Available online at www.sciencedirect.com
SCIENCE DIRECT®
Virus Research 108 (2005) 213–220
Virus Research
www.elsevier.com/locate/virusres

Short communication

Nucleotide sequence and transcriptional analysis of the *pif* gene of *Spodoptera frugiperda* nucleopolyhedrovirus (SfMNPV)☆

Oihane Simón[a], Serafín Gutiérrez[b], Trevor Williams[a], Primitivo Caballero[a], Miguel López-Ferber[b,*]

[a] *Depto. de Producción Agraria, Universidad Pública de Navarra, 31006 Pamplona, Spain*
[b] *Laboratoire de Pathologie Comparée, UR 1231 INRA, 30380 Saint Christol les Alès, France*

Received 24 February 2004; received in revised form 20 July 2004; accepted 26 July 2004
Available online 18 October 2004

Abstract

Defective viruses, not transmissible alone, increase the transmissibility of complete genotypes in natural populations of *Spodoptera frugiperda* multicapsid nucleopolyhedrovirus (SfMNPV). The defective phenotype is associated with a 15 kb deletion, which includes the *pif* (per os infectivity factor) gene. The sequence of a 2.4 kb fragment that includes *pif* was determined. Multiple transcripts encompassing *pif* were detected by Northern blot analysis. RT-PCR and nuclease protection analysis demonstrated the presence of run-through transcripts starting upstream of *pif*. A 2.0 kb messenger started from a CTAAG promoter motif located 11 nt upstream of the *pif* start codon, and ended 450 nt downstream from the *pif* stop codon. This *pif* mRNA included a small downstream ORF (homologous to Se37). A transcript of 0.8 kb was detected that may correspond to a specific transcript from this small ORF. This transcript would start at a late consensus motif internal to *pif* coding sequences, ending at the same polyadenylation signal as the *pif* transcript. These transcription features resemble those of *pif* transcription in *Spodoptera littoralis* NPV, although the genomic location of *pif* is not equivalent in the two viruses. SfMNPV *pif* can encode a protein of 529 amino acids, closely related to *Spodoptera exigua* MNPV PIF.

Keywords: *Spodoptera frugiperda* MNPV; *pif* gene; Transcription; Sequence

Baculoviruses (family *Baculoviridae*) are enveloped viruses with double-stranded DNA genomes. These viruses are pathogenic to invertebrates, particularly to insects of the Order Lepidoptera. Baculovirus replication is biphasic, involving two phenotypically distinct forms of virions. The occlusion-derived virion (ODV) is adapted for infection of the insect midgut, and is responsible for the transmission of infection from insect to insect in nature. ODVs are embedded in occlusion bodies (OBs). The budded virion (BV) buds out of infected cells and transmits the infection from cell to cell, within the infected host. Both forms of virions are usually produced in a sequential manner in the same cell. When susceptible insect larvae ingest OBs present on contaminated foliage, the OBs dissolve in the alkaline conditions in the insect midgut, and the virions they contain are released. ODVs pass through the peritrophic membrane, attach to columnar epithelial cells and infect them (Granados and Lawler, 1981).

Entry of the ODVs into epithelial cells can occur in two stages: binding of ODVs to the cell membrane followed by fusion of the membranes, suggesting that at least two virion proteins may be involved in the process (Horton and Burand, 1993). Three genes have been identified as responsible for the oral infectivity of ODVs. The product of *Autographa californica multiple nucleopolyhedrovirus* (Genus *Nucleopolyhedrovirus*) (AcMNPV) open reading frame (ORF) 138 (Ayres et al., 1994), named P74 (Kuzio et al., 1989; Faulkner et al., 1997; Haas-Stapleton et al., 2004), and ORF7, named *pif* (per

☆ The sequence of the SfMNPV *pif* gene was deposited with GenBank Data Library under Accession No. AY531110.

* Corresponding author.
E-mail address: lopez@ensam.inra.fr (M. López-Ferber).

0168-1702/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2004.07.013

os infectivity factor) of *Spodoptera littoralis nucleopolyhedrovirus* (SpliNPV) (Kikhno et al., 2002) and *pif-2*, identified in *Spodoptera exigua MNPV* (SeMNPV) (Pijlman et al., 2003). These three genes are conserved among all sequenced baculoviruses. Deletion or disruption of either *p74*, *pif* or *pif-2* results in a complete absence of per os infectivity of ODVs, whereas injection into the haemocoel results in polyhedrosis and insect death. The *pif* gene product is present in very low amounts in the ODV envelope of SpliNPV (Kikhno et al., 2002). A previous study of *pif* transcription in SpliNPV (Gutiérrez et al., 2004) concluded that the low levels of PIF protein may be due to a very low *pif* transcription in that virus. A complex unit of transcription with overlapping mRNA was observed in SpliNPV. It therefore seemed appropriate to determine whether another baculovirus with a different genomic organisation would show a similar pattern of expression for such an essential gene.

Insects of the genus *Spodoptera* spp. are important polyphagous pests in various regions of the world. Various baculoviruses infect *Spodoptera* spp. larvae. Of these, SeMNPV is the most highly specific. Productive infections can only be achieved in *S. exigua* larvae, while others, such as *Spodoptera frugiperda MNPV* (SfMNPV) or SpliNPV can infect several *Spodoptera* species, to variable degrees. SfMNPV has a host range that overlaps that of SpliNPV. Using DNA hybridisation kinetics, Kelly (1977) concluded that among the four NPVs infecting *Spodoptera* spp., SpliNPV and SfMNPV were the least closely related.

The genotypic diversity of a Nicaraguan strain of SfMNPV (SfNIC) (Escribano et al., 1999) has recently been described by López-Ferber et al. (2003). Of the nine genotypes isolated from SfNIC, three are not infectious per os. They survive in the virus population by transcomplementation with other genomes. Heterologous hybridisation indicated that, although *p74* is conserved in all three genotypes, two of them (SfNIC-C and SfNIC-D) have deletions of ∼15 kb in the region of *pif* (López-Ferber et al., 2003); they also lack *pif-2* (Simón et al., unpublished). These variants contribute positively to the overall pathogenicity of the viral population (López-Ferber et al., 2003).

The region containing *pif* in SfMNPV B genotype was determined by heterologous hybridisation (López-Ferber et al., 2003). Hybridisations with SpliNPV *pif* did not allow detection, whereas a clear band was observed using a SeMNPV *pif* probe. The *Eco*RI-N restriction fragment of the SfNIC-B variant containing *pif* was cloned in the pSP70 vector (Promega) (p264.306) (Table 1) and its sequence determined by primer walking (Fig. 1). The extremity of the contiguous fragment, *Eco*RI-I (p264.100) was also sequenced using the same approach.

A schematic representation of the SfMNPV *pif* region is shown in Fig. 1. Five ORFs (ORF2, ORF3, ORF5, ORF6 and ORF7) longer than 150 nt were found, plus two partially sequenced ORFs in the extremities (ORF1 and ORF4). An ORF, ORF2, which contained 1589 nucleotides (from nt 170 to 1759) (Fig. 2), encoded a gene homologous to SpliNPV *pif* (Kikhno et al., 2002), that corresponds to AcMNPV ORF119 or SeMNPV ORF36. This gene was orientated positively according to the physical map of the SfNIC isolate (Simón et al., unpublished). The nucleotide sequence of the gene encoding PIF predicts a protein of 529 aa.

Genes homologous to *pif* are found in all sequenced baculoviruses, and in *Helicoverpa zea* virus 1 (Hzv-1). The taxonomic status of Hzv-1 is not clear, but this virus also retains *p74* and a *pif-2* homologue, suggesting that is related to the Baculoviridae. SfMNPV PIF shares the highest degree of identity at the amino acid level with the SeMNPV PIF, 64% identity (78% similarity). Twenty-five residues are completely conserved in all PIF proteins, among them, 14 cysteines. If Hzv-1 PIF is excluded, the number of absolutely conserved residues increases to 46.

SfMNPV *pif* presented a hydrophobic domain (aa 1–15), that could act as a putative secretion signal or a N-terminal signal anchor (Fig. 3). In addition, AnTheProt (Deléage et al., 1988) predicts a single transmembrane domain between residues 136 and 152, while four putative transmembrane domains were detected for SpliNPV *pif* (Kikhno et al., 2002). Prosite detected four potential N-glycosylation sites, various phosphorylation sites (both casein kinase II and tyrosine kinase), and lipid binding consensus sequences (N-myristoylation and membrane lipoprotein lipid attachment consensus) (Fig. 3). The putative myristoylation sites would only be used if the Gly residue is located at the carboxy terminus, which implies that the protein is cleaved. No such cleavage was detected by Kikhno et al. (2002) in SpliNPV PIF. These sites are therefore unlikely to be functional. Kikhno et al. (2002) did not detect experimentally any glycosylation for SpliNPV PIF, suggesting that the protein does not enter into the lumen of the endoplasmic reticulum, and thus, the N-terminus hydrophobic domain would act like a membrane anchor sequence. They reported that PIF extraction from the membranes was difficult. This observation is compatible with a lipid anchor.

The most conserved portion of the protein, between residues 144 and 174 (13 residues perfectly conserved in all lepidopteran baculoviruses), partially overlaps both the transmembrane domain and the conserved lipoprotein membrane attachment site. This consensus pattern is found in the databases at a low frequency (4.75×10^{-15}). Its conservation in all PIF sequences suggests a functional relevance of this motif, probably related to the role of PIF in midgut infection. An analysis of its role using mutations that disrupt this site would be required. Both the putative transmembrane region and the lipid anchoring would likely contribute to understanding the role of PIF in midgut infection. SpliNPV PIF has been shown to be a structural protein of the ODV envelope. Deletion of SfMNPV *pif* leads to the same phenotype; absence of per os infectivity of ODVs, suggesting that its location may be the same.

A late baculovirus consensus promoter motif, TAAG (Blissard and Rohrmann, 1990) was identified at nt 156. This core sequence was preceded by a C. The SfMNPV *pif* gene

Table 1
The plasmids developed and the location of their inserts in the SfMNPV isolate physical map (indicated by map units, m.u.) or in the described sequence, Gb AY531110 (in nt)

Plasmid	Location	Vector	Purpose
p264.306	24.1–26.8 m.u.	pSP70	Sequencing of *pif*
p264.100	26.8–31.5 m.u.	pSP70	Sequencing of *pif*
p284.1	24.1–25.7 m.u.	pUC19	Sequencing of *pif*
p284.2	25.7–26.8 m.u.	pUC19	Sequencing of *pif*
p293.3	1–2247 nt	pGEM-T	Plasmid construction for PCR amplifications
p290.2	1263–1759 nt	pGEM-T	Riboprobe generation, Northern blot
p291.10	1–236 nt	pGEM-T	Riboprobe generation, RNase protection

is preceded by a partially sequenced ORF (ORF1), homologous to SeMNPV ORF35 characterised as *pif*-2 (Pijlman et al., 2003). The 3′ end of SfMNPV *pif*-2 gene was sequenced. Only 18 nt separate *pif*-2 and *pif*. Downstream from *pif*, a small ORF of 85 aa (ORF3) homologous to SeMNPV ORF37 and a partially sequenced ORF (ORF4) homologous to SeMNPV ORF38 (*fgf*) were found, respectively. The ORF3 was preceded by a late baculovirus consensus motif at nt 1533. Putative polyA signals (ATTAAA or AATAAA) are present at nt 1964 and 2185 in the leading strand and 2219 in the complementary strand. Some other almost perfect consensus (ATAAA) motifs were found at 165, 1921 and 2067 in the leading strand, 203, 1307 and 2082 in the complementary strand. A TATAAA sequence at 165 may correspond to the *pif-2* polyadenylation signal. SfMNPV *pif* may use the 1921 or 1964 signal, whereas ORF3 may use the signals located at 2067 or 2185. Both poly(A) consensus motifs overlap *fgf* coding sequences.

Three minor ORFs are included within *pif* coding sequences, ORFs 5 (nt 865–641), 6 (nt 738–938) and 7 (nt 1387–1184), with coding capacity for 74, 66 and 67 aa, respectively. ORFs 5 and 7 run in the complementary strand, while ORF6 runs in the same strand as *pif* but in another frame. No significant homologies were found for these putative ORFs in the Swiss-Prot database. Three late baculovirus consensus motifs were located in the complementary strand (nt 885, 1875 and 2043). Upstream the translation start codon of the ORF5 a late consensus motif was identified. No such promoter consensus sequences were found in the vicinity of ORFs 6 and 7. In the complementary strand, the consensus

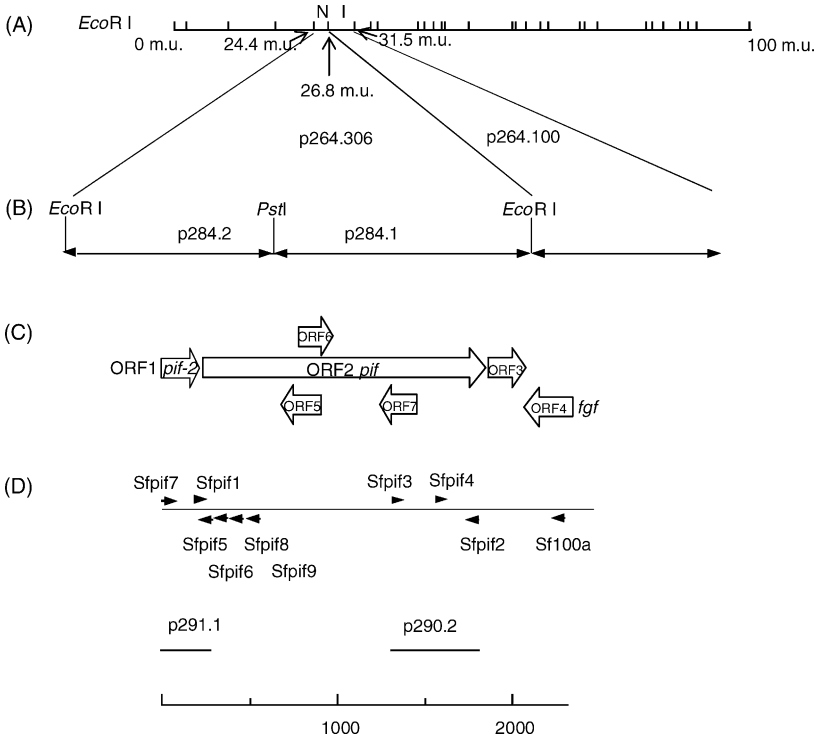

Fig. 1. Position of the *pif* gene on the SfMNPV genomic map and sequencing strategy. (A) *Eco*RI restriction map of SfMNPV-NIC (SfNIC) genome. Numbers below indicate the map units. (B) Detailed physical map of *Eco*RI-N and -I fragments, and plasmid constructions. *Eco*RI-N fragment was cloned into pSP70 plasmid (p264.306), and fragments generated by *Pst*I digestion of *Eco*RI-N fragment into pUC19 (p284.1 and p284.2). *Eco*RI-I was cloned into pSP70 (p264.100) and its left extremity was sequenced. (C) ORFs located in the region are indicated by an arrow. (D) Position of the different primers and probes used.

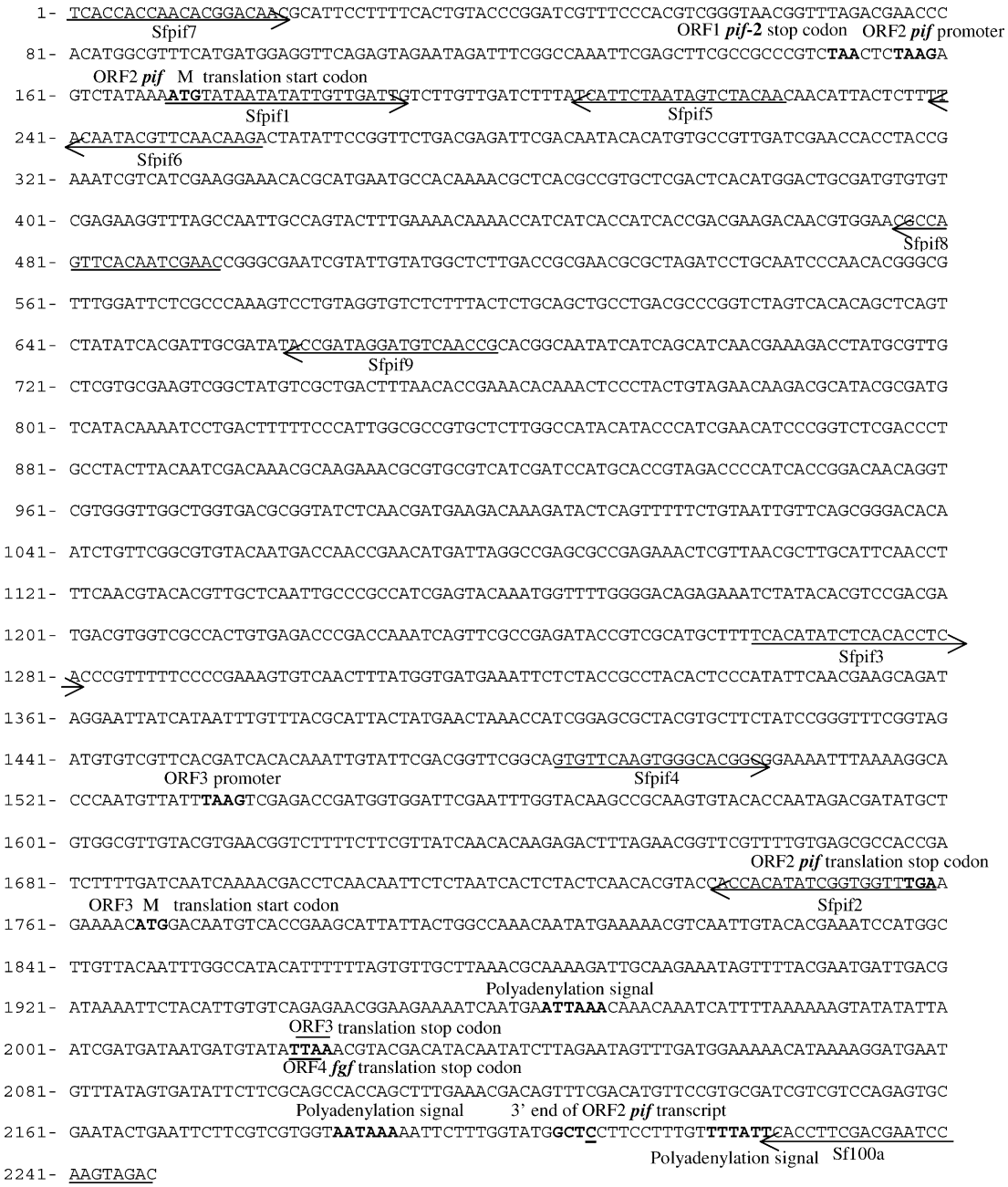

Fig. 2. Sequence of the SfMNPV *pif* region. Partial sequence of the SfMNPV genome *Eco*RI-N and -I fragments starting from Sfpif7 primer, located in the 3′ end of the *pif-2* gene, to the Sf100a primer in the *fgf* gene. Baculovirus consensus late initiation motifs TAAG, polyadenylation signals, start and stop codons SfMNPV *pif-2*, *pif* and the ORF homologue to Se37 (Sf37) and *fgf* transcripts are shown in bold. The sequences of the primers used for transcription analysis are underlined.

located at 2082 or 2219 cannot be used by *fgf*, that ends at 2021. No consensus polyadenylation signals were observed in the vicinity of ORFs 5, 6 or 7.

The putative sequence of SfMNPV PIF clearly places SfMNPV close to SeMNPV, while SpliNPV or AcMNPV are more distantly related. In addition to the sequence data, the gene organization in this region also confirms the degree of relatedness between SeMNPV and SfMNPV. The ORF cluster (*pif* + ORF3 + *fgf*), observed in the majority of the Group II NPVs, is also conserved in SfMNPV, whereas the presence of *pif-2* just upstream, is not a general feature (Gutiérrez et al., 2004). The location of *pif* (and ORF3) in the genomes varies between baculoviruses. In lepidopteran NPVs, *pif* is closely preceded by a late consensus TAAG, often in a non-optimal environment (preceded by C or T). The distance between this putative promoter and the first ATG it is always less than 20 nt. Only in CuniNPV is this distance larger (104 nt). The *H. zea* virus 1 (Hzv-1) is the only virus to lack such a consensus.

Gutiérrez et al. (2004) estimated the relative quantity of the transcript generated by the SpliNPV *pif* gene to be 300

Fig. 3. Protein sequence of SfMNPV PIF (deduced from nucleotide sequence). The residues corresponding to predicted transmembrane domains are highlighted in pink. Residues absolutely conserved in all PIF proteins are in red, those conserved in 75%, in blue, and those over 50%, in green. Underlined residues indicate consensus N-myristoylation sites. Double underlined, consensus membrane lipoprotein attachment site. Baculovirus sequences used in the comparative analysis were GenBank (accession numbers included): *Spodoptera frugiperda* (Sf) MNPV (AY531110), *Spodoptera exigua* (Se) MNPV (NC002169), *Spodoptera littoralis* (Spli) NPV (AF527603), *Spodoptera littura* (Splt) NPV (NC003102), *Mamestra configurata* (Maco) MNPV A (NC003529), *Mamestra configurata* (Maco) MNPV B (NC004117), *Heliothis armigera* (Ha) SNPV (NC003094), *Heliothis armigera* (Ha) SNPV G (NC002654), *Helicoverpa zea* (Hz) SNPV (NC003349), *Autographa californica* (Ac) MNPV (NC001623), *Bombyx mori* (Bm) NPV (NC001962), *Lymantria dispar* (Ld) MNPV (NC001973), *Orgyia pseudotsugata* (Op) MNPV (NC001875), *Epiphyas postvittana* (Eppo) NPV (NC003083), *Rachiplusia ou* (Ro) MNPV (NC004323), *Xestia c-nigrum* (Xc) GV (NC002331), *Phthorimaea operculella* (Phop) GV (NC004062), *Plutella xylostella* (Px) GV (NC002593), *Cydia pomonella* (Cp) GV (NC002816), *Culex nigripalpus* NPV (CuniNPV) and *Helicoverpa zea* virus 1 (Hzv-1, NC004156). The sequences were aligned using ClustalX version 1.7 (Thompson et al., 1997).

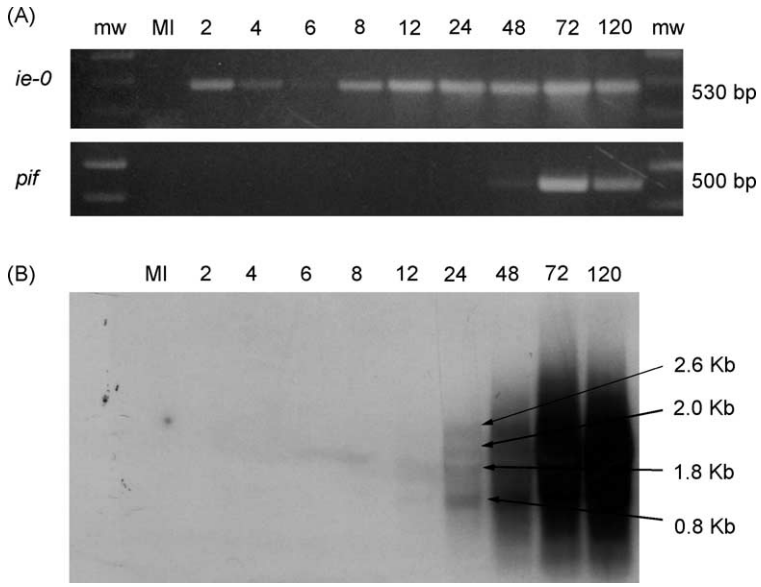

Fig. 4. Temporal expression of SfMNPV *pif* transcripts. Total RNA was extracted from uninfected Sf9 cells (lane MI, mock infected) and SfMNPV infected Sf9 cells at the indicated time (h p.i.). Cells were homogenised in TRIzol (Invitrogen) and the RNA was then extracted following the manufacturer's instructions. The time 0 h p.i. was defined as the moment the virus inoculum was allowed to adsorb to the cells, or the moment that insects consumed the virus inoculum. RT-PCR was performed using the Access RT-PCR System (Promega). RNA samples were treated with DNase prior to RT-PCR. (A) RT-PCR analysis of *pif* gene performed on total RNA extracted from uninfected Sf9 cells (lane MI, mock infected) and SfMNPV infected Sf9 cells at indicated times (h p.i.). Sfpif2 and Sfpif3 primers were used for the amplification, RNA was previously treated with DNase. As control, RT-PCR of *ie-0*, an immediate-early transcribed gene, was performed under the same conditions described above. (B) Northern blot analysis of SfMNPV *pif* transcription using a riboprobe obtained from p290.2 plasmid insert (Table 1). The size of specific hybridisation bands is indicated to the right.

times less than that of the *polyhedrin* gene transcript. The similarity in the promoter environment between *pif* genes in SpliNPV and in SfMNPV strongly suggests that the observations made for SpliNPV concerning the low levels of expression (Kikhno et al., 2002) and transcription (Gutiérrez et al., 2004) are likely to be general features of this gene in other baculoviruses.

The temporal regulation of *pif* transcripts was examined by RT-PCR and Northern blot analysis using total RNA isolated from SfNIC-B infected Sf9 cells (Fig. 4). The very early transcribed gene *ie-0* (Pullen and Friesen, 1995) of SfMNPV was included as reference.

Two primers internal to the *pif* ORF, Sfpif3 and Sfpif2 (Figs. 1 and 2), were designed to amplify a 500 bp fragment internal to *pif*. Total RNA extracted at different points was treated with DNase prior to the RT-PCR assay. Control amplifications were performed to ensure the absence of contaminant DNA in the samples. A single RT-PCR product was obtained of the expected size (500 bp). Amplifications were detected from 48 h p.i., increased at 72 h p.i., and remained at a steady-state level up to 120 h p.i. (Fig. 4A). The same RNA samples were used to amplify a 530 bp fragment of the *ie-0* transcript using two internal primers. The *ie-0* mRNAs were detected from 2 h p.i. Northern blot analysis using the 3′ end of PIF (nt 1263–1759, insert of plasmid p290.2, Table 1) as a probe revealed four transcripts overlapping *pif* that appear late in infection. Both methods confirmed that SfMNPV *pif* behaves as a late gene. The 2.0 kb and 1.8 kb transcripts had a size compatible with that of the hypothetical *pif* messenger (1.6–2.0 kb). However, these transcripts were not the most abundant transcripts detected; a hybridisation band of 0.8 kb was the most abundant transcript.

Nuclease protection assay was carried out using a 256 nt fragment (the insert of p291.10 plasmid, Table 1) as probe. This fragment includes 87 nt complementary to the 5′ end of the *pif* gene and 149 nt complementary to the *pif* upstream sequences. The probe was completely protected, indicating the presence of mRNAs starting well upstream from *pif*. As an alternative technique to map *pif* transcripts, 5′ RACE was employed. Four different primers (Sfpif5, Sfpif6, Sfpif8 and Sfpif9) internal to *pif* ORF were used with total RNA isolated at 120 h p.i. The amplifications obtained with Sfpif8 and Sfpif9 are shown in Fig. 5A. For each 5′ RACE reaction, several products were obtained. The amplified fragments were cloned and several plasmids were sequenced containing amplified fragments from each of the four primers. In all cases, the sequences obtained that included the 5′ region of *pif* (i.e. the specific amplifications), revealed a start of transcription at the C residue, just upstream from the late consensus TAAG (nt 155 in Fig. 2), confirming that this motif acts indeed as the *pif* promoter. This sequence is also conserved in SeMNPV (nt 34402–34406). In SpliNPV, the transcript start was mapped by primer extension to the T located at the equivalent position. The work by Pijlman et al. (2003) suggests that in SeMNPV, the conserved CTAAG motif actually acts as a *pif* promoter. Late promoters have been found to initiate

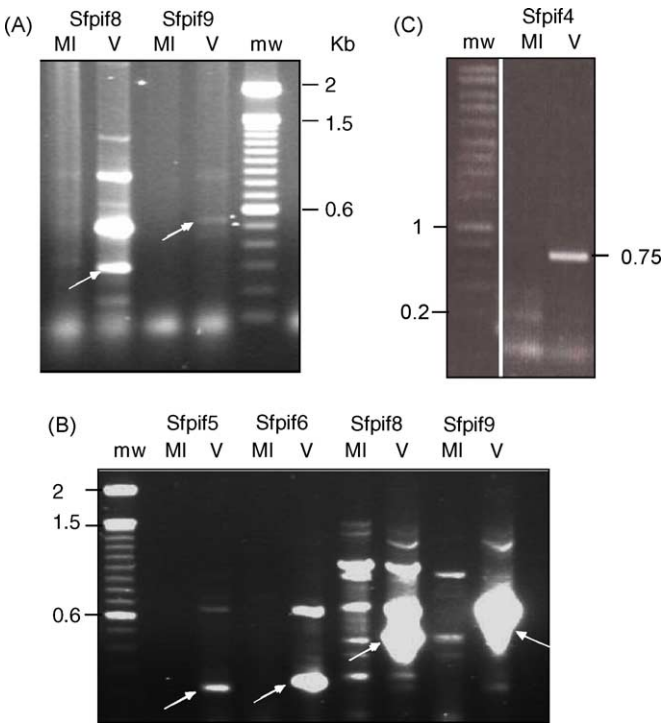

Fig. 5. Transcriptional mapping of the 5′ and 3′ end of the *pif* messenger. Total RNA was extracted from uninfected *S. frugiperda* larvae (MI) and SfMNPV infected larvae at 120 h (V). 5′ RACE RT-PCR was performed using the SMART™ RACE cDNA amplification kit (Clontech, BD Biosciences) employing 1 μg purified total RNA as template. 3′ RACE was performed using the Improm-II™ reverse transcriptase (Promega) and the oligo(dT) anchor primer, following manufacturer's instructions. An aliquot of the reaction was then subjected to PCR amplification with a *Taq* DNA polymerase (Promega). Arrows indicate the amplification observed of expected size. Size standards (mw) are indicated in kb. (A) 5′ RACE analysis using Sfpif8 and Sfpif9 primers. (B) RT-PCR analysis using cDNA mixtures from 5′ RACE and Sfpif7 primer located upstream *pif* promoter, and primers internal to *pif* ORF (Sfpif5, Sfpif6, Sfpif7 and Sfpif8) (see Fig. 1C). (C) 3′ RACE analysis on total RNA extracted from infected larvae, oligodT and Sfpif4 were used in the amplification at 65 °C.

from (A/G/T)TAAG, but to date, none had been found to initiate from CTAAG (Lu and Miller, 1997). This represents the first evidence that the late promoter core TAAG, preceded by a C residue, acts as a functional promoter. In the SeMNPV genome, one unique late promoter motif upstream from the *pif* start codon is present at 13 nt upstream. This late promoter consensus is also preceded by a C. A similar situation is found for the *Mamestra configurata* NPV *pif*, and for other genes, such as SeMNPV *pif-2*.

Both NPA and Northern blot analysis indicated the existence of transcripts initiating upstream from *pif*. To investigate this, PCR was performed using cDNA mixtures obtained from the 5′ RACE experiments, with an specific primer located upstream from the transcription start detected by 5′ RACE (Sfpif7 forward) and within the *pif* gene (Sfpif5, Sfpif6, Sfpif8 and Sfpif9, reverse primers). All amplifications resulted in fragments corresponding to the size expected if mRNAs encompassed both the upstream ORF and the 5′ region of *pif* (Fig. 5B), demonstrating the existence of run-

through transcripts, although their characterization was beyond the scope of this study.

The 3′ end of the *pif* gene transcript was determined by 3′ RACE. RNA was reverse-transcribed using oligo(dT). PCR was then performed with oligo(dT) and Sfpif3 or Sfpif4 at different annealing temperatures to obtain a specific amplification. The 0.75 kb fragment, corresponding to the amplification using Sfpif4 and oligo(dT) primers at 65 °C (Fig. 5C), was cloned into the pGEM-T Easy Vector and sequenced. The 3′ end of the transcript was located at nt 2208, 449 nt downstream from the *pif* stop codon and 24 nt downstream of a consensus polyadenylation signal, AATAAA. The amplification obtained with oligo(dT) and Sfpif3 was not sequenced as it showed a length compatible with the above observation (data not shown). Taken together, 5′ and 3′ terminal mapping predict a *pif* mRNA of approximately 2.0 kb, which is in agreement with the 2.0 kb transcript size determined by Northern blot analysis.

The *pif* mRNA includes the downstream ORF3 (homologous to Se37), that is conserved in all NPVs, giving a bicistronic messenger of 2 kb. However, the presence of the small 0.8 kb messenger suggests that the downstream ORF3 is transcribed alone from a baculovirus late consensus internal to the *pif* ORF, as it has been observed for SpliNPV *pif* (Gutiérrez et al., 2004). The riboprobe used could have detected a putative ORF3 transcript if it started at the late consensus located within *pif* (nt 1533). Such a transcript would end at the poly(A) signal located at nt 2185.

The 1.8 kb transcript observed by Northern blot might be a single *pif* mRNA. However, no putative consensus transcription starts alternative to the one experimentally demonstrated were found in the sequence. The experimental design with four primers would allow the detection of any transcription start located before nt 640. Alternatively, the use of the polyadenylation signal located at nt 1964 would account for a 1.8 kb messenger, but the 3′ RACE experiments did not allow the detection of this hypothetical transcription stop. Alternatively, it could represent ORF3 transcripts not stopping at the polyadenylation signal at nt 2185–2190. The largest transcript (over 2.6 kb) might be the result of a read through from *pif-2*, stopping at the same polyadenylation signal as *pif*, but this has not been characterized.

Two pieces of evidence indicate that *pif* expression does not rely on transcripts starting in upstream genes. First, the difference in *pif* localisation between SfMNPV and other viruses, and in particular SpliNPV. In SpliNPV, transcription originating from upstream genes has also been detected, although the upstream genes are not the same and thus, their regulation is probably different from that of SfMNPV. Second, the structure of the *pif* region in SfMNPV is similar to that of SeMNPV. The *pif-2* gene is located upstream from *pif*. SeMNPV *pif-2* deletion mutants, that conserve the CTAAG motif just upstream from *pif* ATG, do not require *pif* complementation (Pijlman et al., 2003), suggesting that the translation of *pif* does not rely on polycistronic *pif-2* + *pif* mRNAs. More striking is the conservation of the association between *pif* and the downstream ORF, which suggests that their regulation is linked.

The *pif* gene is a conserved gene present in all baculovirus species. However, at the population level, the wild-type SfNIC isolate is naturally diverse in *pif*; not all virus genomes in the population contain this gene (López-Ferber et al., 2003). Infection of individual cells by multiple genotypes is common in baculoviruses (Bull et al., 2001, 2003). The persistence of genomes defective in *pif* can be explained by rescue with undeleted genotypes. However, the presence of the deleted mutants did not reduce the pathogenicity of the SfNIC population. In addition to the low efficiency late promoter driving *pif* transcription, the presence of *pif* defective genomes would contribute to an even greater reduction in the PIF/genome ratio. The results presented here characterise a gene clearly important not only at the level of the individual genotype, as its absence leads to a non infectious ODV, but also at the level of the viral population.

Acknowledgements

We thank B. Limier, L. Croizier, N. Grard and J. Jolivet (INRA) and N. Gorria and I. Maquirriain (UPNA) for technical assistance. O.S. received a Spanish Ministry of Education and Culture studentship and S.G. received a “Formación de Investigadores” fellowship from the Gobierno Autonómico Vasco.

References

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

Blissard, G.W., Rohrmann, G.F., 1990. Baculovirus diversity and molecular biology. Annu. Rev. Entomol. 35, 127–155.

Bull, J.C., Godfray, H.C., O’Reilly, D.R., 2001. Persistence of an occlusion-negative recombinant nucleopolyhedrovirus in *Trichoplusia ni* indicates high multiplicity of cellular infection. Appl. Environ. Microbiol. 67, 5204–5209.

Bull, J.C., Godfray, H.C., O’Reilly, D.R., 2003. A few-polyhedra mutant and wild-type nucleopolyhedrovirus remain as a stable polymorphism during serial coinfection in *Trichoplusia ni*. Appl. Environ. Microbiol. 69, 2052–2057.

Deléage, G., Clerc, F.F., Roux, B., Gautheron, D.C., 1988. ANTHEPROT: a package for protein sequence analysis using a microcomputer. Cabios 4, 351–356.

Escribano, A., Williams, T., Goulson, D., Cave, R.D., Chapman, J.W., Caballero, P., 1999. Selection of a nucleopolyhedrovirus for control of *Spodoptera frugiperda* (Lepidoptera: Noctuidae): structural, genetic, and biological comparison of four isolates from the Americas. J. Econ. Entomol. 92, 1079–1085.

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

Granados, R.R., Lawler, K.A., 1981. In vivo pathway of *Autographa californica* baculovirus invasion and infection. Virology 108, 297–308.

Gutiérrez, S., Kiknho, I., López-Ferber, M., 2004. Transcriptional and promoter analysis of the per os infectivity factor gene of *Spodoptera littoralis* nucleopolyhedrovirus. J. Gen. Virol. 85, 331–341.

Haas-Stapleton, E.J., Washburn, J.O., Volkman, L.E., 2004. P74 mediates specific binding of *Autographa californica* M nucleopolyhedrovirus occlusion derived virus to primary cellular targets in the midgut epithelia of *Heliothis virescens* larvae. J. Virol. 78, 6786–6791.

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

Kelly, D.C., 1977. The DNA contained by nuclear polyhedrosis viruses isolated from four *Spodoptera* sp. (Lepidoptera, Noctuidae): genome size and homology assessed by DNA reassociation kinetics. Virology 76, 468–471.

Kikhno, I., Gutiérrez, S., Croizier, L., Croizier, G., López-Ferber, M., 2002. Characterization of *pif*, a gene required for the per os infectivity of *Spodoptera littoralis* nucleopolyhedrovirus. J. Gen. Virol. 83, 3013–3022.

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

López-Ferber, M., Simon, O., Williams, T., Caballero, P., 2003. Defective or effective? Mutualistic interactions between virus genotypes. Proc. R. Soc. London Ser. B 270, 2249–2255.

Lu, A., Miller, L.K., 1997. Regulation of baculoviruses late and very late gene expression. In: Miller, L.K. (Ed.), The Baculoviruses. Plenum Press, New York, pp. 193–216.

Pijlman, G.P., Pruijssers, A.J., Vlak, J.M., 2003. Identification of *pif-2*, a third conserved baculovirus gene required for per os infection of insects. J. Gen. Virol. 84, 2041–2049.

Pullen, S.S., 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. Virol. 69, 156–165.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.