

Open reading frame 122 of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus encodes a novel structural protein of occlusion-derived virions

Gang Long,^{1,2} Xinwen Chen,¹ Dick Peters,² Just M. Vlak² and Zhihong Hu¹

Correspondence

Just Vlak

just.vlak@viro.dpw.wau.nl

¹Joint Laboratory of Invertebrate Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China

²Laboratory of Virology, Wageningen University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

Helicoverpa armigera single nucleocapsid nucleopolyhedrovirus (HaSNPV) and its closely related variant *H. zea* SNPV (HzSNPV) contain 20 open reading frames (ORFs) unique among baculoviruses. In this report, the function of HaSNPV ORF 122 (Ha122) is investigated. Ha122 was transcribed as a polyadenylated transcript from 8 h post-infection in infected *H. armigera* insect cells. 5'RACE analysis indicated that Ha122 transcription starts predominantly in the consensus major late transcription initiation motif DTAAG, around 47 nt upstream of the putative translation start codon, with a minor start at position –89. Using 3'RACE, the transcription stop site mapped 27 nt downstream of the putative translation stop codon. By Western blot analysis using a chicken-derived polyclonal antibody, the product of Ha122 was found in infected cells to be a 21 kDa protein, close to the theoretical size of 21·6 kDa. The Ha122 protein, when fused to green fluorescent protein, was observed in the nuclei of *H. armigera* cells but only in conjunction with wild-type HaSNPV infection. The 21 kDa protein was located specifically in the nucleocapsid of occlusion-derived virions (ODVs) and not in that of budded virus. The available data suggest that Ha122 is a functional ORF of HaSNPV and HzSNPV and that the 21 kDa protein is a novel specific component of baculovirus ODVs.

Received 26 March 2002

Accepted 11 September 2002

INTRODUCTION

The *Baculoviridae* family is a large group of occluded, rod-shaped viruses with circular, supercoiled, double-stranded DNA genomes of 100–180 kb depending on the virus species. Baculoviruses are infectious to arthropods, primarily insects, and are natural control agents of insect pests in agriculture and forestry. The *Baculoviridae* family is taxonomically subdivided into two genera: *Nucleopolyhedrovirus* (NPV), which have multiple virions present in large polyhedron-shaped occlusion bodies, and *Granulovirus* (GV), which have a single virion with a single nucleocapsid embedded in a granular occlusion body (Blissard *et al.*, 2000). Based on the phylogenetic analysis of up to 63 genes that are conserved among all sequenced lepidopteran NPV baculoviruses, a more recent subdivision that recognizes two distinct groups has been adopted (Zanotto *et al.*, 1993; Herniou *et al.*, 2001). The baculoviruses *Autographa californica* (Ac) multinucleocapsid (M)NPV, *Bombyx mori* (Bm) NPV and *Orgyia pseudotsugata* (Op) MNPV, for example, belong to group I, and *Spodoptera exigua* (Se) MNPV, *Lymantria dispar* (Ld) MNPV and *Helicoverpa armigera* (Ha) single (S)NPV are

members of group II. This grouping is supported further by the observation that viruses of these two groups have distinct budded virus envelope fusion proteins, GP64-like (group I) and F-like (group II) (Pearson *et al.*, 2000; IJkel *et al.*, 2000).

Baculoviruses are used widely as bioinsecticides to control heliothine insects in cotton and vegetable crops around the world (Ignoffo, 1973; Cunningham, 1998). *Helicoverpa zea* SNPV (HzSNPV) was registered as one of the first commercial baculovirus pesticides (Virion-H, Biocontrol-VHZ, Elcar) in the 1970s and has been used extensively to control the cotton bollworm in the USA and elsewhere (Hsieh, 1989; Cunningham, 1998). HaSNPV, isolated in 1975 in the province Hubei, People's Republic of China, has been used successfully in China for over almost 25 years to control *H. armigera* in cotton and vegetable crops in areas of well over 100 000 hectares (Zhang, 1994).

HaSNPV has a DNA genome of 135 kb and its sequence has been determined in its entirety (Chen *et al.*, 2001). Some 135 open reading frames (ORFs) of 50 aa or larger in size have been determined for HaSNPV. Twenty ORFs were present

only in HaSNPV and have elusive functions; some may be responsible for the unique features of this virus, such as host specificity. These 20 unique ORFs were also present in the closely related HzSNPV, a genotypic variant (Chen *et al.*, 2002).

This report is part of an effort to elucidate the function of the unique ORFs of HaSNPV. Eight ORFs were pre-selected for analysis based on their size (> 100 aa) and the presence of known baculovirus consensus transcription initiation and termination signals (Blissard *et al.*, 2000). In this paper, we

analysed the putative ORF 122 (Ha122) (Fig. 1A) and demonstrated by transcriptional analysis and protein identification and localization that Ha122 is a functional gene expressed at late times post-infection (p.i.). Ha122 encodes a 21 kDa protein and is present in nucleocapsids of occlusion-derived virus (ODV) but not in budded virus (BV).

METHODS

Virus, insect and cell lines. HaSNPV strain G4 (Sun *et al.*, 1998) was used for infections and was propagated in *Helicoverpa* cell lines.

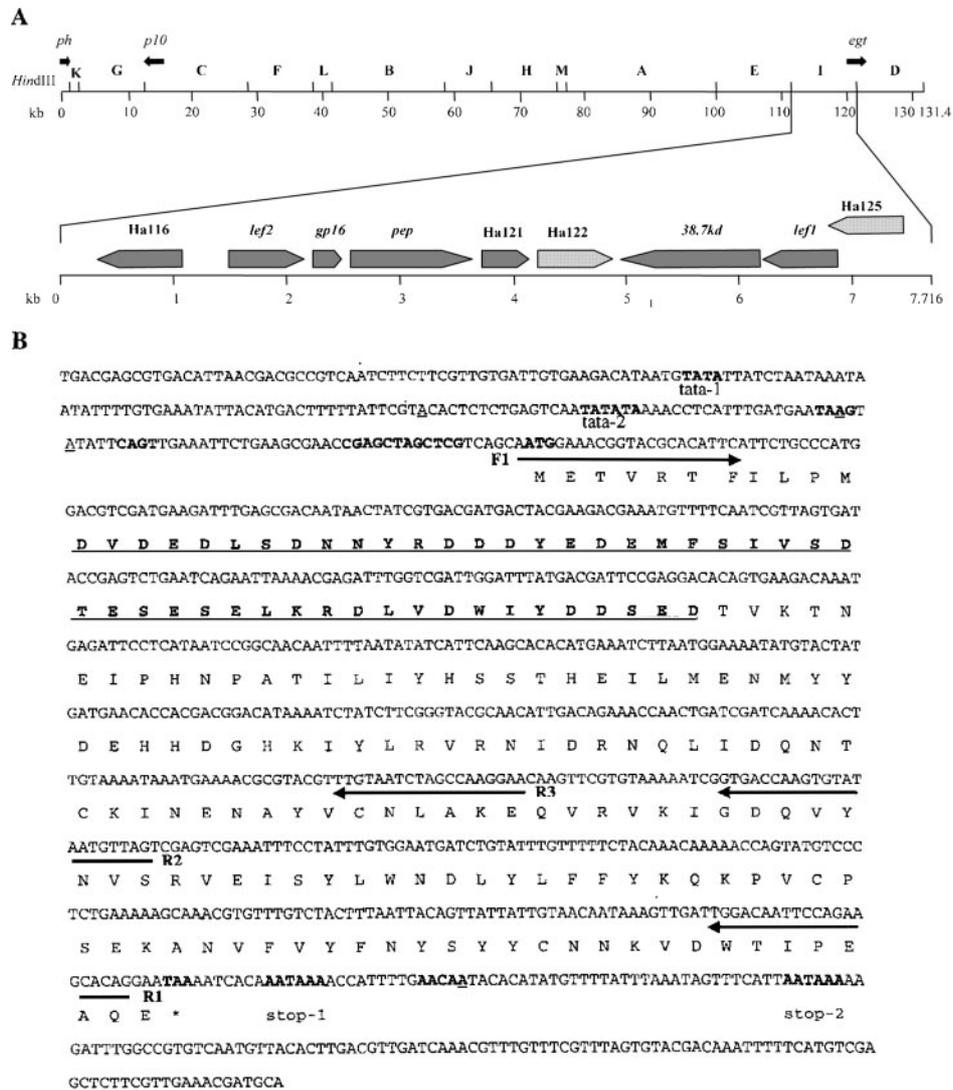


Fig. 1. (A) Physical map (*HindIII*) of the HaSNPV genome and genomic organization of the *HindIII* I fragment (Chen *et al.*, 2001). Arrows represent the individual ORFs and point to their directions of transcription. Dashed ORFs (Ha122 and Ha125) are unique to HaSNPV. *ph*, polyhedrin; *egt*, ecdysteroid UDP-glucosyltransferase; *lef1* and *lef2*, late essential factors 1 and 2; *pep*, polyhedral envelope protein. (B) Sequence and motifs of the HaSNPV ORF122 genomic region from nt 115324 to 116283. Location of TATA boxes (tata-1 to -3), baculovirus consensus early initiation motif CAGT, baculovirus consensus late initiation motif TAAG, polyadenylation signals AATAAA (stop-1 and -2) and translation start and stop codons are denoted in bold. The transcription start and stop sites are underlined. The predicted Asp-rich region in the amino acid sequence is bold and underlined.

H2E5 cells (McIntosh *et al.*, 1999) were grown in CCM3 medium supplemented with 5% foetal bovine serum (FBS) and HzAM1 cells (McIntosh & Ignoffo, 1983) were grown in Grace's medium, pH 5.9–6.1, supplemented with 10% FBS. A culture of *H. armigera* insects was maintained according to Sun *et al.* (1998) for HaSNPV polyhedra production.

Computational analysis. Ha122 was analysed using the software EXPASY (Appel *et al.*, 1994) for the prediction of domains, motifs, transmembrane regions and subcellular localization domains (Reinhardt & Hubbard, 1998). Protein comparisons with entries in the updated GenBank/EMBL, SWISS-PROT and PIR databases were performed with BLASTP and PSI-BLAST programs (Altschul *et al.*, 1997). Sequence alignment was performed with MEGALIGN and alignment editing was performed with the GENEDOC software.

Transcription analysis. Total RNA was isolated from mock-infected and HzAM1 cells infected with BVs of HaSNPV strain G4 (m.o.i. of 5 TCID₅₀ units per cell) at 0, 4, 8, 12, 24, 48 and 72 h p.i. Cells were resuspended in 500 µl Trizol (Gibco-BRL) and 100 µl chloroform, incubated for 8 min and centrifuged at 14000 g for 15 min at 4 °C. The RNA present in the water fraction was precipitated using isopropanol, centrifuged at 14000 g for 10 min at 4 °C, washed with 70% ethanol and resuspended in 50 µl water. The RNA solutions were incubated at 55 °C for 10 min and the RNA quantified by optical density measurements at 260 nm.

RT-PCR was performed using a 3' RACE kit (Roche) employing 1 µg of purified total RNA as template per time-point. First-strand cDNA synthesis was performed using AMV reverse transcriptase and an oligo(dT) anchor primer, according to the manufacturer's instructions. The cDNA mixtures were amplified with PCR using the PCR oligo(dT) anchor primer and the gene-specific forward primer F1 (5'-ATGGAACGGTACGCACATTCATTCTG-3') (Fig. 1B). The PCR products obtained were analysed by agarose gel electrophoresis. The amplified RT-PCR products from 48 h p.i. were gel-purified, cloned into pGEM-T (Gibco-BRL) (Sambrook *et al.*, 1989) and sequenced with T7 primers to determine the 3' end of the Ha122 transcript.

The 5' end of the Ha122 transcript was determined using the 5' RACE kit (Roche) employing 1 µg of purified total RNA as template and isolated at 4 and 48 h p.i. Briefly, first-strand cDNA synthesis was performed with the gene-specific reverse primer R1 (5'-GGG-GAATTCCTGCTTTCTGGAATTGTCCAATCAACTT-3') (Fig. 1B). The cDNA was purified with the High Pure PCR Purification kit (Roche) and a poly(A) tail was added to the 5' end using terminal transferase and dATP. The tailed cDNAs were amplified by PCR using the oligo(dT) anchor primer (forward) and the nested, gene-specific reverse primer R2 (5'-ACTAACATTATACACTTGGTCACCG-3') (Fig. 1B). A second PCR was performed using the PCR 5' oligo(dT) anchor primer and the nested reverse primer R3 (5'-CTTGTTCTTGGCTAGATTACAAACG-3') (Fig. 1B). The PCR products obtained were gel-purified, cloned into pGEM-T and sequenced with T7 primers.

Generation of anti-Ha122 antiserum. The complete Ha122 ORF (588 nt) was amplified by PCR and subcloned into the vector pGEX-KG, in-frame with the C-terminal glutathione S-transferase (GST) tag present in this plasmid (Guan & Dixon, 1991). The plasmid, designated pGST-Ha122, was transfected into *Escherichia coli* BL21 cells and fusion protein expression was induced by incubation in the presence of 1 mM IPTG when the optical density measurement at 600 nm reached ~0.7. After 3 h, the cells were harvested and lysed by sonication in ice-cold PBS. The lysate was centrifuged at 3000 g for 5 min. The pellets (inclusion bodies containing the highly expressed fusion protein) were washed twice with 100 mM Tris/HCl, pH 8.0, and were dissolved in the above buffer containing 5% SDS (sonication was required to dissolve the pellets). This

fusion protein solution was separated on a continuous-elution, preparative 12.5% SDS-PAGE gel (Model 491 PrepCell) (BioRad). The eluted fractions were collected and electrophoresed and the bands were visualized by Coomassie brilliant blue staining. The proteins were concentrated with a Centrprep 10 kDa filter device (Amicon) and the concentrated solution was dialysed overnight with two changes of PBS. This dialysed protein sample was used as an immunogen to raise Ha122-specific antiserum in 8-month-old, egg-laying chickens.

Chickens were injected intramuscularly, each with 100 µg purified GST-Ha122 fusion protein using Freund's incomplete adjuvant. The chickens were boosted after 2 weeks with 100 µg of purified protein. Eggs were collected every day for 4 weeks and serum was collected 12 weeks after the boost injection. Western blot analyses using *E. coli* BL21 cell extracts expressing GST-Ha122 and column-purified GST-Ha122 were used to test the batches of specific antisera.

Western blot analysis. Monolayers of HzAM1 cells were mock- and HaSNPV-infected at an m.o.i. of 5 TCID₅₀ units per cell. Cells were harvested at 0, 24, 48 and 72 h p.i., pelleted at 3000 g, resuspended in PBS and lysed in SDS-PAGE loading buffer by boiling for 10 min. Protein samples were separated by SDS-PAGE and transferred onto an Immobilon-P nitrocellulose membrane (Millipore) by semi-dry electrophoresis transfer (Ausubel *et al.*, 1994). The membranes were incubated overnight in 2% skimmed milk powder in TBS at 4 °C. The membranes were allowed to react with Ha122 antiserum diluted 1:1000 for 1 h at room temperature and treated further as described by IJkel *et al.* (2000).

HzAM1 cells (10⁶ per Petri dish in 2 ml medium) were infected with HaSNPV at an m.o.i. of approximately 5 TCID₅₀ units per cell and incubated at 27 °C. Cells were harvested at 0, 8, 16, 24, 48 and 72 h p.i., washed three times in PBS and lysed in SDS-PAGE loading buffer. Uninfected cells were taken as a control. Proteins were separated by SDS-PAGE and blotted onto nitrocellulose membranes. Duplicate blots were prepared and processed for immunoblotting. One blot was incubated in the presence of pre-immune serum (dilution 1:1000) and the other in the presence of the anti-Ha122 antiserum (dilution 1:1000), as detailed by IJkel *et al.* (2000).

Purification of HaSNPV BV and ODV fractions. HzAM1 cells were infected with BVs of HaSNPV G4 at an m.o.i. of 0.1 TCID₅₀ units per cell. After 3 days, the cell culture supernatant was collected and clarified at 2000 g for 10 min at 4 °C. The supernatant containing the BVs was passed through a 0.45 µm pore-size filter. BVs in the filtrate were pelleted through a 25% (wt/wt) sucrose cushion made up in 0.1 TE (10 mM Tris/HCl, pH 7.5, and 1.0 mM EDTA) by centrifugation at 100000 g for 90 min at 4 °C, and resuspended in 0.1 TE. ODVs were purified from HaSNPV G4 polyhedra derived from infected *H. armigera* fourth-instar larvae as described previously (Chen *et al.*, 2001). Separation of envelope and nucleocapsid fractions after treatment with NP-40 was done as described by IJkel *et al.* (2000). The purity of the nucleocapsids was checked by electron microscopy.

Construction of a transient expression plasmid. To construct plasmids that allow gene expression in *H. zea* cell lines upon transfection, a vector (pHaie-1-GFP) with a HaSNPV *ie-1* promoter-driven green fluorescent protein (GFP) gene was constructed. The HaSNPV *ie-1* promoter (nt 11170–11355) (Chen *et al.*, 2001) was amplified by PCR using two specific primers (primer1, 5'-CGCGAGCTCATAACAGTATTTGGTCCTTCGTTCTG-3', and primer2, 5'-CGAGGATCCCTTGCCGTAAGTTTAAATTTCAATAA-3'). The restriction sites for *Bam*HI and *Sst*I are shown in italics. The PCR product was cloned into the multiple cloning site of pGEM-T EASY (Gibco-BRL) and the HaSNPV *ie-1* promoter insert was sequenced to make sure that there were no mutations in the

promoter region. *SstI/BamHI* digestion was used to isolate the promoter fragment, which was then cloned into a p166AcV5-GFP (IJkel *et al.*, 2000) derivative, where the *Opgp64* promoter and the AcV5 segment were removed by *SstI/BamHI* digestion, to give pHaie-1.

The Ha122-encoding sequence containing the ATG but not the TAA codon was amplified by PCR. Two restriction sites, *BamHI* and *EcoRI*, were introduced at either end by the primers. These two restriction sites were used to insert Ha122 downstream of the *Haie-1* promoter and upstream of GFP to render the fusion vector pHaie-1-ORF122-GFP.

Fluorescence microscopy. Hz2E5 cells (1×10^5) were grown on glass cover slips in Petri dishes and transfected with 5 µg of plasmid DNA using Cellfectin (Gibco-BRL). At 60 h p.i., the cells were examined under a Zeiss LSM510 (confocal) laser scanning microscope for fluorescence using an excitation wavelength of 488 nm and an emission band pass filter of 505–530 nm.

RESULTS

Characterization of Ha122

The Ha122 gene was located in a tail-to-tail configuration with a 38.7 kb gene (Fig. 1A). The latter ORF also occurs in other baculoviruses (IJkel *et al.*, 1999; Herniou *et al.*, 2001). The Ha122 gene theoretically encodes a 21.6 kDa protein. Searches of protein databases, GenBank and SWISS-PROT, showed that the putative Ha122 protein is unique to HaSNPV with no immediate homologues in other biological systems, including baculoviruses (Chen *et al.*, 2001). Motif scan analysis showed that there is an Asp-rich region located in the N-terminal part of the protein sequence (Fig. 1B, underlined sequence). Some short amino acid sequence homologues were found by protein BLAST searches based only on this region. Gap alignment revealed that the predicted protein sequence of Ha122 shares low similarity with a putative DNA-derived RNA polymerase subunit from members of the genus *Leishmania*. Two putative N-glycosylation sites (aa 142–145 and 178–181) were found through the prediction of SCANPROSITE. No signal peptide, transmembrane region, mitochondrial targeting sequence, nuclear localization signal or membrane retention signal was shown by other motif search engines. According to the result of PSORT II analysis, the putative Ha122 protein is probably localized in the cytoplasm or associated with the cytoskeleton.

The regions upstream of Ha122 were analysed for the presence of putative transcription start sites, such as baculovirus consensus early CA(G/T)T and late DTAAG promoter motifs, downstream activating elements and host factor binding sites, such as GATA and CACGTG. A late gene transcription motif (TATAA–N–ATAAG, *tata-1*) was found 74 nt upstream of the putative translation start codon ATG of Ha122. Another TATA box (*tata-2*) was located 175 nt upstream of the start codon. An early transcription motif CAGT was observed 39 nt upstream of the ATG codon of Ha122. Additionally, a palindrome of 12 nt was found 18 nt upstream of the putative translation start site but downstream of putative transcription start sites. This palindrome might be involved in the transcriptional or translational

control of Ha122. Two polyadenylation signal sequences (*stop-1* and *stop-2*) were located 8 and 59 nt downstream, respectively, of the putative translation stop codon TAA.

Transcriptional analysis

To see whether Ha122 was transcribed, 3'RACE analysis was performed with total RNA purified from mock-infected and HaSNPV-infected HzAM1 cells at various times p.i. (Fig. 2). A product of the expected size (600 bp) was detected as early as 8 h p.i., which increased in amount up to 72 h p.i., suggesting that Ha122 is transcriptionally active predominantly late in infection. The 3'RACE products from 48 h p.i. were cloned into pGEM-T and sequenced. The obtained sequences mapped the 3' end of the Ha122 transcript 27 nt downstream of the putative translation stop codon TAA and 14 nt downstream of the last A in the first AATAAA sequence. No evidence was obtained to show that the downstream transcription termination signal *stop-2* (Fig. 1B) was used.

To determine the 5' end of Ha122 transcripts, 5'RACE was performed using total RNA extracted 4 and 48 h p.i. from HaSNPV-infected HzAM1 cells. The lack of a PCR product using the first reverse primer (R1) (Fig. 1B) required the use of a nested primer (R2) to detect a PCR product. cDNA was then detected only at 48 h p.i. but not at 4 h p.i., in concordance with the abundant presence of the 3'RACE product at 48 h p.i. (Fig. 2). Five clones were sequenced and putative start sites were identified in two regions upstream of the translational start site ATG. One start site is 89 nt upstream of the ATG codon, the other two are located at two A residues in or near the consensus baculovirus late RTAAG motif (49 and 45 nt upstream of the ATG codon) (Fig. 1B). These transcription start sites are downstream of a TATA box (*tata-1*) and upstream of a putative early transcriptional

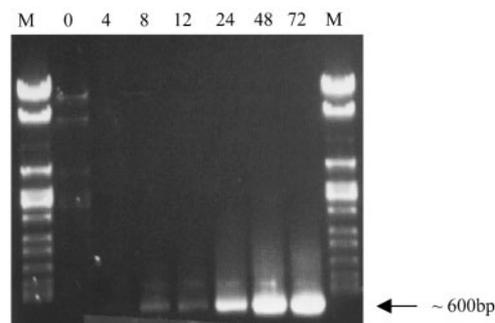


Fig. 2. Temporal expression of Ha122 in HzAM1 cells infected with HaSNPV by 3'RACE. Total RNA was extracted from HaSNPV-infected HzAM1 cells at 0, 4, 8, 12, 24, 48 and 72 h p.i. An oligo(dT) anchor primer was used for reverse transcription. PCR was performed using purified cDNA as template with a Ha122-specific probe (see Methods) and the anchor primer. The size of the expected band is around 600 bp. Size standards (M, marker) are used.

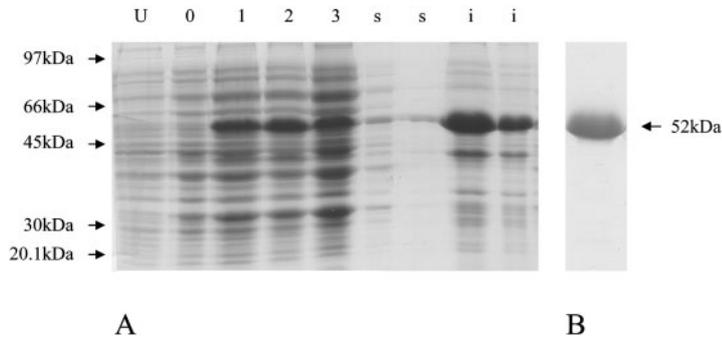


Fig. 3. (A) Expression of the Ha122 protein in BL21 cells. BL21 cells transfected with pGEX-Ha122-GST fusion plasmid were induced for 0, 1, 2 and 3 h and the extracts separated by SDS-PAGE. Cells induced for 3 h were fractionated in a soluble (s) and insoluble (i) fraction prior to SDS-PAGE. (B) Purification of the GST-Ha122 fusion product by column chromatography and analysis by SDS-PAGE.

start site CAGT (Fig. 1B). The collective results suggest that Ha122 is transcriptionally active and most likely a late gene. The 3' and 5'RACE results also suggest that the Ha122 transcript may be around 0.7 kb in size, assuming a poly(A) tail of approximately 100 A residues.

Immunodetection of the Ha122 protein in infected cells

The Ha122 protein has a predicted molecular mass of 21.6 kDa. To see whether a protein encoded by this ORF is made and to study its possible function, a polyclonal antibody was prepared by immunization of a chicken with a purified Ha122 protein. This protein was overexpressed in bacteria as a GST-Ha122 fusion (52 kDa) and purified by column chromatography to almost homogeneity (Fig. 3). The chicken antiserum obtained reacted well against the purified fusion protein and not against bacterial proteins (data not shown).

Protein extracts of HaSNPV-infected HzAM1 cells isolated at 0, 24, 48 and 72 h p.i. were separated by SDS-PAGE and subjected to Western blot analysis using the Ha122-specific

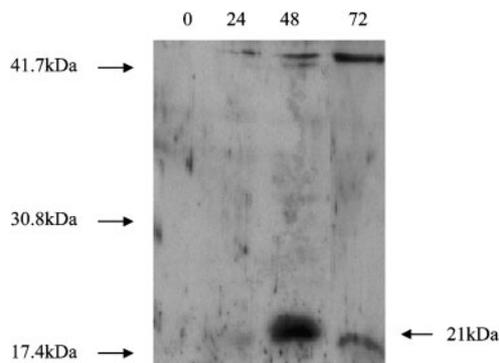


Fig. 4. Western blot analysis of the Ha122 protein in HaSNPV-infected HzAM1 cells (7×10^4 cells per lane). Cells were harvested at 0, 24, 48 and 72 h p.i. Total cell proteins were separated by SDS-PAGE on a 12.5% polyacrylamide gel and blotting onto Immuno-P membrane. Proteins were detected using anti-Ha122 polyclonal chicken antiserum (1:1000 dilution) and detected by chemiluminescence. Molecular mass standards are shown on the left.

chicken antiserum. The antibody did not react to uninfected cells (Fig. 4, lane 1) but reacted strongly with a 21 kDa protein at 48 and 72 h p.i (Fig. 4, lanes 3 and 4). A faint reaction was observed at 24 h p.i. (Fig. 4, lane 2), suggesting that the Ha122 protein was synthesized predominantly late after infection. This is in tune with the analysis of transcription (Figs 1 and 2). A second band reacting with the antiserum was observed at about twice the size of the 21 kDa protein. This may be a dimer and the result of incomplete denaturation.

Immunodetection of the Ha122 protein in BV and ODV

To investigate whether the Ha122 protein is a structural protein, Western blot analysis was carried out on HaSNPV polyhedra, BVs and ODVs (Fig. 5). The Ha122 protein was

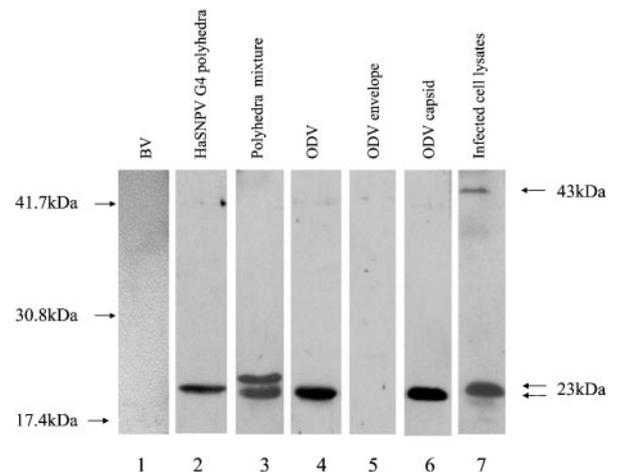


Fig. 5. Immunodetection of the Ha122 protein in HaSNPV. BV (lane 1) and ODV (lane 4) (10 μ g per lane each), HaSNPV polyhedra strain G4 (lane 2) and HaSNPV original isolate (lane 3) (100 μ g per lane each), and HaSNPV-infected HzAM1 cell extract (72 h p.i.) (lane 7), ODV envelopes (lane 5) and nucleocapsids (lane 6) were analysed by SDS-PAGE and Western blot analysis, as described in Fig. 3. Molecular mass standards are shown on the left. Sizes (in kDa) of Ha122 are shown on the right.

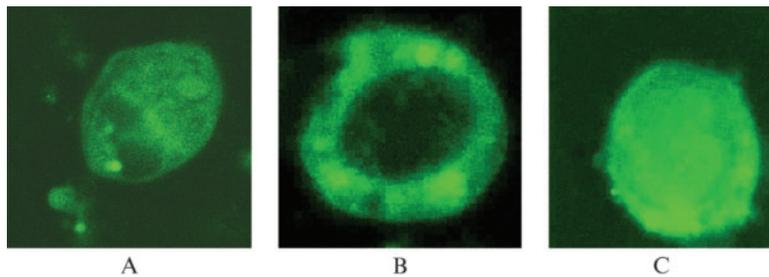


Fig. 6. Localization of the GFP–Ha122 fusion protein in Hz2E5 cells either without (B) or with (C) HaSNPV infection. Hz2E5 cells were transfected with the control plasmid pHaie-1-GFP (A) or with pHaie-1-122GFP (B, C). Cells were then superinfected (4 h later) with HaSNPV BV (C) or not (B). At 48 h post-transfection, the cells were examined for fluorescence by confocal laser scanning microscopy.

not detected in a preparation of BVs (Fig. 5, lane 1). In contrast, the 21 kDa protein could be detected easily in polyhedra from strain G4 (Fig. 5, lane 2) and ODVs (Fig. 5, lane 4), suggesting that the Ha122 protein is a constituent of ODVs. Since equivalent amounts of BV and ODV proteins were loaded onto SDS-polyacrylamide gels (Fig. 5, lanes 1 and 4), it is unlikely that substantial amounts of Ha122 protein are present in BVs. Interestingly, in the original isolate of HaSNPV (genotypic mixture), from which the strain G4 was derived (Sun *et al.*, 1998), a second band reacted with the Ha122 antiserum (Fig. 5, lane 3). This band had a molecular mass of approximately 23 kDa.

To determine a more precise location of the Ha122 protein in ODVs, the virion envelopes were separated from the nucleocapsids by treatment with detergent followed by sucrose gradient purification of nucleocapsids in the presence of protease inhibitors (Ijkel *et al.*, 2000). The nucleocapsid and envelope fractions were analysed by SDS-PAGE (Fig. 5). The purity of nucleocapsids was checked by electron microscopy. The 21 kDa protein was present in the nucleocapsid fraction (Fig. 5, lane 6) but not in the ODV envelope fraction (Fig. 5, lane 5).

Localization of the Ha122 proteins in insect cells

The subcellular localization of the Ha122 protein was investigated using an N-terminal GFP–Ha122 fusion construct. The fusion construct was placed under the control of the *Haie-1* promoter. Hz2E5 cells were transfected with plasmid constructs with GFP alone and with the GFP–Ha122 fusion insert, incubated for 48 h and examined for fluorescence by confocal laser scanning microscopy (Fig. 6). GFP alone showed homogeneous fluorescence in the cytoplasm and nucleus (Fig. 6A). The GFP–Ha122 fusion protein was localized primarily in the cytoplasm (Fig. 6B) and hardly present in the nucleus. However, when cells transfected with the GFP–Ha122 fusion construct were superinfected with HaSNPV, the fusion protein was localized both in the cytoplasm and in the nucleus (Fig. 6C). A similar result was obtained with HzAM1 cells but the fluorescence was not as abundant as that seen in Hz2E5 cells.

DISCUSSION

In this paper, we have described the identification of a novel structural protein of HaSNPV ODVs. This protein, encoded

by Ha122, is hitherto unique to this baculovirus and its close relative HzSNPV (Chen *et al.*, 2001, 2002) and has not been found in any other baculovirus to date. The coding region of Ha122 potentially encodes a 21.6 kDa protein, which is close to the actual size of the protein determined by SDS-PAGE (21 kDa). This suggests that the Ha122 protein is not extensively modified post-translationally, for example, by glycosylation, despite the presence of two potential glycosylation sites. This is in line with other characteristics of Ha122, as it does not have a signal sequence or a Golgi retention signal. Treatment of HaSNPV-infected HzAM1 cells with tunicamycin also showed no alteration in the mobility of the 21 kDa protein in SDS-PAGE (G. Long, personal communication). An interesting observation is the presence of a 43 kDa band in the HaSNPV-infected cell lysates (Fig. 4). This band is also recognized by the antiserum and may be a dimer of the 21 kDa Ha122 protein or a Ha122 protein associated with another viral or cellular protein. Since this band is not detected in polyhedra or ODVs, it may be a cell extract artefact.

Both Ha122 transcripts and the Ha122 product (21 kDa) are detected predominantly relatively late in the infection cycle and this is compatible with the observation that the 21 kDa protein is present in ODVs and thus required when ODVs are assembled. Analysis of the 5' ends of the Ha122 transcript support this conclusion further, as the major transcript starts close to the baculovirus consensus late motif RTAAG. This is a common feature of structural virion proteins (Funk *et al.*, 1997). However, it cannot be excluded that Ha122 is expressed at earlier times after infection, as there is a putative early transcription motif (CAGT) preceded by two TATA boxes (Fig. 1B) upstream of the baculovirus consensus late motif RTAAG. The necessity to use a nested PCR at late times after infection to detect the 5' end of the transcript suggests that the Ha122 transcript is not present in abundance in HaSNPV-infected cells. The Ha122 gene contains two downstream polyadenylation signals (AATAAAA) at 8 and 59 nt relative to the TAA translation stop site. It is not clear why the second stop is preferred.

The functions and precise location of Ha122 in the virion have not been elucidated fully. The protein is associated with the ODV capsid but not with the BV capsid. This may suggest that the Ha122 protein either plays a role in the assembly of ODVs or contributes to the infection of midgut

epithelial cells by ODVs. A similar situation exists for SeMNPV, where one of the unique ORFs, Se117, also encodes an ODV-specific protein (IJkel *et al.*, 2001), whereas the Se117 product is present both in the nucleocapsid and in the envelope fractions. As for Ha122, this protein is associated predominantly with the ODV nucleocapsid. Further localization of the Ha122 protein could be obtained by immunogold electron microscopy and protein-protein interaction studies. Ha122 lacks a nuclear import signal and this explains the inability of this protein to enter the nucleus on its own (Fig. 6). Only in conjunction with a HaSNPV infection does Ha122 end up in the cell nucleus, suggesting that Ha122 needs most likely a viral chaperone protein or a host-induced/modified protein to get there. It would be of interest to learn which HaSNPV protein functions as a chaperone for Ha122, as this will enhance our understanding of the scaffolding of baculovirus ODVs. Further information of the function of Ha122 will come from deletion and mutagenesis studies using a HaSNPV bacmid constructed recently (H. Wang & J. M. Vlak, personal communication).

In the original HaSNPV isolate (Fig. 5, lane 3), another strong band was detected besides the 21 kDa band. This 23 kDa band reacts equally strongly with the 21 kDa antiserum. This may be a reflection of the polymorphism of the Ha122 gene, since the original isolate contains at least eight different genotypes (Sun *et al.*, 1998). Sequence analysis of the Ha122 regions of the various HaSNPV genotypes should give support to this supposition. However, the biological meaning of the two bands remains unclear. The antiserum against ORF122 might be used to identify or classify different strains or subtypes of HaSNPV.

ACKNOWLEDGEMENTS

This research was supported in part by grants from the National Natural Science Foundation of China and the Royal Academy of Sciences of The Netherlands (KNAW).

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Appel, R. D., Bairoch, A. & Hochstrasser, D. F. (1994). A new generation of information retrieval tools for biologists: the example of EXPASY WWW server. *Trends Biochem Sci* **19**, 258–260.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1994). *Current Protocols in Molecular Biology*. New York: Wiley Interscience.
- Blissard, G., Black, B., Crook, N., Keddie, B. A., Possee, R., Rohrmann, G., Theilmann, D. & Volkman, L. (2000). *Baculoviruses*. In *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*, pp. 195–202. Edited by M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle & R. B. Wickner. San Diego: Academic Press.
- Chen, X., IJkel, W. F. J., Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, D., Lankhorst, R. K., Vlak, J. M. & Hu, Z. (2001). The sequence of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus genome. *J Gen Virol* **82**, 241–257.
- Chen, X., Zhang, W.-J., Wong, J., Chun, G., Lu, A., McCutchen, B. F., Presnail, J. K., Herrmann, R., Dolan, M., Tingey, S., Hu, Z. H. & Vlak, J. M. (2002). Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *J Gen Virol* **83**, 673–684.
- Cunningham, J. T. (1998). North America. In *Insect Viruses and Pest Management*, pp. 313–331. Edited by F. R. Fujita-Hunter, Ph. F. Entwistle, H. F. Evans & N. E. Crook. New York: John Wiley.
- Funk, C. J., Braunagel, S. C. & Rohrmann, G. F. (1997). Baculovirus structure. In *The Baculoviruses*, pp. 7–32. Edited by L. K. Miller. New York: Plenum.
- Guan, K. L. & Dixon, J. E. (1991). Eukaryotic proteins expressed in *Escherichia coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione *S*-transferase. *Anal Biochem* **192**, 262–267.
- Herniou, E. A., Luque, T., Chen, X., Vlak, J. M., Winstanley, D., Cory, J. S. & O'Reilly, D. R. (2001). Use of whole genome sequence data to infer baculovirus phylogeny. *J Virol* **75**, 8117–8126.
- Hsieh, T. R. (1989). Industrial production of viral pesticides. *Adv Virus Res* **36**, 315–343.
- Ignoffo, C. M. (1973). Development of a viral insecticide: concept to commercialization. *Exp Parasitol* **33**, 380–406.
- IJkel, W. F. J., van Strien, E. A., Heldens, J. G. M., Broer, R., Zuidema, D., Goldbach, R. W. & Vlak, J. M. (1999). Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *J Gen Virol* **80**, 3289–3304.
- IJkel, W. F. J., Westenberg, M., Goldbach, R. W., Blissard, G. W., Vlak, J. M. & Zuidema, D. (2000). A novel baculovirus envelope fusion protein with a proprotein convertase cleavage site. *Virology* **275**, 30–41.
- IJkel, W. F. J., Lebbink, R.-J., Op den Brouw, M. L., Goldbach, R. W., Vlak, J. M. & Zuidema, D. (2001). Identification of a novel occlusion derived virus-specific protein in *Spodoptera exigua* multicapsid nucleopolyhedrosis virus. *Virology* **284**, 170–181.
- McIntosh, A. H. & Ignoffo, C. M. (1983). Characterization of five cell lines established from species of *Heliothis*. *Appl Entomol Zool* **18**, 262–269.
- McIntosh, A. H., Christian, P. D. & Grasele, J. J. (1999). The establishment of heliothine cell lines and their susceptibility to two baculoviruses. *In Vitro Cell Dev Biol Anim* **35**, 94–97.
- Pearson, M. N., Groten, C. & Rohrmann, G. F. (2000). Identification of the *Lymantria dispar* nucleopolyhedrovirus envelope fusion protein provides evidence for a phylogenetic division of the *Baculoviridae*. *J Virol* **74**, 6126–6131.
- Reinhardt, A. & Hubbard, T. (1998). Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Res* **26**, 2230–2236.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sun, X. L., Zhang, G. Y., Zhang, Z. X., Hu, Z. H., Vlak, J. M. & Arif, B. M. (1998). *In vivo* cloning of *Helicoverpa armigera* single nucleocapsid nuclear polyhedrosis virus genotypes. *Virol Sin* **13**, 83–88.
- Zanotto, P. M., Kessing, B. D. & Maruniak, J. E. (1993). Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. *J Invertebr Pathol* **62**, 147–164.
- Zhang, G. (1994). Research, development and application of *Heliothis* viral pesticide in China. *Resour Env Yangtze Valley* **3**, 1–6.