The *p10* gene of *Bombyx mori* nucleopolyhedrosis virus encodes a 7.5-kDa protein and is hypertranscribed from a TAAG motif

VIKAS B. PALHAN and KARUMATHIL P. GOPINATHAN*

Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India

Abstract

In baculovirus-based high-level expression of cloned foreign genes, the viral very late gene promoters of *polyhedrin (polh)* and *p10* are extensively exploited. Here we report the cloning and characterization of the *p10* gene from a local isolate of *Bombyx mori* nucleopolyhedrosis virus (BmNPV). The gene harbours a 213-bp open reading frame encoding a protein of 70 amino acids with a predicted molecular mass of 7.5 kDa. The BmNPV *p10* showed deletion of a single A at +210 nucleotide compared to the prototype baculovirus, *Autographa californica* multinucleocapsid nucleopolyhedrosis virus (AcMNPV), *p10* gene, resulting in a translational frameshift to generate a termination codon and consequently a truncated polypeptide instead of the 10-kDa protein. This protein P7.5 from BmNPV has a putative leucine zipper dimerization motif towards the N-terminal end and the central nuclear disintegration domain but the carboxy-terminal domain implicated in protein association for fibrillar structure formation is absent. Phylogenetic analysis revealed that *p10* is highly conserved among baculoviruses and the BmNPV strains are more closely related to AcMNPV than other baculoviruses. The transcription of *p10* is regulated in a temporal manner, reaching maximal levels by 72 h post-infection. RNAase protection and primer extension analysis mapped the transcription start sites at −70 and −71 nt with respect to the ATG, within the conserved baculovirus late gene motif TAAG. The upstream region showed complete homology to the strong promoter of the AcMNPV *p10*, suggesting that this promoter from BmNPV could also be exploited for high-level expression of cloned foreign genes in silkworm cells or larvae.

[Palham V. B. and Gopinathan K. P. 2000 The *p10* gene of *Bombyx mori* nucleopolyhedrosis virus encodes a 7.5-kDa protein and is hypertranscribed from a TAAG motif. *J. Genet.* **79**, 33–40]

Introduction

Nucleopolyhedrosis viruses in the very late phase of infection produce two polypeptides, polyhedrin (Polh) and P10, at high levels. These polypeptides are associated with the formation of the viral occlusion bodies and are synthesized maximally after the release of extracellular virus. At 48 h post-infection (hpi) these two proteins constitute nearly 50% of the total protein in virus-infected cell lines (Smith *et al*. 1982). Polyhedrin is the major constituent of the occlusion bodies that are formed exclusively in the nucleus of the baculovirus-infected cells. As occlusion proceeds in the nucleus, large arrays of fibrillar structures begin to accumulate primarily in the nucleus and sometimes in the cytoplasm as well. The P10 protein is associated with the large fibrillar structures in both nuclei and cytoplasm (Harrap 1970; Van der Wilk *et al*. 1987; Williams *et al*. 1989). Disruption of the *p10* gene results in disappearance of the fibrous material but formation of the polyhedra containing normal enveloped virion is not affected (Vlak *et al*. 1988).

The *polh* and *p10* genes have been well characterized from the prototype baculovirus, *Autographa californica* multinucleocapsid nucleopolyhedrosis virus (AcMNPV), and transfer vectors based on their promoters and flanking sequences are most extensively exploited in baculovirus-based expression systems. The *p10* of AcMNPV codes for a 10-kDa protein of 94 amino acids (Kuzio *et al*. 1984). Insertion of the bacterial *lacZ* in frame with the coding sequence of *p10* in AcMNPV leads to the synthesis of copious amounts of the P10–β-galactosidase fusion protein (Vlak *et al*. 1988). Evidently, the *p10* gene is nonessential

* For correspondence. E-mail: kpg@mcbl.iisc.ernet.in.

Keywords. baculovirus; *Bombyx mori*; BmNPV; late gene expression; *p10* gene; transcriptional regulation.

for viral replication and its promoter, therefore, could be effectively utilized for producing large quantities of heterologous proteins.

Both *p10* and *polh* belong to the class of late genes being hypertranscribed during the very late phase of infection but their temporal expression kinetics are slightly different. *p10* is activated a few hours earlier than *polh* although its maximum expression levels are lower. At earlier times post-infection, therefore, *p10* promoter is more active than *polh* (Roelvink *et al*. 1992). By the time the expression from *polh* promoter reaches maximum, the cellular post-translational processing machinery is highly compromised due to the load of viral infection. Hence the *p10*-promoter-based expression vectors offer a 12–24-h advantage over the *polh*-promoter-based vectors for efficient processing of the recombinant proteins.

Although the *p10* promoter from AcMNPV has been exploited for construction of expression vectors, such studies using the *Bombyx mori* nucleopolyhedrosis virus (BmNPV) *p10* promoter have been lacking. Our group (Sriram *et al.* 1997; Sehgal and Gopinathan 1998) has exploited the AcMNPV *p10* promoter for generating recombinant BmNPV through homeologous recombination. Such recombinant BmNPV expressed large quantities of reporter genes under the control of AcMNPV *p10* promoter in *Bombyx*-derived cell lines and *B. mori* larvae. Search for the *p10* gene from a local isolate of BmNPV (Palhan and Gopinathan 1996) was therefore undertaken with the aim of characterizing the gene and its promoter and its transcription profile. The *p10* genes from BmNPVs, the Chinese (Yaozhou 1992) and the Japanese isolates (Hu *et al*. 1994) differed from each other in the size of the encoded protein. We report here the detailed characterization of *p10* from the local isolate of the virus, BmNPV-BGL.

Methods

Virus and cell culture: The *B. mori*-derived cell line BmN was grown in TC100 medium supplemented with 10% foetal calf serum and 50 μg/ml gentamycin (Maeda 1989). The local isolate of BmNPV, BmNPV-BGL (Palhan *et al*. 1995; Palhan and Gopinathan 1996), was used in this study.

Isolation of BmNPV p10: Genomic DNA was isolated from BmNPV derived from the viral polyhedral bodies as described by Maeda (1989). A 2-kb *Kpn*I 'D' fragment from the BmNPV genomic DNA was cloned at the *Kpn*I site in plasmid pTZ18R to generate the clone pNZK9. For sequencing purposes three subclones of pNZK9, designated as K9BI harbouring the 900-bp *Bam*HI fragment, K9δB harbouring the 1.1-kb *Kpn*I–*Bam*HI fragment, and δBX harbouring a 450-bp *Bam*HI to *Xho*I deletion (figure 1), were constructed. These clones were made in pBSKS$^+$ at the appropriate restriction sites.

DNA sequencing: Double-stranded or single-stranded DNA were sequenced using either the universal or reverse sequencing primers and the T7 sequencing kit (Pharmacia). Single-stranded DNA templates were prepared using the helper phage M13K07 (Sambrook *et al.* 1989). Sequences were aligned making use of DNASIS (Hitachi) and Clone Manager software.

RNA isolation, primer extension and RNAase protection: Total RNA was isolated from BmN cells infected with BmNPV (multiplicity of infection, moi, = 10) at 12, 18, 24, 36, 48 and 72 hours post-infection (hpi) using the single-step guanidinium isothiocyanate method (Chomczynski and Sacchi 1987). Primer extension reactions (Sambrook *et al.* 1989) were carried out with 10 μg total RNA and a synthetic 31-mer oligonucletoide primer (5′-CGCGTCAAAACGTT-GGGCTTTGACATGATAG-3′) at 50°C using SuperScript II$^+$ reverse transcriptase (GIBCO-BRL). RNAase protection assay (Sambrook *et al*. 1989) was done with 10 μg total RNA and *p10* antisense riboprobe, which was generated by the *in vitro* transcription reaction using T7 RNA polymerase from the linearized plasmid δBX DNA template. After hybridization (80% formamide overnight at 50°C), the samples were treated with a mixture of RNAase A (1 unit) and RNAase T1 (100 units) for 1 h at 37°C and extracted with phenol–chloroform. The protected RNA was precipitated with 3 volumes of ethanol, dissolved in the gel-loading buffer, and subjected to electrophoresis on an 8% acrylamide – 6 M urea sequencing gel.

For slot blot hybridization, 10 μg total RNA isolated at different times following viral infection was blotted onto nylon membranes, UV-crosslinked and hybridized with the *p10* antisense probe at 45°C for 12 h. The membrane was washed sequentially for 30 min with 6 × SSC, 2 × SSC and 0.1 × SSC at 65°C, in presence of 0.1% sodium dodecyl sulphate, and X-ray film was exposed to it. The autoradiogram was quantified on an LKB Ultra Scan laser densitometric scanner.

Protein structure analysis: Protein secondary structure analysis was done using the UWGCG software (Wisconsin Package, version 8, September 1994, Madison, USA) PEPPLOT, ISOELECTRIC and MULTIALIGN programs. The helical wheel draw analysis was done using the software developed by Staden for MRC. The leucine zipper motif search was done using PROSITE. Phylogenetic analysis was done by a distance method using PHYLIP program (version 1.6.1). All the computer analyses were carried out at the Bioinformatics Centre at Indian Institute of Science.

Results

Identification and cloning the BmNPV p10 gene

The local isolate of BmNPV, BmNPV-BGL, showed nearly identical restriction profiles, except for some polymorphic

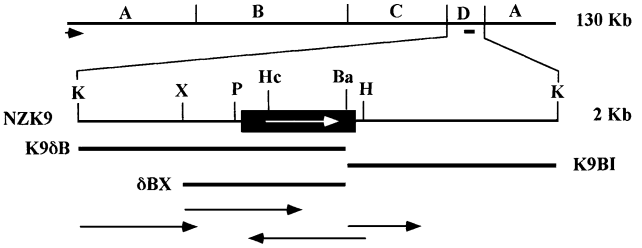

Figure 1. Characterization of the *p10* gene from BmNPV. The location of the *p10* on the linearized physical map of BmNPV genome is shown. The *Kpn*I restriction fragment D (2 kb) was cloned into plasmid pTZ18R at the *Kpn*I site (marked NZK9) and mapped using different restriction enzymes. The location of P10 ORF is indicated by the black box with an arrow showing its orientation. The location of the previously identified *polh* gene is shown by the arrowhead as a reference on BmNPV genomic DNA. The thick lines marked K9*δ*B, K9B1 and *δ*BX designate the subclones (in plasmid pBSKS^{+}) constructed for sequencing as well as antisense RNA isolation purposes (the arrows below indicate the sequencing strategy). The restriction sites marked are as follows: Ba, *Bam*HI; H, *Hin*dIII, Hc, *Hin*cII; K, *Kpn*I; P, *Pac*I; X, *Xho*I.

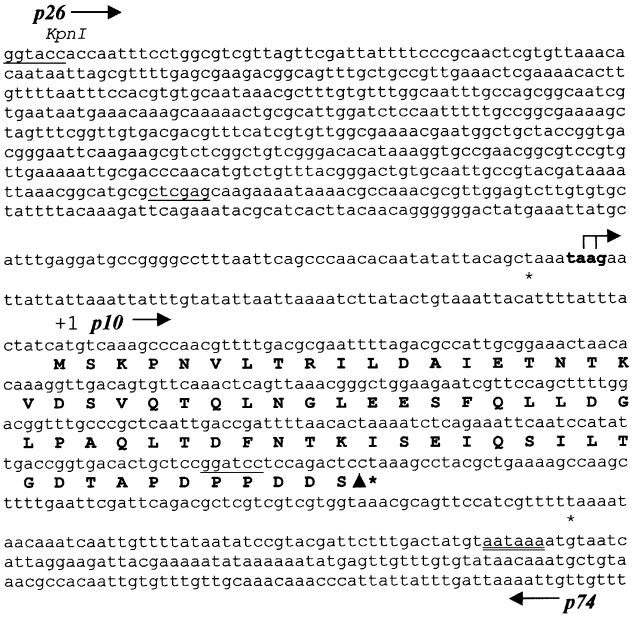

Figure 2. The nucleotide sequence of the BmNPV *p10* region. The P10 coding region is shown in bold and the predicted amino acid sequence is indicated below by the one-letter code. The transcriptional start signal (TAAG) is marked in bold with the arrows indicating the start sites. The polyadenylation signal for *p10* is double-underlined. The *Kpn*I, *Xho*I and *Bam*HI restriction sites are underlined. Partial sequences of the adjacent *p74* (downstream to *p10* and in the opposite orientation) and *p26* (upstream of *p10* and in the same orientation) genes are also given. The stop codon for P26 and P7.5 ORF as well as the one corresponding to AcMNPV-P10 in the absence of 'A' deletion are marked by *. ▲ at the C-terminal end of P10 ORF indicates the site where a single nucleotide (A) deletion was located between the two C residues. The GenBank accession number for this sequence is U46757.

sites for a few enzymes (Palhan and Gopinathan 1996), when compared to the prototype BmNPV-T3 strain (Maeda and Majima 1990). The location of *p10* on the genome of BmNPV-BGL was identified by Southern blot hybridization using the AcMNPV *p10* as probe, and the corresponding genomic fragment (the 2-kb 'D' fragment from *Kpn*I digestion) was cloned into plasmid pTZ18R and sequenced (figure 1).

BmNPV *p10* harbours an open reading frame (ORF) coding for a small polypeptide of 70 amino acids with a predicted mass of 7.5 kDa (figure 2). The sequences are highly homologous to the AcMNPV *p10*, with 93.5% and 90% homology at the DNA and protein levels respectively. However, the C-terminal end of BmNPV-P10 lacks 23 amino acids compared to the AcMNPV-P10. The deletion of a single nucleotide (A) at position +210 (with respect to the ATG) generates a stop codon in BmNPV *p10*, resulting in a truncated protein of 7.5 kDa. Henceforth, the protein encoded by BmNPV *p10* is referred to as P7.5. The *p10* sequences are flanked by the genes *P26* in the 5′ upstream region (in the same orientation) and *p74* in the 3′ downstream region in the opposite orientation (figure 2).

The secondary structure predictions of the putative P7.5 polypeptide done by the PEPPLOT program of the UWGCG package based on the Chou–Fasman and Kyte–Doolittle analysis revealed that the protein is generally hydrophobic, lacks glycosylation sites, and has an estimated *p*I of 3.84.

Analysis of BmNPV p10 promoter sequences

The sequences upstream to the *p10* coding region, located within the short intergenic region (73 nt between *p26* and *p10*), are AT-rich and highly homologous to the AcMNPV *p10* upstream promoter sequences. The conserved baculovirus late gene and very late gene motif 'TAAG' harbouring the transcription start site was located at −72 nt with respect to the A of ATG. The other late gene conserved motif 'TTTGTA' was located at −52 nt.

A TATA-box-like sequence, TATATTA, was present at −87 nt with an overlapping CAAT box immediately upstream to it in the highly AT-rich upstream region. The BmNPV P7.5 ORF had the sequence A^{-3}UC<u>A^{+1}UG</u>U around the start codon, a perfect Kozak consensus required for efficient translation of late baculovirus genes (Rohrmann 1986). The conserved A at −3 nt would predict very efficient translation of the *p10* mRNA.

Temporal nature of p10 transcription

The transcription profile of the BmNPV *p10* was analysed by RNA slot blot analysis employing an antisense riboprobe (figure 3a). The temporal nature of *p10* transcription was evident with its level increasing as the infection proceeded, peaking at 72 hpi. No signals due to *p10* transcripts were detectable until 18 hpi. The burst in *p10* expression at late time points post-infection was typical of the hypertranscribed very late baculovirus genes. The steady-state *p10*

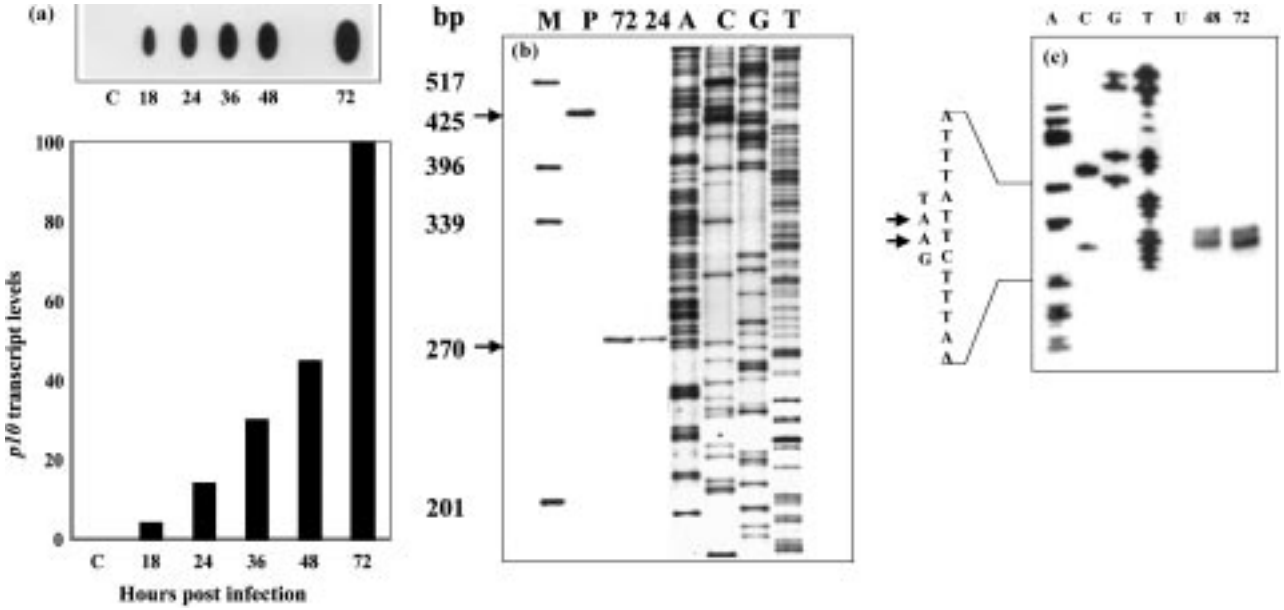

Figure 3. The transcription profile of BmNPV *p10*. (a) Transcription of the *p10* gene was monitored by RNA slot blot using radiolabelled *p10* gene probe or antisense riboprobe. Total RNA isolated from virus-infected cells (moi = 10) at different times (18, 24, 36, 48 and 72 hpi; C, uninfected control) was slot-blotted and probed using the ^{32}P-labelled antisense riboprobe (see text, under Methods). The histogram represents the quantitation of the autoradiogram by laser densitometric scanning, taking the maximal expression at 72 hpi as 100. (b) The *p10* transcription start site. The transcription start site of *p10* was mapped by RNAase protection assay. Using a 425-nt antisense riboprobe (lane P), a 270-nt protected fragment was detected in the RNA isolated from the BmNPV infected cells at 24 and 72 hpi (indicated by lower arrow). Lanes: M, pTZ18 DNA digested with *Hin*fI (size marker); A, C, G, T, the sequencing ladder generated from the *Xho*I end of the double-stranded δBX template DNA using the reverse sequencing primer, to serve as the size marker. (c) Primer extension analysis of *p10* exprssion in BmN cells infected with BmNPV. Total RNA was isolated from uninfected (lane U) BmN cells and cells infected with BmNPV at 48 and 72 hpi. A 31-nt oligodeoxyribonucleotide complementary to the *p10* mRNA (5′-CGCGTCAAAACGTTGGGCTTTGACATGATAG-3′) was end-labelled with ^{32}P, annealed to 10 μg total RNA, and reverse-transcribed to generate the cDNA. The same primer was used in sequencing reaction using δBX dsDNA as template (lanes A, C, G, T). Following electrophoresis on 8% acrylamide – 6 M urea sequencing gel, the bands were visualized by autoradiography. The primer extension products, corresponding to transcription initiation from the two adenines within the consensus baculovirus very-late promoter TAAG motif are marked.

RNA level at 72 hpi compared to 24 hpi was about 5–8 times higher, confirming the very late temporal nature of *p10* expression. By 96 hpi, the lysis of the infected cells was almost complete at the moi used.

Mapping the BmNPV p10 transcription start site

The transcription start site of the *p10* mRNA was mapped by an RNAase protection assay and confirmed by primer extension analysis. The exact size of the RNAase-protected fragment (270 nt) was estimated from the known sequencing ladder run alongside (figure 3b). Thus, the *p10* transcription start site was mapped to the A residue located at −71 nt (with respect to the ATG) within the conserved late gene motif TAAG. Analysis of primer extension products on a high-resolution sequencing gel revealed the presence of two bands in the RNA samples isolated at late time points post-infection (figure 3c, lanes marked 48 and 72). Comparing the size of these bands with a known sequencing ladder generated using the same primer on the δBX DNA template indicated that transcription was initiated from both the A residues (−70 and −71) within the TAAG motif.

Comparison of P10 protein sequences

The protein sequences of P10 from a variety of NPVs are summarized in figure 4, and the extent of their homology to P7.5 from our local virus isolate, BmNPV-BGL, is presented in table 1. While *p10*s from several baculoviruses encode a 10-kDa protein, some of them, including the Japanese strain T3 of BmNPV and the local isolate BmNPV-BGL, encode a protein of only 7.5 kDa. This is in contrast to the Chinese isolate of BmNPV, which encodes a 10.2-kDa protein.

Evolutionary conservation of baculovirus P10 proteins

Comparison of the P10 protein sequences from several baculoviruses revealed that they are highly conserved (summarized in table 1). The BmNPV and AcMNPV proteins showed 89% identity. The phylogenetic analysis

```
Bm-BGL   MSKPNVLTRILDAIAETNTKVDSVQTQLNGLEESFQLLDGLPAQLTDFNTKISEIQSILT 60
Bm-CHN   MSKPNVLTRILDAIAETNTKVDSVQTQLNGLEESFQLLDGLPAQLTDFNTKISEIQSILT 60
Bm-T3    MSKPNVLTRILDAIAETNTKVDSVQTQLNGLEESFQPLDGLPAQLTDFNTKISEIGSILT 60
Bm-TA    MSKPNVLTRILDAIAETNTKVDSVQTQLNGLEESFQPLDGLPAQLTDFNTKISEIGSILT 60
Bm-D1    MSKPNVLTRILDAIAETNTKVDSVQTQLNGLEESFQPLDGLPAQLTDFNTKISEIGSILT 60
AcMNPV   MSKPNVLTQILDAVTETNTKVDSVQTQLNGLEESFQLLDGLPAQLTDLNTKISEIQSILT 60
OpMNPV   MSKPSILTQILDAVRAVDSKVTALQTQVDQLVEDSKTLEALTDQLGELDNKVSDIQSMLS 60
PnNPV    MSKPSILTQILDAVRAVDSKVTALQTQVDQLGEDSKTLEALTDQLGDVDNKVSDIQSMLS 60
CfNPV    MSKPSILQQILTAVQDVDTKVDALQAQLTELDGKVQPLDGLSEQLTALDTKVTTIQDILG 60
LdMNPV   MSQ-NILLVIRADIKALSDKVDAVQQEVQDLAANAPDVSALTAKIDAQTAALAAVQTALD 59
SeNPV    MSQ-NILLLIRADIKAVDEKVDALQQAVNDVSANLPDTSELSAKLDAQATTLDTIVTQVN 59
         **: .:*  *   :   . ** ::*  :  :  .    . *. ::      :  :   :

Bm-BGL   -GDTAPDPPDS------------------------ 70
Bm-CHN   -GDTAPDPPDS------------------------ 70
Bm-T3    -GDTAPDPPDS------------------------ 70
Bm-TA    -GDTAPDPPDS------------------------ 70
Bm-D1    -GDTAPDPPSP------------------------ 70
AcMNPV   -GDIVPDLPDSLKPKLKSQAFELDSDARRGKRSSK 94
OpMNPV   VEEELPEPPAPAPEPELPEIPDVPGLRRSRKQ--- 92
PnNPV    IEEELPEPPAPAPEPELPEIPDVPGLRRSRKQ--- 92
CfNPV    -GAEVPDVPLPDNPLNKTRSRK------------- 81
LdMNPV   KIEAVLNPEIP----ALRKKAK------------- 77
SeNPV    NINDVLNPDLPDVPGNLQKQQQQKKSNKK------ 88
               :   .
```

Figure 4. Multiple alignment of baculovirus P10 sequences. The multiple alignment of P10 sequences from a variety of baculoviruses was done using the CLUSTAL W (1.7) program. The single-letter code for amino acids is used. Note the conserved leucine zipper (repeat of 4 Leu residues and one Leu/Ile/Val with a space of seven residues starting from residue 31), marked in bold, and the conserved Pro residues around 65–70 towards the C-terminal end of all the P10 ORFs.

indicated that the BmNPV strains, as anticipated, were closely related to each other. Their nearest neighbour on the evolutionary tree was AcMNPV (figure 5). However, about seven ORFs such as *pcna* in the *lef8* region and the ORF 603 in the *polh* upstream region were missing in BmNPV-BGL (unpublished observations from the authors' laboratory) and BmNPV-D1 and T3 strains (Gomi *et al*. 1999). In fact, the total genomic sequence comparison also revealed that BmNPV genome was over 90% identical to about three-fourths of the AcMNPV genome (Gomi *et al*. 1999).

Discussion

Functional domains of BmNPV P7.5

Although the complete nucleotide sequences of AcMNPV (Ayres *et al*. 1994), OpMNPV (Ahrens *et al*. 1997), BmNPV T3 strain (Gomi *et al*. 1999) and LdMNPV (Kuzio *et al*. 1999) have been reported and the P10 sequences have been identified, the gene has been characterized in detail mainly from AcMNPV. The P10 protein of AcMNPV has been implicated in the formation of fibrillar structures within the virus-infected cells at late times but disruption of this nonessential gene by itself does not affect the formation of polyhedral bodies (Vlak *et al*. 1988). The aggregation and polymerization functions have been attributed to a hydrophobic region of P10 in AcMNPV (Van Oers *et al*. 1993). The P10 of BmNPV-BGL encodes a 7.5-kDa protein identical to the proteins of the T3, TA and D1 strains of the virus (Hu *et al*. 1994) but different from that of the Chinese strain (Yaozhou 1992). A characteristic difference between the BmNPV P7.5 and AcMNPV P10 is the absence of 23 amino acid residues from the carboxy-terminal region in the former. The positively charged, hydrophobic carboxy terminus was considered to be a common feature of baculovirus P10 proteins (Zuidema *et al*. 1993). This hydrophobic region of AcMNPV P10 is the one implicated in the formation of the fibrillar structures (Van Oers *et al*. 1993) possibly through the interaction of P10 aggregates with

Table 1. Comparison of baculovirus P10 proteins.

Baculovirus	Protein size		Identity	Reference
	AA	kDa	%	
AcMNPV	94	10	88.5	Kuzio *et al*. 1984
BmNPV-BGL*	70	7.5	100	This study
BmNPV-CHN	93	10	100**	Yaozhou 1992
BmNPV-D1	70	7.5	95.7	Hu *et al*. 1994
BmNPV-TA and -T3	70	7.5	98.5	Hu *et al*. 1994 Gomi *et al*. 1999
CfNPV	81	8.8	50	Wilson *et al*. 1995
LdMNPV	77	8.1	24	Kuzio *et al*. 1999
OpMNPV	92	10	42.9	Leisy *et al*. 1986 Ahrens *et al*. 1997
PnNPV	92	10	44.2	Mingchou *et al*. 1992
SeNPV	88	9.6	28.9	Zuidema *et al*. 1993

The amino acid sequence percentage identity with BmNPV-BGL was calculated using the UWGCG software package with a gap weight of 3.0 and a gap length weight of 0.10. The P10 sequences were retrieved/extracted from the EMBL bank or from the literature cited.

*BmNPV-BGL (Bangalore), CHN (China), DI, T3 (Japan) and TA (Taiwan) refer to different isolates of BmNPV. CfNPV, *Choristoneura fumiferana* nucleopolyhedrosis virus; LdMNPV, *Lymantria dispar* multinucleocapsid nucleopolyhedrosis virus; OpMNPV, *Orygia pseudotsugata* multinucleocapsid nucleopolyhedrosis virus; PnNPV, *Penna nuda* nucleopolyhedrosis virus; SeNPV, *Spodoptera exigua* nucleopolyhedrosis virus.

**The identity is for the 70 amino acids present in both strains.

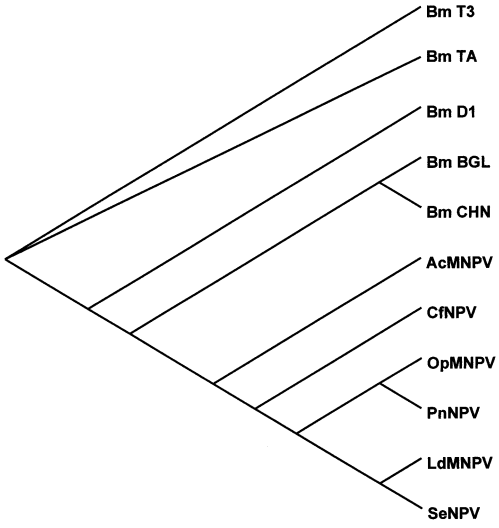

Figure 5. Phylogenetic tree of baculoviruses. A phytogenetic tree based on the P10 sequence similarities of several baculoviruses has been generated. Alignments of the P10 were generated using CLUSTAL protein alignment. Phylogenetic relationships were inferred by a distance method using the PHYLIP program (version 1.6.1). Bm denotes the different BmNPV isolates.

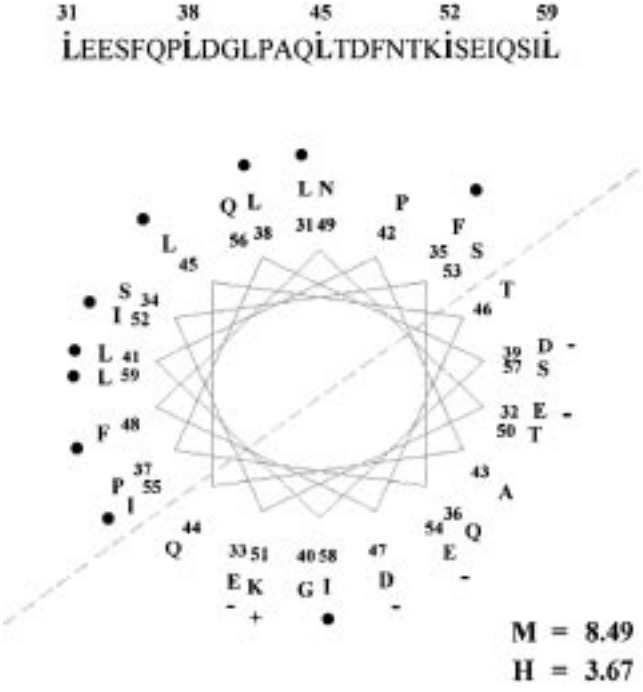

Figure 6. Helical wheel analysis of P7.5. The helical wheel analysis of P7.5 from BmNPV-BGL was done with a window size of 29 residues at a 100° axial angle. The hydrophobic residues (marked with a dot) aligned on one phase of the helix while the negatively charged residues assembled on the opposite phase. The amphipathic helix had a hydrophobic moment (*M*) of 8.49 and total hydrophobicity (*H*) of 3.67.

tubulin (Volkman and Zaal 1990; Cheley *et al*. 1992). It is significant that fibrillar structures are not generally seen in BmN cells infected with BmNPV (Inoue and Mitsuhashi 1984). The absence of the hydrophobic, positively charged

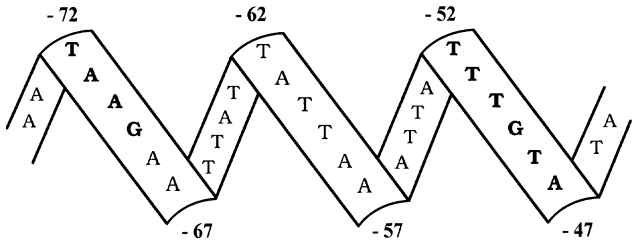

Figure 7. *p10* promoter consensus elements viewed on the DNA helix. The element TAAG harbouring the transcription start site of BmNPV-BGL *p10* is separated from the TTTGTA element by two turns of the helix. Only the partner strand is shown. Nucleotide numbering is in relation to the ATG initiation codon, with A as +1 (not shown).

carboxy-terminal domain in BmNPV P7.5 thus supports the suggestion that this region of P10 is indeed involved in the protein–protein interaction required for its assembly into fibrillar structures.

The amino-terminal region of P7.5 also harbours such a domain. The middle domain of P10, from amino acid residue 52 to residue 78, is involved in nuclear disintegration and polyhedra release. In case of BmNPV P7.5 with a total size of 70 amino acids, the nuclear disintegration domain should thus be limited within residues 52–70 to achieve the same function.

The P7.5 from BmNPV-BGL harbours a leucine zipper dimerization motif (amino acid residues 31–59), bearing five Leu repeats (except one which is Leu/Ile/Val) with a periodicity of seven residues. Helical wheel analysis of this region generated with an axial angle of 100° and a window of 29 residues (figure 6) showed the distribution of hydrophobic amino acids on one phase of the helix (marked with a dot in figure) and negatively charged residues on the other phase of the helix. This stretch of 29 residues had a very high hydrophobic moment of 8.49, with total hydrophobicity of 3.67. The amphipathic nature of the helix would facilitate the dimerization of P7.5 as well as interaction with other proteins necessary for its aggregation and nuclear membrane disintegration. The biological relevance of this domain is obvious considering the fact that it is conserved across 10 baculovirus P10 protein sequences reported so far in the literature (figure 4).

Transcription of baculovirus late genes

Analysis of the BmNPV DNA in the conserved *p10* promoter region, from positions −72 to −42 nt with respect to ATG, when aligned with conserved untranslated leader sequence elements from other baculoviruses (Zanotto *et al*. 1992), indicated that both the consensus elements (TAAG and TTTGTA) were located on the same face of the double helix (figure 7). The core TAAG element was separated by two complete turns of the DNA helix from the conserved motif TTTGTA. The conserved elements in the *polh* promoter were also similar in sequence and orientation (Zanotto

et al. 1992). They resembled the universal consensus of the internal control regions, comprising boxes A and C of the 5S rRNA, or boxes A and B of the tRNA genes (Lassar *et al*. 1983), which foster transcription by RNA polymerase III. The hypertranscribed baculovirus late genes (*polh* and *p10*) are in fact transcribed by a unique, α-amanitin-resistant RNA polymerase (reviewed by Blissard and Rohrmann 1990; Huh and Weaver 1990). There is no consensus among baculoviruses in the region upstream to the conserved 12-mer associated with the mRNA start site in the polyhedrins and granulins. Their promoters were predicted to be internal, as is the case of class III genes, and transcribed by a modified or virus-encoded polymerase, or a modified transcription factor – polymerase-III-like complex (Yang *et al*. 1991). More recently, however, a purified preparation of RNA polymerase has been isolated from AcMNPV-infected Sf9 cells (Guarino *et al*. 1998). This polymerase preparation initiated transcription from the viral very late (*polh*) and late (39K) promoters but not from the viral early promoters. The polymerase was essentially made up of four virally encoded proteins (LEF8, LEF9, LEF4 and P47), but these results await confirmation from other systems.

Compiling the sequence and transcription analyses, it is evident that the BmNPV-BGL *p10* upstream sequences did harbour a strong promoter that could be exploited for heterologous gene expression in the BmNPV-based system because the *p10* promoter was functional in *B. mori*-derived cell lines as well as in the larval caterpillar (Palhan *et al*. 1995; Sriram *et al*. 1997; Sehgal and Gopinathan 1998). Reporter gene fusion constructs under the BmNPV *p10* promoter also showed high levels of expression in BmN cells in transient transfection assays (unpublished observations). Thus, in combination with the *polh* promoter, the *p10* promoter can also be utilized for simultaneous expression of multiple genes, in *B. mori*-derived cell lines or in silkworm larvae for economic large-scale production of biomolecules of commercial importance.

Acknowledgements

We thank Ravinarayan Raghupathy and Satinder Rawat for their help in sequencing; S. Sriram, Asha Acharya and Sangeeta Dhawan in the preparation of the manuscript; and Y.N. Shyamala of Bioinformatics Centre for computer analysis. V.B.P. was a research fellow of the Council for Scientific and Industrial Research (India). This work was supported by grants from Department of Biotechnology, Government of India.

References

Ahrens C. H., Russel R., Funk C. J., Evans J. T., Harwood S. H. and Rohrmann G. F. 1997 The sequence of *Orygia pseudotsugata* multinucleocapsid nucleopolyhedrosis virus genome. *Virology* **229**, 381–399.

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

Blissard G. S. and Rohrmann G. E. 1990 Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* **53**, 127–155.

Cheley S., Kosik K. S., Paskevich P., Bakalis S. and Baylev H. 1992 Phosphorylated baculovirus p10 is a heat-stable microtubule-associated protein associated with process formation in SF9 cells. *J. Cell Sci.* **102**, 739–752.

Chomczynski P. and Sacchi N. 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–159.

Gomi S., Majima K. and Maeda S. 1999 Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrosis virus. *J. Gen. Virol*. **80**, 1323–1337.

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

Harrap K. A. 1970 Cell infection by a nuclear polyhedrosis virus. *Virology* **42**, 311–318.

Hu N. T., Lu Y. F., Hashimoto Y., Maeda S. and Hou R. F. 1994 The p10 gene of natural isolates of *Bombyx mori* polyhedrosis virus encodes a truncated protein with an M_r of 7700. *J. Gen. Virol.* **75**, 2085–2088.

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. Virol.* **71**, 195–202.

Inoue H. and Mitsuhashi J. 1984 A *Bombyx mori* cell line susceptible to a nuclear polyhedrosis virus. *J. Sericulture Sci. Jpn.* **53**, 108–113.

Kuzio J., Rohel D. Z., Curry C. J., Krebs A, Carstens E. B. and Faulkner P. 1984 Nucleotide sequence of the p10 polypeptide gene of *Autographa californica* nuclear polyhedrosis virus. *Virology* **139**, 414–418.

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.

Lassar A. B., Martin P. L. and Roeder R. G. 1983 Transcription of class III genes: formation of preinitiation complexes. *Science* **222**, 740–748.

Leisy D. J., Rohrmann G. F., Nesson M. H. and Beaudreau G. 1986 Nucleotide sequencing and transcriptional mapping of the *Orygia pseudotsugata* multiple nuclear polyhedrosis virus p10 gene. *Virology* **153**, 157–167.

Maeda S. 1989 Gene transfer vectors of a baculovirus, *Bombyx mori* nuclear polyhedrosis virus, and their use for expression of foreign genes in insect cells. In *Invertebrate cell systems and applications* (ed. J. Mitsuhashi), vol. I, pp. 167–182. CRC Press, Boca Raton.

Maeda S. and Majima K. 1990 Molecular cloning and physical mapping of the genome of *Bombyx mori* nuclear polyhedrosis virus. *J. Gen. Virol.* **71**, 1851–1855.

Mingchou J., Lo C. F., Huang C. J. and Wang C. H. 1992 Isolation and nucleotide sequence of *Penna nuda* multiple nucleocapsid nuclear polyhedrosis virus (PnMNPV) two late genes: polyhedrin and p10. In Proceedings of the XIX International Congress of Entomology, Beijing, China, Abstract No. 21, p. 280.

Palhan V. B. and Gopinathan K. P. 1996 Characterization of the local isolate of *Bombyx mori* nuclear polyhedrosis virus. *Curr. Sci.* **70**, 147–153.

Palhan V. B., Sumathy S. and Gopinathan K. P. 1995 Baculovirus mediated high level expression of luciferase in silkworm cells and larvae. *BioTechniques* **199**, 97–104.

Roelvink P. W., van Meer M. M. M., de Kort C. A. D., Possee R. D., Hammock B. D. and Vlak J. M. 1992 Dissimilar expression of *Autographa californica* multicapsid nuclear polyhedrosis virus polyhedrin and p10 genes. *J. Gen. Virol.* **73**, 1481–1489.

Rohrmann G. F. 1986 Polyhedrin structure. *J. Gen. Virol.* **67**, 1499–1513.

Sambrook J., Fritsch E. F. and Maniatis T. 1989 *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sehgal D. and Gopinathan K. P. 1998 Recombinant *Bombyx mori* nucleopolyhedrosis virus harboring green fluorescent protein. *BioTechniques* **25**, 997–1006.

Smith G. E., Vlak J. M. and Summers M. D. 1982 *In vitro* translation of *Autographa californica* nuclear polyhedrosis virus early and late mRNAs. *J. Virol.* **44**, 199–208.

Sriram S., Palhan V. B. and Gopinathan K. P. 1997 Heterologous promoter recognition leading to high-level expression of cloned foreign genes in *Bombyx mori* cell lines and larvae. *Gene* **190**, 181–189.

Van Der Wilk F., Van Lent J. W. M. and Vlak J. M. 1987 Immunogold detection of polyhedrin, p10 and virion antigens in *Autographa californica* nuclear polyhedrosis virus-infected *Spodoptera frugiperda* cells. *J. Gen. Virol*. **68**, 2615–2624.

Van Oers M. M., Elipsen J. T. M., Reusken C. B. E. M., Sliwinsky E. L., Goldbach R. W. and Vlak J. M. 1993 Functional domains of the p10 protein of *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* **74**, 563–574.

Vlak J. M., Klinkenberg F. A., Zaal K. J. M., Usmany M., Klinge-Roode E. C., Geervliet J. B. F., Roosien J. and Van Lent J. W. M. 1988 Functional studies on the p10 gene of *Autographa californica* nuclear polyhedrosis virus using a recombinant expressing a p10-β-galactosidase fusion gene. *J. Gen. Virol.* **69**, 765–777.

Volkman L. E. and Zaal K. J. M. 1990 *Autographa californica* M nuclear polyhedrosis virus: Microtubules and replication. *Virology* **175**, 292–302.

Williams G. V., Rohel D. Z., Kuzio J. and Faulkner P. 1989 A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion-deletion mutants. *J. Gen. Virol.* **70**, 187–202.

Wilson J. A., Hill J. E., Kuzio J. and Faulkner P. 1995 Characterization of the baculovirus *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus p10 gene indicates that the polypeptide contains a coiled-coil domain. *J. Gen. Virol.* **76**, 2923–2932.

Yang C. L., Stetler D. A. and Weaver R. F. 1991 Structural comparisons of the *Autographa californica* nuclear polyhedrosis virus-induced RNA polymerase and the 3 nuclear RNA polymerases from the host, *Spodoptera frugiperda*. *Virus Res.* **20**, 251–264.

Yaozhou Z. 1992 Studies on the p10 gene of *Bombyx mori* nuclear polyhedrosis virus. Ph.D. thesis, Graduate School of Chinese Academy of Agricultural Sciences, Shanghai Institute of Biochemistry, Academia Sinica, Shanghai, China.

Zanotto P. M. de A., Sampaio M. J. A., Johnson D. W., Rocha T. L. and Marunaik J. E. 1992 The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *J. Gen. Virol.* **73**, 1049–1056.

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. Virol.* **74**, 1017–1024.

Received 18 August 2000