NUCLEOCAPSID PROTEMN GENE OF CHINESE ISOLATE OF BOMBYX MORI NUCLEOPOLYHEDROVIRUS

D. LIU[1,2], Y. QI[1*], X. SUN[2], Z. LIU[1], S. HE[1]

[1]Institute of Virology, Wuhan University, Wuhan 430072, P.R. China;
[2]College of Life Sciences, Central China Normal University, Wuhan, P.R. China

Received November 11, 1999; accepted June 28, 2000

Summary. – The nucleocapsid protein gene (*vp39*) of a Chinese isolate of Bombyx mori nucleopolyhedrovirus (BmNPV-Ch), namely an open reading frame (ORF) of 1050 bp that codes for a polypeptide of 39 K (VP39) consisting of 350 amino acids was sequenced. The homology of the nucleotide (nt) and amino acid sequences of *vp39* and VP39, respectively, of BmNPV-Ch and a Japanese isolate of BmNPV (BmNPV-Ja) were found to be 97.5% and 97.1%, respectively. The BmNPV-Ch *vp39* is nine nucleotides longer than that of BmNPV-Ja *vp39* due to insertion of CGA at nt 625 and GTCGGC at nt 985–910. There are differences in 17 nucleotides causing a few substitutions of amino acids which slightly modify the secondary structure of BmNPV-Ch. It indicates that the main part of the secondary structure of VP39 is a folded structure containing high proportion of β-sheet and β-turn units. A dot blot hybridization analysis revealed the existence of a homologous transcript of BmNPV-Ch *vp39* in Sf9 cells infected with Autographa californica multiple nucleocapsid nucleopolyhedrovirus (AcMNPV).

Key words: nucleocapsid protein; VP39; gene; BmNPV; Chinese isolate; Japanese isolate; secondary structure

Introduction

The molecular biology of baculoviruses is one of the most dynamic research fields. Great progress has been made on AcMNPV, a prototype baculovirus. Ayers *et al.* (1994) published the sequence of AcMNPV DNA genome (134 kbp). More than 50 important genes including the nucleocapsid protein gene *vp39* have been identified. Blissard *et al.* (1989) studied the structure and transient expression of *vp39* of Orgyia pseudosugata nucleopolyhedrovirus (OpMNPV). Recent studies have found that the *vp39* gene product, the VP39 protein of AcMNPV has an important role in the infection (Chariton *et al.*, 1993; Slack *et al.*, 1995). VP39 can bind to the actin of the host, which leads to a change of cytoskeleton and its rearrangement to a cable. It enables transport of the nucleocapsid to the cell nucleus for the package of progeny virions (Chariton *et al.*, 1993; Slack *et al.*, 1995). Therefore, *vp39* is a very important baculosvirus gene that should be studied intensively.

BmNPV is one of the baculoviruses that has been studied intensively employing advantage of use of the silkworm expression system. Maeda *et al.* (1996) have sequenced the BmNPV-Ja genome. Another molecular biological study on BmNPV was reported by Deng *et al.* (1999).

In this study, we determined the sequence of BmNPV-Ch *vp39* and compared it with that of BmNPV-Ja. It showed that their general characteristics are similar but their detailed

*Corresponding author. E-mail: mvlwu@edu.cn; fax: 08627-87682938.

Abbreviations: AcMNPV = Autographa californica multiple nucleocapsid nucleopolyhedrovirus; BmNPV = Bombyx mori nucleopolyhedrovirus; BmNPV-Ch = Chinese isolate of BmNPV; BmNPV-Ja = Japanese isolate of BmNPV; OpMNPV = Orgyia pseudosugata nucleopolyhedrovirus; ORF = open reading frame; SDS-PAGE = polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate

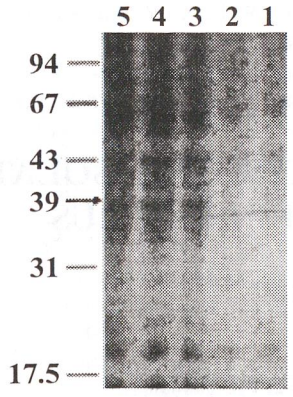

Fig. 1
SDS-PAGE of VP39 protein expressed in *E. coli*
E. coli non-transformed with the vector (control) (lane 1); *E. coli* transformed with pRSET-Bm39 vector (lane 2); *E. coli* transformed with pRSET-Bm39 vector, 2 hrs after induction with IPTG (lane 3); *E. coli* transformed with pRSET-Bm39 vector, 4 hrs after induction with IPTG (lane 4); *E. coli* transformed with pRSET-Bm39 vector, 6 hrs after induction with IPTG (lane 5). Size markers (K values) are shown on the left margin.

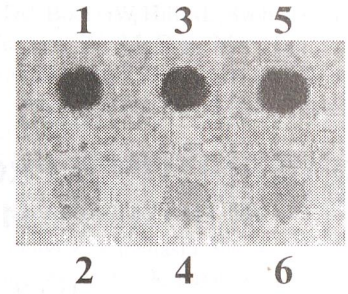

Fig. 2
Dot blot hybridization analysis of RNA with BmNPV-Ch *vp39* probe
RNA from BmNcells infected with BmNPV-Ch (lane 1); RNA from non-infected BmN cells (lane 2); RNA from Sf9 cells infected with AcMNPV (lane 3); RNA from non-infected Sf9 cells (lane 4); pGEM-Bm39 DNA (lane 5); pGEM3zf DNA (lane 6).

sequences and secondary structure are different. These results could be used in the study of structure and function of the BmNPV *vp39* gene and serve as basis for exploring the relationship between the virus and host.

Materials and Methods

Viruses, vectors, host cells, enzymes and reagents. Restriction endonucleases, T4 DNA ligase, Taq DNA polymerase, the BcaBEST™ dideoxy sequencing and silver staining kit were purchased from Promega. BmNPV-Ch and the silkworm BmN cells were provided by Dr. D. Xiaozhao, Institute of Medicine, Nanjing, P.R. China. The expression vector pRSET-A was purchased from Invitrogen. *E. coli* TG1 and BL-21 DE-3 strains were maintained in our laboratory.

Polymerase chain reaction (PCR) and molecular cloning. BmN cells were infected with BmNPV. When the polyhedrosis was apparent in about 50% of cells, they were harvested and the virus DNA was extracted. PCR primers were designed according to the AcMNPV *vp39* gene (the upstream primer P1: 5'-GCCGGATCCAACAATATGGCGCTAGT-3' and the downstream primer P2: 5'-CCGGAGCTCGTTTTCTAATCTTGCGT-3'). The *vp39* gene (1050 bp) was amplified from BmNPV-Ch DNA, cloned into the pGEM-3zf vector at *Bam*HI-*Sac*I sites, and subcloned into the expression vector pRSET-A at *Bam*HI-*Eco*RI sites. In this way the recombinant pRSET-Bm39 vector was constructed.

Expression of VP39 protein and polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE). Recombinant pRSET-Bm39 was transformed into *E. coli* BL-21 DE-3 strain. After 4 hrs of incubation in a shaking incubator, cells were induced with IPTG for 2, 4, and 6 hrs. The expressed target product was identified by SDS-PAGE.

Dot blot hybridization analysis and nucleotide sequencing. The BmNPV *vp39* gene was labeled as the probe with a DIG-labeling kit and hybridized with RNA isolated from insect cells infected with the virus as described earlier (Liu *et al.*, 1999). The nucleotide sequence of the *vp39* gene cloned in pGEM-Bm39 was analyzed by use of a kit from GendaTech Co., Canada. The nucleotide sequence of BmNPV-Ch *vp39* was submitted to the GenBank database (Acc. No. AF063104). DNASIS and PROSIS programs were used to analyze the homology of the *vp39* nucleotides and deduced VP39 amino acids. The hydrophilicity and hydrophobicity of VP39 proteins were analyzed with PROSIS program.

Results

Expression and transcription of BmNPV-Ch vp39 gene

The 1.3 kbp fragment of BmNPV-Ch *vp39* gene amplified by PCR was cloned into a cloning vector, pGEM-3zf (+) and then recloned into an expression vector pRSET-A. The obtained new construct, pRSET-Bm39 was used for transformation of *E. coli* BL21 DE-3 strain in which the *vp39* gene was expressed under the control of T7 promoter. SDS-PAGE showed that the VP39 protein of 39 K was produced (Fig. 1).

The BmNPV-Ch *vp39* gene was used as probe to confirm its functional homology with the *VP39* genes of BmNPV-Ja and AcMNPV. The results of the dot blot hybridization analysis of *vp39* transcripts (Fig. 2) indicated that the BmNPV-Ch *vp39* gene (fragment) is a functional homologue of the AcMNPV *vp39* gene (fragment) that can be transcribed to produce the mRNA of nuclear capsid protein of BmNPV-Ch in BmN cells.

```
          M A L V P V G M A P R Q M R V N R C I F  20
BMVP39-C
ATGGCGCTAGTGCCCGTGGGTATGGCGCCGCGACAAATGAGAGTTAACCGCTGCATTTTC  60
BMVP39-J  .....A.....A.....  60
          S S I V S F D A C I T Y K S P C S P D A  40
BMVP39-C
GCGTCCATCGTGTCGTTCGACGCGTGCATAACATACAAGTCGCCGTGTTCGCCCGACGCG 120
BMVP39-J  .....A.....  120
          Y H D D G W F I C N S H L I K R F K M S  60
BMVP39-C
TATCATGACGATGGATGGTTTATCTGCAACAGCCACCTCATAAAACGTTTTAAAATGTCA 180
BMVP39-J  .....C.....C.....  180
          K M V L P I F D E D D N Q F K M T I A R  80
BMVP39-C
AAAATGGTTTTGCCCATTTTCGACGAAGACGACAATCAATTCAAAATGACGATCGCTAGG 240
BMVP39-J  .....  240
          H L V G N K E R G I K R I L I P S A A N 100
BMVP39-C
CATTTAGTTGGAAATAAAGAAAGGGGTATCAAGCGAATTTTAATTCCAAGCGCAGCCAAT 300
BMVP39-J  .....A.....A.....  300
          Y Q E V F N L N S M M Q A E Q L I F H L 120
BMVP39-C
TACCAAGAGGTGTTTAATCTAAACAGTATGATGCAAGCCGAACAGCTAATCTTTCATTTG 360
BMVP39-J  .....A.....  360
          I Y N N E A A V N V I C D N L K Y T E G 140
BMVP39-C
ATATATAACAACGAAGCGGCGGTTAACGTTATATGCGACAATCTAAAATATACCGAAGGT 420
BMVP39-J  .....  420
          F T S G T Q R V I H S V Y A T T R S I L 160
BMVP39-C
TTCACAAGCGGCACGCAACGCGTTATACACAGCGTTTACGCAACTACAAGAAGCATCCTA 480
BMVP39-J  .....A.....T...  480
          D T T N P N T F C S R V S R D E L R F F 180
BMVP39-C
GACACCACAAACCCGAACACGTTTTGTTCGCGGGTGTCGCGCGACGAATTGCGTTTTTTT 540
BMVP39-J  .....T.....C  540
          D V T N A R T G R G G V G D Q L F N N Y 200
BMVP39-C
GACGTGACCAACGCCCGAACGGGTCGAGGTGGTGTTGGCGATCAATTATTTAACAATTAC 600
BMVP39-J  .....C.....  600
          S G F L Q N L I R R A V A P E Y L Q I D 220
BMVP39-C
AGTGGATTTTTGCAAAATTTGATTCGACGCGCAGTAGCGCCCGAGTACTTGCAAATCGAC 660
BMVP39-J  .....—.....  657
          T E E L R F R N S A T C I I D E T G L V 240
BMVP39-C
ACGGAGGAATTGAGATTTAGAAATAGCGCCACGTGTATAATTGACGAAACGGGCCTGGTG 720
BMVP39-J  .....T.....  717
          A S V P D G P E L Y N P I R S S D I M K 260
BMVP39-C
GCGTCTGTGCCCGACGGCCCCGAGTTGTACAACCCGATAAGAAGCAGTGACATCATGAAA 780
BMVP39-J  .....  777
          S Q P N R L Q I R N V L K F E G D T R E 280
BMVP39-C
AGTCAACCCAATCGTTTGCAAATTAGAAACGTTTTGAAATTTGAAGGCGACACACGTGAG 840
BMVP39-J  .....  837
          L D R T L S G Y E E Y P T Y V P L F L G 300
BMVP39-C
CTGGACAGAACGCTTAGCGGATACGAAGAATACCCGACGTACGTTCCGCTGTTTTTGGGA 900
BMVP39-J  .....G.....  897
          Y Q I I N S E N N F L R N D F I S R A N 320
BMVP39-C
TACCAAATAATTAATTCAGAAAACAACTTTTTGCGAAACGACTTTATATCAAGAGCAAAT 960
BMVP39-J  .....  957
          P N A T L G G G V G A L A G P A P G V V 340
BMVP39-C
CCGAACGCTACTTTGGGCGGCGGCGTCGGCGCACTGGCAGGTCCTGCGCCTGGTGTTGTT 1020
BMVP39-J  .....——T.....  1011
          L G E A S G S V A A  351
BMVP39-C
CTCGGCGAAGCAAGTGGAAGTGTAGCCGCCTAAAAATGGAGTTTGTCAAATTGCAATGCA 1080
BMVP39-J  .....  1044
BMVP39-C
ACATTTGTTTTTCGGTTGCAGAAATTAAAAATTATTTTATGCAACCAATAGACAGATTGA 1140
BMVP39-C
CTATGATACCAGTATTAGAACTAGACACGTGCAAACATCAATTATGCTCAATGTGCATAC 1200
BMVP39-C  GCAAGATTAGAAACGAGCTC 1220
```

Fig. 3

Comparison of sequences of nucleocapsid genes and proteins of BmNPV-Ch and BmNPV-Ja

Top line represents the deduced amino acid sequence of BmNPV-Ch VP39 protein. The nucleotide sequence of BmNPV-Ch *vp39* gene (BMVP39-C, second line). The nucleotide sequence of BmNPV-Ja *vp39* gene (the GenBank Database Acc. No. L33180) (BMVP39-J, third line). Only changes and deletions (broken lines) are shown. The poly-A signal sequence is underlined.

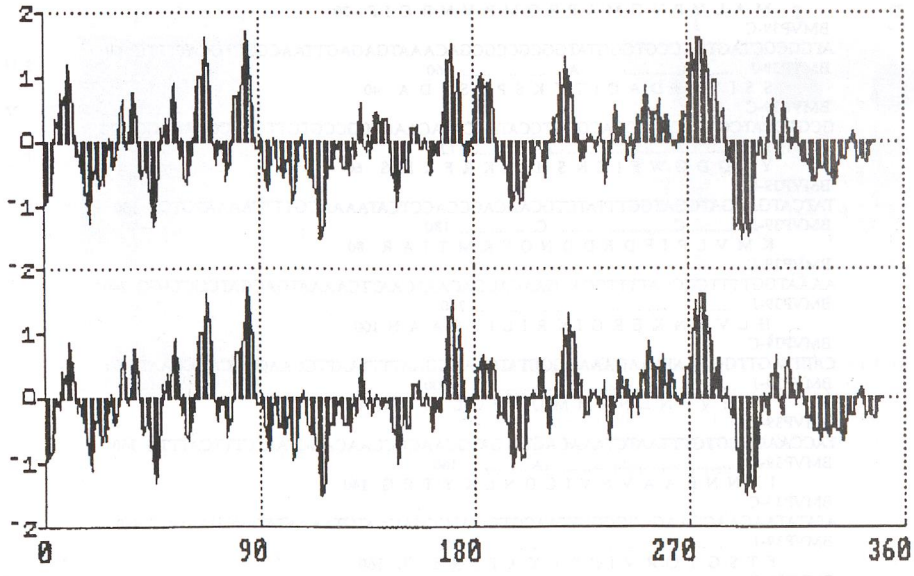

Fig. 4
The hydrophilicity and hydrophobicity of BmNPV-Ch and BmNPV-Ja VP39 proteins
The curves of BmNPV-Ch VP39 protein (upper part) and BmNPV-Ja VP39 protein (lower part).

Sequencing of BmNPV-Ch vp39 gene

The BmNPV-Ch nucleotide sequence of *vp39* gene (1230 bp) was determined. The ORF contains 1050 bp and codes for 350 amino acids (Fig. 3). The homology of BmNPV-Ch and BmNPV-Ja nucleotide sequences of respective *vp39* genes was very high (about 97.5%) and the BmNPV-Ch *vp39* gene was longer by 9 bp than that of BmNPV-Ja due to insertion of CGA added at nt 625 and GTCGGC at nt 985. The above two insertions indicate that the coding frame was maintained. In addition, there are 17 more differences in the BmNPV-Ch nucleotide sequences of *vp39* gene (Fig. 3). A poly-A signal AATTAAA is present at the 3'-teminal region of BmNPV-Ch *vp39* gene.

Analysis of the deduced amino acid sequence of BmNPV-Ch vp39 protein

The BmNPV-Ch *vp39* gene codes for a 39 K protein containing 350 amino acids. This VP39 protein is by 3 amino acids longer than the BmNPV-Ja VP39 protein because of insertion of 9 nucleotides in BmNPV-Ch *vp39* gene.

Fig. 4 shows that BmNPV-Ch VP39 has 142 hydrophilic and 136 hydrophobic amino acids, which form a major hydrophilic domain (aa 210–288) and a major hydrophobic domain (aa 87–170). The N-terminus begins with 10 hydrophobic amino acids, while the C-terminus (aa 289–350) is mainly composed of hydrophilic amino acids. This figure clearly shows that the nucleotide change at nt 434 (A to C) lead to the amino acid change at nt 145 (Lys to Thr). This change caused disappearance of the small hydrophilic and hydrophobic domains. The insertion of CGC (nt 625–627) into the gene leading to the insertion of Arg into the protein resulted to an increase of the small hydrophilic domain and decrease of two small acidic domains.

Comparison of secondary structure of VP39 proteins of BmNPV-Ja and BmNPV-Ch

Table 1 shows that both proteins have a lot of β-sheets and β-turn elements. Apparent differences are in α-helix and random coil elements. β-sheets are mainly in aa 12–32 and aa 54–66 of the N-terminus; a large α-helix region is in aa 102–121. β-turn elements (about 31) are uniformly distributed in the whole protein. This makes the protein molecule to form a folded structure.

Discussion

The nucleocapsid protein *vp39* gene of BmNPV is not only a necessary structural gene associated with the assembly of the nucleocapsid, but also a gene closely related to the host cell splitting and primary fluidification (Rohrmann 1992; Todd *et al.*, 1996; Liu *et al.*, 1999). The newly synthesized VP39 protein is located in the cytoplasm, but after 12 hrs most of it is transferred into the nucleus to form

viral particles (Chariton *et al.*, 1993). The secondary structure of VP39 protein showed that the β-sheet and β-turn elements are its main structural elements accounting for 79.6%. This folded structure is important for the formation of the nucleocapsid and the release of the virus progeny (Chen *et al.*, 1997). This study proves that the *vp39* gene amplified by PCR from BmNPV-Ch is a functional homologue of the AcMNPV *vp39* gene on the basis of their sequences, transcripts, and expressed proteins. The homology of nucleotides between BmNPV-Ch *vp39* and AcMNPV *vp39* genes is 93.5%. Compared with the *vp39* gene of BmNPV-Ja, the nucleotide and the amino acid homologies are 97.5% and 97.1%, respectively. The ORF of BmNPV-Ch *vp39* is 1050 bp long and codes for 350 amino acids. The M_r of the expressed protein in *E. coli* is about 39 K, which is consistent with the deduced M_r of 350 amino acids.

The nomenclature of baculoviruses is closely related to that of their insect host. But it was proved (Fields *et al.*, 1990) that a certain insect can be infected by different baculoviruses and that certain baculoviruses isolated from different geographical areas and environments from the same host are of different serotypes and ecotypes. Therefore, it is important to investigate the genotype and phenotype of the isolates from different regions. This has been ignored in insect virology in the past. BmNPV-Ja and BmNPV-Ch are two geographical isolates that evolved apparently independently in different environments and the differences between them arose during evolution.

In this study we compared the *vp39* gene of both isolates in nucleotide sequence, amino acid sequence, hydrophilicity, hydrophobicity, and the secondary structure. It was found that there are 17 nucleotide differences in the gene leading to 7 amino acid differences. The changes in amino acids resulted in a small modification in hydrophilicity and hydrophobicity, namely the α-helix region slightly increased, while the random coil region slightly decreased. The secondary structure of the protein keeps using β-sheet and β-turn elements as the main structural elements. The effect of these differences on the function of the BmNPV VP39 protein needs further study.

Acknowledgement. This study was supported by the National Natural science Foundation of P. R. China.

Table 1. Secondary structure elements of VP39 proteins of BmNPV-Ch and BmNPV-Ja

Secondary structure element	BmNPV-Ja		BmNPV-Ch	
	%	No. of amino acids	%	No. of amino acids
α-helix	13.2	46	15.4	54
β-sheet	34.6	120	34.2	120
β-turn	45.0	156	44.7	156
Random coil	7.2	25	5.8	20

References

Ayres MD, Howard SC, Kuzio J, Lopez-Ferber M, Possee RD (1994): The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**, 586–605.

Blissard GW, Quant-russell RL, Rohrmann GF, Beaudreau GS (1989): Nucleotide sequence, transcriptional mapping, and temporral expression of the gene encoding p39, a major structural protein of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*. *Virology* **169**, 354–362.

Chariton CA, Volkman LE (1993): Penetration of Autographa californica nuclear polyhedrosis virus nucleocapsids into IPLB Sf21 cells induces actin cable formation. *Virology* **197**, 245–254.

Chen JG, Teng JL, Zhai ZH(1997): Morphogenesis of AcMNPV nucleocapsid. *Acta Microbiol. Sin.* **37**, 83–86.

Deng XZ, Zhu Y, Diao ZY, Qi YP (1999): Construction of bi-shuttle vector for *Bombyx mori* nuclear polyhedrosis virus and expression of a GFP/HBeAg fusion protein. *J. Wuhan Univ.* **45**, 220–224.

Fieds BN, Knipe DM, Chanock RM, Hirsch MS, Melnick JL, Monath TP, Roizman B (1990): *Fields Virology.* 2nd ed., Raven Press, New York.

Lanier LM, Volkman LE (1998): Actin binding and nucleation by *Autographa californica* M nucleopolyhedrovitus. *Virology* **243**, 167–177.

Liu DL, Jin TQ, Qi YP, Li L (1999): Sequencing of a 3.3 kb Fragment from LsNPV Genome and Comparison of its Genome Structure with Other Two NPVs. *Acta Biochem. Biophys. Sin.* **31**, 98–102.

Liu DL, Qi YP, Sun XJ, Li XF, Zhu Y (1999): Cloning and expression of the *vp*39 gene from *Autographa calofornica* nuclear polyhedrosis virus and its function. *Chin. J. Biochem. Mol. Biol.* **31**, 215–219.

Rohrmann GF (1992): Baculovirus structural proteins. *J. Gen. Virol.* **73**, 749–761.

Slack JM, Kuzio J, Faulkner P (1995): Characterization of v-cath, a cathepsin L-like proteinase expressed by the baculovirus *Autographa californica* multiple nuclear polyhedrosis virus. *J. Gen. Virol.* **76**, 1091–1098.

Todd JW, Passarelli AL, Lu A, Miller LK (1996): Factors regulating baculovirus late and very late gene expression in transient-expression assays. *J. Virol.* **70**, 2307–2317.

Zhu Y, Qi YP, Liu DL, Joshua MN, Wang YF (1998): A novel expression system based on host-range expansion of baculovirus. *J. Virol. Methods* **76**, 101–108.