) and BmNPV (Gomi *et al.*, 1999) but much lower than that of OpMNPV (Ahrens *et al.*, 1997) and LdMNPV (Kuzio *et al.*, 1999) (Table 2.2). The smaller number of SeMNPV ORFs compared with AcMNPV and OpMNPV, the genomes of which are similar in size, is caused by the absence of a number of small putative ORFs. In AcMNPV and OpMNPV, these small putative ORFs are located between larger ORFs. The larger ORFs have homologues in many other baculoviruses. SeMNPV does not possess these smaller ORFs, but does possess the larger ORFs. This suggests that the numbers of ORFs in AcMNPV and OpMNPV may be overestimates and that these smaller putative ORFs may not be functional. The frequency of different temporal consensus promoter elements showed considerable variation between baculoviruses, except for the late promoter motif (Table 2.2).

Table 2.2 Characteristics of different baculovirus genomes.
Data for baculoviruses other than SeMNPV taken from Ayres *et al.* (1994) (AcMNPV), Gomi *et al.* (1999) (BmNPV), Ahrens *et al.* (1997) (OpMNPV) and Kuzio *et al.* (1999) (LdMNPV).

Characteristic	AcMNPV	BmNPV	OpMNPV	LdMNPV	SeMNPV
Size (kbp)	133.9	128.4	132.0	161.0	**135.6**
G+C content (mol%)	41	40	55	58	**44**
ORFs (total)	154	136	152	163	**139**
Unique ORFs	14	4	26	47	**20**
Number of *hr* sequences	8	7	5	13	**6**
Promoters:					
Early	65	12	61	12	**34**
Late	72	78	64	79	**72**
Early + Late	29	7	26	6	**14**
Not identified	47	35	58	78	**53**

Of the 139 SeMNPV ORFs identified, 103 have homologues in AcMNPV (Table 2.3) and a further 16 have homologues in other baculoviruses (Table 2.1). SeMNPV shares the largest number of homologues with AcMNPV and LdMNPV. The mean amino acid identity between SeMNPV and AcMNPV, BmNPV, OpMNPV and LdMNPV homologues is 41, 41, 40 and 45%, respectively (Table 2.1). The most conserved ORF in all five baculoviruses is that of *polyhedrin* (SeMNPV compared with AcMNPV, BmNPV, OpMNPV and LdMNPV: 85, 82, 84 and 82% amino acid identity, respectively), closely followed by *ubiquitin* (73, 73, 77 and 74% identity).

Table 2.3 Number of ORFs with homologues in other baculoviruses.

	AcMNPV	BmNPV	OpMNPV	LdMNPV
SeMNPV	**103**	**99**	**102**	**104**
LdMNPV	94	91	95	
OpMNPV	126	121		
BmNPV	115			

The SeMNPV ORF 99 (Se99) homologue in AcMNPV (*p94*) is not essential for virus replication in cell culture but may be involved in the prevention of apoptosis (Clem *et al.*, 1994; Friesen and Miller, 1987). Se99 was probably acquired by an independent insertion from a different source, because Se99 is located in a different region in the SeMNPV genome compared with AcMNPV *p94* (see Fig. 2.2A; Se99 is inserted in cluster 11), or could be the result of a single gene rearrangement.

Fifty-three AcMNPV genes had no homologues in SeMNPV (Table 2.4). Most of these genes are also absent in at least one of the other three baculoviruses compared. However, SeMNPV also lacks AcMNPV ORFs 4, 11, 38, 111 and 115, which are present in the four other baculoviruses. To date, no functions have been assigned to these ORFs. In addition to the AcMNPV ORFs that have no SeMNPV homologues, a number of BmNPV (1), OpMNPV (17) and LdMNPV (33) ORFs without AcMNPV homologues are also absent in SeMNPV (Table 2.4). The total number of AcMNPV, BmNPV, OpMNPV and LdMNPV ORFs without homologues in SeMNPV is 53, 7, 40 and 78, respectively. Although most of the baculovirus ORFs not found in SeMNPV have not yet been characterized, information is available for some (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Kuzio *et al.*, 1999). The presence or absence of baculovirus ORF homologues in SeMNPV and their implications for SeMNPV characteristics are discussed below.

Table 2.4 Baculovirus ORFs without homologues in SeMNPV.

The AcMNPV ORFs that have no homologue in SeMNPV are shown. ORFs from BmNPV, OpMNPV and LdMNPV that have no homologue in either AcMNPV or SeMNPV are also shown.

AcMNPV[*]	1 *ptp-1*[c]	20[bc]	48 *etm*[ac]	85[c]	115	132[c]
	2 *bro*	27 *iap-1*[c]	49 *pcna*[ac]	86 *pnk/pnl*[abc]	116[bc]	135 *p35*[bc]
	3 *conotoxin*[a]	30[c]	58[b]	87[c]	118[bc]	140[bc]
	4	38	63[bc]	91[c]	121[bc]	149[bc]
	5[c]	39 *p43*[bc]	70 *hcf-1*[abc]	97[bc]	122[c]	151 *ie-2*[c]
	7 *orf603*[abc]	41[c]	72[c]	105 *he65*[bc]	123 *pk-2*[bc]	152[bc]
	11	42 *gta*[c]	73[c]	111	124[c]	153 *pe38*[c]
	12[b]	45[bc]	74[c]	112[b]	125 *lef-7*[c]	154[bc]
	16[c]	47[c]	84[bc]	113[b]	128 *gp64*[c]	
BmNPV	111					
OpMNPV	4	33	68	110 Ld30[†]	143 *hrf-1* Ld67[†]	135
	5	36	98	113	147 *opep32*	149 *p8.9*
	28	37	106 *iap4* Ld140[†]	118	148 *opep25*	
LdMNPV	4 *mucin-like*	10	25	50 *helicase-2*	121	152
	5	11	26	52	126	160 *vef-2*
	6	12	28	59	132	163
	7 *g22*	13	31	65 *vef-1*	133	
	8	22 *dna-ligase*	34	69	134	
	9	24	49	77	135	

[*] ORFs also absent from BmNPV (a), OpMNPV (b) or LdMNPV (c).

[†] LdMNPV homologues are indicated.

Structural genes: two copies of *odv-e66*, absence of *gp64*

An interesting characteristic of SeMNPV is the presence of two *odv-e66* homologues. One copy, Se114, is flanked by homologues of Ac40 and Ac43. The other copy, Se57, is located between Se58 and Se56, which are homologues of Ac108 and the *Leucania separata* (Ls) NPV *p13* gene, respectively. The latter *odv-e66* homologue is located close to SeMNPV *hr3*. The identity between the two SeMNPV *odv-e66* genes, Se57 and Se114, is only 32%, which is not significantly higher than that to their AcMNPV, BmNPV or OpMNPV homologues. The identity between Se57 and its LdMNPV homologue (Ld131) is higher (44%) and the most conserved homologue present in GenBank is the LsNPV homologue of Se57 (56%). The Se57 region has probably

undergone extensive rearrangement, because homologues of different baculoviruses are present at this location, whereas Se114 is surrounded only by AcMNPV homologues (see Fig. 2.2 for positions of Se57 and Se114 in the gene parity analysis). Therefore, it is likely that Se57 was acquired independently from a second source that was more closely related to LdMNPV and LsNPV than to the Se114 copy of the *odv-e66* gene. The presence of two copies of *odv-e66* with both late and early baculovirus consensus promoters may be related to the two forms of *odv-e66* found in mature AcMNPV (Hong *et al.*, 1994).

Surprisingly, SeMNPV lacks a homologue of the budded virus (BV) surface glycoprotein *gp64* (Ac128). A similar situation exists for LdMNPV (Kuzio *et al.*, 1999). GP64 is a major envelope glycoprotein that is acquired by virions during budding through the plasma membrane. GP64 is required for efficient virion budding in AcMNPV; deletion of the cytoplasmic tail domain resulted in a reduction in progeny BV and in a virus that was incapable of efficient propagation in cell culture (Oomens and Blissard, 1999). It has been suggested that Ld130, which shares 22% identity with Ac23, could substitute for the lack of GP64 (Kuzio *et al.*, 1999). This supposition was based on the presence of N-terminal signal and transmembrane domains, which are indicative of transmembrane receptor-like proteins. SeMNPV has a homologue of Ld130, Se8, which has twice the identity to LdMNPV (41%) than to Ac23, Bm14 and Op21 (~22%).

DNA replication genes

The genes essential for DNA replication were only moderately conserved: *DNA pol*, *helicase*, *lef-2* and *lef-1* were 44% identical, whereas *lef-3* showed approximately 30% identity. Se126, a homologue of Ac25 and Bm16, which encodes a putative DNA-binding protein (DBP) (Okano *et al.*, 1999; Mikhailov *et al.*, 1998), showed approximately 30% identity to its homologues. The relatively low identity of these SeMNPV proteins to their homologues may explain the specificity of the virus DNA replication process (Heldens *et al.*, 1997b).

None of the *ie-2*, *pe38*, *lef-7* or *p35* genes, found to be stimulatory in AcMNPV and BmNPV DNA replication assays, had a homologue in the SeMNPV genome or the LdMNPV genome (Kuzio *et al.*, 1999). The *ie-2*, *lef-7* and *p35* genes are also non-essential for BmNPV virus replication, since functional deletion by insertion experiments resulted in viable virus mutants (Gomi *et al.*, 1997). A reduction of viral DNA synthesis, however, was demonstrated in only two of the three cell lines infected with AcΔ*lef-7* (Chen and Thiem, 1997).

Genes regulating gene expression

The genes required for transactivation of early gene transcription, such as *ie-1*, *ie-0* and *me53*, were poorly conserved in their amino acid sequence (~35%) among baculoviruses, whereas the late transcription activators including the RNA polymerase, *lef-4* (identity to AcMNPV homologue 51%), *lef-8* (64%), *lef-9* (63%), *p47* (56%) and *vlf-1* (65%) were, in contrast, very well conserved. This is compatible with the supposition that specificity is already displayed early in infection because baculoviruses have to adapt to the host transcription system. This could also be explained as the result of a higher constraint in the late and virus-encoded transcription system.

SeMNPV had a CG30 (Ac88) homologue, Se76, that is absent from the LdMNPV genome. This ORF contains features characteristic of a transcription-regulatory protein: (i) two nucleic acid-binding sites, (ii) a zinc finger and (iii) a leucine zipper. It is considered to be a prime candidate in the regulation of genes at late times in infection. The SeMNPV homologue was extended compared with the AcMNPV, BmNPV and OpMNPV genes, as the sequence ACC(G/A)TCGACATCG-GC(C/T)GG was repeated seven times. The zinc finger and the leucine zipper were present in Se76, although one of the four leucines was changed to a methionine, as was also found for the OpMNPV homologue.

Inhibitors of apoptosis

Baculoviruses have genes involved in the inhibition of apoptosis: *p35*-like genes and *iap*-like genes. SeMNPV lacked homologues of *iap-1*, *iap-4* and *p35*. The AcMNPV annihilator mutant (Ac*p35Δ*) causes cell line-specific apoptosis after infection. This is in contrast to AcMNPV wild-type and *iap-1* or *iap-2* deletion recombinants. This suggests that *iap-1* and *iap-2* are not required for prevention of apoptosis in these cell lines (Griffiths *et al.*, 1999). OpMNPV and *Cydia pomonella* GV *iap-3* have proven to be inhibitors of apoptosis in different cell lines upon infection with Ac*p35Δ* recombinant virus (Vucic *et al.*, 1998; Seshagiri and Miller, 1997; Ahrens and Rohrmann, 1995; Lu and Miller, 1995; Clem *et al.*, 1994; Clem and Miller, 1994; Birnbaum *et al.*, 1994; Crook *et al.*, 1993). SeMNPV possessed *iap-2* and *iap-3* homologues. The *iap-3* gene product shared 48% similarity with its OpMNPV homologue (Table 2.1) and may be involved in the prevention of apoptosis in *S. exigua* larvae and different *S. exigua* cell lines.

Nucleotide metabolism

SeMNPV possesses a number of previously described baculovirus genes involved in

nucleotide metabolism. Genes encoding the large and small subunits of ribonucleotide reductase (*rr1* and *rr2*) and a dUTPase were present in SeMNPV, as well as in OpMNPV and LdMNPV. By means of these proteins, SeMNPV may promote deoxyribonucleotide synthesis in non-dividing cells and conversion of dUTP to dUMP, which serves as a precursor for dTTP (Elledge *et al.*, 1992).

The *rr1* gene of SeMNPV, Se139, has been described previously (van Strien *et al.*, 1997). The SeMNPV *rr2* gene, Se45, which was found distal from *rr1*, was more closely related to Ld120 (*rr2b*) than to the Op34 (*rr2*) or Ld147 (*rr2a*) homologues in terms of protein identity and location on the genome (Fig. 2.2C, D). This is in contrast to the *rr1* gene, which is equally related to its OpMNPV and LdMNPV homologues. In contrast to OpMNPV (ORF32 and 34) and LdMNPV (ORF 147 and 148), the SeMNPV *rr1* and *rr2* as well as the LdMNPV *rr2b* genes appear to have been acquired from a source more closely related to eukaryote than prokaryote homologues (van Strien *et al.*, 1997; Kuzio *et al.*, 1999).

The *dutpase* gene of SeMNPV, Se55, differed more from the *dutpase* of OpMNPV (Op31) than from the LdMNPV (Ld116) homologue. The locations of the *dutpase* gene on the genomes of LdMNPV and SeMNPV are quite similar, whereas the OpMNPV homologue is located in a different region on the genome (Fig. 2.2C, D). This suggests that the SeMNPV and LdMNPV *dutpase* genes were acquired from the same source, whereas the OpMNPV *dutpase* may have been acquired independently from a different source. The different location of OpMNPV *dutpase* could also have resulted from gene duplication and rearrangement.

Genes with auxiliary functions
The auxiliary genes (O'Reilly, 1997) *superoxide dismutase* (*sod*) (identity to AcMNPV homologue 67%), *chitinase* (65%), *cathepsin* (55%) and *ecdysteroid UDP-glucosyltransferase* (50%) were quite well conserved, whereas the *fibroblast growth factor* (*fgf*) gene (identity to AcMNPV homologue 24%) was quite different from the other baculovirus *fgf* genes.

Ac1 encodes a protein tyrosine/serine phosphatase with dual specificity (dsPTP) (Tilakaratne *et al.*, 1991; Kim and Weaver, 1993). This protein removes phosphates specifically from both tyrosine and serine/threonine residues and regulates the phosphorylation status of a variety of proteins, including growth factors, which in turn regulate developmental processes in living cells (Wishart *et al.*, 1995). The absence of a *ptp-1* homologue in SeMNPV may not necessarily result in loss of PTPase function, however, because a *ptp-2* (Op9) homologue was present in the SeMNPV genome. The SeMNPV PTP-2 homologue, Se26, contained the conserved domain

[HCXXGXXR(S/T)] encoding the dsPTP catalytic loop. It is therefore likely that the PTP-2 homologue encodes an active tyrosine/serine phosphatase.

A protein kinase enzyme activity also appeared to have been retained in the SeMNPV genome, since a *pk-1* homologue of AcMNPV was present (Se3). However, a *pk-2* homologue was absent. *pk-2* was shown to be non-essential for AcMNPV, since a *pk-2* deletion mutant had no detectable effect on AcMNPV replication in cell cultures (Chen and Thiem, 1997). Although AcMNPV *pk-2* is non-essential, its presence favours virus gene expression by inhibiting a host stress response in infected cells (Dever *et al.*, 1998).

A homologue of the actin rearrangement-inducing factor-1 (*arif-1*) was present in the SeMNPV genome. This *arif-1* gene induces rearrangements of the actin cytoskeleton after infection, but the functional significance of these conformational changes remains to be elucidated (Roncarati and Knebel-Mörsdorf, 1997).

No homologues of the LdMNPV *viral enhancing factors* (*vef*)-1 and -2 have been identified in SeMNPV (Bischoff and Slavicek, 1997; Hashimoto *et al.*, 1991). In GVs, the *vef* gene products increase virus potency by disrupting the peritrophic membrane, thereby allowing virions access to the surface of midgut epithelial cells (Wang and Granados, 1998; Derksen and Granados, 1988). The *vef* genes encode metalloproteases that specifically degrade the mucin protein component of the peritrophic membrane (Wang and Granados, 1997; Lepore *et al.*, 1996). *vef* homologues are absent in group I baculoviruses and may, therefore, be unique to LdMNPV among the NPVs.

SeMNPV homologue ORFs of unknown function: two *p26* homologues

ORFs without assigned functions, but well conserved among the four baculoviruses, include Ac106/107 (identity to AcMNPV homologue 58%), Ac38 (58%), Ac22 (57%), Ac92 (55%) and Ac103 (54%). The high percentage identities between the baculovirus homologues suggest that these ORFs have essential functions in virus multiplication and pathology, for which a certain degree of conservation is required.

Notable is the 'pairwise' conservation of Ac76 between AcMNPV and OpMNPV (81%) and between LdMNPV and SeMNPV (71%), which may suggest that Ac76 homologues have been acquired twice during evolution from two different sources. Other pairwise alignments yielded identities no higher than 45%. To this end, it can be speculated that SeMNPV and LdMNPV have a more recent baculovirus ancestor in common than SeMNPV and AcMNPV or OpMNPV.

Some ORFs that were previously unique to LdMNPV have homologues in the SeMNPV genome (SeMNPV ORFs 15, 28, 30, 33, 49, 51, 52 and 107) (Table 2.1).

The SeMNPV genome also contained a homologue of the previously described LdMNPV ORF4 (Bjornson and Rohrmann, 1992), although this ORF was not included in the LdMNPV genome analysis (Kuzio *et al.*, 1999) due to overlap with Ld137. A similar situation occurs for the LdMNPV homologue of Se37 that overlaps with Ld155.

gp37 (Se25), named *spindle-like protein* or *fusolin* because of its obvious homology to the entomopoxvirus spindle-shaped proteins, is a conserved NPV gene (identity to AcMNPV homologue 56%; Liu and Carstens, 1996). The *gp37/fusolin* gene family may be essential for virus replication, based on the failure to construct an insertion mutant for this gene in AcMNPV (Wu and Miller, 1989). Furthermore, studies have suggested that *gp37/fusolin* is involved in enhancement of virus infection *in vivo* (Phanis *et al.*, 1999).

Unlike any baculovirus genome so far analysed, SeMNPV possessed two copies of *p26* (Se87 and 129). Se87 was located in the proximity of the non-*hr* (Fig. 2.1). This region is organized differently compared with AcMNPV, in contrast to the Se129 region (Fig. 2.2; see position of *p26* in all GeneParityPlots). It is possible that Se87 was acquired independently from a different source than Se129. This view is further supported by the 8% less identity of Se87 to its AcMNPV homologue than Se129. It is equally possible that Se129 has diverged from Se87 and has been rearranged following duplication. Transcripts have been identified for the AcMNPV homologue, which are synthesized by the host polymerase II both early and late in infection (Huh and Weaver, 1990). The P26 protein was localized primarily to the cytoplasm and is present in the membrane fraction of BV (Goenka and Weaver, 1996). A function is not yet been assigned to the P26 protein, but its conservation in all MNPV genomes analysed so far suggests a function basic to baculoviruses.

Baculovirus repeated ORFs (*bro* genes)
bro genes, present in a number of other baculoviruses and to date of no known function, were not identified in SeMNPV. Five copies of a homologue of Ac2 were identified in BmNPV (Gomi *et al.*, 1999) and there were 16 copies in LdMNPV (Kuzio *et al.*, 1999). In OpMNPV, a truncated version and two smaller *bro*-related ORFs are present (Ahrens *et al.*, 1997). Similarity searches revealed that Se13 showed weak homology (~25%) to some *bro* genes, particularly to BmNPV *bro-d* (Gomi *et al.*, 1999) and LdMNPV *bro-j* (Kuzio *et al.*, 1999) (Table 2.5). However, Se13 had higher homology (~33%) to Ac13 and its homologues Bm5, Op12 and Ld122 (Table 2.1). Furthermore, Se13 is located adjacent to a homologue of Ac14 (Se14) and these two genes are clustered in all baculoviruses compared. Therefore, we consider that

Se13 is not a *bro* gene *sensu stricto.*

Table 2.5 Comparison of Se13 with homologues of Ac13 and with *bro* genes.

The predicted amino acid sequence of SeMNPV ORF13 is compared with those of homologous of AcMNPV ORF13 and other baculovirus *bro* genes, *bro* group nomenclature is according to Kuzio *et al.* (1999).

ORF	Identity (%)	Similarity (%)	Length (aa)	*bro* group
	Amino acid sequence			
Ac13	33.6	44.5	327	
Bm5	33.0	45.3	331	
Op12	27.5	38.6	320	
Ld122	33.3	43.6	200	
Ac2	24.7	36.7	328	I a
Bm *bro-d*	28.4	38.4	349	I a
Ld *bro-n*	25.6	36.6	338	I a
Ld *bro-j*	29.0	40.4	403	I a
Bm *bro-a*	24.7	38.5	317	I c
Bm *bro-c*	25.3	39.7	318	I c
Ld *bro-b*	25.3	36.5	323	I c
Ld *bro-p*	22.1	33.8	337	I c

SeMNPV ORFs with homologues in other baculoviruses: LsNPV *p13* and *Helicoverpa zea* (Hz) SNPV *hoar*

SeMNPV possessed 22 ORFs that have no homologues in AcMNPV, BmNPV, OpMNPV or LdMNPV. Of these, two showed homology to other baculovirus ORFs. A homologue of the LsNPV *p13* gene (Se56) was found in SeMNPV, but the SeMNPV homologue is C-terminally extended (Wang *et al.*, 1995). The two leucine zipper-like structures present in LsNPV P13 (Wang *et al.*, 1995) were also conserved in Se56. The function of this ORF in LsNPV is unknown.

The SeMNPV genome contained an ORF (Se4) with homology to the *hoar* ORF of HzSNPV and *Helicoverpa armigera* NPV (Le *et al.*, 1997). In SeMNPV and HzSNPV, the *pk-1* gene is downstream of the *hoar* gene (Table 2.1; Le *et al.*, 1997). However, the upstream flanking ORFs of the HzSNPV *hoar* gene had no homologues in SeMNPV (HzSNPV ORF480 and ORF321) or were present in different locations in SeMNPV (HzSNPV ORF1–6 corresponding to Se138–132) (Le *et al.*, 1997; Table

2.1). The SeMNPV *hoar* ORF, like the *Heliothis* sp. homologues, contained a complex, A+T-rich, triplet repeat region (RAT-repeats) distributed over 330 bp and a C_3HC_4 (RING finger) zinc-binding motif.

Unique SeMNPV ORFs

Twenty ORFs in the SeMNPV genome were completely unique to this virus and did not exhibit significant homology to any sequence in the GenBank. Hence no putative functions could be assigned to these ORFs. The functions of these ORFs are being investigated. This number is roughly proportional to the size of the genome (Table 2.2).

Organization of the SeMNPV genome

The genomic organization, i.e. the order of genes, is similar in AcMNPV, BmNPV and OpMNPV, except for a small number of rearrangements (Ahrens *et al.*, 1997; Hu *et al.*, 1998; Gomi *et al.*, 1999). To investigate whether the organization in SeMNPV was collinear with these viruses and to the recently sequenced LdMNPV (Kuzio *et al.*, 1999), a comparison was made between the SeMNPV genome organization and those of AcMNPV, BmNPV, OpMNPV and LdMNPV by using GeneParityPlot analysis (Fig. 2.2; Hu *et al.*, 1998). The gene organization was most conserved in the 'central' region (ORFs 30–70) of the linearized baculovirus genomes, confirming the assumption of Heldens *et al.* (1997b). The 'left' part of the SeMNPV genome (ORFs 1–30) displayed a considerable number of gene inversions and translocations in the GeneParityPlot analyses. The 'right' part (ORFs 70–100) showed a high degree of gene scrambling (Fig. 2.2). From these analyses, it is concluded that the organization of SeMNPV is highly characteristic and distinct from those of AcMNPV, BmNPV, OpMNPV and LdMNPV.

By convention, the orientation of a circular baculovirus genome is determined by the relative position of two genes, *polyhedrin* at map unit 0 and *p10* approximately at map unit 90 (Vlak and Smith, 1982). In the initial GeneParityPlot analysis, the orientation of the SeMNPV genome appeared to be inverted for more than 50% of the ORFs compared with AcMNPV, BmNPV, OpMNPV and LdMNPV. This led to perpendicularity in the graph where collinearity was known to occur, i.e. in the conserved, central part of the genome. To facilitate convenient comparison and interpretation of the different genomes, the SeMNPV gene order was reversed before it was subjected to GeneParityPlot analysis. The previously satisfactory choices of *polyhedrin* and *p10* for the zero point and directional orientation, respectively, were not convenient for GeneParityPlot analysis in this case because both genes are located
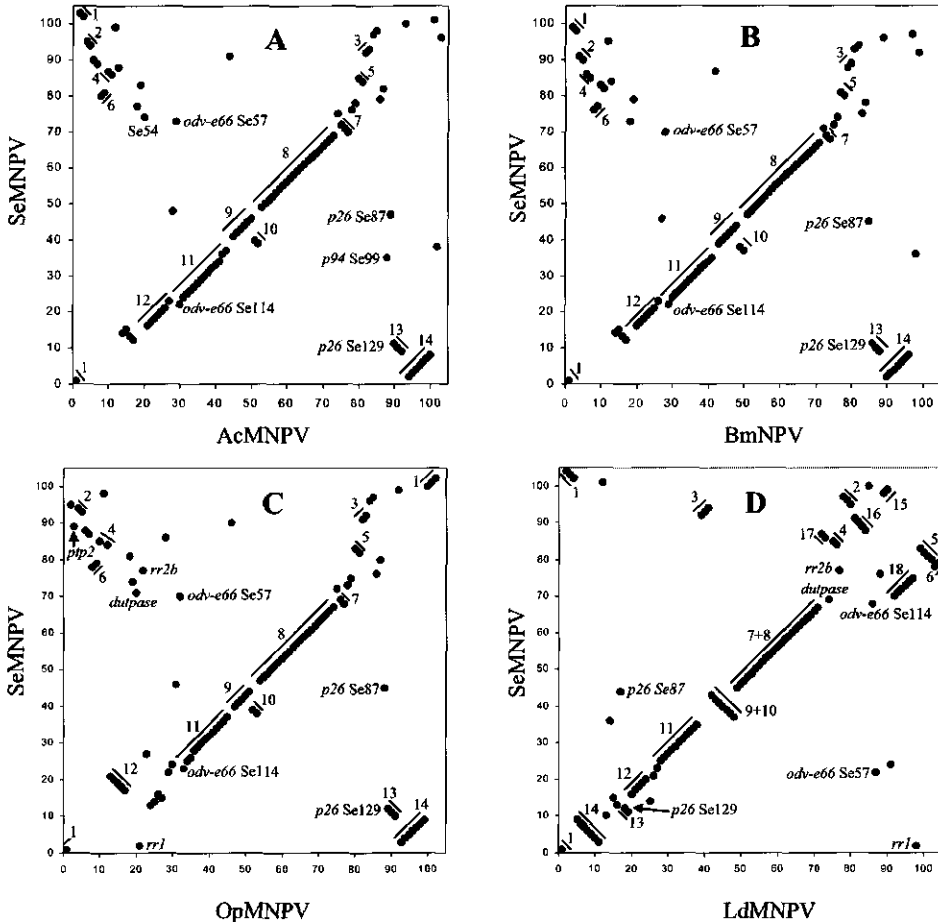
Fig. 2.2 GeneParityPlots of SeMNPV versus AcMNPV (A), BmNPV (B), OpMNPV (C) and LdMNPV (D). The plots are graphic representations of the collinearity of baculovirus genomes obtained by GeneParityPlot analysis (see Methods). Fourteen putative gene clusters of the SeMNPV genome that are similar to those of other baculoviruses are numbered (1–14) and indicated by lines. Four additional putative gene clusters were identified between SeMNPV and LdMNPV (D), numbered 15–18. The positions of the *p94, odv-e66, p26, ptp2, rr1, rr2b* and *dutpase* genes are indicated.

in regions that show extensive rearrangements (Fig. 2.2).

Comparison of the relative gene order between SeMNPV and AcMNPV, BmNPV, OpMNPV and LdMNPV revealed the presence of certain gene clusters that are conserved in all baculovirus genomes compared. These clusters were numbered according to their sequential appearance in the GeneParityPlots. Fourteen clusters conserved in all five baculoviruses were identified (Fig. 2.2, Table 2.1). Cluster 3 was

interrupted in SeMNPV by the insertion of Se17 and Se18, which are unique to SeMNPV. Cluster 12 is discontinuous in LdMNPV because five copies of the *bro* gene and two other ORFs are inserted in this cluster. Four additional clusters were identified in the GeneParityPlot of SeMNPV versus LdMNPV (Fig. 2.2D; Table 2.1; clusters 15–18). Furthermore, clusters 2 and 5 were extended to include genes Se15 and Se38+Se41, respectively. Clusters 7 and 8 and clusters 9 and 10 were present as two contiguous clusters in LdMNPV and SeMNPV. This is in contrast to the other three baculoviruses, where the positions of genes of these clusters in the gene parity plot were perpendicular to each other due to inversion of one of the clusters (Ayres *et al.*, 1994; Gomi *et al.*, 1999; Ahrens *et al.*, 1997, Kuzio *et al.*, 1999). The additional and the enlarged clusters of SeMNPV and LdMNPV suggest that the genomic organization of SeMNPV is more closely related to that of LdMNPV than to that of AcMNPV, BmNPV and OpMNPV. This agrees with the phylogenetic analysis of single genes such as *egt*, *lef2* and *rr1*, which shows that SeMNPV is more closely related to LdMNPV than to AcMNPV, BmNPV or OpMNPV (Chen *et al.*, 1997, 1999; Hu *et al.*, 1997; van Strien *et al.*, 1997). Thus, juxtaposition of ORFs can be used as a phylogenetic marker to study the ancestral relationship of baculoviruses, independent of the evolution of individual genes.

Between- and within-baculovirus genome rearrangement

Comparison of SeMNPV with AcMNPV, BmNPV, OpMNPV and LdMNPV showed that baculovirus genomes may vary due to deletions, (gene) insertions, inversions and duplications (Ayres *et al.*, 1994; Gomi *et al.*, 1999; Ahrens *et al.*, 1997; Kuzio *et al.*, 1999). The mechanisms underlying these rearrangements are still unclear. Transposable elements that may play a role in rearrangements of baculovirus genomes have been identified in several baculoviruses (Friesen, 1993; Jehle, 1996; Jehle *et al.*, 1997). Furthermore, there is evidence to suggest that *hr* sequences are related to the generation of variant baculovirus genotypes (Muñoz *et al.*, 1999).

Genome rearrangements also occur within one baculovirus species, as is the case for SeMNPV. A mutant SeMNPV, containing a single deletion of approximately 25 kb, was obtained within the first passage in cell culture (Heldens *et al.*, 1996). This deletion is located approximately between 17.5 and 42.0 kb (±1.0 kb) and encompasses Se15 to 41 (Table 2.1). So far, none of these ORFs has been shown to be essential for virus replication. However, deletion mutant SeMNPV polyhedra produced *in vitro* do not cause any pathological effect *in vivo* nor does the injection of BV into the haemolymph. In contrast, mutant SeMNPV BV was highly infectious for Se-UCR1 cells and resulted in polyhedron production (Heldens *et al.*, 1996).

Therefore, at least one gene located in the deleted sequences contains information that is important for virulence *in vivo*.

In conclusion, sequencing revealed that the genome of SeMNPV is distinct from those of other baculoviruses both in gene content and arrangement. Two, probably independently acquired, *p26* and *odv-e66* genes are present. Notably, SeMNPV lacks homologues of the *gp64*, *ie-2* and multiple *bro* genes. Furthermore, SeMNPV and LdMNPV may have a recent common ancestor, whereas they are more distantly related to AcMNPV, BmNPV and OpMNPV on the basis of gene homology and genomic organization. The gene order in the 'central' part of baculovirus genomes is highly conserved, whereas the gene order in the other segments has been subjected to multiple rearrangements. The GeneParityPlot analyses demonstrate that this method can be used as an independent means of phylogenetic study and can provide an initial view of the conservation of gene clusters and how viruses may have obtained additional genes. The genomic sequences absent in the deletion mutant of SeMNPV contain information that is important for virulence *in vivo*. Further studies will concentrate on the functional analysis of the ORFs that are unique to SeMNPV. These studies will provide insight in the roles these ORFs may play in the high virulence and narrow host-range of SeMNPV.

Methods

SeMNPV DNA isolation, cloning, PCR and sequence determination
The SeMNPV isolate (Gelernter and Federici, 1986b) was originally obtained from B. A. Federici (Department of Entomology, University of California, Riverside, CA, USA) in the form of polyhedra and was called SeMNPV-US1 (Muñoz *et al.*, 1998). The polyhedra were propagated in fourth-instar *S. exigua* larvae (Smits *et al.*, 1988).

The SeMNPV *Xba*I plasmid and *Sau*3AI cosmid libraries were described previously (Heldens *et al.*, 1996). The *Xba*I-A and *Xba*I-B fragments were too large to be cloned into pUC19. The *Xba*I-A fragment was subcloned from SeMNPV cosmids 24 and 17 into plasmids Se*Bgl*II-H, Se*Pst*I-M (cosmid 24) and Se*BSpe*I-5.4, Se*Spe*I-H and Se*BP*-5.6 (cosmid 17). In addition, the *Xba*I-B fragment was subcloned from SeMNPV cosmid 22 into plasmids Se*Spe*IH-3.2, Se*BSpe*I-6.3, Se*Spe*IH-2.8, Se*HBgl*II-6.2 and Se*Eco*RI-2.2. Some of these clones were described previously (Broer *et al.*, 1998; van Strien *et al.*, 1996).

Four regions of the SeMNPV genome were difficult to clone, as restriction fragments or sequencing attempts resulted in premature termination. The four regions

were located at positions 16530–19106, 32831–36808, 44839–46417 and 54758–54937 on the SeMNPV genome. These regions were amplified by PCR and cloned into pGEM-T vectors (Promega). Template DNA for sequencing was purified from plasmids by using Jetstar columns according to manufacturer's protocol (ITK Diagnostics).

Sequencing was done by using plasmid, cosmid and PCR products from both strands of the viral genomic DNA as templates. Sequence reactions were performed at the Sequencing Core Facilities of Wageningen Agricultural University and Queen's University (Kingston, Ontario, Canada) by primer walking.

DNA sequence analysis

Genomic DNA composition, structure, repeats and restriction enzyme pattern were analysed with the Wisconsin Genetics Computer Group programs (Devereux *et al.*, 1984) and DNASTAR. ORFs consisting of more than 50 amino acids were considered to encode proteins. Relevant ORFs (119 of 139) were checked for maximum alignment with known baculovirus gene homologues from GenBank; ORFs with significant overlap of *hr* sequences were excluded. The overlap between any two ORFs with known baculovirus homologues was set to a maximum of 25 amino acids; otherwise the largest ORF was selected. DNA and protein comparisons with entries in the genetic databases were performed with FASTA and BLAST programs (Pearson, 1990; Altschul *et al.*, 1990). Multiple sequence alignments were performed with the GCG PileUp and Gap computer programs with gap creation and extension penalties set to 9 and 2, respectively (Devereux *et al.*, 1984). Percentage identity indicates the percentage of identical residues between two complete sequences. Motif searches were done against the Prosite release 14 database (Fabian *et al.*, 1997; Bairoch *et al.*, 1997). Prediction of transmembrane domains was accomplished with SignalP and PHD software (Nielsen *et al.*, 1997; Rost and Sander, 1993). GeneParityPlot analysis was performed on the SeMNPV genome versus the genomes of AcMNPV, BmNPV, OpMNPV and LdMNPV as described previously (Hu *et al.*, 1998).

Genbank accession

The GenBank accession number of the SeMNPV genomic sequence reported in this paper is AF169823.

Chapter 3

Identification of a novel occlusion derived virus-specific protein in *Spodoptera exigua* multicapsid nucleopolyhedrovirus

Understanding the molecular basis of the distinct biological properties of *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV), such as its narrow host range and high virulence, requires detailed information on the temporal expression and subcellular localization of SeMNPV gene products. The expression of two unique SeMNPV ORFs, 116 (Se116) and 117 (Se117), which show 45% amino acid similarity, was analyzed. Se116 and Se117 were expressed both in cultured cells and in larvae of *S. exigua*, as polyadenylated transcripts of 0.80 and 0.75 kb, respectively. These transcripts initiated from ATCA(G/T)T promoter motifs, commonly found for baculovirus early genes. Se116 transcripts were detected with increasing abundance from 8 h to 48 h post infection (p.i.), whereas Se117 transcripts were present from 4 h p.i. and most abundantly at 24 h p.i. Western blot analysis of infected Se301 cells revealed 27 and 23 kDa proteins for Se116 and Se117, respectively. C-terminal GFP-fusion proteins of Se116 and Se117 were primarily localized in the nucleus of Se301 cells. When Se301 cells were infected with SeMNPV both GFP-fusion proteins were localized in the virogenic stroma of the nucleus. While the function of the Se116 protein is still enigmatic, the Se117 protein appeared to be a structural protein associated with nucleocapsids of occlusion-derived SeMNPV virions, but not of budded virus.

Introduction

The baculovirus *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV) is highly pathogenic for the beet army worm. This insect causes significant crop damage and economic losses (Federici and Maddox, 1996). SeMNPV differs from many other NPVs infectious for this insect in having an extremely narrow host-range, limited to *S. exigua*, and a relatively high virulence (Smits, 1987). These biological characteristics make SeMNPV a succesful alternative for chemical insecticides. To optimize its application, it is important to understand the molecular basis of host specificity and virulence, and hence the relation of specific SeMNPV genes with these characteristics.

The complete sequence and genetic organization of the SeMNPV genome have recently been elucidated (IJkel *et al.*, 1999). Comparison of the coding potential of SeMNPV with other completely sequenced baculoviruses (Ayres *et al.*, 1994; Gomi *et al.*, 1999; Ahrens *et al.*, 1997; Kuzio *et al.*, 1999; Chen *et al.*, 2001; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000) revealed a total of sixteen ORFs unique to SeMNPV. None of the unique SeMNPV genes has yet been studied in detail. Two of these, SeMNPV ORF116 (Se116) and ORF117 (Se117), are likely active genes since they (i) contain baculovirus early promoter motifs (IJkel *et al.*, 1999), indicative for possible transcription by host and / or viral RNA polymerases and (ii) are not located in the genome region that possesses non-essential genes since it is rapidly and frequently deleted during passage of SeMNPV in cell culture (Heldens *et al.*, 1996).

Se116 is located between nt 114,465 and 115,142 and encodes a putative protein of 225 amino acids (aa) with a predicted molecular weight of 26.3 kDa (IJkel *et al.*, 1999). Se117 is located next to Se116, between nt 115,249 and 115,824, and encodes a putative protein of 191 aa with a molecular weight of 22.5 kDa. Both genes have the same polarity as the *polyhedrin* gene. The genes are located in between SeMNPV homologs of *Autographa californica* (Ac) MNPV ORF38 (*p47*) and ORF40. An AcMNPV ORF39 homolog is absent in SeMNPV (Ayres *et al.*, 1994; IJkel *et al.*, 1999) as well as in many other baculoviruses. To date, the function and significance of AcMNPV ORF39 or its homolog in BmNPV are unknown (Gomi *et al.*, 1999).

To elucidate the molecular basis of the distinct biological properties of SeMNPV with respect to other baculoviruses detailed information on the temporal expression, subcellular localization and function of the SeMNPV-specific genes is essential. Here, we report on the transcriptional and translational analysis of the unique SeMNPV ORFs 116 and 117. We have raised antibodies against the Se116 and Se117 proteins and this allowed the analysis of their expression and subcellular localization. The Se117 protein appeared to be a novel occlusion-derived virion protein, whereas a function could not (yet) be assigned to Se116.

Results

SeMNPV ORF116 and ORF117

Appropriate searches of protein databases showed that the putative Se116 and Se117 proteins are unique to SeMNPV and have no homologs in other biological systems. Surprisingly, Psi-blast searches (Altschul *et al.*, 1997) and Gap alignment revealed that the predicted amino acid sequence of Se116 shares 45% similarity and 29% identity with Se117 (Fig. 3.1). Except for the N-termini the conserved residues are equally distributed throughout the Se116 and Se117 sequences. Furthermore, the Se116 protein has, in contrast to the Se117 protein, an additional hydrophilic C-terminal domain, which is proline- and glutamine-rich. Analysis of both predicted amino acid sequences did not identify any recognizable motifs, transmembrane regions, GPI-anchors or signal peptides. According to the NNCN score analysis (Reinhardt and Hubbard, 1998) both proteins are probably localized in the nucleus of the cell. Blast searches revealed that the identified homology is present only in the amino acid sequence and not in the nucleotide sequence, suggesting that a recent gene duplication is not the likely cause of existence of these two homologous genes next to one another. More likely is a tandem insertion at this position. Both ORFs are present in the forward orientation in the genome (IJkel *et al.*, 1999).

```
Se116 : MKVNTGTLCLAIDSVALVVVKIH-RSVMRARSCMDPEINLTTMSDM-HGNLKVAQYDIQY :  58
Se117 : --MDGGEVKLAIDSVLKDLRFVNKGTLLLP-HYARRDEKMPDM-DIFSGTIDSVKKSLDH :  56


Se116 : QINYA-NKTVDKMDMTVNNMMSETITNDLMIMIKNPAEQINGDQQ----MSHIEASQHIF : 113
Se117 : BENVMLRKITTN-KMSIDDLMSRHTRSFLTFMLMKYVDSDRNDTQPDFDMVYIKTMVYMV : 115


Se116 : QTNDAFDINNFYLYBEQNKFDYMLTMVNITNSSVDB-SGHMFSYLTDKLLFLRRLGNKMY : 172
Se117 : RHLKQMKMKHYGQQMKCFKNDHVKABIHVVD-INDBRDBSCWNHLLQRINRTKELGKRMD : 174


Se116 : IHIHBVDAVAALIYMQKQQQLPNPPIVFEPPPILYRPHQRQANNPHRQQGIFN : 225
Se117 : ARLABFRDLQGR--MBTSV-------------------------------- : 191
```

Fig. 3.1 Alignment of the predicted amino acid sequences of the Se116 and Se117 proteins. Gaps, introduced to optimize the alignment, are indicated with dashes. Shading is used to indicate the occurrence of identical (black) or substitutional (grey) amino acids.

The regions upstream of Se116 and Se117 were analyzed for the presence of possible transcription start sites [TATAA], baculovirus consensus early promoter motifs [ATCA(G/T)T and CGTGC] (Blissard and Rohrmann, 1990; Pullen and Friesen, 1995; Carstens *et al.*, 1993), baculovirus late transcription start site [DTAAG] (Blissard and Rohrmann, 1989), downstream activating elements [(A/T)CACNG] (Friesen, 1997) and host factor binding sites [GATA and CACGTG] (Kogan and Blissard, 1994). An early gene transcription motif (TATAA-N$_{23}$-ATCAGT) was found 64 nt upstream of the putative translational start of Se116 (Fig. 3.2). A polyadenylation signal sequence (AATAAA) was identified only 2 nt downstream of the TAA stop codon of Se116 (Fig. 3.2). The Se117 5' upstream sequence includes two early gene transcription motifs (TATAA-N$_{13}$-TGCATT) starting at -87 and (TATAA-N$_{24}$-ATCATT) at -69 nt relative to the translational start codon (Fig. 3.2). A polyadenylation signal sequence was identified 32 nt downstream of the TGA stop codon of Se117 (Fig. 3.2). This computational analysis suggests that both ORFs may be active genes.

```
114361 ttaaagattaacagcaatgcttatcgtaatgtaccaagtaTATAAacggcgtagaaaact
       L  K  D

114421 gcatcaatATCAGTgaaagcgtagctattgccacaatcgagtcaatgaaaatgaacacgg
                                                    ORF116 M  K  M  N  T

114481 gaacattgtgcctggccatcgacagcgtcgccctcgacgtgcgaaagattcatcagtccg
       G  T  L  C  L  A  I  D  S  V  A  L  D  V  R  K  I  H  Q  S

                       Se116SP1
114541 tgatgagggcgcgctcttgcatggactttgaaataaatctgcccgatctaagcgacattc
       V  M  R  A  R  S  C  M  D  F  E  I  N  L  P  D  L  S  D  I

114601 attgcaacctcaaggtggcccagtacgatattgattatttgatcaattacgcgaataaag
       H  C  N  L  K  V  A  Q  Y  D  I  D  Y  L  I  N  Y  A  N  K

                                                      Se116SP4
114661 atgtcgacaagatggacatgaccgtcaacaacatgataagcgagacaataaccaacgagc
       D  V  D  K  M  D  M  T  V  N  N  M  I  S  E  T  I  T  N  E

114721 tagagattatgctgaaaaactttgcggagcaaattaacggtgatcaacagtatagtcaca
       L  E  I  M  L  K  N  F  A  E  Q  I  N  G  D  Q  Q  Y  S  H

114781 ttaaggcgtgtcagcacattttttcaaactaacgacgcctttgacattaacaactttatt
       I  K  A  C  Q  H  I  F  Q  T  N  D  A  F  D  I  N  N  F  Y

114841 tgtatttggaacaaaacaaattcgactacgtgctgacttttgtgaacattacaaactcta
       L  Y  L  E  Q  N  K  F  D  Y  V  L  T  F  V  N  I  T  N  S

                                                      Se116SP3
114901 gcgtcttgccgtcgagccatatgtttctctatttgacggataaactgttgttcttgcgac
       S  V  L  P  S  S  H  M  F  S  Y  L  T  D  K  L  L  F  L  R

114961 gtctctgcaataagatttacatacatattcacgaagtggacgccgtcgccgctctcattt
       R  L  C  N  K  I  Y  I  H  I  H  E  V  D  A  V  A  A  L  I
```

```
115021 acatgcagaaacagcagcaattgcctaatcctcctattgtttttgaacctccaccaattc
       Y  M  Q  K  Q  Q  Q  L  P  N  P  P  I  V  F  E  P  P  P  I

                        Se116SP2
115081 tgtatcgtcctcatcaacgtcaagcgaacaatccacatcgacaacaaggaatttttaatt
       L  Y  R  P  H  Q  R  Q  A  N  N  P  H  R  Q  Q  G  I  F  N

115141 aatAATAAAattttgtattttTATAAtttgaatttttatTGCATTttacgtcaccaataa
                                                    ┌─►
115201 tttagtaTATAAatcgtcgaggaatacgacatcgacATCATTcacataatggacggcgga
                                                   ORF117 M  D  G  G

115261 tcggtgtgtcttgctctcgacagcgttctcaaagacttgcgatttgtcaacaagcaaacg
       S  V  C  L  A  L  D  S  V  L  K  D  L  R  F  V  N  K  Q  T

115321 ctcttgcttccgcactatgcggactttgacgaaaagatgcccgatttggatattttttcc
       L  L  L  P  H  Y  A  D  F  D  E  K  M  P  D  L  D  I  F  S

115381 tgcactttggatagtgtaaaaaaatcgctagaccatctcgaaaatgtcatgttgcgcaag
       C  T  L  D  S  V  K  K  S  L  D  H  L  E  N  V  M  L  R  K

       Se117SP1
115441 gacaccaccaacaagatgagtctcgacgatttgatatctcggcacacgcgcagcgagcta
       D  T  T  N  K  M  S  L  D  D  L  I  S  R  H  T  R  S  E  L

                                        Se117SP4
115501 gagtttatgctcatgaattatgttgatagcgacaggaacgacacgcaacccgactttgac
       E  F  M  L  M  N  Y  V  D  S  D  R  N  D  T  Q  P  D  F  D

115561 tatgtatatataaagacgtgcgtgtacatggtcagacatctcaagcaaatgaagatgaaa
       Y  V  Y  I  K  T  C  V  Y  M  V  R  H  L  K  Q  M  K  M  K

115621 cactatggtcagcaattaaagtgttttaaaaatgatcatgtcaaggcgtttatacacgtc
       H  Y  G  Q  Q  L  K  C  F  K  N  D  H  V  K  A  F  I  H  V

               Se117SP3
115681 gtcgacataaatctgcccagggactcgtcgtgctggaaccacttgctacaaaagataaat
       V  D  I  N  L  P  R  D  S  S  C  W  N  H  L  L  Q  K  I  N

                 Se117SP2
115741 cgtacgcgcgaattgtgtaaaaggatcgacgctagactggcagagtttagagatttgcaa
       R  T  R  E  L  C  K  R  I  D  A  R  L  A  E  F  R  D  L  Q

115801 ggccgtatggaaacgtctgtttgattATAAGttattgtacaatgattctatataaAATAA
       G  R  M  E  T  S  V

115861 Atgatacatttatatgcacaacagtttattttctttttatcatgcgttgcgccggtttgtt
       M  H  N  S  L  F  S  F  I  M  R  C  A  G  L  F
```

Fig. 3.2 The SeMNPV ORF116 and ORF117 genomic region, from nucleotides 114,361 to 115,920 (IJkel *et al.*, 1999). Location of TATA-boxes, baculovirus consensus early initiation motif ATCA(G/T)T, polyadenylation signal, start and stop codons are denoted in bold. The sequences of the primers used for RT-PCR and 3' mapping (Se116SP1, Se117SP1) and 5' mapping (Se116SP2, 3, 4 and Se117SP2, 3, 4) are underlined. The determined transcriptional start site for the Se116 and Se117 transcripts are indicated with an arrow and their poly(A) chain attachment sites are double underlined.

41

Northern blot and RT-PCR analysis of Se116 and Se117 transcripts

Temporal regulation of the Se116 and Se117 transcripts was examined by Northern blot analysis and RT-PCR using total RNA isolated from SeMNPV infected Se301 cells as template. This analysis would also provide information about the number and sizes of the Se116 and Se117 transcripts. For Se116 a single transcript of 0.8 kb was detected at 8 h p.i. and remained detectable until 48 h p.i. (Fig. 3.3A). Northern analysis of Se117 revealed one major transcript at 0.75 kb (Fig. 3.3B). This transcript was detectable at 4 h p.i., reached maximal transcription levels at 24 h p.i. and steady state levels remained high until 48 h p.i. The transcript sizes of 0.8 kb and 0.75 kb are compatible with predicted ORF sizes of 678 nt for Se116 and 576 nt for Se117, respectively (IJkel *et al.*, 1999). The size of the Se116 and Se117 transcripts suggests that both transcripts were most likely polyadenylated, assuming that the putative early promoter sequences were used as transcription start sites. Twenty-four h p.i., a second 1.4 kb transcript appeared and remained visible until 48 h p.i. using a Se117-specific riboprobe. The latter transcript is interpreted as a read-through from the Se117 gene, since no 1.4 kb transcript was detected at similar time points p.i. using the Se116-riboprobe. The observed bands at 1.9 kb (Fig. 3.3A, B) and 0.2 kb (Fig. 3.3A) are non-viral, since they also appear in the mock infected lane and are probably derived from rRNA (18S and 5S).

RT-PCR was performed to confirm independently the results of the Northern analysis and to obtain further information about the temporal regulation of the transcripts at earlier times. One primer internal to Se116 (Se116SP1; Fig. 3.2) or Se117 (Se117SP1; Fig. 3.2) and the PCR anchor primer were used to amplify fragments of 650 and 475 bp, respectively (Fig. 3.3C, D). The obtained RT-PCR fragments appeared to be specific for Se116 and Se117 upon hybridization with a Se116 and Se117-specific riboprobe, respectively (Fig. 3.3C, D). Consistently, Se116 gene sequences were amplified at 4 h p.i., which was increased at 8 h p.i. and remained at steady state level up to 72 h p.i. (Fig. 3.3C). Se117 gene sequences were also amplified at 4 h p.i., peaked at 24 h p.i. and remained detectable until 72 h p.i. (Fig. 3.3D). Thus, both Se116 and Se117 are most likely early genes according to the current view on early vs. late transcription (Friesen, 1997).

To investigate if Se116 and Se117 were also transcribed during SeMNPV infection in *S. exigua* larvae (*in vivo* infection), RT-PCR was performed on RNA isolated from fat body tissue. The RT-PCR products obtained were cloned into pGEM-T and two clones of each were sequenced. The obtained sequences matched, as expected, the Se116 or Se117 sequences. Both Se116 and Se117 gene sequences were amplified at 48 and 72 h p.i. (Fig. 3.3E, F), indicative of the presence of Se116 and
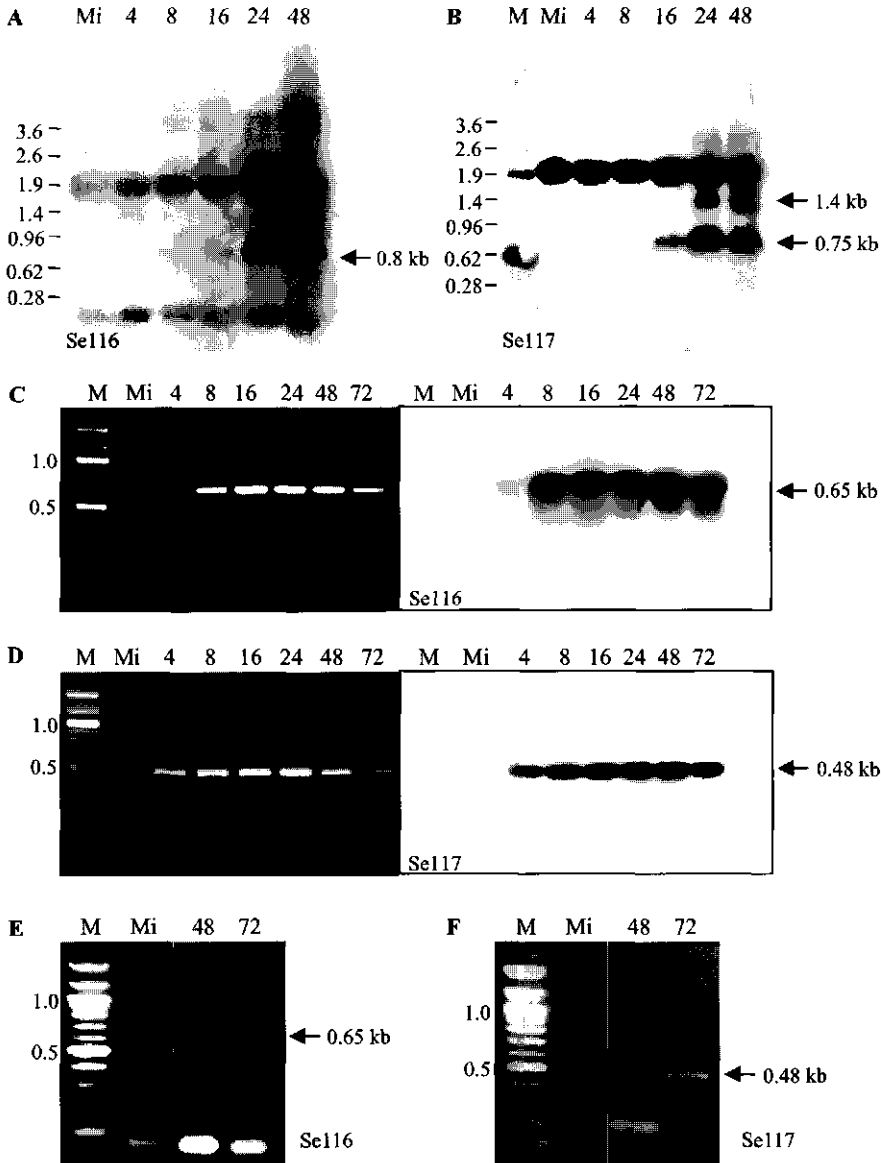
Fig. 3.3 Temporal expression of the Se116 and Se117 transcripts.

(A and B) Northern analysis of Se116 and Se117, respectively. The size of specific hybridization bands is indicated on the right. (C and D) The left panels show RT-PCR analysis of Se116 and Se117, respectively, performed on total RNA extracted from SeMNPV-infected Se301 cells obtained under the same conditions as those in A and B. The right panels of C and D show hybridization patterns of the obtained RT-PCR bands using Se116 and Se117 riboprobes, respectively. (E and F) RT-PCR analysis performed on total RNA extracted from SeMNPV-infected *S. exigua* fat body tissue. Times p.i. are indicated above the lanes (Mi, mock infected). Size standards (M, marker) are indicated in kb.

Se117 transcripts at those times p.i. Thus, both ORFs were transcribed not only upon infection of cultured insect cells but also upon in vivo infection. Furthermore, the obtained RT-PCR fragments indicate that both transcripts contain a poly(A) tail, since the oligo-dT anchor primer was successfully used to synthesize first-strand cDNA of these ORFs.

Transcriptional mapping of the 5' end of the Se116 and Se117 transcripts
The 5' end of the Se116 and Se117 transcripts was determined by 5' RACE analysis with total RNA isolated at an early and late time point p.i. from Se301 cells. A single cDNA was detected at all early and late times tested for both Se116 and Se117 transcripts. The start site of Se116 transcription maps 36 nt upstream of the ATG translation initiation codon at first A in the sequence ATCAGT (Fig. 3.2). The start site of Se117 transcription was located 12 nt upstream of the ATG start codon, at first A in the sequence ATCATT (Fig. 3.2). Thus, the Se116 and Se117 transcripts initiated within a baculovirus consensus early promoter motif.

The Se116 and Se117 early promoters were further investigated using cycloheximide to inhibit protein synthesis. In the presence of this inhibitor in the medium of infected cells only immediate early promoters are transcribed i.e. by host RNA polymerases. Both Se116 and Se117 transcripts were absent in RT-PCR analysis at 4, 8 and 16 h p.i. in the presence of 100 µg/ml cycloheximide (results not shown), a concentration at which protein synthesis is largely inhibited (Ross and Guarino, 1997). This was in contrast to cells infected in parallel, but not treated with cycloheximide. So, Se116 and Se117 are not immediate early genes, since immediate early protein synthesis was required for their transcription, but most likely delayed-early.

Transcriptional mapping of the 3' end of the Se116 and Se117 transcripts
The 3' end of the Se116 and Se117 transcripts was determined by sequencing the 3' specific RACE-PCR fragments obtained after RT amplification of total RNA purified from Se301 cells or *S. exigua* fat body tissue at various times p.i. The obtained sequences indicated that the 3' ends of the transcripts isolated from cultured insect cells are identical to those isolated from fat body tissue and showed little or no variation in their poly(A) attachment sites. The 3' end of the Se116 transcript was located 25 nt downstream of the stop codon at the second T in the sequence TAAT (Fig. 3.2). A conventional mammalian polyadenylation signal, consisting of an AATAAA motif and 20-30 nt downstream a diffuse (G)U-rich sequence, is located 2 nt downstream of the stop codon. The 3' end of the Se117 transcript was mapped 53 nt downstream of the stop codon at the C in the sequence ATGC (Fig. 3.2). A single

conventional mammalian polyadenylation motif is present 32 nt downstream of the stop codon. Hence, the Se116 and Se117 transcripts ended, respectively, 18 and 16 nt downstream of the AATAAA motif. These results suggest that the conventional mammalian polyadenylation signal was used for the termination of SeMNPV transcripts in both Se301 and *S. exigua* fat body cells. Taken together, the data of the 5' and 3' end mapping predict transcript sizes of 739 and 641 nt for Se116 and Se117 (excluding the poly(A) tail), respectively. The expected sizes are in agreement with the 0.8 and 0.75 kb sizes determined by Northern analysis for the major transcripts of Se116 and Se117, respectively, assuming a poly(A) tail of 100-150 nucleotides.

Immunodetection of the Se116 and Se117 proteins in infected cells

The Se116 and Se117 proteins have predicted molecular weights of 26.3 and 22.5 kDa, respectively. Antibodies were prepared by immunization of rabbits with PREP-cell purified Se116 or Se117 protein produced in *E. coli*. Western analysis of extracts of SeMNPV infected Se301 insects cells showed a specific protein of 27 kDa at 48 and 72 h p.i. for the Se116 antiserum (Fig. 3.4A). The size of the 27 kDa protein is in agreement with the predicted 26.3 kDa size of the putative Se116 translation product, suggesting that no major posttranslational modification occurred. Overexposure did not reveal this specific 27 kDa protein earlier than 48 h p.i. A ~46 kDa protein was detected from 4 until 48 h p.i. (Fig. 3.4A) using the same Se116 antiserum. This protein should be considered to be of non-viral origin, since it also appeared in mock-infected cells.

Western analysis of extracts from SeMNPV infected Se301 cells revealed a specific polypeptide with an apparent size of 23 kDa, when using the Se117 antiserum (Fig. 3.4B). This size is in agreement with the predicted 22.5 kDa size of the putative Se117 translation product, suggesting that no major posttranslational modification occurred. The protein was detected from 8 until 72 h p.i., with a maximum at 48 h p.i. This is in agreement with transcription data for Se117 (Fig. 3.3B). The Se117 antiserum showed some cross-reactivity to a ~30 kDa protein that should be considered to be non-viral, since it also appeared in mock-infected cells (Fig. 3.4B).

Localization of the Se116 and Se117 proteins in insect cells
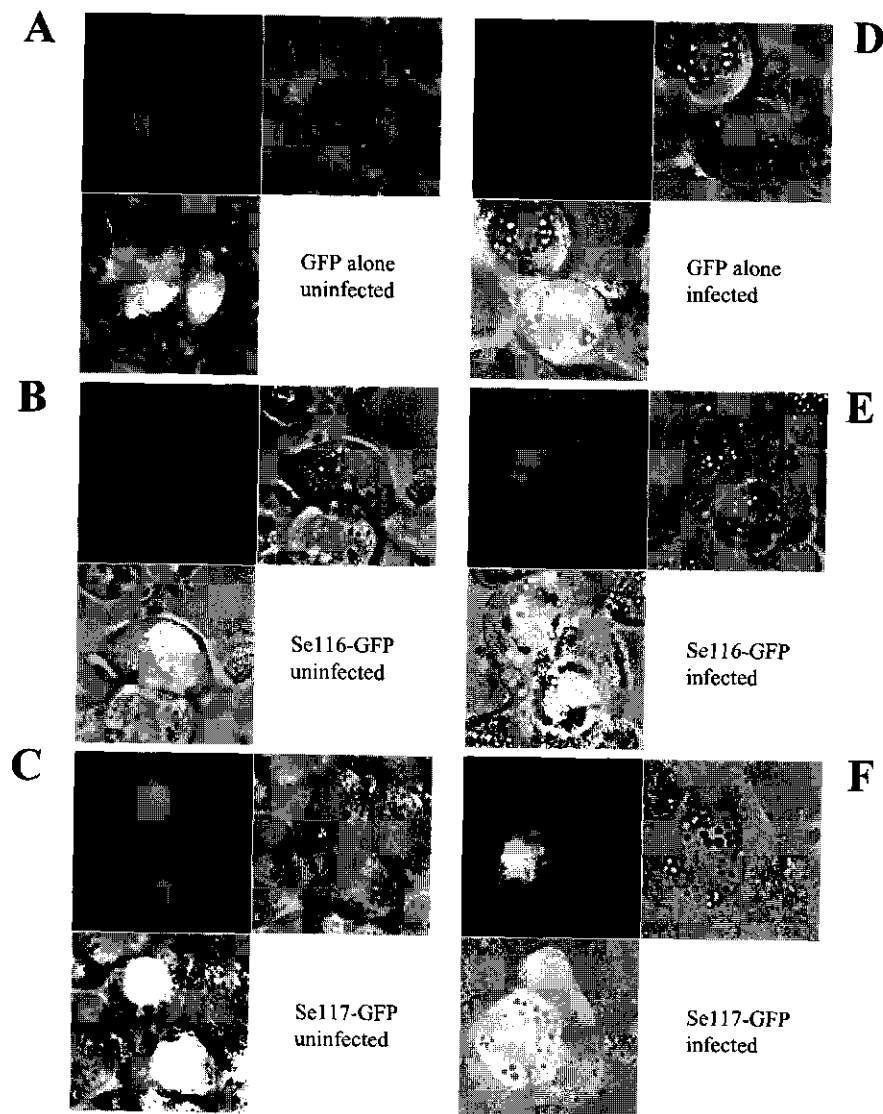
The subcellular localization of the Se116 and Se117 proteins was investigated using C-terminal GFP-fusion constructs. These GFP-fusion constructs were made in plasmid p166BRNX-AcV5 (IJkel *et al.*, 2000). As a negative control, GFP alone was cloned in the same vector. Se301 cells were transfected with 5 μg plasmid DNA, incubated for 48 h and examined for fluorescence by confocal laser scanning

45

A   Mi   4   8   16   24   48   72

```
97 —
67 —

45 —

30 —                                          ◄—— 27 kDa

20 —

14 —  Se116
```

B   Mi   4   8   16   24   48   72

```
97 —
67 —

45 —

30 —

                                              ◄—— 23 kDa
20 —

14 —
     Se117
```
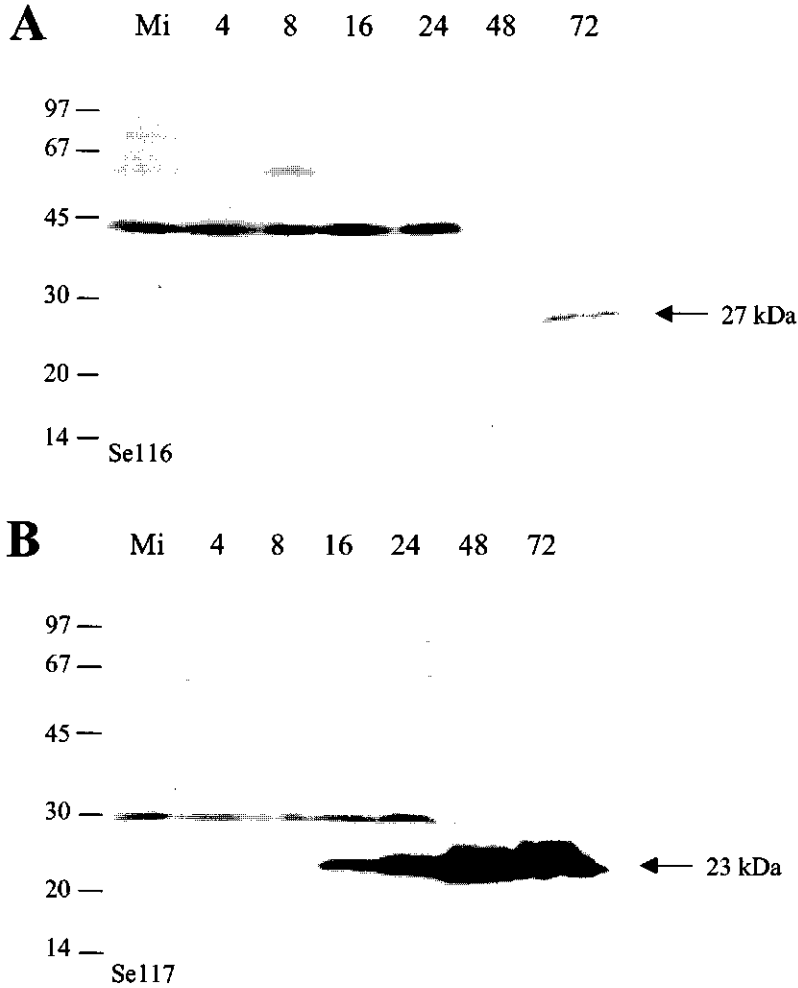
Fig 3.4 Western blot analysis of the Se116 (A) and Se117 (B) proteins in SeMNPV infected Se301 cells (70,000 cells/lane). The corresponding times p.i. are indicated above the lanes (Mi, mock infected). The Se116 and Se117 proteins were identified using the Se116 and Se117 polyclonal antiserum, respectively, and detected with a chemiluminescent substrate. Size standards are indicated in kDa and immunoreactive proteins are indicated by arrows.

microscopy. The non-fused GFP protein showed homogeneous fluorescence in the cytoplasm and nucleus (Fig. 3.5A). However, the Se116 and the Se117 GFP-fusion proteins were primarily localized in the nucleus (Fig. 3.5B, C). Thus, the discrete patterns of fluorescence are due to the linked Se116 and Se117 sequences and are consistent with their computer-predicted nuclear localization (Reinhardt and Hubbard,

1998). A few discrete foci of fluorescence were also observed in the cytoplasm for the Se116 GFP-fusion protein (Fig. 3.5B).

Fig. 3.5 Localization of the Se116 and Se117 GFP-fusion proteins in Se301 cells without or with SeMNPV infection. Se301 cells were transfected with the control plasmid p166AcV5-GFP (A), the plasmid p166AcV5-Se116GFP (B) or with p166AcV5-Se117GFP (C). At 48 h after transfection the cells were examined by confocal laser scanning microscopy for fluorescence. Se301 cells were transfected as in A (D), B (E) or C (F) and infected 24 h post transfection. At 48 h p.i. the cells were examined by confocal laser scanning microscopy for fluorescence. Phase contrast micrographs are shown to the right of the fluorescence graph. Overlay micrographs of the fluorescence and phase contrast micrographs are shown below the fluorescence micrographs.

The localization of the Se116 and Se117 proteins during infection was also investigated to obtain further insight into their function. Se301 cells were transfected with p166Se116-GFP or -Se117-GFP and 24 h later infected by SeMNPV with a MOI of 10. Cells were incubated for 48 h p.i. and examined for fluorescence by confocal laser scanning microscopy. When GFP was not linked to either the Se116 or the Se117 protein, fluorescence was uniformly present throughout the cytoplasm and nucleus (Fig. 3.5D). The Se116 GFP-fusion protein, however, was primarily localized in the center of the nucleus of infected cells in a network of granular material, known as the virogenic stroma (Fig. 3.5E). Furthermore, fluorescence was also observed in the periphery of the nucleus along the nuclear membrane. This often ring-shaped fluorescence colocalizes with that of heterochromatin, which is displaced and marginalized by the emerging stroma (Williams and Faulkner, 1997). No fluorescence was observed in polyhedra. A few discrete foci of intense fluorescence were observed in the cytoplasm (Fig. 3.5E), as was also observed in the absence of SeMNPV infection (Fig. 3.5B). For Se117-transfected and infected Se301 cells, intense fluorescence was observed in the center of the nucleus cells in the virogenic stroma (Fig. 3.5F). Although more difficult to observe, due to absorption and / or distortion of the fluorescence by the electron-dense polyhedra components, the Se117 GFP-fusion protein was also detected in polyhedra (Fig. 3.5F).

Immunodetection of the Se116 and Se117 proteins in BV and ODV

To investigate further if the Se116 and Se117 proteins are structural components of SeMNPV, Western analysis of budded virus (BV) and occlusion derived virus (ODV) was conducted. Equivalent amounts of BV or ODV-derived protein were loaded on SDS-PAGE (data not shown). The Se116 protein was not detected in preparations of BV or ODV (Fig. 3.6A). The Se117 protein (23 kDa), however, was detected in ODV but not in BV (Fig. 3.6B). Thus, the 23 kDa Se117 translation product appeared to be an ODV-specific protein.

The location of the Se117 protein in ODV was determined by Western analysis of nucleocapsid (NC) and envelope (E) fractions, prepared in the presence of protease inhibitors. The purity of the ODV NC and E fractions was tested by SDS-PAGE analysis only (Fig. 3.6C), since no antibodies for SeMNPV ODV NC- or E-specific proteins are available. The banding pattern of the NC and the E fraction was distinct (Fig. 3.6C). Furthermore, the two major SeMNPV ODV NC proteins (31 and 37 kDa) were absent in the E fraction, whereas the two major E proteins (41 and 46 kDa) were absent in the NC fraction (Fig. 3.6C). Therefore, both fractions were considered to be sufficiently pure.
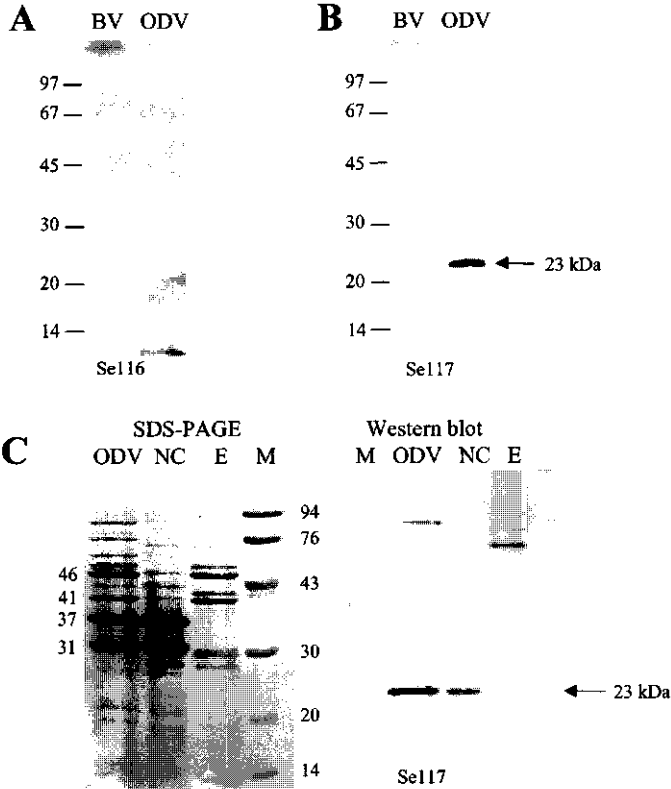
Fig. 3.6 Western blot analysis of the Se116 (A) and Se117 (B) proteins in BV and ODV. BV and ODV (10 µg/lane) represent isolated SeMNPV budded virus and alkali-disrupted occlusion derived virus, respectively. (C) SDS-PAGE and Western blot analysis of the Se117 protein in SeMNPV ODV, ODV nucleocapsid (NC) and envelope (E) fractions (10 µg/lane). The sizes of two NC- (30 and 37 kD) and two E- (43 and 48 kDa) specific proteins are indicated on the left. The Se116 and Se117 proteins were identified using the Se116 and Se117 polyclonal antiserum, respectively, and detected with a chemiluminescent substrate. Size standards (in the centre) are indicated in kDa and immunoreactive proteins are indicated by arrows.

When ODV fractions were prepared in the absence of protease inhibitors, the Se117 protein could be detected neither in the NC nor in the E fraction, but was present in unfractionated ODV (results not shown). This suggests that upon NP-40 fractionation of SeMNPV ODV the Se117 protein was degraded probably by a proteolytic activity. When ODV was fractionated in the presence of a cocktail of protease inhibitors, a major 23 kDa band was detected in the NC fraction, while only a very small amount was present in the E fraction (Fig. 3.6C). Thus, the 23 kDa Se117 protein was predominantly present in the NC fraction.

Discussion

In this study we report the temporal transcription and translation of Se116 and Se117 and the subcellular localization of their products. Both genes are unique to SeMNPV and may encode proteins which explain the unique characteristics of this virus. Northern analysis indicated that Se116 and Se117 transcript levels increased significantly from 4 till 48 h p.i., and remained at steady state levels throughout infection (Fig. 3.3). This suggests that the early nature of the ATCA(T/G)T promoter does not exclude transcription by the viral RNA polymerase at time points late in infection. Similar transcription patterns were observed for other baculovirus early genes, such as the AcMNPV *pnk/pnl* and *lef4* genes, the BmNPV *bro* genes, and the *ie1* genes of AcMNPV, OpMNPV and LdMNPV (Durantel *et al.*, 1998a and b; Kang *et al.*, 1999; Pearson and Rohrmann, 1997; Pullen and Friesen, 1995; Theilmann and Stewart, 1993).

Transcriptional mapping of Se116 and Se117 (Fig. 3.2) showed that both genes have transcription start and stop sites similar to those of other baculovirus early genes, such as the LdMNPV genes *ie1*, *g22* and *vPK* (Pearson and Rohrmann, 1997; Bishoff and Slavicek, 1994 and 1995), the AcMNPV *ie1*, *pnk/pnl* and *lef4* genes (Pullen and Friesen, 1995; Durantel *et al.*, 1998a and b) and the OpMNPV *ie1* and *opep-2* genes (Theilmann and Stewart, 1993; Shippam *et al.*, 1997). Although a second putative TGCATT early promoter motif was present at -87 nt from the Se117 start codon (Fig. 3.2), transcripts using this promoter were not detected. This could be caused by the somewhat unusual spacing, only 13 nt, between the TATA-box and the TGCATT sequence (Roeder, 1991; Lu and Miller, 1995). Furthermore, this putative early promoter is not completely consistent with the arthropod initiator cap site consensus [A(A/C/T)CA(G/T)T]) (Cherbas and Cherbas, 1993).

Western analysis detected a 26 kDa protein specific for Se116 from 48 h p.i. (Fig. 3.4A). This is somewhat unexpected because Se116-specific transcripts could be detected as early as 8 h p.i. (Fig. 3.3). The inability to detect the Se116 protein earlier than 48 h p.i. may be due to the low level of Se116 protein present at earlier time points and or the low affinity of the Se116 antiserum. The cross-reactivity of the Se116 antiserum with a cellular protein of ~46 kDa could be either due to the presence of a non-related protein or to a putative cellular cognate of Se116, which may be downregulated upon infection.

The Se117 translation product appeared to be a structural ODV-specific protein expressed already from 8 h p.i. onwards as detected by Western analysis (Figs. 3.4B and 3.6B). Such an early expression start is not often found for structural baculovirus proteins as they usually contain baculovirus consensus late promoters (Funk *et al.*,

1997). However, Se117 is probably not the only baculovirus gene, which is expressed early in infection and encodes a structural virion protein, since also the SeMNPV *odv-e66*, *vp80capsid*, *vp39capsid* and *p74* genes (IJkel *et al.*, 1999) possess early promoter motifs as do other baculovirus structural genes. It remains to be investigated, however, whether the latter genes use their early promoter motifs as transcription start sites. An alternative possibility is that the Se117 protein possesses different functions early and late in infection.

NP-40 fractionation of SeMNPV ODV revealed that the Se117 protein is predominantly associated with the nucleocapsid fraction. This technique has previously been used to recover intact nucleocapsids (Thiem and Miller, 1989). Since, the Se117 protein was only detected in the presence of protease inhibitors upon fractionation, it resides most likely at the outside of the ODV nucleocapsids, where it is easy accessible for proteases. The localization in ODV nucleocapsids is in agreement with (i) the computer predicted absence of a potential signal peptide, GPI-anchor or hydrophobic transmembrane domain in the Se117 sequence, indicative for envelope localization (Hong *et al.*, 1997), (ii) the sensitivity of the AcMNPV major nucleocapsid protein for proteolytic degradation only upon NP-40 treatment in contrast to the ODV P74 envelope protein (Faulkner *et al.*, 1997) and (iii) the absence of the AcMNPV tegument protein, GP41, in the nucleocapsid fraction of ODVs upon NP-40 fractionation (Whitford and Faulkner, 1992). The small amount of Se117 protein detected in the envelope/tegument fraction supports the supposition that it resides at the outside of the nucleocapsid.

All baculovirus nucleocapsid proteins studied to date are present in both BV and ODV (Funk *et al.*, 1997). Based on the results of SDS-PAGE analyses of AcMNPV BV and ODV nucleocapsids, it was hypothesized that their protein compositions would be different (Braunagel and Summers, 1994). The 23 kDa Se117 protein is a good candidate of an ODV-specific nucleocapsid protein. The absence of the Se117 protein in BV (Fig. 3.6B) supports this conclusion. It is possible that the Se117 protein is present in BV in a non-detectable form when using the Se117 antiserum. The latter explanation, however, is unlikely since no fluorescence was observed at the plasma membrane for the Se117 GFP-fusion protein in SeMNPV infected Se301 cells (Fig. 3.5F).

The region between the gene homologs of Ac38 and Ac40 differs in many baculoviruses (Lapointe *et al.*, 2000). Like in SeMNPV, an Ac39 homolog is also absent in OpMNPV, LdMNPV and the granuloviruses PxGV and XcGV, where respectively, an extensive inversion (Ahrens *et al.*, 1997), replacement with a second AcMNPV ORF25 (*dbp*) homolog (Kuzio *et al.*, 1999) and two deletions were

observed (Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000) compared to the AcMNPV genome. Therefore, this baculovirus region can be considered as a 'hot spot' for genome rearrangement. SeMNPV inserted two unique genes, Se116 and Se117, in this hypervariable region, that share a significant degree of amino acid homology but not nucleotide sequence homology. It is intriguing that both Se116 and Se117 antisera showed cross-reactivity to a cellular protein (Fig. 3.4), which supports a hypothesis that both genes could be derived from cellular origin. If so, Se117 has evolved to a structural SeMNPV ODV nucleocapsid protein, while the function of Se116 still remains enigmatic. Also the function of the highly homologous N-terminus of both proteins may be important for their putative functions.

Upon SeMNPV infection the Se116 and Se117 GFP-fusion proteins were primarily localized in a network of granular material, known as the virogenic stroma. The virogenic stroma is considered a *de novo* product of baculovirus infection in which progeny virions are assembled (Williams and Faulkner, 1997). Since both the Se116 and Se117 proteins are only present (as of yet) in SeMNPV they may play a specific role in virion assembly process or virogenic stroma arrangement. As a minor component of the ODV nucleocapsid, the Se117 protein may function as a scaffold protein in ODV nucleocapsids or could be involved in early infection events, like nucleocapsid entry or transport once fusion occurred. Future studies using SeMNPV site-specific and null mutants will determine the significance of these genes in the SeMNPV infection cycle and shed light on their potential role in the host specificity and virulence.

Material and methods

Computer-assisted analysis

Se116 and Se117 (IJkel *et al.*, 1999) were analyzed using software of the Predict Protein server (Rost, 1996) and the ExPASy server (Appel *et al.*, 1994) for the prediction of domains, motifs, transmembrane regions and subcellular localization (Reinhardt and Hubbard, 1998). DNA and protein comparisons with entries in the updated GenBank/EMBL, SWISS-PROT and PIR databases were performed with BLASTn, FASTA and Psi-BLAST programs (Pearson, 1990; Altschul *et al.*, 1997). Multiple sequence alignments were performed with the GCG PileUp computer programs with gap creation and extension penalty set to 8 and 2, respectively (Devereux *et al.*, 1984). Alignment editing was performed with Genedoc Software.

Plasmid constructions

The complete coding regions of Se116 and Se117 were amplified by high fidelity "Expand" long template PCR (Boehringer Mannheim) from plasmid pHBg6.2 (IJkel *et al.*, 1999) using primers containing 5' *Bam*HI and 3' *Hind*III restriction sites and cloned into pGEM-T. The plasmids were named pGEMSe116 and pGEMSe117 and used for production of gene-specific riboprobes. The plasmids pTriExSe116 and pTriExSe117 were obtained by cloning the *Bam*HI/*Hind*III fragments from pGEMSe116 and pGEMSe117 into the *Bam*HI and *Hind*III sites of the expression vector pTriEx-1 (Novagen) in frame with the HSV and His-tag sequences. These plasmids were used for overexpression of the Se116 and Se117 proteins in *E. coli.* To determine the localization of Se116 and Se117 in insect cells, GFP-fusion constructs were made. The complete coding regions of Se116 and Se117 were amplified by high fidelity "Expand" long template PCR (Boehringer Mannheim) from plasmid pHBg6.2 (IJkel *et al.*, 1999) using primers containing 5' *Bam*HI and 3' *Eco*RI restriction sites and cloned into the *Bam*HI and *Eco*RI sites of the previously described p166AcV5-GFP vector (IJkel *et al.*, 2000) and named p166Se116-GFP and p166Se117-GFP, respectively. Plasmid DNA was purified using Jetstar columns according to manufacturer's protocol (ITK Diagnostics). For each construct, the nucleotide sequence was checked using an automated DNA sequencer (Wageningen University, The Netherlands).

Cells, Insects and Viruses

The *Spodoptera exigua* cell-line Se301 (Hara *et al.*, 1995b) and the SeMNPV-US1 isolate (Gelernter and Federici, 1986b; Muñoz *et al.*, 1998) were maintained and propagated as described previously (IJkel *et al.*, 2000). A culture of *S. exigua* insects was maintained according to Smits and Vlak (1988a). SeMNPV protein synthesis was inhibited by adding cycloheximide (100 µg/ml, Sigma) to the Se301 cells 30 min before infection.

Total RNA isolation, Northern blot, RT-PCR, 3' and 5' RACE analysis

Total RNA was isolated from 2×10^6 mock-infected and SeMNPV-US1-infected Se301 cells (MOI of 5 $TCID_{50}$ units/cell) at 0, 4, 8, 16, 24, 48 and 72 h p.i. Total RNA was also isolated from fat body tissue obtained after dissection of six mock-infected and SeMNPV-US1 infected *S. exigua* larvae (1.7×10^4 polyhedra/larva) at 0, 48 and 72 h p.i. Cells and tissue were resuspended in 500 µl Trizol (GibcoBRL) and 100 µl chloroform, incubated for 8 min and centrifuged at 14,000 g for 15 min at 4°C. The RNA in the water fraction was precipitated using isopropanol, centrifuged at 14,000 g

for 10 min at 4°C, washed with 70% ethanol and resuspended in 50 μl water. The RNA solutions were incubated at 55°C for 10 min and quantified by absorbance at 260 nm.

For Northern analysis 8 μg of total RNA was electrophoresed in 1.5% agarose in the presence of glyoxal (Ausubel *et al.*, 1994) and blotted onto a Hybond N nylon membrane (Amersham). The northern blot was hybridized with a Se116 or Se117 messenger-specific riboprobe generated by T7 and SP6 polymerase, respectively, with [α-32P]CTP and *Bam*HI digested pGEMSe116 and pGEMSe117. Fragment sizes were determined by staining the molecular weight marker (Promega RNA marker) with methylene blue after transfer onto the membrane.

RT-PCR was performed using the 5'/3' RACE kit (Roche) employing 2 μg purified total RNA as template per time point. First-strand cDNA synthesis was performed using AMV reverse transcriptase and the oligo-dT anchor primer according to manufacturer's instructions. The cDNA-mixtures were amplified by PCR using the PCR anchor primer and the gene specific primers Se116SP1 or Se117SP1 (Fig. 3.2). The obtained PCR-products were analyzed in 1.2% agarose gels.

The amplified RT-PCR products of 8 and 16 h p.i. were used to determine the 3' end of the Se116 *in vitro* messenger, while the RT-PCR products of 4 and 24 h p.i. were used for the Se117 *in vitro* messenger. The 3' ends of the Se116 and Se117 transcripts isolated from fat body tissue were determined using the amplified RT-PCR products of 48 and 72 h p.i., respectively. All PCR-products were gel purified, cloned into pGEM-T and sequenced with T7 or SP6 primers.

The 5' ends of the Se116 and Se117 transcripts were determined using the 5'/3' RACE kit (Roche) employing 2 μg purified total RNA as template per time point. (Se116: 8 and 48 h p.i.; Se117: 4 and 48 h p.i.) Briefly, first strand cDNA synthesis was performed with the gene specific primers Se116SP2 and Se117SP2 (Fig. 3.2). The cDNAs were purified with the High Pure PCR purification kit (Roche) and a poly(A) tail was added to the 3' ends using the terminal transferase with dATP. The tailed cDNAs were amplified by PCR using the oligo dT-anchor primer and the nested gene specific primer Se116SP3 or Se117SP3 (Fig. 3.2). A second PCR was performed using the PCR anchor primer and the nested primer Se116SP4 or Se117SP4 (Fig. 3.2). The obtained PCR products were gel purified, cloned into pGEM-T and sequenced with T7 or SP6 primers.

Production of polyclonal antibodies

Cultures of *E. coli* B121 containing pTriExSe116 and pTriExSe117 were grown to an optical density at 600 nm of 0.5 and induced with 1 mM isopropyl-β-D-

thiogalactopyranoside (IPTG). After 4 h at 37°C, cells were collected by centrifuging at 6,500 *g* for 15 min at 4°C, resuspended in 20 mM Tris-HCl pH7.5, lysed with lysozyme, sonicated and centrifuged at 14,000 *g* for 10 min at 4°C. The insoluble fractions were washed with 20 mM Tris-HCl pH7.5 and resuspended in 1% SDS. Aliquots of the soluble and insoluble fractions were electrophoresed in 12.5% SDS polyacrylamide gels according to Laemmli (1970) and stained with Coomassie brilliant blue.

The insoluble fractions containing the Se116 or Se117 proteins were purified by continuous-elution electrophoresis using the Model 491 Prep Cell (Bio-Rad) according to manufacturer's protocol. Elution fractions were collected, electrophoresed in 12.5% SDS polyacrylamide gels and the protein bands visualized by silver staining (Morrisey, 1981). Fractions containing Se116 or Se117 were pooled and dialyzed for 36 h against running buffer without SDS. The samples were concentrated by freeze-drying and dissolved in H_2O. Protein concentrations were determined with the Bio-Rad protein assay (Bio-Rad). Two rabbits were injected intramuscularly with 100 µg purified Se116 or Se117 protein in Specol (ID-Lelystad, The Netherlands). The rabbits were boosted after seven days with 300 µg purified protein. Serum was collected 2, 4, 6 and 8 weeks after the boost injection. Western analysis using *E. coli* BL21 extracts expressing pTriExSe116 or pTriExSe117 and PREP-cell purified Se116 or Se117, which were verified by Western analysis with a His-tag antibody (Clontech), were used to test the production of specific antisera.

Western analysis
Monolayers of Se301 cells were mock- or SeMNPV-US1-infected at a multiplicity of infection (MOI) of 5 $TCID_{50}$ units/cell. Cells were harvested at 0, 4, 8, 16, 24, 48 and 72 h p.i., pelleted, resuspended in PBS and lysed in SDS-PAGE loading buffer by boiling for 5 min. Protein samples were then separated by SDS-PAGE and transferred on Immobilon-P nitrocellulose membrane (Millipore) by semi-dry electrophoresis transfer (Ausabel *et al.*, 1994). The membranes were incubated overnight in 10% block solution (Boehringer Mannheim) in TBS-T buffer (50 mM Tris-HCl, 200 mM NaCl, 0.1% Tween-20, pH7.6) at 4°C. The membranes were allowed to react in TBS-T with Se116 or Se117 antiserum diluted 1:5000 for 1 h at room temperature. After washing in TBS-T (three times 15 min), the membranes were incubated for 1 h at room temperature with horse radish peroxidase-conjugated donkey anti-rabbit immunoglobulin (Amersham) diluted 1:5000 in TBS-T. After washing in TBS-T (three times 15 min) the signal was detected by ECL technology as described by the manufacturer (Amersham).

Purification of SeMNPV BV and ODV

Eight-hundred *S. exigua* fourth-instar larvae were infected by contamination of artificial diet with SeMNPV-US1 polyhedra (10x LD_{99}) (Smits and Vlak, 1988a). To purify BVs, 15 ml of hemolymph was collected 3 days p.i. in 0.5 ml 0.1 x TE (TE is 10 mM Tris, pH 7.5, 1.0 mM EDTA) containing 5 mM phenylthiocarbamide to inhibit prophenoloxidase activity. Hemolymph was clarified twice at 3,000 *g* for 10 min at room temperature. The supernatant was filtered (0.45 µm filter) and the filtrate overlaid onto a 35 ml 25-50% continuous sucrose gradient in 0.1 x TE. Gradients were centrifuged at 100,000 *g* for 120 min at 4°C (Beckman SW28, 24,000 rpm). The BV band was collected, diluted twice and centrifuged at 100,000 *g* for 90 min at 4°C (Beckman SW28, 24,000 rpm). The virus pellet was resuspended in 200 µl 0.1 x TE.

Polyhedra were purified from larvae as described previously (IJkel *et al.*, 2000). Briefly, ODVs were liberated from polyhedra (40 mg/ml) by incubating at room temperature for 15 min in 0.1 M Na_2CO_3, 166 mM NaCl, 10 mM EDTA, pH 10.5. Undissolved polyhedra were removed by low-speed centrifugation for 5 min (500 *g*). The supernatant (5 ml) was layered onto a 35 ml, 25-56% (w/w) continuous sucrose gradient in 10 mM Tris-HCl, pH 7.5 and centrifuged at 100,000 *g* for 90 min at 4°C (Beckman SW28, 24,000 rpm). The multiple virus bands were collected, washed by dilution in 0.1 x TE, concentrated by centrifugation at 55,000 *g* for 60 min at 4°C (Beckman SW41, 18,000 rpm) and resuspended in 0.1 x TE. The purity and integrity of BVs and ODVs were checked by electron microscopy.

Fractionation of virions into envelope and nucleocapsid

Virus was fractionated into envelope and nucleocapsid using a modification of the protocol of Braunagel and Summers (1994). In a 250-µl reaction, 250 µg of ODV was incubated in 1.0% NP-40, 10 mM Tris, pH 8.5, at room temperature for 30 min with gentle agitation. The solution was then layered onto a 5-ml 30% (v/v) glycerol/10 mM Tris, pH 8.5 cushion and centrifuged at 150,000 *g* for 60 min at 4°C (Beckman SW55, 35,000 rpm). The envelope proteins were recovered from the top of the gradient, acetone-precipitated and concentrated by centrifugation (4000 *g*, 30 min), and the pellet was dissolved in 10 mM Tris, pH 7.4. The pelleted nucleocapsids were resuspended in 10 mM Tris, pH 7.4. This fractionation procedure was also carried out in the presence of the protease inhibitor cocktail Complete according to manufacturer's protocol (Boehringer Mannheim).

Fluorescence microscopy

Se301 cells (3×10^5) were grown on glass cover slips and transfected with 5 µg of plasmid DNA using Cellfectin (GibcoBRL). Cells were superinfected with SeMNPV with a MOI of 5 $TCID_{50}$ units/cell 24 h post transfection. At 48 h post transfection or 48 h p.i., the cells were examined with a Zeiss LSM510 (confocal) laser scanning microscope for fluorescence using an excitation wave length of 488 nm and an emission band pass filter of 505 - 530 nm.

Chapter 4

Characterization of *Spodoptera exigua* multicapsid nucleopolyhedrovirus ORF17/18, a homolog of *Xestia c-nigrum* granulovirus ORF129

Spodoptera exigua multicapsid nucleopolyhedrovirus (SeMNPV) contains a number of genes with a homolog found so far only in a distantly related baculovirus. One of these, SeMNPV ORF17/18 (Se17/18) shares 55% amino acid similarity to ORF129 of *Xestia c-nigum* granulovirus (XcGV). To gain insight in the significance of this gene, its expression was analyzed. Se17/18 was transcribed in cultured *S. exigua* 301 cells, as a polyadenylated transcript of 1.1 kb. 5' RACE analysis demonstrated that the Se17/18 transcription start sites mapped at 134, 131 and 126 nt upstream of the putative translational start codon. These sites overlapped with a baculovirus consensus early promoter motif. Se17/18 transcripts were detected by Northern and RT-PCR with increasing abundance from 8 h to 24 h post infection (p.i.) and reseded until 72 h p.i. A chicken polyclonal antiserum was raised that reacted specifically to Se17/18 protein produced in *E. coli*. However, no immunoreactive protein was detected in SeMNPV-infected Se301 cells. A C-terminal GFP-fusion protein of Se17/18 was primarily localized in the cytoplasm of Se301 and Sf21 cells. Based on low homology of Se17/18 to (methyl) transferases its possible role in transcription regulation is discussed.

Introduction

The beet army worm (*Spodoptera exigua*; Lepidoptera, Noctuidae) is an agricultural important pest insect in (sub)tropical regions of the world and in greenhouses (Federici and Maddox, 1996). *S. exigua* multicapsid nucleopolyhedrovirus (SeMNPV) is a highly pathogenic baculovirus for the beet army worm and differs from many other NPVs in that it is infectious for *S. exigua* larvae only and has a relatively high virulence (Smits, 1987). These biological characteristics make SeMNPV an attractive biological alternative for chemical insecticides. Detailed information on the expression and function of specific SeMNPV genes is important to gain insight in baculovirus biology.

In previous studies, a mutant SeMNPV was obtained within the first passage in cell culture that lacked virulence *in vivo* (Heldens *et al.*, 1996). This mutant contained a contiguous 25 kb deletion encompassing SeMNPV ORF15 to ORF41 (IJkel *et al.*, 1999). So far, none of these ORFs has been experimentally shown to be essential for biological activity or virulence *in vivo*. Of these, SeMNPV ORF17 (Se17) and ORF18 (Se18) were previously characterized as unique to SeMNPV (IJkel *et al.*, 1999). Upon resequencing they appeared to be linked into a single ORF (Se17/18). Surprisingly, Se17/18 has a homolog in *Xestia c-nigrum* granulovirus (XcGV; Hayakawa *et al.*, 1999), which is only very distantly related to SeMNPV. The function and significance of Se17/18 and its XcGV homolog are unknown. Therefore, the functionality of Se17/18 is investigated by determining its transcriptional and translational activity as well as subcellular localization.

Results

SeMNPV ORF17/18 and its XcGV homolog

Upon resequencing ORF17 and ORF18 (IJkel *et al.*, 1999), these ORFs appeared to be linked, forming one contiguous ORF within the SeMNPV genome. The occurrence of an additional G, present after nt 20,512, resulted in this correction, leading to one continuous ORF further denoted as Se17/18. Se17/18 is located between nt 20,121 and 20,989 and encodes a putative protein of 289 amino acids (aa) with a predicted molecular weight of 33.7 kDa. Appropriate searches of protein databases showed that the putative Se17/18 protein is highly homologous to XcGV ORF129 (Xc129). Psi-blast searches (Altschul *et al.*, 1997) and Gap alignment revealed that the predicted Se17/18 aa sequence shares 55% similarity and 38% identity with Xc129 (Fig. 4.1). The N-terminal part of the Se17/18 protein is 25 aa smaller than that of the Xc129

```
Se1718 :    MSS  I  LF----------------------LFGL   K LK T H T A   R :  39
Xc129  : MKR  VF   AASSIKRCYLKDEDTDLLPPTNLTQEITNK   K Y  E  S V L  G E :  65


Se1718 : VGA   L  KQFIKY   RFN  INV YR TEY  R----DYVR   V GL Q  VK HLN : 100
Xc129  : GRA   L  RNYKQ     SED  LEK FK ARP NCTAA -LT     RQ  M QPK : 129


Se1718 :    MAD V  K  GNK  NL SL ID  CK QRAMH KGNGT DPAN AL  Q- V  G  ED : 164
Xc129  :    INN   I  EKI SR  AY  NF KKSOP -----E EDSD -I  MA  TN DKA : 189


Se1718 : NI-V  DR      -MI  V DVEY IGK A RRIK KS -Y  SRHRI     LM MN  A  : 226
Xc129  : TRG  -    D   Q  VH -TNL GK FF NS N L NE FY -KLNV  V  YQ LY R  : 249


Se1718 : V IT  RCFESAY MLS      QPP  GEESI K  ND-  E   YV   SK TY   NL YS : 289
Xc129  : I AN  ESIDGEH KVQ       SAS AMP V FLQ  R   SN AS  RI M :  312
```

Fig. 4.1 Alignment of the predicted amino acid sequences of the Se17/18 and Xc129 proteins. Gaps, introduced to optimize the alignment, are indicated with dashes. Shading is used to indicate the occurrence of identical (black) or substitutional (grey) amino acids.

protein. The conserved aa residues are equally distributed throughout the Se17/18 and Xc129 sequences. Most of their cysteine residues are conserved and located at the C-terminal part of the proteins. Analysis of both predicted aa sequences did not reveal motifs for transmembrane regions, GPI-anchors nor signal peptide sequences. According to the NNCN score analysis (Reinhardt and Hubbard, 1998) both proteins are probably localized in the cytoplasm of the cell. Blast searches revelaed that the identified homology is also present on nucleotide level and resembles the similarity found on amino acid level. Se17/18 is present in the reverse orientation in the genome relative to the *polyhedrin* gene (IJkel *et al.*, 1999).

The regions upstream of Se17/18 and Xc129 were analyzed for the presence of possible transcription start sites, baculovirus consensus early CA(G/T)T and late DTAAG promoter motifs (Blissard and Rohrmann, 1990; Pullen and Friesen, 1995; Blissard and Rohrmann, 1989), downstream activating elements (Friesen, 1997) and host factor binding sites, such as GATA and CACGTG (Kogan and Blissard, 1994). Two early gene transcription initiation motifs (TATA-N_1-GGCATT) and (TATAA-N_{14}-GACAGT) were found, respectively, 142 and 50 nt upstream of the putative Se17/18 translational start codon (Fig. 4.2 and 4.3). An additional TATA-box was located 117 nt upstream of the start codon, which could be linked to an AACATT

sequence mapped 56 nt further downstream. A GATA host factor binding site mapped 88 nt upstream of the putative ATG of Se17/18 (Fig. 4.2 and 4.3). A polyadenylation signal sequence (AATAAA) was identified starting at the last A of the TAA stop codon of Se17/18 (Fig. 4.2).

```
                                            ⌐►  ⌐►      ⌐►
21160 caccacgcaagccaagaacagcgtaataaTATAgGGCATTatctaattaaacctTATAAa
      V  V  C  A  L  F  L  T  I  I  Y  P  M

21100 ctatcaatgttaatattcttctaGATAgtttatcaacgctaatcagaccgAACATTttgt

21040 aTATAAactttgacaaacttGACAGTgctagtaatttgcttaaccgatgaaatgtcgagc
                                      ORF17/18    M  S  S

20980 gtcatcgaaaatcaattgtttctgttcggtctgttgccgatggaaatgaaattgaaaata
      V  I  E  N  Q  L  F  L  F  G  L  L  P  M  E  M  K  L  K  I

20920 ctaaaatacacgggtcacgacacatatttcgccgtgacggaggatcgcgtcggcgcagat
      L  K  Y  T  G  H  D  T  Y  F  A  V  T  E  D  R  V  G  A  D

         Se17/18SP1
20860 accattttgttgactgaaaaacaatttataaaatatttcgatacgcgccgtttcaactat
      T  I  L  L  T  E  K  Q  F  I  K  Y  F  D  T  R  R  F  N  Y

                                              Se17/18SP4
20800 tgcataaacgtcgactatagattgactgaatattttaaaagacaagactacgtacgactg
      C  I  N  V  D  Y  R  L  T  E  Y  F  K  R  Q  D  Y  V  R  L

                                           Se17/18SP3
20740 ccgttgctcgtcgacggcctgtttcaaacgtacgtcaagactcatctcaatggttatgtg
      P  L  L  V  D  G  L  F  Q  T  Y  V  K  T  H  L  N  G  Y  V

20680 atggcagacgatgtgttgctaaaaacgttttttggcaataaagacattacgaatttgtct
      M  A  D  D  V  L  L  K  T  F  F  G  N  K  D  I  T  N  L  S

20620 ctgttaatcgatggacaattttgcaaataccaaagagctatgcacgaaaaaggaaacgga
      L  L  I  D  G  Q  F  C  K  Y  Q  R  A  M  H  E  K  G  N  G

20560 acatacgacccgccaatgatgcgctgttgttgcaatatctgggatcgggagcaaatgag
      T  Y  D  P  A  N  D  A  L  L  L  Q  Y  L  G  S  G  A  N  E

20500 gacaatattgtggaatgttttgatcgtttaattgatttcgacaaaatgctgagattagtc
      D  N  I  V  E  C  F  D  R  L  I  D  F  D  K  M  L  R  L  V

20440 gccgatgttgaatacactatcggcaaattatttgctctgcgcagaataaaaacaaagtct
      A  D  V  E  Y  T  I  G  K  L  F  A  L  R  R  I  K  T  K  S

20380 ttatacaaattctttagccgacaccgaatcacattgccagaatgtttgatgcccatgaat
      L  Y  K  F  F  S  R  H  R  I  T  L  P  E  C  L  M  P  M  N

          Se17/18SP2
20320 ttggccaaagtttttgattacatacacgcgttgctttgaaagtgcgtacttgatgctaagt
      L  A  K  V  L  I  T  Y  T  R  C  F  E  S  A  Y  L  M  L  S

20260 ggagaatgtgtcaattgtcaggaaccctcaattggcgaggaaagtattttaaagctgctg
      G  E  C  V  N  C  Q  E  P  S  I  G  E  E  S  I  L  K  L  L

20200 gataatgattgggaagatttgtatgtaattttgtttgttcgaaatgtacatattgtttg
      D  N  D  W  E  D  L  Y  V  I  L  F  C  S  K  C  T  Y  C  L
```

```
20140 ttgaacttaccgtattcgtaAATAAAtttgtataataataataattttgttttattttat
       L  N  L  P  Y  S
```

```
20080 cgttgttttcgacgaaaaagtgtcgttaagagcgcaaacattagtgtatttaagtatcta
```

```
20020 ttgtttattgaatttataatgaacaaattactcatactttttctgctattaaacgcggcg
                        M  N  K  L  L  I  L  F  L  L  L  N  A  A
```

Fig. 4.2 The SeMNPV ORF17/18 genomic region, from nucleotides 19,961 to 21,160 (IJkel *et al*., 1999). Location of TATA-boxes, baculovirus consensus early initiation motif ATCA(G/T)T, host factor binding site GATA, polyadenylation signal, start and stop codons are denoted in bold. The sequences of the primers used for RT-PCR and 3' mapping (Se17/18SP1 and 5' mapping (Se17/18SP2, 3, 4) are underlined. The determined transcriptional start sites for the Se17/18 transcripts are indicated with arrows and its poly(A) chain attachment site is double underlined.

The promoter region of Se17/18 was compared to the Xc129 promoter region (Fig. 4.3). Several putative Se17/18 promoter elements were also present in the Xc129 promoter region. Two additional GATA-sequences were present in the promoter region of Xc129. The early promoter element most closely to the start codon of Xc129 contained two CATT sequences, which were more distally located from the TATA-box as compared to the TATAA-N_{14}-GACAGT promoter motif of Se17/18. The presence of putative promoter elements and polyadenylation signals suggests that both ORFs may be active genes.

```
                              GATA
Se17/18 : ATAGTTTTGAATCTGAGCAATCGATAGTGGCAACGCCGGGTACTTCTGCTTTCGTGGACAA :  60
Xc129   : AAACCAAGGGGCCCCGGCCGTCAAACGATGGATCTTTAAAATACACTGGGCTAGCCATAG :  60


                              TATA    CATT                 TATA
Se17/18 : CACCACGCAAGCCAAGAACAGCGTAATAATAAAGGGCATTATCTAATCAAAGCTAAAAA : 120
Xc129   : TTAGTTTAAAATCAATTTTTGATTGGCTTCCTTTTATAC-ACAATTTGTATGACGACATC : 119


                         GATA                           CATT
Se17/18 : CTATCAAAGTTAAGATATACCTTGATAGTCAATCAACGCTAATCAGACCGAAGATTT--- : 177
Xc129   : TGAGGGTGAAAGAGAAGGAGGACAAACAAGTGTTTATGAACTAGAT--AACATTGAT : 176
          GATA                                GAT  A    GAT


            TATAA        CAGT                          ATG
Se17/18 : -TGTAAATAAACTT-TGAGAAACTTGACAGAGCTAGTAATGGCTAACCGATGAAAGG :  234
Xc129   : AGGTATATAAAGGAAAGCCAAAAT--ATTAGTCA-GCACAGTCAGTAAT------AATG :  226
          A                                     CATTCATT
```

Fig. 4.3 Alignment of the nucleotide sequences upstream of the putative translational start codons of Se17/18 and Xc129. Gaps, introduced to optimize the alignment, are indicated with dashes. Shading is used to indicate the occurrence of identical nucleotides.

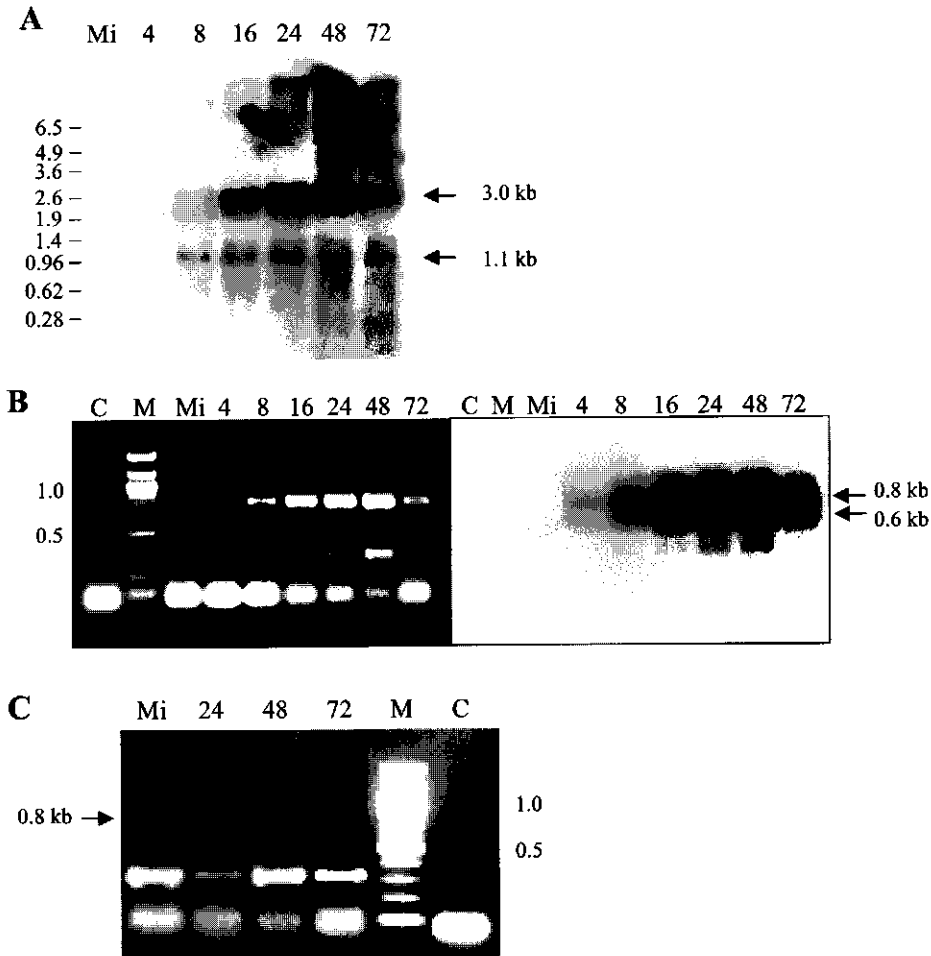**Northern blot and RT-PCR analysis of Se17/18 transcripts**

The temporal regulation of Se17/18 transcripts was examined by Northern blot analysis and RT-PCR, using total RNA isolated from SeMNPV-infected Se301 cells. Northern analysis of Se17/18 revealed a band of 1.1 kb that could represent Se17/18 transcripts (Fig. 4.4A). These transcripts were detected at 4 h p.i., reached maximal transcription levels at 24 h p.i. and reseded until 72 h p.i. The transcript size of 1.1 kb is in good agreement with the predicted ORF size of 870 nt for Se17/18. The size of the Se17/18 transcript suggests that it was most likely polyadenylated, assuming that one of the two putative early promoter sequences were used as transcription start sites. A second transcript of 3 kb was detected from 16 h p.i. onwards. This transcript could be a read through of Se17/18 in the *cathepsin* gene (Se16) and was more abundantly present than the 1.1 kb. Several transcripts larger than 6.6 kb were detected from 24 h p.i. onwards.

Using the more sensitive RT-PCR technique, further information was obtained about the temporal regulation of the 1.1 kb Se17/18 transcripts. A fragment of 790 bp was amplified (Fig. 4.4B) using a primer internal to Se17/18 (Se17/18SP1; Fig. 4.2) and the PCR anchor primer. The obtained RT-PCR fragment of 0.8 kb appeared to be specific for Se17/18 upon hybridization with a Se17/18-specific riboprobe (Fig. 4.4B). However, hybridization showed an additional Se17/18-specific band of 0.65 kb that was not detected on agarose gel. Although several attempts were performed to clone this smaller fragment in pGEM-T this remained unsuccessful. The 0.8 kb Se17/18 fragment was amplified from 8 h till 72 h p.i., with maximum abundancy at 24 h and 48 h p.i. (Fig. 4.4B). Therefore, Se17/18 can be considered as gene that is transcribed from early till late in infection.

To investigate if Se17/18 was also transcribed during SeMNPV infection in *S. exigua* larvae (*in vivo* infection), RT-PCR was performed on RNA isolated from fat body tissue. A RT-PCR product was obtained at 72 h, but not at 24 h and 48 h p.i (Fig. 4.4C). Sequencing of two clones obtained after cloning this RT-PCR product into pGEM-T confirmed that the 0.8 kb RT-PCR product was derived from the Se17/18 sequence. Thus, Se17/18 was transcribed upon infection of cultured insect cells as well as during infection of insect larvae. Furthermore, the obtained RT-PCR fragments indicate that the Se17/18 transcripts contain a poly(A) tail, since the oligo-dT anchor primer was successfully used to synthesize first-strand cDNA.

Transcriptional mapping of the 5' and 3' ends of the Se17/18 transcript

The 5' end of the Se17/18 transcript was determined by 5' RACE analysis using total RNA isolated at 16 h and 48 h p.i. from Se301 cells. A single cDNA was detected at

Fig. 4.4 Temporal expression of the Se17/18 transcript in SeMNPV-infected Se301 cells. (A) Northern analysis of Se17/18. The size of specific hybridization bands is indicated on the right. (B) The left panel shows RT-PCR analysis of Se17/18 performed on total RNA extracted form SeMNPV-infected Se301 cells obtained under the same conditions as in A. The right panel of B shows the hybridization pattern of the obtained RT-PCR bands using a Se17/18 riboprobe. (C) RT-PCR analysis performed on total RNA extracted from SeMNPV-infected *S. exigua* fat body tissue. Times p.i. are indicated above the lanes (Mi, mock infected; C, negative control: PCR without RT on 48 h p.i. sample). Size standards (M, marker) are indicated in kb.

both times tested. Three clones each time point were sequenced. The obtained transcription start sites varied for each timepoint within a small but identical region. The most distal start site of Se17/18 transcription mapped at nt 134, while the most proximal start site is located 126 nt upstream of the ATG translation initiation codon, respectively, at the A in the sequence GGCATT and the second A in the sequence

TCTAAT (Fig. 4.2). An intermediate start site mapped at nt 131 upstream of the putative translation start codon (Fig. 4.2). Thus, the Se17/18 transcripts initiated within or directly after an eight nt region that contains a baculovirus consensus early promoter motif. The most distal start site of Se17/18 overlaps with the ATG start codon of the *chitinase* gene encoded by the complementary DNA strand.

The 3' end of the Se17/18 transcript was determined by sequencing the specific 0.8 kb RACE-PCR fragments obtained after RT amplification of total RNA purified from Se301 cells or *S. exigua* fat body tissue at various times p.i. The obtained sequences indicated that the 3' ends of the transcripts isolated from cultured insect cells are identical to those isolated from fat body tissue and showed no variation in their poly(A) attachment sites. The 3' end of the Se17/18 transcript was located 17 nt downstream of the stop codon at the last A in the sequence TAAT (Fig. 4.2). A convential mammalian polyadenylation signal, consisting of an AATAAA motif and 20-30 nt downstream a diffuse (G)U-rich sequence, is located directly downstream of the stop codon. Hence, the Se17/18 transcript ended 12 nt downstream of the AATAAA motif. These results suggest that the convential mammalian polyadenylation signal was used for the termination of Se17/18 transcripts in both Se301 and *S. exigua* fat body cells. Taken together, the data of the 5' and 3' end mapping predict a minimal transcript size of 1,011 nt for Se17/18 (excluding the poly(A) tail). The expected size is in good agreement with the 1.1 kb size determined by Northern analysis for the putative transcript of Se17/18, assuming a poly(A) tail of approximately 100 nucleotides.

Localization of the Se17/18 protein in insect cells
The subcellular localization of the Se17/18 gene product was investigated with a C-terminal GFP-fusion construct. The GFP-fusion construct was made in plasmid p166BRNX-AcV5 (IJkel *et al.*, 2000). As a negative control, GFP alone was cloned in the same vector. Sf21 and Se301 cells were transfected with 5 µg plasmid DNA, incubated for 48 h at 27°C and examined for fluorescence by confocal laser scanning microscopy. The non-fused GFP protein showed homogeneous fluorescence in the cytoplasm and nucleus (Fig. 4.5A, B). The Se17/18 GFP-fusion protein, however, was mainly localized in the cytoplasm (Fig. 4.5C, D). This distinct pattern of fluorescence is consistent with the computer-predicted cytoplasmatic localization of the Se17/18 protein (Reinhardt and Hubbard, 1998). Notably, the number of fluorescent cells upon transfection with the Se17/18 GFP-fusion construct was considerably lower in Sf21 as well as Se301 cells (~10%) than for the free GFP-control in these cell lines (~30%). Furthermore, the majority of cells transfected with Se17/18 GFP-fusion construct that
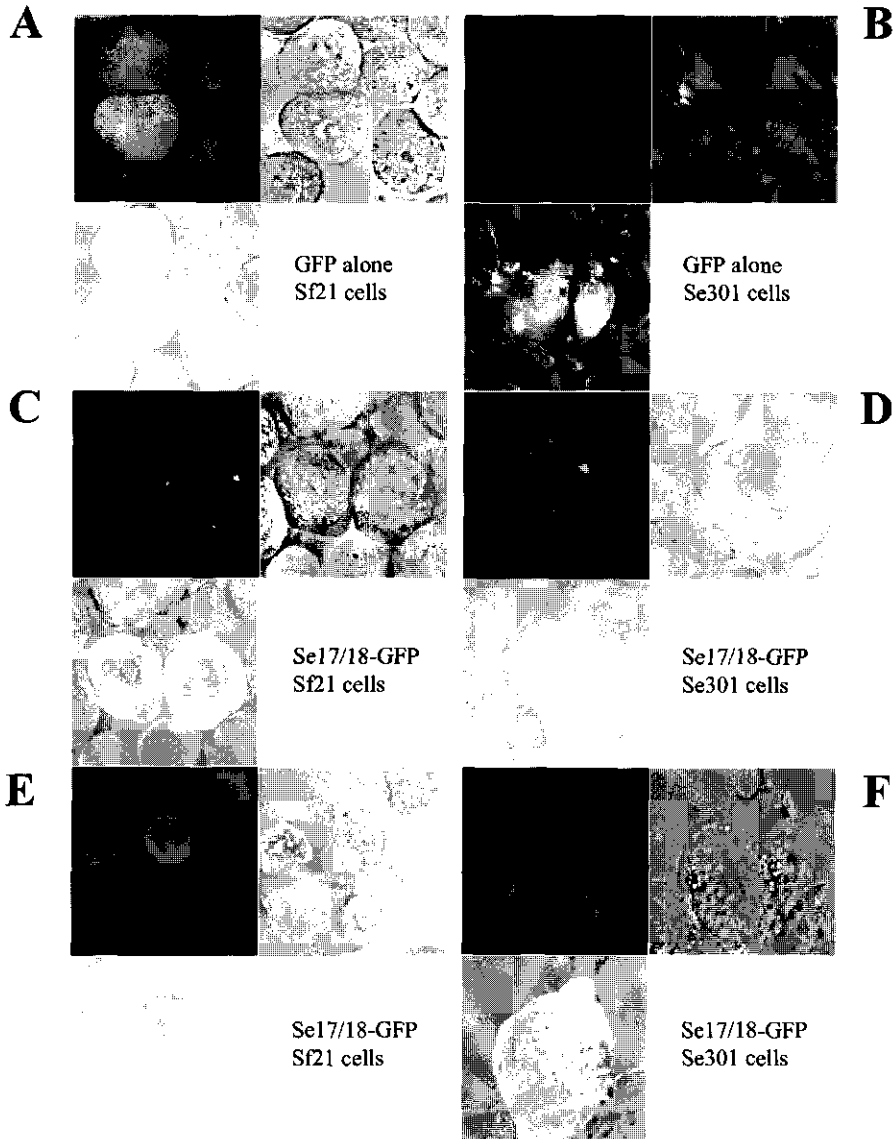
Fig. 4.5 Localization of the Se17/18 GFP-fusion protein in Sf21 and Se301 cells.
Sf21 and Se301 cells were transfected with the control plasmid p166AcV5-GFP (A and B) or with plasmid p166AcV5-Se17/18GFP (C, D, E and F). At 48 h after transfection the cells were examined by confocal laser scanning microscopy for fluorescence. Phase contrast micrographs are shown to the right of the fluorescence graph. Overlay micrographs of the fluorescence and phase contrast micrographs are shown below the fluorescence micrographs.

showed fluorescence were in bad condition, as observed by their shape (Fig. 4.5E). Some of these fluorescent cells even appeared to be apoptotic (Fig. 4.5F) in contrast to fluorescent cells of the negative control (data not shown).

Immunodetection of the Se17/18 protein in infected cells

The Se17/18 protein has a predicted molecular weight of 33.7 kDa. Antibodies were prepared by immunization of chicken with PREP-cell purified Se17/18 protein produced in *E. coli*. The obtained chicken antiserum reacted strongly against this purified Se17/18 protein (Fig. 4.6). Extracts of SeMNPV-infected Se301 insect cells, however, showed no immunoreactive protein when using this antiserum (Fig. 4.6). The absence of an immunoreactive Se17/18 protein was unexpected since Northern and RT-PCR demonstrated that transcripts were present at various times p.i. (Fig. 4.4).
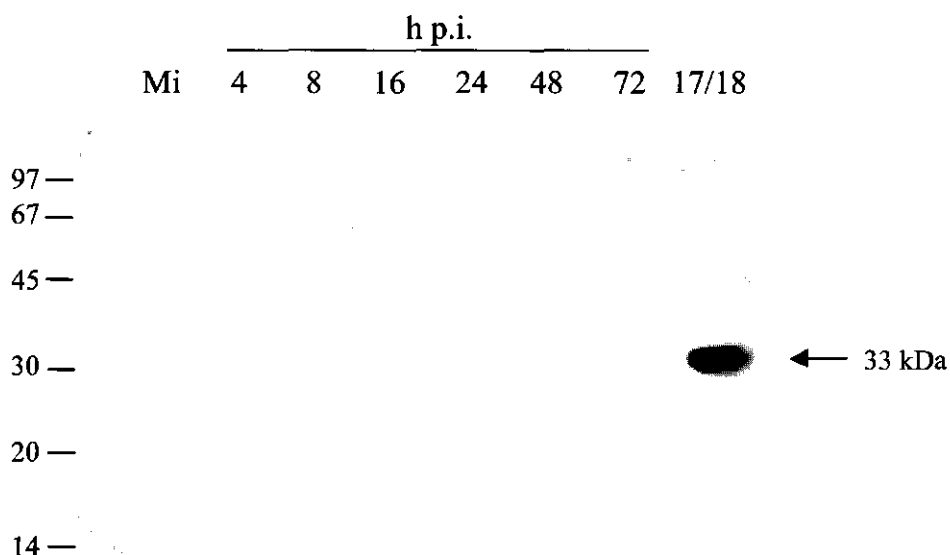
Fig. 4.6 Western blot analysis of SeMNPV infected Se301 cells (70,000 cells/lane) using a polyclonal antiserum against the Se17/18 protein. The corresponding times p.i. are indicated above the lanes (Mi, mock infected; Se17/18, 10 ng of purified Se17/18 protein). The Se17/18 protein was identified using a chicken Se17/18 polyclonal antiserum and detected with a chemiluminescent substrate. Size standards are indicated in kDa and immunoreactive protein is indicated by an arrow.

Similarity of Se17/18 to other (viral) proteins

Since Se17/18 was not detected in SeMNPV-infected Se301 cells and appeared to have a negative effect on insect cell viability when highly expressed, it may be present in spurious amounts and possess a regulatory function. A computer-assisted analysis was performed to elucidate if patterns of residues conserved between Se17/18 and Xc129 were also present in other (viral) proteins. The N-terminal 100 aa of Se17/18 appeared to have low (~25-35%) similarity to various proteins, such as the RNA-directed RNA polymerase (L protein) of Ebola virus [NP_066251], the histone acetyltransferase (Gcn5p) of *Saccaromyces cerevisiae* [NP_011768], the methylated-DNA protein cysteine methyltransferase (cmt) of *Pseudomonas aeruginosa* [AAG04384], the histidine decarboxylase (hdc) of *Morganella morhanii* [AAA25321], the positive transcriptional regulator (Xys2) of *Pseudomonas putida* [Q05092] and the DNA polymerase III α-subunit of *Helicobacter pylori* [P56157] (Fig. 4.7A).

The C-terminal part (aa 200-289) of the Se17/18 protein showed similarity (~25-40%) to the small subunit of the *E. coli* hydrogenase-2 [A55516], the Human papillomavirus type 82 E6 protein [NP_038151] and the Zyx102 protein of *Drosophila melanogaster* [AAF33231] (Fig. 7B). Notably, special conservation was observed for some cysteines in the compared protein sequences (Fig. 4.7B). The C-terminal part of Se17/18 showed also similarity to the LdMNPV ORF53 (Ld53) protein (Kuzio *et al.*, 1999). Ld53 is a homolog of BmNPV ORF41 (Bm41), HaSNPV ORF42 (Ha42), AcMNPV ORF52 and SeMNPV ORF109 (Ayres *et al.*, 1994; IJkel *et al.*, 1999; Kuzio *et al.*, 1999; Gomi *et al.*, 1999; Chen *et al.*, 2001). Strikingly, two of these four homologs, Ac52 and Se109, have smaller C-termini and as a consequence lack the cysteine residues conserved between Ld53 and Se17/18.

Most of the proteins that show (low) homology to the N-terminal as well as to the C-terminal parts of the Se17/18 and Xc129 proteins have two features in common (i) they possess a transferase/reductase activity and (ii) they are involved in the transcription. The middle part (aa 100-200) of the Se17/18 protein also showed low (~20% aa similarity) homology to several proteins with different transferase activities (data not shown).

A

```
Se17/18   : ------LFGLIPMEMKLKILK---YTGHDTYPAVTEDR-VGADTILLTEKQFIK-YFDTR :  49
Xc129     : TNLTQEITNKLPVEMKLYIIE---YLGSDVYYLVTGDE-GRADTLLLQERNYKQ-YFDTR :  55
Lprotein  : LFSIQKHNGHPVLHSETAIQK---VKKHATVLKALRPI-VIFETYCVFKYSIAKHYFDSQ :  56
Gcn5p     : STEQVRGYGAHLMNHLKDYVR---NTSNIKYFL-LY-----ADNYAI--GYFKKQGI-LK :  48
Cmt       : PVGTLTIVARDDAFLVAILWQ---HERPNRVPLDEMRLSEDSSLLAETIRQL-RIYISGI :  56
Hdc       : --LNSFDEEKEVMEYFADLFK---IPFEQSAGYVINGGTBG---NMFGCYLG-REINPDG :  51
Xys2      : EHYTKIHATKLLETLSSNVSRKGFVEGNPCEERVVQFIEEN-VKRSISLEQLAELALMSP :  59
DNApol    : NLVKSIDIEELEIGSQAYLIGKI-VEVKKKIGKRSGKPYGIAD-ILDRYGKE-ELMLFEK :  57
```

```
Se17/18   : RFNMCINVDYRLTEYFK----RQDYVRLPLIVDGLFQTYVKTHL-NGYVMADDVLLKTFF : 104
Xc129     : SEDYCLEKDFKLARPFKNCT-AAQLTRLPMIADRQELTYLMTQP-KGYVINNDTLLKTYF : 113
Lprotein  : GSWMYSVTSDRNLTPGLNSYIKRNQFPPLPMIKELLNEFYHLDHPPLFSTKIISDLSI--F : 114
Gcn5p     : -EITLDKSIWMGYIKDYEGGTLMCCSMUPRLRYLDAGKILLLQEAAL-RRKIRTISKSHI : 106
Cmt       : RSRFELPLDFQGTEPOK--KVWSALLTIPFGETRSYTEIAVQTGSPNAVRAVGAANGRNP : 114
Hdc       : TLYMSKDTHYSVAKIVKLLRIKSLVVESQPNGEIDMDDLMKKIAD--DK-EAHPII---F : 105
Xys2      : RSLYTVFEKHTGITPMNYIPNEKLECVRACLSNPTTNIRSITEVALDYGFLHLGRFAEKV : 119
DNApol    : QLNALEELDINKPLVFK-CKIEEQEEVV-RLR--ILE-ILDLESA-----REVKIPKARM : 107
```

B

```
Se17/18   : TRCFESAYLMLSGEGVNCOEP-SIGEESILKLL-DND--WEDLYVILFCSKCTYCI
Xc129     : TESIDGEHLKVOGECVNCGET-SISASTAMPLV-DFLC-WFNLRVIVFCSNCASCL
Hydro2    : ---EENGVAYIKKQCMHCVDPNCVSVCPVSALKKDPKTGIVHYDKDV-CTGCRYCM
Ld53      : VYILNAARGDNELVCVQCQRL--IDRRACEQFH-DAPCQDFN--VNLFCSHCLFPL
Zyx102    : CTAMDQIMHIFCFTCTDCQIN--LQGKPFYALDGKPYCEYDYLQTLEKCSVCMEPI
E6        : AITNKSLMELL-IRCHRCCRPLGPEEKQKVVDDKKRFH-EIAGRWTGQCAMCRKP-
```

Fig. 4.7 Amino acid homology of (A) the N-terminal 100 amino acids of Se17/18 with Xc129, the RNA-directed RNA polymerase (L protein) of Ebola virus [NP_066251], the histone acetyltransferase Gcn5p of *Saccaromyces cerevisiae* [NP_011768], the methylated-DNA protein cysteine methyltransferase (cmt) of *Pseudomonas aeruginosa* [AAG04384], the histidine decarboxylase (hdc) of *Morganella morhanii* [AAA25321], the positive transcriptional regulator (Xys2) of *Pseudomonas putida* [Q05092] and the DNA polymerase III α-subunit of *Helicobacter pylori* [P56157]. (B) Amino acid relationship of the C-terminal 89 amino acids of Se17/18 with Xc129, the small subunit of the *E. coli* hydrogenase-2 [A55516], the Human papillomavirus type 82 E6 protein [NP_038151], the Zyx102 protein of *Drosophila melanogaster* [AAF33231] and the LdMNPV ORF53 protein. Gaps, introduced to optimize the alignment, are indicated with dashes. Shading is used to indicate the occurrence of identical (black) or substitutional (grey) amino acids.

Discussion

In this study we report the temporal transcription, translation and subcellular localization patterns of Se17/18. Northern analysis showed that the amount of the 1.1 kb Se17/18 transcripts increased from 4 h until 24 h p.i. and declined till 72 h p.i. (Fig. 4.4). A similar transcription pattern was previously observed for other baculovirus genes, such as the AcMNPV *pnk/pnl* and *lef4* genes, and the BmNPV *bro* genes (Durantel *et al.*, 1998a and b; Kang *et al.*, 1999).

Hybridization of the RT-PCR products revealed a second smaller band that reacted specific to the Se17/18 riboprobe (Fig. 4.4B). This smaller product, and most likely favourable in PCR amplification, was present in much lower amounts than the anticipated 0.8 kb RT-PCR product. Though, it is likely that the smaller RT-PCR product is a PCR artefact, it can not be excluded that it represents a 3' truncated Se17/18 transcript. Cloning and subsequent sequencing of the smaller RT-PCR product would reveal the nature of its origin.

Mapping of Se17/18 (Fig. 4.2) transcripts by 5' RACE showed that its transcription initited in a region rather than at one specific nucleotide. This could either resemble the natural variation in Se17/18 transcription start sites or be caused by RNA degradation at the 5' end of the transcripts. Another explanation for the variation in the Se17/18 transcription start site could lie in reduced or 'slippery' recognition of the the GGCATT motif since it is not completely consistent with the arthropod initiator cap site consensus [A(A/C/T)CA(G/T)T]) (Cherbas and Cherbas, 1993).

The spacing of only 1 nt between the upstream TATA-box and the GGCATT sequence (Fig. 4.2), is very unusual (Roeder, 1991; Lu and Miller, 1995). Therefore, it is unlikely that this TATA-box is functional. Since another TATA-box is not present in the 100 nt upstream of the Se17/18 transcription start site, its transcription could be TATA-independent. Previously, TATA-independent transcription was reported for the OpMNPV *gp64* gene (Kogan *et al.*, 1995). Four elements, the host factor binding sites GATA and CACGTG, the transcription start site CA(G/T)T and a CA-rich sequence in the 5' untranslated region (UTR) were found to be required for TATA-independent transcription of *gp64*. Three of these, the transcription start site, the GATA sequence and CA-rich sequence are also present in the 5' UTR of Se17/18, supporting the assumption that Se17/18 transcription is TATA-independent and may involve host factors.

Comparison of the promoter regions of Se17/18 and Xc129 revealed no apparent nucleotide conservation at the determined Se17/18 transcriptional start sites (Fig. 4.3). Furthermore, the spacing of 20-25 nt between the TATA-box and the two CATT

motifs in the Xc129 promoter region (Fig. 4.3) is in good agreement with the spacings detected in other baculovirus early promoters that are active (Cherbas and Cherbas, 1993). Therefore, Xc129 transcription initiation would be possible from this putative proximal motif, while the unusual spacing in Se17/18 apparantly leads to the use of more distal transcription start sites.

By Western analysis no Se17/18 translation product could be detected in SeMNPV-infected Se301 cells (Fig. 4.6). This is somewhat unexpected since Se17/18-specific transcripts could be detected as early as 4 h p.i. in Se301 cells (Fig. 4.4). The inability to detect the Se17/18 protein may be due to the low affinity of the Se17/18 antiserum and/or the low steady state levels of Se17/18 protein during infection. The latter is more or less supported by the lower percentage of insect cells that, upon transfection with a construct resulting in high Se17/18-GFP protein expression, showed fluorescence. Combined with the high mortality of cells transfected with Se17/18-GFP this may suggest that the Se17/18(-GFP) protein has a toxic effect on insect cells. The observation that a mutant SeMNPV, obtained within the first passage in insect cell culture (Heldens *et al.*, 1996), lacks Se17/18 among others suggests that it is non-essential in cell culture and may require a co-factor to become non-toxic.

The genomic region of SeMNPV, where Se17/18 is located - between the gene homologs of *cathepsin* and *chitinase*- differs in many baculoviruses. Besides Se17/18, SeMNPV contains another gene, Se21, with a homolog only in XcGV (ORF128; Hayakawa *et al.*, 1999). So, either SeMNPV inserted, between Se17/18 and Se21, Se19 (*chitinase*) and Se20 or these sequences were deleted in the XcGV genome assumming an identical sequence source (Fig. 4.8).
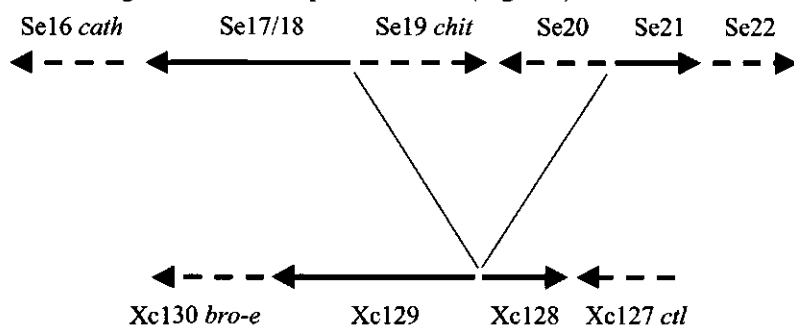
Fig. 4.8 Comparison of the Se17/18 - Se21 gene region of SeMNPV to the Xc128 - Xc129 gene region of XcGV. Arrows indicate the orientation and location of the ORFs. The name of the ORF is given above the arrow. Se17/18 and Se21 have corresponding homologs in Xc129 and Xc128, respectively. Dashed arrows indicate that no corresponding homologs are present in this genome region between SeMNPV and XcGV.

The sequence similarity (Fig. 4.7) of Se17/18 to proteins that all have transferase /reductase activity and of which some are known to be transcription regulators suggests that Se17/18 may have a function in transcription activation. A possible function of Se17/18 may be that it is involved in regulation of transcription by catalyzing the removal of methyl groups from methylated DNA sequences. If Se17/18 plays a role in transcription regulation it might also explain the toxic effect on insect cells since its high expression level could distort the balance in host transcription. However, the cytoplasmic localization of the Se17/18 protein does not support a direct role in transcription regulation, which is likely to take place in the nucleus of infected insect cells.

It is hypothesized, based on homology to the *E. coli hydrogenase* and *ferredoxin* genes, that the conserved cysteines in the C-terminal part of the Se17/18 protein (Fig. 4.7B) may form an iron-sulphur (Fe-S) cluster (Menon *et al.*, 1994; Sargent *et al.*, 1998). Fe-S clusters are cofactors found in many proteins that have important redox, catalytic or regulatory functions (Beinert and Holm, 1997). Although most Fe-S proteins are found in the mitochondria, some are present in the cytosol (Tong and Rouault, 2000).

The observation that the conserved cysteine residues are present in Ld53, Bm41 and Ha42 but not in their Ac52 and Se109 homologs together with the absence of Se17/18 as well as Se109 homologs in OpMNPV (Ahrens *et al.*, 1997) and PxGV (Hashimoto *et al.*, 2000) suggests that the function of Se17/18 is either non-essential in OpMNPV and PxGV or can be substituted by other baculovirus genes. As the Se17/18 protein is highly homologous only (so far) to the Xc129 protein, they may play a role in specific virus host interactions. Future studies in SeMNPV infected larvae as well as site-specific and null mutants will determine the significance of this gene in the SeMNPV infection cycle and may shed light on its function.

Materials and Methods

Computer-assisted analysis
Se17/18 (IJkel *et al.*, 1999) and Xc129 (Hayakawa *et al.*, 1999) were analyzed using software of the Predict Protein server (Rost, 1996) and the ExPASy server (Appel *et al.*, 1994) for the prediction of domains, motifs, transmembrane regions and subcellular localization (Reinhardt and Hubbard, 1998). DNA and protein comparisons with entries in the updated GenBank/EMBL, SWISS-PROT and PIR databases were performed with BLASTn, FASTA and Psi-BLAST programs

(Pearson, 1990; Altschul *et al.*, 1997). Multiple sequence alignments were performed with the GCG PileUp computer programs with gap creation and extension penalty set to 8 and 2, respectively (Devereux *et al.*, 1984). Alignment editing was performed with Genedoc Software.

Plasmid constructions

The complete coding region of Se17/18 including the stop codon was amplified by high fidelity "Expand" long template PCR (Boehringer Mannheim) from cosmid 17 (Heldens *et al.*, 1996) using primers containing 5' *Bam*HI and 3' *Hind*III restriction sites and cloned into pGEM-T. The plasmid was named pGEMSe17/18 and used for production of a gene-specific riboprobe. The plasmid pTriExSe17/18 was obtained by cloning the *Bam*HI/*Hind*III fragment from pGEMSe17/18 into the *Bam*HI and *Hind*III sites of the expression vector pTriEx-1 (Novagen). This plasmid was used for overexpression of the Se17/18 protein in *E. coli*.

To determine the localization of Se17/18 protein in insect cells, a GFP-fusion construct was made. The complete coding region of Se17/18 was amplified by high fidelity "Expand" long template PCR (Boehringer Mannheim) from cosmid 17 (Heldens *et al.*, 1996) using primers containing 5' *Bam*HI and 3' *Eco*RI restriction sites and cloned into the *Bam*HI and *Eco*RI sites of the previously described p166AcV5-GFP vector (IJkel *et al.*, 2000) and named p166Se17/18-GFP. Plasmid DNA was purified using Jetstar columns according to manufacturer's protocol (ITK Diagnostics). For each construct, the nucleotide sequence was checked using an automated DNA sequencer (Wageningen University, The Netherlands).

Cells, Insects and Viruses

The *Spodoptera exigua* cell-line Se301 (Hara *et al.*, 1995b), the *S. frugiperda* cell-line IPLB-Sf-21 (Vaughn *et al.*, 1977) and the SeMNPV-US1 isolate (Gelernter and Federici, 1986b; Muñoz *et al.*, 1998) were maintained and propagated as described previously (IJkel *et al.*, 2000). A culture of *S. exigua* insects was maintained according to Smits and Vlak (1988a).

Total RNA isolation, Northern blot, RT-PCR, 3' and 5' RACE analysis

Total RNA was isolated from 2×10^6 mock-infected and SeMNPV-US1-infected Se301 cells (MOI of 5 $TCID_{50}$ units/cell) at 4, 8, 16, 24, 48 and 72 h p.i. Total RNA was also isolated from fat body tissue obtained after dissection of six mock-infected and six SeMNPV-US1 infected *S. exigua* larvae (1.7×10^4 polyhedra/larva) at 24, 48 and 72 h p.i. Cells and tissue were resuspended in 500 μl Trizol (GibcoBRL) and 100

µl chloroform, incubated for 8 min and centrifuged at 14,000 *g* for 15 min at 4°C. The RNA in the water fraction was precipitated using isopropanol, centrifuged at 14,000 *g* for 10 min at 4°C, washed with 70% ethanol and resuspended in 50 µl water. The RNA solutions were incubated at 55°C for 10 min and quantified by absorbance at 260 nm.

For Northern analysis 8 µg of total RNA was electrophoresed in 1.5% agarose in the presence of glyoxal (Ausubel *et al.*, 1994) and blotted onto a Hybond N nylon membrane (Amersham). The northern blot was hybridized with a Se17/18 messenger-specific riboprobe generated by SP6 polymerase with [α-32P]CTP and *Bam*HI digested pGEMSe17/18. Fragment sizes were determined by staining the molecular weight marker (Promega RNA marker) with methylene blue after transfer onto the membrane.

RT-PCR was performed using the 5'/3' RACE kit (Roche) employing 2 µg purified total RNA as template per time point. First-strand cDNA synthesis was performed using AMV reverse transcriptase and the oligo-dT anchor primer according to manufacturer's instructions. The cDNA-mixtures were amplified by PCR using the PCR anchor primer and the gene specific primer Se17/18SP1 (Fig. 4.3). The obtained PCR-products were analyzed in 1.2% agarose gels.

The amplified RT-PCR products of 16 h and 24 h p.i. were used to determine the 3' end of the Se17/18 *in vitro* messenger, while the RT-PCR product of 72 h p.i. was used for the 3' ends of the Se17/18 transcripts isolated from fat body tissue. All PCR-products were gel purified, cloned into pGEM-T and sequenced with T7 or SP6 primers.

The 5' end of the Se17/18 transcript was determined using the 5'/3' RACE kit (Roche) employing 2 µg purified total RNA as template per time point, isolated at 16 h and 48 h p.i. Briefly, first strand cDNA synthesis was performed with the gene specific primer Se17/18SP2 (Fig. 4.3). The cDNAs was purified with the High Pure PCR purification kit (Roche) and a poly(A) tail was added to the 3' end using the terminal transferase with dATP. The tailed cDNAs were amplified by PCR using the oligo dT-anchor primer and the nested gene specific primer Se17/18SP3 (Fig. 4.3). A second PCR was performed using the PCR anchor primer and the nested primer Se17/18SP4 (Fig. 4.3). The obtained PCR products were gel purified, cloned into pGEM-T and sequenced with T7 or SP6 primers.

Production of polyclonal antibodies

A culture of *E. coli* Bl21 containing pTriExSe17/18 was grown to an optical density at 600 nm of 0.5 and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG).

After 4 h at 37°C, cells were collected by centrifuging at 6,500 g for 15 min at 4°C, resuspended in 20mM Tris-HCl pH7.5, lysed with lysozyme, sonicated and centrifuged at 14,000 g for 10 min at 4°C. Aliquots of the insoluble fractions were washed with 20 mM Tris-HCl pH7.5 and resuspended in 1% SDS. Soluble and insoluble fractions were electrophoresed in 12.5% SDS polyacrylamide gels according to Laemmli (1970) and stained with Coomassie brilliant blue.

The insoluble fractions containing the Se17/18 protein was purified by continuous-elution electrophoresis using the Model 491 Prep Cell (Bio-Rad) according to manufacturer's protocol. Elution fractions were collected, electrophoresed in 12.5% SDS polyacrylamide gels and the protein bands visualized by silver staining (Morrisey, 1981). Fractions containing Se17/18 were pooled and concentrated using Centriprep-10 kDa filter devices (Amicon). Protein concentrations were determined with the Bio-Rad protein assay (Bio-Rad).

Two chickens were injected intramuscularly each with 12.5 µg purified Se17/18 protein using a water in oil adjuvant. The chickens were boosted after 6 weeks with 25 µg purified protein. Eggs were collected every day for 4 weeks and serum was collected 12 weeks after the boost injection. Western analysis using *E. coli* BL21 extracts expressing pTriExSe17/18 and PREP-cell purified Se17/18 were used to test the production of specific antisera.

Western analysis

Monolayers of Se301 cells were mock- or SeMNPV-US1-infected at a multiplicity of infection (MOI) of 5 $TCID_{50}$ units/cell. Cells were harvested at 4, 8, 16, 24, 48 and 72 h p.i., pelleted, resuspended in PBS and lysed in SDS-PAGE loading buffer by boiling for 5 min. Protein samples were then separated by SDS-PAGE and transferred to an Immobilon-P nitrocellulose membrane (Millipore) by semi-dry electrophoresis transfer (Ausabel *et al.*, 1994). The membranes were incubated overnight in 10% block solution (Boehringer Mannheim) in PBS buffer at 4°C. The membranes were allowed to react in PBS with Se17/18 antiserum diluted 1:5000 for 1 h at room temperature (RT). After washing in PBS (three times 15 min), the membranes were incubated for 1 h at RT with horse radish peroxidase-conjugated rabbit anti-chicken immunoglobulin (Sigma) diluted 1:50,000 in PBS. After washing in PBS (three times 15 min) the signal was detected by ECL technology as described by the manufacturer (Amersham).

Fluorescence microscopy

Se301 cells (3 x 10^5) and Sf21 cells (1 x 10^5) were grown on glass cover slips and transfected with 5 µg of plasmid DNA using Cellfectin (GibcoBRL). At 48 h post transfection the cells were examined with a Zeiss LSM510 (confocal) laser scanning microscope for fluorescence using an excitation wave length of 488 nm and an emission band pass filter of 505 - 530 nm.

Chapter 5

A novel baculovirus envelope fusion protein with a proprotein convertase cleavage site

The entry mechanism of *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV), a group II NPV, in cultured cells was examined. SeMNPV budded virus (BV) enters by endocytosis as do the BVs of the group I NPVs, *Autographa californica* (Ac) MNPV and *Orgyia pseudotsugata* (Op) MNPV. In group I NPVs, upon infection acidification of the endosome triggers fusion of the viral and endosomal membrane, which is mediated by the BV envelope glycoprotein GP64. However, the SeMNPV genome lacks a homolog of GP64 envelope fusion protein (EFP). A functional homolog of the OpMNPV GP64 EFP was identified in SeMNPV ORF8 (Se8; 76 kDa) and appeared to be the major BV envelope protein. Surprisingly, a 60-kDa cleavage product of this protein is present in the BV envelope. A furin-like proprotein convertase cleavage site (R-X-K/R-R) was identified immediately upstream of the N-terminus of the mature Se8 protein and this site was also conserved in the *Lymantria dispar* (Ld) MNPV homolog (Ld130) of Se8. Syncytium formation assays showed that Se8 and Ld130 alone were sufficient to mediate membrane fusion upon acidification of the medium. Furthermore, C-terminal GFP-fusion proteins of Se8 and Ld130 were primarily localized in the plasma membrane of insect cells. This is consistent with their fusogenic activity and supports the conclusion that the Se8 gene product is a functional homolog of the GP64 EFP.

Introduction

The *Baculoviridae* are a family of large, enveloped double-stranded DNA (80 to 180 kbp) viruses that almost exclusively infect insects (Adams and McClintock, 1991). The family is taxonomically subdivided into two genera, *Nucleopolyhedrovirus* (NPV) and *Granulovirus* (GV), distinguished by occlusion body (OB) morphology (Volkman *et al.*, 1995). The NPVs produce large polyhedron-shaped OBs called polyhedra that contain many virions, whereas the GVs have smaller OBs called granules that normally contain a single virion. The NPVs are designated as single (S) or multiple (M) depending on the potential number of nucleocapsids (NC) packaged in a virion. A more recent subdivision, based on the phylogenetics of Polyhedrin / Granulin (Zanotto *et al.*, 1993), EGT (Chen *et al.*, 1997), LEF-2 (Chen *et al.*, 1999) and DNA polymerase (Bulach *et al.*, 1999) proteins, has been proposed that distinguishes lepidopteran NPVs into two distinct groups, named groups I and II. The baculoviruses AcMNPV (Ayres *et al.*, 1994), *Bombyx mori* (Bm) NPV (Gomi *et al.*, 1999) and OpMNPV (Ahrens *et al.*, 1997) are members of group I, whereas SeMNPV (IJkel *et al.*, 1999) and LdMNPV (Kuzio *et al.*, 1999) belong to group II.

A NPV infection starts with the uptake of polyhedra by the insect larvae. Upon ingestion the OBs dissolve in the alkaline environment of the larval midgut, liberating numerous infectious virions, which are termed "occlusion derived virus" (ODV). After direct membrane fusion of ODVs with the midgut columnar epithelial cells (Granados, 1978; Granados and Lawler, 1981; Horton and Burand, 1993), the virions are uncoated and transported to the nucleus, where gene expression, DNA replication and assembly of progeny NCs occur. Newly assembled NCs then migrate from the nucleus towards the plasma membrane. During infection of group I NPVs, a viral encoded major envelope glycoprotein, GP64, is synthesized and transported to the plasma membrane (Volkman and Goldsmith, 1984; Volkman *et al.*, 1984; Blissard and Rohrmann, 1989; Oomens *et al.*, 1995). When the NCs bud from the basal side of the epithelial cells into the hemocoel, they acquire a loosely adhering plasma membrane envelope containing the GP64 protein. This protein is required for efficient budding (Monsma *et al.*, 1996; Oomens and Blissard, 1999). This second virion phenotype is termed "budded virus" (BV) and is believed to be essential for systemic infection, mediating movement of the virus from midgut to other tissues and propagating the infection from cell to cell within the infected animal (Flipsen, 1995; Keddie *et al.*, 1989; Granados and Lawler, 1981). BVs of AcMNPV and OpMNPV belonging to group I NPVs, infect insect cells, other than midgut epithelial cells, through the endocytotic pathway (Volkman and Goldsmith, 1985). After BV binding to the cell membrane and uptake into an endosome (Hefferon *et al.*, 1999) the

acidification of the endosome triggers GP64-mediated fusion of the viral and endosomal membrane (Blissard and Wenz, 1992; Kingsley *et al.*, 1999; Plonsky *et al.*, 1999). Then the nucleocapsids are released into the cytoplasm and are transported to the nucleus, where viral transcription and DNA replication occur.

Recently, the complete genome of SeMNPV has been sequenced and its genome analyzed (IJkel *et al.*, 1999). Surprisingly, SeMNPV lacks a homolog of the AcMNPV *gp64* BV envelope fusion protein (EFP) gene. A similar situation exists for LdMNPV (Kuzio *et al.*, 1999) and HaSNPV (W. F. J. IJkel and J. M. Vlak, unpublished data). Since, the GP64 protein of AcMNPV and OpMNPV plays an essential role in spreading the infection in the insect and is required for efficient virus budding, we asked whether SeMNPV infects insects by endocytosis and has a functional GP64 homolog. In the current study, the mechanism of entry of SeMNPV, a group II NPV, is examined in its target cells. A functional homolog of the OpMNPV GP64 EFP is identified in the SeMNPV ORF8 (Se8) protein, and was shown to be the major BV envelope protein. In addition, evidence is provided to demonstrate that the Se8 protein is present in the BV envelope as a cleavage product.

Results

Budded virions of SeMNPV enter insect cells by endocytosis

BVs of AcMNPV and OpMNPV, group I NPVs, enter host cells by endocytosis (Volkman and Goldsmith, 1985; Wang *et al.*, 1997; Hefferon *et al.*, 1999). To examine the mechanism of entry by BVs of SeMNPV, a member of the group II NPVs, we used the lysosomotrophic reagent, ammonium chloride, to demonstrate that SeMNPV BV enters cells by endocytosis. This lipophilic amine buffers the endosomal pH and inhibits acid-triggered membrane fusion in the endosome during viral entry by endocytosis (Helenius *et al.*, 1982; Lenard and Miller, 1982).

Se301 cells were infected with SeMNPV BVs in the presence of a final concentration of 50 mM ammonium chloride. As positive control, Sf21 and Se301 cells were infected with AcMNPV BVs. Polyhedra production was used as an indicator of infectivity. The number of cells that produced polyhedra was counted at 72 h post infection (p.i.) (Table 5.1). When cells were infected in the presence of ammonium chloride, a severe reduction (>95 %) in the number of cells containing polyhedra was observed. Control experiments performed in parallel showed that ammonium chloride did not affect polyhedra formation when ammonium chloride was added up to 6 h p.i. These results provide strong evidence that BVs of SeMNPV enter

insect cells primarily via endocytosis, like the BVs of group I NPVs.

Table 5.1 Effect of ammonium chloride as endosome acidification preventive on the percentage of infected cells as determined by polyhedra production.[a]

	SeMNPV		AcMNPV	
	0 mM	50 mM	0 mM	50 mM
Se301	75	1	78	1
Sf21	0	0	98	2

[a] The experiments were performed in triplicate, with 200 cells scored per analysis.

Membrane fusion of SeMNPV BV is triggered by acidification
Previous studies of AcMNPV and OpMNPV infection showed that acidification triggers fusion of the viral and cellular membranes, and that GP64 is the acid-triggered membrane fusion protein (Volkman and Goldsmith, 1985; Blissard and Wenz, 1992). To determine whether fusion of the SeMNPV BV envelope is triggered by acidification, Sf21 and Se301 cells were infected with SeMNPV and AcMNPV and syncytium formation assays were used as a direct measure of the acid-triggered fusogenic activity of the BVs.

Cell-to-cell fusion of both Sf21 and Se301 cells infected with either SeMNPV or AcMNPV was observed only after the pH of the tissue culture medium was lowered to pH 5 (Fig. 5.1). If the acidic treatment was excluded from the assay, no cell-to-cell fusion was observed (results not shown). Cell-to-cell fusion of Se301 or Sf21 cells induced upon infection with AcMNPV BVs showed no detectable differences from those infected with SeMNPV (Fig. 5.1). SeMNPV BVs were thus capable of inducing pH-dependent membrane fusion in susceptible insect cells.

Identification of the SeMNPV major envelope protein
The previous experiments demonstrated that representative viruses of group I and group II NPVs enter host cells by endocytosis, and that acidification triggers membrane fusion. In AcMNPV and OpMNPV, fusion is mediated by the viral encoded GP64/67 protein, which is the major envelope protein of these BVs (Blissard and Wenz, 1992; Hohmann and Faulkner, 1983; Carstens *et al.*, 1979). However, SeMNPV lacks a GP64 homolog (IJkel *et al.*, 1999). To identify the envelope proteins of SeMNPV, BVs were isolated from hemolymph derived from SeMNPV-infected fourth-instar *S. exigua* larvae. Sucrose gradient purified BVs were separated on SDS-PAGE and the proteins were visualized by Coomassie brilliant blue (Fig. 5.2). A
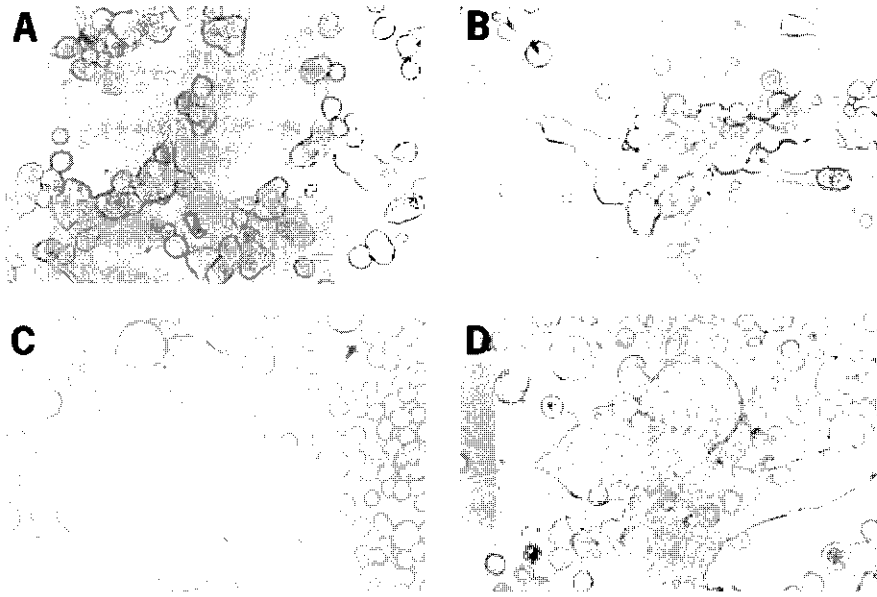
Fig. 5.1 Baculovirus-mediated pH-dependent membrane fusion of insect cells.
Se301 (A, B) and Sf21 cells (C, D) were infected at a m.o.i. of 10 with either SeMNPV (A, C) or AcMNPV (B, D). Forty-eight hours after infection, cells were treated for 2 min with Grace's medium, pH 5.0. Syncytium formation was scored 4 h after dropping the pH by phase-contrast microscopy.

number of proteins were detected in BV preparations (Fig. 5.2 lane 2), which were not present in a sample from mock-infected insects (Fig. 5.2 lane 1), and are therefore related to proteins of the SeMNPV virions. Proteins appearing in the mock-infected sample are probably derived from the insect hemolymph and copurify with the BVs.

The major protein band present in the SeMNPV BV preparation had a size of approximately 60 kDa. This is approximately 4-7 kDa smaller in size than the AcMNPV and OpMNPV major envelope proteins, respectively (Blissard and Rohrmann, 1989; Whitford *et al.*, 1989; Carstens *et al.*, 1979). To investigate whether this protein is present in the envelope or the NC, purified BVs were treated with a non-ionic detergent (NP40), fractionated and analyzed by SDS-PAGE. A 60-kDa protein was the major component of the SeMNPV BV envelope fraction (Fig. 5.2 lane 3). A 60-kDa protein was also present in the NC fraction (Fig. 5.2 lane 4) and this probably resulted from incomplete separation of the envelope and the NCs.

To determine whether the major 60-kDa protein of SeMNPV BV was also present in ODV, ODVs were purified from isolated polyhedra and its proteins separated by SDS-PAGE. A major protein band of 60 kDa was not detected in ODV (Fig. 5.2 lane 5) although conclusive evidence of the possible presence of lower
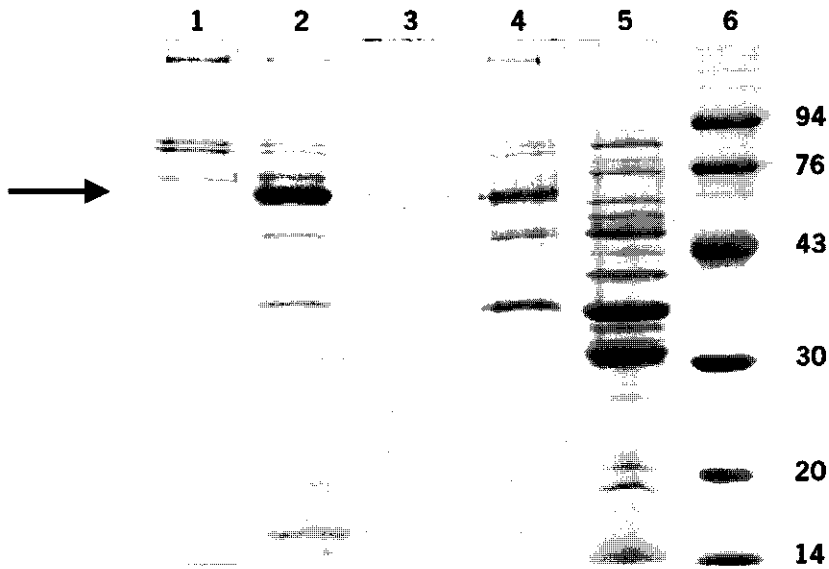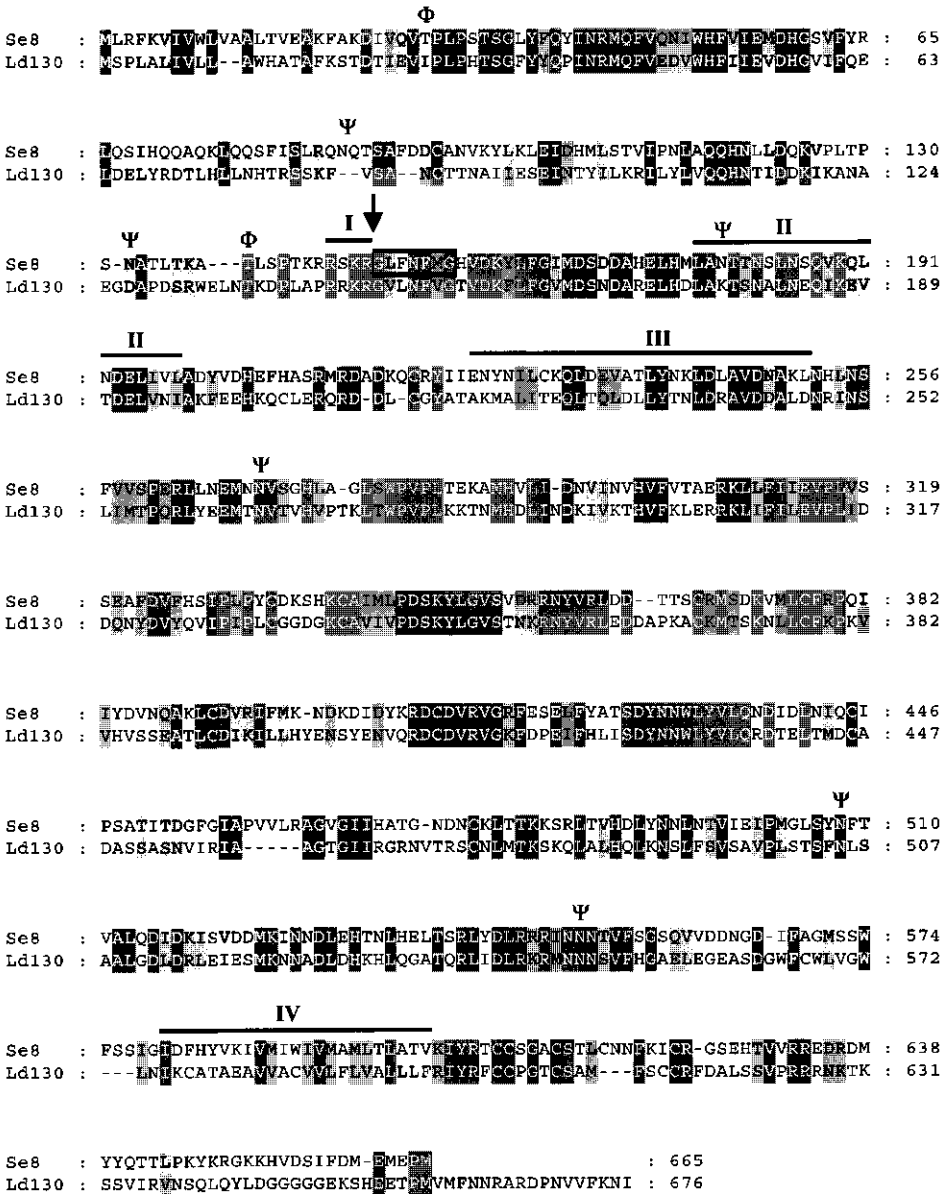
Fig. 5.2 Coomassie brilliant blue-stained SDS-PAGE gel (12%) of purified SeMNPV. Lane 1: Mock purification of BVs from hemolymph of uninfected *S. exigua* larvae. Lane 2: SeMNPV BVs. Lane 3: SeMNPV BV envelopes. Lane 4: SeMNPV BV nucleocapsids. Lane 5: SeMNPV ODVs. Lane 6: Low-molecular-weight marker. The major BV envelope protein (60 kDa) is indicated with an arrow.

quantities of this protein in ODV will require analysis with an antibody directed against the 60-kDa protein. However, it was concluded that SeMNPV BV contains a major envelope protein with an apparent molecular weight of 60 kDa that does not appear to be present in the ODV.

The BV major envelope protein is a proprotein convertase cleavage product
To determine the SeMNPV ORF encoding the BV major envelope protein, BVs were purified and its proteins were separated by SDS-PAGE. The proteins were blotted onto a PVDF membrane and the major 60-kDa envelope protein was N-terminally sequenced. The sequence obtained, GLFNFMG, matched that derived from translation of the Se8 ORF but was located 150 amino acids downstream of the putative N-terminus (Fig. 5.3). This strongly suggested that the product of Se8 was present in the form of a cleavage product in BVs. The C-terminal cleavage product has a predicted molecular weight of 59 kDa, which is in close agreement with the observed size of the BV major envelope protein in SDS-PAGE. The uncleaved Se8 protein has a predicted

```
                                  Φ
Se8   : MLRFKVIVWLVAALTVEAKFAKDIVQVTPLPSTSGLYFQYINRMQFVQNIWHEVIEMDHGSVFYR :  65
Ld130 : MSPLALIVLL--AWHATAFKSTDTIEVIPLPHTSGFYYQPINRMQFVEDVWHFIIEVDHGVIFQE :  63


                     Ψ
Se8   : MQSIHQQAQKIQQSFISLRQNQTSAFDDCANVKYLKLEIDHMLSTVIPNIAQQHNLLDQKVPLTP : 130
Ld130 : IDELYRDTLHLLNHTRSSKF--VSA--NCTTNAIIESEINTYILKRILYLVQQHNTIDDKIKANA : 124


            Ψ         Φ          I  ↓                          Ψ    II
Se8   : S-NATLTKA---ILSETKRESKRELFNRAGHVDKYHFGIMDSDDAHELHMLANHTNSLNSFVKQL : 191
Ld130 : EGDAPDSRWELNIKDELAPERRKREVLNFVGTIDKQLFGVMDSNDARELHDLAKTSNALNEQIKEV : 189


         II                              III
Se8   : NDELIVIADYVDHEFHASRMRDADKQGRHIIENYNILCKQLDEVATLYNKLDLAVDNAKLNHLNS : 256
Ld130 : TDELVNIAKFEEHKQCLERQRD-DL-CGHATAKMALITEQLTQLDLLYTNLDRAVDDALDNRINS : 252


                    Ψ
Se8   : FVVSPERILNEMNNVSGHLA-GLQATVFETEKARHVFI-DNVINVHVFVTAERKLLFIIIAVIVS : 319
Ld130 : LIMTPQRLYEEMTNVTVEVPTKLTWDVPEKKTNIHDLINDKIVKTHVFKLERRKLIFILEVPIID : 317


Se8   : SEAFDVFHSIPLFYGDKSHKCAIMLPDSKYLGVSVDRRNYVRLDD--TTSQRMSDKVMIGFRPQI : 382
Ld130 : DQNYDVYQVIPILLGGDGKCAVIVPDSKYLGVSTNKRNFVRLEDDAPKACKRISKNLICFKPKV : 382


Se8   : IYDVNQAKICDVRIFMK-NDKDIDYKRDCDVRVGRFESELFYATSDYNNWLRVVIQNDIDFNIQCI : 446
Ld130 : VHVSSEATLCDIKILLHYENSYENVQRDCDVRVGKFDPEIEHLISDYNNWIYVIQRDTEITMDCA : 447


                                                             Ψ
Se8   : PSATITDGPGIAPVVLRAGVGIIHATG-NDNCKETTKKSRITVHDLYRNLNIVIEIPMGLSVNFT : 510
Ld130 : DASSASNVIRIA-----AGTGIIRGRNVTRSQNLMTKSKQLALHQLKNSLFSVSAVPLSTSFNLS : 507


                                            Ψ
Se8   : VALQDIDKISVDDMKINNDLEHTNLHELTSRIYDLRNRIINNNIVFSGSQVVDDNGD-IFAGMSSW : 574
Ld130 : AALGDLDRLEIESMKNNADLDHKHLQGATQRLIDLRKRNNNSVPHGAELEGEASDGWPCWLVGW : 572


              IV
Se8   : FSSIGIDFHYVKIVMIWIVMAMLTIATVKIYRTCGSGACSTLCNNGKICR-GSEHTVVRREDRDM : 638
Ld130 : ---LNIKCATAEAWACVVLFLVALLLFRIYRFCCPGTCSAM---ESCCRFDALSSVPRERNKTK : 631


Se8   : YYQTTLPKYKRGKKHVDSIFDM-EMEPM : 665
Ld130 : SSVIRVNSQLQYLDGGGGGEKSHEETEMVMFNNRARDPNVVFKNI : 676
```

Fig. 5.3 Alignment of the predicted amino acid sequences of the Se8 and Ld130 proteins. Gaps introduced to optimize the alignment are indicated with dashes. Shading is used to indicate the occurrence of identical (black) or substitutional (gray) amino acids. The box represents the obtained N-terminal amino acid sequence of the 60-kDa BV envelope protein; the vertical arrow indicates the putative endoprotease cleavage site. Line I indicates the furin-like cleavage consensus sequence. Line II, III and IV represent the predicted leucine zipper, the coiled-coil domain and the C-terminal transmembrane domain, respectively. Computer-predicted consensus N- and O-linked glycosylation sites are indicated by the symbols Ψ and Φ, respectively.

molecular weight of 76 kDa (74 kDa minus the signal peptide) (IJkel *et al.*, 1999) and contains a furin-like proprotein convertase cleavage site (Arg-Xaa-(Arg/Lys)-Arg↓) (Hosaka *et al.*, 1991; Nakayama, 1997) located at the amino acids 146 to 149. This cleavage site in the Se8 product is therefore likely to be responsible for the occurrence of the 59-kDa protein in BVs.

Computer analysis of Se8 predicted the occurrence of a C-terminal transmembrane domain (amino acids 580-602), a coiled coil structure (amino acids 222-249), a leucine zipper (amino acids 177-198) and a positively charged cytoplasmic tail domain (CTD) (Fig. 5.3). In the AcMNPV and OpMNPV major envelope fusion proteins similar domains were identified. A predicted amphipathic alpha helix and a small hydrophobic domain have been identified as necessary for membrane fusion by GP64 (Monsma and Blissard, 1995; Kingsley *et al.*, 1999). Although GP64 proteins have a small charged cytoplasmic tail domain, it was not essential for efficient budding of AcMNPV BVs (Oomens and Blissard, 1999).

Previously, Se8 was identified as a homolog to the Ld130 protein encoded by ORF 130 of LdMNPV, with 41% amino acid identity and 60% similarity (IJkel *et al.*, 1999). The furin-like proprotein convertase consensus cleavage sequence was also present in Ld130 (amino acids 144-147) at an analogous position compared to Se8 (Fig. 5.3). In conclusion, Se8 encoded the major envelope protein of SeMNPV BV and the Se8 protein appears to be present as a cleavage product in the BV. Se8 contained several protein domains similar to those identified in the AcMNPV and OpMNPV EFPs.

The Se8 encoded protein mediates pH-dependent membrane fusion
To determine whether Se8 and its LdMNPV homolog (Ld130) could mediate pH-dependent membrane fusion, we constructed plasmids in which these ORFs were cloned under the control of an optimized OpMNPV early *gp64* promoter, to facilitate expression in insect cells (Blissard and Rohrmann, 1991). As a positive control, we used an OpMNPV *gp64* gene which contained the OpMNPV *gp64* ORF behind the same *gp64* promoter. The empty plasmid vector (p166BRNX-AcV5) was used as a negative control. Sf21 cells were transfected with these constructs and syncytium formation assays were performed. Because Se301 cells are difficult to efficiently transfect (in contrast to Sf21 cells) and because cell-to-cell fusion was difficult to monitor in Se301 cells (multimorphic cell phenotypes are found in this line), we used transfected Sf21 cells to examine cell-to-cell fusion. The examination of cell-to-cell fusion of SeMNPV infected Sf21 cells (Fig. 5.1C), indicated that this cell line was suitable for studies of membrane fusion mediated by SeMNPV proteins. Cells

transfected with the Se8, the Ld130 or the OpMNPV *gp64* construct clearly showed low pH-dependent cell-to-cell fusion (Fig. 5.4). Although some membrane fusion was observed at earlier times after the pH drop, a substantial degree of fusion occurred by 4 h after shifting the pH to 5.0. In cells transfected with the empty p166BRNX-AcV5 construct, no cell-to-cell fusion was detectable upon lowering the pH to 5 (Fig. 5.4). Thus, the Se8 protein or the LdMNPV homolog (Ld130) protein alone was sufficient to mediate pH-dependent membrane fusion in Sf21 cells.
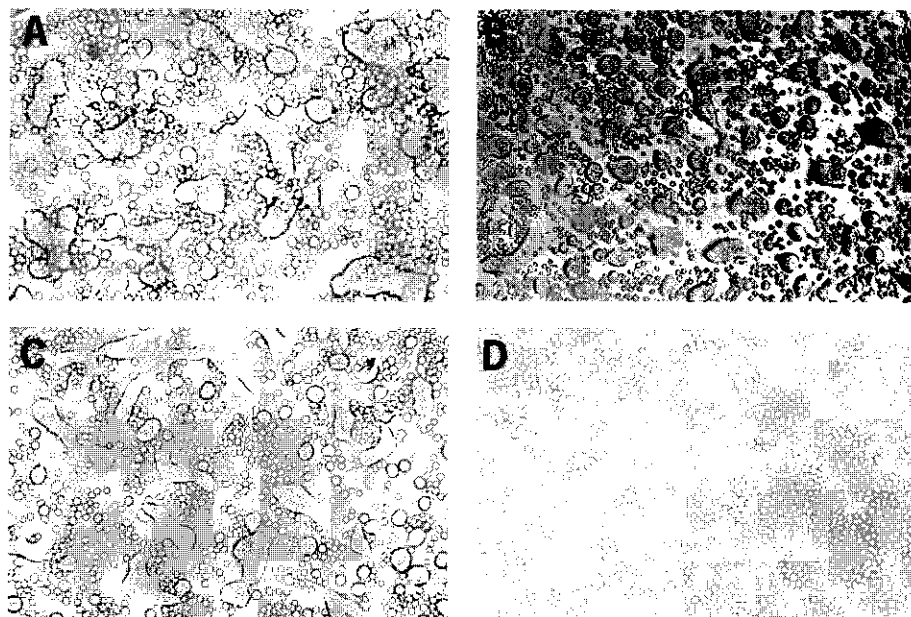
Fig. 5.4 Cell-to-cell fusion of Sf21 cells transfected with putative baculovirus EFP. Cells (1×10^{5}) were transfected with 5 µg of plasmid p166AcV5-Se8 (A), p166AcV5-Ld130 (B), p166-OpGP64 (C), or a control p166BRNX-AcV5 plasmid (D). At 48 h after transfection, cells were treated for 2 min with Grace's medium, pH 5.0. Syncytium formation was scored 4 h after dropping the pH by phase-contrast microscopy.

The Se8 protein is localized in the plasma membrane

To investigate the cellular localization of the Se8 and the Ld130 proteins, C-terminal GFP-fusion constructs were made in plasmid p166BRNX-AcV5. As a negative control, GFP alone was cloned in the same vector. Sf21 cells were transfected with 5 µg plasmid DNA, incubated for 48 h, and examined for fluorescence by confocal laser scanning microscopy. The Se8 and the Ld130 GFP-fusion proteins were primarily localized in the plasma membrane (Fig. 5.5A, B). Diffuse patches of fluorescence were observed in the cytoplasm (Fig. 5.5). No fluorescence was observed in the

nucleus of transfected insect cells. This is in agreement with the previously observed localization of the OpMNPV GP64 in the plasma membrane (Blissard and Rohrmann, 1989). When expressed alone, the GFP protein showed homogeneous fluorescence in the cytoplasm and nucleus (Fig. 5.5C). Similar results were obtained when Se301 cells were transfected with the GFP-fusion constructs. Thus, the Se8 and Ld130 GFP-fusion proteins were primarily localized in the plasma membrane of Sf21 and Se301 cells and this localization of Se8 and Ld130 is consistent with their fusogenic activity.
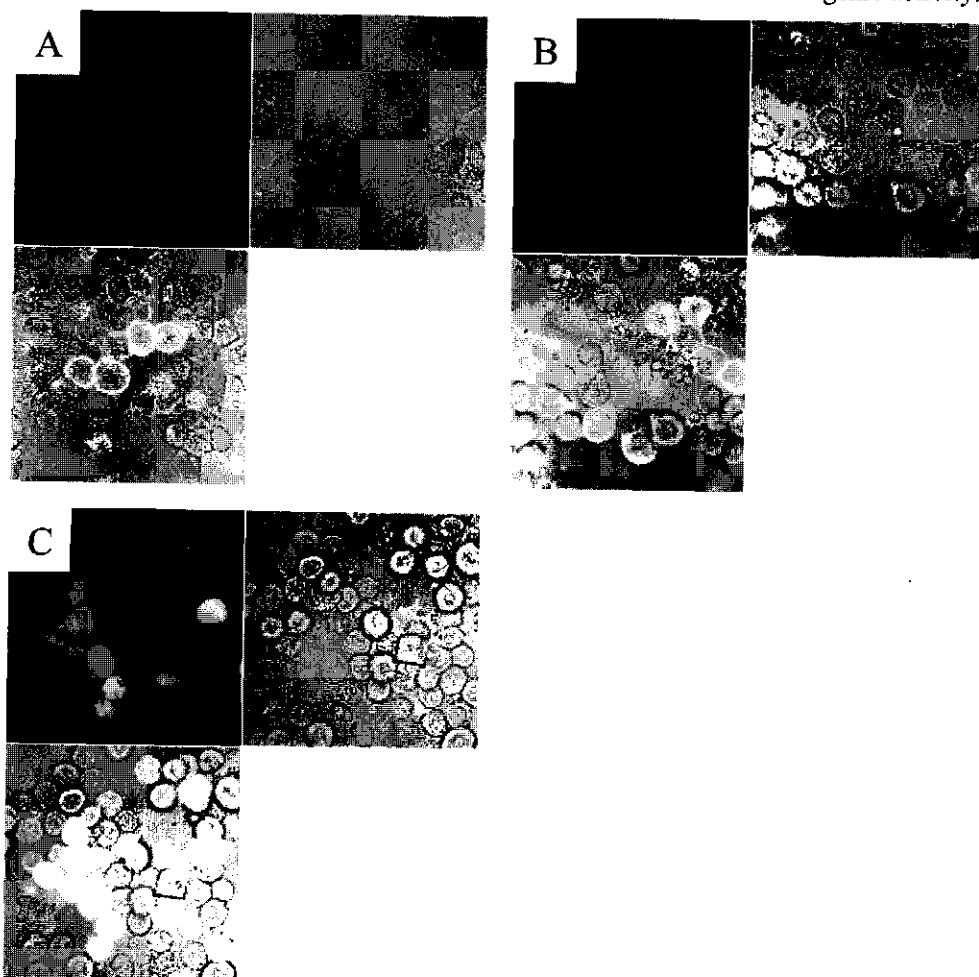
Fig. 5.5 Localization of the Se8 and Ld130 GFP-fusion proteins in Sf21 cells.
Sf21 cells were transfected with the plasmid p166AcV5-Se8GFP (A), p166AcV5-Ld130GFP (B), or with the control plasmid p166AcV5-GFP (C). At 48 h after transfection the cells were examined by confocal laser scanning microscopy for fluorescence. Phase-contrast micrographs are shown to the right of the fluorescence graph. Overlay micrographs of the fluorescence and phase-contrast micrographs are shown below the fluorescence micrographs.

Discussion

SeMNPV BV entry into host cells

The major envelope glycoprotein (GP64) of the AcMNPV BV is involved in host receptor binding and fusion with the host cell membrane during viral entry. Deletion of the AcMNPV *gp64* gene resulted in a virus unable to move from cell to cell, nonlethal to infected larvae, and incapable of efficient budding (Monsma *et al.*, 1996; Oomens and Blissard, 1999). Homologs of the AcMNPV *gp64* gene (Whitford *et al.*, 1989) have been found in *Anagrapha falcifera* NPV (Federici and Hice, 1997), *Anticarsia gemmatalis* MNPV (Oomens, 1999), BmNPV (Gomi *et al.*, 1999), *Choristoneura fumiferana* MNPV (Hill and Faulkner, 1994), *Epiphyas postvittana* NPV (AF061579), *Galleria mellonella* NPV (Blinov *et al.*, 1984) *Hyphantria cunea* NPV (AF190124), and OpMNPV (Blissard and Rohrmann, 1989). These viruses are all phylogenetically placed within subgroup I of the NPVs (Zanotto *et al.*, 1993). No *gp64* homolog has been identified so far in group II NPVs and GVs. In this study, we provide evidence that group II baculoviruses enter host cells in a manner similar to that of group I baculoviruses, and we have identified a functional homolog of the group I GP64 protein from the group II virus, SeMNPV.

We examined the BV entry mechanism of SeMNPV, a group II NPV, using ammonium chloride as an inhibitor of the endocytotic pathway. A $\geq 95\%$ reduction in polyhedra production was observed when Se301 cells were infected with SeMNPV in the presence of ammonium chloride. This indicates that BVs of SeMNPV enter insect cells primarily by endocytosis. Similar reduction rates were observed when Se301 or Sf21 cells were infected with AcMNPV in the presence of ammonium chloride and these values are in agreement with those reported for AcMNPV in Sf21 and Sf9 cells (Volkman and Goldsmith, 1985; Hefferon *et al.*, 1999). These data suggest that the entry mechanisms of AcMNPV and SeMNPV BVs are similar in this regard for Se301, Sf21 and Sf9 cells. Thus, BVs from both group I and II NPVs enter cells by endocytosis, although the group II NPVs lack a GP64 protein.

SeMNPV infection of Sf21 cells did not lead to a productive infection, as evidenced by the absence polyhedra. However, the observation that SeMNPV infection and also Se8 transfection of Sf21 cells can result in cell-to-cell fusion suggests that some viral proteins are expressed from either an early or late SeMNPV promoter. This observation further suggests that the abortive infection of Sf21 cells by SeMNPV may not be the result of inhibition of early gene expression but is likely the result of a block in later events, since BV and polyhedra were not produced from such infections (Shirata *et al.*, 1999).

The SeMNPV major envelope protein of BVs

The SeMNPV structural proteins of BVs and ODVs were isolated and examined by SDS-PAGE. The major BV envelope protein of SeMNPV had an apparent molecular weight of 60 kDa and a major protein of similar size was not detected in ODVs. Sequencing of the N-terminus of the 60-kDa BV major envelope protein from SeMNPV revealed that this protein was encoded by Se8 and was present in BVs as a truncated product. The amino acid sequence (RRSKR) that immediately precedes the N-terminal amino acid (G) of the 60-kDa protein resembles a furin-like proprotein convertase cleavage site (R-X-R/K-R). We therefore hypothesize that the N-terminal 150 amino acids are cleaved from the full-length Se8 protein during Se8 synthesis and processing. This processing would likely result in cleavage products with predicted molecular weights of approximately 15 kDa (the N-terminal fragment minus the signal peptide) and 59 kDa (the C-terminal fragment).

The size of the 59-kDa cleavage product is in close agreement with the estimated 60 kDa of the major envelope protein isolated from SeMNPV BVs. The presence of the 60-kDa major envelope protein in the BV envelope is in accordance with the predicted C-terminal transmembrane domain of the Se8 protein. The 59-kDa product possesses a number of putative phosphorylation, N-glycosylation, myristoylation and amidation sites. It is not known whether these modifications actually occur, although the close similarity between measured and predicted size would suggest that the C-terminal fragment is not heavily glycosylated. Several protein bands around 15-17 kDa were detected in SeMNPV BVs (Fig. 5.2) and we might speculate that one of these may be the 15-kDa cleavage product of the Se8 encoded protein.

Se8 may be similar to proteins such as HIV GP160 which requires cleavage to form GP120 and GP41 subunits that remain associated, but noncovalently. Other proteins such as the influenza hemagglutinin (HA) protein also require cleavage for function, but HA subunits remain covalently associated by disulfides. Protein sequencing of small protein bands associated with the SeMNPV BV may reveal whether a Se8 cleavage product is present in BVs. Whether Se8 cleavage is required for fusion activity and whether a cleavage product is necessary for this or other functions remain to be determined.

Function and localization of the Se8 and Ld130 proteins

To determine whether the major envelope protein from SeMNPV and LdMNPV BV could mediate pH-dependent membrane fusion, a previously described syncytium formation assay of *Lymantria dispar* cells was adopted for Sf21 cells (Blissard and Wenz, 1992). The low pH value (5.0) was selected because similar values have been

reported for eucaryotic endosomes after acidification (Helenius *et al.*, 1982) and because this pH was shown to be sufficient to trigger AcMNPV, OpMNPV, and AgMNPV *gp64*-mediated membrane fusion (Blissard and Wenz, 1992; Oomens, 1999). The syncytium formation assays with cells expressing Se8 and the LdMNPV homolog (Ld130) both showed membrane fusion and syncytium formation upon acidification of the medium. This is in agreement with the observed pH-dependent cell-to-cell fusion observed in Sf21 and Se301 cells infected with SeMNPV BVs. Thus, the SeMNPV major BV envelope protein (Se8) and the homolog from LdMNPV (Ld130) are each able to independently mediate low pH-triggered membrane fusion. They therefore represent functional homologs of the OpMNPV GP64 protein.

To mediate cell-to-cell fusion, the Se8 and Ld130 proteins must be localized at the plasma membrane. GFP-fusion constructs of Se8 and Ld130 proteins were observed at the periphery of transfected cells using confocal fluorescence microscopy (Fig. 5.5). Because Se8 and Ld130 both contain a predicted hydrophobic signal peptide at the N-terminus and a predicted C-terminal transmembrane domain, it is likely that both proteins are found at the surface of infected cells.

Similarities to baculovirus GP64 proteins

Like GP64 proteins, Se8 and Ld130 appear to be type I membrane glycoproteins. In the AcMNPV and OpMNPV GP64 proteins, 4-3 heptad repeats of leucines and/or methionine residues have been identified. These repeats are predicted to form amphipathic α-helices (Monsma and Blissard, 1995). Disruption of the OpMNPV heptad repeat resulted in defective trimerization (Monsma and Blissard, 1995). Mutational analysis of the AcMNPV heptad repeat revealed that it was essential at a stage after the initiation of the individual trimer conformational change and before the functioning of the putative fusion complex (Kingsley *et al.*, 1999). Hydrophobic 4-3 heptad repeats and a leucine zipper motif (LANTTNSLNSQVKQLNDELIVL at amino acids 177-198 of Se8) have also been identified in the Se8 and Ld130 proteins (Fig. 5.3). Coiled coil regions are predicted for amino acids 222 – 249 of the Se8 protein and amino acids 165-247 of the Ld130 protein. It is thus possible that these predicted coiled coil regions may play a similar role in the membrane fusion process.

Homologs of Se8 in other baculoviruses

Se8 is a homolog of the Ld130 protein, with 41% amino acid identity and 60% similarity (IJkel *et al.*, 1999). Both ORFs also showed very low amino acid identity (~22%) to AcMNPV ORF23 (Ac23), BmNPV ORF14 and OpMNPV ORF21 (IJkel *et*

al., 1999). The latter ORFs also lacked a proprotein convertase consensus cleavage sequence in their predicted amino acid sequences. Also, in preliminary syncytium formation experiments, we were unable to detect membrane fusion mediated by the Ac23 protein (W.F.J. IJkel and J. Mangor, unpublished results). The higher degree of relatedness among the group I NPV *gp64* genes, combined with a much larger degree of divergence among the group II NPVs and GVs and their Se8 homologs, suggest that *gp64* may have been acquired more recently, perhaps resulting in the branching off of the group I NPVs from the group II NPVs and the GVs.

Cleavage of viral glycoproteins by furin

Cleavage of viral envelope glycoproteins seems to be a general mechanism used by viruses to activate envelope fusion proteins. A number of viral envelope proteins are cleaved by furin and these include the human immunodeficiency virus gp160 (Hallenberger *et al.*, 1992), human cytomegalovirus glycoprotein B (Vey *et al.*, 1995), Mouse mammary tumor virus-7 superantigen (Park *et al.*, 1995), Measles virus F_0 (Watanabe *et al.*, 1995), Newcastle disease virus F_0 (Gotoh *et al.*, 1992), Sindis virus gpE2 (Gotoh *et al.*, 1992), Human parainfluenza virus type 3 F_0 (Ortmann *et al.*, 1994) and Avian Influenza virus hemagglutinin A (HA) (Stieneke-Grober *et al.*, 1992).

Furin is a member of a family of subtilisin-like endoproteases called "preprotein convertases" that function in proteolytic processing of a large variety of precursor proteins. Recently, furin was identified and characterized from *S. frugiperda* cells (Sf9) (Cieplik *et al.*, 1998). Sf9 cells are a clonal isolate of Sf21 cells and it is likely that Se301 cells also contain a subtilisin-like endoprotease like Sf furin. The Sf furin was localized mainly in the *trans*-Golgi network (TGN) but was also present at the plasma membrane. The cellular localization of Sf furin in the TGN and the presence of the Sf furin recognition motif (R-X-K/R-R) at the Se8 cleavage site (and at a similar site in Ld130) strongly suggest that both proteins are cleaved by an Sf furin-like endoprotease. Although this strong circumstantial evidence suggests Se8 cleavage by a cellular furin-like endoprotease, we cannot yet exclude the possibility that Se8 may be processed by a viral gene product.

Comparison of Se8 with other viral envelope fusion proteins

The avian influenza HA protein most closely resembles the Se8 protein, when the amino acid sequence of its cleavage site and the size of the protein are considered. HA is synthesized as a precursor (HA_0) of 75 kDa. The precursor, HA_0, is post-translationally cleaved at a conserved arginine residue into two subunits. The two

subunits, HA_1 and HA_2, are linked by a single disulfide bond. Cleavage of HA_0 is necessary for virus infectivity (Klenk *et al.*, 1975; Lazarowitz and Choppin, 1975) because it activates the membrane fusion potential of the HA (Maeda and Ohnishi, 1980; Vey *et al.*, 1992; Steinhauer, 1999). The X-ray crystal structure of HA_0 revealed that its cleavage site (amino acids 323-341 of HA_0) can be seen as a prominent surface loop that protrudes out into solution and is accessible to proteases (Steinhauer, 1999). The amino acid sequence (PTKRRSKR↓GLFNFM) of the putative Se8 cleavage site has 71% similarity to the HA_0 protein cleavage site (PQRKRKKR↓GLFGAI) and the five amino acids at the cleavage site (KR↓GLF) are identical. It is tempting to speculate that cleavage of the Se8 and Ld130 proteins by a furin-like endoprotease may be necessary for virus infectivity and may activate the membrane fusion potential of these proteins, similar to that of the HA_0 protein (Klenk *et al.*, 1975; Lazarowitz and Choppin, 1975). We are currently pursuing a mutational analysis of the cleavage site and its potential function in membrane fusion and virion infectivity.

Materials and Methods

Cells, Insects and Viruses
The *Spodoptera frugiperda* cell-line IPLB-SF-21 (Vaughn *et al.*, 1977) and the *Spodoptera exigua* cell-line Se301 (Hara *et al.*, 1995b) were cultured in plastic tissue culture flasks (Nalge Nunc International, naperville, IL) in Grace's insect medium, pH 5.9-6.1 (GibcoBRL, Gaithersburg, MD), supplemented with 10% foetal bovine serum (FBS). A culture of *Spodoptera exigua* insects was maintained according to Smits and Vlak (1988a). The SeMNPV isolate (Gelernter and Federici, 1986b) was originally obtained from B.A. Federici (Department of Entomology, University of California, Riverside, CA) and was called SeMNPV-US1 (Muñoz *et al.*, 1998). The AcMNPV-E2 strain (Smith and Summers, 1978) was originally obtained from M. D. Summers (Texas A&M University, College Station, TX).

Endocytosis assay
Se301 and Sf21 cells were incubated 30 min prior to infection in Grace's insect medium with 10% FBS containing 50 mM ammonium chloride as lysosomotrophic reagent, to inhibit acidification of endosomes (Lenard and Miller, 1982), or without the latter compound (control). AcMNPV and SeMNPV (m.o.i. of 5 $TCID_{50}$ units/ml) were incubated with cells for 2 h at 4°C in Grace's containing 50 mM ammonium chloride to synchronize the binding. Subsequently, the cells were washed gently twice

in 2 ml Grace's insect medium with 10% FBS with or without (control) 50 mM ammonium chloride and then incubated in the presence or absence (control) of ammonium chloride throughout the infection at 27° for 72 h. The number of infected cells compared to the control (without ammonium chloride) was determined.

Purification of SeMNPV BV and ODV

S. exigua fourth-instar larvae were infected by contamination of artificial diet with polyhedra (10x LD$_{99}$) (Smits and Vlak, 1988a). To purify BVs, 5 ml hemolymph was collected 3 days p.i. in 0.5 ml 0.1x TE (TE is 10 mM Tris, pH 7.5, 1.0 mM EDTA) containing 5 mM phenylthiocarbamide to inhibit prophenoloxidase activity. Hemolymph was clarified at 2,000 *g* for 10 min at 4°C. The supernatant was filtered (0.45-μm filter) and the filtrate overlaid onto a 35-ml, 25-56% continuous sucrose gradient in 0.1x TE. Gradients were centrifuged at 100,000 *g* for 90 min at 4° (Beckman SW28, 24,000 rpm). The BV band was collected and dialyzed overnight at 4° against 0.1x TE. The virus suspension was concentrated by overlaying onto a 1.5-ml, 25/56% discontinuous sucrose gradient. Gradients were centrifuged at 100,000 *g* for 90 min at 4° (Beckman SW55, 30,000 rpm). Bands were collected and dialyzed overnight at 4° against 0.1x TE.

Polyhedra were purified from larvae using the method of Braunagel and Summers (1994). ODVs were purified from polyhedra using a combination of the methods described by Caballero *et al.* (1992) and Braunagel and Summers (1994). Briefly, ODVs were liberated from polyhedra (40 mg/ml) by incubating at RT for 15 min in 0.1 M Na$_2$CO$_3$, 166 mM NaCl, 10 mM EDTA, pH 10.5. Undissolved polyhedra were removed by low-speed centrifugation for 5 min (500 *g*). The supernatant (5 ml) was layered onto a 35-ml, 25-56% (w/w) continuous sucrose gradient in 10 mM Tris-HCl, pH 7.5, and centrifuged at 100,000 *g* for 90 min at 4°C (Beckman SW28, 24,000 rpm). The multiple virus bands were collected, washed by dilution in 0.1x TE, concentrated by centrifugation at 55,000 *g* for 60 min at 4° (Beckman SW41, 18,000 rpm), and resuspended in 0.1x TE. The purity and integrity of BVs and ODVs were checked by electron microscopy.

Fractionating of virions into envelope and nucleocapsids

BVs were incubated in 0.1x TE containing 1% NP-40, at RT for 30 min with gentle agitation. NCs were sedimented by centrifugation at 150,000 *g* for 60 min at 4° (Beckman SW55, 35,000 rpm). The pellet was resuspended in 0.1x TE. The envelope proteins in the supernatant were acetone-precipitated, concentrated by centrifugation (4,000 *g*, 30 min) and the pellet dissolved in 0.1x TE.

SDS-PAGE and protein sequencing

Purified SeMNPV BV proteins were analyzed in a 12 % SDS-PAGE gel, according to Laemmli (1970), and stained with Coomassie brilliant blue. Semidry blotting was performed onto a polyvinyl difluoride (PVDF) membrane (Bio-Rad, Richmond, CA) using CAPS buffer (10 mM CAPS, 10 % (v/v) methanol, pH 11). Proteins were visualized on the PVDF membrane using Coomassie brilliant blue. The major protein band from SeMNPV BV was N-terminally sequenced (Protein Research Facility Amsterdam, The Netherlands).

Computer-assisted analysis of fusion proteins

The Se8 (IJkel *et al.*, 1999) and the Ld130 proteins were analyzed using software of the Predict Protein server (Rost, 1996) for the prediction of transmembrane domains, N-terminal signal sequences, leucine zippers, coiled coil structures and the pI of the cytoplasmic tail domain (CTD). Motif searches were done against the Prosite release 14 database (Fabian *et al.*, 1997; Bairoch *et al.*, 1997). DNA and protein comparisons with entries in the updated GenBank/EMBL, SWISS-PROT, and PIR databases were performed with FASTA and Psi-BLAST programs (Pearson, 1990; Altschul *et al.*, 1997). Multiple sequence alignments were performed with the GCG PileUp and Gap computer programs with gap creation and extension penalty set to 8 and 2, respectively (Devereux *et al.*, 1984). Alignment editing was performed with Genedoc Software (http://www.psc.edu/biomed/genedoc).

Plasmid constructions

To construct plasmids that allow expression in insect cells upon transfection, a plasmid vector, p166BRNX-AcV5 that contains the OpMNPV *gp64* early promoter plus a multiple cloning site and an AcV5 epitope tag (provided by G. Lin), was used to construct vectors containing Se8 and Ld130 encoding putative viral fusion proteins. Primers containing 5' *Bam*HI and 3' *Eco*RI restriction sites were designed for the directional PCR-cloning of SeMNPV ORF 8 and Ld130. These ORFs were amplified by high fidelity "Expand" long template PCR (Boehringer Mannheim, Mannheim, Germany) using the plasmids pSe *Bgl*II-H (IJkel *et al.*, 1999) and genomic LdMNPV DNA (Riegel *et al.*, 1994), respectively, as template DNA. The PCR products were ligated into the *Bam*HI and *Eco*RI sites of p166BRNX-AcV5. The plasmids were named p166AcV5-Se8 and p166AcV5-Ld130, respectively, and tested in syncytium formation assays. The complete OpMNPV *gp64* gene used as a positive control in transfection experiments was from plasmid p64-166 and was described previously (Blissard and Wenz, 1992).

To determine the localization of the putative fusion proteins in insect cells, GFP-fusion constructs were made. The red-shifted GFP ORF (Davis and Vierstra, 1996) was amplified by high fidelity "Expand" long template PCR (Boehringer Mannheim) using primers containing 5' *Eco*RI and 3' *Xba*I restriction sites and cloned in frame as an *Eco*RI/*Xba*I fragment into the plasmids p166AcV5-Se8 and p166AcV5-Ld130 to give C-terminal fusions with the putative BV fusion proteins. Plasmid DNA was purified using Jetstar columns according to manufacturer's protocol (ITK Diagnostics, Uithoorn, The Netherlands) and used for transfections. All constructs were sequenced to confirm the in-frame cloning of the AcV5 tag or GFP with the putative fusion ORFs. Sequencing was carried out using an Applied Biosystems automated DNA sequencer (Eurogentec, Belgium).

Syncytium formation assay

Syncytium formation (Sf21-Sf21 fusion or Se301-Se301 fusion) assays were performed by either transfection of 5×10^5 Sf21 or Se301 cells, with 5 µg of plasmid DNA using Cellfectin (GibcoBRL) or infection with SeMNPV or AcMNPV BVs (m.o.i. of 10 $TCID_{50}$ units/ml). As a negative control for the syncytium formation assay, we used the empty p166BRNX-AcV5 plasmid vector. Forty-eight hours after transfection, cells were washed three times with 2 ml Grace's medium (pH 6.1) without FBS, and afterwards cells were treated for 2 min in 1 ml acidic Grace's medium at pH 5.0. The acidic medium was removed and replaced with 2 ml Grace's (pH 6.1) with 10% FBS. Syncytium formation was scored and observed by light microscopy 4 h after treatment with the acidic medium. Syncytium formation was recorded when at least 4 nuclei were present in each syncytial mass.

Fluorescence microscopy

Sf21 cells (1×10^5) were grown on glass cover slips and transfected with 5 µg of plasmid DNA. At 48 h posttransfection the cells were examined with a Zeiss LSM510 (confocal) laser scanning microscope for fluorescence using an excitation wave length of 488 nm and an emission band pass filter of 505 - 530 nm.

Chapter 6

Isolation of a *Spodoptera exigua* baculovirus recombinant with a 10.6 kbp deletion that retains biological activity

When *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV) is grown in insect cell culture, defective viruses are generated. These viruses lack about 25 kbp of sequence information and are no longer infectious for insects. This makes the engineering of SeMNPV for improved insecticidal activity or as expression vectors difficult to achieve. Recombinants of *Autographa californica* MNPV have been generated in insects after lipofection with viral DNA and a transfer vector into the haemocoel. In the present study a novel procedure to isolate SeMNPV recombinants was adopted by alternate cloning between insect larvae and cultured cells. The *S. exigua* cell line Se301 was used to select the putative recombinants by following a green fluorescent protein marker inserted in the *p10* locus of SeMNPV. Polyhedra from individual plaques were fed to larvae to select for biological activity. In this way a SeMNPV recombinant (SeXD1) was obtained with the speed of kill improved by about 25%. This recombinant lacked 10,593 bp of sequence information, located between 13.7 and 21.6 map units of SeMNPV and including *ecdysteroid UDP glucosyl transferase, gp37, chitinase* and *cathepsin* genes, as well as several genes unique to SeMNPV. The result indicated, however, that these genes are dispensable for virus replication both *in vitro* and *in vivo*. A mutant with a similar deletion was identified by PCR in the parental wild-type SeMNPV isolate, suggesting that genotypes with differential biological activities exist in field isolates of baculoviruses. The generation of recombinants *in vivo*, combined with the alternate cloning between insects and insect cells, is likely to be applicable to many baculovirus species in order to obtain biologically active recombinants.

Introduction

The beet army worm *Spodoptera exigua* causes extensive economic losses in many cultivated crops throughout the temperate and subtropical regions of the Northern hemisphere and in greenhouses. The insect is resistant to many commonly used chemical insecticides. *S. exigua* multicapsid nucleopolyhedrovirus (SeMNPV) is an attractive bio-insecticide since the virus is monospecific to the beet army worm and highly virulent as compared to other baculoviruses (Smits *et al.*, 1988). It has also been commercialized as a bio-insecticide (Smits and Vlak, 1994). However, further improvements in the biological activity of SeMNPV are sought, either by strain selection (Muñoz *et al.*, 1998) or by genetic engineering.

The molecular genetics of SeMNPV have been relatively well studied. A detailed physical map has been constructed (Heldens *et al.*, 1996) and a number of SeMNPV genes have been characterized in detail (Van Strien *et al.*, 1992; Zuidema *et al.*, 1993; Van Strien *et al.*, 1996; Van Strien *et al.*, 1997; Heldens *et al.*, 1997b). Recently the complete sequence and gene organization of the SeMNPV genome have been reported (IJkel *et al.*, 1999). However, the molecular basis for specificity and virulence has not yet been revealed.

Several cell lines have been derived from *Spodoptera exigua*, such as SeUCR (Gelernter and Federici, 1986a), Se301 (Hara *et al.*, 1995b) and IZD2109 (B. Möckel, personal communication), and susceptibility to SeMNPV has been reported (Hara *et al.*, 1993; 1995a). However, when SeMNPV is grown in insect cell culture defective viruses are quickly generated (Heldens *et al.*, 1996). The majority of these viruses lack about 25 kbp of sequence information and are no longer infectious for insects. The deletion is located approximately between 12.9 and 32.3 map units (m.u.) and encompasses the SeMNPV open reading frames (ORFs) 15 to 41 (IJkel *et al.*, 1999). This makes the engineering of SeMNPV for improved insecticidal activity or as expression vectors difficult to achieve. The generation of defective viruses in cell culture limits the structural and functional analysis of the SeMNPV genes and the isolation of recombinants with adequate infectivity *in vivo* and in cell culture.

SeMNPV has been isolated from many different geographical regions throughout the world (Vlak *et al.*, 1981; Gelernter and Federici, 1986b; Hara *et al.*, 1995a; Muñoz *et al.*, 1998). Wild type (wt) SeMNPV isolates are frequently found consisting of several genotypic variants. This is typically indicated by the presence of submolar bands in restriction endonuclease digestion profiles of viral DNA (Muñoz *et al.*, 1998; 1999). Isolation of individual genotypic variants by *in vivo* cloning methods (Smith and Crook, 1988) has allowed the evaluation of the relative virulence of the different genotypic variants (Muñoz *et al.*, 1998; 1999). Since multiple passaging of

SeMNPV in cultured insect cells results in the generation of defective viruses (Heldens *et al.*, 1996), cloning of genotypic variants of SeMNPV is difficult to obtain by conventional plaque purification techniques. Hence, a novel strategy was adopted in this study to generate genotypic variants of SeMNPV by cloning alternately *in vivo* and *in vitro*.

We previously reported that recombinants of *Autographa californica* (Ac) MNPV were successfully generated in *S. exigua* larvae by transfection of viral and transfer vector DNA into the hemocoel by lipofection (Hajós *et al.*, 1998). In this study we used a similar strategy to generate recombinants of SeMNPV. We also adopted a novel procedure to isolate SeMNPV recombinants by cloning alternately between *S. exigua* larvae and Se301 cultured cells to secure *in vivo* and *in vitro* infectivity. By this strategy a SeMNPV recombinant (SeXD1) was generated using GFP as a screening marker. This recombinant had a similar genetic make-up to one of the variants observed in the SeMNPV isolates and was able to replicate both *in vivo* and in cultured insect cells. However, it lacked 10,6 kbp of nucleotide sequence as compared to the complete SeMNPV genome. Bioassays indicated that the recombinant has superior speed of kill as compared to the wt isolate.

Results

Generation of $p10^-$ recombinant SeXD1

To generate a $p10^-$ SeMNPV recombinant, the transfer vector pSeXD1 carrying a green fluorescent protein (GFP) marker gene was constructed (Fig. 6.1). The size of transfer vector pSeXD1 was 4.6 kbp, and it contained 503 bp upstream (including *p10* promoter) and 673 bp downstream [including the *p10* poly(A) motif; van Oers *et al.*, 1999] sequences of the SeMNPV *p10* ORF, and the 747 bp GFP gene driven by the authentic *p10* promoter.

Seventeen fourth instar larvae were injected with wt SeMNPV and pSeXD1 DNA at a ratio of 1:30 µg, corresponding to a molar ratio of approximately 1:800. Sixteen larvae survived the injection treatment. The hemolymph from these 16 larvae was transferred to 5 ml of Grace's medium without FBS. The controls included larvae injected with only viral DNA, only transfer vector DNA, only Cellfectin and untreated larvae. Plaque assays indicated that the total virus titre of the hemolymph was 3.7×10^4 p.f.u./ml. The percentage of recombinants was approximately 3.3% in agreement with data obtained previously for AcMNPV (Hajós *et al.*, 1998). The GFP gene driven by the *p10* promoter of SeMNPV induced bright fluorescence, as observed with an UV
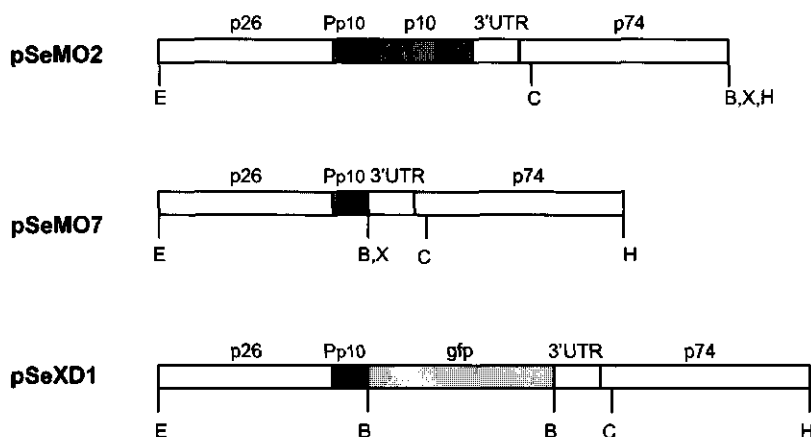
Fig. 6.1 Schematic representation of the *p10* locus in SeMNPV and in the *p10*-based transfer vectors pSeMO7 and pSeXD1. Plasmid pSeMO2 contains a 1448 bp *Eco*RI-*Bam*HI fragment derived from the *Xba*I-H fragment of SeMNPV containing the *p10* locus and its flanking sequences. Plasmid pSeMO7 is the empty SeMNPV *p10* promoter-based transfer vector. In plasmid pSeXD1 the open reading frame for the green fluorescent protein (GFP) is present downstream of the *p10* promoter. Pp10: *p10* promoter; 3'UTR: *p10* 3' untranslated region including a poly(A) motif. B: *Bam*HI, C: *Cla*I, E: *Eco*RI, H: *Hin*dIII, X: *Xba*I.

microscope. With the help of GFP, it was easy to screen by fluorescence and pick recombinant plaques from Se301 cells. Several recombinant viruses were isolated by three rounds of alternate cloning between third instar *S. exigua* larvae and Se301 cells (see Methods). Finally, recombinant SeXD1 was amplified in fourth instar *S. exigua* larvae and analysed.

To confirm the location of the GFP gene insertion, recombinant SeXD1 DNA was examined by *Spe*I restriction endonuclease digestion. The SeMNPV *p10* gene is located on the 4.5 kbp *Spe*I I fragment corresponding to nt 122,885 to 127,355 of the genome (Fig. 6.2A) (IJkel *et al.*, 1999). When the 264 bp *p10* ORF (corresponding to nt 123,740 to 124,006) is replaced with the 747 bp GFP gene ORF, the *Spe*I I fragment would become 5.0 kbp (Fig. 6.2A). As shown in Fig. 6.2B, the *Spe*I restriction endonuclease pattern of SeXD1 confirmed the insertion of the GFP ORF into the *p10* locus in SeXD1. SDS-PAGE and Western analysis showed the absence of the P10 protein while GFP was expressed in both Se301 and SeUCR cells infected with SeXD1 (Fig. 6.2C, D: lane 3 and lane 6).

Uninfected Se301 cells are shown in Fig. 6.3A. At 16 h p.i., polyhedra were observed in about 20% of the Se301 cells infected with either wt SeMNPV or SeXD1 at a m.o.i. of 10 (data not shown). At 48 h p.i., polyhedra were observed in about 90% of the Se301 cells infected with wt SeMNPV (Fig. 6.3B) and in almost 100% of
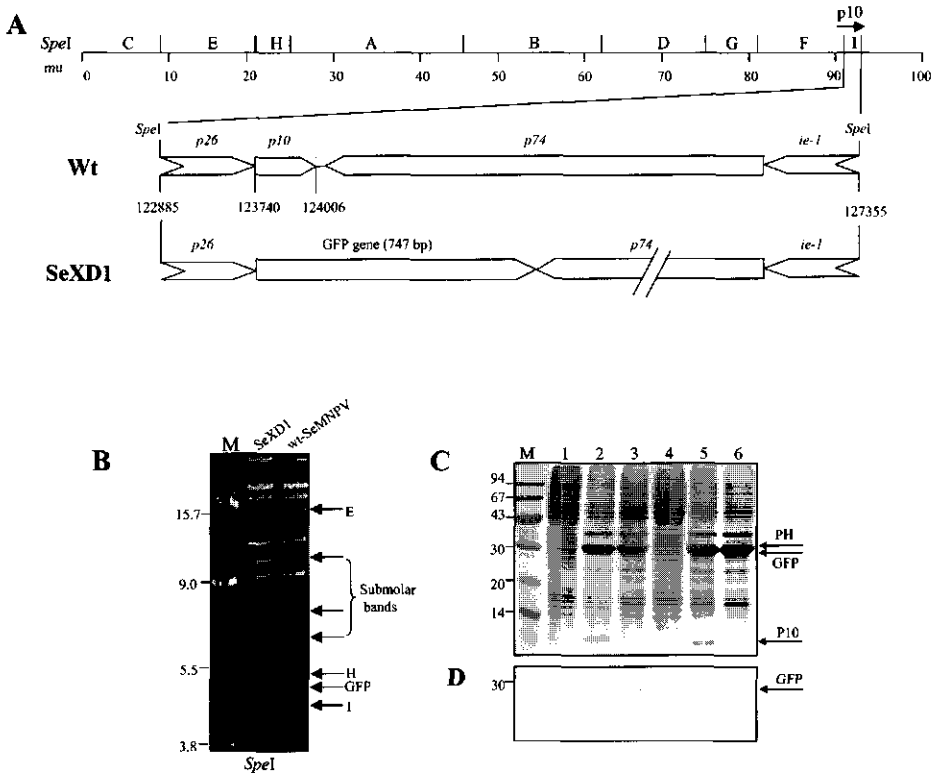
Fig. 6.2 Construction and analysis of SeXD1, a *p10⁻* SeMNPV recombinant expressing GFP. (A) Construction of SeXD1. The top line represents the physical map of the wt SeMNPV genome for *Spe*I restriction endonuclease. *p10* is located in the 4.5 kb *Spe*I I-fragment (corresponding to nt 122,885 to 127,355). The 264 bp *p10* ORF (corresponding to nt 123,740 to 124,006) was replaced with the 747 bp GFP gene; the *Spe*I I-fragment became 5.0 kb in SeXD1. (B) *Spe*I restriction endonuclease analysis of genomic DNAs from wt SeMNPV and SeXD1. (C) and (D) SDS-PAGE and Western analysis using anti-GFP antibodies. Uninfected Se301 cells (lane 1), Se301 cells infected with wt SeMNPV (lane 2) and with SeXD1 (lane3); uninfected SeUCR cells (lane 4), SeUCR cells infected with wt SeMNPV (lane 5) and with SeXD1 (lane 6). PH, polyhedrin.

Se301 cells infected with SeXD1 (Fig. 6.3C). Bright fluorescence was observed in SeXD1-infected Se301 cells under the UV microscope (Fig. 6.3D). No fluorescence was observed either in wt SeMNPV-infected or in uninfected Se301 cells (data not shown).

Thus, UV microscopy, restriction enzyme analysis, SDS-PAGE and Western blot analysis demonstrated that the recombinant virus SeXD1 lacked the *p10* gene and expressed GFP. This recombinant was capable to complete its replication cycle both in *S. exigua* larvae and in the cultured cell lines Se301 and SeUCR.

103
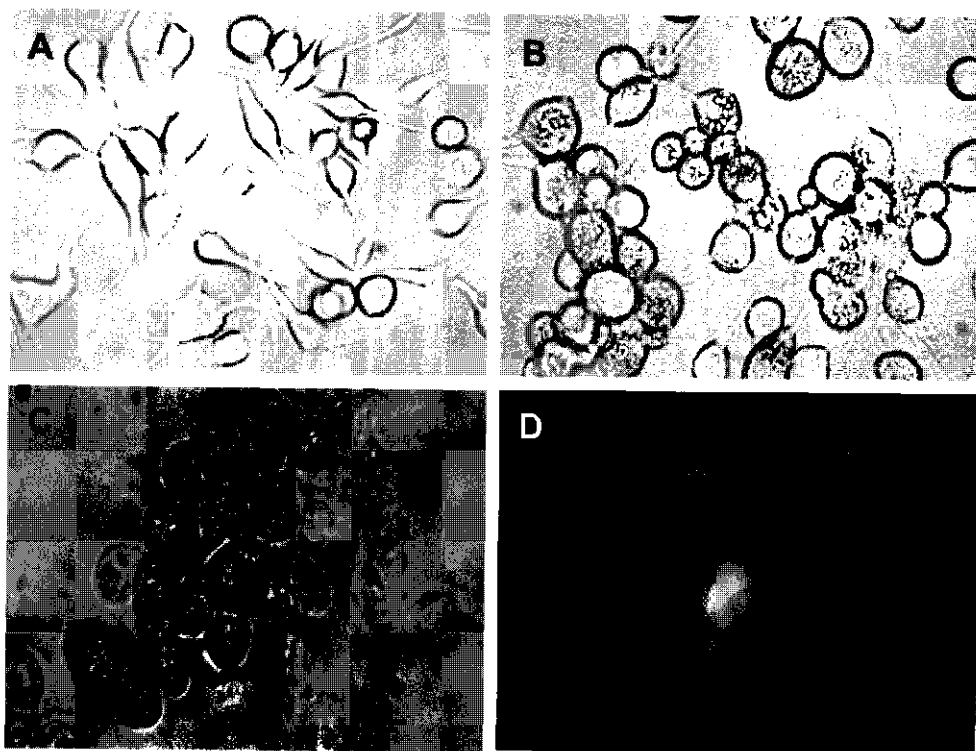
Fig. 6.3 Phase-constrast and UV micrographs of the *S. exigua* cell line Se301.
Se301 cells (A) were infected with wt SeMNPV (B) and with SeXD1 (C and D) at 48 h p.i.
Polyhedra were observed in phase-contrast images of wt SeMNPV (B) and SeXD1-infected cells (C).
The expression of GFP in SeXD1-infected cells is shown by irradiation with UV-light (D).

Analysis of deletion mutants

Wt SeMNPV is made up of several genotypic variants (Muñoz *et al.*, 1998; 1999) and replication of SeMNPV in cultured cells often results in the generation of deletion mutants (Heldens *et al.*, 1996). To determine whether the recombinant SeXD1 is one of these variants, SeXD1 as well as wt SeMNPV were analysed with restriction endonucleases. The *Spe*I and *Pst*I digestions showed several submolar bands in wt SeMNPV (Fig. 6.2B and 6.4B), indicating that the wt SeMNPV isolate is a mixture of genotypes. No submolar bands were found in the *Spe*I and *Pst*I digestion patterns of SeXD1 DNA (Fig. 6.2B and 6.4B). However, the *Pst*I-D, *Spe*I-E and *Spe*I-H fragments were absent in SeXD1 (Fig. 6.2B and 6.4B), suggesting that although SeXD1 is genetically homogeneous it might be a deletion mutant. One of the submolar bands found after *Spe*I digestion in SeMNPV wild type (Fig. 6.2B) is a molar band in the recombinant, suggesting that a variant with a similar deletion is present in wt SeMNPV.

To determine in more detail which region was absent, both SeXD1 and wt SeMNPV DNA were examined by PCR amplification. The restriction analysis had shown the absence of the *Pst*I D-fragment in SeXD1, while the neigbouring fragments L and C were retained (Fig. 6.4A, B). Therefore, PCR primers were designed annealing approximately 100 bp up- and downstream of the *Pst*I D-fragment in fragments L and C, respectively (see Methods). In an amplification from complete genomic SeMNPV DNA the PCR product should be 11,289 bp, and a product of this size was indeed observed (Fig. 6.4A). Amplification from SeXD1 DNA, however, resulted in a single ~700 bp product (Fig. 6.4C), suggesting that about 10 kbp was deleted from SeXD1. The PCR analysis also indicated that SeXD1 most likely contained a single genotype. Amplification using wt SeMNPV DNA as template resulted in at least five products, including a 11 kbp, 2.8 kbp, 2.0 kbp, 1.2 kbp, and 700 bp product (approximate sizes; Fig. 6.4C). These results suggested that the wt SeMNPV is a mixture containing several deletion mutant variants in this locus. Conclusions about the relative amounts of the variants can not be drawn from this analysis, however, since smaller fragments are likely to be amplified more efficiently than larger ones.

The ~700 bp product was observed in both SeXD1 and wt SeMNPV (Fig. 6.4C), implying that SeXD1 might have originated from one particular genotypic variant in the wt SeMNPV isolate. To exactly locate the deleted region and to compare SeXD1 with wt SeMNPV, the ~700 bp fragments from both SeXD1 and wt SeMNPV were cloned into pGEM-T and sequenced. Sequence analysis showed the presence of both primers in the PCR products and mapped the deletion of SeXD1 from 13.7 to 21.6 m.u. (10,593 bp, from nt 18,513 to nt 29,106) (Fig. 6.4A). The deletion in a genotypic variant of wt SeMNPV was also from nt 18,513 to nt 29,106, a total of 10,593 bp (Fig. 6.4A). A total of 12 ORFs was completely deleted, encompassing SeMNPV ORF 16 to 27 and including *ecdysteroid UDP glucosyl transferase* (*egt*), *gp37*, *chitinase* (*chiA*), *cathepsin* (*v-cath*), *ptp-2* and nine others. Two ORFs, ORF 15 and 28, were partially deleted. Therefore, the sequences maintained in SeXD1 and in one of the wt SeMNPV variants were the same, suggesting that SeXD1 is derived from an existing genotypic variant of wt SeMNPV.

SeXD1 was passaged in Se301 cells several times when purified but still retained the same deletion as its parental wt SeMNPV. The result indicated that the genotypic variant with a deletion of 10,593 bp was quite stable. The result also indicated that naturally *egt*-deleted, *gp37*-deleted, *chiA*-deleted and *v-cath*-deleted genotypes existed in the wt SeMNPV population and that none of the deleted genes are required for viral DNA replication either *in vivo* or *in vitro*.

A

B   **C**
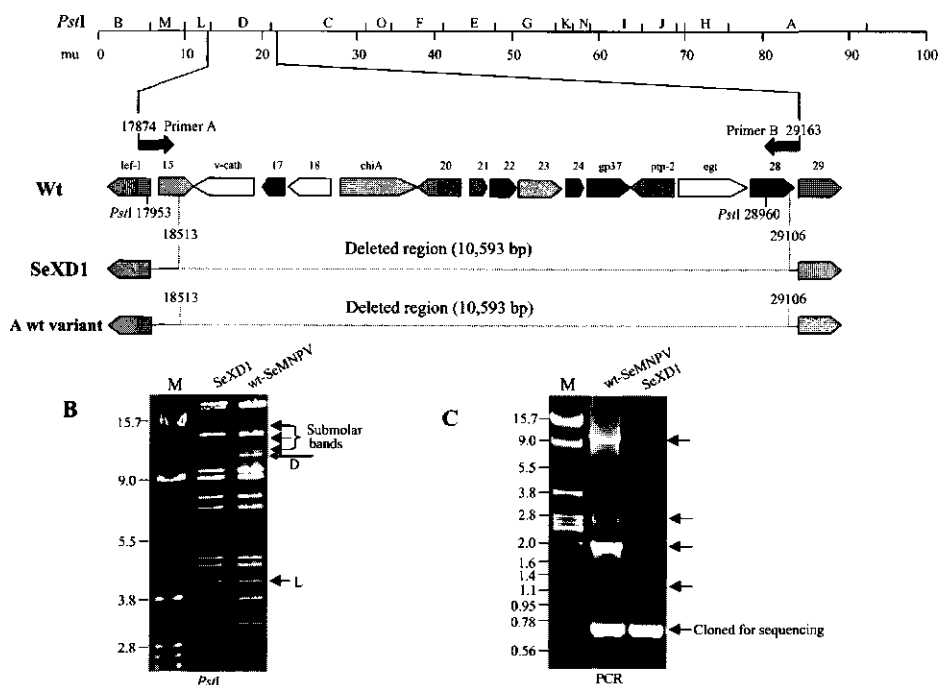
Fig. 6.4 Analysis of the deletion mutants.

(A) The top line represents the physical map of the wt SeMNPV genome for *Pst*I restriction endonuclease. A 10,593 bp fragment, including *egt*, *v-cath*, *gp37*, *chiA*, etc. (corresponding to nt 18,513 to 29,106; dashed horizontal line), was deleted in the recombinant SeXD1. The deletion in a genotypic variant of wt SeMNPV was also from nt 18,513 to nt 29,106 (a total of 10,593 bp). (B) *Pst*I restriction endonuclease analysis of genomic DNAs from wt SeMNPV and SeXD1. The *Pst*I-D fragment was absent in SeXD1. (C) PCR analysis of genomic DNAs from wt SeMNPV and SeXD1. The primers A and B correspond to nt 17,874 to 17,904 and 29,135 to 29,163, respectively.

Biological activity and symptomatology of virus-infected *S. exigua* larvae

The insecticidal activities of the recombinant SeXD1 and wt SeMNPV were determined for third instar *S. exigua* larvae in terms of LD_{50} and ST_{50} (Table 6.1). The ST_{50} value of SeXD1 (70.2 h) was 25% lower than that of wt SeMNPV (93.1 h). The ST_{50} value was significantly different (P<0.05). The slopes of the filled time-mortality relationships were not significantly different for both viruses.

The LD_{50} value of SeXD1 [403 occlusion bodies (OBs)/larva] was approximately three times higher than that of wt SeMNPV (125 OBs/larva), but this was not significantly different (P=0.094) (Table 6.1). The slopes of the filled dose-mortality curves were not significantly different (P=0.05).

There were some differences in symptoms of wt SeMNPV and SeXD1-infected *S. exigua* larvae. The larvae infected with wt SeMNPV became pale and creamy in color prior to death. After death infected insects rapidly liquefied. A small proportion of the wt SeMNPV-infected larvae first turned black before liquefaction. The larvae infected with SeXD1 also became pale prior to death but all larvae turned black. In addition, the SeXD1-infected larvae did not liquefy after death and remained physically intact (data not shown), a typical phenotype of infection with a baculovirus lacking *cathepsin* and/or *chitinase* (Slack *et al.*, 1995; Hawtin *et al.*, 1997).

Table 6.1 Dose mortality (LD_{50}) and lethal time mortality (ST_{50}) of wt SeMNPV and recombinant SeXD1 for third instar *S. exigua* larvae.

The data in the table came from the statistical analysis. The LD_{50} was determined in three repetitions by a leaf disc bioassay and the ST_{50} in four repetitions by a droplet-feeding bioassay.

Viruses	LD_{50} Log LD_50	(OBs/larva)	Slope	ST_{50} (h)	Slope
Wt SeMNPV	$4.83^a \pm 0.68$	125^a	$1.50^a \pm 0.32$	$93.1^b \pm 5.9$	$10.92^a \pm 3.86$
SeXD1	$6.00^a \pm 1.15$	403^a	$1.26^a \pm 0.34$	$70.2^b \pm 6.7$	$9.17^a \pm 2.16$

[a], No significant difference; [b] significantly different.

Discussion

Replication of SeMNPV in cultured cells results in the generation of deletion mutants, which are not infectious to *S. exigua* larvae (Heldens *et al.*, 1996). This is the major reason why engineering of SeMNPV has been difficult to achieve in the past several years. Based on the successful generation of AcMNPV recombinants by cotransfection of viral and transfer vector DNA into the haemocoel of *S. exigua* larvae (Hajós *et al.*, 1998) and the supposition that a few intact SeMNPV would survive one or two passages in cultured cells, we adopted a procedure to engineer SeMNPV by alternate cloning between insect larvae and cultured cells. When the molar ratio between viral DNA and transfer vector was 1:30, recombinants were observed at 3.3%. This is in the same order of magnitude as in the case of AcMNPV, where ~2% has been recorded (Hajós *et al.*, 1998). Although the same amount of viral DNA per larva (0.4 µg) was used in the injection, the total virus titre in the haemolymph of the cotransfected larvae is much lower (3.7×10^4 p.f.u./ml) than found for AcMNPV (5.2×10^8 p.f.u./ml) (Hajós *et al.*, 1998). The result suggests that the transfection with SeMNPV DNA is less efficient than with AcMNPV DNA, but that the relative

proportion of recombinants is more or less similar.

A wt SeMNPV isolate is made up of several genotypic variants; some of these contain large deletions and are helper dependent (Muñoz *et al.*, 1998; 1999). PCR and sequence analysis showed that in the recombinant SeXD1 10,593 bp of SeMNPV sequence was deleted (Fig. 6.4A, C). The same procedure revealed the presence of a genotype with a deletion of the same size in wt SeMNPV (Fig. 6.4C). The question is whether the SeMNPV deletion naturally exists in the wt SeMNPV population or results from the passages in cultured cells. Since it was reported that extensive deletions in the SeMNPV genome occurred very quickly in the SeUCR cell line (Heldens *et al.*, 1996), it was generally thought that SeMNPV would loose its pathogenic effect *in vivo* after just one passage with multiple replication cycles in cultured cells and that it would be difficult to obtain SeMNPV variants that retained biological activity *in vivo* from cultured cells. However, with this novel approach we successfully selected several SeMNPV recombinants infectious *in vivo* and *in vitro*, one of which, SeXD1, was analyzed in detail.

Restriction endonuclease and PCR analysis showed the presence of several other genotypes in wt SeMNPV (Fig. 6.2B and 6.4B, C). After the first round of plaque purification using Se301 cells and the haemolymph of cotransfected larvae, we observed several plaques containing polyhedra that were not infectious for *S. exigua* larvae (X. Dai, unpublished data). However, most plaques were pathogenic for *S. exigua* larvae. In our study we picked plaques in Se301 cells 3 days p.i. and then amplified the plaques in Se301 cells for another 3 days before harvesting the polyhedra-containing cells. Thus, recombinant SeXD1 grown in Se301 cells for about two passages still retained its biological activity and consisted of a single genotype. Apparently in Se301 cells the deletion in SeMNPV does not happen as quickly as in SeUCR cells. Hara *et al.* (1993) reported that SeMNPV produced in Se301 cells was still infectious for larvae. Recently, Choi *et al.* (1999) generated a SeMNPV *polyhedrin⁻* recombinant in these cells, but its infectivity for insects and its genetic make-up was not studied. Hence, there might be differences in the induction of defective viruses of SeMNPV between Se301 and SeUCR cells and some cell factors might be involved in the generation of deletion mutants.

SeXD1 lacked the *p10* gene of SeMNPV and expressed GFP. SeXD1 also lacked 10,593 bp of additional sequence information of SeMNPV, including *egt*, *gp37*, *chiA*, *v-cath*, and ten other genes located in this region (IJkel *et al.*, 1999). Bioassays showed that the ST_{50} value of SeXD1 was 25% lower than of wt SeMNPV, but that the LD_{50} value of SeXD1 was approximately the same as for wt SeMNPV (Table 6.1). The result suggests that the absence of one or more genes may be responsible for the

enhanced speed of kill. Various studies showed that deletion of *p10* did not lead to an increased speed of kill (Martens *et al.*, 1995; Bianchi *et al.*, 2000). Recent results also indicated that GFP does not affect the biological activity of *Helicoverpa armigera* SMPV (Chen *et al.*, 2000). It has been reported that EGT is a key enzyme in abrogating the regulation of host insect metamorphosis (O'Reilly and Miller, 1989). It conjugates ecdysteroids with sugars and hence blocks molting of the insect. Insects infected with an *egt*-deleted virus exhibit reduced feeding and earlier mortality compared to wt virus-infected larvae (O'Reilly and Miller, 1991; O'Reilly, 1995; Flipsen *et al.*, 1995). Another study has shown that the LT_{50} value of *egt*-deleted *Lymantria dispar* MNPV was about 33% lower than that of wt LdMNPV for fifth instar *L. dispar* larvae (Slavicek *et al.*, 1999). Our findings are thus consistent with these studies on *egt* deletion mutants.

Of those ORFs deleted from SeXD1, ORFs 17, 18 and 21 have homologues in *Xestia c-nigrum* granulovirus (XcGV) (Hayakawa *et al.*, 1999). ORFs 15 and 28 have homologues in LdMNPV (Kuzio *et al.*, 1999). ORFs 20, 22, 23 and 24 are unique to SeMNPV (IJkel *et al.*, 1999), but their function is unknown. SeXD1 was able to replicate in *S. exigua* larvae as well as in the cultured Se301 and SeUCR cells, so all the deleted genes are dispensable for virus replication both *in vivo* and *in vitro*.

Baculovirus *gp37* encodes a spindle-like protein, clearly related to fusolin of entomopoxviruses (EPVs) (Dall *et al.*, 1993; Liu and Carstens, 1996; Mitsuhashi *et al.*, 1997). There is accumulating evidence that fusolin of EPVs can enhance NPV infection in insects (Mitsuhashi *et al.*, 1998; Hayakawa *et al.*, 1996). Baculovirus *gp37* might also be involved in enhancing virus infection in insects (Phanis *et al.*, 1999) and the *gp37/fusolin* gene family might be essential for virus replication (Wu and Miller, 1989). In the present study, the absence of *gp37* did not affect virus replication in a detectable way either in cell culture or in insects. Thus, it remains enigmatic what the function of *gp37* is in the biology of baculovirus infection.

The baculovirus-infected insect host liquefies after death (Volkman and Keddie, 1990) and polyhedra are released. This process plays an important role in ensuring the efficient dissemination of virus by physical forces such as wind and rain splash. It has been reported that *chiA* and *v-cath* are involved in the liquefaction process of virus-infected insect larvae (Ohkawa *et al.*, 1994; Rawlings *et al.*, 1992; Slack *et al.*, 1995; Hawtin *et al.*, 1997). Recombinant SeXD1 with a *chiA* and *v-cath* deletion could not liquefy *S. exigua* larvae, consistent with previous reports. Gopalakrishnan *et al.* (1995) reported that a recombinant AcMNPV containing a *Manduca sexta chiA* gene required less time to kill *Spodoptera frugiperda* fourth instar larvae when injected into the haemocoel. However, Hawtin *et al.* (1997) reported that deletion of *chiA* or *v-cath*

from AcMNPV had no significant effect on LD_{50} or ST_{50} of the recombinant. It is not clear whether the absence of *chiA* and *v-cath* has any effect on the LD_{50} value of SeXD1.

As a result of fluoresence microscopic studies using GFP as a marker, we observed that upon cotransfection of insect larvae SeMNPV recombination took place predominantly in fat body cells. In contrast, with AcMNPV, the recombination upon cotransfection was found to take place typically in the haemocytes (data not shown). GFP also proved to be a helpful marker in the screening of SeMNPV recombinants. This marker will also be useful in analysing the pathological effects of this virus in target and non-target hosts using, for example, confocal laser scanning microscopy. The procedure to generate recombinant viruses followed in this paper is applicable for many baculovirus species, for instance, to generate recombinants with improved insecticidal characteristics. The method applied in this paper may also be useful for the investigation of naturally occurring genotypic variants in virus isolates and their insecticidal properties. The isolation of SeXD1 confirms a previous observation by *in vivo* cloning of SeMNPV (Muñoz *et al.*, 1998; 1999) that genotypes with different biological and insecticidal properties exist in natural baculovirus isolates.

Materials and Methods

Virus, insects and cells
The SeMNPV-US1 isolate (Gelernter and Federici, 1986b) was originally obtained from Dr. B. A. Federici (Department of Entomology, University of California, Riverside CA) in the form of polyhedra and propagated in fourth instar *S. exigua* larvae (Smits *et al.*, 1988). Cultures of *S. exigua* were reared on artificial diet at 27°C, 70% humidity and a 16:8 h photoperiod. The *S. exigua* cell lines Se301 (Hara *et al.*, 1995b) and SeUCR (Gelernter and Federici, 1986a) were donated by Dr. T. Kawarabata (Institute of Biological Control, Kyushu University, Japan) and Dr. B. A. Federici, respectively. All cells were propagated at 27°C in Grace's supplemented medium containing 10% foetal calf serum (FCS; Gibco). Viral DNA used for the generation of recombinant viruses and restriction endonuclease analysis was extracted from polyhedra produced in *S. exigua* larvae by standard methods (O'Reilly *et al.*, 1992).

Construction of an SeMNPV *p10* promoter-based transfer vector for GFP expression

The SeMNPV 5.1 kbp *Xba*I H-fragment containing the *p10* gene flanked by *p26* and *p74* sequences (Zuidema *et al.*, 1993; IJkel *et al.*, 1999) was used as a basis for the construction of a SeMNPV *p10* promoter-based transfer vector (Fig. 6.1). A 1448 bp *Eco*RI/*Bam*HI fragment was derived from the *Xba*I H-fragment and cloned into pUC19 (pSeMO2). The *Bam*HI and *Xba*I sites located at one end of the insert were both removed by filling in with Klenow resulting in plasmid pSeMO4. The 5' flanking sequence of the SeMNPV *p10* locus, containing the 3' end of the *p26* gene and the *p10* promoter, was isolated by PCR with the forward M13 and a specific antisense primer (5' TCTAGACCTAAG<u>GGATCC</u>TAATGTATAATATAATTAC 3') using pSeMO4 as template. With this PCR a *Bam*HI site was introduced immediately downstream of the adenosine residue of the *p10* translational start codon. The PCR product was cloned into pUC19 as a 513 bp *Eco*RI/*Bam*HI fragment (pSeMO5) and its identity was verified by sequence analysis. A second PCR was performed on pSeMO4 with the reverse M13 primer and a sense primer (5'GGATCCCTTAG G<u>TCTAGA</u>TAAAACTTAACGACGACG 3') to generate the 3' flanking region of the transfer vector containing the *p10* 3' untranslated region and the 3' end of the *p74* gene. With this PCR a *Xba*I site was generated immediately upstream of the *p10* translational stop codon TAA. The PCR product was cloned into pUC19 as a 680 bp *Xba*I/*Hin*dIII fragment (pSeMO6). Sequence analysis showed the correct sequence between the introduced *Xba*I site and the internal *Cla*I site. A three-point ligation was performed to bring the 5' and 3' flanking regions of the *p10* gene together, separated by *Bam*HI and *Xba*I sites. An approximately 3.2 kbp *Eco*RI-*Cla*I fragment of pSeMO4 containing pUC19 sequences and part of the 3' flanking sequence, was combined with the 513 bp *Eco*RI-*Bam*HI fragment of pSeMO5 and the 130 bp *Bam*HI-*Cla*I fragment of pSeMO6 to give pSeMO7. In this new vector the *Bam*HI site is juxtaposed to the *Xba*I site. Finally, a 747 bp *Bam*HI fragment containing the green fluorescent protein (GFP) open reading frame (ORF), derived from pUC19 GFP (Reiländer *et al.*, 1996) was cloned into the *Bam*HI site of pSeMO7 to give the transfer vector pSeXD1.

Generation of a SeMNPV *p10* minus recombinant expressing GFP

A SeMNPV recombinant was generated by injection of viral and transfer vector DNA into the hemocoel of fourth instar *S. exigua* larvae according to Hajós *et al.* (1998) followed by alternate cloning between *S. exigua* larvae and Se301 cells. The injection into insect larvae was performed using a 1.5 ml volume B-D Pen (Becton &

Dickinson) and 28 gauge half-inch NovoFine needles (Novo Nordisk). The injection solution was added to 1.5 ml injector cartridges (Eli Lilly) in a sterile hood (Hajós *et al.*, 1998). Twenty µl of the cotransfection solution containing 0.4 µg circular SeMNPV DNA and 12 µg transfer vector pSeXD1 DNA, and 30% Cellfectin (Gibco-BRL), were injected into the haemocoel of each larva. Haemolymph was obtained from a cut proleg 3 days post transfection and added to 5 ml of serum free Grace's medium containing a few crystals of phenylthiourea, filtered through a 0.45 µm filter (Schleicher & Schuell) and stored at -80°C. The haemolymph filtrate was tested for virus titre and the relative proportion of wild type and recombinant SeMNPV by plaque assay determinations (O'Reilly *et al.*, 1992). The assays were scored for fluorescence under a UV microscope.

Recombinant plaques were selected by their GFP expression and each plaque was diluted with 200 µl Grace's medium without FBS to elute extracellular virus. The virus was amplified in a 24-well plate by adding 100 µl of the plaque eluate to a well with approximately 2×10^4 Se301 cells. Wells with polyhedra-containing cells were harvested 3 days post infection (p.i.) and the cells were suspended in 12 µl distilled water. *S. exigua* third instar larvae were then orally fed after adding the cell suspensions of each well onto *Chrysanthemum* leaf discs with a diameter of 4 mm and placed in 6-well tissue culture plates containing 1 ml 1.5% agarose layer to prevent desiccation. One larva was put in each well with one leaf disc. After consumption of the leaf disc (approx. 16 h) the larvae were placed on artificial diet. Haemolymph was collected at 3 p.i. from larvae showing infection symptoms (lethargy, impared locomotion, pale appearance, no food consumption) and used to measure virus titre and to perform a second round of plaque purification. After three rounds of alternate *in vivo* and *in vitro* cloning, the SeMNPV recombinants were amplified in fourth instar *S. exigua* larvae.

SDS-PAGE and Western analysis
Se301 cells were infected at a m.o.i. of 10 with wt SeMNPV and the recombinant (SeXD1), respectively. Infected cells were harvested at 48 h p.i. and the proteins were analysed by electrophoresis in a 12.5% SDS-polyacrylamide gel using a Bio-Rad Mini-Protein II apparatus. Western blot analysis was performed with a GFP antibody (Molecular Probes) (1:2000 diluted) by standard methods (Sambrook *et al.*, 1989).

PCR, cloning and sequencing
To analyse deletions in the SeMNPV *Pst*I-D fragment, a PCR was performed with the Expand Long Template PCR system (Boehringer Mannheim) using forward primer A

(5' GTAGGGGACGCGAATTTGACTGTTGTTGCAG 3') and reverse primer B (5' CGCACGCTCCACGCTACTCGACTTTGATA 3'), corresponding to nt 17,874 to 17,904 and 29,135 to 29,163 of the SeMNPV genome (IJkel *et al.*, 1999), respectively. The PCR products were cloned into pGEM-T (Promega) and sequence reactions were performed at the Sequencing Core Facility of Eurogentec using universal primers.

Bioassays

The infectivities of wt SeMNPV and recombinant SeXD1 were determined in a leaf disc bioassay as described by Bianchi *et al.* (2000). *Chrysanthemum* leaf discs were prepared using a cork borer with a diameter of 9 mm and placed individually in a 12-well tissue culture plate containing 1 ml 1.5% agarose. Droplets (3 µl) of polyhedra suspensions containing 0 (control), 3×10^3, 10^4, 3×10^4, 10^5, 3×10^5 polyhedra/ml were applied to each leaf disc and dried using a fan. One third instar *S. exigua* larva was added per well. For each dose 36 larvae were used. Larvae that consumed whole leaf disc within 24 h were transferred to a 12-well tissue culture plate containing fresh artificial diet and were further reared at 27°C. Mortality was recorded daily until all larvae had either pupated or died due to SeMNPV infection. The bioassay was performed in three repetitions.

The speed of action of wt SeMNPV and the recombinant SeXD1 was determined in a modified droplet-feeding bioassay (Hughes and Wood, 1981). Third instar *S. exigua* larvae were starved for 16-20 h at 27°C prior to bioassaying. The larvae were allowed to drink from an aqueous suspension containing 10% (w/v) sucrose, 0.001% (w/v) SÄURE-blue and polyhedra at concentrations of 0 (control), 10^3, 3×10^3, 10^4, 3×10^4, 10^5 polyhedra/ml. The first 36 larvae that drank from the solution within 10 min were transferred to individual wells of three 12-well tissue culture plates with fresh artificial diet. Larvae were reared at 27°C, and mortality was recorded every 12 h until all larvae had either pupated or died. The bioassay was performed in four repetitions. Dose mortality data were analyzed with the computer program "POLO" (Russell *et al.*, 1977). For the calculation of LD_{50} values, median ingested volumes of 0.55 µl for third instar *S. exigua* larvae were used as measured by Bianchi *et al.* (2000). Median survival times (ST_{50}) were calculated using the Vistat program (Version 2.1; Boyce Thompson Institute, Cornell University, Ithaca, New York). Log LD_{50} and ST_{50} values were analyzed by regression analysis and t-tests of pairwise differences between treatments with Genstat (Payne *et al.*, 1993).

Chapter 7

General Discussion

Introduction

The research presented in this thesis centered around the unravelling of the genetic and biological properties of the baculovirus SeMNPV. To gain insight in the structure and genomic organization, the full sequence of its DNA genome (135,612 bp) was determined. Furthermore, the role of some selected genes in this genome during the viral infection process has been investigated. Over the past 5 years an increasing number of complete genomic sequences of other large eukaryotic DNA viruses has become available. This allows a comparative genetic and phylogenetic analysis of these viruses on the level of genome structure, gene content and gene arrangement, and thus to determine their genetic relatedness. This chapter will provide a brief overview on the genetic characteristics of large eukaryotic DNA viruses, with special reference to the *Baculoviridae* family, and discuss which types of genes are widely conserved and which genes are more specific on the level of virus family or even virus species to gain insight in common and distinct features of DNA virus genomes.

Characteristics of large viral DNA genomes

Eukaryotic double stranded DNA viruses appear to form a diverse group of pathogens with a wide varation in host range, virulence, virion morphology and genome complexity. Among these viruses a distinct set of families with genomes larger than 100 kb can be discerned (Table 7.1; van Regenmortel *et al.*, 2000). These families comprise viruses infecting vertebrates, invertebrates and algae. The presently unclassified white spot syndrome virus of shrimp (WSSV; Yang *et al.*, 1997) also falls in this category. The *Adenoviridae* and Rhizidovirus with genomes of 25 to 45 kb in size take an intermediate position between the large and small DNA viruses, the latter with genomes of less than 20 kb. Since baculoviruses are large DNA viruses (>100 kb), the DNA viruses with small and intermediate genomes are not further discussed.

The genomes of large DNA viruses generally contain tightly packed, predominantly non-overlapping ORFs, lacking introns, preceded by virus-specific promoters that temporally regulate transcription of an early and a late gene class in a cascaded manner. The early genes are expressed prior to DNA replication and encode

many non-structural proteins, including enzymes involved in genome replication, viral transcription initiation and host response modulation. The late genes are dependent on expression of early genes and mainly encode structural virion proteins. In general, there is no correlation between the transcription polarity or function of a gene and its position in the genome. This is in contrast to the regulation in bacterial genomes, where functionally-related genes tend to cluster along the genome (Huynen *et al.*, 2000).

Table 7.1 Double stranded DNA viruses: genome and virion characteristics.

Family	Genome			Virion		
	Size	Topology	Rep	Morphology	Env	Host
Phycodnaviridae	160-380	linear	N?	Isometric	–	Algae
Iridoviridae	140-303	linear	C	Isometric	–	V/I
Poxviridae	130-375	linear	C	Pleomorphic	+	V/I
WSSV	293	circular	N	Bacilliform	+	I
Polydnaviridae	150-250	circular	N?	rod fusiform	+	I
Herpesviridae	125-240	linear	N	Isometric	+	V
Asfarviridae	170-190	circular	C	Spherical	+	V
Ascoviridae	100-180	linear	N	Reniform	+	I
Baculoviridae	80-180	circular	N	Bacilliform	+	I
Adenoviridae	28-45	linear	N	Isometric	–	V
Rhizidiovirus	27	linear	N	Isometric	–	F

Other double stranded DNA viruses with genome sizes smaller than 15 kb and phages are not listed. Genome sizes in kb. (Env, envelope; I, Invertebrates; V, Vertebrates; F, Fungi; Rep, Replication site; C, Cytoplasm; N, Nucleus; ?, not known).

Genes conserved among all large DNA viruses

Some DNA viruses replicate in the host cell nucleus, while others replicate in the cytoplasm (Table 7.1). The nuclear viruses generally utilize the host's transcriptional apparatus for early gene expression with modulation of virus specified factors, while the cytoplasmatic viruses encode their own transcriptional enzymes. Hence, the latter have to be present in the virus particle to enable the initiation of viral gene expression. The difference in replication site, however, is generally not reflected by the absence of transcription enzymes, such as RNA polymerase and transcription regulator genes, in the nuclear replicating viruses (Hannenhalli *et al.*, 1995; Hayakawa *et al.*, 2001). In

contrast to cellular RNA polymerases, the viral RNA polymerases are capable of enhancing the expression of late genes, which are generally expressed at very high levels as compared to cellular and early viral genes. Thus, genes encoding the RNA polymerase and transcription regulators are present in all DNA viruses and, therefore, can be considered as 'core' genes for large DNA viruses.

Also DNA replication enzymes are commonly conserved between large DNA viruses. These include a DNA polymerase, DNA polymerase processivity factor(s), a major DNA binding protein and components that make up a helicase-primase complex (Hannenhalli *et al.*, 1995; Hayakawa *et al.*, 2001). The viral DNA polymerase is one of the very few gene products, which has sequence motifs conserved among all large DNA viruses and host organisms. Therefore, this gene is frequently used to study the phylogenetic relatedness of viruses (Bulach *et al.*, 1999; Knopf, 1998; Tidona and Darai, 2000; Moser *et al.*, 2001).

All large DNA viruses also possess structural virion core, capsid and tegument proteins as well as proteins involved in cell attachment and entry (Hannenhalli *et al.*, 1995; Hayakawa *et al.*, 2001). They are considered 'core' genes because they are functionally conserved, but usually do not show similarity on amino acid level. A striking example, even within a single genus, is observed in the *Baculoviridae*, where group II NPVs contain a functional homolog of the group I NPVs envelope fusion protein not displaying amino acid similarity (IJkel *et al.*, 2000; Chapter 5). The diversity in these structural proteins reflects the distinct features of virus particles, the different mechanisms used to enter host cells, and the diversity in host ranges.

Genes conserved among the *Baculoviridae*

Nowadays, the complete DNA sequences of three group I NPVs (Ayres *et al.*, 1994; Ahrens *et al.*, 1997, Gomi *et al.*, 1999), of three group II NPVs (IJkel *et al.*, 1999; Kuzio *et al.*, 1999; Chen *et al.*, 2001) and of three GVs (Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, personal communication) have been reported. The overall characteristics of these viruses are shown in Table 7.2. The variation in the baculovirus genome sizes is mainly due to differences in the number of repeated ORFs, unique ORFs and homologous regions (*hrs*) (Table 7.2). The coding density is higher in baculoviruses (on average 1 ORF per kb; Table 7.2) than those found in herpesviruses (0.5 till 0.9 ORF per kb) and poxviruses (on average 0.4 till 0.6 ORF per kb) (Hannenhalli *et al.*, 1995; Isegawa *et al.*, 1999; Vink *et al.*, 2000).

Table 7.2 Baculoviral genome characteristics.

	Px	Cp	Bm	Ha	Op	Ac	Se	Ld	Xc
Size (kb)	101	124	128	131	132	134	**136**	161	179
Coding (%)	86	nk	90	87	89	90	**90**	87	88
Total ORFs	120	143	136	135	152	154	**139**	163	181
Unique ORFs	16	26	1	23	18	11	**16**	31	52
Repeated ORFs	10	nk	9	5	13	4	**9**	32	30
hrs	4	nk	7	5	5	9	**4**	13	9
GC (mol%)	40	nk	40	39	55	41	**44**	58	41
Classification	GV	GV	NPV I	NPV II	NPV I	NPV I	NPV II	NPV II	GV

Se, SeMNPV, IJkel *et al.*, 1999; Ld, LdMNPV, Kuzio *et al.*, 1999; Ha, HaSNPV, Chen *et al.*, 2001; Ac, AcMNPV, Ayres *et al.*, 1994; Bm, BmNPV, Gomi *et al.*, 1999; Op, OpMNPV, Ahrens *et al.*, 1997; Xc, XcGV, Hayakawa *et al.*, 1999; Px, PxGV, Hashimoto *et al.*, 2001, Cp, CpGV, Luque *et al.* personal communication. nk, not known.

Comparison of the baculovirus gene content revealed sixty-three genes conserved among all based on amino acid similarity (Table 7.3). These genes can be considered as 'core' genes of lepidopteran baculoviruses. Approximately half of these genes have assigned functions either in DNA replication and gene expression or in virus structure (Table 7.3), which is consistent with the functions of conserved genes among large DNA viruses in general. The total number of 63 ORFs conserved among the nine baculoviruses compared, is considerably higher than for the 13 herpesviruses which share about 25 conserved ORFs (Hannenhalli *et al.*, 1995; Montague and Hutchison, 2000). Besides the 'core' genes all baculoviruses contain an *alkaline exonuclease*, a *fibroblast growth factor*, a *superoxide dismutase* and an *ubiquitin* gene, which may affect different host cell processes. The functions of the other conserved genes are unknown.

Table 7.3 Functions of SeMNPV genes conserved among all baculoviruses*.

Repl. / Trans.	Virion structure	Auxiliary	Unknown
39K, dbp, helicase, lef1, lef2, lef3, lef4, lef5, lef6, lef8, lef9, lef11, dnapol, ie1, me53, p47, vlf1	*Se8, fp25K, gp41, p74, odv-e18, pk1, odv-e25, p95, odv-ec27, p6.9, odv-e56, vp1034, pol, odv-e66, vp39*	*alk-exo, fgf, sod, v-ubi*	*38K, 38.7K, se35, se36, se50, se53, se59, se60, se69, se72, se73, se78, se79, se81, se90, se92, se94, se95, se108, se118, se128, se133 se134, se137, p12, p40, p45*

Abbreviations of gene names are described in Chapter 2. Repl., Replication; Trans., Transcription.
* Following the classification of O'Reilly in The Baculoviruses (1997).

Genes conserved on genus and group level within the *Baculoviridae*

In addition to the 63 'core' genes, the six NPVs and three GVs share a further set of 15 and 27 genes, respectively, which are not present in the other genus (Table 7.4). These genus-specific genes may underscore the distinct phenotypic characteristics of NPVs and GVs, such as occlusion body morphology and nuclear disintegration. Only a few of these genus-specific genes are characterized – *vp80*, *pp34* and *orf1629* as structural proteins (Funk *et al.*, 1997), *iap2* and *iap4* as inhibitors of apoptosis (Clem and Miller, 1994), *mnpase* as a metalloproteinase (Hayakawa *et al.*, 1999), *pkip1* as a protein kinase interacting factor (Fan *et al.*, 1998) and *arif1* as a cytoskeleton rearrangement inducing factor (Roncarati and Knebel-Mörsdorf, 1997) – and their functions may be associated with the biological differences between NPVs and GVs. A similar correlation is found in NPVs belonging to either group I or group II (Table 7.4).

Table 7.4 ORFs conserved within baculovirus genera and groups.

GVs	NPVs	NPV group I	NPV group II
2, 7, 8, 17, 18, 19, 25, 26, 29, 34, mpnase, 47, 54, 85, 86, 90, 113, 116, 136, iap4, 142, 143, 165, 169, 172, 173, 178	*orf1629, 29, pkip, arif1, 42, 58, 101, 102, 104, 124, p26, 125, iap2, vp80capsid, pp34*	*1, ptp1, odv-e26, iap1, 30, gta, ets, 72, 73, p87, 114, 122, 124, lef7, gp64/67, 132, ie2*	*30, 52, 107*

GVs; ORF numbers are from XcGV. NPVs and NPV group II; ORF numbers are from SeMNPV. NPV group I; ORF numbers are from AcMNPV. Abbreviations of genes are described in IJkel *et al.* (1999) for SeMNPV, in Ayres *et al.* (1994) for AcMNPV and in Hayakawa *et al.* (1999) for XcGV.

Thus, the taxonomic classification of baculoviruses, based on morphological differences (NPV versus GV) and single gene comparisons (group I versus group II NPV), correlates with the presence of a set of genes specific for each group. These group-specific genes are possibly involved in the distinct biological properties of lepidopteran baculovirus groups. The occurrence of the M NPV morphotype, does correlate to genetic relatedness, as one gene Se43 is present in all MNPV genomes, whereas it is absent in HaSNPV. Since, the sequence of only one S NPV is known (Chen *et al.*, 2001), a correlation between the genes specific to the S NPV type can not be investigated yet. It remains to be investigated if the Se54 gene is related to the M NPV morphotype. It is equally possible that the S and M NPV morphotypes are caused by subtile changes in proteins encoded by genes conserved among these viruses.

A correlation between the taxonomic classification and the presence of group-specific genes is also observed for the mammalian and avian herpesviruses (Montague

and Hutchison, 2000). However, the herpesviruses of fish and amphibians, which are very distantly related to those of mammals and birds on the basis of single gene phylogeny, do not seem to have group-specific genes (Davison, 1992 and 1998; Davison *et al.*, 1999). Thus, it seems that the correlation of group-specific genes with the taxonomic classification is not a common feature for all large DNA viruses and so far an exceptional feature for baculoviruses and some herpesviruses. Since alpha- and beta-herpesviruses likely cospeciated with their hosts (McGeoch and Cook, 1994; McGeoch *et al.*, 1995; McGeoch *et al.*, 2000), this could suggest that baculoviruses also co-evolved with their hosts.

Prediction of protein function by comparative genome analysis

The phylogenetic profile of a gene indicates which of the genomes compared code for homologs of this gene (Tatusov *et al.*, 1997). For each gene a phylogentic profile can be constructed. Comparison of these profiles could reveal clusters of genes with identical profiles. Bacterial genes with identical phylogenetic profiles tend to have related functions (Pellegrini *et al.*, 1999). Furthermore, it was speculated that some herpesvirus genes with identical phylogenetic profiles may determine the cell-type specificity of the latent state of beta-herpesviruses (Montague and Hutchison, 2000). Therefore, baculovirus genes with an identical phylogenetic profile may also have related functions based on the assumption that proteins that function together in a pathway or complex are likely to evolve in a correlated fashion. The baculovirus genes which are conserved among all, on genus or on group level, were excluded. Comparison of the conservation of the other baculovirus genes revealed eight identical phylogenetic profiles that occured at least twice (Table 7.5). It is well possible that the genes within each profile have biochemically related functions. In line with this supposition, an interaction of *chitinase* and *v-cathepsin* has recently been demonstrated (Hom and Volkman, 2000). Furthermore, an interaction between *ribonucleotide reductase-1* (*rr*) *1* and *rr2* as well as between *dnaligase* and *helicase2* is likely. Thus, phylogenetic profiles can be helpful to predict the function of some but not all baculovirus genes.

Table 7.5 Genes with special phylogenetic profiles.

Genetic Profile	Genes
All except PxGV	*gp37, lef10, v-cath, chit*
All except CpGV	*p10, ie0*
NPVs group I and SeMNPV	*se37, se84*
NPVs group I and LdMNPV	*ac4, ac11*
NPVs group I and HaSNPV	*lef12, ac74*
NPVs except LdMNPV	*se103, se111, se113, se47, cg30, gp16*
In OpMNPV, SeMNPV, LdMNPV and CpGV only	*se54, rr1, rr2*
GVs and LdMNPV	*dna-ligase, helicase2*

ORF numbers and abbreviations in Ayres *et al.* (1994) for AcMNPV and in IJkel *et al.* (1999) for SeMNPV.

Genes unique to a single virus species

All baculoviruses sequenced so far contain genes that are unique to a single virus species. Their number ranges between 1 for BmNPV and 52 for XcGV (Table 7.2) and might be somewhat biased as three closely 'related' group I and three less 'related' group II NPVs have been sequenced. Comparison of the unique SeMNPV genes to the recently sequenced baculovirus genomes (XcGV, PxGV and HaSNPV) revealed that Se17/18, Se21 and Se68 share homology to Xc129, Xc128 and Ha83, respectively. Thus, SeMNPV contains as yet 16 unique genes instead of 20 as reported by IJkel *et al.* (1999). Approximately half of the total of ~425 baculovirus genes identified are genes unique to baculovirus species. It is tempting to speculate that acquisition of these unique genes from the host may contribute to baculovirus speciation by causing alterations in host range and or virulence.

Gene order conservation among baculoviruses

In addition to single gene comparisons, the arrangement of overall conserved genes was also used to study qualitatively the relatedness of lepidopteran baculoviruses (Chapter 2). Several clusters of conserved genes were identified between baculoviruses. The genes within a cluster are conserved in order and direction of transcription, but the clusters can have different orientation and arrangement among
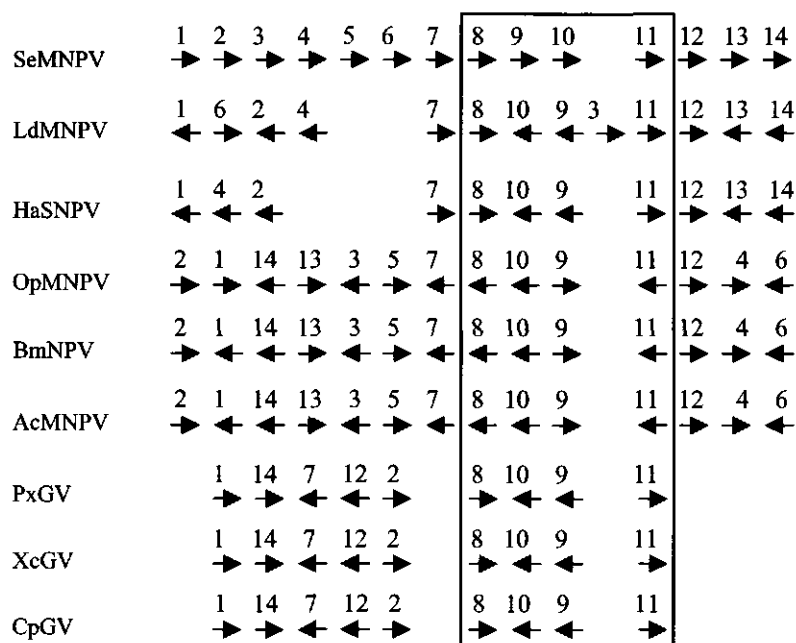
Fig. 7.1 Arrangement of gene cluster in baculoviruses.

The numbers of the clusters and the ORFs within are derived from IJkel *et al.* (1999). Arrows indicate the orientation of cluster. The box indicates the 'core' clusters of baculoviruses.

baculoviruses (Fig. 7.1). One of these clusters (cluster 14), encompassing Se132 - Se138, contains the only spliced baculoviral gene (*ie-1*; Se132; van Strien *et al.*, 2000). The conservation of this region may be due to preservation of either splicing signals or the exon of the gene (Se138) to which it is spliced.

Based upon the number of gene clusters, their length and relative order it was concluded that SeMNPV is most closely related to LdMNPV (IJkel *et al.*, 1999). Comparison of the genomic organization between SeMNPV and HaSNPV reveals that these viruses are also very closely related (Fig. 7.1; Chen *et al.*, 2001). This is in agreement with the phylogenetic analysis of individual genes such as *egt*, *lef-2*, *dnapol* and *rr* (Chen *et al.*, 1997 and 1999; Bulach *et al.*, 1999; van Strien *et al.*, 1997). When the order of gene clusters is taken to represent the baculovirus genome organization, the common structure of group II baculoviruses becomes apparent (Fig. 7.1). Within each group, the structural difference is relatively small and predominantly determined by inversions of gene clusters as well as inversions of individual genes. The latter is surprising, considering that the non-coding region between most baculovirus genes is on average only 100 nt and random inversion would therefore, in most cases result in disruption of genes.

Although the baculovirus genomes show a high degree of variability in their relative cluster order, one region of partial conservation appears within their genomes. This region consists of four clusters (8-9-10-11) and is located in the 'central' region (Se62 – Se108) of baculovirus genomes. The maintenance of this gene cluster conservation suggests that this region may play a critical role in gene expression and genome replication. Comparison of the two NPV groups showed extensive genomic translocations in addition to cluster inversions. A common genome structure for group I and II viruses can be derived, showing a major inversion of a genomic segment containing the clusters 4-6-2-1-13-14 (Fig. 7.1). The relative order of gene clusters seems to be more scrambled among group II NPVs than among group I NPVs and GVs (Fig. 7.1). Whether this gene cluster variablity is a special feature for group II NPVs or occurred by mere chance will become clear when more group II NPVs genomes are analyzed. Furthermore, statistical programmes need to be developed or applied to quantify the relationship on the basis of gene order.

Gene clusters are not an exceptional feature of baculoviruses since also herpesviruses contain blocks of genes conserved in order and polarity. Among the herpesviruses of mammals and birds a subset of about 25 genes is conserved, which is arranged in seven gene clusters (Hannenhalli *et al.*, 1995). Like in baculovirus clusters, the gene clusters have different orders and orientations in different herpesvirus subfamilies but genes within a given cluster maintain order and transcriptional polarity (Gompels *et al.*, 1995). The conserved herpesvirus genes encode capsid proteins, components of the DNA replication complex, transcription regulators, nucleotide-modifying enzymes, membrane proteins, and tegument proteins, hence most of them are 'core' genes of large DNA viruses. However, the herpesviruses of fish and amphibians are very distant from those of mammals and birds and analysis of the sequences of these viruses has so far not revealed common gene clusters (Minson *et al.*, 2000). Thus, gene clustering is observed in several DNA viruses, but seems not to be a common feature of all large DNA viruses. Nonetheless, the relatedness between baculoviruses based on gene clusters is in agreement with their taxonomic classification and with single gene comparisons and therefore, can be used as an independent marker of baculovirus relatedness.

Genes with scrambled genome positions

Comparison of the position of conserved genes in baculovirus genomes revealed a number of genes, whose genomic position is extremely scrambled. These genes can be divided in two categories. The first category includes genes with multiple copies in a

single genome, such as *odv-e66, p26, p10, ptp, ctl, rr2, dbp, fgf, helicase, enhancins, iaps* and *bro* genes (Hayakawa *et al.*, 2001). At least one of the copies will show a scrambled genomic position. These repeated genes could either have been acquired independently or have been duplicated during multiplication followed by diversion and rearrangement. The second category possesses single copy genes, such as *pk-2, pkip, p35, sod, odv-ec27, v-ubi, rr1, dutpase, alk-exo, lef-2, cathepsin, chitinase, gp37, egt, arif1, pp34, glycogenin, p94, lef8, p47* and *me53*, with a wide variety in function.

A number of the genes with scrambled genomic positions has a cellular homolog (Elledge *et al.*, 1992; Guarino, 1990; van Strien *et al.*, 1997). Since some of these genes are more closely related to eukaryotic than prokaryotic homologs, it is well possible that they have been acquired from cellular sources (van Strien *et al.*, 1997; Kuzio *et al.*, 1999). A similar situation is observed in herpesviruses, where the majority of genes having cellular homologs is located in between the seven conserved gene clusters (Megaw *et al.*, 1998; Minson *et al.*, 2000). However, in the latter the *rr* gene is located in a conserved gene cluster indicating that exceptions to this broad generalization certainly exist (Minson *et al.*, 2000).

Genes possessing properties of cellular homologs are observed in all large DNA viruses and their wide spectrum most likely reflects the extent and diversity of the host responses (Tidona and Darai, 2000; Raftery *et al.*, 2000; Bugert and Darai, 2000). Since baculoviruses in contrast to other large DNA viruses lack genes that display homology to the humoral and cellular immune system genes in vertebrates, it seems likely that such systems, if present, do not play an important role in the invertebrate host response upon baculovirus infection.

Baculovirus genes as possible determinants of specific biological properties

Comparative baculovirus genomics may reveal gene determinants for specific biological properties, such as host range and virulence. An AcMNPV mutant containing the LdMNPV host range factor-1 (*hrf-1*) gene showed an expanded host range towards otherwise non-permissive Ld-652Y cells (Thiem *et al.*, 1996; Du and Thiem, 1997). Furthermore, the AcMNPV *p35* deletion mutant causes cell line-specific apoptosis after infection in contrast to the wild-type (Griffiths *et al.*, 1999). These observations support the view that gene products can influence the host range of a baculovirus.

To investigate possible correlations between the presence of specific genes and biological characteristics of baculoviruses, the gene content of AcMNPV that can

infect *S. exigua*, *O. pseudotsugata*, *H. armigera* and *P. xylostella* larvae (Adams and McClintock, 1991) was compared to SeMNPV, HaSNPV, OpMNPV and PxGV, which only can infect a single insect species within the AcMNPV host range. Twenty genes are present in AcMNPV, while absent in the other viruses compared, and these may correlate with the wider host range of AcMNPV (Table 7.6). The observed functions for some of these genes, such as a cell-specific factor required for transcription (*hcf-1*; Lu and Miller, 1996), the polynucleotidekinase / RNA ligase (*pnk/pnl*; Durantel *et al.*, 1998a), the protein kinase (*pk2*; Dever *et al.*, 1998) and the antiapoptotic inhibitor *p35* (Griffiths *et al.*, 1999), indicate that they could play a role in modulating host stress responses. The narrow host ranges of SeMNPV, HaSNPV, OpMNPV and PxGV can not be correlated to the presence of genes in these four viruses. Whether the limited host range of SeMNPV infecting only *S. exigua* and its relatively high virulence can be correlated to the presence of unique genes or to conserved genes with altered properties remains to be investigated.

Table 7.6 Possible gene determinants of the broad AcMNPV host range.

AcMNPV ORFs	7, 12, 39 (p43), 45, 58, 70 (hcf-1), 84, 86 (pnk/pnl), 97, 107, 112, 113,
Ayres *et al.* (1994)	118, 121, 123 (pk2), 135 (p35), 140, 149, 152, 154

Although comparative genomics may identify candidate genes involved in specific biological properties of baculoviruses, caution should be taken since amino acid changes in overall conserved proteins can also influence host range as shown for the AcMNPV helicase gene (Kondo and Maeda, 1991; Croizier *et al.*, 1994). In addition, functional homologs that do not display similarity on amino acid level, such as the envelope fusion proteins (EFP) of group I and II baculoviruses, will be misinterpreted. Finally, host factors may play an important role in the determination of the virus host range, as observed for human herpesvirus-4 where cell surface receptors are involved (Aubry *et al.*, 1993). Similarly, the tissue and or host range of group II NPVs could be restricted by host factors compared to group I NPVs, as the group II EFP may require cleavage by a cellular furin-like proprotein convertase (Chapter 5).

Future aspects

The elucidation of the SeMNPV genome and comparison to other baculoviruses as described in this thesis contribute significantly to the emerging genome-wide view on their patterns of organization and structure (Chapter 2 and 7). The more focussed

study on two genes unique to SeMNPV, Se116 and Se117 (Chapter 3), the latter being an ODV-specific nucleocapsid protein, and on Se17/18 having a homolog in a distantly related granulovirus (Chapter 4), is an initial step to understand their function(s). The function of the novel BV envelope fusion protein identified in SeMNPV is characterized and an interaction with a cellular furin-like proprotein convertase is reported (Chapter 5). It is anticipated that this envelope fusion protein is a good marker for group II NPVs (Chapter 5; Pearson *et al.*, 2000). To investigate the functions of these genes in more detail, a major technical difficulty hampering the generation of SeMNPV recombinants was resolved (Chapter 6). This novel approach can now be applied to characterize the functions of all SeMNPV genes by constructing site-specific or null mutants and study their phenotype(s) *in vivo* as well as in cell culture.

Mutation analysis and biochemical techniques applied towards proteins, such as the yeast two-hybrid system and two-dimensional gel electrophoresis, are unfortunately often tedious, labour intensive, expensive and incomplete. This can be enlighted using computational methods that assign function by sequence similarity (Fischer and Eisenberg, 1997; Koonin *et al.*, 1998; Bork and Koonin, 1998; Bork *et al.*, 1998; Przytycka *et al.*, 1999) or other shared characteristics, such as identical phylogenetic profiles (Tatusov *et al.*, 1997; Pellegrini *et al.*, 1999; Montague and Hutchison, 2000), domain-fusion analysis (Marcotte *et al.*, 1999) and correlated mRNA expression profiles under the same series of conditions (Eisen *et al.*, 1998). However, even a combination of these approaches will not reveal the function for the majority of the SeMNPV genes nor their interaction with host factors as there are many exceptions to the assumptions underlying these methods.

An approach that is a step closer to assign the biological function of baculovirus genes is proteomics. A promising technique is mass-spectrometric identification of gel-separated proteins (Pandey and Mann, 2000). This sensitive technique could be applied to identify the presence of multiple baculovirus proteins at different times post infection and in different tissues. This would provide insight at the timing of protein function and possible protein-protein interactions. The interaction of virus proteins with each other or with host proteins could be investigated by purifying entire multi-protein complexes by affinity-based methods, such as glutathione S-transferase-fusion proteins, antibodies, DNA, RNA, or a small molecule binding specifically to a cellular target. For instance, a bait protein carrying a tag or epitope can be overexpressed in cells and afterwards be immunnoprecipitated by an antibody against the epitope. The precipitated factors can than be identified by mass spectrometry. With the information on the complete SeMNPV genome as well as other baculovirus genomes available,

these approaches can be applied to study the function of the encoded proteins. This will give further insight in the baculoviral strategies to modulate host response pathways and in the principles underlying their specific biological properties. Ultimately, these functional annotations can be used to further improve SeMNPV as a biological control agent and more recently as a gene delivery factor in gene therapy (Merrihew *et al.*, 2001; Van Loo *et al.*, 2001).

References

Adams, J. R., and McClintock, J. T. (1991). Nuclear polyhedrosis viruses of insects. *In* "Atlas of Invertebrate Viruses", pp. 87-204. Edited by J.R. Adams and J.R. Bonami. Boca Raton, FL: CRC Press.

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

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

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25, 3389-3402.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.

Appel, R. D., Bairoch, A., and Hochstrasser, D. F. (1994). A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends Biochem. Sci.* 19, 258-260.

Aubry, J. P., Shields, J. G., Jansen, K. U., and Bonnefoy, J. Y. (1993). A multiparameter flow cytometric method to study surface molecules involved in interactions between subpopulations of cells. *J. Immunol. Meth.* 159, 161-171.

Ausabel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994). *In* "Current protocols in molecular biology." Greene and Wiley-Interscience, New York.

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

Bairoch, A., Buchner, P., and Hofmann, K. (1997). The PROSITE database; its status in 1997. *Nucl. Acids Res.* 25, 217-221.

Beinert, H., and Holm, R. H. (1997). Iron-sulphur clusters: nature's modular multipurpose structures. *Science* 277, 653-659.

Bianchi, F. J., Snoeijing, I., Van der Werf, W., Mans, R. M., Smits, P. H., and Vlak, J. M. (2000). Biological activity of SeMNPV, AcMNPV, and three AcMNPV deletion mutants against *Spodoptera exigua* larvae (Lepidoptera: Noctuidae). *J. Invert. Pathol.* 75, 28-35.

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

Bischoff, D. S., and Slavicek, J. M. (1997). Molecular analysis of an *enhancin* gene in the *Lymantria dispar* nuclear polyhedrosis virus. *J. Virology* 71, 8133-8140.

Bishoff, D. S., and Slavicek, J. M. (1995). Identification and characterization of an early gene in the *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus. *J. Gen. Virology* 76, 2933-2940.

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

Bjornson, R. M., and Rohrmann, G. F. (1992). Nucleotide sequence of the *polyhedron envelope protein* gene region of the *Lymantria dispar* nuclear polyhedrosis virus. *J. Gen. Virology* 73, 1499-1504.

Blinov, V. M., Gutorov, V. V., Holodilov, N. G., Iljichev, A. A., Karginov, V. A., Mikrjukov, N. N., Mordvinov, V. A., Nikonov, I. V., Petrov, N. A., Urmanov, I. H., and Vasilenko, S. K. (1984). Nucleotide sequence of the *Galleria mellonella* nuclear polyhedrosis virus origin of DNA replication. *FEBS Lett.* 167, 254-258.

Blissard, G. W., Black, B., Crook, N., Keddie, B. A., Possee, R., Rohrmann, G., Theilmann, D. A., and Volkman, L. E. (2000). Family *Baculoviridae*. *In* "Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses" (M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wicker, Eds.), pp. 195-202. Academic press, New York.

Blissard, G. W., and Wenz, J. R. (1992). Baculovirus GP64 envelope glycoprotein is sufficient to mediate pH-dependent membrane fusion. *J. Virology* 66, 6829-6835.

Blissard, G.W., and Rohrmann, G. F. (1991). Baculovirus *gp64* gene expression: analysis of sequences modulating early transcription and transactivation by IE1. *J. Virology* 65, 5820-5827.

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

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

Bork, P., and Koonin, E. V. (1998). Predicting functions from protein sequences--where are the bottlenecks? *Nat. Genet.* 18, 313-318.

Bork, P., Dandekar, T., Diaz-Lazcoz, Y., Eisenhaber, F., Huynen, M., and Yuan, Y. (1998). Predicting function: from genes to genomes and back. *J. Mol. Biol.* 283, 707-725.

Braunagel, S. C., and Summers, M. D. (1994). *Autographa californica* nuclear polyhedrosis virus, PDV, and ECV viral envelopes and nucleocapsids: structural proteins, antigens, lipid and fatty acid profiles. *Virology* 202, 315-328.

Broer, R., Heldens, J. G. M., van Strien, E. A., Zuidema, D., and Vlak, J. M. (1998). Specificity of multiple homologous genomic regions in *Spodoptera exigua* nucleopolyhedrovirus DNA replication. *J. Gen. Virology* 79, 1563-1572.

Bugert, J. J., and Darai, G. (2000). Poxvirus homologues of cellular genes. *Virus Genes* 21, 111-133.

Bulach, D. M., Kumar, C. A., Zaia, A., Liang, B. F., and Tribe, D. E. (1999). Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences. *J. Invert. Pathol.* 73, 59-73.

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

Carstens, E. B., Lu, A. L., and Chan, H. B. (1993). Sequence, transcriptional mapping, and overexpression of *p47*, a baculovirus gene regulating late gene expression. *J. Virology* 67, 2513-2520.

Carstens, E. B., Tjia, S. T., and Doerfler, W. (1979). Infection *of Spodoptera frugiperda* cells with *Autographa californica* nuclear polyhedrosis virus: Synthesis of intracellular proteins after virus infection. *Virology* 99, 368-398.

Chen, C. J., and 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.

Chen, X., IJkel, W. F. J., Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, D., Lankhorst, R. K., Vlak, J. M., and Hu, Z. (2001). The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. *J. Gen. Virology* 82, 241-257.

129

Chen, X., Sun, X., Hu, Z.H., Li, M., O'Reilly, D.R., Zuidema, D., and Vlak, J.M. (2000). Genetic engineering of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrocirus as an improved pesticide. *J. Invert. Pathol.* 76, 140-146.

Chen, X., IJkel, W. F. J., Dominy, C., Zanotto, P. de, Hashimoto, Y., Faktor, O., Hayakawa, T., Wang, C. H., Krell, P. J., Hu, Z., and Vlak, J. M. (1999). Identification, sequence and phyologeny of the *lef-2* gene of *Heliothis armigera* single-nucleocapsid baculovirus. *Virus Research* 65, 21-32.

Chen, X., Hu, Z, Jehle, J. A., Zhang, Y., and Vlak, J. M. (1997). Analysis of the *ecdysteroid UDP-glucosyltransferase* gene of *Heliothis armigera* single-nucleocapsid baculovirus. *Virus Genes* 15, 219-225.

Cherbas, L., and Cherbas, P. (1993). The arthropod initiator: the capsite consensus plays an important role in transcription. *Insect Biochem. Mol. Biol.* 23, 81-90.

Choi, J. Y., Woo, S. D., Je, Y. H., and Kang, S. K. (1999). Development of a novel expression vector system using *Spodoptera exigua* nucleopolyhedrovirus. *Molecules and Cells* 9, 504-509.

Cieplik, M., Klenk, H. D., and Garten, W. (1998). Identification and characterization of *Spodoptera frugiperda* furin: A thermostable subtilisin-like endoprotease. *Biological Chemistry* 379, 1433-1440.

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

Clem, R. J., Robson, M., and 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*. *J. Virology* 68, 6759-6762.

Croizier, G., Croizier, L., Argaud, O., and Poudevigne, D. (1994). Extension of *Autographa californica* nuclear polyhedrosis virus host range by interspecific replacement of a short DNA sequence in the *p143* helicase gene. *Proc. Natl. Acad. Sci. USA* 91, 48-52.

Crook, N. E., Clem, R. J., and Miller, L. K. (1993). An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J. Virology* 67, 2168-2174.

Dall, D., Sriskantha, A., Vera, A., Lai-Fook, J., and Symonds, T. (1993). A gene encoding a highly expressed spindle body protein of *Heliothis armigera* entomopoxvirus. *J. Gen. Virology* 74, 1811-1818.

Davis, S. J., and Vierstra, R. D. (1996). Soluble derivatives of green fluorescent protein (GFP) for use in *Arabidopsis thaliana*. *Weeds World* 3, 43-48.

Davison, A. J., Sauerbier, W., Dolan, A., Addison, C., and McKinnell, R. G. (1999). Genomic studies of the Lucke tumor herpesvirus (RaHV-1). *J. Cancer Res. Clin. Oncol.* 125, 232-238.

Davison, A. J. (1998). The genome of salmonid herpesvirus 1. *J. Virology* 72, 1974-1982.

Davison, A. J. (1992). Channel catfish virus: a new type of herpesvirus. *Virology* 186, 9-14.

Derksen, A. C., and Granados, R. R. (1988). Alteration of a lepidopteran peritrophic membrane by baculoviruses and enhancement of viral infectivity. *Virology* 167, 242-250.

Dever, T. E., Sripriya, R., McLachlin, J. R., Lu, J., Fabian, J. R., Kimball, S. R., and Miller, L. K. (1998). Disruption of cellular translational control by a virus truncated eukaryotic translation initiation factor 2α kinase homolog. *Proc. Natl. Acad. Sci. USA* 95, 4164-4169.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* 12, 387-395.

Du, X., and Thiem, S. M. (1997). Responses of insect cells to baculovirus infection: protein synthesis shutdown and apoptosis. *J. Virology* 71, 7866-7872.

Durantel, D., Croizier, L., Ayres, M. D., Croizier, G., Possee, R. D., and Lopez-Ferber, M. (1998a). The *pnk/pnl* gene (ORF86) of *Autographa californica* nucleopolyhedrovirus is a non-essential, immediate early gene. *J. Gen. Virology* 79, 629-637.

Durantel, D., Croizier, G., Ravallec, M., and Lopez-Ferber, M. (1998b). Temporal expression of the AcMNPV *lef-4* gene and subcellular localization of the protein. *Virology* 241, 276-284.

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA.* 95, 14863-14868.

Elledge, S. J., Zhou, Z., and Allen, J. B. (1992). Ribonucleotide reductase: regulation, regulation, regulation. *Trends Biochem. Sci.* 17, 119-123.

Fabian, P., Murvai, J., Hatsagi, Z., Vlahovicek, K., Hegyi, H., and Pongor, S. (1997). The SBASE protein domain library, release 5.0: a collection of annotated protein sequence segments. *Nucl. Acids Res.* 25, 240-243.

Fan, X., McLachlin, J. R., and Weaver, R. F. (1998). Identification and characterization of a protein kinase-interacting protein encoded by the *Autographa californica* nuclear polyhedrosis virus. *Virology* 240, 175-182.

Faulkner, P., Kuzio, J., Williams, G. V., and Wilson, J. A. (1997). Analysis of *p74*, a PDV envelope protein of *Autographa californica* nucleopolyhedrovirus required for occlusion body infectivity *in vivo. J. Gen. Virology* 78, 3091-3100.

Federici, B. A. (1997). Baculovirus pathogenesis. *In* "The baculoviruses" (L.K. Miller, Ed.), pp. 33-59. Plenum Press, New York.

Federici, B. A., and Hice, R. H. (1997). Organization and molecular characterization of genes in the *polyhedrin* region of the *Anagrapha falcifera* multinucleocapsid NPV. *Arch. Virology* 142, 333-348.

Federici, B. A. and Maddox, J. V. (1996). Host specificity in microbe-insect interactions. *Bioscience* 46, 410-421.

Fischer, D., and Eisenberg, D. (1997). Assigning folds to the proteins encoded by the genome of *genitalium. Proc. Natl. Acad. Sci. USA* 94, 11929-11934.

Flipsen, J. T. M., Mans, R. M., Kleefsman, A. W., Knebel-Morsdorf, D., and Vlak, J. M. (1995). Deletion of the baculovirus *ecdysteroid UDP-glucosyltransferase* gene induced early degeneration of Malpighian tubules infected insects. *J. Virology* 69, 4529-4532.

Flipsen, J. T. M. (1995). Systemic infection of *Spodoptera exigua* larvae using a marked *Autographa californica* nuclear polyhdrosis virus recombinant. *In* "Pathogenesis induced by (recombinant) baculovirus in insects", pp. 51-64. Phd Thesis, Wageningen University.

Friesen, P. D. (1997). Regulation of baculovirus early gene expression. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 141-170. Plenum, New York.

Friesen, P. D. (1993). Invertebrate transposable elements in the baculovirus genome: characterization and significance. *In* "Parasites and Pathogens of Insects: Parasites", pp. 147-178. Edited by N. E. Beckage, S. N. Thompson and B. A. Federici. London: Academic Press.

Friesen, P. D., and Miller, L. K. (1987). Divergent transcription of early 35- and 94-kilodalton protein genes encoded by the *Hind*III K genome fragment of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virology* 61, 2264-2272.

Funk, C. J., Braunagel, S. C., and Rohrmann, G. F. (1997). Baculovirus structure. *In* "The Baculoviruses" (L. K. Miller, Ed.), pp. 7-32. Plenum, New York.

Gelernter, W. D., and Federici, B. A. (1986a). Continuous cell line from *Spodoptera exigua* (Lepidoptera: Noctuidae) that supports replication of nuclear polyhedrosis viruses from *Spodoptera exigua* and *Autographa californica. J. Invert. Pathol.* 48, 199-207.

Gelernter, W. D., and Federici, B. A. (1986b). Isolation, identification and determination of virulence of a nuclear polyhedrosis virus from the beet armyworm, *Spodoptera exigua* (Lepidoptera: Noctuidae). *Environ. Entomol.* 15, 240-245.

Goenka, S., and Weaver, R. F. (1996). Cellular localization of P26, an early AcMNPV gene product. In *"Abstracts of the 15th Annual Meeting of the American Society of Virology"*, p. 154.

Gomi, S., Majima, K., and Maeda, S. (1999). Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus. *J. Gen. Virology* 80, 1323-1337.

Gomi, S., Zhou, C. E., Yih, W., Majima, K., and Maeda, S. (1997). Deletion analysis of four of eighteen late gene expression factor gene homologues of the baculovirus, BmNPV. *Virology* 230, 35-47.

Gompels, U.A., Nicholas, J., Lawrence, G., Jones, M., Thomson, B. J., Martin, M. E., Efstathiou, S., Craxton, M., and Macaulay, H. A. (1995). The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution. *Virology* 209, 29-51.

Gopalakrishnan, B., Muthukrishnan, S., and Kramer, K. J. (1995). Baculovirus-mediated expression of a *Manduca sexta chitinase* gene: Properties of the recombinant protein. *Insect Biochem. Mol. Biol.* 25, 255-265.

Goto, C., Minobe, Y., and Iizuka, T. (1992). Restriction endonuclease analysis and mapping of the genomes of granuloviruses isolated from *Xestia c-nigrum* and five other noctuid species. *J. Gen Virology.* 73, 1491-1497.

Gotoh, B., Ohnishi, Y., Inocencio, N. M., Esaki, E., Nakayama, K., Barr, P. J., Thomas, G., and Nagai, Y. (1992). Mammalian subtilisin-related proteinases in cleavage activation of the paramyxovirus fusion glycoprotein: superiority of furin/PACE to PC2 or PC1/PC3. *J. Virology* 66, 6391-6397.

Granados, R. R. (1978). Early events in the infection of *Heliothis zea* midgut cells by a baculovirus. *Virology* 90, 170-174.

Granados, R. R., and Lawler, K. A. (1981). *In vivo* pathway of *Autographa californica* baculovirus invasion and infection. *Virology* 108, 297.

Griffiths, C. M., Barnett, A. L., Ayres, M. D., Windass, J., King, L. A., and Possee, R. D. (1999). *In vitro* host range of *Autographa californica* nucleopolyhedrovirus recombinants lacking functional *p35*, *iap1* or *iap2*. *J. Gen. Virology* 80, 1055-1066.

Guarino, L. A., Xu, B, Jin, J., and Dong, W. (1998). A virus-encoded RNA polymerase purified from baculovirus-infected cells. *J.Virology* 72, 7985-7991.

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

Hajós, J. P., Zuidema, D., Kulcsár, P., Heldens, J. G. M., Závodszky, P., and Vlak, J. M. (1998). Recombination of baculovirus DNA following lipofection of insect larvae. *Arch. Virology* 143, 2045-2050.

Hannenhalli, S., Chappey, C., Koonin, E., and Pevzner, P. A. (1995). Genome sequence comparison and scenarios for gene rearrangements: a test case. *Genomics* 30, 299-311.

Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H. D., and Garten, W. (1992). Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. *Nature* 360, 358-361.

Hara, K., Funakoshi, M., and Kawarabata, T. (1995a). *In vivo* and *in vitro* characterization of several isolates of *Spodoptera exigua* nuclear polyhedrosis virus. *Acta Virology* 39, 215-222.

Hara, K., Funakoshi, M., and Kawarabata, T. (1995b). A cloned cell line of *Spodoptera exigua* has a highly increased susceptibility to the *Spodoptera exigua* nuclear polyhedrosis virus. *Can. J. Microbiology* 4, 1111-1116.

Hara, K., Funakoshi, M., Tsuda, K., and Kawarabata, T. (1993). New *Spodoptera exigua* cell lines susceptible to *Spodoptera exigua* nuclear polyhedrosis virus. *In Vitro Cellular Developmental Biology* 29A, 904-907.

Hashimoto, Y., Hyakawa, T., Ueno, Y., Fujita, T., Sano, Y., and Matsumoto, T. (2000). Sequence analysis of the *Plutella xylostella* granulovirus genome. *Virology* 275, 358-372.

Hashimoto, Y., Corsaro, B. G., and Granados, R. R. (1991). Location and nucleotide sequence of the gene encoding the viral enhancing factor of the *Trichoplusia ni* granulosis virus. *J. Gen. Virology* 72, 2645-2651.

Hawtin, R. E., Zarkowska, T., Arnold, K., Thomas, C. J., Gooday, G. W., King, L. A., Kuzio, J. A., and Possee, R. D. (1997). Liquefaction of *Autographa californica* nucleopolyhedrovirus-infected insects is dependent on the integrity of virus-encoded *chitinase* and *cathepsin* genes. *Virology* 238, 243-253.

Hayakawa, T., Rohrmann, G. F., and Hashimoto, Y. (2001). Patterns of genome organization and content in lepidopteran baculoviruses. *Virology* 278, 1-12.

Hayakawa, T., Ko, R., Okano, K., Seong, S.-I., Goto, C., and Maeda, S. (1999). Sequence analysis of the *Xestia c-nigrum* granulovirus genome. *Virology* 262, 277-297.

Hayakawa, T., Xu, J., and Hukuhara, T. (1996). Cloning and sequencing of the gene for an enhancing factor from *Pseudaletia separata* entomopoxvirus. *Gene* 177, 269-270.

Hefferon, K. L., Oomens, A. G., Monsma, S. A., Finnerty, C. M., and Blissard, G. W. (1999). Host cell receptor binding by baculovirus GP64 and kinetics of virion entry. *Virology* 258, 455-468.

Heldens, J. G. M., Broer, R., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1997a). Identification and functional analysis of a non-*hr* origin of DNA replication in the genome of *Spodoptera exigua* multicapsid nucleopolyhedrovirus. *J. Gen. Virology* 78, 1497-1506.

Heldens, J. G. M., Liu, Y., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1997b). Characterization of a putative *Spodoptera exigua* multicapsid nucleopolyhedrovirus *helicase* gene. *J. Gen. Virology* 78, 3101-3114.

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

Helenius, A., Marsh, M., and White, J. (1982). Inhibition of Semliki forest virus penetration by lysosomotropic weak bases. *J. Gen. Virology* 58, 47-61.

Hill, J. E., and Faulkner, P. (1994). Identification of the *gp67* gene of a baculovirus pathogenic to the spruce budworm, *Choristoneura fumiferana* multinucleocapsid nuclear polyhedrosis virus. *J. Gen. Virology* 75, 1811-1813.

Hohman, A. W., and Faulkner, P. (1983). Monoclonal antibodies to baculovirus structural proteins: determination of specificities by Western blot analysis. *Virology* 125, 432-444.

Hom, L. G., and Volkman, L. E. (2000). *Autographa californica* M Nucleopolyhedrovirus *chiA* is required for processing of *v-cath*. *Virology* 277, 178-183.

Hong, T., Summers, M. D., and Braunagel, S. C. (1997). N-terminal sequences from *Autographa californica* nuclear polyhedrosis virus envelope proteins ODV-E66 and ODV-E25 are sufficient to direct reporter proteins to the nuclear envelope, intranuclear microvesicles and the envelope of occlusion derived virus. *Proc. Natl. Acad. Sci. USA* 94, 4050-4055.

Hong, T., Braunagel, S. C., and Summers, M. D. (1994). Transcription, translation, and cellular localization of PDV-E66: a structural protein of the PDV envelope of *Autographa californica* nuclear polyhedrosis virus. *Virology* 204, 210-222.

Horton, H. M., and Burand, J. P. (1993). Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion. *J. Virology* 67, 1860-1868.

Hosaka, M., Nagahama, M., Kim, W. S., Watanabe, T., Hatsuzawa, K., Ikemizu, J., Murakami, K., and Nakayama, K. (1991). Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. *J. Biol. Chem.* 266, 12127-12130.

Hu, Z. H., Arif, B. M., Jin, F., Martens, J. W. M., Chen, X. W., Sun, J. S., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1998). Distinct gene arrangement in the *Buzura suppressaria* single-nucleocapsid nucleopolyhedrovirus genome. *J. Gen. Virology* 79, 2841-2851.

Hu, Z. H., Broer, R., Westerlaken, J., Martens, J. W. M., Jin, F., Jehle, J. A., Wang, L. M., and Vlak, J. M. (1997). Characterization of the *ecdysteroid UDP-glucosyltransferase* gene of a single nucleocapsid nucleopolyhedro-virus of *Buzura suppressaria*. *Virus Research* 47, 91-97.

Hughes, P. R., and Wood, H. A. (1981). A synchronous peroral technique for the bioassay of insect viruses. *J. Invert. Pathol.* 37, 154-159.

Huh, N. E., and Weaver, R. F. (1990). Identifying the RNA polymerases that synthesize specific transcripts of the *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virology* 71, 195-201.

Hunter, D. K., and Hall, I. M. (1968). Cytopathology of a nuclear polyhedrosis virus of the beet army worm, *Spodoptera exigua*. *J. Invert. Pathol.* 12, 93-97.

Huynen, M., Snel, B., Lathe, W., and Bork, P. (2000). Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome Res.* 10, 1204-1210.

IJkel, W. F. J., Lebbink, R.-J., Op den Brouw, M. L., Goldbach, R. W., Blissard, G. W., Vlak, J. M., and Zuidema, D. (2001). Identification of a novel occlusion derived virus-specific protein in *Spodoptera exigua* multicapsid nucleopolyhedrovirus. Virology, in press.

IJkel, W. F. J., Westenberg, M., Goldbach, R. W., Blissard, G. W., Vlak, J. M., and Zuidema, D. (2000). A novel baculovirus envelope fusion protein with a proprotein cleavage site. *Virology* 275, 30-41.

IJkel, W. F. J., van Strien, E. A., Heldens, J. G., Broer, R., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1999). Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *J. Gen. Virology* 80, 3289-3304.

Isegawa, Y., Mukai, T., Nakano, K., Kagawa, M., Chen, J., Mori, Y., Sunagawa, T., Kawanishi, K., Sashihara, J., Hata, A., Zou, P., Kosuge, H., and Yamanishi, K. (1999). Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B. *J. Virology* 73, 8053-8063.

Jehle, J. A., van der Linden, I. F., Backhaus, H., and Vlak, J. M. (1997). Identification and sequence analysis of the integration site of transposon TCp3.2 in the genome of *Cydia pomonella* granulovirus. *Virus Research* 50, 151-157.

Jehle, J. A. (1996). Transmission of insect transposons into baculovirus genomes: an unusual host–pathogen interaction. In *Transgenic Organisms – Biological and Social Implications*, pp. 81-97. Edited by J. Tomiuk, K. Wohrmann and A. Sentker. Basel: Birlhauser Verlag.

Kang, W., Suzuki, M., Zemskov, E., Okano, K., and Maeda, S. (1999). Characterization of baculovirus repeated open reading frames (*bro*) in *Bombyx mori* nucleopolyhedrovirus. *J. Virology* 73, 10339-10345.

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

Kim, D., and Weaver, R. F. (1993). Transcription mapping and functional analysis of the *protein tyrosine/serine phosphatase* (PTPase) gene of the *Autographa californica* nuclear polyhedrosis virus. *Virology* 195, 587-595.

Kingsley, D. H., Behbahani, A., Rashtian, A., Blissard, G. W., and Zimmerberg, J. (1999). A discrete stage of baculovirus GP64-mediated membrane fusion. *Molecular Biol. Cell* 10, 4191-4200.

Klenk, H. D., Rott, R., Orlich, M., and Blodorn, J. (1975). Activation of influenza A viruses by trypsin treatment. *Virology* 68, 426-439.

Knopf, C. W. (1998). Evolution of viral DNA-dependent DNA polymerases. *Virus Genes* 16, 47-58.

Kogan, P. H., Chen, X., and Blissard, G. W. (1995). Overlapping TATA-dependent and TATA-independent early promoter activities in the baculovirus *gp64* envelope fusion protein gene. *J. Virology* 69, 1452-1461.

Kogan, P. H., and Blissard, G. W. (1994). A baculovirus *gp64* early promoter is activated by host transcription factor binding to CACGTG and GATA elements. *J. Virology* 68, 813-822.

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

Kool, M., Ahrens, C. H., Vlak, J. M., and Rohrmann, G. F. (1995). Replication of baculovirus DNA. *J. Gen. Virology* 76, 2103-2118.

Koonin, E. V., Tatusov, R. L., and Galperin, M. (1998). Beyond complete genomes: from sequence to structure and function. *Curr. Opin. Struct. Biol.* 8, 355-363.

Kuzio, J., Pearson, M. N., Harwood, S. H., Funk, C. J., Evans, J. T., Slavicek, J. M., and Rohrmann, G. F. (1999). Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar*. *Virology* 253, 17-34.

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

Lapointe, R., Back, D. W., Ding, Q., and Carstens, E. B. (2000). Identification and molecular characterization of the *Choristoneura fumiferana* multicapsid nucleopolyhedrovirus genomic region encoding the regulatory genes *pkip, p47, lef-12*, and *gta. Virology* 271, 109-121.

Lazarowitz, S. G., and Choppin, P. W. (1975). Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide. *Virology* 68, 440-454.

Le, T. H., Wu, T., Robertson, A., Bulach, D., Cowan, P., Goodge, K., and Tribe, D. (1997). Genetically variable triplet repeats in a RING-finger ORF of *Helicoverpa* species baculoviruses. *Virus Research* 49, 67-77.

Lenard, K., and Miller, D. K. (1982). Uncoating of enveloped viruses. *Cell* 28, 5-6.

Lepore, L. S., Roelvink, P. R. and Granados, R. R. (1996). Enhancin, the granulosis virus protein that facilitates nucleopolyhedrovirus (NPV) infections, is a metalloprotease. *J. Invert. Pathol.* 68, 131-140.

Liu, J. J., and Carstens, E. B. (1996). Identification, molecular cloning, and transcription analysis of the *Choristoneura fumiferana* nuclear polyhedrosis virus *spindle-like protein* gene. *Virology* 223, 396-400.

Lu, A., and Miller, L. K. (1996). Species-specific effects of the *hcf-1* gene on baculovirus virulence. *J. Virology* 70, 5123-5130.

Lu, A., and Miller, L. K. (1995). The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *J. Virology* 69, 975-982.

Maeda, T., and Ohnishi, S. (1980). Activation of influenza virus by acidic media causes hemolysis and fusion of erythrocytes. *FEBS Lett.* 122, 283-287.

Marcotte, E. M., Pellegrini, M., Thompson, M. J., Yeates, T. O., and Eisenberg, D. (1999). A combined algorithm for genome-wide prediction of protein function. *Nature* 402, 83-86.

Martens, J. W. M., Van Oers, M. M., Van de Bilt, B., Vlak, J. M., and Oudshoorn, P. (1995). Development of a baculovirus vector that facilitates the generation of *p10*-based recombinants. *J. Virological Methods* 52, 15-19.

McGeoch, D. J., Dolan, A., and Ralph, A. C. (2000). Towards a comprehensive phylogeny for mammalian and avian herpesviruses. *J. Virology* 74, 10401-10406.

McGeoch, D. J., Cook, S., Dolan, A., Jamieson, F. E., and Telford, E. A. (1995). Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J. Mol. Biol.* 247, 443-458.

McGeoch, D. J., and Cook, S. (1994). Molecular phylogeny of the alphaherpesvirinae subfamily and a proposed evolutionary timescale. *J. Mol. Biol.* 238, 9-22.

Megaw, A. G., Rapaport, D., Avidor, B., Frenkel, N., and Davison, A. J. (1998). The DNA sequence of the RK strain of human herpesvirus 7. *Virology* 244, 119-132.

Menon, N. K., Chatelus, C. Y., Dervartanian, M., Wendt, J. C., Shanmugam, K. T., Peck, H. D. Jr., and Przybyla, A. E. (1994). Cloning, sequencing, and mutational analysis of the hyb operon encoding *Escherichia coli* hydrogenase 2. *J. Bacteriol.* 176, 4416-4423.

Merrihew, R. V., Clay, W. C., Condreay, J. P., Witherspoon, S. M., Dallas, W. S., and Kost, T. A. (2001). Chromosomal integration of transduced recombinant baculovirus DNA in mammalian cells. *J. Virology* 75, 903-909.

Mikhailov, V. S., Mikhailova, A. L., Iwanaga, M., Gomi, S., and Maeda, S. (1998). *Bombyx mori* nucleopolyhedrovirus encodes a DNA-binding protein capable of destabilizing duplex DNA. *J. Virology* 72, 3107-3116.

Miller, L. K. (1996). Insect Viruses. *In* "Fields Virolgy" (B. N. Fields, D. M. Knipe, P. M. Howley, Eds.), pp. 33-59. Lippincott-Ravem Publishers, New York.

Minson, A. C., Davinson, A., Eberle, R., Desrosiers, R. C., Fleckenstein, B., McGeoch, D. J., Pellet, P. E., Roizman, B., and Studdert, D. M. F. (2000). Family *Herpesviridae*. *In* "Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses" (M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wicker, Eds.), pp. 203-225. Academic press, New York.

Mitsuhashi, W., Furuta, Y., and Sato, M. (1998). The spindles of an entomopoxvirus of coleoptera (*Anomala cuprea*) strongly enhance the infectivity of a nucleopolyhedrovirus in lepidoptera. *J. Invert. Pathol.* 71, 186-188.

Mitsuhashi, W., Saito, H. and Sato, M. (1997). Complete nucleotide sequence of the *fusolin* gene of an entomopoxvirus in the cupreous chafer, *Anomala cuprea* Hope (Coleoptera: Scarabeidae). *Insect Biochem. Mol. Biol.* 27, 869-876.

Monsma, S. A., Oomens, A. G., and Blissard, G. W. (1996). The GP64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. *J. Virology* 70, 4607-4616.

Monsma, S. A., and Blissard, G. W. (1995). Identification of a membrane fusion domain and an oligomerization domain in the baculovirus GP64 envelope fusion protein. *J. Virology* 69, 2583-2595.

Montague, M. G., and Hutchison, C. A. (2000). Gene content phylogeny of herpesviruses. *Proc. Natl. Acad. Sci. USA* 97, 5334-5339.

Morrisey, J. H. (1981). Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. *Analyt. Biochem.* 117, 307-310.

Moser, B. A., Becnel, J. J., White, Afonso, C., Kutish, G., Shanker, S., and Almira, E. (2001). Morpholigical and molecular evidence that *Culex nigripalpus* baculovirus is an unusual member of the family *Baculoviridae*. *J. Gen. Virology* 82, 283-297.

Muñoz, D., Murillo, R., Krell, P. J., Vlak, J. M., and Caballero, P. (1999). Four genotypic variants of a *Spodoptera exigua* nucleopolyhedrovirus (Se-SP2) are distinguishable by a hypervariable genomic region. *Virus Research* 59, 61-74.

Muñoz, D., Castillejo, J. I., and Caballero, P. (1998). Naturally occurring deletion mutants are parasitic genotypes in a wild-type nucleopolyhedrovirus population of *Spodoptera exigua. Appl. Environm. Microbiol.* 64, 4372-4377.

Nakayama, K. (1997). Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *J. Biochem.* 327, 625-635.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* 10, 1-6.

Ohkawa, T., Majima, M., and Maeda, S. (1994). A cysteine proteinase encoded by the baculovirus *Bombyx mori* nuclear polyhedrovirus. *J. Virology* 68, 6619-6625.

Okano, K., Mikhailov, V. S., and Maeda, S. (1999). Colocalization of baculovirus IE-1 and two DNA-binding proteins, DBP and LEF-3, to viral replication factories. *J. Virology* 73, 110-119.

Oomens, A. G. P., and Blissard, G. W. (1999). Requirement for GP64 to drive efficient budding of *Autographa californica* multicapsid nucleopolyhedrovirus. *Virology* 254, 297-314.

Oomens, A. G. P. (1999). Characterization of the *lef7-gp64-p74* locus of *Anticarsia gemmatalis* multicapsid nucleopolyhedrovirus. *In* "Functional and structural analysis of GP64, the major envelope glycoprotein of the budded virus phenotype of Autographa californica and Orgyia pseudotsugata multicapsid nucleopolyhedroviruses", pp. 103-122. Phd Thesis, Wageningen University.

Oomens, A. G. P., Monsma, S. A., and Blissard, G. W. (1995). The baculovirus GP64 envelope fusion protein: synthesis, oligomerization, and processing. *Virology* 209, 592-603.

O'Reilly, D. R. (1997). Auxiliary genes of baculoviruses. *In* "The Baculoviruses", pp.267-295. Edited by L.K. Miller, Plenum Press, New York.

O'Reilly, D. R. (1995). Baculovirus-encoded ecdysteroid UDP-glucosyltransferase. *Insect Biochem. Mol. Biol.* 26, 541-550.

O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992). Baculovirus expression vectors: A Laboratory Mannual. New York: W.H. Freeman.

O'Reilly, D. R., and Miller, L. K. (1991). Improvement of a baculovirus pesticide by deletion of the *egt* gene. *Biol. Technology* 9, 1086-1089.

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

Ortmann, D., Ohuchi, M., Angliker, H., Shaw, E., Garten, W., and Klenk, H. D. (1994). Proteolytic cleavage of wild type and mutants of the F protein of human parainfluenza virus type 3 by two subtilisin-like endoproteases, furin and Kex2. *J. Virology* 68, 2772-2776.

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. *Nature* 405, 837-846.

Park, C. G., Jung, M. Y., Choi, Y., and Winslow, G. M. (1995). Proteolytic processing is required for viral superantigen activity. *J. Exp. Med.* 181, 1899-1904.

Payne, R. W., Lane, P. W., Digby, P. G. N., Harding, S. A, Leech, P. K., Morgan, G. W., Todd, A. D., Thompson, R., Tunnicliffe W. G., Welham, S. J., and White, R. P. (1993). *In* "Genstat 5 Release 3 Reference Manual." Clarendon Press, Oxford, U.K.

Pearson, M. N., Groten, C., and Rohrmann, G. F. (2000). Identification of the *Lymantria dispar* nucleopolyhedrovirus envelope fusion protein provides evidence for a phylogenetic division of the *Baculoviridae. J. Virology* 74, 6126-6131.

Pearson, M. N., and Rohrmann, G. F. (1997). Splicing is required for transactivation by the *immediate-early gene 1* of the *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus. *Virology* 235, 153-165.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Met. Enzymol.* 183, 63-98.

Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D., and Yeates, T. O. (1999). Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc. Natl. Acad. Sci. USA* 96, 4285-4288.

Phanis, C. G., Miller, D. P., Cassar, S. C., Tristem, M., Thiem, S. M., and O'Reilly, D. R. (1999). Identification and expression of two baculovirus *gp37* genes. *J. Gen. Virology* 80, 1823-1831.

Plonsky, I., Cho, M. S., Oomens, A. G., Blissard, G., and Zimmerberg, J. (1999). An analysis of the role of the target membrane on the GP64-induced fusion pore. *Virology* 253, 65-76.

Przytycka, T., Aurora, R., and Rose, G. D. (1999). A protein taxonomy based on secondary structure. *Nat. Struct. Biol.* 6, 672-682.

Pullen, S. S., and Friesen, P. D. (1995). Early transcription of the *ie-1* transregulator gene of *Autographa californica* nuclear polyhedrosis virus is regulated by DNA sequences within its 5' noncoding leader region. *J. Virology* 67, 5776-5785.

Raftery, M., Müller, A., and Schönrich, G. (2000). Herpesvirus homologues of cellular genes. *Virus genes* 21, 65-75.

Ranjan, A., and Hasnain, S. E. (1995). Codon usage in the prototype baculovirus – *Autographa californica* nuclear polyhedrosis virus. *Indian J. Biochem. Biophysics* 32, 424-428.

Rawlings, N. D., Pearl, L. H., and Buttle, D. J. (1992). The baculovirus *Autographa californica* nuclear polyhedrosis virus genome includes a papain-like sequence. *Biological Chemistry Hoppe-Seyler* 373, 1211-1215.

Riegel, C. I., Lanner-Herrera, C., and Slavicek, J. M. (1994). Identification and characterization of the *ecdysteroid UDP-glucosyltransferase* gene of the *Lymantria dispar* multinucleocapsid nuclear polyhedrosis virus. *J. Gen. Virology* 75, 829-838.

Reiländer, H., Haase, W., and Maul, G. (1996). Functional expression of the *Aquorea victoria* green fluorescent protein in insect cell using the baculovirus expression system. *Biochem. Biophysics Res. Comm.* 219, 14-20.

Reinhardt, A., and Hubbard, T. (1998). Using neural networks for prediction of the subcellular location of proteins. *Nucl. Acids Res.* 26, 2230-2236.

Roeder, R. G. (1991). The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *Trends Biochem. Sci.* 16, 402-408.

Roncarati, R., and Knebel-Mörsdorf, D. (1997). Identification of the early *actin-rearrangement-inducing factor* gene, *arif-1*, from *Autographa californica* multicapsid nuclear polyhedrosis virus. *J. Virology* 71, 7933-7941.

Ross, L., and Guarino, A. G. (1997). Cycloheximide inhibition of delayed early gene expression in baculovirus-infected cells. *Virology* 232, 105-113.

Rost, B. (1996). PHD: predicting one-dimensional protein structure by profile based neural networks. *Met. Enzymol.* 266, 525-539.

Rost, B., and Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.* 232, 584-599.

Russell, R. M., Robertson, J. L., and Savin, N. E. (1977). POLO: A new computer program for probit analysis. *Bulletin Entomol. Society of America* 23, 209-213.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989*). Molecular cloning: A Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory press.

Sargent, F., Ballantine, S. P., Rugman, P. A., Palmer, T., and Boxer, D. H. (1998). Reassignment of the gene encoding the *Escherichia coli* hydrogenase 2 small subunit - identification of a soluble precursor of the small subunit in a hypB mutant. *Eur. J. Biochem.* 255, 746-754.

Seshagiri, S., and Miller, L. K. (1997). Baculovirus inhibitors of apoptosis (IAPs) block activation of Sf-caspase-1. *Proc. Natl. Acad. Sci. USA* 94, 13606-13611.

Shippam, C., Wu, X., Stewart, S., and Theilmann, D. A. (1997). Characterization of a unique OpMNPV-specific early gene not required for viral infection in tissue culture. *Virology* 227, 447-459.

Shirata, N., Ikeda, M., Kamiya, K., Kawamura, S., Kunimi, Y., and Kobayashi, M. (1999). Replication of nucleopolyhedroviruses of *Autographa californica* (Lepidoptera : Noctuidae), *Bombyx mori* (Lepidoptera : Bombycidae), *Hyphantria cunea* (Lepidoptera : Arctiidae), and *Spodoptera exigua* (Lepidoptera : Noctuidae) in four lepidopteran cell lines. *Appl. Entomol. Zool.* 34, 507-516.

Slack, J. M., Kuzio, J., and Faulkner, P. (1995). Characterization of *v-cath*, a cathepsin L-like proteinase expressed by the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virology* 76, 1091-1098.

Slavicek, J. M., Popham, H. J. R., and Riegel, C. I. (1999). Deletion of the *Lymantria dispar* multicapsid nucleopolyhedrovirus *ecdystroid UDP-glucosyl transferase* gene enhances viral killing speed in the last instar of the gypsy moth. *Biological Control* 16, 91-103.

Smith, I. R. L., and Crook, N. E. (1988). *In vivo* isolation of baculovirus genotypes. *Virology* 166, 240-244.

Smith, G. E., and Summers, M. D. (1978). Analysis of baculovirus genomes with restriction endonucleases. *Virology* 98, 517-527.

Smits, P. H., and 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 van de Faculteit Landbouwwetenschappen Rijks Universiteit Gent* 59, 385-392.

Smits, P. H., Van de Vrie, M., and Vlak, J. M. (1988). Nuclear polyhedrosis virus control of *Spodoptera exigua* on glasshouse crops. *Entomol. Experim. Appl.* 43, 73-80.

Smits, P. H., and Vlak, J. M. (1988a). Quantitative and qualitative aspects in the production of a nuclear polyhedrosis virus in *Spodoptera exigua* larvae. *Ann. Appl. Biol.* 112, 249-257.

Smits, P. H. and Vlak, J. M. (1988b). Selection of nuclear polyhedrosis viruses as biological control agents of *Spodoptera exigua* (Lepidoptera: Noctuidae). *Entomophaga* 33, 299-308.

Smits, P. H. (1987). *In* "Nuclear polyhedrosis virus as biological control agent of *Spodoptera exigua*". PhD. Thesis, Agric. Univ. Wageningen, The Netherlands, 127 p.

Stieneke-Grober, A., Vey, M., Angliker, H., Shaw, E., Thomas, G., Roberts, C., Klenk, H. D., and Garten, W. (1992). Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease. *EMBO J.* 11, 2407-2414.

Steinhauer, D. A. (1999). Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology* 258, 1-20.

Summers, M. D. (1977). Characterization of shrimp baculovirus. EPA Ecological Research Series November 1977, EPA-600/3-77-130, EPA, Gulf Breeze, FL.

Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997). A genomic perspective on protein families. *Science* 278, 631-637.

Thiem, S. M., Du, X., Quentin, M. E., and Berner, M. M. (1996). Identification of baculovirus gene that promotes *Autographa californica* nuclear polyhedrosis virus replication in a nonpermissive insect cell line. *J. Virology* 70, 2221-2229.

Thiem, S. M., and Miller, L. K. (1989). Identification, sequence, and transcriptional mapping of the major capsid protein of *Autographa californica* nuclear polyhedrosis virus. *J. Virology* 63, 2009-2018.

Theilmann, D. A., and Stewart, S. (1993). Analysis of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus trans-activators IE1 and IE2 using monoclonal antibodies. *J. Gen. Virology* 74, 1819-1826.

Tidona, C. A., and Darai, G. (2000). Iridovirus homologues of cellular genes – implications for the molecular evolution of large DNA viruses. *Virus Genes* 21, 77-81.

Tilakaratne, N., Hardin, S. E., and Weaver, R. F. (1991). Nucleotide sequence and transcript mapping of the *Hind*III F region of the *Autographa californica* nuclear polyhedrosis virus genome. *J. Gen. Virology* 72, 285-291.

Tong, W. H., and Rouault, T. (2000). Distinct iron-sulphur cluster assembly complexes exist in the cytosol and mitochondria of human cells. *EMBO J.* 19, 5692-5700.

van Loo, N. D., Fortunati, E., Ehlert, E., Rabelink, M., Grosveld, F., and Scholte, B. J. (2001). Baculovirus infection of nondividing mammalian cells: mechanisms of entry and nuclear transport of capsids. *J. Virology* 75, 961-970.

van Oers, M. M., Vlak, M. J., Voorma, H. O., and Thomas, A. A. M. (1999). Role of the 3' untranslated region of baculovirus *p10* mRNA in high-level expression of foreign genes. *J. Gen. Virology* 80, 2253-2262.

van Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L., Carstens, E. B., Estes, M. K.,. Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R., and Wicker, R. B. (Eds.) (2000). *In* "Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses", pp. 1-1162. Academic press, New York.

van Strien, E. A., IJkel, W. F. J. Gerrits, H. Vlak, J. M., and Zuidema, D. (2000). Characteristics of the transactivator gene *ie1* of *Spodoptera exigua* multiple nucleopolyhedrovirus *Arch. Virology* 145, 2115-2133.

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

van Strien, E. A., Jansen, B. J. H., Mans, R. M. W., Zuidema, D., and Vlak, J. M. (1996). Sequence and transcriptional analysis of the *ubiquitin* gene cluster in the genome of *Spodoptera exigua* nucleopolyhedrovirus. *J. Gen. Virology* 77, 2311-2319.

van Strien, E. A., Faktor, O., Hu, Z. H., Zuidema, D., Goldbach, R. W., and Vlak, J. M. (1997). Baculoviruses contain a gene for the large subunit of ribonucleotide reductase. *J. Gen. Virology* 78, 2365-2377.

Vaugh, J. L., Goodwin, R. H., Tompkins, G. J., and McCawley, P. (1977). The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro* 13, 213-217.

Vey, M., Schafer, W., Reis, B., Ohuchi, R., Britt, W., Garten, W., Klenk, H. D., and Radsak, K. (1995). Proteolytic processing of human cytomegalovirus glycoprotein B (gpUL55) is mediated by the human endoprotease furin. *Virology* 206, 746-749.

Vey, M., Orlich, M., Adler, S., Klenk, H. D., Rott, R., and Garten, W. (1992). Hemagglutinin activation of pathogenic avian influenza viruses of serotype H7 requires the protease recognition motif R-X-K/R-R. *Virology* 188, 408-413.

Vink, C., Beuken, E., and Bruggeman, C. A. (2000). Complete DNA sequence of the rat cytomegalovirus genome. *J. Virology* 74, 7656-7665.

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

Vlak, J. M., Frankenhuyzen, K., and Peters, D. (1981). Identification of a new nuclear polyhedrosis virus from *Spodoptera exigua*. *J. Invert. Pathol.* 38, 297-298.

Volkman, L. E., Blissard, G. W., Friesen, P., Keddie, B. A., Possee, R., and Theilmann, D. A. (1995). Family *Baculoviridae*. *In* "Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses" (F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial. A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds.), pp. 104-113. Vienna and New

York: Springer-Verlag.

Volkman, L. E., and Keddie, B. A. (1990). Nuclear polyhedrosis virus pathogenesis. *Seminars in Virology* 1, 249-254.

Volkman, L. E., and Goldsmith, P. A. (1985). Mechanism of neutralization of budded *Autographa californica* nuclear polyhedrosis virus by a monoclonal antibody: inhibition of entry by adsorptive endocytosis. *Virology* 143, 185-195.

Volkman, L. E., Goldsmith, P. A., Hess, R. T., and Faulkner, P. (1984). Neutralization of budded *Autographa californica* NPV by a monoclonal antibody: Identification of the target antigen. *Virology* 133, 354-362.

Volkman, L. E., and Goldsmith, P. A. (1984). Budded *Autographa californica* NPV 64K protein: further biochemical analysis and effects of postimmunoprecipitation sample preparation conditions. *Virology* 139, 295-302.

Vucic, D., Kaiser, W. J., and Miller, L. K. (1998). A mutational analysis of the baculovirus inhibitor of apoptosis Op-IAP. *J. Biol. Chemistry* 273, 33915-33921.

Wang, J. W., Qi, Y. P., Huang, Y. X., and Li, S. D. (1995). Nucleotide sequence of a 1446 base pair *SalI* fragment and structure of a novel early gene of *Leucania separata* nuclear polyhedrosis virus. *Arch. Virology* 140, 2283-2291.

Wang, P., and Granados, R. R. (1998). Observations on the presence of the peritrophic membrane in larval *Trichoplusia ni* and its role in limiting baculovirus infection. *J. Invert. Pathol.* 72, 57-62.

Wang, P., Hammer, D. A., and Granados, R. R. (1997). Binding and fusion of *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virology* 63, 1393-1399.

Wang, P., and Granados, R. R. (1997). An intestinal mucin is the target substrate for a baculovirus enhancin. *Proc. Natl. Acad. Sci. USA* 94, 6977-6982.

Watanabe, M., Hirano, A., Stenglein, S., Nelson, J., Thomas, G., and Wong, T. C. (1995). Engineered serine protease inhibitor prevents furin-catalyzed activation of the fusion glycoprotein and production of infectious measles virus. *J. Virology* 69, 3206-3210.

Whitford, M., and Faulkner, P. (1992). A structural polypeptide of the baculovirus *Autographa californica* nuclear polyhedrosis virus contains O-linked N-acetylglucosamine. *J. Virology* 66, 3324-3329.

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

Williams, G. V., and Faulkner, P. (1997). Cytological changes and viral morphogenesis during baculovirus infection. *In* "The baculoviruses", (L.K. Miller, Ed.), pp. 61-108. Plenum, New York.

Wishart, M. J., Denu, J. M., Williams, J. A., and Dixon, J. E. (1995). A single mutation converts a novel phosphotyrosine binding domain into a dual-specificity phosphatase. *J. Biol. Chem.* 270, 26782-26785.

Wu, J., and Miller, L. K. (1989). Sequence, transcription of a late gene of the *Autographa californica* nuclear polyhedrosis virus encoding a 34.8K polypeptide. *J. Gen. Virology* 70, 2449-2459.

Yang, F., Wang, W., Chen, R. Z., and Xu, X. (1997). A simple and efficient method for purification of prawn baculovirus DNA. *J. Virol. Meth.* 67, 1-4.

Zanotto, P. M., Kessing, B. D., and Maruniak, J. E. (1993). Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. *J. Invert. Pathol.* 62, 147-164.

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

Summary

The *Baculoviridae* are a family of rod-shaped viruses with large circular double-stranded DNA genomes (Chapter 1). The family is subdivided into two genera, *Granulovirus* (GV) and *Nucleopolyhedrovirus* (NPV) on the basis of the type of body occluding the virions. NPVs are further subdivided in group I and II based on phylogenetic evidence of the DNA polymerase protein. Baculoviruses almost exclusively infect insects and are, therefore, attractive biological alternatives to chemical insecticides for insect pest control. The baculovirus *Spodoptera exigua* multicapsid nucleopolyhedrovirus (SeMNPV) infects the beet army worm *S. exigua* (Lepidoptera: Noctuidae) and has been successfully used as a bioinsecticide to control this world-wide insect pest of agricultural importance. SeMNPV differs from many other baculoviruses in that it is mono-specific and highly virulent for *S. exigua* larvae. The research described in this thesis aimed at the molecular characterization of the baculovirus SeMNPV to gain insight in its gene content and organization in comparison to those of other baculoviruses. At the same time this study will support or reject its current taxonomic position using gene and genome phylogeny analyses and might reveal insight in the molecular mechanisms associated with the biological properties of SeMNPV.

As a start the complete nucleotide sequence of the DNA genome of SeMNPV, a putative group II NPV, was determined and analyzed (Chapter 2). The genome was composed of 135,612 bp containing 138 putative genes or open reading frames (ORFs). Major differences in SeMNPV gene content and arrangement were found compared with the group I NPVs *Autographa californica* (Ac), *Bombyx mori* (Bm), *Orgyia pseudotsugata* (Op) and the group II NPV *Lymantria dispar* (Ld). Sixteen ORFs were unique to SeMNPV, while the remaining ORFs (122) all had a homolog in one or more of the nine baculoviruses sequenced to date (Chapter 7). Sixty-three ORFs were conserved among all nine baculoviruses and are likely to be essential for NPV multiplication and survival. Strikingly, two of these NPV 'core' genes, *odv-e66* and *p26*, were found duplicated in SeMNPV. Gene parity analysis of baculoviral genomes indicated that SeMNPV and LdMNPV are closely related and that they are only distantly related to group I NPVs. Therefore, SeMNPV can be considered as a group II NPV.

Two of the 16 unique SeMNPV genes, Se116 and Se117, share similarity on amino acid level, but are not related on nucleotide level. To investigate the function, if

any, of the unique SeMNPV genes in general, Se116 and Se117 were analyzed and characterized (Chapter 3). Se116 and Se117 were expressed from early till late in infection both in cultured cells and in larvae of *S. exigua*. Their transcripts were polyadenylated and initiated from typical baculovirus early promoter motifs. Se116 and Se117 encoded proteins of 27 and 23 kDa, respectively, which were localized in the virogenic stroma of the nucleus. While the function of the Se116 protein remains enigmatic, the Se117 protein appeared to be a structural protein associated with nucleocapsids of occlusion-derived virus (ODV), but not of budded virus (BV). Further investigation will reveal if and how these proteins are involved in the SeMNPV virulence or host range determination.

The research on unique SeMNPV genes was extended (Chapter 4) by the characterization of another gene, Se17/18, unique among NPVs, but strikingly having a homolog (ORF129) in the granulovirus *Xestia c-nigum* (XcGV), which is only distantly related to SeMNPV. Se17/18 was transcribed in cultured *S. exigua* 301 cells from early till late in infection. However, *in vivo* transcripts could only be detected late in infection. These polyadenylated transcripts started in a region containing a baculovirus consensus early promoter motif. In contrast to the Se116 and Se117 proteins, the Se17/18 protein was primarily localized in the cytoplasm. A chicken polyclonal antiserum was raised that reacted specifically to Se17/18 protein expressed in *E. coli*. However, no immunoreactive protein was detected in SeMNPV-infected insect cells. The absence of immunoreactive Se17/18 protein implies that it is rapidly turned over in insect cell culture or that the gene is only active in larvae and possibly has a regulatory function.

A thorough analysis of the complete SeMNPV genome revealed that it lacked a homolog of the major budded virus glycoprotein gene *gp64*, that is found in AcMNPV and other group I NPVs. Upon infection, by representatives of this group, acidification of the endosome triggers fusion of the viral and endosomal membrane, which is mediated by the BV envelope glycoprotein GP64. Therefore, the entry mechanism of SeMNPV in cultured cells was examined. SeMNPV budded virus (BV) entered insect cells by endocytosis like BVs of group I NPVs. Furthermore, a functional homolog of the envelope fusion protein GP64 was identified in Se8 (76 kDa) and appeared to be the major envelope protein of SeMNPV BVs. Surprisingly, a 60 kDa cleavage product of this protein was present in the BV envelope. A furin-like proprotein convertase cleavage site was identified immediately upstream of the N-terminus of the mature Se8 protein and this site was also conserved in the LdMNPV homolog (Ld130) of Se8. Syncytium formation assays showed that Se8 and Ld130 alone were sufficient to mediate membrane fusion. Both proteins were primarily localized in the plasma

membrane of insect cells, which was consistent with their fusogenic activity. If Se8 is cleaved by a cellular convertase the host could also play a role in the determination of virus host range and virulence.

The research on function of single SeMNPV genes and also the engineering of this virus for improved insecticidal activity or as expression vector have been hampered as defective viruses are quickly generated when using insect cell culture. These defective viruses lack 25 kb sequence information and are no longer active *in vivo* upon oral feeding. A novel procedure to isolate SeMNPV recombinants was adopted by alternate cloning between insect larvae and cultured cells. In this way a SeMNPV recombinant (SeXD1) was obtained infectious both *in vivo* and in cell culture and with an improved speed of kill. This recombinant lacked 10.6 kb of sequence information, including *ecdysteroid UDP glucosyl transferase (egt), gp37, chitinase* and *cathepsin* genes, as well as several genes unique to SeMNPV. One of these unique genes was Se17/18. The result indicated, however, that these genes are dispensable for virus replication both in cell culture and *in vivo*. A mutant with a similar deletion was identified by PCR in the parental wild type SeMNPV isolate suggesting that genotypes with differential biological activities exist in field isolates of baculoviruses.

The research on SeMNPV described in this thesis, has provided a complete overview of its coding potential and insight in several features common to lepidopteran baculoviruses, such as 'core' genes, unique genes and clustering of conserved genes (Chapter 7). The initial characterization of several SeMNPV genes resulted in the identification of a novel ODV-specific nucleocapsid protein unique to SeMNPV and a novel major BV envelope fusion protein. The latter is the first baculovirus protein reported to be cleaved by a cellular furin-like proprotein convertase. The development of a novel procedure to generate recombinants *in vivo* is presumably applicable to many baculovirus species in order to obtain biologically active recombinants. Exploitation of this technique will enable the further characterization of (unique) SeMNPV genes by deletion, insertion and mutation by *in vivo* recombination. Understanding the function of SeMNPV genes will ultimately lead to the unravelling of the molecular basis underlying the mono-specificity and high virulence of SeMNPV for the beet army worm *Spodoptera exigua*.

Samenvatting

Baculovirussen vormen een familie (*Baculoviridae*) van staafvormige DNA-virussen, die alleen geleedpotigen (Arthropoda) infecteren. Op grond van de verschijningsvorm van de virusdeeltjes, als polyeders of als granula, is deze familie in twee *genera* verdeeld: de kernpolyedervirussen (NPV, nucleopolyhedrovirus) en de granulovirussen (GV). Daarvan zijn de NPVs weer onderverdeeld in twee groepen (I en II) op basis van fylogenetische verwantschap in het DNA polymerase eiwit. Veel baculovirussen hebben een beperkt gastheerbereik en zijn daarom een aantrekkelijk alternatief voor breed werkende chemische insecticiden. Het baculovirus *Spodoptera exigua* MNPV (SeMNPV) infecteert het larvale stadium van de floridamot (*S. exigua*) en wordt in de praktijk met succes toegepast als biologisch bestrijdingsmiddel tegen dit plaaginsect. SeMNPV heeft twee opvallende biologische kenmerken: het is alleen infectieus voor *S. exigua* en is in vergelijking tot andere baculovirussen relatief virulent. Het onderzoek, beschreven in dit proefschrift, had als doel om de genetische organisatie van het baculovirus SeMNPV volledig in kaart te brengen en te vergelijken met die van andere baculovirussen. Dit onderzoek zou tevens uitsluitsel kunnen geven over de juistheid van de voorlopige indeling van SeMNPV als groep II NPV en een eventuele genetische basis van de specifieke biologische eigenschappen aan het licht kunnen brengen.

Daartoe werd allereerst de gehele nucleotidenvolgorde van het SeMNPV DNA-genoom bepaald en geanalyseerd (Hoofdstuk 2). Dit genoom bleek uit 135.612 nucleotiden te bestaan met daarin 138 mogelijke genen, ook wel 'open reading frames' (ORFs) genoemd. De genensamenstelling van SeMNPV bleek zeer afwijkend te zijn van die van *Autographa californica* (Ac) MNPV, *Bombyx mori* (Bm) NPV en *Orgyia pseudotsugata* (Op) MNPV (groep I NPV) en ook van die van *Lymantria dispar* (Ld) MNPV (groep II NPV). Zestien ORFs bleken uniek te zijn voor SeMNPV, terwijl de overige 122 ORFs een homoloog gen hadden in één of meer van de negen baculovirussen, waarvan de volledige nucleotidenvolgorde ondertussen is bepaald (Hoofdstuk 7). Van deze 122 ORFs waren 63 ORFs aanwezig in alle negen baculovirussen. Deze 63 ORFs worden derhalve beschouwd als 'core'-genen en zijn waarschijnlijk essentieel voor de basale processen van deze virussen, zoals genoom replicatie en transcriptie regulatie. Twee van deze 'core'-genen, *odv-e66* en *p26,* bleken elk gedupliceerd te zijn in SeMNPV. Vergelijking van de gen-organisatie in diverse baculovirussen bracht aan het licht dat SeMNPV het meest verwant is aan

LdMNPV en minder verwant aan groep I NPVs. Hieruit werd geconcludeerd dat SeMNPV inderdaad als een NPV van groep II kan worden beschouwd.

Twee van de 16 unieke SeMNPV-genen, Se116 en Se117, vertonen onderling verwantschap in hun translatieproduct, maar niet in hun nucleotidenvolgorde. Om de functie van deze genen te bestuderen, werden hun transcriptie- en translatiepatronen geanalyseerd (Hoofdstuk 3). Se116 en Se117 transcripten komen van 'vroeg' tot 'laat' in infectie voor in zowel celculturen als in *S. exigua*-larven. De Se116- en Se117-transcripten zijn gepolyadenyleerd en initiëren beide in een promotermotief, dat vaak voorkomt in baculovirusgenen die 'vroeg' tot expressie komen. De translatieproducten van Se116 en Se117, eiwitten van respectievelijk 27 en 23 kDa, werden gelokaliseerd in het virogene stroma van de celkern. Met betrekking tot de functie van het Se116-eiwit konden geen verdere aanknopingspunten gevonden worden, maar het Se117-eiwit bleek een nieuw structureel eiwit te zijn dat voorkomt in de nucleocapsiden van virions uit polyeders (ODV, occlusion derived virus) maar niet in virions uit 'budded virus' (BV). Verder onderzoek zal moeten uitwijzen of, en zo ja hoe, deze eiwitten een rol spelen in de biologische eigenschappen van SeMNPV, zoals gastheerbereik en virulentie.

Het onderzoek naar de functie van unieke SeMNPV-genen werd voortgezet door het Se17/18-gen te bestuderen (Hoofdstuk 4). Weliswaar komt er geen homoloog van dit gen voor in andere NPVs, maar opvallend genoeg wel in het granulovirus *Xestia c-nigum* (XcGV ORF129), dat verder weinig genetische verwantschap heeft met SeMNPV. In Se301 cellen komen gepolyadenyleerde transcripten van Se17/18 zowel in de 'vroege' als 'late' stadia van infectie voor, maar in *S. exigua*-larven alleen 'laat'. In tegenstelling tot de Se116- en Se117-eiwitten bleek het Se17/18-eiwit niet aantoonbaar te zijn in geïnfecteerde insectencellen met behulp van een tegen dit eiwit gericht antiserum. Door het Se17/18-eiwit te fuseren met GFP kon door middel van fluorescentie aangetoond worden dat Se17/18 een cytoplasmatisch eiwit is. Dit zou kunnen betekenen dat Se17/18 eiwit òf weinig wordt aangemaakt òf snel wordt afgebroken in insectencellen en derhalve een regulerende functie zou hebben. Een andere verklaring zou kunnen zijn dat het Se17/18-eiwit alleen actief is in larven.

Uit een nauwgezette analyse van het SeMNPV-genoom bleek dat er geen homoloog voor het BV glycoproteïne GP64 aanwezig is, terwijl dit gen geconserveerd is in de groep I NPVs. Het GP64-eiwit speelt daar een cruciale rol in het bewerkstelligen van de fusie tussen de virale en cellulaire membraan bij de start van de infectie. Omdat GP64 in SeMNPV afwezig is, werd gezocht naar het fusie-eiwit in SeMNPV virions en het coderende gen (Hoofdstuk 5). Uit onderzoek bleek dat SeMNPV eenzelfde mechanisme (endocytose) gebruikt als groep I NPVs om

insectencellen binnen te dringen, maar daarbij geen GP64-gerelateerd eiwit aanwendt. In plaats daarvan werd een ander eiwit, het translatieproduct van Se8 (76 kDa) geïdentificeerd als celfusie-eiwit. Verrassenderwijs, werd niet het volledige Se8 eiwit maar een kleiner (60 kDa) klievingsproduct in BVs van SeMNPV gevonden. Dit 60 kDa-eiwit is het meest voorkomende virale eiwit in de envelop van SeMNPV BV. Het kon aannemelijk gemaakt worden dat een tot de furine (een klasse van proteasen) behorend 'proprotein convertase' van de gastheer betrokken was bij de klieving. Dit bleek ook op te gaan voor de LdMNPV (groep II NPV) homoloog (Ld130) van Se8. Met syncytium-inductie-experimenten werd vervolgens aangetoond dat de Se8- en Ld130-eiwitten zelfstandig in staat zijn om deze fusie te bewerkstelligen. Tenslotte werden deze twee eiwitten voornamelijk in de celmembraan van insectencellen gedetecteerd, hetgeen in overeenstemming is met hun fusogene activiteiten. Als het Se8-eiwit wordt gekliefd door een cellulaire 'convertase', een nieuw fenomeen bij baculovirussen, kan dit proces een belangrijke rol spelen als determinant van gastheerbereik en virulentie van SeMNPV.

Het onderzoek naar de functie van specifieke SeMNPV-genen en de toepassing van het virus als bio-insecticide of als expressievector, wordt gehinderd doordat het SeMNPV-genoom instabiel is in celculturen. Als gevolg daarvan ontstaan deletiemutanten, die ongeveer 25.000 nucleotiden aan sequentie-informatie missen en daardoor hun infectiositeit voor *S. exigua*-larven hebben verloren. Daarom werd een nieuwe procedure aangewend om infectieuze SeMNPV-recombinanten te isoleren (Hoofdstuk 6). Door beurtelings *S. exigua*-larven en celcultuur te gebruiken werd een eerste SeMNPV-recombinant (SeXD1) verkregen, die infectieus was voor zowel *S. exigua*-larven als insectencellen en deze zelfs nog sneller doodde. De SeXD1-recombinant miste slechts 10.600 nucleotiden aan sequentie-informatie, waaronder de genen voor *ecdysteroid UDP glucosyl transferase (egt), gp37, chitinase, cathepsin* en ook een aantal unieke SeMNPV-genen, waaronder Se17/18. De resultaten impliceren dat al deze genen dus niet essentieel zijn voor vermenigvuldiging van het virus in insectencellen en *S. exigua*-larven. Een mutant met eenzelfde deletie als in SeXD1 werd met behulp van PCR-technieken gedetecteerd in een veldisolaat van SeMNPV. Blijkbaar bestaan er dus dergelijke varianten met verschillende biologische activiteiten in natuurlijke veldisolaten.

Samenvattend kan gesteld worden dat het onderzoek beschreven in dit proefschrift een compleet overzicht verschaft met betrekking tot de genensamenstelling van het SeMNPV-genoom en tevens inzicht geeft in de genetische eigenschappen van baculovirussen in het algemeen, zoals het bestaan van 'core'-genen, unieke genen, en de wijze waarop de genen gegroepeerd liggen in een beperkt

aantal genclusters (Hoofdstuk 7). De karakterisering van een aantal geselecteerde SeMNPV-genen resulteerde in de ontdekking van een nieuw ODV-specifiek nucleocapsid-eiwit, uniek voor SeMNPV en een nieuw type fusie-eiwit voor baculovirussen. De ontwikkeling van een nieuwe procedure om recombinanten te isoleren uit *S. exigua*-larven zal verder onderzoek naar de functie van (unieke) SeMNPV-genen via gerichte mutagenese mogelijk maken. Inzicht in de functie van de individuele genen zal vervolgens kunnen uitwijzen wat de moleculaire basis is van de karakteristieke biologische eigenschappen van SeMNPV, zoals de specificiteit en hoge virulentie voor de floridamot *Spodoptera exigua*.

Curriculum vitae

Willem Frederik Johannes IJkel werd geboren op 20 januari 1973 in Amersfoort. In 1991 behaalde hij het VWO diploma aan de Gereformeerde Scholengemeenschap in Amersfoort. In datzelfde jaar werd aangevangen met de studie Moleculaire Wetenschappen, specialisatie chemische biologisch, aan de Landbouwuniversiteit in Wageningen. In augustus 1996 werd het ingenieursdiploma behaald. Tijdens de doctoraalfase heeft hij afstudeeronderzoek verricht in de Biochemie aan de Landbouwuniversiteit Wageningen (prof. dr. C. Veerman en dr. W.M.A.M. van Dongen), in de Bacteriële Genetica (prof. dr. W.M. de Vos) aan de Landbouw Universiteit Wageningen en in de Endocrinologie en Voortplanting aan de faculteit Geneeskunde van de Erasmus Universiteit Rotterdam (prof. dr. J.A. Grootegoed en dr. A.P.N. Themmen). Van januari 1997 tot april 2001 was hij werkzaam als Onderzoeker in Opleiding (OIO) bij het laboratorium voor Virologie aan de Wageningen Universiteit onder begeleiding van prof. dr. J.M. Vlak en prof. dr. R.W. Goldbach. Het onderzoek dat in deze periode is uitgevoerd, werd gefinancierd door de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO) en staat beschreven in dit proefschrift.

Account

W.F.J. IJkel, R.-J. Lebbink, M.L. Op den Brouw, R.W. Goldbach, G.W. Blissard, J.M. Vlak and D. Zuidema. (2001). Identification of a novel occlusion derived virus-specific protein in *Spodoptera exigua* multicapsid nucleopolyhedrovirus.
Virology 283, in press.

X. Chen, **W.F.J. IJkel**, R. Tarchini, X. Sun, H. Sandbrink, H. Wang, S. Peters, D. Zuidema, R.K. Lankhorst, J.M. Vlak and Z. Hu. (2001). The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome.
Journal of General Virology 82, 241-257.

W.F.J. IJkel, M. Westenberg, R.W. Goldbach, G.W. Blissard, J.M. Vlak and D. Zuidema (2000). A novel baculovirus envelope fusion protein with a proprotein convertase cleavage site.
Virology 275, 30-41.

X. Dai, J.P. Hajós, N.N. Joosten, M.M. van Oers, **W.F.J. IJkel**, D. Zuidema, Y. Pang and J.M. Vlak. (2000). Isolation of a *Spodoptera exigua* baculovirus recombinant with a 10.6 kbp genome deletion that retains biological activity.
Journal of General Virology 81, 2545-2554.

E.A. van Strien, **W.F.J. IJkel**, H. Gerrits, J.M. Vlak and D. Zuidema (2000). Characteristics of the transactivator gene *ie1* of *Spodoptera exigua multiple nucleopolyhedrovirus*.
Archives of Virology 145, 2115-2133.

W.F.J. IJkel, E.A. van Strien, J.G.M. Heldens, R. Broer, D. Zuidema, R.W. Goldbach and J.M. Vlak (1999). Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome.
Journal of General Virology 80, 3289-3304.

X. Chen, **W.F.J. IJkel**, C. Dominy, P.M.A. Zanotto, Y. Hashimoto, O. Faktor, T. Hayakawa, C.H. Wang, A. Prekumar, S. Mathavan, P.J. Krell, Z. Hu and J.M. Vlak. (1999). Identification, sequence analysis and phylogeny of the *lef-2* gene of *Helicoverpa armigera* single-nucleocapsid baculovirus.
Virus Research 65, 21-32.