Journal of Biochemistry and Molecular Biology, Vol. 35, No. 6, November 2002, pp. 595-603

JOURNAL OF
Biochemistry and Molecular Biology

Identification and Characterization of a Conserved Baculoviral Structural Protein ODVP-6E/ODV-E56 from *Choristoneura fumiferana* Granulovirus

Kianoush Khajeh Rashidan*, Nasha Nassoury†, Paresa N. Giannopoulos and Claude Guertin
Institut national de la recherche scientifique-Institut Armand-Frappier 531 Blvd des Prairies, H7V 1B7, Laval-Canada
†Université de Montréal, 4101 Sherbrook east, H1X 2B2 Montreal, Canada

Received 6 August 2002, Accepted 25 September 2002

A gene that encodes a homologue to baculoviral ODVP-6E/ODV-E56, a baculoviral envelope-associated viral structural protein, has been identified and sequenced on the genome of *Choristoneura fumiferana* granulovirus (ChfuGV). The ChfuGV *odvp-6e/odv-e56* gene was located on an 11-kb *BamH*I subgenomic fragment using different sets of degenerated primers, which were designed using the results of the protein sequencing of a major 39 kDa structural protein that is associated with the occlusion-derived virus (ODV). The gene has a 1062 nucleotide (nt) open-reading frame (ORF) that encodes a protein with 353 amino acids with a predicted molecular mass of 38.5 kDa. The amino acid sequence data that was derived from the nucleotide sequence in ChfuGV was compared to those of other baculoviruses. ChfuGV ODVP-6E/ODV-E56, along with other baculoviral ODVP-6E/ODV-E56 proteins, all contained two putative transmembrane domains at their C-terminus. Several putative *N*- and *O*-glycosylation, *N*-myristoylation, and phosphorylation sites were detected in the ChfuGV ODVP-6E/ODV-E56 protein. A similar pattern was detected when a hydrophobicity-plots comparison was performed on ChfuGV ODVP-6E/ODV-E56 with other baculoviral homologue proteins. At the nucleotide level, a late promoter motif (GTAAG) was located at −14 nt upstream to the start codon of the ChfuGV *odvp-6e/odv-e56* gene. A slight variant of the polyadenylation signal, AATAAT, was detected at the position +10 nt that is downstream from the termination signal. A phylogenetic tree for baculoviral ODVP-6E/ODV-E56 was constructed using a maximum parsimony analysis. The phylogenetic estimation demonstrated that ChfuGV ODVP-6E/ODV-E56 is most closely related to those of *Cydia pomonella* granulovirus (CpGV) and *Plutella xylostella* granulovirus (PxGV).

Keywords: *Choristoneura fumiferana* granulovirus, Envelope-associated protein, ODVP-6E/ODV-E56, Phylogeny, Protein analysis

Introduction

Granuloviruses (GVs) are members of the Baculoviridae. Baculoviruses have relatively large double-stranded DNA genomes (80 to 165 kb). Baculoviruses are frequently used as bio-pesticides against phytophagous insects, belonging mainly to the orders *Lepidoptera*, *Hymenoptera*, and *Diptera* (Federici, 1999). *Choristoneura fumiferana* granulovirus (ChfuGV) is pathogenic to the spruce budworm, the most devastating forest defoliation pest in eastern Canada and the United States.

Like other members of the family, the replication of ChfuGV in insects is biphasic, and involves two morphological-distinct forms of virion; occlusion-derived virus (ODV) and budded virus (BV). The dissociation of ingested-viral-occlusion bodies (a highly organized paracrystalline matrix that contains the ODV), which is under the alkaline environment that is inside the midgut of the susceptible-insect larvae, yield ODVs. These bind to midgut epithelium cells and initiate the primary infection. During the early infection stages in midgut cells, the BV phenotype is produced. This phenotype (with a high potency to infect a variety of cell types in the insect body) derives its envelope by budding through a modified-plasma membrane to start a systemic infection. Later in the infection, the ODV form of the virus is produced in various tissues of the infected insect, and released upon the dead insect. The ODV phenotype obtains its envelope from intracellular microvesicules within the nucleoplasm of infected cells (Hong *et al*., 1994; Braunagel *et*

The nucleotide sequence data that is reported in this paper has been submitted to the GenBank nucleotide sequence database, and assigned the accession number AF389081.

*To whom correspondence should be addressed.
Tel: 204-230-0210; Fax: 204-237-3240
E-mail: kiarashidan@juno.com

al., 1996). The invaginated-inner nuclear membrane has been hypothesized to be the source of the baculovirus-induced intracellular microvesicles (Braunagel *et al*., 1996). It has been shown that one of the ODV-specific envelope proteins that can be found in these intracellular microvesicle structures is ODVP-6E/ODV-E56 (Braunagel *et al*., 1996).

An analysis of the structural proteins by SDS-PAGE revealed that baculoviruses have a complex structure with more than 25 polypeptides in their BV and ODV phenotypes (Braunagel and Summer, 1994). Both phenotypes have complex structures; several proteins were identified as specific to ODV or BV (Rohrmann, 1992). Among these structural proteins, at least eight have been known to be specific to the ODV phenotype. These are as follows: VP17 (Funk and Consigli, 1993), ODV-E25 (Russell and Rohrmann, 1993), ODV-E35 (Braunagel *et al*., 1996), GP41 (Whitford and Faulkner, 1993), P74 (Kuzio *et al*., 1089), ODV-E18 (Braunagel *et al*., 1996), ODV-E66 (Hong *et al*., 1994), and ODVP-6E/ODV-E56 (Braunagel *et al*., 1996, Theilmann *et al*., 1996). The proteins that could participate in adsorption, fusion, and penetration have not been identified. However, these polypeptides [except GP41, which has been localized to the tegument region of the virion (Whitford and Faulkner, 1993)] could be involved in this process.

In the present study, we report the identification and characterization of a gene homologue to the baculoviral *odvp-6e/odv-e56* in ChfuGV. The nucleic acid and deduced amino acid sequence of ChfuGV is presented. In order to search for the conserved motifs, the ChfuGV ODVP-6E/ODV-E56 predicted amino acid sequence was compared to those of other baculoviruses. The hydrophobicity plots and membrane-spanning regions of ChfuGV ODVP-6E/ODV-E56 were predicted and compared to those of other baculoviruses. All potential post-translational modifications on ChfuGV ODVP-6E/ODV-E56 were predicted using various bioinformatics tools. A phylogenetic analysis was performed in order to determine the evolutionary relationship between ChfuGV ODVP-6E/ODV-E56 proteins with those of other baculoviruses.

Materials and Methods

***In vivo* Production and Purification of Virus and DNA Extraction** *Choristoneura fumiferana* fourth-instar larvae were infected with ChfuGV using a virus-contaminated-artificial diet (Forté *et al*., 1999). Next, the virus isolation from infected larvae by DNA extraction was carried out, as previously described (Bah *et al*., 1997).

SDS-PAGE and Protein microsequencing Enveloped-nucleocapsids of ChfuGV were purified. To isolate the enveloped-nucleocapsids, occlusion bodies of ChfuGV were solubilized under alkaline conditions (0.5 M Sodium Carbonate pH 10.5 for 60 min on ice); the undissolved granules were separated by centrifugation (5,000 *g* for 10 min.). The supernatants were then layered on a sucrose gradient (10% to 50%) and centrifuged at 25,000 *g* (Beckman SW-41) for 60 min at 4°C. The enveloped-nucleocapsids were then collected, and after they were washed in distilled water, they were centrifuged at 100,000 *g* (Beckman SW-41) for 60 min at 4°C and resuspended in a TE buffer (10 mM-Tris-HCl, 1 mM-EDTA, pH 7.4). The purity and integrity of the enveloped-nucleocapsids were examined by a transmission-electron microscope. Purified-enveloped-nucleocapsids were disrupted in an equal volume of a Laemmli sample buffer (Laemmli, 1970) in the presence of β-mercaptoethanol (Bio-Rad, Hercules, USA). The samples were then boiled for 5 min and clarified at 13,000 *g* for 5 min before electrophoresis on a 3% stacking/12.5% separating-sodium-dodecyl-sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were visualized using light Coomassie Brilliant Blue (Invitrogen, San Diego, USA) staining. The molecular mass (*Mr*) of the virion proteins was determined by comparing them to those of the standards (Bio Rad).

The gel-purified-ChfuGV-structural proteins were used for the protein sequencing. Major protein bands from the ChfuGV enveloped-nucleocapsids preparation, including a 39 kDa band, were excised from the gel and subjected to a protein-sequencing analysis. The sequencing was performed by the Harvard Microchemistry Facility using either microcapillary reverse-phase HPLC nano-electrospray-tandem-mass-spectrometry (μLC/MS/MS) on a Finnigan LCQ quadrupole-ion-trap mass-spectrometer or chemical sequencing (Edman). The obtained sequences of several stretches of amino acids were used to design the degenerated primers (with the least redundancy) that were to be used in the PCR reactions on the ChfuGV genomic DNA. This was done in order to locate the gene that codes for this protein on the ChfuGV genome.

Cloning and DNA sequencing Using different combinations of degenerated primers, PCR amplifications were performed using ChfuGV genomic DNA as the template. A 758 bp fragment from the gene that encodes the ODVP-6E/ODV-E56 was amplified with a combination of the following pair of degenerated primers: 39-FRI: TTT/C ACA/T GGT/A TTA/G AGG/A and 39-RVII: G/ACG G/ACA G/AAC G/ACT C/TTG. All of the PCR reactions were carried out using *Taq* DNA polymerase (Amersham Pharmacia Biotech, Uppsala, Sweden) in a buffer that was supplied by the manufacturer. It was supplemented with 200 μM of each dNTP and $MgCl_2$ to a final concentration of 2.5 mM. The PCR reactions were subjected to heat for 10 min at 95°C prior to the start of the amplification (30 cycles of 95°C, 30 s; 45°C 30 s; and 72°C, 30 s). The PCR amplicon was cloned in a PCR2.1 (Invitrogen) cloning vector and sequenced. The sequence was compared to homologues in GenBank/EMBL using BLAST algorithm (Altshul *et al*., 1992). In order to locate the *odvp-6e/odv-e56* gene on the ChfuGV genome, subgenomic fragments were generated by *BamH*I endonuclesase and fractioned in 1% agarose gel. The gel-extracted fragments were then used as templates for the PCR reactions using the same set of primers. The PCR reaction on the 11 kb ChfuGV *BamH*I fragment yielded an amplicon with the expected size. This 11kb restriction fragment was cloned in pBluescript-SK^+ cloning vector (Stratagene, La Jolla, USA), and transformed into *E. coli* XL-1 Blue (Stratagen). All of the manipulations (including the restriction enzyme digestion, agarose gel electrophoresis, transformation, colony-lifting, and plasmid purification) were carried out according to standard protocols (Sambrook *et al*., 1989).

The fragment was sequenced from both extremities on either strand using the primer-walking technique. Automatic sequencing was carried out using an Applied Biosystems automated DNA sequencer model 377XL (Applied Biosystems, USA).

Computer Analysis Sequence-data assembly and analysis were performed with the Sequencher software version 4.0.5 (Gene Codes Corp.) and the MacVector program version 4.5.0 (Eastman Kodak). The nucleotide sequence and its predicted amino acid sequence were compared to homologues in GenBank/EMBL and SWISSPROT using BLAST (Altshul *et al*., 1992). CLUSTALX (1.81) (Thompson *et al*., 1997) was used for multiple amino-acid-sequence alignments. BOXSHADE was used for similarity shading and scoring among the aligned sequences. The alignment was used as the input to construct the phylogenetic tree with Branch and Bound search settings of PAUP 4.0b4a (Swofford, 2000) with 100 bootstrap replicates. Accession numbers for the sequences that were used in this study are as follows: *Cydia pomonella* granulovirus (CpGV), AAB39099; *Plutella xylostella* granulovirus (PxGV), AAG27314; *Xestia c-nigrum* granulovirus (XcGV), AAF05129; *Helicoverpa armigera* nuclear polyhedrovirus (HaNPV), AAK96270; *Spodoptera exigua* nucleopolyhedrovirus (SeNPV), AAF33537; *Orgyia pseudotsugata* multicapsid polyhedrovirus (OpMNPV), AAC59145; *Choristoneura fumiferana* multicapsid nuclear polyhedrovirus (CfMNPV), AAA46698; *Epiphyas postvittana* nucleopolyhedrovirus (EpNPV), AAK85694, *Bombyx mori* nuclear polyhedrovirus (BmNPV), AAC63814; *Autographa californica* multicapsid nuclear polyhedrovirus (AcMNPV), AAA98967; *Lymantria dispar* multicapsid nuclear polyhedrovirus (LdMNPV), AAC70199; *Spodoptera litura* nucleopolyhedrovirus (SlNPV), AAL01703; and *Culex nigripalpus* baculovirus (CnBV), AAK13276. The amino acid composition was calculated using Peptide Statistic (http://web. umassmed.edu/cgi-bin/biobin/pepstats), Pepinfo (www.ebi.ac.uk/servicestmp/1598281014310854.html) and ProtParam (http://www. expasy.ch/tools/ protparam.html) tools. Hydropaty plots were obtain using Kyte and Doolittle hydropathy values (−4.5 to 4.5) via Nixon web tools (http://www.bmb.psu.edu/ nixon/ webtools.html). The secondary-structure-prediction method

```
   1 - GGGGGGATCCCAATAATAAATATTCTATAATTAATGGTTTGTTGTATGATATGCGTGACG - 60

  61 - AAACTGTTGTTGTAAAGATTAATACAGAAAGCACCGTCAACTCCAATTTATATATTCATG - 120

 121 - GTTCTTTTTAAAATTGTTTTACCTTGTAAGTTAAATTAAATGGCTTCTTTTTTTACGGGT - 180
     -                                       M  A  S  F  F  T  G
 181 - CTAAGACGCACTAATAAAGTATATCCAAACACAAACAGTTTTTTGTCTGACCATGCGTTA - 240
     - L  R  R  T  N  K  V  Y  P  N  T  N  S  F  L  S  D  H  A  L
 241 - TTTATTAGAAACCAAACACCTGTAGGTTTTAATTTAAACAATCCTACCACAATGGGTATT - 300
     - F  I  R  N  Q  T  P  V  G  F  N  L  N  N  P  T  T  M  G  I
 301 - GCGGGAGGTAATGTTACGCCGGGTTATAATATAAACGGTACCTTTGTTAGCAACGCCAAT - 360
     - A  G  G  N  V  T  P  G  Y  N  I  N  G  T  F  V  S  N  A  N
 361 - GTCAATTCAGTGCTTAGGAACAACGATGTTGTTGGTATGAGGCAAATTTTTCCGGAGGCA - 420
     - V  N  S  V  L  R  N  N  D  V  V  G  M  R  Q  I  F  P  E  A
 421 - TCCAATAATCAAATGAATGGTTTGACAAACTTGAGGAGAGCCGACAATATTCCAGATTCT - 480
     - S  N  N  Q  M  N  G  L  T  N  L  R  R  A  D  N  I  P  D  S
 481 - ACTATTCACAGTTTAAAAACAAGAAAAAATAACGTTAAACAATCACATCCCGAAACAGTA - 540
     - T  I  H  S  L  K  T  R  K  N  N  V  K  Q  S  H  P  E  T  V
 541 - GTTAGAGATAGAGCGGGTGTCGAAAATGCACTAGCGCAGAATCCGCGTTTAGGCGACTAT - 600
     - V  R  D  R  A  G  V  E  N  A  L  A  Q  N  P  R  L  G  D  Y
 601 - TTAAGGGGTGCCGGATATGTGACATTGTTTGGAGTGGGTGTTTATTTAGTTATTAATGTC - 660
     - L  R  G  A  G  Y  V  T  L  F  G  V  G  V  Y  L  V  I  N  V
 661 - GCCGATCTCGTCGGGTCAATAGTTGACGCGATAAACCGCACTGGTGGTAGCTGGTATTTT - 720
     - A  D  L  V  G  S  I  V  D  A  I  N  R  T  G  G  S  W  Y  F
 721 - AGAGGCAATAATGGGGGTGATAATTTTAACAACATTCAATCGTGCATATTGCGCTACAGA - 780
     - R  G  N  N  G  G  D  N  F  N  N  I  Q  S  C  I  L  R  Y  R
 781 - ACTTGCGGCGTACCATTTACAGATATTCAAGATTCGGTATGTGTTCTAGATCCACACGAC - 840
     - T  C  G  V  P  F  T  D  I  Q  D  S  V  C  V  L  D  P  H  D
 841 - GCTACAAACGTTGACCCACTAATGACTTTAGAAGAAGCAAGTGATTTTTGCAATAACTAT - 900
     - A  T  N  V  D  P  L  M  T  L  E  E  A  S  D  F  C  N  N  Y
 901 - AATCATGCAACAGAACAAAGTGTTTGTCGTGGTTCAGATCCAAACGCAAATCCTACAAGT - 960
     - N  H  A  T  E  Q  S  V  C  R  G  S  D  P  N  A  N  P  T  S
 961 - TTACAATACCTTGACATAAGTCTTCTTGCATCCAATCAAACTATAGAATGTATCGAACCC - 1020
     - L  Q  Y  L  D  I  S  L  L  A  S  N  Q  T  I  E  C  I  E  P
1021 - TACGATTTTGGTGATTTAATAGGAGATTTAGGATTAGATTGGTTATTAGGCGATAATGGT - 1080
     - Y  D  F  G  D  L  I  G  D  L  G  L  D  W  L  L  G  D  N  G
1081 - ATTGTCACCGCCAGCTCTAACAGTTTATTAAGTGTATCGGACAATTTTTTGACCATAATA - 1140
     - I  V  T  A  S  S  N  S  L  L  S  V  S  D  N  F  L  T  I  I
1141 - TTAGTGATAGGGGGAATTTTAGTTTTGCTTTTCATTGGATTTATAATCTACAAAGTAACT - 1200
     - L  V  I  G  G  I  L  V  L  L  F  I  G  F  I  I  Y  K  V  T
1201 - ATAACGAAAAGAAGTATATGAACACGAGTTAATAATGAGCGTGAGTTAAAATTAAAAATT - 1260
     - I  T  K  R  S  I  *
1261 - TTAATTTTAAAAATATTTGCTATTTTATGTCGCGCTCATTATTTGCTAAAAATACACAAA - 1320

1321 - TTTGTATATCGATAAGCGCGAGTTAAAACAGCAAAAATTTTTAAATTTTTAATTTGCTGT - 1380

1381 - TTTAACTCGCGCTCATTTCTAGAAACAAACAAGCGAAGAGCGAGCGGTGTAACCCAAATT - 1440

1441 - GAGATTAAAGCTCATTTGTAAATGTGTATAATCAATAATATAGTAACATATTAAGTGATG - 1500

1501 - TAAAATAAGTATT - 1513
```

Fig. 1. Figure shows nucleotide and protein sequence of ChfuGV ODVP-6E/ODV-E56. All possible *cis*-acting elements located within the noncoding leader region are in underlined italic. Putative polyadenylation signal is in underlined bold, astrix (*) shows the stop codon. The position of degenerated primers is in bold and italics.

that was used in this study was GORIV (Garnier *et al*., 1996). Potential transmembrane regions were identified using the Tmpred (Hofmann and Stoffel, 1993) and TMHMMver. 2.0 (Moller *et al*., 2001) programs. The *N*- and *O*-linked glycosylation, *N*-myristoylation, and phosphorylation sites were predicted using the NetOglyc (Hansen *et al*., 1998) and Proscan (Bairoch *et al*., 1997) programs.

Results and Discussion

In this paper, we described the identification and characterization of a gene on the ChfuGV genome that encodes a 39kDa protein homologue to ODVP-6E/ODV-E56 from CpGV, PxGV, XcGV, HaNPV, SeNPV, OpMNPV, CfMNPV, EpNPV, BmNPV, AcMNPV, LdMNPV, SlNPV, and CnBV. The results of protein sequencing of the major 39 kDa protein in the ChfuGV enveloped-nucleocapsid yielded several stretches of amino acids. These amino acid sequences were quite similar to the highly-conserved-ODV-specific envelope protein (known as ODVP-6E/ODV-E56) after being compared to homologue proteins in GenBank/EMBL using BLAST algorithm (Altshul *et al*., 1992). Based on the protein sequence of ODVP-6E/ODV-E56, different sets of degenerated primers were developed. DNA coding for ODVP-6E/ODV-E56 was amplified using a pair of these degenerated primers (39-FRI and 39-RVII). The gene that encodes this protein was located in an 11-kb *BamH*I fragment. The open-reading frame (ORF) is 1062 nt, potentially encoding 353 amino acids, with an estimated molecular mass of 38.5 kDa (the actual size of the protein on SDS-PAGE is 39 KDa). The size of baculoviral ODVP-6E/ODV-E56 proteins varies in ranges between 351 aa in PxGV to 379 aa in CfMNPV. The localization and orientation of the *odvp-6e/odv-e56* gene on the ChfuGV genome is similar to that described for CpGV (Luque *et al*., 2001). The 5 noncoding region of *odvp-6e/odv-e56* was analyzed with the aim of detecting all possible *cis*-acting elements and possible transcription start sites (tss). A TATA box and CCAAT element were centered, respectively, at 134 nt and 150 nt from the putative-start triplet. A late promoter motif (GTAAG) was located at 14 nt upstream of the first ATG. The TAAG motif is a strong promoter of the late genes in baculoviruses. The gene contained a slight variant of a polyadenylation signal, AATAAT, at the 3 noncoding region at the position +10 nt downstream from the termination signal. The nucleotide sequence that surrounds the translational start codon (TAAATGG) was conformed partly to Kozaks rule for efficient eukaryotic-translation initiation with the presence of a purine base only at the +4 position and not at the −3 (Kozak, 1986) (Fig. 1).

The deduced-amino-acid composition in terms of non-polar (ACFGILMPVWY), polar (DEHKNQRST), basic (HRK), and acidic residues (DE) indicates non-polar (54.4%), polar (45.6%), acidic (8.8%), and basic (8.2%) contents. The protein is particularly rich in asparagine (11.6%) residues. The protein has three potential *N*- glycosylation N-{P}-[ST]-{P} sites that are located at residues 59, 179, and 279. Glycosylation is important for the proper-folding or targeting of some of the polypeptides (Darvey, 1989). Six possible *O*-glycosylation sites, eight *N*-myristoylation (G-{EDRKHPFYW}-X-X-[STAGCN]-{P}) sites, and nine phosphorylation ([ST]-X-[RK] or [ST]-X-X-[DE]) sites were predicted (Table 1). An analysis of the ODVP-6E/ODV-E56 hydrophobicity plots suggested the presence of two hydrophobic domains within the C-terminal that were conserved in all of the baculoviral ODVP-6E/ODV-E56 proteins. These highly-hydrophobic regions that encompass two membrane-spanning regions are presented in ChfuGV and all other baculoviral ODVP-6E/ODV-E56 proteins. The locations of these two membrane-spanning regions were almost the same in all baculoviral ODVP-6E/ODV-E56 proteins (Figs. 2 and 3).

Besides these two regions, the ChfuGV ODVP-6E/ODV-E56 protein shared almost similar hydrophobicity plots with other baculoviral ODVP-6E/ODV-E56 proteins (Fig. 4). On the other hand, the overall secondary-structure comparison showed a relatively high degree of similarities between

Table 1. List and the location of all putative N- glycosylation, O-glycosylation, N-myristoylation phosphorylation sites within the ChfuGV ODVP-6E/ODV-E56 protein.

Modification	Residue	Sequence
N-glycosylation	59	NGTF
	179	NRTG
	279	NQTI
O-glycosylation	64	S
	107	S
	181	T
	254	S
	267	S
	352	S
Phosphorylation	11	TNK
	111	SLK
	114	TRK
	349	TKR
	122	SHPE
	173	SIVD
	229	TNVD
	236	TLEE
	318	SVSD
N-myristoylation	7	GLRRTN
	46	GIAGGN
	60	GTFVSN
	133	GVENAL
	189	GNNGGD
	210	GVPFTD
	258	GSDPNA
	307	GIVTAS

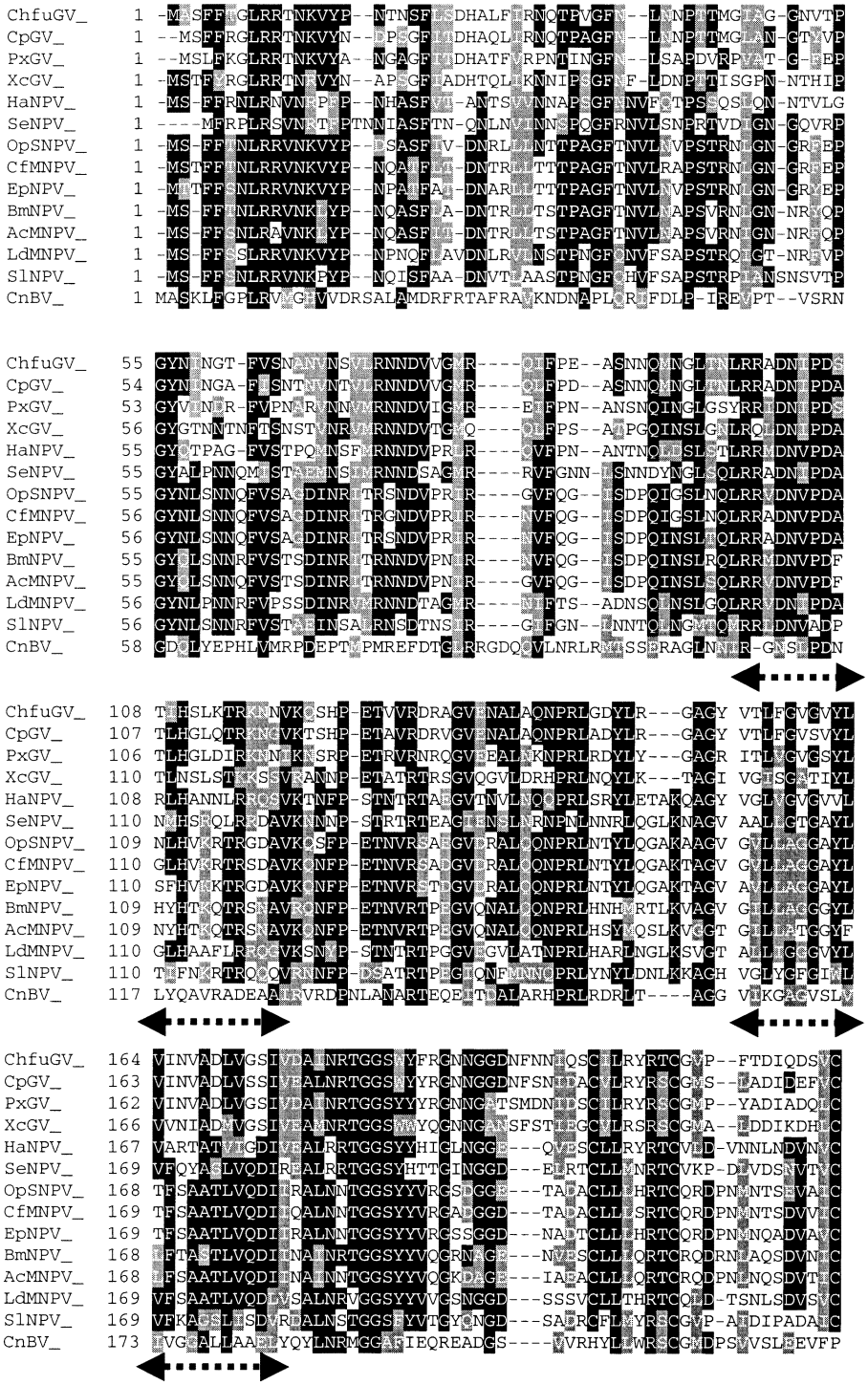

Fig. 2. Alignment of ChfuGV ODVP-6E/ODV-E56 with other baculoviral ODVP-6E/ODV-E56 proteins. The alignment was processed using CLUSTAL X software and BOX SHADE tool. Two shading levels were set: gray for more than 50% and black for 100% identity. Arrows demonstrate location of transmembrane domains. Accession numbers for the sequences used for this alignment are as follow: CpGV, AAB39099; PxGV, AAG27314; XcGV, AAF05129; HaNPV, AAK96270; SeNPV, AAF33537; OpMNPV, AAC59145; CfMNPV, AAA46698; EpNPV, AAK85694, BmNPV, AAC63814; AcMNPV, AAA98967; LdMNPV, AAC70199; SlNPV, AAL01703; and CnBV, AAK13276.

ChfuGV and other known baculoviral ODVP-6E/ODV-E56 proteins (data not shown).

The deduced-amino-acid sequence of ChfuGV ODVP-6E/ODV-E56 was compared to other known baculoviral ODVP-

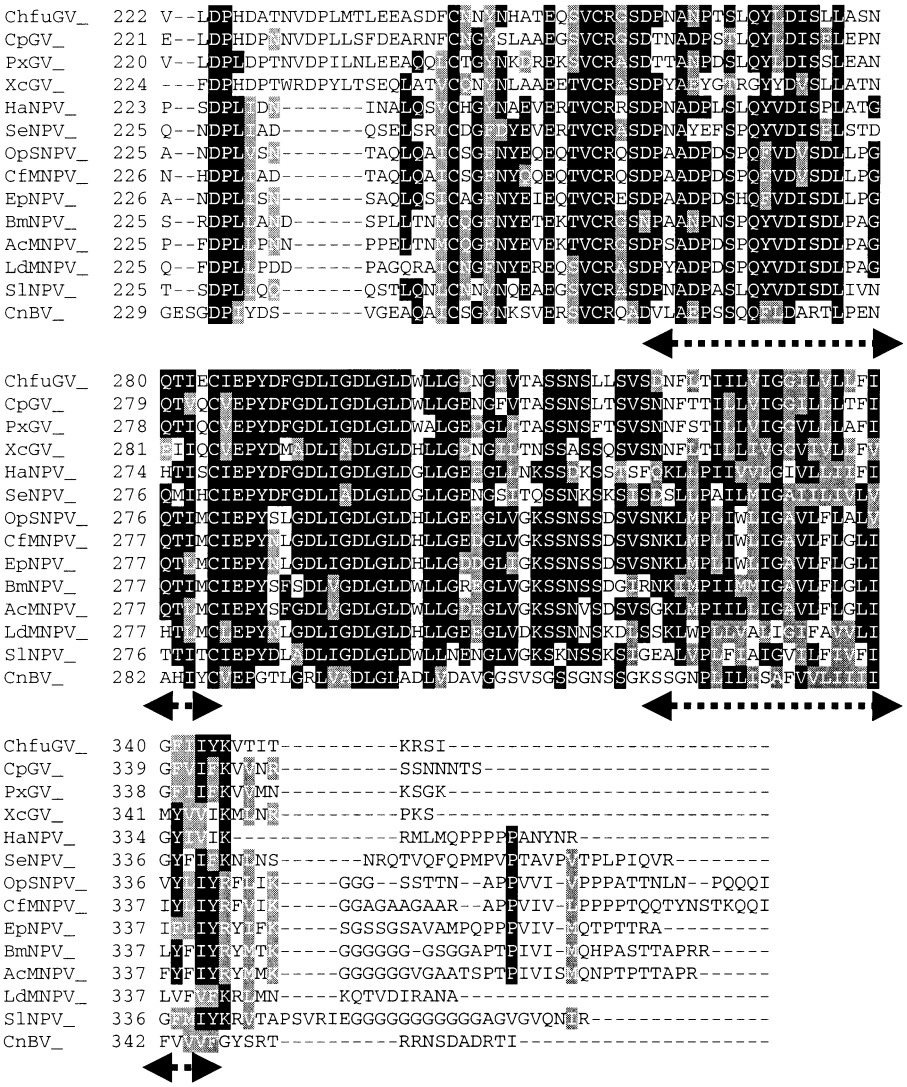

Fig. 2. Continued.

6E/ODV-E56 proteins. These results revealed a significant similarity between this protein sequence in ChfuGV and other granuloviral and nucleopolyhedral ODVP-6E/ODV-E56 proteins (Table 2). ODVP-6E/ODV-E56 in ChfuGV has a 57% nucleotide-sequence identity and 74% amino acid sequence identity to its closest relative, CpGV. CnBV ODVP-6E/ODV-E56 was the most distant, showing a 24% amino acid sequence identity. A high degree of conservation in amino acid sequences was observed within the two membrane-spanning regions of the granuloviral ODVP-6E/ODV-E56 proteins (Fig. 2). Six other conserved motifs were also detected in the ODVP-6E/ODV-E56 proteins from ChfuGV and other GVs. Three such motifs were located within the N-terminal region, where the same number was observed within the C-terminal. These motifs were (^{8}LRRTN[K/R]VY^{15}, ^{71}V[L/M]RNNDV[V/I/T]GMR[Q/E][L/I]F^{84}, ^{102}DNIPD[S/A]T[I/L]^{109}, ^{174}V[D/E]A[I/L/M]NRTGGS[W/Y][Y/W]^{186}, ^{254}[S/T]VCR[G/A]SD^{260}, and ^{284}C[I/V]EPYD[F/M][G/A]DLI[G/A]DLGLD^{300}. ChfuGV ODVP-6E/ODV-E56 shared a distinct motif with all of the other baculoviral ODVP-6E/ODV-E56 (except for CnBV). This motif was located between the two membrane-spanning regions within the C-terminal region of the protein (^{292}DL[I/V][G/A]DLGLD^{300}). ChfuGV, CpGV, and PxGV shared a highly-conserved stretch of 31 amino acids that covered a region from residue 162 to 193 of ChfuGV ODVP-6E/ODV-E56. A close look at this region in other baculoviral ODVP-6E/ODV-E56 revealed that other baculoviral ODVP-6E/ODV-E56, more or less, maintain the same pattern, although some residues are shifted.

The presence of six conserved-cysteine residues between the two hydrophobic membrane-spanning regions was revealed by performing a multiple alignment on the deduced-amino-acid sequences of ChfuGV ODVP-6E/ODV-E56, as well as those of other baculoviruses (Fig. 2). These highly-conserved-cysteine residues might be implicated in the

Table 2. Comparison of deduced amino acid sequence of ChfuGV ODVP-6E/ODV-E56 with that of three other GVs and ten NPVs ODVP-6E/ODV-E56 proteins. Pair wise identity values (%) between taxa are shown.

	CpGV	PxGV	XcGV	SeNPV	HaNPV	BmNPV	AcMNPV	CfNPV	OpMNPV	EpNPV	LdMNPV	SlNPV	CnBV
ChfuGV	74.44	66.1	52.99	47.2	46.59	46.04	45.32	45.91	45.46	46.78	45.19	46.63	23.96
CpGV		66.38	56.86	44.44	45.13	42.86	44.48	45.03	46.33	46.51	44.61	42.86	25.59
PxGV			49.57	42.31	44.48	42.77	44.12	44.12	43.36	43.53	44.57	43.07	27.38
XcGV				40.12	43.45	37.24	37.13	38.3	38.71	37.43	39.65	36.37	23.96
SeNPV					50.57	44.8	46.74	47.67	47.38	47.54	50	44.05	23.92
HaNPV						49.43	49.43	48.72	48.72	48.41	50.29	45.2	26.39
BmNPV							83.42	66.13	66.3	65.68	52.82	49.31	24.85
AcMNPV								68.82	69.92	68.73	55.49	47.94	25.36
CfNPV									86.1	83.51	56.21	47.24	26.88
OpMNPV										82.02	54.83	46.67	26.09
EpNPV											55.37	46.15	24.42
LdMNPV												48.02	26.32
SlNPV													22.25

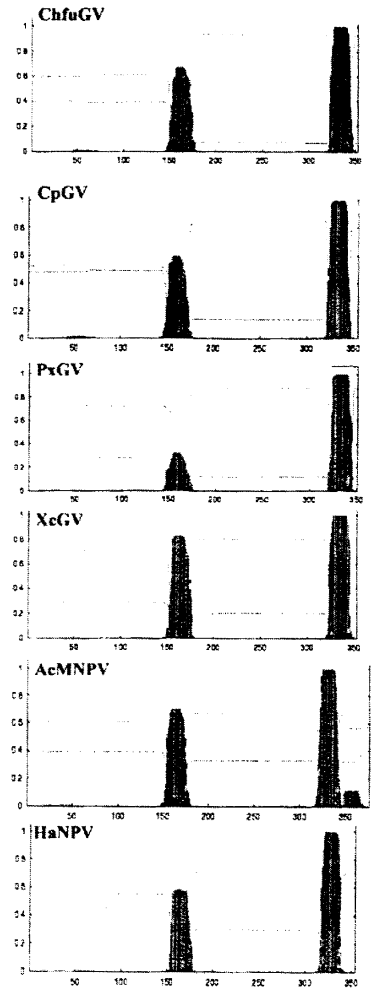

Fig. 3. Position of potential transmembrane domains in ChfuGV ODVP-6E/ODV-E56 and other baculoviral ODVP-6E/ODV-E56 proteins. The amino acid residue numbers are on the X-axis and probability of the existence of transmembrane domains are on Y-axis.

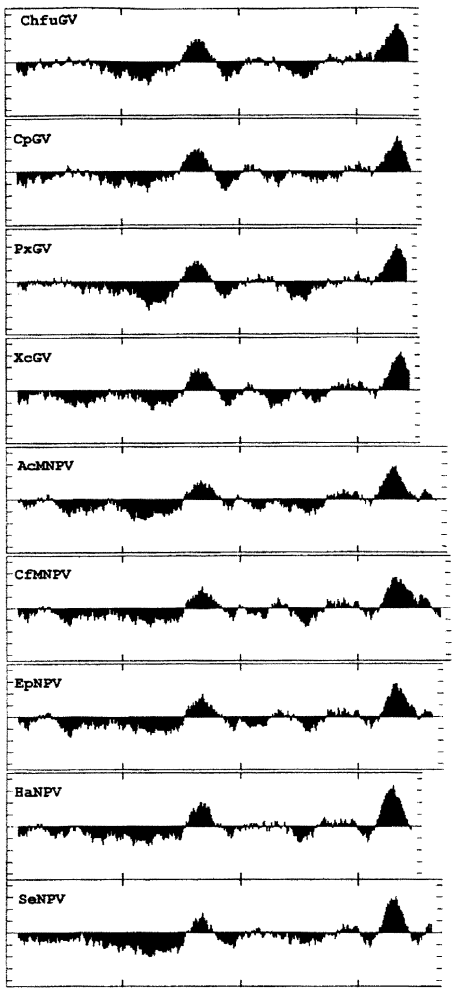

Fig. 4. Comparison of the hydrophilicity plots of ODVP-6E/ODV-E56 from ChfuGV to CpGV, PxGV, XcGV, AcMNPV, SeNPV and HaNPV ODVP-6E/ODV-E56 proteins. Above the axis (+) denotes hydrophilic regions and below the axis (–) indicates hydrophobic regions.

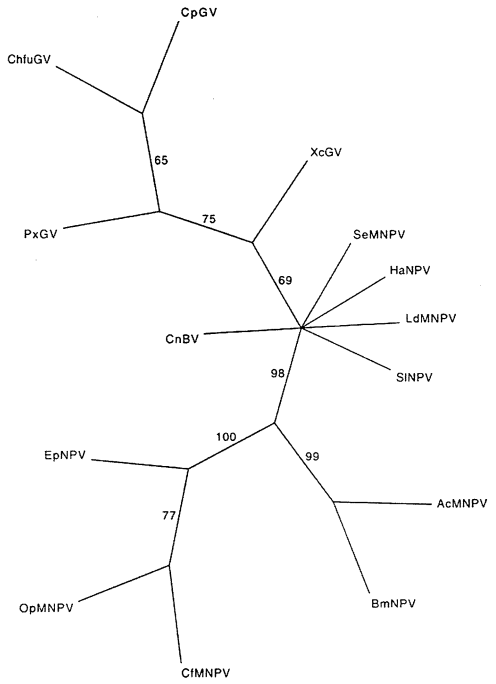

Fig. 5. Phylogenetic analysis performed on baculoviral ODVP-6E/ODV-E56 proteins. The unrooted tree was produced by a Branch and Band search using PAUP4.0b4a. Bootstrap values (100 replicates) are shown. Tree was constructed according to the following settings: 1-uninformative characters ignored. 2-Branch-and -bound search options: an initial upper bound: compute via stepwise; keep minimals only; collapse zero-length branches ON; MULPARS ON; addition sequences: furthest. Accession numbers for the sequences used are as follow: CpGV, AAB39099; PxGV, AAG27314; XcGV, AAF05129; HaNPV, AAK96270; SeNPV, AAF33537; OpMNPV, AAC59145; CfMNPV, AAA46698; EpNPV, AAK85694, BmNPV, AAC63814; AcMNPV, AAA98967; LdMNPV, AAC70199; SlNPV, AAL01703; and CnBV, AAK13276.

formation of disulfide bridges that eventually lead to correct folding of the baculoviral ODVP-6E/ODV-E56 proteins.

A phylogenetic analysis of the ODVP-6E/ODV-E56 proteins is shown in Fig. 5. A tree was produced via maximum parsimony to estimate the evolutionary relationship between the baculoviral ODVP-6E/ODV-E56 proteins. These analyses showed a clear division between the GV and NPV ODVP-6E/ODV-E56 proteins. ChfuGV is positioned alongside CpGV in the same cluster, supported by a 65% bootstrap value.

The data presented in this paper provide evidence to demonstrate the fact that the ODVP-6E/ODV-E56 protein is a highly-conserved protein in GVs and NPVs. The high degree of conservation, even between two distantly-related members of the baculoviridae family, indicates the importance of this protein, and can be translated to a possible significant function that is accomplished by this protein throughout the infection cycle. We are currently conducting studies that are aimed at further characterizing ChfuGV *odvp-6e/odv-e56* at the transcriptional and translational level.

References

Altshul, S. F., Gish, W., Miller, W., Meyers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

Bah, A., Bergeron, J., Arella, M., Lucarotti, C. J. and Guertin, C. (1997) Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. *Arch. Virol.* **142**, 1577-1584.

Bairoch, A., Bucher, P. and Hofmann, K. (1997) The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25**, 217-221.

Braunagel, S. C., He, H., Ramamurthy, P. and Summers, M. D. (1996) Transcription, translation, and cellular localization of three *Autographa californica* nuclear polyhedrosis virus structural proteins: ODV-EI8, ODV-E35, and ODV-EC27. *Virology* **222**, 100-114.

Braunagel, S. C. and Summers, M. D. (1994) *Autographa californica* nuclear polyhedrosis virus, PDV, and ECV viral envelopes and nucleocapsids: structural proteins, antigens, lipid and fatty acid profiles. *Virology* **202**, 315-328.

Darvey, J. (1989) Sorting out the secretory pathway. *Bioassays* **11**, 185-187.

Federici, B. A. (1999) Naturally occurring baculoviruses for insect pest control. *Methods in Biotechnology* **5**, 301-305.

Forté, A. J., Guertin, C. and Cabana, J. (1999) Pathogenicity of a granulovirus towards *Choristoneura fumiferana*. *Canadian Entomologist* **131**, 725-727.

Funk, C. J. and Consigli, R. A. (1993) Temporal expression and immunogold localization of *Plodia interpunctella* granulosis virus structural proteins. *Virus Res.* **28**, 57-66.

Garnier, J., Gibrat, J. -F. and Robson, B. (1996) GOR secondary structure prediction method version IV; in *Methods in Enzymology* vol. 266, Doolittle, R. F. (ed), pp. 540-553, Academic Press, San Diego, Carlifornia.

Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L. and Brunak, S. (1998) NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility.*Glycoconj. J.* **15**, 115-130.

Hofmann, K. and Stoffel, W. (1993) TMbase-A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler* **374**, 166.

Hong, T., Braunagel, S. C. and Summers, M. D. (1994) Transcription, translation, and cellular localization of PDV-E66: a structural protein of the PDV envelope of *Autographa californica* nuclear polyhedrosis virus. *Virology* **204**, 210-222.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulate translation by eukaryotic ribosomes. *Cell* **44**, 282-292.

Kuzio, J., Jaques, R. and Faulkner, P. (1989) Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. *Virology* **173**, 759-763.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Luque, T., Finch, R., Crook, N., O'Reilly, D. R. and Winstanley, D. (2001) The complete sequence of the Cydia pomonella

granulovirus genome. *J. Gen. Virol.* **82**, 2531-2547.

Moller, S., Croning, M. D. R. and Apweiler, R. (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**, 646-653.

Rohrmann, G. F. (1992) Baculovirus structural proteins, *J. Gen. Virol*. **73**, 749-761.

Russell, R. L. and Rohrmann, G.F. (1993) A 25-kDa protein is associated with the envelopes of occluded baculovirus virions. *Virology* **195**, 532-540.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Swofford, D. L. (2000) PAUP Version 4, Sinauer Associates, Sunderland, Massachusetts.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**, 4876-4882.

Whitford, M. and Faulkner, P. (1993) Nucleotide sequence and transcriptional analysis of a gene encoding gp41, a structural glycoprotein of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **67**, 2427.