Nucleotide sequence and transcriptional analysis of the polyhedrin gene of *Spodoptera exigua* nuclear polyhedrosis virus

Elisabeth A. van Strien, Douwe Zuidema, Rob W. Goldbach and Just M. Vlak*

Department of Virology, Wageningen Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

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

Introduction

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

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

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

Polyhedrin is the most abundant protein of polyhedra and present in high amounts in infected cells. For many baculoviruses it has been the subject of detailed molecular and genetic studies. It is approximately 30K in size and highly conserved among baculoviruses (Vlak & Rohrmann, 1985). Elucidation of the genome structure and expression of the SeMNPV polyhedrin gene is a starting point for the design and engineering of this virus to improve its insecticidal properties. In this paper the nucleotide sequence organization and transcriptional analysis of the polyhedrin gene of SeMNPV is reported. The nucleotide and amino acid sequences of SeMNPV have been assigned accession number X67243 by EMBL and GenBank.
polyhedrin are compared with those of polyhedrin genes of other MNPVs.

**Methods**

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

Haemolymph from SeMNPV-infected insects was used as a source of extracellular virus (ECV) for infection of cultured S. exigua cells (Gelernter & Federici, 1986b). This cell line (UCR-SE-1), obtained from Dr B. A. Federici, was maintained in plastic tissue culture flasks in TNN-FH medium (Hink, 1970) supplemented with 10% foetal calf serum. Infectivity assays of the ECV solutions were performed using the endpoint dilution method (Vlak, 1979) and the titres were expressed as TCID₅₀ units per ml.

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

**Southern blot hybridization, cloning and sequencing.** To locate polyhedrin gene-containing sequences, the SeMNPV DNA was digested with several restriction enzymes, separated in 0-7% agarose gels, transferred to Hybond N filters (Amersham) and hybridized with the ³²P-labelled HindIII V fragment of AcMNPV DNA according to procedures described by Sambrook et al. (1989). This AcMNPV DNA fragment contained polyhedrin gene sequences (Smith et al., 1983). SeMNPV DNA fragments hybridizing with the AcMNPV HindIII V probe were isolated from agarose gels by the freeze–squeeze method (Sambrook et al., 1989) and (sub)cloned into the plasmids pTZ18R (Promega) and pUC. The SeMNPV DNA inserts were sequenced using T7 DNA polymerase (Promega) and [³²P]dATP according to Sanger et al. (1977).

**RNA isolation.** For Northern blot hybridization, primer extension and SI nuclease analysis, total RNA was isolated from infected S. exigua cells at 24 or 48 h post-infection (p.i.). Cells were lysed in 4 M-guanidine isothiocyanate, 25 mM-sodium acetate, pH 6-0, and 83% 2-mercaptoethanol and stored frozen until centrifugation of the RNA through a CsCl cushion (5-7 M-CsCl, 25 mM-sodium acetate, pH 6-0) by centrifugation for 23 h at 32000 r.p.m. (Beckman ultracentrifuge). The pelletted RNA was dissolved in sterilized double-distilled water and ethanol-precipitated after the addition of 0-1 volume of 1-5 M-sodium acetate (pH 5-0). After washing the RNA pellet in 70% (v/v) ethanol the pellet was redissolved in sterilized water. The RNA concentration was determined by measuring the u.v. absorbance at 260 nm.

**Northern blot hybridization.** Total RNA was electrophoresed in a 1% (w/v) agarose gel in 10 mM-phosphate buffer (5 mM-Na₂HPO₄, pH 6-5 to 7-0), with constant recirculation of the running buffer. Prior to electrophoresis the RNA was denatured in 50% (w/v) DMSO, 1 M-glyoxal (deionized), 10 mM-sodium phosphate buffer for 1 h at 50 °C and a 0-2 volume of loading buffer (50% glycerol, 10 mM-sodium phosphate, 0-4% bromophenol blue) was added.

The separated RNAs were transferred to GeneScreen Plus (Dupont) or Hybond N (Amersham) in 10 × SSC (1-5 M-NaCl, 0-15 M-sodium citrate, pH 7-0). The filter was prehybridized for 1 h at 48 °C in a volume of 0-2 ml hybridization buffer [6 × SSC, 1% SDS, 200 µg denatured herring sperm DNA, 1% Ficoll-400, 1% polyvinylpyrrolidone, 1% BSA (fraction V)] per cm² filter. The 5' end-labelled oligonucleotide used in the primer extension assay (see below) was used as a probe in the Northern blot hybridization and added to the prehybridization solution. After hybridization for 16 h at 48 °C the filter was washed for 5 min in 4 × SSC, 1% SDS and then for 5 min in two changes of 2 × SSC, 0-5% SDS at room temperature. The filter was autoradiographed for several days.

**Primer extension.** To identify the transcriptional start site of the SeMNPV polyhedrin gene, a 17-mer oligonucleotide (5'-TCGTACAAGTAAGTGC-3') complementary to the polyhedrin mRNA was synthesized and used in primer extension assays. The oligonucleotide was 5' end-labelled with [γ-³²P]dATP using T4 polynucleotide kinase and denatured at 90 °C for 10 min. The labelled oligonucleotide was purified on a 1 ml Sephadex G25 column and ethanol-precipitated after addition of a 0-1 volume of 1-5 M-sodium acetate pH 5-0. Fifteen ng labelled primer was added to 2 µg total RNA and this mixture was denatured for 10 min at room temperature in 25 mM-methylmercury. After addition of diluted 2-mercaptoethanol, the reverse transcriptase reaction was carried out for 30 min at 37 °C in a volume of 15 µl with 0-5 µl of each of the four dNTPs, 1 µl Moloney murine leukaemia virus reverse transcriptase (Gibco-BRL) in the buffer supplied by the manufacturer. The reaction was stopped by addition of 5 µl of a solution containing 95% (v/v) formamide, 0-01% xylene cyanol, 0-01% bromophenol blue. Five µl of this reaction mixture was analysed in a 6% polyacrylamide sequencing gel. After drying the gel was subjected to autoradiography.

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

Five µg total RNA and 1000 c.p.m. labelled DNA were suspended in 80% formamide, 40 mM-PIPES pH 6 to 7, 1 mM-EDTA, 0-4 M-NaCl and denatured at 85 °C for 15 min. After hybridization for 4 h at 48 °C, 300 µl ice-cold SI nuclease digestion buffer [0-28 M-NaCl, 50 mM-sodium acetate, 20 µg/ml denatured herring sperm DNA, 0-1 or 0-5 units/µl SI nuclease (Gibco-BRL)] was added and the mixture was incubated for 30 min at 37 °C. The reaction was terminated by addition of 80 µl termination buffer (25 M-ammonium acetate, 50 mM-EDTA) and the nucleic acids were precipitated by addition of two volumes of ethanol. SI nuclease-protected fragments were electrophoresed on a 6% polyacrylamide sequencing gel. After drying, the gel was subjected to autoradiography.

**Results**

**Localization, cloning and sequencing of the polyhedrin gene**

The location of the polyhedrin gene on the SeMNPV genome was investigated by hybridization of various restriction enzyme digests of SeMNPV DNA with AcMNPV fragment HindIII V as a probe under non-stringent conditions (Fig. 1). This fragment contains approximately 500 nucleotides (nt) from the 3' end of the AcMNPV polyhedrin gene (Smith et al., 1983). An XbaI fragment of 1-7 kb (lane 2) and a SphI fragment of 11-4 kb (lane 3) of SeMNPV DNA gave positive signals with the
SeMNPV polyhedrin gene characterization

Fig. 1. Localization of the SeMNPV polyhedrin gene. Restriction enzyme analysis (a) of AcMNPV DNA digested with EcoRI (lane 1) and SeMNPV DNA digested with XbaI and SphI (lanes 2 and 3, respectively), and Southern blot hybridization (b) under non-stringent conditions with AcMNPV HindIII V as a probe. The sizes of AcMNPV fragment EcoRI I (7.4) and SeMNPV fragments XbaI R (1.7) and SphI D are indicated (in kbp).

Fig. 2. Physical map of an 11.4 kbp segment (a) of SeMNPV DNA and its 2.2 kbp subfragment (b) for restriction enzymes SphI, SspI, XbaI, EcoRI and SalI. The polyhedrin gene is indicated by the solid bar. The start (ATG) and stop (TAA) codons of the polyhedrin gene are indicated. The arrows indicate the direction in which (sub)clones were sequenced, except 1 and 2 where specific primers were used.

AcMNPV probe and were further analysed. The probe also hybridized with EcoRI fragment I of AcMNPV confirming the specificity of the HindIII V (lane 1). SeMNPV XbaI fragment R (1.7 kb) and SphI D are indicated (in kbp).

Fig. 3. Nucleotide sequence of the SeMNPV polyhedrin gene and its flanking regions. The predicted amino acids are indicated with one-letter code designations. Putative transcription initiation and termination signals are underlined.
Fig. 4. Transcriptional analysis of the SeMNPV polyhedrin gene. (a) Northern blot analysis of polyhedrin transcripts. Total RNA was extracted from uninfected *S. exigua* cells (lane 1) and SeMNPV-infected *S. exigua* cells 48 h p.i. (lane 2) and 24 h p.i. (lane 3) and analysed in a 1% agarose gel (5 μg RNA per lane). The polyhedrin transcripts were detected with a **3**P-labelled SeMNPV polyhedrin gene-specific oligonucleotide probe (17-mer) complementary to the mRNA strand. The size of the mRNA is indicated (right); a RNA ladder was used as size markers (left). (b) Size analysis of cDNA prepared by primer extension of an oligonucleotide complementary to the SeMNPV polyhedrin mRNA. The oligonucleotide was 5' end-labelled with [**3**P]dATP, annealed to total RNA from SeMNPV-infected *S. exigua* cells 48 h p.i. (lane 1) and uninfected cells (lane 2) and elongated with reverse transcriptase. The sequence of the polyhedrin promoter (mRNA strand) is indicated by the lettering at the right. The sizes of the extension products were determined against a sequence ladder (AGCT) obtained from an SeMNPV plasmid clone using the same 17-mer extension oligonucleotide as in (a). (c) Nuclease S1 analysis of the 3’ end of the SeMNPV polyhedrin mRNA. End-labelled probes were hybridized to total RNA (5 μg per lane) isolated from uninfected (lanes 1 and 4) or SeMNPV-infected *S. exigua* cells 48 h p.i. treated with 0.1 units (lanes 3 and 6) or 0.5 units (lanes 2 and 5) of S1 nuclease, analysed on denaturing sequencing gels. The sizes of the protected fragments were determined by comparison with a sequencing ladder of M13 ssDNA (M13).

*M* (29K) was similar to the apparent size of SeMNPV polyhedrin as determined by SDS-PAGE (Caballero *et al.*, 1992).

In the 5’ non-coding region two putative transcriptional start sites containing the canonical core TAAG (Vlak & Rohrmann, 1985) are found, one around residue −70 and another around residue −45 with respect to the translational start codon (Fig. 3). At residue −45 this core sequence is part of a dodecanucleotide sequence TGTAAGTAATTT, which is highly conserved among all polyhedrin genes sequenced so far and always contains the start signal for transcription (Rohrmann, 1992). No canonical TATA boxes are found in the region upstream from the two putative transcriptional start sites. Two short, imperfect tandem repeats with the sequence TATC(T)ATCGA were found at residues −85 and −112. Downstream from the putative transcription start site two other conserved motifs (Zanotto *et al.*,...
In the 3' non-coding region of the polyhedrin gene the canonical poly(A) signal sequence (AATAAA) was not found (Fig. 3). However two signal-like sequences were observed, one (ATTAAA) as part of the translational stop signal and a second (AAATTTAA) approximately 10 nt downstream from the first putative signal. Downstream from the polyhedrin-coding sequence the 3' end of another ORF was found.

Transcriptional analysis of the polyhedrin gene

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

Northern blot analysis of total RNA from infected cells isolated at 24 and 48 h p.i. showed (Fig. 4a) an abundant mRNA of approximately 1000 nt that a polyhedrin-specific oligonucleotide (5' TCGTACACG-TAAGTGGC 3') was used as probe (lanes 2 and 3). Longer transcripts were not detected. The specificity of the probe was confirmed by the absence of a signal when RNA from uninfected cells was used (lane 1).

Primer extension was employed to determine the transcriptional start site of the polyhedrin mRNA. A 17-mer oligonucleotide, corresponding to the complement of the coding sequence from residues 36 to 53 (Fig. 3), was used in the primer extension assay of polyhedrin mRNA (Fig. 4b, lane 1). RNA from uninfected cells was used as a negative control (lane 2). An internal *NruI* subclone of fragment *SphI* D served as a positive control for the assay and the sequence (other lanes). The reverse transcription extended as far as the T and A at positions −45 and −43, respectively.

### Table 1. Amino acid and nucleotide sequence identities (%) of 10 NPV polyhedrins

<table>
<thead>
<tr>
<th></th>
<th>SeMNPV</th>
<th>MbMNPV</th>
<th>PfMNPV</th>
<th>AcMNPV</th>
<th>OpMNPV</th>
<th>AgMNPV</th>
<th>BmMNPV</th>
<th>LdMNPV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SeMNPV</strong></td>
<td>98</td>
<td>92</td>
<td>89</td>
<td>85</td>
<td>84</td>
<td>84</td>
<td>82</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>86</td>
<td>82</td>
<td>81</td>
<td>76</td>
<td>77</td>
<td>75</td>
<td>74</td>
<td>73</td>
</tr>
<tr>
<td><strong>MbMNPV</strong></td>
<td>92</td>
<td>89</td>
<td>85</td>
<td>83</td>
<td>83</td>
<td>81</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>82</td>
<td>81</td>
<td>79</td>
<td>77</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td><strong>PfMNPV</strong></td>
<td>82</td>
<td>81</td>
<td>79</td>
<td>77</td>
<td>74</td>
<td>75</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>88</td>
<td>86</td>
<td>84</td>
<td>83</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AcMNPV</strong></td>
<td>90</td>
<td>80</td>
<td>84</td>
<td>84</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>81</td>
<td>77</td>
<td>76</td>
<td>75</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OpMNPV</strong></td>
<td>89</td>
<td>88</td>
<td>86</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>79</td>
<td>76</td>
<td>75</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AgMNPV</strong></td>
<td>95</td>
<td>91</td>
<td>83</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>91</td>
<td>80</td>
<td>80</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BmMNPV</strong></td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Bold and italic lettering denotes amino acid and nucleotide sequence identities, respectively.
AcMNPV  TGGAAATGTCTATCAATATATA---GTTGCTGATATCATGGAGATAATTA~TGATAACCATCTCGCAA
SeMNPV  CATATCTATGGATTAATGCTA---TTGTGTTTATCATGCAGCAGCATATGATATGATGTTGGG
SfMNPV  CATATCTATGGATTAATGCTA---TTGTGTTTATCATGCAGCAGCATATGATATGATGTTGGG

Fig. 6. Computer alignment of the DNA sequences of the 5' sequence flanking the polyhedrin genes of SeMNPV, SfMNPV (Gonzales et al., 1989) and AcMNPV (Smith et al., 1983; Howard et al., 1986). The tandem repeats and TAAG consensus sequence are underlined. Asterisks indicate nucleotide identity.

Comparison of the SeMNPV polyhedrin gene with those of other baculoviruses

The amino acid sequence of the SeMNPV polyhedrin was compared to eight other MNPV polyhedrins (Fig. 5) and found to be most similar to S. frugiperda MNPV (SfMNPV) polyhedrin with 98% amino acid identity and most distantly related to Lymantria dispar MNPV (LdMNPV), only 81% identity (Table 1). The nucleotide sequences of the promoter region of SeMNPV and SfMNPV up to residue -58 were identical (Fig. 6). The sequences upstream from this region showed no similarity.

Discussion

In this report the nucleotide sequence and transcriptional analysis of the SeMNPV polyhedrin gene are presented. Polyhedrin genes are highly conserved among baculoviruses (Vlak & Rohrmann, 1985) and the SeMNPV polyhedrin is no exception. The SeMNPV polyhedrin gene was most closely related to the polyhedrin gene of SfMNPV (Gonzales et al., 1989), differing by only five amino acids (Fig. 5). DNA sequence comparison of SeMNPV and SfMNPV polyhedrins showed a difference in identity in 102 nt, of which only nine were not in the wobble position (data not shown). SeMNPV was most distantly related to LdMNPV (Smith et al., 1988; Chang et al., 1989) which were different with respect to 46 amino acids (Fig. 5).

The similarity between SeMNPV and SfMNPV polyhedrins is higher than with those of any of the other MNPVs (Table 1). The differences are located at sites known to be particularly variable among polyhedrins (Rohrmann, 1986) and are functionally conservative. This suggests that the polyhedrin genes of SeMNPV and SfMNPV have evolved from a common ancestor distinct from other MNPVs. Polyhedrins of SeMNPV and SfMNPV on the one hand are more closely related to Mamestra brassicae MNPV (MbMNPV; Cameron & Possee, 1989), Panolis flammea MNPV (PfMNPV; Oakey et al., 1989) and Orgyia pseudotsugata single NPV (Leisy et al., 1986a) than to O. pseudotsugata MNPV (OpMNPV; Leisy et al., 1986b), Anticarsia gemmatalis MNPV (AgMNPV; Zanotto et al., 1992), AcMNPV (Hooft van Iddekinge et al., 1983), Bombyx mori NPV (BmMNPV; Iatrou et al., 1985) or LdMNPV (Smith et al., 1988; Chang et al., 1989) (Table 1). Availability of polyhedrin protein sequences of other baculoviruses may aid in their classification and may help define baculovirus species.

By primer extension analysis the transcription initiation site for the polyhedrin gene mRNA of SeMNPV was mapped in the promoter within the TAAG motif at approximately residue -44 (Fig. 4c). This motif is conserved in all baculovirus late genes (Rohrmann, 1986; Blissard & Rohrmann, 1990). The major start site
was at the A at -45, but the T at -46 and the A at -44 serve as minor start sites. This heterogeneity at the 5' end of the polyhedrin mRNA also exists in polyhedrin transcripts from other MNPVs as determined for AcMNPV (Hooft van Iddekinge et al., 1983) and OpMNPV (Leisy et al., 1986c) by S1 analysis and for AcMNPV (Howard et al., 1986) and PfMNPV (Oakey et al., 1989) by primer extension analysis. It is therefore unlikely that the upstream TAAG motif at residue -72 is used for SeMNPV polyhedrin gene transcription. Whether this start site is used much earlier in infection, as demonstrated for other late genes such as the gp64 gene of OpMNPV (Blissard & Rohrmann, 1989), requires further investigation.

The promoter regions of the SeMNPV and SfMNPV polyhedrin gene were identical as far as residue -58 (with respect to the translational start). Upstream from the transcriptional start site the sequence was highly diverged (Fig. 6). A similar situation exists in the 5' flanking region of polyhedrins of other closely related MNPVs, such as MbMNPV (Cameron & Possee, 1989) and PfMNPV (Oakey et al., 1989). Two imperfect tandem repeats with the sequence TATC(T)ATCGA were found upstream from the transcriptional start site at residues -85 and -112. Similar repeats have been found in AcMNPV (Hooft van Iddekinge et al., 1983) and SfMNPV (Gonzales et al., 1989). There is, however, little evidence that these repeats contribute to the promoter activity of baculovirus polyhedrins as suggested from (deletion) mutagenesis experiments (Possee & Howard, 1987), although they may have an auxiliary function (Ooi et al., 1989).

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

The canonical poly(A) signal AATAAA (Birnstiel et al., 1985) is not present in the 3' end of the SeMNPV polyhedrin gene. However there are two similar motifs in proximity to the translational stop signal. One overlaps with the TAA (ATTAAA) of the translational stop and the other is located approximately 10 nt downstream. Both may serve as a signal for polyadenylation. The 3' non-coding sequence of the mRNA contained at least an additional 230 nt before polyadenylation could occur (Fig. 4a). This puts a transcriptional stop signal within an AT-rich region, which is almost invariably present in the transcription termination sequence of other baculovirus polyhedrin genes. The variation at the 3' end of the message (Fig. 4c) may be due to microheterogeneity in the transcripts or to a slight 'breathing' of AT base pairs in this region during the S1 nuclease treatment.

In the process of mRNA isolation we were unable to isolate SeMNPV polyhedrin mRNA by oligo(dT)-cellulose chromatography. It is possible that the poly(A) tail, if present at all, is too short to bind efficiently to the oligo(dT). This is supported by the observation that the mRNA size of about 1000 nt as observed in Northern blot analysis is close to the value of the primary transcript of 1020 nt as deduced from primer extension and S1 analysis (Fig. 4). If the SeMNPV polyhedrin mRNA is polyadenylated, then the poly(A) tail is either very small or rapidly degraded in the infected cell. The observation that the poly(A) motif is not in the canonical format (Fig. 3) favours the hypothesis that the SeMNPV polyhedrin mRNA is not polyadenylated. This is in contrast to AcMNPV polyhedrin mRNA, which is polyadenylated (Smith et al., 1983; Howard et al., 1986).

There is a potential downstream ORF on the opposite strand terminating at residue 848 with a putative poly(A) signal at residue 782 (Fig. 3). In AcMNPV an ORF of 1629 residues is present downstream of the polyhedrin gene which is running in the opposite orientation and which is essential for virus replication (Possee et al., 1992). Comparison of our sequencing data indicates that the downstream sequences have no homology with the 1629 nt gene of AcMNPV. Transcripts from the polyhedrin gene and a putative downstream ORF have a considerable overlap in both AcMNPV and SeMNPV. Further sequencing, transcriptional mapping and deletion mutagenesis may elucidate whether the ORF in the downstream region is functionally similar to the AcMNPV gene.

The transcription of SeMNPV polyhedrin mRNA terminated in an AT-rich region within the sequence AATTTTTTTTTT (AAT₉) at residue 970. Similar AT sequences are frequently found at the 3' end of other polyhedrin mRNAs. This AAT₉ motif is observed in SfMNPV (Gonzales et al., 1989), MbMNPV (Cameron & Possee, 1989), PfMNPV (Oakey et al., 1989) and AgMNPV (Zanotto et al., 1992) in roughly the same position, but not in OpMNPV (Leisy et al., 1986b) or AcMNPV polyhedrin mRNA (Possee et al., 1992), although the 3' end of the polyhedrin mRNA in the latter two MNPVs is in an AT-rich region. Computer-assisted alignment of the 3' non-translated region of SeMNPV polyhedrin gene with other polyhedrins also showed limited sequence homology (data not shown). Some structural features, however, have been maintained, such
as the location of (putative) poly(A) signals and transcription termination sequences.

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

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

References

VLAK, J. M. (1979). The proteins of nonoccluded nuclear polyhedrosis virus produced in an established cell line of Spodoptera frugiperda. Journal of Invertebrate Pathology 34, 110–118.


(Received 7 May 1992; Accepted 16 July 1992)