Open reading frame 94 of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus encodes a novel conserved occlusion-derived virion protein, ODV-EC43

Minggang Fang,1 Hanzhong Wang,1 Hualin Wang,1 Li Yuan,1 Xinwen Chen,1 Just M. Vlak2 and Zhihong Hu1

**Correspondence**
Zhihong Hu
huzh@pentium.whiov.ac.cn
1Join-Laboratory of Invertebrate Virology and Key Laboratory of Molecular Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, People’s Republic of China
2Laboratory of Virology, Wageningen University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands

Received 12 April 2003
Accepted 11 July 2003

Open reading frame 94 (Ha94) of *Helicoverpa armigera* single nucleocapsid nucleopolyhedrovirus (HaSNPV) is 1086 bp long and a homologue of *Autographa californica* multiple NPV ORF109. The gene is conserved among all baculoviruses whose genomes have been completely sequenced so far and is thus considered a baculovirus core gene. Ha94 transcripts were detected from 24 to 96 h post-infection (p.i.) of HzAM1 cells with HaSNPV. Polyclonal antiserum raised to a GST–HA94 fusion protein recognized a 43 kDa protein, HA94, in infected cell lysates from 36 to 96 h p.i., suggesting that Ha94 is a late gene. Western blot analysis of proteins present in budded virus and occlusion-derived virus (ODV) showed that Ha94 encodes a structural component of ODV. When ODVs were fractionated further into nucleocapsid and envelope components, Western blot analysis indicated that the encoded protein was associated with both the nucleocapsid and the envelope. In summary, data available indicated that Ha94 encodes a novel ODV-specific protein of HaSNPV, designated ODV-EC43.

**INTRODUCTION**

The Baculoviridae, a diverse family of more than 600 viruses, encompasses two genera, the Nucleopolyhedroviruses (NPVs) and the Granuloviruses (GVs) (Blissard et al., 2000). These viruses are very specific and mainly infect insects of the orders Lepidoptera, Hymenoptera and Diptera, but always within the phylum Arthropoda. The family is characterized by the occlusion of virions into large proteinaceous capsules or occlusion bodies (OBs). NPVs have two morphotypes, SNPV and MNPV, depending on the single (S) or multiple (M) packaging of the nucleocapsids (NCs) into the envelope (E). Based on phylogenetic analysis, NPVs have been divided into groups I and II (Zanotto et al., 1993; Herniou et al., 2001). In a single infection cycle, two forms of virions are produced, the budded virus (BV) and the occlusion-derived virus (ODV). BVs are responsible for the spread of the virus from cell to cell in the larva host and ODVs initiate the infection in a susceptible host and are responsible for the horizontal spread of the virus in an insect population (Blissard et al., 2000). The structural components of BVs and ODVs, except for the circular double-stranded DNA, are different to accommodate their respective functions in an infection cycle (Braunagel & Summers, 1994).

The molecular biology of baculoviruses has developed rapidly in the last two decades. To date, the complete sequences of 17 baculovirus genomes have been reported; they range in size from 101 to 179 kbp and are predicted to encode 109 to 181 open reading frames (ORFs). These viruses are *Autographa californica* MNPV (AcMNPV) (Ayres et al., 1994), *Bombyx mori* NPV (BmNPV) (Gomi et al., 1999), *Culex nigripalpus* NPV (CuniNPV) (Afonso et al., 2001), *Epiphyas postvittana* NPV (CuniNPV) (Afonso et al., 2001), *Epiphyas postvittana* NPV (EpNPV) (Hyink et al., 2002), *Helicoverpa armigera* SNPV (HaSNPV) (Chen et al., 2001), *Helicoverpa zea* SNPV (HzSNPV) (Chen et al., 2002), *Lymnantria dispar* MNPV (LdMNPV) (Kuzio et al., 1999), *Mamestra configurata* NPV A (MacoNPV A) (Li et al., 2002b), MacoNPV B (Li et al., 2002a), *Orgyia pseudotsugata* MNPV (OpMNPV) (Ahrrens et al., 1997), *Rachiplusia ou* NPV (RaouNPV) (Harrison & Bonning, 2003), *Spodoptera exigua* MNPV (SeMNPV) (IJkel et al., 1999), *Spodoptera litura* NPV (SpltNPV) (Pang et al., 2001), *Cydia pomonella* GV (CpGV) (Luque et al., 2001), *Phthorimaea operculella* GV (PhopGV) (GenBank accession no. NC_004062), *Plutella xylostella* GV (PxGV) (Hashimoto...
et al., 2000) and Xestia c-nigrum GV (XcGV) (Hayakawa et al., 1999). Genome comparisons of all of the baculoviruses completely sequenced so far revealed 30 conserved genes (Table 1), most of which are related either to DNA replication/gene expression or to virion structure (Herniou et al., 2003). However, there are still several genes whose functions remain unknown, including Ac109.

Helicoverpa armigera SNPV (HearNPV, also called HaSNPV) was isolated initially from diseased larvae in the province of Hubei. It has been developed as a commercial pesticide and has been used successfully to control cotton bollworm in China (Zhang, 1994). The genome of HaSNPV contains 135 ORFs of 50 aa or more (Chen et al., 1997a), and its gene product is an ODV protein. Ha122 is unique to HaSNPV and its gene product is an ODV protein.

HaSNPV ORF94 (Ha94), a homologue of Ac109, is one of the conserved or core genes among baculoviruses (Herniou et al., 2003), although its function remains elusive. By transcriptional analysis, protein identification and localization, we present evidence that HA94 is a newly assigned structural protein of baculoviruses.

### Table 1. Genes conserved among members of the Baculoviridae

The conserved genes were the ORFs that exist in all 17 baculoviral genomes sequenced so far. The sources were: Ayres et al. (1994) for AcMNPV; Gomi et al. (1999) for BmNPV; Afonso et al. (2001) for CamiNPV; Hyink et al. (2002) for EppoNPV; Chen et al. (2001) for HaSNPV; Chen et al. (2002) for HzSNPV; Kuzio et al. (1999) for LdMNPV; Li et al. (2002b) for MacoNPV A; Li et al. (2002a) for MacoNPV B; Ahrens et al. (1997) for OpMNPV; Harrison & Bonning (2003) for RaouNPV; IJkel et al. (1999) for SeMNPV; Pang et al. (2001) for SpltNPV; Luque et al. (2001) for CpGV; GenBank accession no. NC_004062 for PshG; Hashimoto et al. (2000) for PxGV; and Hayakawa et al. (1999) for XcGV.

<table>
<thead>
<tr>
<th>Expression/replication-related genes</th>
<th>Structural genes</th>
<th>Others</th>
<th>Unknown genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>lef-1 (Ac14)</td>
<td>pif-2 (Ac22)</td>
<td>alk-exo(Ac133)</td>
<td>Ac68</td>
</tr>
<tr>
<td>lef-2 (Ac6)</td>
<td>ld130 (Ac23)</td>
<td></td>
<td>Ac81</td>
</tr>
<tr>
<td>lef-4 (Ac90)</td>
<td>vp1054 (Ac54)</td>
<td></td>
<td>Ac92</td>
</tr>
<tr>
<td>lef-5 (Ac99)</td>
<td>gp41 (Ac80)</td>
<td></td>
<td>Ac96</td>
</tr>
<tr>
<td>lef-8 (Ac50)</td>
<td>vp91 (Ac83)</td>
<td></td>
<td>Ac98</td>
</tr>
<tr>
<td>lef-9 (Ac62)</td>
<td>vp39 (Ac89)</td>
<td></td>
<td>Ac109 (Ha94)</td>
</tr>
<tr>
<td>dna-pol (Ac65)</td>
<td>p6.9 (Ac100)</td>
<td></td>
<td>Ac115</td>
</tr>
<tr>
<td>helicase (Ac95)</td>
<td>pif-1 (Ac119)</td>
<td></td>
<td>Ac142</td>
</tr>
<tr>
<td>p47 (Ac40)</td>
<td>p74 (Ac138)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vlf-1 (Ac77)</td>
<td>odr-cc27 (Ac144)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>odr-e56 (Ac148)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Expression of Ha94 in Escherichia coli and generation of an HA94-specific antibody. Two primers, 94up and 94down (5'-CTCGAGGCTAAATACCTGTA-3') were designed to amplify the entire Ha94 ORF from the HaSNPV-G4 genome. The PCR product was cloned first into pGEM-T-Easy (Promega) and then into the expression vector pGEX-KG (Guan & Dixon, 1991) as an EcoRI-Xhol fragment. This generated plasmid pGEX-Ha94, in which Ha94 is in-frame and fused with GST at the C terminus. E. coli DH5α cells containing pGEX-Ha94 were grown to an OD₆₀₀ of 0.4 and then induced with 1 mM IPTG. After 3 h at 37 °C, cells were harvested and lysed with lysozyme, sonicated and centrifuged at 5000 g for 10 min at 4 °C. The fusion protein present in the pellet was separated in 12 % SDS-polyacrylamide gels and purified. Antisera were generated by immunizing rabbits with purified protein (Sambrook et al., 1989) and tested by Western blot analysis.

**Western blot analysis.** HaAM1 cells were infected with HaSNPV-G4 at an m.o.i. of 5 p.f.u. per cell. Samples of total cell proteins were prepared from infected cells harvested at 0, 12, 16, 24, 36, 48, 72 and 96 h p.i. The BVs and ODVs of HaSNPV were purified and the E and NC fractions of ODVs were separated according to Ikel et al. (2001). Protein samples were separated by 12 % SDS-PAGE and the separated proteins transferred onto Hybond-N membranes (Amersham) by semi-dry electrophoresis transfer (Ausubel et al., 1987). HA94-specific antiserum and alkaline phosphatase-conjugated goat anti-rabbit immunoglobulin (Sino-American) were used as the primary and secondary antibodies, respectively. The signal was detected using a BCIP/NBT kit (Sino-American).

**RESULTS**

**Sequence analysis of Ha94 and its homologues**

Ha94 is located in the HindIII-A fragment of the HaSNPV genome (nt 87,300–88,385) (Chen et al., 2001). Sequence analysis indicated that Ha94 is 1086 bp and encodes a protein of 362 aa with a predicted molecular mass of 41.5 kDa. A late baculoviral transcription initiation motif, ATAAAG, was found 65 nt upstream of the putative translational start site of Ha94, suggesting that Ha94 may be a 'late' gene of HaSNPV. A polyadenylation signal (AATAAA) was identified 14 nt downstream of the TAA stop codon.

Searches of databases showed that Ha94 is conserved among baculoviruses and is shared by all baculoviruses whose complete genome sequences have been published so far. The homologues of Ha94 are AcMNPV ORF109, BmNPV ORF92, CuniNPV ORF69, EppoNPV ORF95, HsSNPV ORF97, LdMNPV ORF107, MacoNPV A ORF79, MacoNPV B ORF79, OpMNPV ORF109, RaounPV ORF104, SeMNPV ORF59, SpltNPV ORF96, CpGV ORF55, PhopGV ORF50, PxGV ORF43, and XcGV ORF53. A homologue was also reported in another baculovirus, Leucania separata NPV (LeseNPV), as LS109. Analysis indicated that, apart from HzSNPV, Ha94 shared amino acid sequence identities ranging from 58 % with LdMNPV ORF107 to 10 % with CuniNPV ORF69. Ha94 is related most closely to HsSNPV ORF97, with 99 % identity, consistent with the previous suggestion that HsSNPV and HaSNPV are different strains of the same virus (Chen et al., 2002). The very low identity of CuniNPV ORF69 with all other baculoviral HA94 homologues (data not shown) indicates that CuniNPV, which was isolated from a dipteran host, is distantly related to baculoviruses isolated from the Lepidoptera.

No signal peptide sequence was predicted on the putative HA94 protein (Fig. 1) by PSORT and SIGNALP and no transmembrane region was identified by TMpred at the EXPASy server (Appel et al., 1994). HA94 contained three potential N-linked glycosylation consensus sequences (N-X-S/T), but we have not investigated whether these sites are glycosylated. Only one of these, at amino acid position 80, was predicted by the EXPASy server to be glycosylated. Alignment of the putative HA94 peptide with its homologues in other baculoviruses indicated that 13 aa were absolutely conserved (Fig. 1). These conserved amino acids may be important for the function of the HA94 homologues.

**Transcription of Ha94**

Transcription of Ha94 was examined by RT-PCR, using total RNA isolated from HzAM1 infected with HaSNPV at different time points as template and primers 94up primer and oligo(dT). A single band of 1.2 kbp was first detected at 24 h p.i. and continued to be present until 96 h p.i., indicating that Ha94 is a late gene (Fig. 2A). The predicted size from 94up to the poly(A) signal site AATAAA is 1114 nt, which is in agreement with the size of the transcript detected.

**Immunodetection of HA94 in infected cells and virus particles**

To obtain a polyclonal rabbit antibody against this protein, HA94 was expressed in E. coli as a GST–HA94 fusion protein (data not shown) and purified by gel separation and electroelution. Specific antisera was generated by immunizing rabbits with purified protein. When HaSNPV-infected HzAM1 cell extracts were analysed by SDS-PAGE and Western blot, the HA94 antibody detected a specific protein of molecular mass 43 kDa (Fig. 2B). This protein was first detectable at 36 h p.i. and continued to accumulate until 96 h p.i. This suggested, in agreement with the analysis of transcripts, that HA94 was synthesized at a late stage of infection, possibly as part of OB morphogenesis. The size of the 43 kDa protein is in agreement with the predicted 41.5 kDa based on its sequence, which suggests that no major post-translational modification occurs.

To investigate if HA94 is a structural component of HaSNPV virions, Western blot analysis of BV and ODV proteins was conducted. HA94 was detected in preparations of ODVs but not BVs (Fig. 3A), indicating that HA94 is a structural component of ODVs. The location of HA94 in ODVs was determined further by Western blot analysis of the NC and E fractions of ODVs. The ODV NC and E fractions were analysed by SDS-PAGE and the purity of nucleocapsids was checked by electron microscopy. The protein profiles of the NC and E were distinct (Fig. 3B). Western blot analysis indicated that HA94 was located...
mostly in the NC fraction and a minor part of this protein was located in the E fraction of ODVs (Fig. 3C).

**DISCUSSION**

The conservation of HA94 in baculoviruses may imply an important function of this protein in baculovirus multiplication. To elucidate its function, we analysed the transcription and translation of Ha94 in HaSNPV-infected *H. armigera* cells and investigated whether it was a structural component of the virions. Transcriptional analysis of Ha94 by RT-PCR showed that Ha94 transcription started at 24 h p.i. and continued until at least 96 h p.i. (Fig. 2A). This result suggests that Ha94 is a late gene and probably uses the late transcription initiation ATAAG at 65 nt upstream of the translational start site. Western blot analysis confirmed this observation, as HA94 was detected predominantly from 36 to 96 h p.i., relatively late in the infection cycle (Fig. 2B). This is compatible with the observation that the 43 kDa protein is present in ODVs but not in BVs (Fig. 3A). Western blot analysis revealed further that HA94 was an ODV-specific structural protein and is present in both the NC and E fractions of ODVs (Fig. 3C). Therefore, HA94 is a novel structural ODV protein, which we have named ODV-EC43. The conservation of ODV-EC43 in baculoviruses suggests that the protein plays an important role in ODV morphogenesis and/or ODV structure, or in the infection process.

Considerable progress has been achieved in the identification of baculovirus genes that are likely to encode virion structural proteins. At the start of our research, nine structural virion proteins had been identified as conserved in all baculoviruses. These include three proteins that exist in both BVs and ODVs: the basic DNA-binding protein p6.9 (Wilson et al., 1987), the major capsid protein VP39 (Blissard et al., 1989) and VP1054 (Olszewski & Miller, 1997). One protein exists in BVs: LD130 (Pearson et al., 2000). In addition, five proteins that are specific for ODVs have been identified: P74 (Kuzio et al., 1989), GP41 (Whitford & Faulkner, 1992a, b), ODV-EC27 (Braunagel et al., 1996b), ODV-E56 (Braunagel et al., 1996a) and VP91 (Russell & Rohrmann, 1997). Recently, two novel structural virion proteins that are involved in oral...
infectivity, PIF-1 (Kikhno et al., 2002) and PIF-2 (Pijlman et al., 2003), have been identified, although their exact location in the ODV has not been established yet. The identification of ODV-EC43 brings the growing number of conserved structural proteins to 12. These 12 proteins are likely to form the core structure of the baculovirus virions or may be important for infection. Deletion of Ha94 from the HaSNPV genome is under way, to allow further functional analysis of Ha94 in terms of its interaction with other structural proteins and its role in virus infection.

ACKNOWLEDGEMENTS

This research was supported in part by grants from the National Natural Science Foundation of China (NSFC) (grant nos 30025003 and 30070034), the Chinese Academy of Sciences (CAS) (grant nos Kscx2-1-02 and Kscx2-SW-301-09), an 863 project of China (grant no. 2001AA214031) and a joint grant from the CAS and the Royal Netherlands Academy of Sciences (grant no. 01CDP023).

REFERENCES


