Characterization of ORF39 from *Helicoverpa armigera* Single-nucleocapsid Nucleopolyhedrovirus, the Gene Containing RNA Recognition Motif

Hai-jun Xu, Yan-He Liu, Zhang-Nv Yang and Chuan-Xi Zhang*

Institute of Insect Sciences, Zhejiang University, Kaixuan Road 268, Hangzhou 310029, Zhejiang, China

Received 19 October 2005, Accepted 20 January 2006

In the genome of *Helicoverpa armigera* single-nucleocapsid nucleopolyhedrovirus, open reading frame 39 (Ha39) is the only gene predicted to encode an RNA recognition protein. Computer analysis revealed that Ha39 homologues were found in 15 NPVs, but not in GVs. Its transcripts were detected from 3 through 72 hours post infection (h p.i.) using RT-PCR and Northern blot analysis. The protein was detected in infected-cell lysates from 6 h p.i. Western blot assay of ODV and BV preparations revealed that Ha39 encodes a structural protein associated with BVs. Additionally, immunofluorescence microscopy demonstrated that the protein was present within cytoplasm in virus-infected cells, but not in the nuclear region.

Keywords: BV associated protein, Ha39, HearNPV, RRM

Introduction

The Baculoviridae family is a large group of occluded double-stranded DNA viruses that are infectious to arthropods, mainly to insects of the order Lepidoptera. Two genera have been recognized, Nucleopolyhedrovirus (NPV) and Granulovirus (GV), and are distinguished by occlusion body morphology. NPVs have been subdivided into two distinct groups based on molecular phylogenies (Zanottto et al., 1993), and they have been designated single (S) or multiple (M), depending on the number of nucleocapsids packaged in a virion. Baculoviruses typically produce two virion phenotypes of progeny virus: occlusion derived virus (ODV) and budded virus (BV). The ODV transmits infection from insect to insect by infecting midgut columnar epithelial cells, whereas the BV is responsible for causing systemic infection within the host (Keddie et al., 1989).

Baculoviruses have been extensively studied, and the devastating effects that they can have on natural populations of insects have made them an obvious choice for use as biological agents to control pests (Entwistle and Evans, 1985). So far, total 29 baculovirus genomes have been determined, including 19 lepidopteran NPVs and 7 GVs, 2 hymenopteran NPVs and 1 dipteran baculovirus. The cotton bollworm, *Helicoverpa armigera*, is a world-wide pest causing serious damage to cotton, tobacco, pepper and tomato plants (Fitt, 1989). *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus (HearNPV), first isolated from cotton bollworm in 1975, has been successful in controlling this pest for over 25 years in China (Long et al., 2003). At present, two strains of HearNPV, HearNPV G4 (Chen et al., 2001) and HearNPV C1 (Zhang et al., 2005), have been isolated with genomes of 131.4 kb and 130.5 kb, respectively, and they share 98% identity in nucleotide sequence. Compared with genomes of other baculoviruses, twenty unique ORFs were present in HearNPV and the genotypic variant genome of *H. zea* SNPV (HzSNPV). The functions of these ORFs remain unknown (Chen et al., 2002).

Numerous genes have been characterized, including basic DNA-binding protein (*bdhp*) (Wang et al., 2001), *odv-ec43* (Fang et al., 2003), *chitinase* (Wang et al., 2004), p10 (Dong et al., 2005), *fp25k* (Wu et al., 2005), Ha29 (Guo et al., 2005) and Ha128 (An et al., 2005). One ORF, Ha33 is restricted to the Lepidopteran baculovirus, was found to be a structural protein for BV (Wang et al., 2005). Ha122, present only in HearNPV and HzSNPV, was determined to be a specific ODV protein (Long et al., 2003).

In this study, a unique ORF (Ha39) in HearNPV genome containing an RNA recognition motif was characterized. Ha39 is transcribed immediately early post infection. The virion structural location coupled with subcellular location show that Ha39 encodes a structural protein of BV, implying that this gene may be involved in virus invasion.

*To whom correspondence should be addressed.
Tel: 86-571-86971697; Fax: 86-571-86971697
E-mail: chxzhang@zju.edu.cn
Materials and Methods

Cells and viruses. HearNPV (C1 strain) virus was propagated in HzAM1 cells, maintained at 27°C in TNM-FH insect medium (Sigma) supplemented with 10% (v/v) fetal bovine serum (Gibco-BRL). The titration of virus and other routine manipulations were performed according to standard protocols (O'Reilly et al., 1992).

Computer-assisted analysis. The functional domains and motifs of Ha39 were predicted by ScanProsite online (http://au.expasy.org/prosite). The data of Ha39 homologues compared in this paper were all cited from GenBank. Sequence alignment was performed with Clustal X software (Thompson et al., 1997) and homology shading was done using GeneDoc software (Nicholas et al., 1997).

Expression of Ha39. The Ha39 coding region was amplified from the HearNPV genomic DNA (C1 isolate) by PCR, using the forward primer F (5′-GGATCCATGTCGAGAACACAGCAAC-3′) and the reverse primer R (5′-ACTCGAGTAGACAGAAGCTCGCA-3′) with a BamHI site and XhoI site (shown in bold), respectively. The Ha39 amplified sequence was subcloned into the expression vector pGEX4t-2 with glutathione S-transferase (GST) at the N-terminus. The fusion protein GST-Ha39 was expressed in E. coli and induced with 1 mM IPTG at 28°C.

Preparation of the antibody. Antibody was prepared using standard techniques (Harlow and Lane, 1988). Purified GST-Ha39 protein (2 mg) in complete Freund’s adjuvant was injected subcutaneously to immunize New Zealand white rabbits, followed by two booster injections in incomplete Freund’s adjuvant within 2 weeks before exsanguinations. The polyclonal rabbit antibody against GST-Ha39 was used for immunobassay.

Northern blot analysis. HzAM1 monolayers were infected with HearNPV C1 at a multiplicity of infection (m.o.i.) of 10. Virus was absorbed for 1 h in serum-free culture and then TNM-FH medium with 10% FBS replaced the virus solution. Total RNA from mock- or HearNPV-infected cells at various time intervals (3, 6, 12, 24, 48 and 72 h p.i.). Total RNA was purified by incubating HzAM1 cells, infected with HearNPV C1 at m.o.i. of 10. Cells were harvested at the designated time points (3, 6, 12, 24, 48 and 72 h p.i.) and washed three times with 1× PBS (0.14 M NaCl, 2.7 mM KCl, 10.1 mM NaHPO4, 1.8 mM KH2PO4, pH 7.4). The protein concentration of the cell extracts was determined by Bradford’s method (Bradford, 1976). Twenty micrograms of the cell lysates was separated on 10% SDS-PAGE and subjected to western blot assay.

ODV and BV preparation. Monolayer of HzAM1 cells was infected with HearNPV C1 at m.o.i. of 10. After three days, the supernatant was collected. The purification of BVs from the supernatant was performed as described by Long et al. (2003). Polyhedral inclusion bodies (PIBs) were isolated from infected H. armigera fifth-instar larvae and purified by sucrose gradient centrifugation as described by Summers and Smith (1975). ODVs were isolated from PIBs by centrifugation as described by Braunagel and Summers (1994) and Caballero et al. (1992).

Western blot analysis. Western blot was performed as described by Towbin et al. (1979). Following SDS-PAGE, the protein samples were transferred onto PVDF membranes (Immobilon-P, Millipore) in cold Towbin buffer. The protein concentration of the cell extracts was determined by Bradford’s method (Bradford, 1976). Twenty micrograms of the cell lysates was separated on 10% SDS-PAGE and subjected to western blot assay.

Confocal laser scanning microscope. HzAM1 monolayers were infected with HearNPV C1 at m.o.i. of 10. At 48 h p.i., cells were rinsed three times with PBS and fixed in methanol: acetone (1 : 1) for 15 min, followed by three washes in PBS. To detect the Ha39 location, cells were incubated with anti-GST-Ha39 polyclonal antibody (1 : 400 dilution) in PBS for 1 h at room temperature. After three washes with 1× PBS, the cells were then incubated with Protein G fused with enhanced green fluorescent protein (EGFP) for 2 h and the nuclear DAPI (Sigma) for 1 h. Cells were directly observed and photographed using a Zeiss LSM 510 confocal laser scanning microscope.
Characterization of *Helicoverpa armigera* SNPV Ha39

**Results**

**Sequence analysis of Ha39.** The Ha39 ORF contains 585 nucleotides and is predicted to encode a 194 amino acid peptide with a molecular weight of 22.5 kDa. A baculovirus early consensus transcriptional start motif (CAG T) and a late transcriptional start motif (TAAG) are observed at 132 nt and 23 nt upstream of the start codon, respectively. Additionally, a typical polyadenylation signal (AATAAA) is found at 15 nt downstream of the translation stop codon TGA.

Computer analysis of Ha39 amino acid sequence predicts a confident eukaryotic RNA recognition motif (RRM) at aa 76-152 (Fig. 1, outlined). No signal peptide or trans-membrane domain was observed. Ha39 is the only gene containing such a RRM motif in the HearNPV genome. Other putative functional sites were predicted, including a N-glycosylation site (NRTD) at
aa 10-13, three possible protein kinase C phosphorylation sites at aa 9-11, aa 29-31 and aa 189-191, four casein kinase II phosphorylation sites at aa 40-42, aa 44-47, aa 100-103 and aa 130-133, and a putative tyrosine kinase phosphorylation site of aa 99-107. Ha39 homologues exist in 14 baculovirus NPVs in addition to HearNPV and HzSNPV. However, these homologues have different protein lengths ranging from 192 to 415 aa, and they are homologous only in the RRM regions with identities from 28% to 43% (Fig. 1). In the RRM regions, three amino acids, Arginine, Cystine and Glutamate, are conserved in 17 NPVs (Fig. 1, shown by arrows).

**Transcription course of Ha39 in infected HzAM1 cells.** Temporal regulation of the Ha39 transcripts was examined by RT-PCR and Northern blot analysis using total RNA isolated from HearNPV-infected HzAM1 cells. Transcripts were detected at 3 h p.i., with levels increasing by 24 h p.i., and were still stable at 72 h p.i. using either RT-PCR or Northern blot analysis (Fig. 2). Thus, these data coupled with the early consensus initiation sequence (CAGT) located in the upstream of the ORF indicate that Ha39 is an early gene, with transcripts being immediately produced after virus infection.

**Temporal expression of Ha39 in infected HzAM1 cells.** To determine the time course of the Ha39 protein expression, cell lysates from mock- and virus-infected cells were sampled at designated time points and analyzed by western blot using anti-GST-Ha39 serum. A distinct band of 23 kDa was
Characterization of *Helicoverpa armigera* SNPV Ha39 detected that correlated with the predicted molecular weight of Ha39. This band could be detected at 6 h p.i., and was still present through 72 h p.i. (Fig. 3A). The protein was detected at a later time point than the transcript. The reason for this might be that the immunodetection was not as sensitive as RT-PCR or Northern blot analysis.

**Immunodetection of the Ha39 protein in BVs.** To determine whether Ha39 protein was associated with BVs, purified BVs were subjected to western blot analysis, and a predominant band was detected with immune antisera (Fig. 3B). The protein size was in line with that detected in cell lysates. Additionally, ODVs were isolated for western blot assay, and purified ODVs showed the absence of Ha39 protein (Fig. 3B). Thus, the Ha39 appeared to be a structural protein specific for BV.

**Subcellular localization of Ha39.** The subcellular localization of the Ha39 was investigated by immunofluorescence using a confocal laser scanning microscope. At 48 h p.i., the infected cells were collected for fluorescence examination. High levels of Ha39 were observed within cytoplasm and not within the nuclear region (Fig. 4). In the control samples, no fluorescence was detected when preimmune serum was used (Fig. 4).

**Discussion**

Many eukaryotic proteins that bind single-stranded RNA contain an RNA-binding domain of about 90 amino acids (Dreyfuss *et al.*, 1988; Bandziulis *et al.*, 1989), known as the RNA recognition motif (RRM). RRMs are found in a number of RNA binding proteins, including various heterogeneous nuclear recognition proteins (hnRNP), proteins implicated in regulation of alternative splicing, and protein components of small nuclear recognition proteins (snRNP) (Améro *et al.*, 1992; Mitunis *et al.*, 1993). This motif also appears in a few single stranded DNA binding proteins (Ikeda *et al.*, 1996). Analysis of the genomic sequence of HearNPV reveals a putative RNA recognition motif encoded by the Ha39 gene. In this study, we performed transcriptional analysis of the Ha39 gene, as well as the localization of the gene product in subcellular compartments and within virions. Transcriptional analyses show that Ha39 is an early gene with a slightly detectable transcript initiated at 3 h p.i. This might indicate that the early transcriptional start motif (CAG T) was used, although an early and a late transcriptional start motif were simultaneously predicted by sequence analysis. The expression profile of Ha39 indicated that this gene product could be detected as early as 6 h p.i. The size of the immunoreactive protein (23 kDa) nearly corresponded with predicted molecular weight of 22.5 kDa. This suggests that the post-translational modification events may be limited, despite the presence of many putative post-translational modification sites. Furthermore, subcellular location was performed to determine if Ha39 is a

**Fig. 2.** RT-PCR and Northern blot analysis of Ha39 transcription in mock- and HearNPV-infected HzAM1 cells. A, RT-PCR analysis of the Ha39 transcripts. RT-PCR analysis was performed using purified RNA isolated from mock- and virus infected cells at appropriate timepoints (3, 6, 12, 24, 48 and 72 h p.i.). The first strand cDNA was synthesized with AMV reverse transcriptase and oligo-p(dT)18 primer, and then nested PCR was amplified with Ha39 ORF specific primer F and R. The products were separated on 1.0% agarose gel. The β-actin RT-PCR was set as the control. B, Northern blot analysis of the Ha39 transcripts. Ten micrograms of total RNA was separated on a 0.8% agarose-formaldehyde gel, transferred onto nylon membrane, and hybridized to a DIG labeled DNA probe. The probe was prepared by PCR product which was previously amplified with primer F and primer R. The size of detected RNA is indicated on the right. The portion of the gel containing rRNA stained with ethidium bromide are shown (lower panel).

**Fig. 3.** Western blot analysis of Ha39 in HzAM1 cells and its location in virion of HearNPV. A, The cells were collected at mock, 3, 6, 12, 24, 48, 72 h p.i. and 20 µg cell lysates at each interval was subjected to western blot analysis using anti-GST-Ha39 serum. The binding was developed with diaminobenzidine (DAB) as a chromogenic substrate. Protein markers are indicated on the left. B, Western blot assay for ODVs and BVs of HearNPV. The purified ODVs and BVs were subjected to western blot analysis using anti-GST-Ha39 serum. The sizes of reactive bands are indicated with arrows.
The structural protein for ODV and BV. The immune-reactive band was only present in BVs, but not in ODVs. Interestingly, the homolog in SpliNPV (Sl39) was found to localize not only with BV but also with ODV (Wang et al., 2002). Sequence comparison revealed that the two sequences were different in length and motifs. Sl39 is 302 aa in length with two additional motifs, a baculovirus J domain and a coiled-coil motif in the N-terminal region. The differences in sequences and locations suggest discriminating events between them. Additionally, we used laser confocal microscopy to visualize the distribution of Ha39 in HearNPV-infected cells. Immunofluorescence clearly demonstrates that this gene product is not present in the nucleus. This indirectly supports the evidence that Ha39 is specific for BV, as the nucleocapsids of BV budded through the plasma membrane for further infection. This localization is similar to another BV-specific protein of Bombyx mori NPV GP64 (BmGP64) in the cytosol and cytoplasmic membrane (Rahman and Gopinathan, 2003). However it is difficult to determine if Ha39 was localized in cytoplasmic membrane. So far, little is known about the RNA recognition specificity of Ha39. Whether the role of Ha39 in RNA regulation depends on its RNA binding specificity is remained to be determined.

Acknowledgments This project was supported by the Foundation for the Author of National Excellent Doctoral Dissertation of P. R. China (Project No. 200055), the National Natural Science Foundation of China (Project No. 3070032, 30770031, 30570074) and the National Basic Research Program of China (Project No. 2003CB114400).

References


Entwistle, P. F. and Evans, H. F. (1985) Insect Control;