

Deletion of a *Helicoverpa armigera* nucleopolyhedrovirus gene encoding a virion structural protein (ORF107) increases the budded virion titre and reduces *in vivo* infectivity

Xiaoyu Pan,^{1,2} Gang Long,^{1,3} RanRan Wang,^{1,2} Songwang Hou,^{1†} Huiyuan Wang,^{1,2} Yueting Zheng,¹ Xiulian Sun,¹ Marcel Westenberg,³ Fei Deng,¹ Hualin Wang,¹ Just M. Vlask³ and Zhihong Hu¹

Correspondence
Zhihong Hu
huzh@wh.iov.cn

¹State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, PR China

²Graduate School of the Chinese Academy of Sciences, Beijing 100039, PR China

³Laboratory of Virology, Wageningen University, 6709 PD Wageningen, The Netherlands

The open reading frame *Ha107* of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus (HearNPV) encodes a putative protein of 51 kDa with homologues in a few group II NPVs and a granulovirus. *Ha107* was transcribed as polyadenylated transcripts in infected HzAM1 insect cells. The transcripts were initiated at two distinct locations, one upstream of *Ha106* (superoxide dismutase gene, *sod*) and the second upstream of *Ha107*. By Western blot analysis, two forms of the HA107 protein were detected in infected cells, a major polypeptide of 48 kDa and a minor one of 51 kDa. Western blot and immunoelectron microscopy analyses further showed that the HA107 protein was associated with the nucleocapsids of both budded virions (BVs) and occlusion-derived virions. A *Ha107* knockout virus expressing enhanced green fluorescent protein and polyhedrin was constructed using bacmid technology. A one-step virus growth curve indicated that the BV titre of the knockout virus was significantly higher than that of the parental virus and a *Ha107* repair virus. Bioassays indicated that the knockout virus was able to infect third-instar *H. armigera* larvae; however, its median lethal dose (LD₅₀) was significantly higher than those of the parental virus and *Ha107* repair virus. These data indicate that *Ha107* encodes a non-essential structural protein of HearNPV virions and that deletion of this gene increases the BV titre and LD₅₀ of the occluded virus.

Received 8 August 2007
Accepted 23 August 2007

INTRODUCTION

The family *Baculoviridae* is a group of occluded, rod-shaped viruses, with circular, supercoiled, double-stranded DNA genomes of a size ranging from 80 to 230 kbp. This family of insect viruses has been divided taxonomically into two genera: *Nucleopolyhedrovirus* (NPV), members of which have multiple virions present within large polyhedron-shaped occlusion bodies, and *Granulovirus* (GV), members of which have a single enveloped nucleocapsid embedded in a granule occlusion body (Theilmann *et al.*, 2004). Phylogenetic studies have indicated that NPVs can be subdivided into two groups, I and II (Bulach *et al.*, 1999; Herniou *et al.*, 2003; Zanotto *et al.*, 1993). During the infection cycle of lepidopteran baculoviruses, two distinct viral phenotypes are produced, the occlusion-derived virus

(ODV) and budded virus (BV) (Volkman & Summers, 1977). ODVs are encapsulated in polyhedra that dissolve in the alkaline environment of the midgut, release the virions and initiate infection in midgut columnar epithelial cells. The ODV is associated with the spread of the virus from insect to insect, whereas BVs are not occluded and are adapted to disseminate infection from cell to cell and are responsible for the systemic infection in larvae and in cell cultures. Although ODVs and BVs differ in the composition of their envelopes, the nucleocapsids of the two viral phenotypes are similar in structure and function (Funk *et al.*, 1997). The functions of many of the structural virion proteins, however, are unknown.

Helicoverpa armigera single nucleocapsid NPV (HearNPV, also called HaSNPV) is a highly infectious pathogen of the cotton bollworm *H. armigera* and related heliothines (Sun *et al.*, 1998). The virus was isolated from diseased *H. armigera* larvae in the Hubei province of China in 1975

†Present address: Pharmacology and Toxicology, Medical College of Wisconsin, Milwaukee, WI 53226, USA.

(Zhang, 1989). HearNPV has been adapted for production as a biological pesticide and has been applied to control insect pests in China and other countries on cotton and vegetables. The genome sequence of the HearNPV G4 strain has been determined in its entirety (Chen *et al.*, 2001). Computer-assisted analysis has revealed 135 putative open reading frames (ORFs) among which 100 ORFs have homologues in the *Autographa californica* multicapsid NPV (AcMNPV), 15 have homologues in other baculoviruses and the remaining 20 ORFs are so far unique to HearNPV and its genotypic variant *Helicoverpa zea* SNPV (HzSNPV) (Chen *et al.*, 2002).

Ha107 is one of the 15 ORFs with homologues in other baculoviruses and is currently uncharacterized. The *Ha107* gene has been predicted to encode a protein of about 51 kDa. A homologue of this ORF is present in the HearNPV Cl strain (Zhang *et al.*, 2005) and in HzSNPV (Chen *et al.*, 2002). Moreover, homologues of *Ha107* are also found in a limited number of baculovirus species, such as *Agrotis segetum* multicapsid nucleopolyhedrovirus (AgseNPV; Jakubowska *et al.*, 2006), *Spodoptera exigua* MNPV (SeMNPV; IJkel *et al.*, 1999), *S. litura* MNPV (SpltNPV; Pang *et al.*, 2001) and *A. segetum* granulovirus (AgseGV; unpublished GenBank accession no. NC_005839), but not in group I NPVs.

In this study, the transcription characteristics of *Ha107* and the expression and location of its encoded protein were investigated. The HA107 protein appeared to be a nucleocapsid-associated protein of both BV and ODV. In addition, *Ha107* was not essential for virus replication, but deletion of this gene appeared to increase the BV titre and the median lethal dose (LD₅₀) of the recombinant virus.

METHODS

Cells, virus and insects. The HzAM1 cell line (McIntosh *et al.*, 1999) was maintained in Grace's medium with 10% fetal bovine serum. HearNPV strain G4 (Sun *et al.*, 1998) was used for experiments and was propagated in HzAM1 cells. A culture of *H. armigera* insects was maintained according to Sun *et al.* (1998) for HearNPV occlusion body production and bioassays.

Total RNA isolation and 3' and 5' rapid amplification of cDNA ends (RACE) analysis. HzAM1 cells were infected with BVs of HearNPV at an m.o.i. of 10. Total RNA was isolated from infected and mock-infected cells at 0, 4, 8, 12, 24, 36, 48, 72 and 96 h post-infection (p.i.) by Trizol extraction (Invitrogen). RT-PCR was performed with 1 µg purified total RNA as template per time point. First-strand cDNA synthesis was performed using AMV reverse transcriptase (Promega) and an oligo(dT) anchor primer to the 3' end of *Ha107* (5'-GACCACGCGTATCGATGTCGACTTTTTTTTTT-TTTTTTV-3'). The cDNA mixtures were amplified using an anchor primer (5'-GACCACGCGTATCGATGTCGAC-3') and *Ha107*-specific forward primer 1 (5'-GTTGCCCACTCGCGACATAAG-3'). The PCR products obtained from RNA at 48 h p.i. were cloned into the pGEM-T Easy vector (Promega) and sequenced to determine the 3' end of the *Ha107* transcripts. The 3'RACE PCR of *ORF106* of HearNPV (*Ha106*) was amplified with the oligo(dT) anchor primer and *Ha106*-specific forward primer 2 (5'-GCGTGATTGCACG-GTATACAC-3').

The 5' initial sites of the *Ha107* transcripts were determined using a 5'/3' RACE kit (Roche) with 2 µg purified total RNA (48 h p.i.) as a template. Briefly, first-strand cDNA synthesis was performed with a gene-specific primer (5'-CACAGAACCCGCGAGCAACG-3'). cDNAs were then purified using a High Pure PCR purification kit (Roche) and a poly(A) tail was added to the 3' ends using terminal transferase with dATP. The tailed cDNAs were amplified using the oligo(dT) anchor primer and the first nested *Ha107*-specific primer (R1: 5'-AACAAAGGGCGGTATAAGTTGAAC-3'). A second PCR was performed using the PCR anchor primer and the second nested primer (R2: 5'-GTCGCTTTGCTAGTGCTGCTAC-3'). The obtained PCR products were gel purified, cloned into pGEM-T Easy and sequenced.

Generation of polyclonal antibody against HA107. A truncated fragment of *Ha107* (nt 101017–102391) (Chen *et al.*, 2001) encoding the hydrophilic loops (aa 136–304) (Fig. 1) was PCR-amplified using two primers, ExF (5'-GGGGATCCGATTCGGTAATTGCTTTAAT-GTAC-3') (*Bam*HI site underlined) and ExR (5'-GGGGAATTC-AGCGAGTGGGCAACATTATCGT-3') (*Eco*RI site underlined). The purified PCR product was cloned into the pGEX-KG expression vector (Guan & Dixon, 1991), giving pGEX-EX107 and expressed as truncated HA107 protein fused to glutathione S-transferase (GST-EX107) in *E. coli* BL21 cells. The purified GST-EX107 proteins from gel slices were injected into a rabbit (300 µg per injection for initial as well as booster) to generate anti-HA107 antibodies.

Western blot analysis of HA107 expression. Monolayers of HzAM1 cells were mock- and HearNPV-infected at an m.o.i. of 5. Cells were harvested at 0, 8, 12, 24, 48, 72 and 96 h p.i. and subjected to SDS-PAGE and Western blot analysis according to the method of Long *et al.* (2003). Membranes were allowed to react with the anti-HA107 antiserum diluted 1:1500 and bands were detected after incubation with horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin (Sigma) using a SuperSignal West Femto Maximum Sensitivity Substrate kit as described by the manufacturer (Pierce).

Purification of BV and ODV fractions and Western blot analysis. The envelope and nucleocapsid fractions of BV and ODV were separated after treatment with NP-40 as previously described (Long *et al.*, 2003). Proteins of purified BVs and ODVs, as well as their nucleocapsid and envelope fractions, were separated by 12% SDS-PAGE and transferred onto Hybond-N membranes (Amersham). Anti-HA107 antiserum was used for Western blot analysis and polyclonal anti-VP39, anti-ODV-E18 and anti-HaF1 antibodies were used as controls for nucleocapsid-, ODV envelope- and BV envelope-specific proteins, respectively.

Immunoelectron microscopy (IEM) analysis. The IEM protocol was according to Tsai *et al.* (2006). Suspensions of purified ODVs or their nucleocapsid fractions were diluted and adsorbed to Formvar-supported, carbon-coated nickel grids (250 mesh) for 30 min at 4 °C. The grids were blocked in blocking buffer [4% BSA, 50 mM Tris/HCl (pH 7.5), 200 mM NaCl] for 30 min at 37 °C and then incubated with anti-HA107 antiserum or pre-immune serum (1:20 dilution) in incubation buffer [0.1% BSA, 50 mM Tris/HCl (pH 7.5), 200 mM NaCl] for 2 h at 37 °C. After incubation and several washes with incubation buffer, the grids were incubated with goat anti-rabbit secondary antibody conjugated with gold particles (1:20 dilution in incubation buffer, 12 nm diameter; Jackson ImmunoResearch) for 1 h at 37 °C. The grids were then washed extensively with incubation buffer and negatively stained with 2% phosphotungstate. Specimens were examined with a transmission electron microscope (H-7000 FA; Hitachi).

Construction of *Ha107* knockout HearNPV bacmid. The *Ha107* knockout HearNPV bacmid was constructed by using a modified λ phage Red recombination system (Datsenko & Wendell, 2000). Two

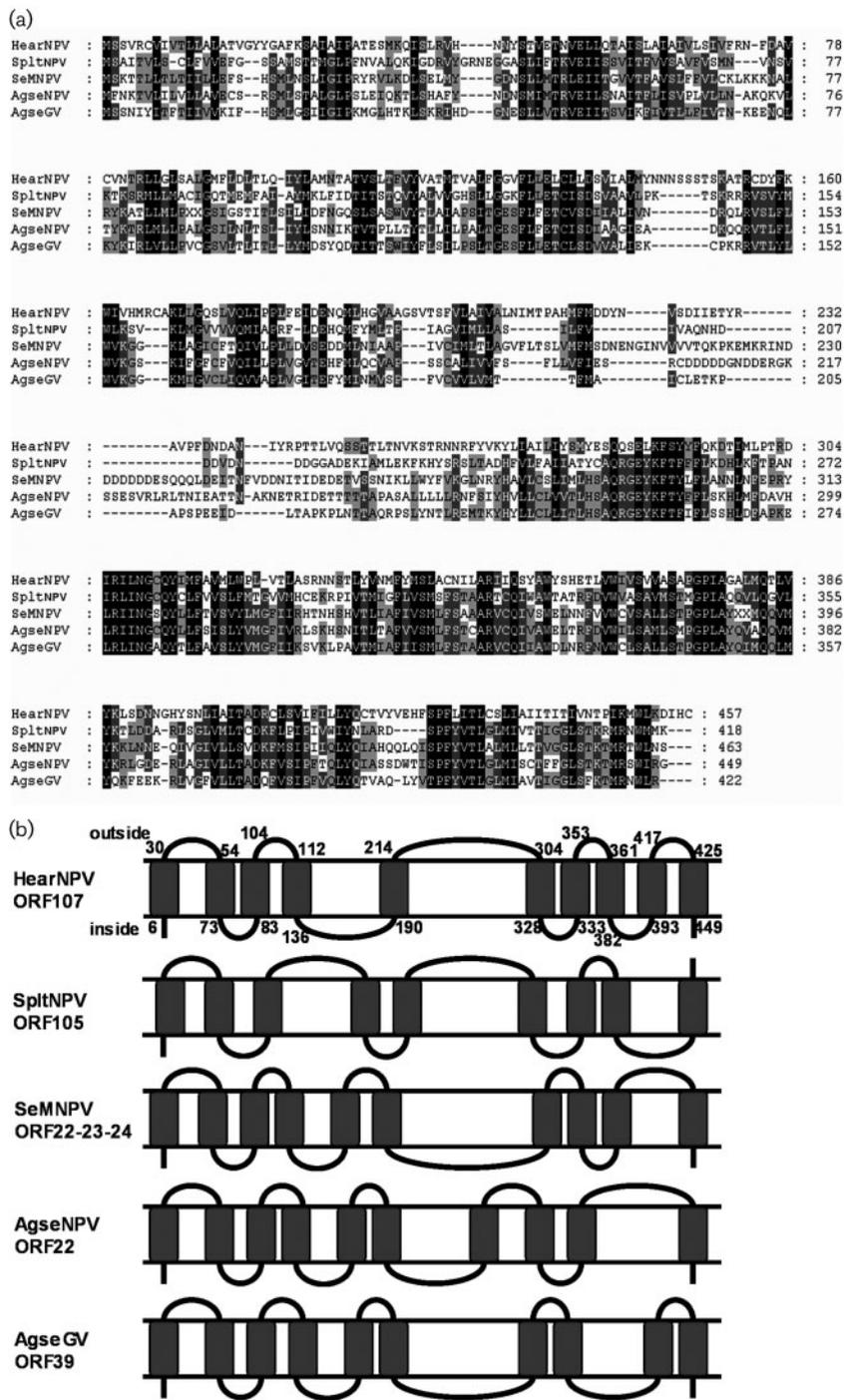


Fig. 1. Comparison of HA107 homologues. (a) Comparison of the predicted amino acid sequences of the HA107, SpltnPV ORF105, AgseNPV ORF22, SeMNPV ORF22-23-24 and AgseGV ORF39 proteins. Alignment was performed using MEGALIGN software with CLUSTAL W and was edited using Gendoc software. (b) Predicted topology of HA107 homologues. Shaded boxes represent predicted transmembrane domains. The figure was drawn according to the TMHMM software-predicted result. The positions of the amino acid sequences are indicated.

primers were designed to generate by PCR a linear fragment containing a chloramphenicol resistance gene (Cm^R) and with 43 bp flanking sequences targeting the *Ha107* region on the HearNPV genome. The forward primer was 5'-TGTACGATGTG-TCATCGTAACGTTATTGGCGCTCGCGACAGTGGGCACCAATA-ACTGCCT-3' (nucleotides homologous to *Ha107* are underlined). A stop codon (in bold) was introduced in the primer. The reverse primer was 5'-TAATTTGTACTAAAGTTTGCATTAAAGCGCC-AGCAATTGGACCTGTGCGACGTTACG-3' (nucleotides homologous to *Ha107* are underlined). The 3' end of the primers anneal to the Cm^R gene of pBeloBac11 (Shizuya *et al.*, 1992). After *DpnI*

digestion, the linear PCR fragment was electrotransferred into competent DH10B cells carrying HearNPV bacmid HaBacHZ8 and the helper plasmid pKD46 (Hou *et al.*, 2002). The resulting recombinant bacmid was designated HaBac Δ Ha107.

The *Ha107* truncation and Cm^R gene insertion in HaBac Δ Ha107 was confirmed by *EcoRI* digestion and PCR analysis with two pairs of primers. The first pair was used to confirm insertion of the Cm^R gene cassette into the *Ha107* locus of the HearNPV bacmid: one primer corresponded to sequences within the inserted Cm^R gene (CM: 5'-GTATGGCAATGAAAGACGGTGAG-3') and the other (EnF:

5'-GGGATGTCTTCTGTACGATGTGTCAT-3', underlined nucleotides correspond to the virus genome nt 101017–101039) was from the HearNPV genome outside the sequence used for recombination. With this pair of primers, a PCR product of 530 bp should be amplified from HaBacΔHa107 but not from the control bacmid HaBacHZ8. The second pair of primers (ExF and ExR) was within the *Ha107* cassette and was used to confirm the absence of the *Ha107* gene. These two primers should not produce a PCR product in HaBacΔHa107 but should produce a 524 bp product in the parental HaBacHZ8. As the inserted *Cm^R* gene contained an extra *EcoRI* site, this enzyme was used to identify the bacmids. According to software-predicted results, a 5.99 kb *EcoRI* fragment from HaBacHZ8 should be replaced by two smaller *EcoRI* fragments of 2.91 and 2.92 kb, respectively, in HaBacΔ107.

Construction of knockout, repair and control HearNPV bacmids containing the polyhedrin and enhanced green fluorescence protein genes. To insert the polyhedrin gene (*ph*) and the enhanced green fluorescence protein gene (*egfp*) into HaBacΔHa107 and HaBacHZ8, a donor plasmid, pFastBac-eGFP-PH, was constructed. The *egfp* ORF was amplified using primer pair GFP-F (5'-GACCTCGAGATGGTGAGCAAGGGCGAGGA-3') and GFP-R (5'-GCCGCTAGCTTACTTGTACAGCTCGTCCATG-3') containing *XhoI* and *NheI* (underlined) sites, respectively, from the pEGFP-N1 vector (Invitrogen). The digested PCR product was cloned into the pFastBacDual vector (Invitrogen), giving pFastBac-eGFP. The HearNPV *ph* gene and its own promoter were PCR amplified with primers BacpolhF (5'-GGGGATATCGTTGAAGGAGCGTACGTGCCA-3') and BacpolhR (5'-GGGGAATTCAATCGCAAGTTTAAATATGCAGGAC-3'), which contained *EcoRV* and *EcoRI* (underlined) sites, respectively. After sequence verification, the *ph* gene was inserted into *Bst*1107I/*EcoRI*-restricted pFastBac-eGFP vector, resulting in pFastBac-eGFP-PH.

To generate a *Ha107* repair HearNPV bacmid, a donor plasmid, pFastBac-eGFP-PH-Ha107, was constructed as follows. A 238 bp *NotI*-*AflIII* fragment containing the SV40 poly(A) tail addition signal was obtained from the vector pEGFP-N1. A 1.7 kb *AflIII*-*XbaI* fragment containing the coding region of *Ha107* and its proximal promoter region was amplified using forward primer 5'-GGGCTTAAGCGACACCAGCAACGGATGTA-3' and reverse primer 5'-GCGTCTAGATTAGCAGTGTATATCTTTTAAACCA-3'. After purification and sequence verification of the PCR product, the *Ha107* insert and digested SV40 poly(A) were inserted into *NotI*/*XbaI*-digested vector pFastBac-eGFP-PH at the same time to give the donor plasmid pFastBac-eGFP-PH-Ha107.

Bacmid transpositions were performed according to the Bac-to-Bac manual (Invitrogen) to insert the donor sequences into the various bacmids. The resulting recombinant bacmids were named: HaBacHZ8-eGFP-PH (the control bacmid HaBacHZ8 containing *ph* and *egfp*), HaBacΔHa107-eGFP-PH (*Ha107* knockout bacmid containing *ph* and *egfp*) and HaBac-rHa107-eGFP-PH (*Ha107*-repaired bacmid containing *ph* and *egfp*) (see Fig. 4a). These recombinant bacmids were used to transfect HzAM1 cells to obtain the corresponding viruses.

One-step viral growth curves. To investigate the growth rate of virus constructs, HzAM1 cells were infected with HaBacHZ8-eGFP-PH, HaBacΔHa107-eGFP-PH or HaBac-rHa107-eGFP-PH at an m.o.i. of 10 and BV titres were determined by end-point dilution assay at 0, 8, 16, 24, 36, 48, 72, 96 and 120 h p.i. Green fluorescence was used as a marker to determine positive viral infection. Each virus infection was carried out in triplicate. BV titres at different time p.i. were analysed by two-way analysis of variance in GLM (SPSS Inc.) with virus and time as factors. Mean BV titres of virus types were separated by Fisher's least significant difference if significant effects were found. BV titres were logarithm transformed before analysis.

Bioassay of recombinant viruses. LD₅₀ values of the recombinant HearNPVs (HaBacHZ8-eGFP-PH, HaBacΔHa107-eGFP-PH and HaBac-rHa107-eGFP-PH) in third-instar *H. armigera* larvae were determined with five different doses (20, 60, 200, 600 and 2000 polyhedra per larva). Second-instar larvae were starved for 16 h at 28 ± 1 °C and allowed to moult into third instar. The starved larvae (4.61 ± 0.34 mg) were fed on a diet plug with 1 μl polyhedra suspension. The larvae that consumed the entire plug within 24 h were transferred to new diet. For each dose of each virus, 48 larvae were inoculated. Inoculated larvae were maintained at 28 ± 1 °C. Mortality was scored daily until surviving larvae in each treatment had either pupated or were in the pre-pupal stage. LD₅₀ values and their standard deviations were determined by probit analysis using SPSS (SPSS Inc.). LD₅₀ values of viruses were further compared using two-sided z-tests (Snedecor & Cochran, 1989).

Electron microscopy of recombinant virus-infected cells. HzAM1 cells were infected with HaBacHZ8-eGFP-PH or HaBacΔHa107-eGFP-PH at an m.o.i. of 5. Infected cells were fixed at 60 h p.i. and processed for electron microscopic analysis as described previously (Wang *et al.*, 2003).

Computational analysis. The HA107 protein sequence was analysed using the ExPASy server (Appel *et al.*, 1994) (<http://us.expasy.org>) and CBS server (Jensen *et al.*, 2003) (<http://www.cbs.dtu.dk>) for prediction of signal peptide, transmembrane domain and potential post-translational modification sites and subcellular localization. Homologues were explored using the BLAST network service on ExPASy in the Swiss-Prot and TrEMBL databases (lambda, 0.327; K, 0.136; H, 0.405). The topology of the HA107 protein homologues was predicted using TMHMM software. The topology was drawn according to the predicted results.

RESULTS

Sequence of *Ha107* and its homologues

The *Ha107* gene is located downstream of and in the same orientation as *Ha106*, a homologue of a superoxide dismutase (*sod*) gene (Chen *et al.*, 2001) (Fig. 2a). The regions upstream of *Ha107* were analysed for the presence of putative transcriptional initiation sites, such as baculovirus consensus early CAG/TT, late DTAAG promoter motifs and other transcriptional regulation elements, such as GATA and CACGTG. No consensus motif was found upstream of the *Ha107* translational start codon ATG. A polyadenylation signal sequence (A)ATAAA was located 20 nt downstream of the putative translation stop codon TAA.

The coding region of *Ha107* is 1374 nt and potentially encodes a 51 kDa protein of 457 aa. Using TMHMM software, ten transmembrane regions (including the signal peptide) were predicted, located at aa 7–29, 55–72, 84–103, 113–135, 191–213, 305–327, 334–352, 362–381, 394–416 and 426–448, and two hydrophilic loops were predicted at aa 136–190 and 214–304 (Fig. 1b). An N-terminal signal peptide was predicted and a potential cleavage site was found between aa 22 and 23 (YGA ↓ FK). Five potential N-linked glycosylation sites (aa 45–48, 145–148, 146–149, 223–226 and 331–334) were found using NetNglyc prediction. According to PSORTII analysis (Horton & Nakai, 1997), the putative HA107 protein is probably

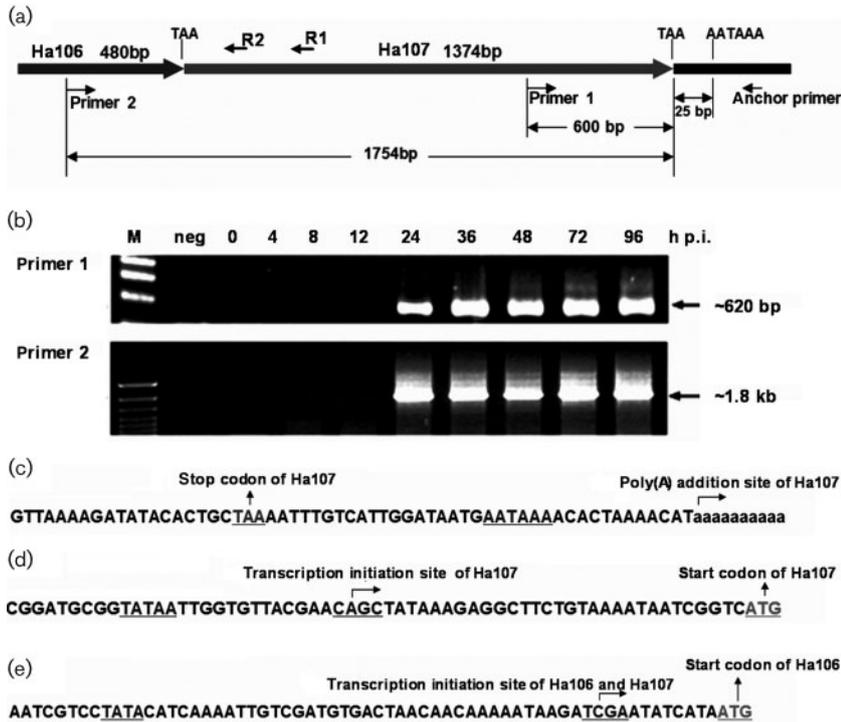


Fig. 2. Transcription of *Ha107* in HearNPV-infected HzAM1 cells. (a) Schematic representation of the HearNPV *Ha107* locus with the relative positions and orientation of *Ha106* and *Ha107*. Primers used for mapping the *Ha107* transcripts by 3'RACE analyses are shown as short arrows below the ORFs. (b) 3'RACE analysis of *Ha107* in HearNPV-infected HzAM1 cells from 0 to 96 h p.i. The location of the specific primers is shown in (a). (c) Sequence at the 3' end of *Ha107*. The poly(A) tail addition signal (AATAAA) and the translation stop codon of *Ha107* are underlined. The arrow points to the poly(A) tail addition site. (d, e) Sequences at the 5' end of *Ha107* and *Ha106*, respectively. The TATA box (TATAA) and the transcription start codon of *Ha107* are underlined. The rightward arrow indicates the location of the transcription initiation sites.

localized in or associated with membrane structures. Only four homologous genes of *Ha107* were found in the available baculovirus genomes (Fig. 1a) including ORF105 of SpltNPV (Pang *et al.*, 2001; GenBank accession no. NC_003102), ORF22 of AgseNPV (Jakubowska *et al.*, 2006; GenBank accession no. NC_007921), ORF22-23-24 of SeMNPV (IJkel *et al.*, 1999; GenBank accession no. NC_002169) and ORF39 of AgseGV (unpublished data, GenBank accession no. NC_005839). The amino acid identity among these homologous proteins was low (20.1–47.4%), but the predicted topological structure was rather conserved among the proteins from NPVs (Fig. 1b).

Transcriptional mapping of the 3' and 5' end of *Ha107* transcripts

To determine whether *Ha107* was transcribed, 3'RACE analysis was performed with total RNA purified from mock-infected and HearNPV-infected HzAM1 cells at various times p.i. (Fig. 2b). A product of the expected size (620 bp) was first detected at 24 h p.i., which increased in amount up to 96 h p.i., suggesting that *Ha107* is transcriptionally active, predominantly late during infection. The 3'RACE products from 48 h p.i. were cloned into pGEM-T Easy and sequenced. The obtained sequences mapped the 3' end of the *Ha107* transcript at 38 nt downstream of the putative translation stop codon TAA and 13 nt downstream of the last A of the polyadenylation signal sequence AATAAA (Fig. 2c).

5'RACE was performed using total RNA extracted at 48 h p.i. from infected HzAM1 cells to determine the 5' end of

the *Ha107* transcripts. The lack of an initial PCR product using the first reverse primer (R1) necessitated the use of a nested primer (R2) to detect a PCR product (Fig. 2a). Five clones were sequenced and putative start sites were identified in two regions upstream of the *Ha107* translational start site ATG. One start site was 33 nt upstream of the ATG codon of *Ha107* at an A residue (Fig. 2d); the other was a C residue located at 11 nt upstream of the ATG codon of *Ha106* (Fig. 2e). The presence of a 1.8 bp 5'RACE product also showed that *Ha107* not only has its own transcription initiation site, but is also transcribed as one transcription unit together with *Ha106* (*sod*).

Time-course analysis of the HA107 protein in HearNPV-infected cells

Synthesis of the HA107 protein in infected HzAM1 cells was followed by Western blot analysis using a polyclonal antiserum raised by immunization of a rabbit with a bacterially expressed truncated HA107 protein (GST-EX107). The results showed that anti-HA107 antiserum recognized two bands of 48 and 51 kDa, respectively, in infected cells (Fig. 3a). The major band (48 kDa) appeared at 24 h p.i. and reached maximum levels at 72 h p.i. The minor 51 kDa band, detected as early as 48 h p.i., was the same as the theoretical size of HA107. The protein levels remained relatively high throughout late times of infection. The size of the major protein form of HA107 was smaller than the predicted size of the putative *Ha107* translation product, possibly as the result of removal of the signal peptide.

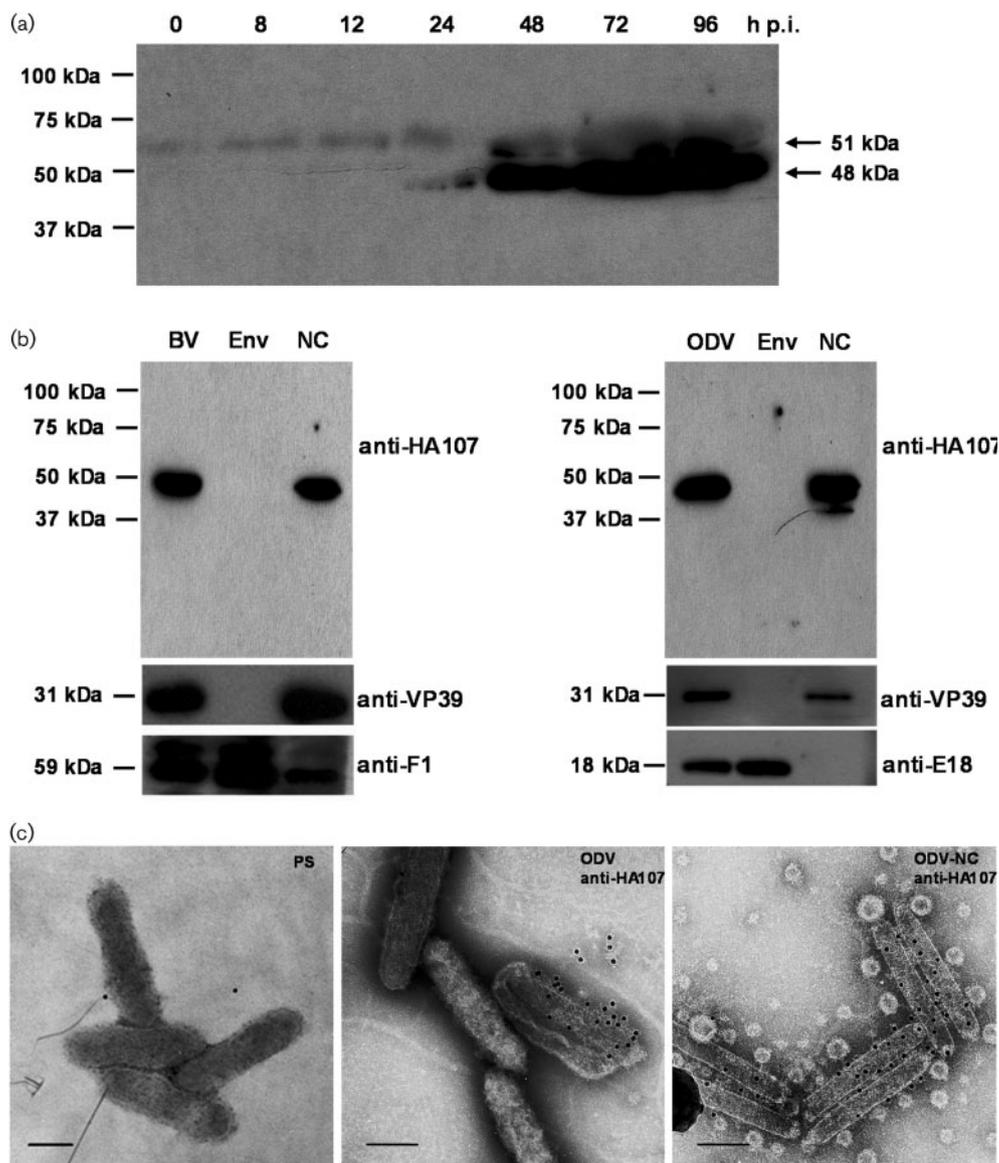


Fig. 3. Immunodetection of the HA107 protein in HearNPV-infected HzAM1 cells and virions. (a) Time-course analysis of HA107 production in HearNPV-infected HzAM1 cells. The corresponding times p.i. are indicated above the lanes. Total cell proteins were separated by SDS-PAGE, blotted onto an Immuno-P membrane and detected using anti-HA107 antiserum by chemiluminescence. The estimated sizes of detected bands are indicated with arrows. (b) Western blot analysis of HearNPV virions and separated fractions using anti-HA107 antiserum. BV, ODV and separated envelope (Env) and nucleocapsid (NC) fractions of BV and ODV were subjected to SDS-PAGE followed by Western blotting using anti-HA107, anti-ODV-E18 (ODV envelope-specific), anti-F1 (BV envelope-specific) and anti-VP39 (nucleocapsid-specific) antisera. (c) IEM analysis of HA107 in ODV. PS, ODV sample incubated with pre-immune rabbit serum; ODV anti-HA107, ODV sample incubated with anti-HA107 serum; ODV-NC anti-HA107, nucleocapsids of ODV sample incubated with anti-HA107 serum. Bars, 100 nm.

Immunodetection of the HA107 protein in virions

To investigate whether or not HA107 is a structural component of HearNPV virions, Western blot and IEM analyses were conducted. Western blot analysis using anti-HA107 antiserum recognized a 48 kDa band in ODV and BV, specifically in the nucleocapsid fractions (Fig. 3b).

These data indicated that the HA107 protein is a structural component of the nucleocapsids of both BV and ODV. To study the location of HA107 further, IEM analysis was used and showed that the anti-HA107 antiserum recognized only the nucleocapsids and not ODVs with intact envelope (Fig. 3c). This observation further confirmed that HA107 is a nucleocapsid-associated protein.

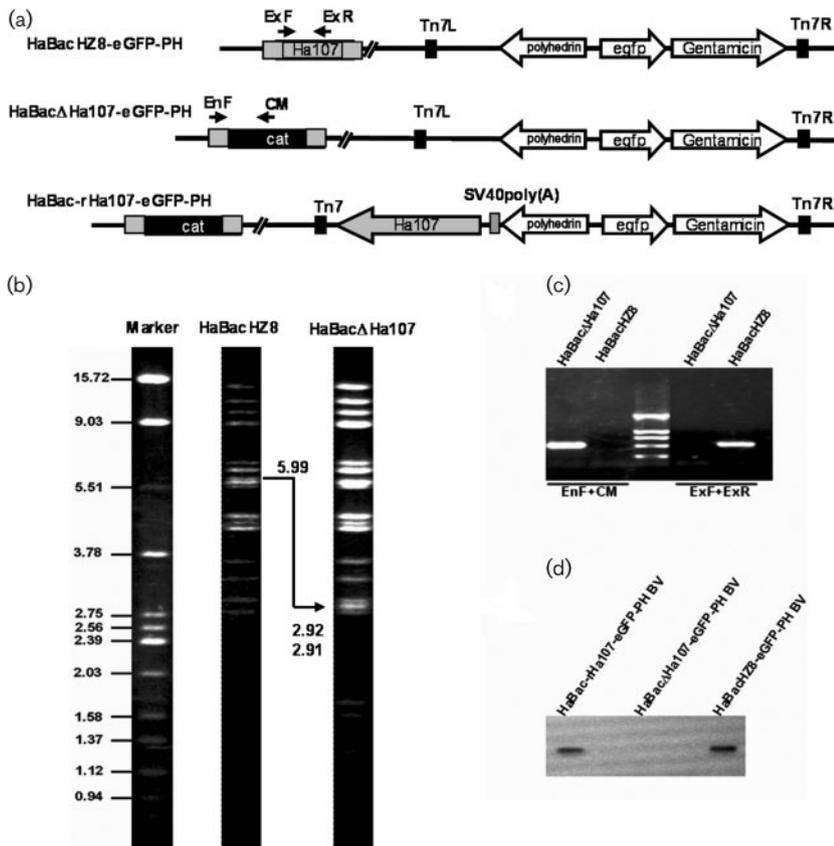


Fig. 4. Construction and identification of *Ha107* knockout, repair and control HearNPV viruses. (a) Schematic diagram of *HaBacΔHa107-eGFP-PH*, *HaBacHZ8-eGFP-PH* and *HaBac-rHa107-eGFP-PH*. The positions of primer pairs used in the analysis to confirm the deletion of *Ha107* and correct insertion of the *cat* gene (*Cm^R*) cassette are indicated. (b) Identification of *HaBacΔHa107* by *EcoRI* digestion. The size (kb) of the bands that were altered is indicated. (c) PCR identification of *HaBacΔHa107*. The bacmids analysed are shown above each lane and the primer pairs used are shown below. (d) Western blot analysis of the presence or absence of the HA107 protein in BVs of the recombinant viruses.

Construction of *Ha107* knockout and repair HearNPV bacmids

To investigate the role of *Ha107* during the viral infection cycle, a *Ha107* knockout virus (*HaBacΔHa107*) was generated from the HearNPV bacmid *HaBacHZ8* (Wang *et al.*, 2003). The λ phage Red homologous recombination system (Hou *et al.*, 2002) was used to obtain the *Ha107* knockout bacmid. PCR analysis and *EcoRI* digestion were used to verify the construct. As expected, with primers *CM* and *EnF*, a PCR product of 530 bp was amplified from *HaBacΔHa107* but not from the control bacmid *HaBacHZ8*, and with primers *ExF* and *ExR*, a 524 bp PCR product was amplified from *HaBacHZ8* but not from *HaBacΔHa107* (Fig. 4c). *EcoRI* digestion profiles revealed that the 5.99 kb *EcoRI* fragment from *HaBacHZ8* containing HA107 was replaced by two smaller *EcoRI* fragments of 2.91 and 2.92 kb in *HaBacΔHa107* (Fig. 4b) indicating correct insertion of the *Cm^R* gene. Therefore, the *Ha107* gene was successfully eliminated in *HaBacΔHa107*.

The *egfp* and *ph* genes were inserted into *HaBacHZ8* and *HaBacΔHa107* to generate *HaBacHZ8-eGFP-PH* and *HaBacΔHa107-eGFP-PH* (Fig. 4a). To rescue the wild-type phenotype of the *Ha107* knockout, a repair bacmid, *HaBac-rHa107-eGFP-PH*, was constructed. The *Ha107* cassette with its own promoter was inserted into the polyhedrin locus of *HaBacΔHa107* by Tn7-mediated transposition to give *HaBac-rHa107-eGFP-PH* (Fig. 4a).

All insertion constructs were confirmed by PCR analysis (data not shown).

Viral replication in HzAM1 cells

To determine whether *Ha107* is essential for viral replication, HzAM1 cells were transfected separately with *HaBacΔHa107-eGFP-PH*, *HaBacHZ8-eGFP-PH* and *HaBac-rHa107-eGFP-PH* bacmids and the transfection assay was monitored by eGFP expression. At 4 days post-transfection, green fluorescence and polyhedra formation were detected in all transfections. The supernatants were collected at 5 days post-transfection and used to infect fresh HzAM1 cells. After 3 days, green fluorescence was clearly observed in cells infected with each of the recombinant viruses (Fig. 5a). Normal polyhedra formation was also observed in the infected cells. Western blot analysis of purified BVs of *HaBacΔHa107-eGFP-PH*, *HaBac-rHa107-eGFP-PH* and *HaBacHZ8-eGFP-PH* was performed to confirm the effective removal of *Ha107* (Fig. 4d). The HA107 protein was detected in BVs of both *HaBac-rHa107-eGFP-PH* and *HaBacHZ8-eGFP-PH*, but not in *HaBacΔHa107-eGFP-PH*. These results indicated that *Ha107* is not essential for viral replication and propagation in cell culture.

A virus growth curve assay was performed to study further the possible effect of deletion of *Ha107* on viral replication.

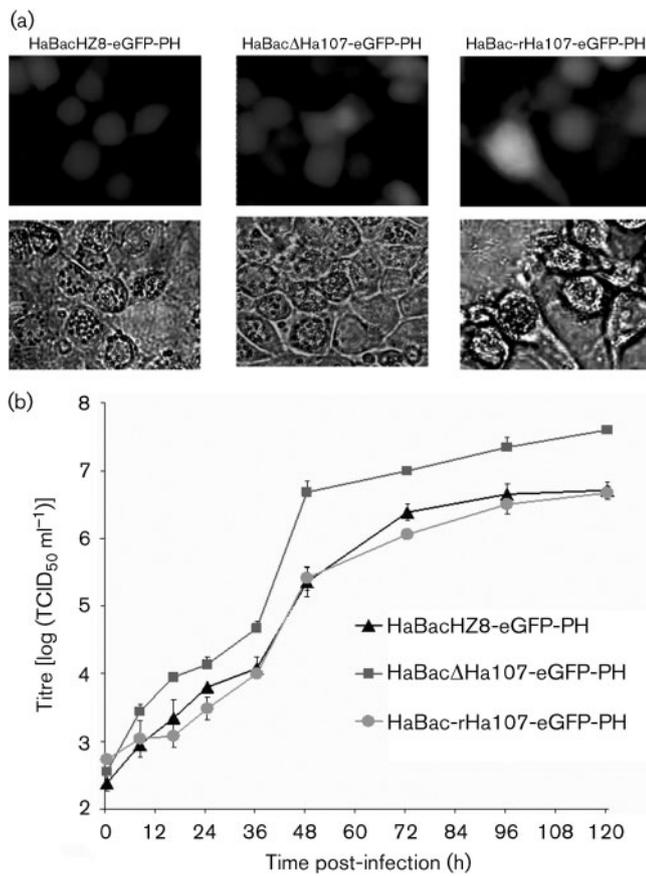


Fig. 5. Analysis of virus replication of HaBacΔHa107-eGFP-PH, HaBacHZ8-eGFP-PH and HaBac-rHa107-eGFP-PH in infected HzAM1 cells. (a) Fluorescence (upper panel) and light microscopy (lower panel) of HzAM1 cells infected with HaBacΔHa107-eGFP-PH, HaBacHZ8-eGFP-PH or HaBac-rHa107-eGFP-PH, analysed at 72 h p.i. (b) Virus growth curves of HaBacΔHa107-eGFP-PH, HaBacHZ8-eGFP-PH and HaBac-rHa107-eGFP-PH in HzAM1 cells. Cells (3.0×10^5) were infected at an m.o.i. of 10 for each virus and the supernatant was harvested at the indicated time points p.i. and assayed for the production of infectious virus by end-point dilution assay. Each point represents the average titre derived from three independent TCID₅₀ assays. Error bars represent SEM.

HzAM1 cells were infected separately with HaBacΔHa107-eGFP-PH, HaBacHZ8-eGFP-PH or HaBac-rHa107-eGFP-PH at an m.o.i. of 10. Statistical analysis revealed that HaBacΔHa107-eGFP-PH produced BV to a significantly higher titre than HaBacHZ8-eGFP-PH ($F=100.84$, d.f.=1, 54; $P<0.01$) and the HA107 rescue virus HaBac-rHa107-eGFP-PH ($F=100.84$, d.f.=1, 54; $P<0.01$) (Fig. 5b).

Polyhedra formation and oral infectivity

At 60 h p.i., polyhedra were seen by light microscopy in the nuclei of HzAM1 cells infected with HaBacΔHa107-eGFP-PH, HaBacHZ8-eGFP-PH and HaBac-rHa107-eGFP-PH (Fig. 5a). Infected cells were harvested and analysed by electron microscopy, which revealed the formation of

polyhedra containing large numbers of occluded virions (Fig. 6). Polyhedra with singly enveloped virions in HaBacΔHa107-eGFP-PH-infected cells had a similar shape and size to those of HaBacHZ8-eGFP-PH. These results indicated that the absence of *Ha107* did not appear to influence the formation and structure of polyhedra or ODVs.

The infectivity of the recombinant viruses was assayed in *H. armigera* larvae (Table 1). Third-instar larvae were fed polyhedra of HaBacΔHa107-eGFP-PH, HaBac-rHa107-eGFP-PH or HaBacHZ8-eGFP-PH and mortality was monitored daily. The results showed that HaBacΔHa107-eGFP-PH was infectious to *H. armigera* larvae, indicating that HA107 is not essential for infectivity *in vivo*. Notably, its LD₅₀ was about 10 times higher than that of the control virus ($z=4.972$, $P<0.01$) and 15 times higher than that of the *Ha107* repair virus ($z=5.181$, $P<0.01$) (Table 1). Statistically, the difference between the LD₅₀ values of the control virus and *Ha107* repair virus was not significant ($z=1.598$, $P>0.05$). The slopes of the regression lines between mortality (probit) and log (dose) of these three viruses were not significantly different from each other ($z=0.885$, $P>0.05$, for HaBacΔHa107-eGFP-PH and HaBacHZ8-eGFP-PH; $z=1.656$, $P>0.05$, for HaBacΔHa107-eGFP-PH and HaBac-rHa107-eGFP-PH; $z=0.663$, $P>0.05$, for HaBac-rHa107-eGFP-PH and HaBacHZ8-eGFP-PH).

DISCUSSION

In this study, we have reported the characterization of the *Ha107* gene and its role in virus infection. The existence of *Ha107* homologues in several NPVs (HearNPV, SpltNPV, SeMNPV and AgseNPV) and their phylogenetic position in

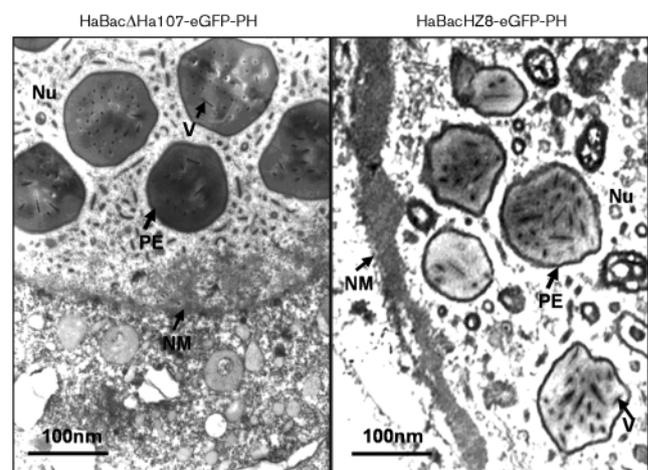


Fig. 6. Electron microscopy of cells infected with the recombinant viruses. HzAM1 cells were infected with HaBacΔHa107-eGFP-PH or HaBacHZ8-eGFP-PH at an m.o.i. of 5. Cells were fixed at 60 h p.i. Nu, nucleus; NM, nuclear membrane; V, virion; PE, polyhedron envelope.

Table 1. LD₅₀ values of HaBacΔ*Ha107*-eGFP-PH, HaBac-r*Ha107*-eGFP-PH and HaBacHZ8-eGFP-PH in third-instar *H. armigera* larvae

Virus	LD ₅₀ (polyhedra per larva)	95 % Confidence limits	
		Lower	Upper
HaBacΔ <i>Ha107</i> -eGFP-PH	307.8	220.8	437.4
HaBacHZ8-eGFP-PH	31.6	20.9	43.3
HaBac-r <i>Ha107</i> -eGFP-PH	20.4	11.1	30.0

a common clade (Jakubowska *et al.*, 2006) suggest that the *Ha107* gene may have been derived from an ancestral virus. The observation that only one GV (AgseGV) has this gene might be explained by the fact that AgseGV and AgseNPV have the same host, *A. segetum*, and horizontal gene transfer may have occurred.

Ha107 potentially encodes a 51 kDa protein, but two forms of this protein (51 and 48 kDa) were found in infected cells. The smaller but major form (48 kDa) was present in nucleocapsids in both BV and ODV as elucidated by Western blot analysis (Fig. 3). This is in line with the possibility that the predicted 22 aa signal peptide is removed upon entry into the endoplasmic reticulum. Whether this is actually the case would require N-terminal sequencing of the HA107 protein and immunolocalization studies in the cell. The size difference could also be attributed to other post-translational modifications such as glycosylation.

On the basis of the predicted protein structure (Fig. 1), HA107 was predicted to have multiple transmembrane regions. However, HA107 was only found in the nucleocapsid fraction and not in the envelope fraction of BV or ODV, as evidenced by Western blot analysis (Fig. 3b). IEM of ODV and its nucleocapsid confirmed that HA107 is associated with the nucleocapsid and not with the envelope of intact virions (Fig. 3c). Whether HA107 is also a transmembrane component needs to be studied further.

Temporal transcription and expression analysis showed that *Ha107* was active predominantly at the late stage in the infection cycle. This is in agreement with the observation that HA107 is a structural component of BV and ODV virions, which assemble relatively late after infection (Fig. 3a). Mapping of *Ha107* transcripts by 5'RACE showed that *Ha107* transcription was initiated at two distinct locations: one located at 33 nt upstream of the ATG of *Ha107*, within a sequence without a clear baculovirus consensus transcriptional initiation motif, and the other located 11 nt upstream of the ATG of *Ha106* (*sod*). This suggests the presence of two types of transcript, a short transcript of *Ha107* and a long transcript encompassing *Ha106* and *Ha107*, co-terminating at the same 3' end. This observation was also supported by the

absence of putative transcriptional termination signal sequences downstream of the *Ha106* gene (Fig. 2e). This 3' co-terminal transcript was identified here for the first time in HearNPV, but is common in group I NPVs such as AcMNPV (Friesen & Miller, 1985). The repair virus HaBac-r*Ha107*-eGFP-PH possessed similar replication characteristics (Fig. 5b) and LD₅₀ values (Table 1) to the control virus HaBacHZ8-eGFP-PH, indicating that the function of *Ha107* can be rescued by reintroduction of the *Ha107* gene at a different locus in the HearNPV genome. These results also indicated that a short *Ha107* transcript is enough to produce a functional protein and that the presence of a long transcript encompassing *Ha106* and *Ha107* does not appear to be necessary for the function of the *Ha107* gene.

The *Ha107* knockout mutant HaBacΔ*107*-eGFP-PH was able to infect and propagate in cell culture producing polyhedra and infectious BV progeny (Fig. 5). Polyhedra were also able to infect *H. armigera* larvae. Interestingly, the BV titre of HaBacΔ*107*-eGFP-PH was much higher than the control virus and the *Ha107* repair virus (Fig. 5b), whilst the LD₅₀ of HaBacΔ*107*-eGFP-PH was significantly higher than that of the control and repair viruses. These results indicate that the HA107 protein is not essential for virus replication *in vivo* and in cell culture, but somehow directly or indirectly modulates virus infectivity.

It remains unknown why the deletion of *Ha107* results in an increased BV titre and LD₅₀ of the *Ha107* knockout virus. It is also not known whether the function of HA107 as a structural component is related to its function affecting infectivity. The existence of *Ha107* homologues in several baculoviruses (Fig. 1) suggest that the gene may have been derived from an ancestral virus, and it may still play a role in infection of viruses containing this gene.

ACKNOWLEDGEMENTS

This research was sponsored by NSFC grant no. 30630002 and 973 project no. 2003CB114202 to Z.H. and by joint grants to Z.H. and J.M.V. (Program Strategic Alliances projects 2004CB720404 and 04-PSA-BD-02 and KNAW project 03CDP012). The authors would like to thank Ms Yanfang Zhang and Mr Liang Jin for technical support and Dr Basil M. Arif for scientific editing of the manuscript.

REFERENCES

- Appel, R. D., Baroch, A. & Hochstrasser, D. F. (1994). A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server. *Trends Biochem Sci* **19**, 258–260.
- Bulach, D. M., Kumar, C. A., Zaia, A., Liang, B. & Tribe, D. E. (1999). Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences. *J Invertebr Pathol* **73**, 59–73.
- Chen, X., IJkel, W. F. J., Tarchini, R., Sun, X., Sandbrink, H., Wang, H., Peters, S., Zuidema, D., Lankhorst, R. K. & other authors (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. & other authors (2002). Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *J Gen Virol* **83**, 673–684.
- Datsenko, K. A. & Wendell, B. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640–6645.
- Friesen, P. D. & Miller, L. K. (1985). Temporal regulation of baculovirus RNA: overlapping early and late transcripts. *J Virol* **54**, 392–400.
- 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 Press.
- 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., Olszewski, J. A., Cory, J. S. & O'Reilly, D. R. (2003). The genome sequence and evolution of baculoviruses. *Annu Rev Entomol* **48**, 211–234.
- Horton, P. & Nakai, K. (1997). Better prediction of protein cellular localization sites with the k nearest neighbors classifier. *Proc Int Conf Intell Syst Mol Biol* **5**, 147–152.
- Hou, S., Chen, X., Wang, H., Tao, M. & Hu, Z. (2002). Efficient method to generate homologous recombinant baculovirus genomes in *E. coli*. *Biotechniques* **32**, 783–788.
- IJkel, W. F., van Strien, E. A., Heldens, J. G., 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.
- Jakubowska, A. K., Peters, S. A., Ziemnicka, J., Vlak, J. M. & van Oers, M. M. (2006). Genome sequence of an enhancin gene-rich nucleopolyhedrovirus (NPV) from *Agrotis segetum*: collinearity with *Spodoptera exigua* multiple NPV. *J Gen Virol* **87**, 537–551.
- Jensen, L. J., Gupta, R., Staerfeldt, H. H. & Brunak, S. (2003). Prediction of human protein function according to gene ontology categories. *Bioinformatics* **19**, 635–642.
- Long, G., Chen, X., Peters, D., Vlak, J. M. & Hu, Z. (2003). Open reading frame 122 of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus encodes a novel structural protein of occlusion-derived virions. *J Gen Virol* **84**, 115–121.
- McIntosh, A. H., Christian, P. D. & Grasela, J. J. (1999). The establishment of heliothine cell lines and their susceptibility to two baculovirus. *In Vitro Cell Dev Biol* **35**, 94–97.
- Pang, Y., Yu, J., Wang, L., Hu, X., Bao, W., Li, G., Chen, C., Han, H., Hu, S. & Yang, H. (2001). Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. *Virology* **287**, 391–404.
- Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**, 8794–8797.
- Snedecor, G. W. & Cochran, W. G. (1989). *Statistical Methods*, 8th edn. Ames: Iowa State University Press.
- Sun, X., Zhang, G., Zhang, Z., 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.
- Theilmann, D. A., Blissard, G. W., Bonning, B., Jehle, J. A., O'Reilly, D. R., Rohrmann, G. F., Thiem, S. & Vlak, J. M. (2004). *Baculoviridae*. In *VIIIth Report of the International Committee on Taxonomy of Viruses*, pp. 177–185. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. San Diego: Academic Press.
- Tsai, J. M., Wang, H. C., Leu, J. H., Wang, A. H., Zhuang, Y., Walker, P. J., Kou, G. H. & Lo, C. F. (2006). Identification of the nucleocapsid, tegument, and envelope proteins of the shrimp white spot syndrome virus virion. *J Virol* **80**, 3021–3029.
- Volkman, L. E. & Summers, M. D. (1977). *Autographa californica* nuclear polyhedrosis virus: comparative infectivity of the occluded, alkali-liberated, and nonoccluded forms. *J Invertebr Pathol* **30**, 102–103.
- Wang, H., Deng, F., Pijlman, G. P., Chen, X., Sun, X., Vlak, J. M. & Hu, Z. (2003). Cloning of biologically active genomes from a *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus isolate by using a bacterial artificial chromosome. *Virus Res* **97**, 57–63.
- 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. Y. (1989). Commercial viral insecticide-*Heliothis armigera* viral insecticide in China. *IPM Pract* **11**, 13.
- Zhang, C. X., Ma, X. C. & Guo, Z. J. (2005). Comparison of the complete genome sequence between C1 and G4 isolates of the *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus. *Virology* **333**, 190–199.