Molecular analysis of the p48 gene of *Choristoneura fumiferana* multicapsid nucleopolyhedroviruses CfMNPV and CfDEFNPV

Xing Li,† Hilary A. M. Lauzon, Sardar S. Sohi, Subba R. Palli, Arthur Retnakaran and Basil M. Arif

Great Lakes Forestry Centre, Canadian Forest Service, PO Box 490, Sault Ste Marie, Ontario, Canada P6A 5M7

Attempts were made to linearize the DNA of *Choristoneura fumiferana* (Cf) multicapsid nucleopolyhedrovirus (MNPV), in order to improve the efficiency of generation of recombinant viruses after transfection. A unique site for the restriction enzyme *Sse*8387I was found in ORF p48. The requirement for this ORF during virus replication was investigated by molecular analyses including sequencing, transcriptional analysis and inactivation by insertion of marker genes. Sequence analysis showed that ORF p48 consists of 1233 nucleotides encoding a potential protein of 47–88 kDa. The proteins encoded by ORF p48 from CfMNPV and *Orgyia pseudotsugata* MNPV contain 411 amino acids while that from CfDEFNPV (a virus that is defective for infection by the *per os* route) is slightly smaller, at 408 amino acids. Transcriptional and primer extension analyses showed that the mRNA is initiated from a typical baculovirus late gene ATAAG motif. The mRNA was detected at 24 h post-infection (p.i.), reached maximum levels at 48 h p.i. and declined by 96 h p.i., which confirmed the late property of the gene. Inactivation of the gene was attempted by inserting a cassette containing either the gene encoding β-galactosidase or that encoding green fluorescent protein. Blue or fluorescent green plaques of infected cells were observed after transfection. Attempts to generate a plaque-purified virus were not successful. Restriction enzyme analysis showed that the marker genes were inserted randomly at positions other than the p48 locus. This indicated that the gene may be needed for virus replication. The gene is relatively well conserved among baculoviruses but its function remains unclear.

Introduction

The eastern spruce budworm, *Choristoneura fumiferana*, is one of the most devastating insect pests of North American forests. The multicapsid nucleopolyhedrovirus (MNPV) isolated from this insect (CfMNPV) is slow acting and, in its present form, is not an economically viable control agent. A second virus, CfDEFNPV, which is genotypically distinct from CfMNPV and is defective in its ability to infect *Choristoneura fumiferana* larvae by the *per os* route, has also been isolated previously from the spruce budworm. Several recombinant CfMNPV have been generated that contain transgenes that are potentially deleterious to the host. In the case of the *Autographa californica* (Ac) MNPV, the generation of recombinant viruses has been dramatically enhanced by linearization of the viral DNA prior to transfection. This strategy depends on the existence or the introduction of a unique restriction site in a non-essential region of the viral DNA (Kitts *et al*., 1990). In an effort to maximize the number of virus recombinants, the CfMNPV DNA was treated with the available restriction endonucleases that recognize 8 bp sites. All but one of these endonucleases were found to cut the virus DNA more than once. Only the enzyme *Sse*8387I cut the virus DNA once, in ORF p48. The fragment containing this ORF was cloned and sequenced on both strands. The p48 promoter contained motifs consistent with it being a late baculovirus gene. Transcriptional analysis indicated that the putative gene was transcribed late in the infection process. Attempts to generate a virus containing a reporter gene, *lacZ* or *gfp*, in place of p48 were not successful, suggesting that the gene may be required for virus replication. The p48 ORF is also present in the genomes of *Orgyia pseudotsugata* (Op) MNPV (Russell &
Methods

Viruses and cell lines. The plaque-purified CfMNPV (Ireland strain) used in this study has been described previously (Arif et al., 1984). FPMI-CF-203 cells permissive to CfMNPV were maintained in Insect-Xpress medium (Bio/Whittaker) supplemented with 10% foetal bovine serum (FBS) (Gibco BRL). CfDEFNPV, a plaque-purified NPV originally isolated from a wild-type CfMNPV population, is defective in its ability to infect larvae by the per os route but is capable replicating in FPMI-CF-70 cell line. CF-70 cells were maintained in Grace’s medium (Gibco BRL) supplemented with 10% FBS and 0.25% tryptose broth.

Restriction enzyme digestions, DNA cloning and sequence analysis. Viral DNA samples, purified from occlusion bodies (King & Possee, 1992), were treated with restriction enzymes under the conditions recommended by the manufacturer and analysed in agarose gels without enzyme treatment (uncut, lane 1) and after Sse8387I digestion (SseI, lane 2). (B) Viral DNA was digested with BglII (lane 3) or double-digested with BglII and Sse8387I (lane 4). (C) Viral DNA was digested with BamHI (lane 5) or double-digested with BamHI and Sse8387I (lane 6). Lanes 5 contain a 1 kb ladder (Gibco BRL). Letters beside the lanes indicate the fragments produced by the enzymes and their sizes are summarized in (D), based on the total genome size of 125.5 kb (Arif et al., 1984).

Primer extension and Northern blot hybridization analyses. CF-203 cells were infected with CfMNPV at an m.o.i. of 8. At various times post-infection (p.i.), mRNA samples were extracted and purified by using the mRNA purification kit (Gibco BRL). A 33 base primer (EP48) was designed to generate an extension product for the identification of the transcriptional start of ORF p48. The primer (5′ AGTATGGGATATCGATCTCATGATGGTACATTTG 3′) binds a region 88–120 nt downstream of the ATG translational start site. Primer extension experiments were carried out as described by Sambrook et al. (1989) and the products were analysed on denaturing 6% polyacrylamide gels. For Northern blot analysis, mRNA samples were electrophoresed on 1% formaldehyde gels and blotted onto a nylon membrane (Amersham). A PCR product (376 bp) amplified from within ORF p48 by a pair of primers (R-6II, 49–66 nt downstream of the ATG start site, 5′ CAGCAGGGGCACCGCTTC 3′; and F-2, 408–425 nt, 5′ CAGGACACTTTGCCCAGG 3′) was labelled with [33P]dATP and hybridized to the Northern blots at 42 °C in 50% formamide and washed at 65 °C in 0.1 × SSC and 0.1% SDS.

Construction of p48 null recombinant viruses. The transfer vector pCF48.lacZ was constructed by digesting the plasmid pCFHindO2 with NsiI (Fig. 2C). A 4.3 kb cassette containing the gene encoding β-galactosidase (lacZ) under the control of the AcMNPV polyhedrin promoter was removed from plasmid pP10 (a gift from C. Richardson, Biotechnology Research Institute, Montréal, Quebec, Canada) by digestion with NsiI and then ligated into the plasmid pCHindO2. CF-203 cells were transfected with pCF48.lacZ and CfMNPV DNA by using lipofectin (Gibco BRL) (King & Possee, 1992). Virus plaques expressing β-galactosidase were stained blue by addition of X-Gal to the low-melting-point agarose overlay.
In order to engineer a virus that expressed green fluorescent protein (GFP), an 1170 bp cassette containing *AgrAI* restriction sites at both ends was amplified by PCR from the plasmid template pCfPoly.GFP (a gift from J. Barrett, Great Lakes Forestry Centre, Sault Ste Marie, Ontario, Canada). In this cassette, the gene encoding GFP is under the control of the CfMNPV polyhedrin promoter (J. Barrett, personal communication).

The product was digested with *AgrAl* and cloned into the same site located near 5' end of pCfEcoRN (Fig. 2C). CF-203 cells were transfected with the vector (pCf48.GFP) and CfMNPV DNA. Green fluorescent plaques were observed under a Leica fluorescence microscope.

At least three rounds of plaque purification were carried out before progeny viruses were amplified. Putative recombinant viral DNA samples were extracted from budded viruses (King & Possee, 1992). *BglII* and *BamHI* digestions were carried out to verify whether the reporter cassette fragments were inserted at the proper positions in the putative recombinant viruses. Southern blot hybridization with *lacZ*- or gfp-specific probes and PCR amplifications were also used to confirm the location of the inserts in progeny recombinant viruses.

**Results and Discussion**

The CfMNPV genome contains a unique restriction enzyme site (*Sse8387I*).

The ideal strategy to generate a linear virus DNA for transfection purposes is to insert a restriction enzyme rec-

---

**Fig. 2.** *BglII* and *BamHI* restriction maps of CfMNPV and the location of ORF p48. (A) Linearized physical maps and orientations of the polyhedrin gene and ORF p48. The *Sse8387I* site is indicated by open arrowheads. (B)–(C) Location of the fragments *HindIII-O2* (B) and *EcoRI-N* (C) within *BamHI-D*. Plasmids containing the fragments *HindIII-O2* (pCfHindO2) and *EcoRI-N* (pCfEcoRN) were used to construct the transfer vectors. The restriction sites indicated in (C) were used to insert the report gene cassettes (see text). (D) The ORFs present in the six possible reading frames of the this part of the genome are shown. The sequence of *EcoRI-N* (1983 bp, GenBank accession number AF067799) contains the complete ORF p48 (solid box) and other ORFs (those > 60 amino acids are shown).

---

**Fig. 3.** Comparison of 5' flanking sequences of ORF p48 from five NPVs. The translation start site of ATG is in boldface and boxed. Identical nucleotides among the five NPVs are shaded. The motifs TATA, ATAAG, ATCGTT, CAGT are boxed. Boldface underlined letters within ATAAG motifs are the transcription initiation sites established by primer extension experiments.
Fig. 4. Primer extension and Northern blot analyses of ORF p48 of CfMNPV. (A) Primer extension experiment. mRNA samples were extracted from cells (0, 6, 12, 24, 36, 48 h p.i.), hybridized to an end-labelled primer and extended products were separated on a denaturing 6% polyacrylamide gel. Lane 5 contains labelled DNA size markers (Promega). The arrow to the right indicates the 141 nucleotide extension product. (B) Northern blot analysis. mRNA samples were prepared from 0–96 h p.i. as indicated, blotted and probed with a randomly labelled PCR product [indicated by DNA in (C)]. The numbers to the left are the sizes (kb) of the RNA markers (Gibco BRL). The arrows to the right indicate the 3·2- and 2·1 kb transcripts. (C)–(E) The cluster of ORFs and transcripts of p48 in three NPVs. In CfMNPV (C), three complete ORFs, p82, p48 and p12, and a partial ORF, p40, are arranged in tandem. The arrow with an asterisk indicates the 141 bp primer extension product. The 3200 bp transcript is shown underneath. In OpMNPV (D), five ORFs, p87, p48, p12, p40 and p6-5, are shown. The complementary RNA (cRNA, with direction shown by arrow) probe for p6-5 was used in Northern blot analysis (Russell & Rohrmann, 1990) and detected a 3200 bp transcript (shown underneath) originating from ORF p48. In AcMNPV (E), the cRNA probe (direction is shown by arrow) of p80 detected a 3100 bp transcript originating from p48 (Lu & Carstens, 1992).
ognition site that is not normally present into a non-essential region of the genome. Digestion of the virus DNA with this specific enzyme would generate a linear molecule suitable for transfections. The presence of rare restriction sites in the CfMNPV DNA was examined by screening a number of enzymes requiring 8 bp recognition sequences. All the enzymes tested digested the DNA at more than one site with the exception of one enzyme, Sse8387I, which generated a single virus DNA fragment, indicating that the enzyme either did not cut the DNA or cut it only once (Fig. 1A, lanes 1 and 2). Lane 1 represents uncut viral DNA but may also contain some nicked material. The DNA was then digested with Sse8387I and BglII (Fig. 1B). The latter enzyme cuts the DNA into four fragments (lane 3). The double digestion experiment showed that the fragment BglII-A was cut further into two fragments by Sse8387I, indicating that the latter cut the virus genome only once (lane 4). The results were also confirmed by double digestion of the DNA with BamHI and Sse8387I (Fig. 1C), in which the BamHI-D fragment was cut into two by Sse8387I (lane 6). The apparent submolar bands in Fig. 1(C), lane 2, can be identified as incomplete digestion products, since the virus had previously been plaque-purified and such bands did not appear consistently. Southern blot hybridization showed that Sse8387I did indeed cut the genome only once in the fragments BglII-A and BamHI-D (Fig. 2A) and that no small fragments were generated (data not shown).

The unique enzyme site is in ORF p48

Further analysis demonstrated that the unique Sse8387I site was in the fragments HindIII-O2 and EcoRI-N (Fig. 2B). These fragments were sequenced entirely in both directions and revealed that the unique site was in ORF p48 (Fig. 2C). ORF p48 contains 1233 nucleotides and can encode a protein with molecular mass of 47–88 kDa. It is in the opposite orientation to the polyhedrin gene in the CfMNPV genome (Fig. 2A). Immediately upstream of ORF p48 is the p82 gene, which encodes a capsid-related virus structural protein (Li et al., 1997). There are two other ORFs, p12 and p40, downstream of p48 in the fragment EcoRI-N (Fig. 2C). The 5′ end of ORF p12 has a 20 nucleotide overlap with the 3′ end of p48. Homologues of ORF p48, p12 and p40 are present in the same tandem arrangement in the genomes of OpMNPV (Russell & Rohrmann, 1990), AcMNPV (Lu & Carstens, 1992) and CDEFNPV (see below). The EcoRI-N fragment contains three other ORFs, 1, 2 and 3, which could potentially encode proteins of more than 60 amino acids (Fig. 2C). The last three ORFs did not share similarity to baculovirus sequence data deposited in the GenBank.

Analysis of the 5′ flanking area of p48 (Fig. 3) showed that the nucleotides at the −3 and +4 positions relative to the ATG translation start did not conform to the Kozak rule (PuNNATGpu) for efficient translation (Kozak, 1986). There was a baculovirus late gene promoter present (ATAAG, −18 to −22 nt), which was confirmed to be the mRNA initiation start site in CfMNPV (see below). There were two TATA boxes upstream of the ATG start site (at −12 to −15 nt and at −104 to −107 nt in CfMNPV, Fig. 3), a cap site (CAGT, at −56 to −59 nt) and an enhancer-like element (ATCGTT, at −98 to −103 nt). The enhancer-like elements are early transcription signals, but they may not play a role in the transcription of a late baculovirus gene (Ayres et al., 1994). There was no polyadenylation motif, AATAAA, within 500 nt downstream of the ORF p48 stop codon.

Transcriptional analysis

Primer extension analysis of CfMNPV ORF p48 was conducted to identify the mRNA initiation site in the 5′ untranslated region. The primer EP48, complementary to the sequence downstream of the p48 ATG start codon (see Methods), was synthesized, end-labelled and hybridized to purified mRNA samples. The extended products were analysed on a denaturing 6% polyacrylamide gel. A 141 bp extension product was detected at 36 and 48 h p.i. (Fig. 4A), indicating that the p48 mRNA starts at the T residue within the late promoter motif ATAAG 21 bp upstream of the ATG start site (position −21) (Fig. 3C, CfMNPV). The temporal pattern of mRNA expression was investigated by collecting samples at various times after inoculating CF-203 cells with CfMNPV and Northern blot analysis was performed by using a p48-specific DNA probe (372 bp). As shown in Fig. 4B, two mRNAs of 3200 bp and 2100 bp appeared at 24 h p.i., peaked at 48 h p.i. and had declined by 96 h p.i. (Fig. 4B). Both mRNAs were larger than the theoretically expected 1500 bp product, suggesting that they terminated beyond ORF p48 (Fig. 4C). In the case of OpMNPV, 3′ S1 mapping experiments of ORF p48 showed that the transcripts co-terminated with ORF p6·5 located about 1500 bp downstream (Russell & Rohrmann, 1990). A p6·5-specific cRNA probe detected a 480 bp transcript of p6·5 as well as another of 3200 bp at 36 h p.i., which was assumed to be from ORF p48 (Fig. 4D; Russell & Rohrmann, 1990). In a study of AcMNPV, two cRNA probes (sense and anti-sense) to the p80 gene were prepared for gene transcription analysis (Lu & Carstens, 1992). The sense cRNA probe detected three transcripts (3·1, 4·7 and 6·8 kb) from 9 to 24 h p.i. The 3·1 kb transcript (Fig. 4E), detected at 9 h p.i. and increasing in abundance by 24 h p.i., was attributed to ORF p48 (Lu & Carstens, 1992). The smaller transcript may have been initiated by p48 but the possibility that it is a product of ORF2 or 3 cannot be discounted here. Transcriptional analyses of p48 from the three NPVs showed that the transcripts were expressed abundantly late p.i. and that the mRNAs always co-terminated with other downstream ORFs. This transcription pattern, conserved among CfMNPV, OpMNPV and AcMNPV, suggests that p48 may play a role in concert with other downstream genes during virus replication. Only small amounts of the 3200 bp transcript were synthesized at 24 h p.i.
Sequence comparisons

Homologues of ORF p48 of CfMNPV have been found and appear to be relatively conserved in the genomes of OpMNPV, AcMNPV and BmNPV. To verify whether this held true for other NPVs, we used a p48-specific probe to screen a CfDEFNPV genomic library under low-stringency hybridization conditions and found a p48 homologue in this virus genome. It was cloned and sequenced in both directions. ORF p48 of CfDEFNPV contains 1224 nucleotides, which is 9 nucleotides shorter those of CfMNPV and OpMNPV (Table 1), and did not contain an Sse8387I site. All five p48 genes from baculoviruses (CfMNPV, CfDEFNPV, OpMNPV, AcMNPV and BmNPV) did not conform to Kozak’s rules at the ATG codon (Fig. 3). The first 25 nucleotides upstream of the ATG translational signal appear to be quite conserved among the five NPVs (Fig. 3), indicating that control of transcription of this gene is conserved. Transcripts are initiated from within the ATAAAG motif of CfMNPV, AcMNPV and OpMNPV. In OpMNPV, transcripts may be initiated one nucleotide upstream of the motif (Russell & Rohrmann, 1990). A potential cap site (CAGT) and an enhancer-like element (ATCGTT) were found in the genes of CfMNPV, AcMNPV and BmNPV (Fig. 3).

Alignment of amino acid sequences from the five p48 ORFs showed 48.0% and 69.3% identity and similarity, respectively, overall (Fig. 5). The conserved amino acids were distributed throughout the ORFs. Pairwise comparisons of the five sequences showed that the amino acid identities were all above 50% (Table 1), implying a conserved gene structure. CfMNPV, CfDEFNPV and OpMNPV appear to be more closely related to each other than to AcMNPV and BmNPV. CfDEFNPV also had the same gene arrangement of p40–p12–p48–p82 in this area (data not shown) as the other NPVs (Fig. 4). All these data suggest that ORF p48 and the gene organization around it are well conserved among baculoviruses.

Requirement for ORF p48 in virus replication

Essential genes are required for virus replication and cannot be inactivated without deleterious effects on the virus. The requirement for ORF p48 in CfMNPV replication was investigated by attempting to insert either a cassette containing the lacZ gene into the NsiI site or one containing the gene encoding GFP into the AgrAI site (Fig. 2C). Both of these large insertions were expected to inactivate the gene. CF-203 cells were co-transfected with viral DNA and a transfer vector and then progeny viruses were collected and plaque purified. In both cases, blue-coloured or green fluorescent plaques appeared in CF-203 cells. Positive recombinant viruses were amplified after a minimum of three rounds of plaque purification. Viral DNA was extracted and analysed by restriction endonuclease digestion and agarose gel electrophoresis. In both cases, the restriction enzyme patterns showed that the reporter gene was not inserted at the expected location in the viral genome (data not shown). All attempts to insert either of the marker genes into ORF p48 were unsuccessful. The plasmids containing the marker genes appeared to have been integrated randomly into the viral genome at loci other than p48 and replicated along with the viral DNA. At least eight plaques were picked for each marker and their virus DNA was analysed. In all cases, insertion of the marker genes into ORF p48 was not seen. A recent study (Wu et al., 1999) on the interaction between plasmids and virus DNA during co-transfection experiments showed that the plasmid DNA was

---

**Table 1. Comparison of the putative P48 polypeptides from CfMNPV, CfDEFNPV, OpMNPV, AcMNPV and BmNPV**

<table>
<thead>
<tr>
<th></th>
<th>CfMNPV</th>
<th>CfDEFNPV</th>
<th>OpMNPV</th>
<th>AcMNPV</th>
<th>BmNPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acids</td>
<td>411</td>
<td>408</td>
<td>411</td>
<td>387</td>
<td>387</td>
</tr>
<tr>
<td>Molecular mass (kDa)</td>
<td>47.9</td>
<td>47.9</td>
<td>47.9</td>
<td>45.3</td>
<td>45.4</td>
</tr>
<tr>
<td>pI</td>
<td>8.7</td>
<td>8.9</td>
<td>8.9</td>
<td>8.9</td>
<td>8.8</td>
</tr>
</tbody>
</table>

(Fig. 3 B), which probably explains why it was not detected at that time by primer extension analysis (Fig. 3 A).
Fig. 5. Alignment of deduced P48 amino acid sequences from CfMNPV, CfDEFNPV, OpMNPV, AcMNPV and BmNPV. Identical amino acid residues are shown by asterisks, conserved amino acid substitutions are shown by dots and gaps are shown by dashes.

The p48 gene of CfMNPV and CfDEFNPV

integrated into the virus genome after transfection. It was not dependent on the presence of any specific virus sequences in cis and these plasmids were stably present after serial passage. Our data corroborate this conclusion and show that plasmid DNA can be stably integrated into a baculovirus genome. Attempts were then made to clone the reporter gene cassette containing GFP directly into the virus DNA. The GFP cassette was ligated directly into virus DNA that had been linearized by Sse8387I. Green fluorescence was observed in the primary transfection cultures but could not be detected after plaque purification. This indicated that no recombinant viruses were produced by the direct cloning method. This direct cloning method has been used for insertion within the AcMNPV genome at non-essential gene loci, such as the polyhedrin or egt genes (Ernst et al., 1994; Lu & Miller, 1996). The green fluorescence observed in our primary transfection cultures was probably due to transient expression of GFP. The collective data presented here suggest that ORF p48 may be needed for virus replication and attempts to inactivate it would result in an abortive infection. A similar conclusion, based on the inability to generate a gp34.8 null AcMNPV, was reached by Wu & Miller (1989).

Even though sequence analysis showed that the mRNA may not be translated efficiently, it was decided to prepare antiserum against the putative protein encoded by p48 of CfMNPV. Part of the p48 ORF was cloned in-frame into pMAL-c2 expression vector (New England Biolabs) and a fusion protein of the expected size was produced and used to generate antibodies in rabbits. At a dilution of 1:1000, the antibody reacted strongly with the bacterial fusion protein (not shown). For immunoblot analysis, cells were infected with CfMNPV and collected at various times p.i. The samples were processed for SDS–PAGE, blotted onto nitrocellulose membranes and incubated with the anti-fusion protein antiserum.
The polyclonal antiserum failed to detect any viral protein produced in infected cells (data not shown). The same immunological method was used successfully to detect the P82 protein in infected cells (Li et al., 1997). Thus, the amount of protein expressed in infected cells was probably below the threshold of immunoblot detection. More sensitive methods, e.g. pulse-labelling combined with immunoprecipitation, may help to detect such a poorly expressed protein, if indeed any protein is expressed.

It is suggested that ORF p48 may be needed for virus replication, since all attempts to inactivate it by inserting reporter gene cassettes were unsuccessful. The results of transcriptional and immunochemical analyses suggest that very little or no protein is produced from the mRNA of ORF p48. The data indicate that p48 may have a regulatory function (L. King, personal communication). Similarly, a CfMNPV virus DNA lacking ORF p48 was generated and was found to be non-infectious to Sf21 cells on its own but could be rescued then be investigated.

The presence or absence of a plasmid containing this ORF could be generated and its infectivity to CF-203 cells in the cation). Similarly, a CfMNPV virus DNA lacking ORF p48 was generated and was found to be non-infectious to Sf21 cells on its own but could be rescued then be investigated.

More definitive proof of the requirement for ORF p48 could be boost the synthesis of late adenovirus proteins (Cann, 1997). The polyclonal antiserum failed to detect any viral protein expressed in infected cells (data not shown). The same immunological method was used successfully to detect the P82 protein expressed in infected cells (Li et al., 1997).

We thank Drs C. Richardson and J. Barrett for the reporter gene cassette plasmids and A. Pang for preparing antibodies. This research was supported by a grant from the Biotechnology Strategy Fund and by the Pest Management Methods Network of the Canadian Forest Service.

References


Received 18 December 1998; Accepted 18 March 1999