

**Université du Québec
INRS-Institut Armand-Frappier**

**THÈSE PRÉSENTÉE À L'INRS-INSTITUT ARMAND-FRAPPIER COMME
EXIGENCE PARTIELLE POUR OBTENTION DU GRADE DE PHILOSOPHIAE
DOCTOR (Ph.D.) EN BIOLOGIE**

**Par
Kianoush Khajeh Rashidan**

**Identification, Characterization and Phylogenetic Analyses of
Genes Encoding Structural and Regulatory Proteins Located
within Two Conserved Regions on the *Choristoneura*
fumiferana Granulovirus Genome**

2005

Remerciements

J'aimerais remercier le Dr Claude Guertin pour m'avoir accueilli dans son laboratoire de recherche et pour la confiance qu'il m'a accordée.

J'aimerais remercier le Dr Yves Mauffette pour sa contribution dans ma formation professionnelle.

J'aimerais remercier le Dr David Bird non seulement pour son professionnalisme mais également pour son amitié.

J'aimerais remercier le Dr Max Arella avec qui j'ai eu des discussions tant au niveau professionnel que personnel durant ces années.

Je remercie Marc Henrichon pour les discussions scientifiques.

Un remerciement à tous les étudiant(e)s avec qui j'ai travaillé durant ces années : André Forté, Annie Bourassa et particulièrement Paresa Giannopoulos.

Je remercie la Fondation Armand-Frappier pour une bourse de 3^e cycle.

Je dédie cette thèse de doctorat à ma chère femme, mes parents et mon frère.

PRÉAMBULE

Les résultats de ce projet de doctorat sont présentés sous la forme de six publications.

Cette thèse comporte sept chapitres. Le premier est une revue bibliographique décrivant la biologie des virus de la famille *Baculoviridae*. Les chapitres 2 à 7 renferment les six publications dont cinq ont été publiées, une a été acceptée.

Chaque article possède sa propre structure (introduction, matériels et méthodes, résultats et discussion). Pour chaque publication, il y a une description de la contribution de chaque auteur. Finalement, une synthèse des résultats obtenus est présentée dans la dernière partie.

TABLE OF CONTENTS

REMERCIEMENTS.....	II
PRÉAMBULE.....	III
TABLE OF CONTENTS.....	IV
LIST OF TABLES.....	VII
LIST OF FIGURES.....	IX
LIST OF ABBREVIATIONS.....	XVI
RESUMÉ (EN FRANÇAIS)	XVIII
PREFACE.....	XXII
 GENERAL INTRODUCTION.....	 1
CHAPTER 1-LITERATURE REVIEW.....	4
1.1. Viral Bioinsecticides.....	5
1.2. Baculovirus.....	6
1.2.1. Budded and Occluded virions.....	7
1.2.2. Genome Structure of Baculovirus.....	8
1.2.3. Genes involved in virus replication in the host cell.....	10
1.2.4. Genes involved in baculovirus DNA replication.....	13
1.3. Granulovirus.....	15
1.3.1. Granulovirus replication.....	17
1.3.2. Virus ultrastructure.....	20
1.4. Structural Proteins in Baculoviruses.....	21
1.4.1. Baculovirus Major Capsid Proteins.....	23
1.4.2. Baculovirus major envelope proteins.....	27
 CHAPTER 2- PUBLICATION #1.....	 31
Identification, Characterization and Phylogenic Analysis of Conserved Genes within the <i>odvp-6e/odv-e56</i> Gene Region of <i>Choristoneura fumiferana</i> Granulovirus	
Contribution of the authors.....	33
Abstract in French:	34
Identification, caractérisation et analyse phylogénétique des gènes conservés dans la région du gène <i>odvp-6e/odv-e56</i> du granulovirus de <i>Choristoneura fumiferana</i>	
2.1. Abstract.....	35
2.2. Introduction.....	36
2.3. Material and Methods.....	38
2.4. Results and Discussion.....	41
2.5. Acknowledgments.....	53
 CHAPTER 3- PUBLICATION #2:	 54
Identification and Characterization of a Conserved Baculoviral Structural Protein (ODVP-6E/ODV-E56) from <i>Choristoneura fumiferana</i> Granulovirus.	
Contribution of the authors	56
Abstract in French:	57

Identification et caractérisation de la protéine structurale ODVP-6E/ODV-E56 du
granulovirus de *Choristoneura fumiferana* granulovirus

3.1. Abstract.....	59
3.2. Introduction.....	61
3.3. Material and Methods.....	64
3.4. Results and Discussion.....	68
3.5. Acknowledgments.....	86

CHAPTER 4- PUBLICATION #3:87

Identification and Characterization of a Putative Baculoviral Transcriptional Factor (IE-1) from *Choristoneura fumiferana* Granulovirus

Contribution of the authors.....89

Abstract in French:90

Identification et caractérisation du facteur transcriptionnel IE-1 du granulovirus de
Choristoneura fumifera

4.1. Abstract.....	92
4.2. Introduction.....	93
4.3. Material and Methods.....	96
4.4. Results and Discussion.....	98
4.5. Acknowledgments.....	119

CHAPTER 5- PUBLICATION #4:120

Identification, Characterization and Phylogenic Analysis of Conserved Genes within the
p74 Gene Region of *Choristoneura fumiferana* Granulovirus Genome

Contribution of the authors.....122

Abstract in French:123

Identification, caractérisation et analyse phylogénétique des gènes conservés dans la
région du gène *p74* du granulovirus de *Choristoneura fumiferana*

5.1. Abstract.....	124
5.2. Introduction.....	126
5.3. Material and Methods.....	128
5.4. Results and Discussion.....	130
5.5. Acknowledgments.....	146

CHAPTER 6- PUBLICATION #5:147

Choristoneura fumiferana Granulovirus *p74* Protein: A Highly Conserved Baculoviral
Envelope Protein

Contribution of the authors.....149

Abstract in French:150

Identification et caractérisation de la protéine structurale *p74* du granulovirus de
Choristoneura fumiferana

6.1. Abstract.....	151
6.2. Introduction.....	152

6.3. Material and Methods.....	155
6.4. Results and Discussion.....	160
6.5. Acknowledgments.....	186
 CHAPTER 7- PUBLICATION #6:	 187
Transcription, Translation, and Immunolocalization of ODVP-6E/ODV-E56 and p74 proteins: Two Highly Conserved ODV-Associated Envelope Proteins of <i>Choristoneura fumiferana</i> Granulovirus.	
Contribution of the authors.....	189
Abstract in French:	190
Transcription, traduction et immunolocalisation de l'ODVP-6E/ODV-E56 et p74: deux protéines hautement conservées chez le phénotype ODV du granulovirus de <i>Choristoneura fumiferana</i>	
 7.1. Abstract.....	 191
7.2. Introduction.....	192
7.3. Material and Methods.....	196
7.4. Results and Discussion.....	201
7.5. Acknowledgments.....	214
 GENERAL DISCUSSION AND CONCLUSION.....	 216
REFERENCES.....	230
 APPENDIX A.....	 252
LIST OF PUBLICATIONS AND COMMUNICATION DURING MY Ph.D. STUDIES.	

LIST OF TABLES

CHAPTER 2

Table1. Properties of ChfuGV ORFs detected in 11 kb restriction fragment and their related ORFs on other granulovirus genomes.

CHAPTER 3

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.

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.

CHAPTER 4

Table 1. Amino acid compositions (%), isoelectric points and molecular mass of ChfuGV and other baculoviral IE-1 proteins. Peptide Statistic and ProtParam tools were used for calculation.

Table 2. Comparison of deduced amino acid sequence of ChfuGV IE-1 with that of three GVs and ten NPVs IE-1 proteins. Pair wise identity values (%) between taxa are shown.

CHAPTER 5

Table1. Properties of ChfuGV ORFs detected in 8.9 kb *Bam*HI restriction fragment and their related ORFs on CpGV, PhopGV, AoGV ClGV, PxGV and XcGV genomes.

CHAPTER 6

Table 1. Comparison of conceptual amino acid sequences of ChfuGV p74 with those of other baculoviruses studied here. Pair wise identity values (%) between taxa are shown.

LIST OF FIGURES

CHAPTER 2

Figure1. Schematic illustration of the gene structure and orientation of a) *odvp-6e/odv-e56*-containing region of various GVs (i-ChfuGV, ii- CpGV, iii-PhopGV, iv-PxGV, v-XcGV); b) ORFs upstream to granulins (i- ChfuGV, ii-PhopGV, iii-AoGV, iv-CpGV, v-XcGV, vi- PxGV). The arrows represent the relative length and direction of the ORFs. White arrows indicate ORFs with no granuloviral homologue.

Figure2. Phylogenetic analysis of the protein product of baculoviral *odv-e18* gene located within the *odvp-6e/odv-e56* region of ChfuGV genome.

CHAPTER 3

Figure 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, asterix (*) shows the stop codon. The position of degenerated primers is in bold and italics.

Figure 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; SINPV, AAL01703; and CnBV, AAK13276.

Figure 3. Position of potential transmembrane domains in ChfuGV ODVP-6E/ODV-E56 and other baculoviral ODVP-6E/ODV-E56 proteins. The amino acid numbers are on the X-axis and probability of the existence of transmembrane (red), inside (blue) and outside (pink) domains are on Y-axis.

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

Figure 5. Phylogenetic analysis performed on baculoviral ODVP-6E/ODV-E56 proteins. The unrooted tree was produced by a Branch and Bound 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; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e 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; SINPV, AAL01703; and CnBV, AAK13276.

CHAPTER 4

Figure 1. Nucleotide sequence and predicted amino acid sequence of *ie-1* gene from ChfuGV. 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.

Figure 2. Alignment of the N-terminal (a) and C-terminal (b) of granuloviral IE-1 protein. 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. Accession numbers for the sequences used are as follow: CpGV AAK70674, XcGV CQ9PZ32; PxGV Q9DW20.

Figure 3. Hydrophobicity profiles using Kyte-Doolittle from C-terminal regions of all known granuloviral and two nucleopolyhedral IE-1 proteins. Above the axis (+) denotes hydrophilic regions and below the axis (-) indicates hydrophobic regions.

Figure 4. (a) HLH-like domain of IE-1. ChfuGV IE-1 residues 387 to 414 were aligned with equivalent IE-1 residues from NPVs and other GVs. Conserved hydrophobic (ϕ) residues are illustrated; (b) Basic residues (+) preceding the HLH-like domain of ChfuGV IE-1 are depicted.

Figure 5. Phylogenetic analysis performed on baculoviral IE-1 proteins. The unrooted tree was produced by a Branch and Bond 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; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e addition sequences: furthest. Accession numbers for the sequences

used are as follow: CpGV AAK70674, XcGV CQ9PZ32; PxGV Q9DW20; HaNPV AAK96269; SINPV AAL01702; HzNPV 010619; AcMNPV P11138; BmNPV P33245; EpNPV AAK85693; OpMNPV 22114; CfNPV P41716; SeNPV Q9J811; LdMNPV 036454; TnSNPV AAL14888.

CHAPTER 5

Figure.1. Schematic illustration of the gene structure and orientation of *p74*-containing region of various GVs (i- ChfuGV, ii- ClGV, iii-CpGV, iv-PhopGV, v-AoGV, vi-PxGV, vii-XcGV). The arrows represent the relative length and the direction of transcription of the ORFs.

Figure 2. Phylogenetic analysis of the conceptual protein products of baculoviral late expression factor 11 (*lef-11*) genes. The unrooted tree was produced employing the Neighbor joining (NJ) method using PAUP4.0b4a. Bootstrap values (1000 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; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e addition sequences: furthest. Accession numbers for the sequences used are as follow: ChfuGV AAN77198; CpGV AAK70718; ClGV AAQ21651; PhopGV AAM70251; XcGV AAF05170; AoGv AAP85687; PxGV AAG27344; McNPV AAM95134; SeNPV AAF33648; AhNPV AP006270; SINPV AAL01720; RoNPV AAN28128; LsNPV AB009614; AcMNPV AAA66667; BmNPV AAC63712; LdNPV AAC70230; OpMNPV BAA02637; EpNPV AAK85585.

CHAPTER 6

Figure 1. Schematic presentation of the 8.9kb BamHI fragment from ChfuGV genomic DNA showing the position and direction of the putative ORFs p74, p10, and *sod*. Only 573nt of p74 *orf* (denoted as p74p, “p” for partial) is on this fragment. The strategy employed for sequencing of complete p74 *orf* is presented, arrows show the position of the primers used for sequencing and the direction of sequencing.

Figure 2. Figure shows nucleotide and protein sequence of ChfuGV p74. Two late promoter motifs and a putative polyadenylation signal for p74 gene are in bold. The positions of degenerated primers are in underlined bold, and the locations of primers used for sequencing are underlined. Asterisk (*) shows the stop codon.

Figure 3. a) Amino acid alignments of deduced p74 protein of ChfuGV along with other p74 protein homologues studied here (see materials and methods for accession numbers). Conserved proline, glycine are depicted by (*), and conserved cysteine residues are illustrated by (+). Two highly conserved regions (C.R.) are shown within two boxes as C.R.I and C.R.II. b) Figure shows the position of conserved cysteine residues and their spacing patterns in baculoviral p74 protein homologues,

Figure 4. Structural analysis of p74: a) Predicted Kyte-Doolittle hydrophobic profiles of ChfuGV p74 and those of other GVs, NPVs, CnBV and Hz-1 virus. Arrows indicate two C-terminal hydrophobic domains. Profiles were generated using Pepwindow hydropathy analysis (7 residue average; Kyte and Doolittle, 1982). Positions of the amino acids are shown on the X-axis and hydrophobicity values are shown on the Y-axis. The list of baculoviral p74s in the right column is as follow: 1-PxGV; 2- AcMNPV; 3- OpMNPV; 4- CfMNPV; 5- LdMNPV; 6- BmNPV; 7- SiNPV; 8- SeNPV; 9- HaNPV; 10- EpNPV; 11-

CnBV; 12- Hz-1 virus. (b) Demonstration of transmembrane, out side, and inside domains of p74 proteins. (c) Schematic presentation of ChfuGV p74. Computer analysis (TMHMM-ver. 2) of the amino acid sequence predicts that the N terminal of the protein is located outside of the envelope followed by two transmembrane regions which results in an external C terminal. Each transmembrane domain spans about 20 amino acids and are shown as cylinders.

Figure 5. Phylogenetic analysis performed on baculoviral p74 proteins. The unrooted tree was produced by a Branch and Bond 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 protein sequences used are shown in material and methods section.

CHAPTER 7

Figure 1. (a) Analysis of *odvp-6e/odv-e56* gene transcripts using RT-PCR using *odvp-6e/odv-e56*-specific primer sets and total RNA purified from infected *Choristoneura fumiferana* larvae or uninfected larvae (Ctl.). The result of PCR when the reverse-transcription step is omitted is shown on the first lane. (b) Northern blot analysis on transcripts from infected larvae extracted 48 hr p.i. performed using a ^{32}P labeled DNA probe specific to *odvp-6e/odv-e56* revealed the presence of a prominent band detected at 1.2 kb. (c) Analysis of p74 gene transcripts using RT-PCR using p74-specific primer sets and total RNA purified from infected *Choristoneura fumiferana* larvae or uninfected

larvae (Ctl.). The result of PCR when the reverse-transcription step is omitted is shown on the last lane. (d) Northern blot analysis on transcripts from infected larvae extracted 48 hr p.i. performed using a ^{32}P labeled DNA probe specific to p74. The Northern blot analysis revealed the presence of a smear extents from 1.7 to almost 2 kb.

Figure 2. Preparation of uninfected (UL) or infected larvae with ChfuGV. Immunoblot analysis of ODVP-6E/ODV-E56 and p74 expression in ChfuGV infected fourth instar larvae. Larvae were harvested at different hours post infection (hr p.i.). Samples were analyzed by immunoblotting with the (a) anti-ODVP-6E/ODV-E56 antiserum. Band at 39kDa and (b) anti-p74 antiserum. Band at 74 kDa. Uninfected larvae (UL) were used as control.

Figure3. Immunoelectron microscopy of the ODV phenotype of ChfuGV. Thin sections of occluded virions were incubated with (a) antisera to ODVP-6E/ODV-E56 and (b) antisera to p74 protein, then the grids were incubated with goat anti-rabbit antibody conjugated to 10 nm gold particles and stained with 3 % uranyl acetate and 5 min with lead citrate. The grids then were viewed on a Philips model 410 transmission electron microscope operating at 80 kV.

LIST OF ABBREVIATIONS

aa:	Amino acid.
Ac(M)NPV:	<i>Autographa californica</i> (multicapsid) nucleopolyhedrovirus.
AgNPV:	<i>Anticarsia gemmatalis</i> nucleopolyhedrovirus.
AoGV:	<i>Adoxophyes orana</i> granulovirus.
bHLH:	Basic helix-loop-helix.
BmNPV:	<i>Bombyx mori</i> nucleopolyhedrovirus.
BSA:	Bovin Serum Albumin
Bs(S)NPV:	<i>Buzura suppressaria</i> single-nucleocapsid nucleopolyhedrovirus.
BV:	Budded virus.
Cf(M)NPV:	<i>Choristoneura fumiferana</i> (multicapsid) nucleopolyhedrovirus.
ChfuGV:	<i>Choristoneura fumiferana</i> granulovirus.
CIGV:	<i>Cryptophlebia leucotreta</i> granulovirus.
CnBV:	<i>Culex nigripalpus</i> baculovirus.
CpGV:	<i>Cydia pomonella</i> granulovirus.
DE:	Delayed early.
DNA:	Deoxyribonucleic acid
EMG:	Electron microscopy immunogold.
EpNPV:	<i>Epiphyas postvittana</i> nucleopolyhedrovirus.
GV:	Granulovirus.
HaNPV:	<i>Helicoverpa armigera</i> nucleopolyhedrovirus.
HbGV:	<i>Harrisina brillians</i> granulovirus.
HcNPV:	<i>Hyphantria cunea</i> nucleopolyhedrovirus.
HLH:	Helix-loop-helix.
hr:	Hour.
hrs:	Homologous regions.
HTH:	Helix-turn-helix.
HLH	Helix-loop-helix.
H _z -1:	<i>Heliothis zea</i> virus1.
H _z NPV:	<i>Helicoverpa zea</i> nucleopolyhedrovirus.
ICTV:	International Committee for Taxonomy of Viruses.
IE:	Immediate-early.
kb:	Kilo base pair.
kDa:	Kilo Dalton.
kV:	Kilo volt.
L:	Late.
Ld(M)NPV:	<i>Lymantria dispar</i> (multicapsid) nucleopolyhedrovirus.
LEF:	Late expression factor.
M:	Molar
McNPV:	<i>Mamestra configurata</i> nucleopolyhedrovirus.
µg:	Microgram.
min:	Minute
ml:	Millilitre.
mM:	Milli molar.
mRNA:	Messenger RNA

nm:	Nanometer.
NPV:	Nucleopolyhedrovirus.
nt:	Nucleotide.
NTR:	Non-translated region.
ODV:	Occlusion-derived virus.
Op(M)NPV:	<i>Orgyia pseudotsugata</i> (multicapsid) nucleopolyhedrovirus.
ORF:	Open Reading Frame.
PCR:	Polymerase Chain Reaction.
PhopGV:	<i>Phthorimaea operculella</i> granulovirus.
PK:	Protein Kinase
PnNPV:	<i>Perina nuda</i> nucleopolyhedrovirus.
PxGV:	<i>Plutella xylostella</i> granulovirus.
p.i.:	Post infection.
RNA:	Ribonucleic acid
RoNPV:	<i>Rachiplusia ou</i> nucleopolyhedrovirus.
RT-PCR:	Reverse transcriptase PCR
SDS:	Sodium dodecyl sulfate
SDS-PAGE:	Sodium dodecyl sulfate- polyacrylamid gel electrophoresis
SeNPV:	<i>Spodoptera exigua</i> nucleopolyhedrovirus.
SINPV:	<i>Spodoptera litura</i> nucleopolyhedrovirus.
SOD:	Superoxide dismutase.
TnGV:	<i>Trichoplusia ni</i> granulovirus.
Tn(S)NPV:	<i>Trichoplusia ni</i> (single) capsid nucleopolyhedrovirus.
UTR:	Untranslated region.
XcGV:	<i>Xestia c-nigrum</i> granulovirus.
VL:	Very late.
UV:	Ultraviolet.

RESUMÉ

L'utilisation des insecticides biologiques à risque réduit est une alternative acceptable aux insecticides chimiques. En effet, les problèmes associés à ces derniers, notamment au niveau de l'émergence d'insectes résistants, la réduction de la biodiversité de l'entomofaune et les risques environnementaux, sont minimisés voir même absents lors de l'utilisation des insecticides biologiques. Ce projet s'intéresse au granulovirus *Choristoneura fumiferana* un granulovirus qui possède des caractéristiques intéressantes comme agent de lutte biologique. Les baculovirus possèdent un spectre d'hôte limité puisqu'ils infectent principalement les insectes de l'ordre des lépidoptères. Le ChfuGV se retrouve spécifiquement chez la tordeuse des bourgeons de l'épinette, le principal ravageur entomologique des forêts résineuses de l'Est du Canada. Cet insecte représente une menace économique importante pour l'industrie canadienne du bois. Les populations de la tordeuse des bourgeons de l'épinette sont caractérisées par des épidémies cycliques. Une meilleure compréhension des mécanismes impliqués lors de l'infection des insectes par le ChfuGV devrait permettre d'élaborer des prescriptions sur son emploi à grande échelle dans la lutte contre les populations de la tordeuse des bourgeons de l'épinette.

Les interactions entre les virus et les cellules hôtes lors des infections virales sont largement dépendantes du rôle de certaines protéines. Dans ce contexte, une caractérisation des protéines impliquées dans le processus d'infection des larves de la tordeuse des bourgeons de l'épinette par le ChfuGV pourrait fournir des éléments importants permettant d'expliquer la spécificité de cette entomopathogène. Cependant, très peu d'informations sont connues sur la composition génomique et protéique du ChfuGV.

Le génome baculoviral est constitué d'ADN double brin super-enroulé, variant entre 80 et 230 kb et codant pour environ 100 protéines. L'ADN est emmagasiné à l'intérieur d'une nucléocapside protéique laquelle est entourée d'une enveloppe phospholipidique contenant des glycoprotéines transmembranaires. De plus, les baculovirus possèdent aussi une couche appelée tégument, située entre la nucléocapside et l'enveloppe.

Dans le cycle viral des baculovirus, deux phénotypes structurellement distincts sont essentiels pour la réplication virale. Il s'agit du phénotype BV qui est un virus bourgeonnant qui est présent dans la première phase d'infection du virus et du phénotype ODV qui est un virus inclus dans une matrice protéique nommée granuline (polyhédrine dans le cas des NPV) qui est produit lors de la deuxième phase de l'infection. Les ODV peuvent persister longtemps dans l'environnement et sont responsables de la propagation du virus d'un insecte à l'autre et les BV sont responsables de la propagation du virus de cellule en cellules à l'intérieur de l'insecte.

Deux protéines majeures de la nucléocapside ont attiré notre attention lors de cette étude. Il s'agit des protéines ODVP-6e/ODV-E56 et p74. Il s'agit de protéines hautement conservées chez les baculovirus et qui sont associées avec l'enveloppe de phénotype ODV. Les protéines de la nucléocapside enveloppée ODV-6e/ODV-E56 et p74 ont été détectées par une migration sur un gel SDS-PAGE et les bandes de masses de 39 et 74 kDa ont été isolées, purifiées et les séquences partielles ont été obtenues par spectrométrie de masse. En se servant du programme BLAST et FASTA 3, ces séquences partielles ont

été comparées aux séquences connues des banques de données. Les séquences primaires similaires récupérées correspondaient à celles des protéines ODVP-6e/ODV-E56 et p74 retrouvées chez d'autres baculovirus. Des amorces dégénérées oligonucléotidiques ont été dessinées à partir des séquences primaires afin d'identifier et isoler, par PCR, les gènes codant pour les protéines ODVP-6e/ODV-E56 et p74 sur les fragments de restriction de l'ADN génomique du ChfuGV. Le gène codant pour la protéine ODVP-6e/ODV-E56 a été localisé sur le fragment subgénomique *Bam*HI de 11kb et le gène codant pour la protéine p74 a été retrouvé sur le fragment subgénomique *Bam*HI de 8,9 kb. Tous les ORF situés à l'intérieur de la région du gènes *odvp-6e/odv-e56* et *p74* du granulovirus de *Choristoneura fumiferana* ont été identifiés par séquençage du fragments de restriction *Bam*HI de 11 kb et 8.9 kb sur le génome du ChfuGV. Les séquences des tous les ORFs conservé à l'intérieur de la région *odvp-6e/odv-e56* et p74 sont comparées aux ORFs équivalents localisés au niveau de la même région dans d'autres GV.

PREFACE

Very little is known about the molecular basics of *Choristoneura fumiferana* granulovirus (ChfuGV); to date, the only group with published results on molecular characterization of ChfuGV is Dr. Guertin's group at INRS-Institut Armand-Frappier. To my knowledge, except this group no other research group has any published work on this subject. The main reason for this infrequency may be explained by the absence of a permissive cell line for ChfuGV, meaning all manipulations must be performed *in vivo*. Such *in vivo* manipulation, required for studying ChfuGV at molecular levels, necessitates a laboratory fully equipped with insect larvae rearing facilities along with the usual molecular virology lab equipments. In addition to the equipment requirements, experimental studies on *in vivo* models are labor intensive and require specialized staff for insect manipulation, which are the two main reasons for considering *in vivo* studies less appealing.

Most of the work done on the members of the Baculoviridae family is focused on viruses for which there is at least one permissive cell line available. All these obstacles have left the molecular aspects of ChfuGV almost untouched. At the time that the present project was launched, only two articles on molecular characterization of ChfuGV existed in the literature. Dr. Guertin's group was eager to continue his studies on molecular characterization of ChfuGV when I joined his group in 1999. The original idea was to explore the proteins in ChfuGV that can play important roles during the infection cycle of the virus. One of the main interests was to characterize the major envelope proteins which may play a role during the attachment stage of the virion to the midgut cell membrane.

The quest began with isolation of major envelope proteins of ChfuGV in order to be used for protein sequencing. Four major baculoviral structural proteins, that were considered to play a role during the initiation stage of virus infection cycle, were chosen and subsequently sequenced. At the same time, a genomic library of ChfuGV was constructed for the first time, in order to be used for localizing and sequencing the genes encoding the abovementioned proteins. The same structural proteins were also used to raise monospecific antibodies, a powerful tool for studies involving protein expression pattern as well as immunolocalization of the proteins in the virus ultrastructure. Monospecific antibodies react with only a single specified antigen or antigenic determinant.

The project began with the microsequencing of four structural proteins. The protein sequences were then used to design degenerate oligonucleotidic primers to screen for clones containing the genes coding for these proteins. After several essays using several different degenerate primers, two clones containing *odv-e56* and p74 genes were determined, and by sequencing these clones (11.5kb and 8.9kb), the sequence of the *odv-e56* and p74 along with other important genes located on the same region were determined.

In order to help the reader to appreciate the complicatedness involved in working with ChfuGV, some of the methods and techniques used in this project will be presented here. One of the major time and labor consuming steps in working with ChfuGV is producing the virus *in vivo*. The most important aspects of the virus production process are: (i) choice of the host, (ii) rearing conditions, (iii) and virus purification. Usually, the natural virus host is the best choice for virus production, but in cases when the natural host is not

suitable for laboratory rearing conditions (for example when the natural host has a special dietary requirement, or when long obligatory diapauses are required), alternative hosts may be considered. In the case of this study, the natural host was used.

Temperature and humidity are the most important aspects in insect rearing, with other key factors including the number of larvae per each diet container and the size of the container. Spruce budworm larvae demonstrate no cannibalism activity, so both large and small size containers can be used in ChfuGV production. In this project, small sized containers with the capacity of 10 larvae per container were employed. The multiplication process of virus consists of three steps. In the first step, the first-instar larvae are placed on diet in incubator in order to reach the fourth-instar, which is the most susceptible instar of *Choristoneura fumiferana* to ChfuGV infection. On the next level, fourth-instar larvae are infected with a suspension of ChfuGV and are re-incubated on fresh nutrition source until the appearance of the first granulose symptoms. In the final step, the virus is extracted from the infected larvae. During the study, more than 100 rounds of insect rearing and virus isolation were performed in order to obtain enough material to continue the work.

The virus production begins always with rearing insect larvae. The first instar larvae are placed on an agar based media containing sources of carbohydrates, proteins, vitamins and a large spectrum antibiotic. Larvae are placed in an incubator at 24 °C, 55% of relative humidity and 16:8 hrs (day:night) photoperiod. An incubation time of 6 to 9 days is required for the larvae to be reached the fourth-instar.

As soon as the larvae are ready infection of insects with ChfuGV must be carried out. Viral suspension at a concentration between 2×10^4 and 4×10^4 granules per ml is prepared by diluting a concentrated stock suspension of ChfuGV in sterilized double distilled water. Plastic containers containing fresh diet media are prepared and each container received 50 μ l of this viral suspension. The media are left under the laminar hood until the viral suspension is fully absorbed to the surface of the media. The fourth-instar larvae are then transferred from old containers to the fresh virus contaminated media, and are incubated as previously described. Usually, 4 to 6 days of incubation are required in order to observe the granulose symptoms and the infected larvae become pale pink. The infected larvae are stored at -70°C until the whole batch is ready for virus isolation.

To isolate and purify the ChfuGV from infected *C. fumiferana* larvae, frozen infected larvae are transferred to a SDS solution (1%) and homogenized using an industrial blender. The homogenized infected larvae suspension in SDS is passed through 3 layers of cheese cotton to eliminate the insect debris. The filtrate is then centrifuged at 300Xg for 5 min in order to further eliminate the debris. The supernatant is then centrifuged at 10000Xg for 25 min to sediment all ChfuGV particles and other remaining larval tissue debris. The pellet is then transferred to a 0.5M NaCl solution and incubated for 30 min at room temperature. Another round of centrifugation at 10000Xg for 25 min is performed, followed by rinsing the pellet with sterilized double distilled water. The viral suspension

in double distilled water is then centrifuged in 10000Xg for 15 min and the washing step is repeated three times.

After these washings steps, the pellet is resuspended in sterilized double distilled water. To purify the granulovirus, the pellet is ultracentrifuged at 19000Xg on a 70-100% glycerol gradient for 45 min. The band containing the ChfuGV (visible in the region between 70-80%) is extracted and transferred to another ultracentrifugation tube containing sterilized double distilled water. To isolate the virus, another round of centrifugation at 1000Xg for a period of 20 min is required. The pellet containing the high number of ChfuGV is resuspended in a minimal volume of sterilized double distilled water. To verify the purity of the final granulovirus suspension and to calculate the concentration of ChfuGV, a thorough examination using transmission electronic microscopy (TEM) is required.

Obtaining clean and pure genomic DNA is one of the chief concerns in any project dealing with the molecular characterization of granuloviral genome and any failure to generate pure genomic DNA may cause unpredictable obstacles for this purpose. Hence, extraction of ChfuGV genomic DNA from occluded virion is an extremely important step in the realization of this project. Several of the existing DNA extraction kits commercially available were tested for extraction of ChfuGV genomic DNA from occluded virion, but none yielded genomic DNA with acceptable purity for cloning. The complex structure of occluded virion and the presence of the relatively substantial amount of matrix protein (granulin) which covers the enveloped nucleocapsid is the main

reason that the usual protocols do not work, and this makes the extraction of ChfuGV genomic DNA a labor intensive process. To ensure the extraction of pure DNA, the enveloped nucleocapsid of ChfuGV must first be fully isolated from the occlusion body before any attempt to isolate the genomic DNA and subsequently the final stage of genomic DNA extraction must be realized on purified ChfuGV enveloped nucleocapsid.

In the purification procedure for ChfuGV enveloped nucleocapsid, the adjustment of an effective method to purify high concentration ChfuGV enveloped nucleocapsids was one of the key concerns of this project. Without enough ChfuGV enveloped nucleocapsid, the process of genomic DNA extraction and isolation of structural proteins for protein sequencing would have encountered a major set-back. After months of trial and error, a protocol was established which significantly increases the yield of enveloped nucleocapsid compared to the existed protocol. ChfuGV enveloped nucleocapsids were purified using the following protocol. Occluded virions were first incubated in 0.5 M sodium carbonate (pH 10) for 60 min. The suspension is then centrifuged at 8000Xg for 15 min and the supernatant is placed on a sucrose gradient (15-35%) before being spun at 25000Xg during 90 min. ChfuGV enveloped nucleocapsid can be isolated from the band located between 25-30%. The band containing ChfuGV enveloped nucleocapsid is transferred to sterilized double distilled water for another round of ultracentrifugation at 100000Xg for 36 h. The pellet is then resuspended in TE buffer with pH=7.5 (Tris HCl 0.01 M and EDTA 0.001 M). The quality of the isolated enveloped nucleocapsid in suspension is then verified using TEM. This suspension is stored at -70 °C before being used for DNA extraction.

To compare the ChfuGV ODV associated enveloped nucleocapsid proteins with BV- (Budded virion) -associated enveloped nucleocapsid proteins, a great deal of effort was expended to isolate the BV phenotype from the infected larvae fat tissue at various hours post-infection. In the case of baculoviruses with a permissive cell line, the BV phenotype can be isolated from infected cell through a relatively easy and routine procedure. In the case of ChfuGV, due to the lack of such a cell line, the only way to obtain the BV phenotype is the isolation from the fat tissue at different hours post-infection during the late infection cycle prior to appearance of the granulose symptoms. Despite several efforts, I never succeeded in isolating BV phenotype from the infected larvae.

In order to carry out the genomic DNA extraction procedure from purified ChfuGV enveloped nucleocapsid, enveloped nucleocapsids are incubated in proteinase K (250 $\mu\text{g/mL}$) in the presence of SDS (0.5%) at 37 °C for 120 min. Subsequently, enveloped nucleocapsid were stripped from their proteinous shell by phenol and chloroform treatment. DNA was then precipitated by ethanol (100%) in the presence of sodium acetate (0.3M, pH 5.2) at -20 °C and centrifugation at 16000Xg for 30 min. The pellet was resuspended in TE buffer (pH 7) and stored at -70 °C.

For preparation of the ChfuGV genomic library, the genomic DNA was digested using several restriction enzymes (*Bam*HI, *Hind* III, *Eco*RI). Digestion of genomic DNA using *Bam*HI produced the minimum number of fragments compared to the other restriction

enzymes, therefore *Bam*HI restriction fragments were chosen to construct the ChfuGV genomic bank in pBlueScript-SK⁺ (Stratagene).

For ChfuGV structural protein isolation and purification enveloped nucleocapsid is isolated from occluded virion using the technique explained above. The enveloped nucleocapsid is then disrupted in an equal volume of Laemmli sample buffer (Laemmli, 1970) in presence of β -mercaptoethanol (BioRad). The samples are then boiled for 5 min and clarified at 13,000Xg for 5 min before electrophoresis on 10% sodium dodecyl sulfate- polyacrylamide gel electrophoresis (SDS-PAGE). Several concentrations of separating gels were used to isolate structural proteins with different molecular weights. Proteins were visualized using light Coomassie Brilliant Blue (Invitrogen) staining and the molecular mass (*Mr*) of virion proteins was determined by comparison to known standards (BioRad).

Gel purified ChfuGV structural proteins were used for protein sequencing. Major protein bands from ChfuGV enveloped nucleocapsid preparations were excised from the gel and subjected to protein sequencing analysis at the Harvard Microchemistry facility. The facility used either microcapillary reverse-phase HPLC nano-electrospray tandem mass spectrometry (μ LC/MS/MS) on a Finnigan LCQ quadrupole ion trap mass spectrometer or chemical sequencing (Edman). The peptide sequences were used to design degenerate primers (with the least redundancy) in order to screen the ChfuGV genomic bank.

To determine the localization and the sequence of the genes coding for the studied structural proteins, using different combinations of degenerate primers, PCR amplifications were performed on ChfuGV genomic DNA. Several rounds of PCR reactions were performed in order to obtain the desirable results. All PCR reactions were carried out using *Taq* DNA polymerase (Amersham Pharmacia Biotech) in the buffer supplied by manufacturer supplemented with 200 μ M of each dNTP and $MgCl_2$ to a final concentration of 2.5 mM. PCR reactions were heated for 10 min at 95°C prior to amplification (30 cycles of 95 °C, 30 sec; 45 °C 30 sec; and 72 °C, 30 sec). Amplified fragments were cloned in a PCR 2.1 (Invitrogene) cloning vector and sequenced, and the sequence was compared to homologues in GenBank/EMBL using the BLAST algorithm (Altshul *et al.*, 1992). In order to locate the genes of interest on ChfuGV genome, subgenomic fragments were generated by *Bam*HI endonuclease and fractionated on 1% agarose gel. Gel extracted fragments were then used as templates for PCR reactions using the designed set of degenerate primers. The restriction fragments were cloned in pBlueScript-SK⁺ cloning vector (Stratagene), and transformed into *E. coli* XL-1 Blue (Stratagene). All the manipulations, including restriction enzyme digestion, agarose gel electrophoresis, transformation, colony lifting and plasmid purification were carried out according to standard protocols (Sambrook *et al.*, 1989). The fragments of interest were subsequently sequenced from both extremities on either strand using a primer walking technique. More than 300 primers have been designed to complete these sequencing.

Four gel purified ChfuGV structural proteins of interest in this project were used for production of monospecific polyclonal antibodies. For preparation of antisera, purified

enveloped nucleocapsids were electrophoresed on SDS-PAGE. Gels were then lightly (10 min) stained in 0.05 % Coomassie brilliant blue (Sigma) in order to locate the positions of the protein bands in the gels. After destaining, bands of interest were excised from the gels and washed in deionized water three times each for 10 min followed by an overnight wash. Following the washes, the preparation of proteins was injected into rabbits as described by Harlow and Lane (1988). Briefly, the gel pieces were extruded through the barrel of one 5ml syringe with an equal volume of deionized water into another 5-ml syringe. Antisera were produced in 6 weeks, male New Zealand rabbits by subcutaneous injection of the each antigen at multiple sites along the back. All initial injections were performed with Freund's complete adjuvant and subsequent injections were done with incomplete Freund's adjuvant.

To adjust the effective concentration of antibodies and test their specificity, all antisera were subjected to ELISA in order to obtain their effectiveness against ChfuGV enveloped nucleocapsids and to establish their effective dilutions. The appropriate dilutions were then used to examine the specificity of the obtained antisera using Western immunoblotting assays in which several ChfuGV enveloped nucleocapsids protein gels were electrophoretically transferred to a nitrocellulose membrane (Millipore). The membranes were then examined for the presence of proteins by staining the membrane with Ponceau S (Sigma). After washing with PBS buffer, membranes were blocked in PBS containing 3% BSA or 5% non fat milk for 1 hr in room temperature followed by 2h incubation with appropriate antisera concentration. All membranes were then washed in

PBS (pH 7.2), and reacted with the peroxidase-conjugated secondary antibody and visualized using a chemiluminescence substrate (Amersham).

In order to verify if antibodies against ODV-E56 and p74 are able to reduce the viral infectivity toward the larvae, different concentrations of each antiserum were separately added to the purified suspension of ODV phenotype ChfuGV enveloped nucleocapsid. The general idea was based on the inactivation of these ODV associated envelope proteins (thought to play major roles in the attachment process) and measure the rate of the mortality in infected cells. These tests usually work with permissive cell lines and once more the lack of an appropriate cell line obliged me to use the fourth-instar larvae as subjects for this experiment. The injection method has been used for establishment of activity of ODV associated non-occluded enveloped nucleocapside neutralized by anti ODV-E56 and anti-p74. This method is very tedious and time consuming but has two main advantages: (i) the amount of inoculums per insect is known, and (ii) the time of the beginning of infection is known. Despite my continued efforts to establish an effective way to perform such an experience *in vivo*, all the efforts were unsuccessful and these assays remained fruitless. To eliminate the possibility that the injection method was the reason for the failure of the experience, another approach was tested where the neutralized ODV associated non-occluded were introduced to the fourth-instar larvae using a diet contaminated with known quantities of GV preparation. Bioassays using contaminated artificial diets are the most commonly used assays for many insect viruses including GVs. In this method, known quantities of GVs are incorporated into or layered on the surface of artificial diets, which allow the evaluation of the LC_{50} . There are two

very important advantages related to this method: (i) early stages of insect larvae, which are generally the most susceptible to viral infection, can be used in large numbers, (ii) insect handling is minimized since the stay in the same container throughout the bioassay. Disadvantage of this method is that the dose of the virus ingested by each insect is not known. These assays did not lead to any relevant results.

An important question that may arise for the reader is the following: Why should one want to study ChfuGV molecular biology while other similar studies has been already done on other members of *Baculaviridae* family such as AcMNPV? The response to this question lies in the development of a new bioinsecticide against spruce budworm using ChfuGV. This lead us to know more about the nature of this virus and to its molecular characterization. Several major differences existed between the baculoviruses belonging to NPV genus and GV genus and the increasing and continuous demand for insecticide agents that are harmless and safe for human and other animals with a very narrow host range are the main reasons that focus our studies on the viruses belonging to this genus. The results of the present study can be considered as the foundation for other researchers to further characterize this virus at a molecular level. The data presented in this study can be, also, considered very important for any worker in this field with an interest in the development of genetically modified ChfuGV particles with more virulence and a shorter infection cycle. The phylogenetic analyses presented here are not only important in molecular evolution and taxonomical perspective, but also they can be considered as valuable tools to be used in any attempt to formulate a novel insecticide that with ChfuGV as its active ingredient by providing valuable comparative studies on ChfuGV

and other GVs that are currently used as active ingredient in commercially available bioinsecticides. Laboratory bioassays with ChfuGV demonstrated its high pathogenicity for spruce budworm population. Formulation of granuviral-based insecticides is a very important part in production. For large-scale applications, different aspects such as storage stability and UV protection must be considered in order to have a stable and high quality product. Studies which shed more light on the molecular basis of ChfuGV infection cycle is highly appreciated by the people who try to determine an effective formulation for ChfuGV as a final product to be use in large scale.

At the end of this preface it seems necessary to point out a few points which will facilitate the reading of this thesis. The term nuclear disintegration used in this text indicates the destruction of insect cell nucleus at the very late stages of baculoviral infection. All bioinformatics studies presented in this study were performed on genomic DNA of ChfuGV, therefore the terms 5' or 3' untranslated /non translated regions (UTR/NTR) used in this text are referred to the nucleotide sequence on genomic DNA and not the mRNA. Baculovirus transcripts do not undergo splicing therefore the nucleotide sequence on genomic DNA can be used as a reference in genomic studies. Any report on the detection of any motifs or promoter regions can be considered simply as reporting a fact indicating the existence of these sequences and that these sequences were detected upstream to the initiation codons of the genes on genomic DNA. To avoid any confusion here I present the definitions by which the genomic results are usually reported: The coding region begins with the initiation codon, which is normally ATG. It ends with one of three termination codons: TAA, TAG or TGA. On either side of the

coding region are DNA sequences that are transcribed but are not translated. These untranslated (non translated) regions or non-coding regions often contain regulatory elements that control protein synthesis. Both UTR (NTR) regions and the coding sequence is transcribed from the template DNA strand and the genomic results usually are reported using the complementary strand which has the same sequence as mRNA. Promoter regions contain factor binding sites and other regulatory elements in the upstream region of the genes in promoter region. Although promoter regions vary, it is usually possible to find a small DNA sequence (called the consensus sequence) which is common between all of them. The report on the detection of baculoviral early and late promoter motifs within the 5'NTR regions different genes on ChfuGV genomic DNA does not prove that all these motifs are definitely involved in transcription process and merely reports their existence in these regions of ChfuGV genomic DNA. I hope that the clarifications presented in this section will help the reader to achieve a better understanding on the results presented in this thesis.

INTRODUCTION

Baculoviridae is a large family of viruses that naturally infects arthropods. Its utilization as a biological control agent made it an interesting family of virus to study during the past several decades. More recently another interesting aspect of their infection cycle, their gene expression, has attracted a major scientific attention and put them in the spot light as one of the most efficient eukaryotic gene expression systems.

In eastern North America, *Choristoneura fumiferana* (spruce budworm) is considered the most destructive insect for conifers. The spruce budworm is a big economic threat to vast forest areas (60 million ha) in Canada and the eastern United States. Maritime Provinces (New Brunswick, Nova Scotia, Newfoundland), Quebec, Ontario, and the Great Lake states are the areas that were affected most extensively by spruce budworm outbreaks. Spruce budworm larvae feed on a number of conifers, but balsam fir (*Abies balsamea* [L.] Mill.), white spruce (*Picea glauca* [Moench] Voss), and red spruce (*Picea rubens* Sarg.) are the major hosts in Eastern North America. Species occasionally attacked include black spruce (*Picea mariana* [Mill.] B. S. P.), eastern hemlock (*Tsuga canadensis* [L.] Carr.), tamarack (*Larix laricina* [Du Roi] K. Koch), and white pine (*Pinus strobus* L.). In Quebec, the outbreaks of *C. fumiferana* usually affect huge forest areas.

Defoliation, inhibition of seed production, cone mortality and tree mortality are the most important impacts of spruce budworm on trees. Defoliation caused by spruce budworm decrease the growth rates of trees and this decline can last several years. When outbreaks occur, the weaker trees usually die after three or four years of heavy defoliation and most

of the trees die between six and ten years after the first attack. Even when the spruce budworm population returns to its endemic level, the damaged trees continue to die (Morin *et al.*, 1993; Bergeron *et al.*, 1995).

Chemical insecticides were the most common possible way to protect spruce-fir forests from the spruce budworm from 1927 up to the 1970s. DDT and Phosphophamidon were used mostly between 1944 and 1970. In the 70s and 80s, DDT was replaced by Organophosphates and Carbamats. Most of these compounds are toxic for human and warm-blood animals. The concern about finding an alternative for chemicals started during 1960s and one of the candidates was the use of natural occurring insect's pathogens. As insects continue to gain resistance to chemical pesticides industrial interest in commercial development of biological pesticides increased in U.S. and Europe (Moscardi, 1999).

Today we have several groups of biological pesticides that are in use against insect pests in agriculture and forestry. *Choristoneura fumiferana* granulovirus (ChfuGV) can be considered an interesting viral agent to be used against spruce budworm population. One of the major tasks to be fulfilled before using this viral agent in the field is to expand our current knowledge about ChfuGV at molecular level.

CHAPTER I – LITERATURE REVIEW

1.1. Viral Bioinsecticides

Natural predators, parasites, competitors and pathogen microorganisms like fungi, bacteria and viruses have been used as biological insecticides. Some of these microbiological agents used as an insecticide like *Bacillus thuringiensis* (B.t.) and baculoviruses are natural occurring insect pathogens. Insect viruses provide one of the most interesting solutions for pest-specific control because of their specificity. Like most other animals, insects are susceptible to diseases caused by viruses and viral diseases is one the various factors that hold insect populations in a natural balance. Viral diseases are detected in both beneficial and pest species, the latter can be used as an alternative in pest control management (Moscardi, 1999).

A variety of viruses differing markedly in morphological, biological, biochemical and molecular characteristics infect insects. There are different families of viruses which have been isolated from insects: *Baculoviridae*, *Poxviridae*, *Reoviridae*, *Picornaviridae*, *Bunyaviridae*, *Polydnaviridae*, *Ascoviruses*, *Tetraviridae* and *Nodaviridae*. These insect viruses may infect several insect species. Entomopoxviruses and nucleopolyhedroviruses (NPVs) infect insects belonging to various insect orders. In contrast infection by granuloviruses (GVs) is mostly limited to lepidopterous species. Generally, insect viruses attack larval forms and cause symptoms ranging from acute infection and insect death to unapparent infection (Adams and McClintock, 1991; Tanada and Hess, 1991; Miller, 2001).

Several insect viruses are closely related to vertebrate viruses both in structure and molecular properties. In particular, the entomopoxviruses show striking similarities to viruses from vertebrate hosts (Bergoin and Dales, 1971). In contrast to other insect virus families, members of the *Baculoviridae* (the GVs and NPVs) seem to be restricted to invertebrate hosts (Blissard *et al.*, 1996; Miller, 2001). No morphologically similar counterpart has been isolated in vertebrates. Of the insect viruses, baculoviruses are highly recommended for field use. A limited host range is an important consideration in the safe implantation of baculoviruses as insecticides. Baculoviruses are significant pathogens of arthropods, especially insects, belonging mainly to the orders *Lepidoptera*, *Diptera* and *Hymenoptera*. Baculoviruses can cause serious and fatal diseases in these insects (Summers, 1975a, b; Adams and McClintock, 1991; Tanada and Hess, 1991; Miller, 2001).

1.2. Baculovirus

Awareness in the use of baculoviruses as pest-control biological agents exist since 1527, when the baculovirus disease of insects was found during studies on the ‘jaundice disease’ of the silkworm, *Bombyx mori*. The viral nature of the disease was recognized by 1947 and before long it became apparent that these viruses were widespread in nature among economically important insect pests, and therefore could be potentially valuable in pest-management in agricultural and forestry practices (Benz, 1986).

The name Baculovirus is derived from the Latin word *baculum*, meaning “rod” or “stick”. They have received considerable attention because of their potential for use as

insecticides. Baculoviruses are large rod-shaped and have an enveloped nucleocapsid which contain a large circular, covalently closed, double-stranded supercoiled DNA genome of 80 to 230 kilobase pairs (kbp) ((Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*,. 2003). The family is subdivided into two genera based on occlusion body morphology: the nucleopolyhedroviruses (NPVs) and granuloviruses (GVs) (Volkman *et al.*, 1995; Miller, 2001). Baculoviruses of both genera are occluded in a crystalline protein matrix composed of a protein of about 29 kDa. This matrix protein is termed polyhedrin for NPVs and granulin for GV. In NPVs, the polyhedral occlusion bodies contain numerous virions but granules in GVs usually contain only one virion. The protein matrix is a feature unique to certain invertebrate viruses and is thought to contribute to their stability. Occlusion within these crystals stabilizes the virion, allowing them to tolerate harsh environmental conditions and remain infectious for many years. The granulin and polyhedrin both are alkali soluble, thus after their ingestion by host larvae they dissolve by the alkaline pH of the insect midgut and release virions. The shape of granules has been described variously as oval, ellipsoidal, ovocylindrical or ovoid (Adams and McClintock, 1991; Tanada and Hess, 1991; Miller, 2001).

1.2.1. Budded and Occluded virions

During baculovirus life cycle, two different phenotypes for virions are observed, each having a specific role during the baculovirus infection cycle. The budded virus or BV

phenotype is the virion that is produced during the early stages of infection. BVs are capable to spread the infection from cell to cell within the insect body. On the other hand, occlusion-derived virion (ODV) phenotype is produced during the late and very late stages of infection in the nucleus of the infected cells. Due to the presence of polyhedrin or granulin, ODVs are capable of persisting for a long period of time in the environment. (Volkman and Summers, 1977; Blissard, 1996; Miller, 2001)

After ingestion by a susceptible insect, occlusion bodies are dissolved in the alkali pH of insect's midgut and release ODVs which penetrate the midgut cells and initiate the infection. Both BVs and ODVs have envelopes that cover the nucleocapsids. These envelopes are derived from two different sources. The major proteins of envelope in BV is a glycoprotein that is synthesized and transported to the host cell plasma membrane during both early and late phase of infection (McIntosh, 1975; Monsma *et al.*, 1996; Oomens and Blissard, 1999; Washburn *et al.*, 2003).

Nucleocapsids destined to become BVs are transported through the cytoplasm toward plasma membrane and then bud out through the modified plasma membrane. During the budding process the modified membrane envelopes the nucleocapsids. On the other hand, ODVs acquire their envelope inside the host cell nucleus and subsequently they are surrounded by granulin (or polyhedrin), become occluded, and are released upon the death of the cell (Blissard and Rohrmann, 1989; Blissard 1996; Miller, 2001).

1.2.2. Genome Structure of Baculovirus

Baculoviruses have large genomes (80-230 kbp) that have the potential to encode about 100 genes (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003). On the contrary of other viruses with big genome like poxvirus that carry an extensive array of enzyme which are essential for early gene transcription, baculoviruses carry no virion-associated proteins that are essential for virus early gene transcription on their genome (Kelly and Wang, 1981; Blissard 1996; Miller, 2001).

The genome of baculoviruses is composed almost entirely of unique DNA sequences, though several small repeated sequences known as homologous regions (hrs) are known in the DNA (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002). The homologous regions are shown to have roles as enhancers for early genes expression (Miller, 2001), and also as origin of DNA replication (Kool *et al.*, 1994; Miller, 2001). The activation property of an early gene known as *ie-1* is enhanced when the genes are linked to homologous region sequences. The basic unit of each homologous region in AcMNPV and BmNPV is a conserved 60-bp motif, which includes a highly conserved 28-bp motif imperfect inverted repeat. The product of *ie-1* gene is one of the most essential regulatory

proteins which play an indispensable role in baculovirus infection cycle (Kovacs *et al.*, 1992; Forsythe *et al.*, 1998; Leisy and Rohrmann, 2000; Olson *et al.*, 2001).

Open reading frames are on either strand of the DNA. Most ORFs are separated by 2-200 bps of DNA rich in A+T. There are also some overlapping ORFs in baculovirus genome, usually termination codon UAA overlaps with the primary polyadenylation signal AAUAAA. Some promoters are located within the neighboring ORFs. Frequently transcripts of one gene initiate within into or through neighboring ORFs. Beside, partial clustering of genes which have assigned roles in early gene regulation, *ie-1*, *ie-2*, and *pe-38*, genes in the genome of baculoviruses do not appear to be clustered. Genes encoding structural proteins are distributed throughout the genome with no obvious pattern to the location (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003).

1.2.3. Genes involved in virus replication in the host cell

Baculovirus genes are expressed in a transcriptional cascade in which each successive phase is dependant on the expression of genes during the previous phase. There are currently four generally accepted categories of baculovirus genes; immediate early, delayed early, late and very late (Adams and McClintock, 1991; Tanada and Hess, 1991; Blissard 1996; Miller, 2001).

During the early phase, the molecular environment and other aspects of intracellular environments of the cell is altered in preparation for the replication and expression of viral DNA. Early genes have promoters that can be recognized by host cell RNA polymerase. There is also a category of delayed early genes that are transactivated by a regulatory protein (IE), these genes need the presence of IE proteins to be efficiently transcribed (Guarino and Summers, 1986; Nissen and Friesen, 1989; Kovacs *et al.*, 1992; Pullen, and Friesen, 1995; Leisy and Rohrmann, 2000). Baculoviruses genes encoding *trans*-activating factors have been identified and enhancer sequence elements are located in several NPV and GV genomes (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Miller, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003).

These regulatory proteins contribute to the speed and facility with which the virus is able to direct the host cell metabolic machinery to transcribe particular genes necessary for viral replication. The homologous regions form a complex directly or indirectly with a multifunctional viral regulatory protein, IE-1 and enhance the early gene transcription (Leisy and Rohrmann, 2000; Olson *et al.*, 2001; Miller, 2001).

Genes that are transcribed only after DNA replication begins are classified as late genes. After replication of DNA most late genes are actively transcribed but the levels of expression decline at later times. Expression of *granulin* and *polyhedrin* genes is initially delayed but subsequently reaches extremely high levels very late in the infection

(Blissard 1996; Miller, 2001). The other protein that is abundantly expressed at both late and very late times post-infection is p10 (Roelvink *et al.*, 1992) which is a small poorly conserved protein that may be involved in occlusion body formation or cell lysis (Gross *et al.*, 1994; van Oers *et al.*, 1993; Zuidema *et al.*, 1993). Generally, the very late phase of gene expression follows late gene expression and is characterized by the hyper expression of two genes, *granulin/polyhedrin* and *p10* (Miller, 2001).

Late genes transcribed by an α -amanitin-resistant RNA-polymerase that recognize the baculovirus late promoter (Yang *et al.*, 1991; Huh and Weaver, 1990), which normally contains the sequence motif ATAAG (Rohrmann, 1986), the first nucleotide of the motif is sometimes a G or rarely a T but TAAG sequence is invariant. The late mRNAs normally initiate at the T or A residue of the TAAG sequence. The very late genes (*granulin/polyhedrin* and *p10*) begin to be transcribed well after the initiation of late genes transcription, and transcription continues very late into the infection after gene transcription ceases (Hill-Perkins and Possee, 1990; Blissard 1996). Very late promoters appear to be akin to late promoters in that they also contain the conserved core TAAG motif and flanking sequences, but vary in that they also need an supplementary sequence called a "burst" sequence. The extraordinarily high levels of transcription from the *granulin/polyhedrin* and *p10* genes appear to be regulated or mediated by binding of viral protein known as very late factor 1 (VLF-1) to the "burst" DNA sequence, downstream of the transcription start site. The burst sequence is so called since it appears to regulate the burst of very late transcription (McLachlin and Miller, 1994; Todd *et al.*, 1995; Yang and Miller, 1998; Yang and Miller, 1999).

1.2.4. Genes involved in baculovirus DNA replication

The same *cis*- and *trans*- acting factors that are involved in stimulation of early transcription are also involved in DNA replication. In particular: *i*) homologous regions (hrs) sequences that serve as enhancers of transcription also appear to function as replication origins and *ii*) IE-1, which is required for baculovirus DNA replication, also transactivates early gene expression (Tanada and Hess, 1991; Adams and McClintock, 1991; Kool *et al.*, 1994; Ahrens and Rohrmann, 1995a, b; Blissard 1996).

At least ten essential and stimulatory genes involved in DNA replication are known in baculoviruses (Kool *et al.* 1994; Ayres *et al.*, 1994; Lu and Miller 1995; Ahrens and Rohrmann 1995 a, b; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Miller, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003). Among those gene products IE-1 is one of the most important regulatory proteins shown to be essential for DNA replication. IE-1 plays this critical role by activation of the expression of other regulatory genes to the levels sufficient to initiate DNA replication (Kovacs *et al.*, 1992; Forsythe *et al.*, 1998; Leisy and Rohrmann, 2000).

Other important regulatory genes involved in baculoviral DNA replication are late expression factors. The transcription of late genes requires a number of early gene products. Using transient late expression assays genes necessary for transient transcription of baculoviral 19 late promoter have been identified. These studies revealed

the existence of 19 late expression factors (*lef*) (Passarelli and Miller, 1993 a, b; Kool *et al.*, 1994; Ayres *et al.*, 1994; Lu and Miller 1995; Hang *et al.* 1995; Ahrens *et al.*, 1997; Rapp *et al.*, 1998; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Li *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002; Lange and Jehle, 2003; Wormleaton *et al.*, 2003). Transient assays for late gene expression and DNA replication have proved to be extremely important tools for the identification of genes associated with DNA replication and late gene expression. However, because viral proteins are expressed transiently from plasmid constructs in these assays, it is likely that the regulation of expression of each protein differs substantially from its expression in the context of a viral infection. *lef-1* and *lef-2* are directly involved in DNA replication and *lef-3* is involved in the production of single-stranded DNA-binding protein that contribute to the opening and unwinding of duplex DNA at an origin of replication (Kornberg and Baker, 1992; Passarelli and Miller, 1993 a, b; Miller, 2001) and *lef-7* contains two single-stranded DNA binding motifs and is shown to stimulate DNA replication (Lu and Miller, 1995). These genes were originally identified as required for the late gene transcription. *lef-11* was originally identified as a gene required for efficient transcription from a late promoter in a transient late expression assay (Todd *et al.*, 1995). More recent studies demonstrate that the *lef-11* gene appears to be indispensable for baculoviral DNA replication throughout the infection cycle (Lin and Blissard. 2002).

The p35 gene is another important gene for replication identified which block cellular apoptosis (Clem *et al.*, 1991). There is other data that suggest p35 may also act as a transcriptional activator of early genes (Gong and Guarino, 1994).

1.3. Granulovirus

Paillot first detected granulosis in 1926, in the larvae *Pieris brassica* (Paillot, 1926; Tanada and Hess, 1991). Since this initial description several lepidopterous species have been reported to be susceptible to GVs, including some of the most serious insect pests of agricultural crops and forests. The high pathogenicity of GVs for these and other economically important insects has been the main reason for the widespread interest in this group of viruses and has led to their consideration for use as biological insecticides (Tanada and Hess, 1991).

Some GVs have already been used as insecticides to control pests. In initial laboratory studies during the 1950s, the GV of *P. rapae* was used successfully against early instar larvae (Kelsey, 1957; Tanada and Hess, 1991). In 1970s the *P.rapae* GV played a significant role in reducing populations of this destructive insect in Canada (Jaques, 1973; Tanada and Hess, 1991). Large-scale applications of GVs in pest management programs also were used for the first time in the former Soviet Union. In this country GVs were used against the cotton cutworm and cabbage looper. In Canada GVs were used for the first time in a large scale against the fir budworm (*C. murinana*) to limit defoliation by this insect (Schonherr, 1969; Tanada and Hess, 1991).

The feasibility of using GVs as insecticides also has been confirmed by comparing the effectiveness of these viruses with that of chemical insecticides. First field studies in 70s with chemicals and GV of the green clover worm (*Plathypena scabra*) on soybeans showed no significant differences between the two methods of control (Beegle *et al.*, 1973; Tanada and Hess, 1991). In addition to being highly virulent to their hosts, GVs have many other characteristics that make them ideally suitable for pest control; as it was mentioned before infection by GVs is mostly limited to the order Lepidoptera and in most cases species-specific (Tanada and Hess, 1991). Because of this selectivity, GVs can be used more precisely in pest management programs, without the determinable effect on nontarget organisms.

GVs are stable and can be stored as aqueous suspensions or dried powders for long periods of time without any loss of activity (David and Gardiner, 1967; Tanada and Hess, 1991). On the other hand, GVs deposited on foliage are inactivated rapidly by exposure to the ultraviolet radiation in direct sunlight (David *et al.*, 1968, David, 1969; Tanada and Hess, 1991). To protect the viruses against the UV radiation the best solution is using optical brighteners with viruses in field applications. Optical brighteners are routinely used in many industrial processes for brightening paints, fibers, coating etc. The UV protection efficiency of some fluorescent brighteners is very high (Dougherty *et al.*, 1996).

1.3.1. Granulovirus replication

GV infection is initiated when an insect ingests an occluded virus and virus is transported to the lumen of the midgut. With the disruption of the protein matrix in alkaline environment of insect gut, enveloped nucleocapsids that are the infectious form of GVs are released and become associated with the microvilli of midgut columnar cells. Summers (1969, 1971) found nucleocapsids without envelopes in microvilli, and then Kawanishi *et al.* (1972) and Tanada *et al.* (1975) suggested fusion between nucleocapsid envelope and microvillar membrane as the probable means of virus entry into the gut cell cytoplasm.

Penetration of GVs into the midgut cells leads to primary infection in which the virus that are responsible for the secondary infection of fat bodies are generated. Studies on *Trichoplusia ni* revealed a unique mechanism for the uncoating of GV nucleocapsids at the nuclear membrane (Tanada and Hess, 1991; Miller, 2001). This study showed, at 2 to 6 hours post-infection (h pi) intact nucleocapsids frequently are associated with nuclear pores. Empty or partially empty capsids are then observed outside the nucleus, suggesting that only the viral genome was inserted into the nucleus through nuclear pores without the capsid. A similar association of nucleocapsids with the nuclear pores was observed in infections caused by NPVs either (Raghow and Grace, 1974; Tanda and Hess, 1976; Tanada and Hess, 1991).

After an eclipse period of approximately 12 to 18 h in GV-infected *T. ni* larvae, progeny nucleocapsids are detected in the midgut cell nuclei in regions of dense material distinct

from host chromatin (Summers, 1969; Tanada and Hess, 1991). Nuclear membranes become disorganized and disrupted in some infected cells. Changes in nuclear membranes lead to release of some nucleocapsids into cytoplasm that have loose-fitting envelopes originated from nuclear membranes (Tanada and Leutenegger, 1970; Tanada and Hess, 1991).

Enveloped nucleocapsids appear to be incorporated into vacuole-like structures in the cytoplasm, in which they are transported to the basal lamina of midgut cells. The enveloped nucleocapsids are then released from the vacuoles into the hemocoel (Hunter *et al.*, 1975). These newly replicated virus particles initiate secondary infection in fat body cells and other susceptible cells. Tracheolar cells, which are situated near the basal lamina of midgut cells, are frequently infected about 48 h after the primary infection (Harrap *et al.*, 1977). Budded baculoviruses enter fat body cells by a typical receptor-mediated endocytosis mechanism. Nucleocapsids are then released from the endosome in the result of the fusion of the BV envelope and endosome membrane. This is a pH-dependant fusion and the major glycoprotein of the BV's envelope seems to play a major role in this process (Volkman and Goldsmith, 1985).

The sequence of events during GV replication in the fat body is different than sequence observed in midgut epithelial cells. The three major differences are following: i- the yields of virus per cell are much higher: ii- the number of infected cells is much higher; iii- occlusion of the virions takes place at the end of the cycle (Tanada and Hess, 1991). After entrance of virus into the fat body cells, the size of nucleus demonstrates an

increase in size and at the mean time host cell chromatin redistributes to the periphery of the nuclear membrane. This is followed by development of an electron-dense network of chromatin-like material in the nucleus (Tanada and Hess, 1991). This network has been first termed the virogenic stroma by Xeros in 1956 (Xeros, 1956; Tanada and Hess, 1991). While this stroma develops the host chromatin degenerates and become finally disappear. Viral DNA synthesis and nucleocapsid assembly are taking place at this site. In some insect species GV-infected fat body cell nuclear membranes may disintegrate and viral DNA mix with cytoplasmic contents. In some species nucleocapsid assembly continues in the former nuclear area of the cell, after the break down of the nuclear membrane (Benz and Wager, 1971).

Several mechanisms of acquisition of envelopes by nucleocapsids in fat body cells have been observed in infected cells. Some nucleocapsids obtain envelopes by *de novo* synthesis (Summer, 1971), others obtain it by budding through nuclear or cytoplasmic membranes. Studies on NPVs demonstrate that the initial association of nucleocapsid and envelope begins at the capped end of the nucleocapsid (Adams *et al.*, 1977).

Formation of granulin on outer surface of enveloped nucleocapsid starts as the stroma is depleted. Protein matrix formation around enveloped nucleocapsids is a selective process, as no other cellular components are occluded in matrix protein. The crystallization appears to begin either at one side or at one end of the envelope and proceeds around the particle (Pinnock and Hess, 1978; Tanada and Hess, 1991). The majority of enveloped nucleocapsids inside fat body cells become occluded in the matrix, but a few of them

remain nonoccluded and though to be the source for further infection, along with nucleocapsids budding through the plasma membrane (Summer and Volkman, 1976).

In the final stages of an infection, infected cells are completely packed with GV and eventually the cells rupture and liberate occluded virions into the hemocoel. The high concentration of GVs changes the color of fat body to opaque white, and color of hemolymph become turbid and milky. The color of infected larvae becomes pinkish to milky. Finally, infected larvae die and the occluded viruses released into the environment (Tanada and Hess, 1991).

1.4. Virus ultrastructure

In baculoviruses ODVs the structure contains one or more nucleocapsids in the shape of a slightly curved rod enclosed in an envelope. This in turn is enclosed in a protein matrix (granulin or polyhedrin) that itself is sometimes enclosed in an outer membrane. The size of GVs is averaging 300 to 500 nm in length by 120 to 350 nm in width. These viruses consist of rod-shaped nucleocapsids surrounded by envelopes, and each enveloped nucleocapsid is embedded within a matrix of protein, which has a regular crystalline lattice. GV particles in their embedded form generally are referred to as occluded enveloped nucleocapsids. The terms “occlusion body” and “capsule” have been used to indicate the entire GV structure, consisting of nucleocapsid, envelope, and protein matrix (Tanada and Hess, 1991, Miller, 2001).

The rod shaped nucleocapsid in GVs is a cylindrical core made of DNA and proteinous capsid and has average dimensions of 30 to 60 by 260 to 360 nm. This structure is capped

at both ends. Helical symmetry such as what we see in plant viruses cannot be seen in GVs and NPVs (Shikata, 1977). Each viral capsid is made up of subunits assembled in a regular lattice. Diffraction studies and electron microscopic measurements have shown that lattice consists of rings of subunits stacked on top of one another (Beaton and Filshie, 1976; Tanada and Hess, 1991). Each ring in the capsid of *Spodoptera frugiperda* granulovirus composed of approximately 12 subunits (Burley *et al.*, 1982). Stacked series of rings of subunits is forming the capsid, these rings aligned perpendicular to the longitudinal axis. A number of reports exist that shows the opposite ends of the nucleocapsids are morphologically distinct, a blunt end and conical end (Tanada and Hess, 1991). Conical end has polarity and contacts with the plasma membrane of infected cells during the process of budding (Hess and Falcon, 1987).

1.5. Structural Proteins in Baculoviruses

Analysis of structural proteins by SDS-PAGE revealed that baculoviruses have a complex structure with approximately 30 polypeptides in their ODV and BV structures (Braunagel and Summers, 1994). Several proteins have been demonstrated to be specific to ODV or BV. Some proteins are common to both ODV and BV structures (Rohrmann, 1992; Hong *et al.*, 1994; Braunagel *et al.*, 1996; Theilmann *et al.*, 1996).

As mentioned before, polyhedrin and granulin are the major component of polyhedra and granule (capsule) in NPVs and GVs, respectively. Polyhedrin and granulin form a protective crystal around the nucleocapsid and resist solubilization except under strongly alkaline conditions similar to those found in the insect midgut. Although polyhedrin and

granulin can vary by about 50% in amino acid sequence, many of their structural features are highly conserved (Bah *et al.*, 1997). This reflects the similar function and biochemical properties of these proteins. The size of these proteins is remarkably stable (29 kDa), and the number of amino acids is 247 to 248 aa in granulin and 243 to 246 aa in polyhedrin. In *Choristoneura fumiferana* granulovirus the *granulin* gene encodes a protein of 248 amino acids with a predicted *Mr* of 29 kDa (Bah *et al.*, 1997).

Polyhedrin is a hyper expressed protein constituting up to 18% or more of total cellular alkali-soluble protein late in infection (Quant *et al.*, 1984). The N-terminal region of the proteins (amino acids 1 to 16) is highly conserved within each but not between baculovirus genera. At the N-terminus the granulins have two or three additional amino acids and a conserved cysteine which are absent in polyhedrin, whereas polyhedrin has a conserved proline (amino acid 11) not found in granulin (Bah *et al.*, 1997). The region near amino acid 210 also differs between the two proteins and contains two conserved prolines in granulin, one of which is shared with polyhedrin (Rohrmann *et al.*, 1982; Rohrmann *et al.*, 1981; Rohrmann *et al.*, 1980; Kozlov *et al.*, 1981). These differences could produce very different N-terminal protein structures between the two genera. Limited degree of glycosylation and phosphorylation has been reported, but evidence for substantial modification is limited. Quantitative studies showed the presence of 0.5 to 1 phosphate group per polyhedrin molecule (Vlak and Rohrmann, 1985; Rohrmann *et al.*, 1980). It has been postulated that the crystallization of polyhedrin occurs by salt bridge formation between acidic and basic residues. Such ionic bonds would be disrupted by a strongly alkaline pH (Vlak and Rohrmann, 1985).

There is a highly conserved sequence of 12 nucleotides (AATAAGTATTTT) is known in the flanking region of *polyhedrin* gene. This highly conserved sequence is followed by an AT-rich sequence consisting of from 75 to 89% A+T whereas the 100 nt upstream from the 12-mer are 47% to 65% A+T. The conserved 12-mer begins 20 to 70 nt upstream the protein start codon. The same highly conserved 12-mer is also known for flanking region of another hyper expressed late protein (p10) in baculoviruses (Rohrmann, 1986). This highly conserved sequence of nucleotides also can be detected in flanking region of *granulin* genes from position -18 to -31 (TTTATAAGGAATTT) of five GVs (ChfuGV, ClGV, CpGV, PbGV and TnGV) a consensus baculovirus late promoter (ATAAG) can be seen within this conserved sequence (Bah *et al.*, 1997). An electron-dense envelope which has been termed the PE or polyhedron calyx surrounds polyhedra. The PE of one NPV is composed of protein (Whitt and Manning 1988).

1.5.1. Baculovirus Major Capsid Proteins

Baculoviral nucleocapsid contains several conserved proteins (Miller, 2001). One of the major baculoviral capsid proteins is p39. This protein is a major capsid component of both ODV and BV phenotypes of baculoviruses. The p39 protein has a predicted molecular weight consistent with that measured on Western blots of SDS-PAGE gels. This protein is slightly charged and contains high percentages of asparagine, leucine, alanine, valine, and arginine residues (Pearson *et al.*, 1987).

Immunoelectron microscopy confirms that the p39 is a component of the capsid and shows that it is randomly distributed over the surface of nucleocapsid (Russell *et al.*,

1991). The gene coding for this protein has been identified in nucleopolyhedroviruses; AcMNPV (Thiem and Miller, 1989a), OpMNPV (Blissard *et al.*, 1989) and Ld (*Lymanttria dispar*) MNPV (Bjornson and Rohrmann, 1992a). The AcMNPV and OpMNPV P39 amino acid sequences are 59% identical, and they are 39% and 47% identical, respectively, to the LdMNPV p39 sequence (Bjornson and Rohrmann, 1992).

During infection p39 can be detected in the cytoplasm of infected cells at 24 h pi, and by 48 h pi the p39 can be detected primarily in the nuclei of infected cells (Thiem and Miller, 1989b). These evidences and other evidences from Northern blot hybridization analysis, suggest that between 18 and 36 h pi the p39 protein is synthesized at a high rate and most p39 remains in the cytoplasm, and between 36 and 48h pi the production of p39 decrease and p39 is transported to the nucleus (Miller, 2001).

Another important capsid associated protein is p87. This protein was first identified as another capsid protein in OpMNPV and AcMNPV (Lu and Carstens, 1991; Muller *et al.*, 1990). The evidences from immunofluorescence microscopy has demonstrated that p87, like p39, is concentrated in the nucleus of OpMNPV in the late stage of the infection. The p87 ORF encodes a protein of predicted *Mr* 71 kDa. It has been speculated that this difference is caused by some unknown post-translational modifications (Muller *et al.*, 1990), but no evidence of glycosylation of this protein has been found and the reason of this difference is not known yet. RNA analysis showed that it was a late gene and immunofluorescence microscopical studies showed that P87 was concentrated in the nucleus late in OpMNPV infection (Muller *et al.*, 1990).

A protein homologue to p87 is a protein with molecular mass of 82 kDa, has been identified in *Choristoneura fumiferana* MNPV. Amino acid sequence comparisons revealed that ORF p82 had high homology with the ORF p87 of OpMNPV (Xing *et al.*, 1997).

The protein p24 is one of the proteins that likely participate in forming the capsid structure (Wolgamat *et al.*, 1993). Both Western blot analysis on AcMNPV and OpMNPV and immunoelectron microscopic examination of OpMNPV infected cells indicate that p24 is associated with nucleocapsid structure. p24 in AcMNPV and OpMNPV is coded by the first open reading frame in a series of five late-expressed ORFs in the polyhedron envelope protein gene region. (Gombart *et al.*, 1989; Oellig *et al.*, 1987). A homologue of the p24 gene has also been identified in LdMNPV. In contrast to AcMNPV and OpMNPV in which p24 is a late-expressed gene, the p24 gene of LdMNPV lacks a late promoter element, evidences show that p24 is described as an early gene in LdMNPV (Bjonson and Rohrmann, 1992b).

Another baculoviral capsid-associated protein is a 91 kDa protein (p91) (Russell and Rohrman, 1997). Immunoelectron microscopy observations indicated that P91 is present in capsids and envelopes surrounding the capsid of ODVs. Fractional studies using a nonionic detergent NP-40 to remove the virion envelopes, demonstrated that p91 behave like other major proteins of the capsid and can not be released from virion-containing pellet fraction by detergent treatment. Therefore, p91 may either be a capsid protein or may be part of envelope that is closely linked to the capsid (Russell and Rohrman, 1997).

Two late promoter elements, GTAAG and TTAAG were found upstream of the p91 ATG in OpMNPV genome. Results from temporal expression of p91 demonstrated that an immunoreactive band of approximately 91 kDa can be detected at 18 hours post infection (hr p.i.). This band increased significantly in intensity by 36 h pi and remained constant for the duration of the infection (Russell and Rohrman, 1997).

The p6.9 protein is an arginine-serine rich protein that is associated with nucleocapsids and may be involved in condensation of viral DNA prior to or during encapsidation (Wilson and Miller, 1986; Wilson, 1988; Wilson *et al.*, 1987). The gene coding for this basic protein has been isolated in NPV (Wilson *et al.*, 1987). Transcription of this gene starts late in infection and it is closely associated with the viral DNA (Tweeten *et al.*, 1980; Wilson *et al.*, 1987). This protein is the substrate for the kinase activity which catalyses the transfer of phosphate to both serine and arginine residues of p6.9. Once the protein kinase is activated the DNA is released from the nucleocapsid (Wilson and Consigli, 1985a, b).

It seems that the activation of protein kinase occurs when the nucleocapsid aligns with the nucleopore. It is speculated that the high concentration of Mn^{2+} within the nuclear membrane of the insect cell is involved in this process (Wilson and Consigli, 1985b). These speculations can only explain the process of uncoating of the nucleocapsid that occurs at the nucleopore, but the uncoating mechanism for the nucleocapsids that enter the nucleus may not be the same.

1.5.2. Baculovirus major envelope proteins

Envelopes of two baculovirus phenotypes (BV and ODV) are different. Both phenotypes have complex structures and several proteins have been identified as specific to ODV or BV (Rohrmann, 1992). BV envelopes are specialized for cell to cell infection, while ODV envelopes seems to have an affinity for granulin and polyhedrin and could form nuclei for granulin or polyhedrin crystallization around enveloped nucleocapsid within infected cell nuclei beside this it seems that ODV envelopes have a role in initiating the primary infection in insects midgut cells. In the insect midgut, ODVs are three to four orders of magnitude more infectious than the BVs, but when injected into the hemocoel BVs act at least two orders of magnitude more infectious than ODVs (Volkman and Summers, 1977; Keddie and Volkman, 1985).

The major protein found in association of BV envelope is termed GP64 or GP67 that has been characterized in baculoviruses (Whitford *et al.*, 1989; Blissard and Rohrmann, 1989). Immunoelectron and immunofluorescence microscopy indicate that GP64 concentrates at the plasma membrane (Blissard and Rohrmann, 1989). It is shown that palmitic acid is ester-linked to the AcMNPV (the prototype of *Baculoviridae* family) GP64 that makes it the major acylated protein associated with the BV. It is believed that acylation of proteins may be involved in anchoring to the membranes, membrane fusion or regulation of intracellular transport (Roberts and Faulkner, 1989). The protein is highly glycosylated and as much as 10 kDa of the *Mr* is contributed by glycosylation residues, and examination of the two amino acid sequences in AcMNPV and OpMNPV showed that they contain between five and seven sites for N-linked glycosylation (Blissard and

Rohrmann, 1989; Whitford *et al.*, 1989). Nucleocapsids migrate from nucleus into the cytoplasm and bud from the cell surface, they become enveloped within GP64 modified plasma membrane during the process of budding.

Transcriptional characterization of this gene in OpMNPV has demonstrated that the gp64 has only one early promoter and four late promoters (Blissard and Rohrmann, 1989). The presence of an early promoter permit the expression of the genes in early stages so the plasma membrane is modified in advance and of the production of nucleocapsids, and the late expression of this gene ensure the production of this protein during entire cycle of virion production. Synthesis of GP64 in late stages of infection allows the replacement of GP64 in plasma membrane after depletion caused by BV production.

At least eight proteins have been known to be specific to ODV phenotype. Those are: VP17 (Funk and Consigli, 1993), ODV-E25 (Russell and Rohrmann, 1993), ODV-E35 (Braunagel *et al.*, 1996a), GP41 (Whitford and Faulkner, 1993), p74 (Kuzio *et al.*, 1989), ODV-E18 (Braunagel *et al.*, 1996a), ODV-E66 (Hong *et al.*, 1994), and ODVP-6E/ODV-E56 (Braunagel *et al.*, 1996b, Theilmann *et al.*, 1996). Proteins that could participate in adsorption, fusion and penetration have not been identified but any of abovementioned 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 baculovirus the gene coding for ODV-associated envelope protein p74 is a species-specific gene (Wu *et al.*, 2003). It has been shown that in AcMNPV p74 protein is

exposed on the virion surface (Faulkner *et al.*, 1997) and has an essential role in the infectivity of the virus (Kuzio *et al.* 1989; Faulkner *et al.*, 1997). These evidences place this protein in special position to be considered as an ODV envelope protein that may play a major role in adsorption, and/or penetration of virion. It has been suggested that the N terminal of p74 is located outside of the ODV envelope and the C terminal act as a transmembrane anchor (Faulkner *et al.*, 1997). The significance of hydrophobic C terminal of p74 in localization and transmembrane anchoring of the protein in AcMNPV was later established (Slack *et al.*, 2001).

Another ODV specific envelope protein that may play a role during the attachment / penetration phase is ODVP-6E/ODV-E56 (Braunagel *et al.*, 1996b; Theilmann *et al.*, 1996). ODVP-6E/ODV-E56 is a highly conserved protein and has been identified in several baculoviruses (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002 Lange and Jehle, 2003; Wormleaton *et al.*, 2003). Study on ODVP-6E/ODV-E56 in NPVs demonstrated that this protein is a transmembrane protein (Braunagel *et al.*, 1996b). ODV-E18 is another conserved baculoviral ODV-associated envelope protein. ORF coding for this protein has been detected on the genome of several members of *Baculoviridae* family (Ayres *et al.*, 1994; Ahrens *et al.*, 1997; Gomi *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Hayakawa *et al.*, 1999; Hashimoto *et al.*, 2000; Luque *et al.*, 2001; Pang *et al.*, 2001; Chen *et al.* 2002; Cheng *et al.*, 2002; Li *et al.*, 2002 Lange and Jehle, 2003; Wormleaton *et al.*, 2003). ODV-E18 has been characterized as a late protein

in AcMNPV associated with both the ODV envelope and capsid (Braunagel *et al.*, 1996a).

Little is known about the ChfuGV regulatory and structural proteins as well as its genomic architecture and there is almost no major publication in regard to genomic characterization of ChfuGV. For us in order to be able to use ChfuGV as an effective biological insecticide we must intensify our efforts to discover more about molecular aspects of this virus. This knowledge may lead to a better understanding about the interaction between the virus and its host which, hopefully, allows us to develop an effective biological agent to control *Choristoneura fumiferana* populations. Results presented in this document can be considered as the most extensive study ever done on characterization of ChfuGV at genomic level and will open several doors to future molecular studies on this virus.

CHAPTER 2 – PUBLICATION #1

Identification, Characterization and Phylogenic Analysis of Conserved Genes within the *odvp-6e/odv-e56* Gene Region of *Choristoneura fumiferana* Granulovirus

Authors: Kianoush Khajeh Rashidan^{1*}, Nasha Nassoury², Paresa N. Giannopoulos¹, Yves Mauffette³ and Claude Guertin¹

1- Institut national de la recherche scientifique-Institut Armand-Frappier, 531 Blvd des Prairies, H7V 1B7, Laval-Canada

2- Université de Montréal, 4101 Sherbrooke east, H1X 2B2, Montreal-Canada

3-Université du Québec a Montréal, Succursale Centre Ville CP8888, Montreal-Canada

The nucleotide sequences data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers AF58690, AF389081, AF406758, AF406760, AY064489.

Published in Journal of Biochemistry and Molecular Biology in 2004. Number 37, pages 206 to 212.

Contribution of the authors

I was the principal investigator in this report while Dr. Nassoury's expertise in screening to locate the clone was used to complete this report. Mrs. Giannopoulos helped me by performing some of the PCR reactions. Dr. Mauffette was the co-director of this project with whom I had several scientific discussions prior the submission of this paper. Dr. Guertin was the director of the project and the majority of the experiences have been realized in his laboratory.

Abstract in French :

Identification, caractérisation et analyse phylogénétique des gènes conservés dans la région du gène *odvp-6e/odv-e56* du granulovirus de *Choristoneura fumiferana*

Les gènes localisés dans la région du gène *odvp-6e/odv-e56* du granulovirus de *Choristoneura fumiferana* ont été identifiés par séquençage d'un fragment de 11 kb d'une banque *Bam*HI contenant le génome entier du ChfuGV. Le contenu en GC, calculé à partir des données obtenues de cette région génomique, était de 34,96%. Les séquences des ORF dans la région *odvp-6e/odv-e56* sont présentées et comparées aux gènes équivalents localisés au niveau de la même région chez d'autres GV. Cette région renferme 14 cadres de lecture ouverts (open reading frames, ORFs). Parmi ceux-ci, 11 ORF semblent correspondre à des ORF homologues granuloviraux tels que les gènes *gramuline*, *CfORF2*, *pk-1*, *ie-1*, *odv-e18*, *p49* et *odvp-6e/odv-e56*. En plus de ces ORF, trois ORF uniques au ChfuGV ont été identifiés, et ne correspondaient à aucun autre ORF de baculovirus. Les ORF conservés codent pour trois protéines structurales et quatre protéines régulatrices. Pour chaque ORF conservé, les arbres phylogénétiques ont été construits en utilisant les séquences conceptuelles des protéines. Les analyses phylogénétiques ont démontré que le ChfuGV semble d'avantage associé au CpGV et au PhopGV formant le I des GV par rapport au groupe II qui regroupe les GV de XcGV et de PxGV). Les études comparatives sur l'arrangement des gènes indiquent une similitude dans l'arrangement des gènes des GV du groupe I.

2.1. Abstract

The genes located within the *odvp-6e/odv-e56* region of *Choristoneura fumiferana* granulovirus (ChfuGV) were identified by sequencing an 11 kb *Bam*HI restriction fragment on the ChfuGV genome. The global GC content calculated from the data obtained from this genomic region was 34.96%. The open reading frames (ORFs) located within *odvp-6e/odv-e56* region are presented and compared to the equivalent ORFs located at the same region in other GVs. This region is composed of 14 ORFs including three ORFs unique to ChfuGV with no obvious homologues in other baculoviruses as well as eleven ORFs with homologues to granuloviral ORFs such as granulin, CfORF2, *pk-1*, *ie-1*, *odv-e18*, *p49* and *odvp-6e/odv-e56*. The conceptual products of seven major conserved ORFs, granulin, CfORF2, IE-1, ODV-E18, p49 and ODVP-6E/ODV-E56, were used in order to construct phylogenetic trees in this study. Our results show that granuloviruses can be grouped in 2 distinct groups. Group I: *Choristoneura fumiferana* granulovirus (ChfuGV), *Cydia pomonella* granulovirus (CpGV), *Phthorimaea operculella* granulovirus (PhopGV) and *Adoxophyes orana* granulovirus (AoGV). Group II: *Xestia c-nigrum* granulovirus (XcGV), *Plutella xylostella* granulovirus (PxGV) and *Trichoplusia ni* granulovirus (TnGV). ChfuGV conserved proteins are most closely related to those of CpGV, PhopGV and AoGV. Comparative studies performed on gene arrangement within this region of genome demonstrated that three GVs from group I maintain similar gene arrangements.

Key words: Baculovirus, *Choristoneura fumiferana* granulovirus, phylogeny, gene arrangement.

2.2. Introduction

At least eight families of insect viruses are known, but the viruses most commonly used as viral bioinsecticides are viruses from *Baculoviridae* family. Baculoviruses are a worldwide distributed group of viruses that are specific to arthropods and are divided in two genera: the nucleopolyhedroviruses (NPVs), and the granuloviruses (GVs). Baculoviruses are characterized by the presence of a large protein matrix or occlusion body, which encase the viral particles. Baculovirus genome is composed of a double stranded circular DNA.

The high pathogenicity of GV's toward different insect pests of agricultural crops and forests, make this group of viruses a very attractive candidate to be used as biological insecticides. Since 1950s, different GV's have been used as biological insecticides in different parts of the globe. In a method of classification based on tissue tropism, GV's are placed in three groups (Federici, 1997). TypeI: GV's infecting the fat body (this group of GV's kill the insect host in a rather long process); TypeII: GV's infecting the most of the insect host tissues in a faster mode than TypeI GV's; TypeIII: To date *Harrisina brillians* granulovirus (HbGV) is the only known member of this type that infects only the midgut cells.

Granuloviruses have large genomes (80-180kbp) with the potential of encoding about 100 genes. Other viruses with big genome like poxvirus carry an extensive array of enzymes which are essential for early gene transcription, on the contrary, granuloviruses like other baculoviruses do not carry any virion-associated proteins.

Genome of granuloviruses is composed almost of unique DNA sequences, though several small repeated sequences known as homologous regions (*hrs*) are known in their DNA. Open reading frames are on either strand of the DNA. Most ORFs are separated by 2-200 bps. There are also some overlapping ORFs in granulovirus genome, usually the termination codon TAA overlaps with the primary polyadenylation signal AATAAA. Some promoters are located within the neighbouring ORFs. Frequently transcripts of one gene initiate within or through neighbouring ORFs (Hayakawa et al, 1999; Hashimoto *et al.*, 2000; Luque *et al.* 2001).

To date only four granuloviral genomes are completely sequenced, *Xestia c-nigrum* granulovirus (XcGV) (Hayakawa et al, 1999), *Plutella xylostella* granulovirus (PxGV) (Hashimoto *et al.*, 2000), *Cydia pomonella* granulovirus (CpGV) (Luque *et al.* 2001) and *Phthorimaea operculella* granulovirus (PhopGV) (Genbank/EMBL accession no.AF499596). Using the data available from GV and NPV genomes it is now possible to establish the relatedness of the members of *Baculoviridae* family. It has been suggested that baculoviruses may have a coevolutionary pathway with their insect host (Rohrmann, 1986) and this might attribute to the diversification of different groups of baculoviruses. *Choristoneura fumiferana* granulovirus (ChfuGV), which has been isolated from infected spruce budworm in several part of eastern Canada, is considered a very interesting alternative to be used as a biological insecticide against spruce budworm larvae.

In this study, the *odvp-6e/odv-e56* region on the ChfuGV genome have been characterized and compared to other GVs. The locations and transcriptional orientations

of ORFs within this region in all GVs have been presented. We used the available sequences of conserved genes located in this area and the encoded gene products to examine the phylogenetic relationship of ChfuGV to other GVs as well as NPVs.

2.3. Material and Methods

2.3.1. In vivo Production and Purification of Virus and DNA Extraction

Choristoneura fumiferana fourth-instar larvae have been infected by ChfuGV using virus contaminated artificial diet (Forté *et al.*, 1999). The virus isolation from infected larvae was carried out as previously described (Bah *et al.*, 1997).

2.3.2. DNA extraction, cloning and sequencing

ChfuGV genomic DNA was extracted from purified enveloped nucleocapsids as we have already described (Rashidan *et al.*, 2002). An 11 kb *Bam*HI restriction fragment were identified and cloned into a plasmidic vector and sequenced as explained previously (Rashidan *et al.*, 2002).

2.3.3. Computer Analysis

Sequence data assembly and analysis were performed with Sequencher software version 4.0.5 (Gene Codes Corp.) and MacVector program version 4.5.0 (Eastman Kodak).

Nucleotide sequence and its predicted amino acid sequence were compared to homologues in GenBank/EMBL and SWISSPROT by using BLAST (Altshul *et al.*, 1990). CLUSTALX (1.81) (Thompson *et al.*, 1997) was used for multiple amino acid sequence alignments. BOXSHADE version 3.21 was used for similarity shading and scoring among the aligned sequences. The alignments were used then as the input to construct the phylogenetic tree with Branch and Bound search settings of PAUP 4.0b4a employing Neighbor joining (NJ) method (Swofford, 2000). The reliability of the phylogenetic trees was evaluated by bootstrap analysis with 1000 bootstrap replicates.

2.3.4. Assignment of putative ORFs

The information obtained from sequencing the 11 kb genomic fragment of ChfuGV was used to detect homologous sequences in the Genbank database (whole database and restricted database containing only virus sequences) using BLASTN (Altshul *et al.*, 1990) and Fasta3 (Pearson and Lipman 1988) programs. Alternatively, putative ORFs were first translated from ChfuGV nucleotide sequences and amino acid sequences were compared with existing sequences in the Genbank database using BLASTP program (Altshul *et al.*, 1990). These data and the one obtained from the gene order from other granulovirus genomes were used to identify ChfuGV ORFs.

2.4. Results and Discussion

An 11 kb *Bam*HI restriction fragment of ChfuGV was cloned and sequenced. The global GC content calculated from the data obtained from this genomic region was 34.96%. Among fourteen detected putative open reading frames (ORFs), eleven ORFs were found to be coding for proteins with known homologues in other GVs while three of them were unique to ChfuGV (Table 1). The data obtained from sequencing of this fragment along with previous sequence data from the region upstream to ChfuGV granulin helped us to obtain useful information over the gene content, transcription orientation (Figure1) and codon preference of ChfuGV. The comparative studies on ORFs present within this region and their transcription orientations in ChfuGV and other GVs proved that closely related GVs maintain almost the same gene arrangement and transcription orientations (Figure1).

2.4.1. CfORF1 a homologue to granulin

Granulin along with polyhedrin are found in proteinous matrix of baculoviruses occlusion bodies. ChfuGV granulin contains 747 nucleotides (nt) encoding a protein of 248 amino acids with a molecular mass equal to 29.299 kDa. ChfuGV granulin seemed to be closely related to *Adoxophyes orana* granulovirus (AoGV, a Tortricidae-infecting GV) and is positioned in a cluster along with PhopGV (a Gelechiidae-infecting GV), CpGV (a Tortricidae-infecting GV), *Cryptophlebia leucotreta* granulovirus (ClGV, a Tortricidae-infecting GV) and HbGV (a Zygaenidae-infecting GV) apart from the cluster containing XcGV (a Noctuidae-infecting GV), PxGV (a Yponomeutidae-infecting GV) and *Trichoplusia ni* granulovirus (TnGV, a Noctuidae-infecting GV).

2.4.2. CfORF2 a homologue to CpORF2

A putative ORF of 381 nt is located immediately downstream of the granulin gene of ChfuGV. CfORF2 encodes 126 amino acids with an estimated molecular weight of 14.2 kDa. At the nucleotide level, the two putative TATA boxes and a baculovirus late promoter motif, ATAAG, were found to precede CfORF2 start codon. N-terminal region of CfORF2 demonstrates similarities with the N-terminus portion of the L11p protein family found in eukaryotic (*Guillardia theta*) and prokaryotic (*Thermotoga maritime*) cells. L11p proteins are ribosomal RNA binding proteins and bind directly to ribosomal RNA. A phylogenetic tree of CfORF2 product and the homologue ORFs in other GVs was constructed and the result places ChfuGV, PhopGV, CpGV and AoGV together in the same group.

2.4.3. CfORF3 a homologue to *pk-1*

ORF number 3 on the ChfuGV is a protein kinase homologue with the potential to encode a putative baculoviral protein kinase homologue with the predicted molecular mass of 32 kDa. Conceptual ChfuGV PK-1 protein akin to its homologues in other baculoviruses contains 11 conserved subdomains. These subdomains include motifs for serine/threonine protein kinases and ATP-binding. At nucleotide level an early element (CAGT) and a late consensus (TAAG) motif were detected within the 5'-non translated region (NTR) of ChfuGV *pk-1* gene. No obvious TATA element was observed in this region of ChfuGV *pk-1* gene. Phylogenetic analysis of protein kinase proteins separated GVs from NPVs, and demonstrates that ChfuGV *pk-1* is most closely related to the PhopGV and CpGV.

2.4.4. CfORF7 a homologue to IE-1

IE1, an immediate-early baculoviral gene product, is a multifunctional transactivator with the capacity to regulate expression from both viral and cellular gene promoters (Kovacs *et al.*, 1992). The ORF potentially encodes a protein of 426 amino acids homologue to baculoviral IE1 with an estimated molecular weight of 50.33 kDa. At the nucleotide level, one TATA element, two putative CCAAT elements, a (A/C/T)CA(G/T)T motif (proposed consensus for arthropod transcriptional initiator elements), and two baculoviral late promoter motifs (TAAG) were detected within the 5' NTR of this gene. A domain similar to basic helix-loop-helix like (bHLH-like) domain in NPVs was detected at the C-terminal region of IE-1 from ChfuGV (Rashidan *et al.*, 2002a). The phylogenetic estimation demonstrated an obvious separation between GVs and NPVs and showed that ChfuGV IE-1 is most closely related to that of CpGV and PhopGV.

2.4.5. CfORF12 a homologue to *odv-e18*

ODV-E18 is a conserved baculoviral ODV-associated envelope protein. ORF coding for this protein have been detected on the genome of several members of *Baculoviridae* family. ODV-E18 has been characterized as a late protein in AcMNPV associated with both the ODV envelope and capsid (Braunagel *et al.*, 1996). ORF number 12 on ChfuGV genome potentially codes for an ODV-E18 homologue. The presumed ChfuGV ODV-E18 contains 88 amino acids with an estimated molecular weight of 9.28 kDa. Two baculoviral late promoter motifs (TAAG) were detected within the 5' NTR of ChfuGV *odv-e18* gene at the positions –17 and –101 nt upstream the start codon. Phylogenetic analysis of baculoviral ODV-E18 protein separated GVs from NPVs and demonstrates

that ChfuGV ODV-E18 is most closely related to those of PhopGV and CpGV (Figure 2).

2.4.6. CfORF13 a homologue to *p49*

Three types of anti-apoptotic genes, *iap*, *p35* and *p49*, are known in baculoviruses. *p35* has the ability to suppress apoptosis induced by virus infection. A baculoviral protein known as *p49*, a homologue to *p35*, is a novel apoptosis suppressor gene with the ability to inhibit insect and human effector caspases by inhibiting an initiator caspase (Zoog *et al.*, 2002; Pei *et al.*, 2002). The *p49* homologue gene in ChfuGV contains 1368 nt encoding 455 amino acids with an estimated molecular weight of 53.720 kDa. ChfuGV *p49* gene contains a TATA element, and two baculoviral late promoter motifs (TAAG) within the 5' NTR at the positions –56 nt and –17 nt, respectively, from the first ATG. Result of phylogenetic studies showed ChfuGV *p49* protein along with CpGV and PhopGV *p49* proteins were placed in the same cluster.

2.4.7. CfORF14 a homologue to *odvp-6e/odv-e56*

ORF 14 on ChfuGV genome contains 1062 nt potentially encoding 353 amino acids, with an estimated molecular mass of 38.5 kDa. This ORF is a homologue to baculoviral ODVP-6E/ODV-E56 proteins. Baculoviral ODVP-6E/ODV-E56 protein is a highly conserved ODV envelope associated protein. At nucleotide level a TATA box and a 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 bp upstream to the first ATG. The gene contained a slight variant of a polyadenylation signal, AATAAT, at 3'

NTR at the position +10 nt downstream from termination signal. ChfuGV ODVP-6E/ODV-E56 contained two potential membrane-spanning regions at C-terminal and many putative *N*- and *O*- glycosylation, *N*-myristoylation, and phosphorylation sites (Rashidan *et al.*, 2002b). The phylogenetic estimation demonstrated that ChfuGV ODVP-6E/ODV-E56 is most closely related to those of CpGV and PhopGV.

2.4.8. Evolutionary Relatedness of granuloviruses

It is crucial to establish a relationship among different members of Baculoviridae family in order to extend our knowledge on the biology and the evolution history of this family of viruses. With the aim of establishing such a relationship between ChfuGV and other GVs and NPVs, we used two different approaches: *i*- phylogenetic study on several conserved genes in GVs and NPVs, *ii*- gene arrangement within a region that contains two genes coding for major highly conserved ODV-associated proteins (granulin and ODVP-6E/ODV-E56).

A true comprehension of the evolutionary history of these conserved structural proteins will help us to achieve a more accurate understanding of their defined role in the virus infection cycle. To date one of the major sources used as information to build phylogenetic trees was based mostly on the sequence of occlusion body matrix proteins granulin and polyhedrin. The validity of this approach was first questioned by Herniou and his associates (2001). Alternatively, a more plausible approach is to apply more than one conserved baculoviral protein in these studies. An approach similar to this has been

employed previously by Mitchell and his associates (2000) combining data from different genes for phylogenetic studies on Noctuoidea family (Insecta: Lepidoptera).

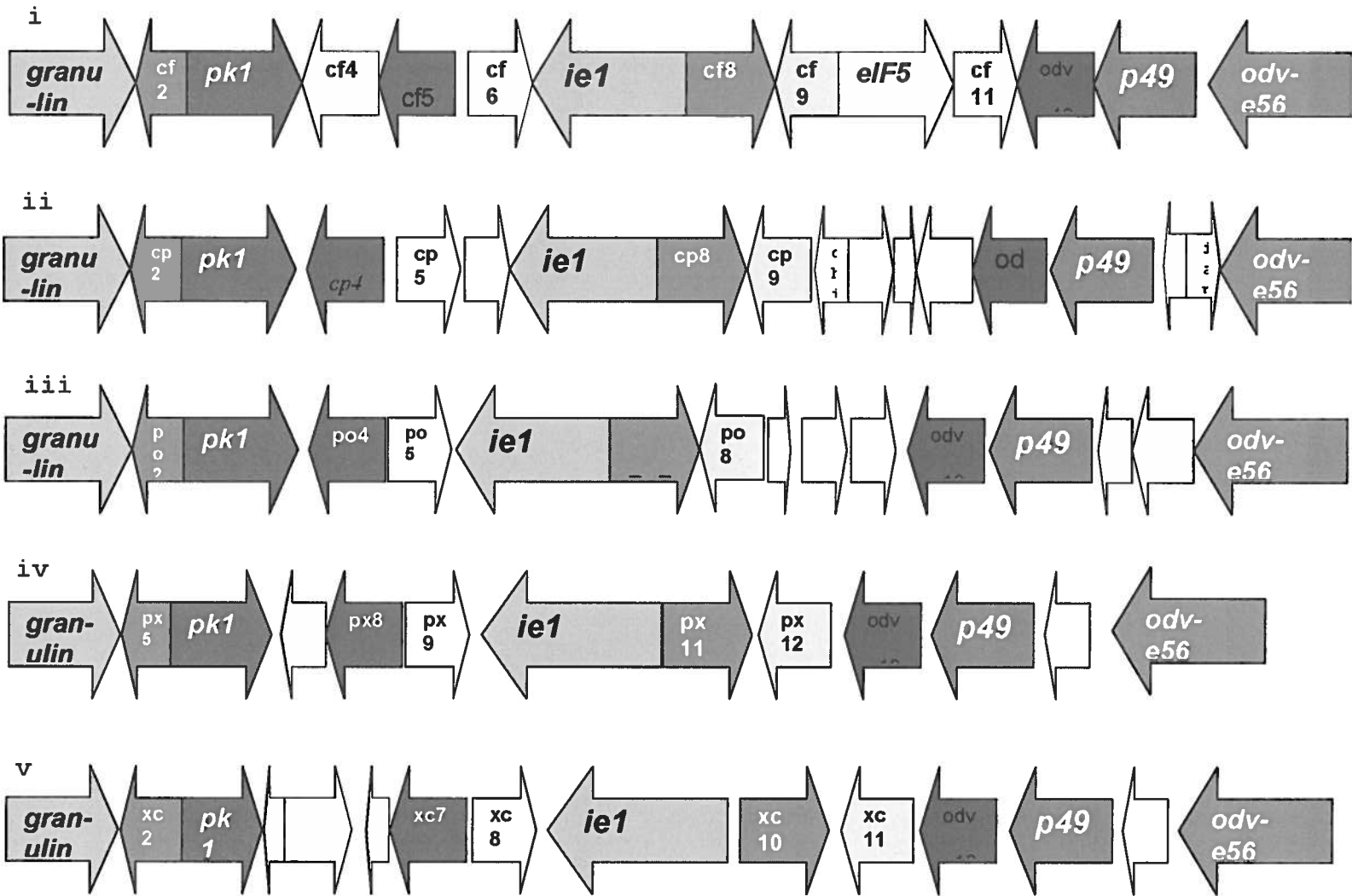
Using this approach several phylogenetic trees were constructed and every one of them agreed on the separation of the GVs and NPVs and the division of NPVs into groups I and II as postulated by other researches (Zanotto *et al.*, 1993; Bulach *et al.*, 1999) supported with high bootstrap values (Figure2).

Here we demonstrate that GVs can be grouped in 2 distinct groups: Group I (ChfuGV, CpGV, PhopGV and AoGV) and Group II (XcGV, PxGV, LoGV and TnGV). ChfuGV conserved proteins are most closely related to those of CpGV, PhopGV and AoGV. ChfuGV along with CpGV, PhopGV and AoGV share a high degree of gene order preservation within the studied genome region (Figure1). GVs and NPVs demonstrated different patterns in their gene order in polyhedrin/granulin, *odvp-6e/odv-e56* and granulin/polyhedrin genes regions (data is not shown).

The search is now intense to determine the complete genome sequences of more GVs. Baculoviruses are rapidly being dissected by sophisticated tools in attempt to illuminate the evolutionary pathway of this family of viruses. Hence, the provided results in this article can offer useful information for future studies on ChfuGV molecular biology and can be used in other baculoviral phylogenetic studies in which the achievement of a better understanding about those biological and ecological characteristics that divide this group of viruses is the major goal.

Figure1. Schematic illustration of the gene structure and orientation of a) *odvp-6e/odv-e56*-containing region of various GVs (i-ChfuGV, ii- CpGV, iii-PhopGV, iv-PxGV, v-XcGV). b) ORFs upstream to granulin (i- ChfuGV, ii-PhopGV, iii-AoGV, iv-CpGV, v-XcGV, vi- PxGV). The arrows represent the relative length and direction of the ORFs. White arrows indicate ORFs with no granuloviral homologue.

a)



b)

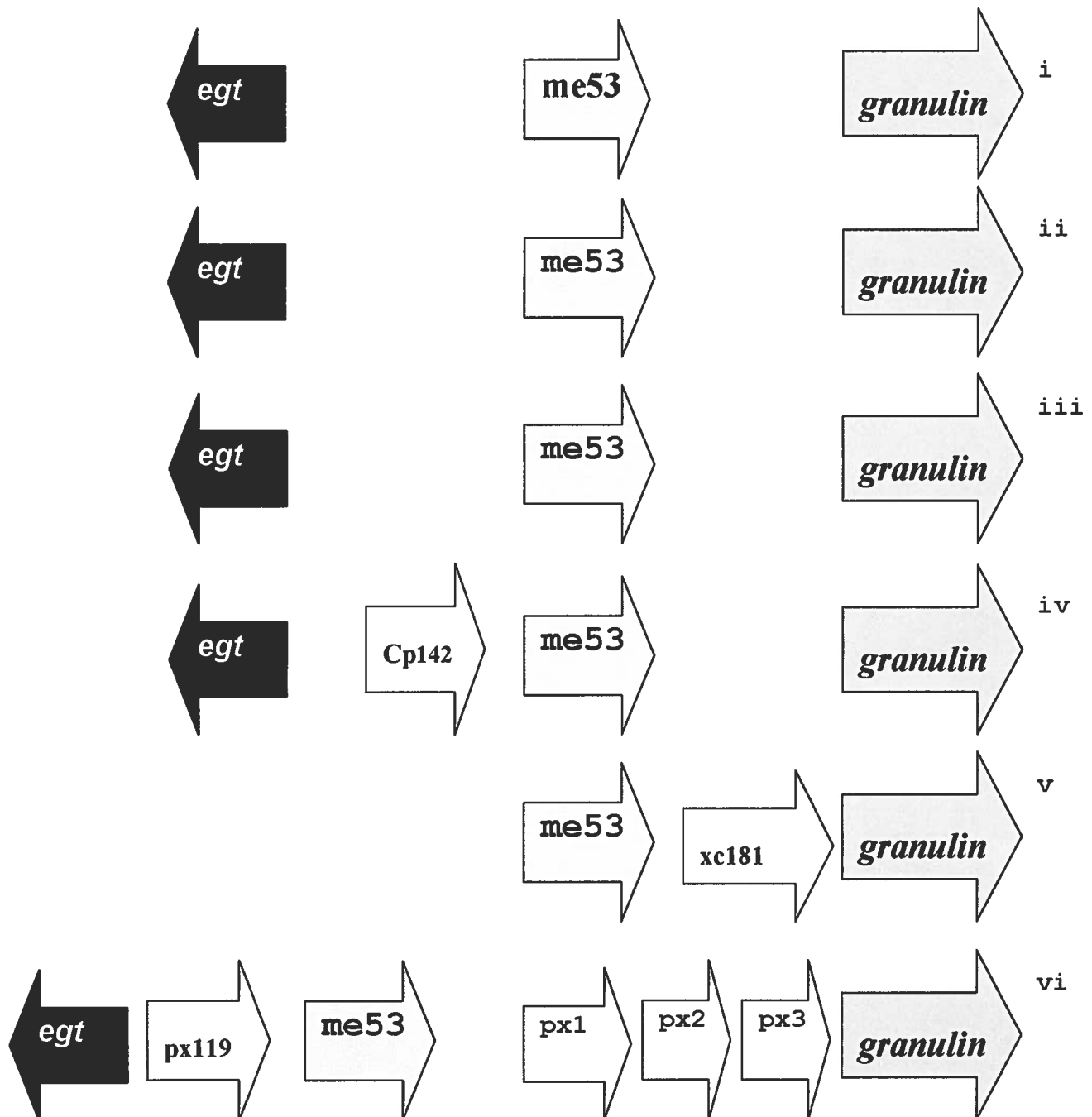
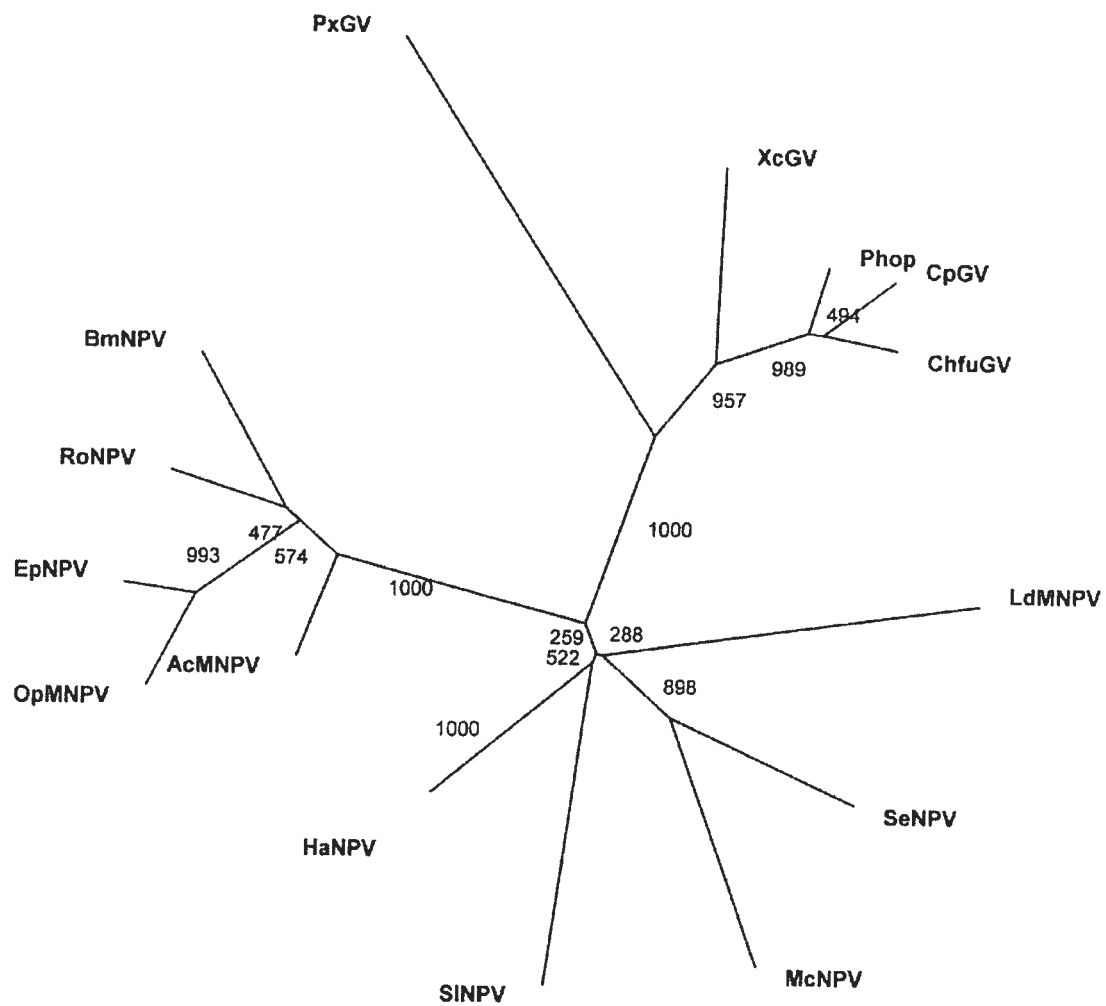


Figure2. Phylogenetic analysis of the protein product of baculoviral *odv-e18* gene located within the *odvp-6e/odv-e56* region of ChfuGV genome. Bootstrap values are shown.



0.1

2.5. Acknowledgements

The authors would like to thank Dr. Anne Bruneau, Department of Biology, University of Montreal for helping us in the phylogenetic studies. We also thank Dr. David Morse, Department of Biology, University of Montreal for his scientific discussions. This research was funded by the *Ministère des Ressources Naturelles du Québec* Grant number 0316-240S to CG.

CHAPTER 3 – Publication #2

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

Authors: Kianoush Khajeh Rashidan^{1*}, Nasha Nassoury², Paresa N. Giannopoulos¹, and Claude Guertin¹

1- Institut national de la recherche scientifique-Institut Armand-Frappier, 531 Blvd des Prairies, H7V 1B7, Laval-Canada

2- Université de Montréal, 4101 Sherbrooke east, H1X 2B2 Montreal-Canada

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

Published in Journal of Biochemistry and Molecular Biology in 2002. Number 35, pages 595 to 603.

Contribution of the authors

I was the principal investigator in this report while Dr. Nassoury's expertise in screening to locate the clone was used to complete this report. Mrs. Giannopoulos helped me by performing some of the PCR reactions. Dr. Guertin was the director of the project and the majority of the experiences have been realized in his laboratory.

Abstract in French:

Identification et caractérisation de la protéine structurale ODVP-6E/ODV-E56 du granulovirus de *Choristoneura fumiferana* granulovirus

Un gène identifié et séquencé sur le génome du *Choristoneura fumiferana* granulovirus (ChfuGV) code pour une protéine baculovirale homologue à l'ODVP-6E/ODV-E56, une protéine structurale associée à l'enveloppe. Le gène odvp-6e/odv-e56 du ChfuGV a été localisé sur le fragment subgénomique *Bam*HI de 11 kb, en utilisant différentes séries d'amorces dégénérées, lesquelles ont été dessinées en se basant sur les résultats du séquençage protéique d'une protéine majeure structurale de 39 kDa, associée avec le virus d'inclusion (occlusion-derived virus, ODV). Le gène possède un cadre de lecture ouvert (open reading frame, ORF) de 1062 nucléotides (nt) codant pour une protéine de 353 acides aminés dont le poids moléculaire prédit est de 38,5 kDa. Les données de la séquence d'acides aminés, déduite de la séquence nucléotidique, ont été comparées à celles de d'autres baculovirus. La protéine ODVP-6E/ODV-E56 du ChfuGV, ainsi que les autres protéines ODVP-6E/ODV-E56 de baculovirus contiennent toutes deux domaines transmembranaires putatifs à leurs extrémités C-terminales. Quelques sites putatifs de N- et O-glycosylation, de N-myristoylation et de phosphorylation ont été détectés dans la protéine ODVP-6E/ODV-E56 du ChfuGV. Un patron similaire a été détecté lorsque les comparaisons de profils d'hydrophobicité ont été réalisées sur ODVP-6E/ODV-E56 du ChfuGV et sur les protéines homologues de d'autres baculovirus. Au niveau nucléotidique, un motif de promoteur tardif (GTAAG) a été localisé à -14 nt en amont du codon d'initiation du gène odvp-6e/odv-e56 chez le ChfuGV. Une variation légère du signal de polyadénylation AATAAT, a été détectée à la position +10 nt en aval du signal

de terminaison. Un arbre phylogénétique des ODVP-6E/ODV-E56 de baculovirus a été construit en utilisant une analyse de parcimonie maximale. L'estimation phylogénétique a démontré que l'ODVP-6E/ODV-E56 du ChfuGV s'associe d'avantage aux protéines retrouvées chez les granulovirus de *Cydia pomonella* (CpGV) et de *Plutella xylostella* (PxGV).

3.1. Abstract

A gene encoding 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 *Bam*HI subgenomic fragment using different sets of degenerate primers, which were designed using the results of protein sequencing of a major 39 kDa structural protein associated with the occlusion-derived virus (ODV). The gene has a 1062 nucleotide (nt) open reading frame (ORF) encoding a protein with 353 amino acids with a predicted molecular mass of 38.5 kDa. The amino acid sequence data derived from 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 ChfuGV ODVP-6E/ODV-E56 protein. A similar pattern was detected when hydrophobicity plots of ChfuGV ODVP-6E/ODV-E56 were compared 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 ChfuGV *odvp-6e/odv-e56* gene. A slight variant of a polyadenylation signal, AATAAT, was detected at the position +10 nt downstream from the termination signal. A phylogenetic tree for baculoviral ODVP-6E/ODV-E56 was constructed using 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, ODVP-6E/ODV-E56, envelope-associated protein, protein analysis, phylogeny.

3.2. Introduction

Granuloviruses (GVs) are members of the Baculoviridae. Baculoviruses have relatively large double-stranded DNA genomes. Baculoviruses are frequently used as bio-pesticides against phytophagous insects mainly belonging 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 the eastern Canada and United States.

Like other members in the family, replication of ChfuGV in insects is biphasic and involves two morphological distinct forms of virion: occlusion-derived virus (ODV), and budded virus (BV). Dissociation of ingested viral occlusion bodies (a highly organized paracrystalline matrix which contains the ODV) under the alkaline environment inside the midgut of susceptible insect larvae yield ODVs which bind to midgut epithelium cells and initiate the primary infection. During the early stages of infection in midgut cells 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 infection, the ODV form of virus is produced in various tissues of the infected insect and released upon the death of the insect. ODV phenotype obtains its envelope from intracellular microvesicles within the nucleoplasm of infected cell (Hong *et al.*, 1994, Braunagel *et al.*, 1996). Invaginated inner nuclear membrane has been hypothesized to be the source for 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.*, 1996b).

Analysis of 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 and several proteins have been identified as specific to ODV or BV (Rohrmann, 1992). Among these structural proteins at least eight are known to be specific to the ODV phenotype. Those are: 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.*, 1989), 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). Proteins that could participate in adsorption, fusion and penetration have not been identified but any of abovementioned 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 homologous to the baculoviral *odvp-6e/odv-e56* in ChfuGV. The nucleic acid and deduced amino acid sequence of ChfuGV is presented. The ChfuGV ODVP-6E/ODV-E56 predicted amino acid sequence was compared to those of other baculoviruses to search for the conserved motifs. 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 has been performed to determine the evolutionary relationship between ChfuGV ODVP-6E/ODV-E56 protein with those of other baculoviruses.

3.3. Materials and Methods

3.3.1. In vivo Production and Purification of Virus and DNA Extraction

Choristoneura fumiferana fourth-instar larvae have been infected by ChfuGV using virus contaminated artificial diet (Forté *et al.*, 1999). The virus isolation from infected larvae followed by DNA extraction was carried out as previously described (Bah *et al.*, 1997).

3.3.2. 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 (0.5M Sodium Carbonate pH 10.5 for 60 min on ice) and undissolved granules were separated by centrifugation (5,000 g for 10 min.). Supernatant 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. Enveloped nucleocapsids were then collected and after being washed in distilled water, centrifuged at 100,000 g (Beckman SW-41) for 60 min at 4°C and resuspended in TE buffer (10mM-Tris-HCl, 1mM-EDTA, pH 7.4). The purity and integrity of enveloped nucleocapsids were examined by transmission electron microscope. Purified enveloped nucleocapsids were disrupted in an equal volume of Laemmli sample buffer (Laemmli, 1970) in presence of β -mercaptoethanol (Bio Rad). 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-polyacrylamid gel electrophoresis (SDS-PAGE). Proteins were visualized using light Coomassie Brilliant Blue (Invitrogen) staining and the molecular mass (*Mr*) of virion proteins was determined by comparing to those of standards (Bio Rad).

Gel purified ChfuGV structural proteins were used for protein sequencing. Major protein bands from ChfuGV enveloped nucleocapsids preparation, including a 39kDa band, were excised from the gel and were subjected to protein sequencing analysis. Sequencing was performed by Harvard Microchemistry facility using either microcapillary reverse-phase HPLC nano-electrospray tandem mass spectrometry (μ LC/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 degenerate primers (with the least redundancy) to be used in PCR reactions on ChfuGV genomic DNA in order to locate the genes coding for the proteins on ChfuGV genome.

3.3.3. Cloning and DNA sequencing

Using different combinations of degenerate primers, PCR amplifications were performed using ChfuGV genomic DNA as template. A 758bp fragment from the gene encoding the ODVP-6E/ODV-E56 was amplified with the combination of following pair of degenerate 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 PCR reactions were carried out using *Taq* DNA polymerase (Amersham Pharmacia Biotech) in the buffer supplied by manufacturer supplemented with 200 μ M of each dNTP and $MgCl_2$ to a final concentration of 2.5 mM. PCR reactions were subjected to heat for 10 min at 95° C prior to the start of the amplification (30 cycles of 95 °C, 30sec; 45 °C 30 sec; and 72 °C, 30sec). The PCR amplicon was cloned in a PCR2.1 (Invitrogene) cloning vector and sequenced, and the sequence was compared to homologues in GenBank/EMBL by using the BLAST algorithm (Altshul *et al.*, 1992). In order to locate *odvp-6e/odv-e56* gene on ChfuGV genome, subgenomic fragments

were generated by *Bam*HI endonuclease and fractionated in a 1% agarose gel. Gel extracted fragments were then used as templates for PCR reactions using the same set of primers. PCR reaction on an 11kb ChfuGV *Bam*HI fragment yielded an amplicon with the expected size. This 11kb restriction fragment was cloned in pBlueScript-SK⁺ cloning vector (Stratagene), and transformed into *E.coli* XL-1 Blue (Stratagen). All the manipulations including 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 on both extremities on either strand using a primer walking technique. Automatic sequencing was carried out using an Applied Biosystems automated DNA sequencer model 377XL (Applied Biosystems CA, USA).

3.3.4. Computer Analysis

Sequence data assembly and analysis were performed with Sequencher software version 4.0.5 (Gene Codes Corp.) and MacVector program version 4.5.0 (Eastman Kodak). Nucleotide sequence and its predicted amino acid sequence were compared to homologues in GenBank/EMBL and SWISSPROT by 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 used in this study are as follow: *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 (SiNPV), AAL01703; and *Culex nigripalpus* baculovirus (CnBV), AAK13276. Amino acid composition was calculated using Peptide Statistic (<http://web.umassmed.edu/cgi-bin/biobin/pepstats>), Pepinfo (www.ebi.ac.uk/service/tmp/1598281014310854.html) and ProtParam (<http://www.expasy.ch/tools/protparam.html>) Tools. Hydropathy plots were obtained 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 used in this study was GORIV (Garnier *et al.*, 1996). Potential transmembrane regions were identified using programs Tmpred (Hofmann and Stoffel, 1993), and TMHMMver. 2.0 (Moller *et al.*, 2001). *N*- and *O*-linked glycosylation, *N*-myristoylation, and phosphorylation sites were predicted using the programs NetOglyc (Hansen *et al.*, 1998) and Proscan (Bairoch *et al.*, 1997).

3.4. Results and Discussion

In this paper we have described the identification and characterization of a gene on the ChfuGV genome encoding a 39kDa protein homologue to ODVP-6E/ODV-E56 from CpGV, PxGV, XcGV, HaNPV, SeNPV, OpMNPV, CfMNPV, EpNPV, BmNPV, AcMNPV, LdMNPV, SINPV, and CnBV. The results of protein sequencing of a major 39 kDa protein in ChfuGV enveloped nucleocapsid yielded several stretches of amino acids. These amino acid sequences were found to be very similar to a highly conserved ODV-specific envelope protein known as ODVP-6E/ODV-E56 after being compared to homologous proteins in GenBank/EMBL using the BLAST algorithm (Altshul *et al.*, 1992). Based on protein sequence of ODVP-6E/ODV-E56 different sets of degenerate primers were developed. DNA coding for ODVP-6E/ODV-E56 was amplified using a pair of these degenerate primers (39-FRI and 39-RVII). The gene encoding this protein was located in an 11-kb *Bam*HI fragment. The open reading frame (ORF) is 1062 nt potentially encoding 353 amino acids, with an estimated molecular mass of 38.5kDa (the actual size of the protein on SDS-PAGE is 39kDa). The sizes of baculoviral ODVP-6E/ODV-E56 proteins vary between 351 aa in PxGV to 379 aa in CfMNPV. The localization and the orientation of the *odvp-6e/odv-e56* gene on the ChfuGV genome are 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 a CCAAT element were centered, respectively, at 134 nt and 150 nt upstream from the predicted start triplet. A late promoter motif (GTAAG) was located at 14 nt upstream

from the first ATG. The TAAG motif is known to be a strong promoter of the late genes in baculoviruses. The gene contained a slight variant of a polyadenylation signal, AATAAT, at 3' noncoding region 10 nt downstream from translational terminator signal. The nucleotide sequence surrounding the translational start codon (TAAATGG) was conformed partly to Kozak's rule for efficient eukaryotic translation initiation with the presence of purine base only at +4 position and not at -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 a content of non-polar (54.4%), polar (45.6%), acidic (8.8%) and basic (8.2%). The protein is particularly rich on asparagine (11.6%) residues. The protein has three potential *N*-glycosylation N-{P}-[ST]-{P} sites located at residues 59, 179, and 279. Glycosylation is important for proper folding or targeting of some 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). Analysis of ODVP-6E/ODV-E56 hydrophobicity plots suggested the presence of two hydrophobic domains within the C-terminal that were conserved in all baculoviral ODVP-6E/ODV-E56 proteins. These highly hydrophobic regions encompassing two membrane-spanning regions 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 (Fig. 2&3).

Besides these two regions ChfuGV ODVP-6E/ODV-E56 protein shared almost similar hydrophobicity plots with others 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 ChfuGV and other known baculoviral ODVP-6E/ODV-E56 proteins (data is not shown).

The deduced amino acid sequence of ChfuGV ODVP-6E/ODV-E56 was compared to other known baculoviral ODVP-6E/ODV-E56 proteins. Results revealed a significant similarity between the sequence of this protein in ChfuGV and other granuloviral and nucleopolyhedral ODVP-6E/ODV-E56 proteins (Table 2). ODVP-6E/ODV-E56 in ChfuGV has 57% nucleotide sequence identity and 74% amino acid sequence identity to its closest relative being CpGV. CnBV ODVP-6E/ODV-E56 was the most distant, showing 24% amino acid sequence identity. A high degree of conservation in amino acid sequences was observed within the two membrane-spanning regions of granuloviral ODVP-6E/ODV-E56 proteins (Fig. 2). Six other conserved motifs were also detected in 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 (⁸LRRTN[K/R]VY¹⁵, ⁷¹V[L/M]RNNDV[V/I/T]GMR[Q/E][L/I]F⁸⁴, ¹⁰²DNIPD[S/A]T[I/L]¹⁰⁹, ¹⁷⁴V[D/E]A[I/L/M]NRTGGS[W/Y][Y/W]¹⁸⁶, ²⁵⁴[S/T]VCR[G/A]SD²⁶⁰, and ²⁸⁴C[I/V]EPYD[F/M][G/A]DLI[G/A]DLGLD³⁰⁰. ChfuGV ODVP-6E/ODV-E56 shared a distinct motif with all other baculoviral ODVP-6E/ODV-E56 (except for CnBV). This motif was located in between the two membrane-spanning regions within the C-terminal

region of the protein (²⁹²DL[I/V][G/A]DLGLD³⁰⁰). ChfuGV, CpGV, and PxGV shared a highly conserved stretch of 31 amino acids covering a region from residue 162 to 193 of ChfuGV ODVP-6E/ODV-E56. A close look to 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.

Presence of six conserved cysteine residues between two hydrophobic membrane-spanning regions was revealed by performing multiple alignments on deduced amino acid sequences of ChfuGV ODVP-6E/ODV-E56 and those of other baculoviruses (Fig. 2). These highly conserved cysteine residues might be implicated in the formation of disulfide bridges that eventually lead to correct folding of baculoviral ODVP-6E/ODV-E56 proteins.

Phylogenetic analysis of ODVP-6E/ODV-E56 proteins is shown in Fig. 5. Tree was produced via maximum parsimony to estimate the evolutionary relationship between 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 along side with CpGV in the same cluster, supported by 65% bootstrap value.

The data presented in this paper provide another evidence to demonstrate the fact that ODVP-6E/ODV-E56 protein is a highly conserved protein in GVs and NPVs. The high degree of the conservation, even between two distantly related members of baculoviridae family, indicates the importance of this protein and can be translated to a possible

significant function accomplished by this protein throughout the cycle of infection. We are currently conducting studies aimed to further characterize the ChfugV *odvp-6e/odv-e56* at transcriptional and translational level.

Figure1. 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.

1 - GGGGGGATC CCAATAATAAATATTC TATAATTAATGGTTTGTGTATGATATGCGTGACG - 60
 61 - AAACGTGTGTGTAAAGATTAATACAGAAAGCACCGTCAACTCCAATTTATATATTCATG - 120
 121 - GTTCTTTTTAAATTTGTTTACCTT GTAAGTTAATTAATGGCTTCTTTT **TTTACGGGT** - 180
 - M A S F F T G
 181 - **CTAAG**ACGCACTAATAAAGTATATCCAAACACAAACAGTTTTTGTCTGACCATGCGTTA - 240
 - L R R T N K V Y P N T N S F L S D H A L
 241 - TTTATTAGAAACCAACACCTGTAGGTTTAAATTTAAACAATCCTACCACAATGGGTATT - 300
 - F I R N Q T P V G F N L N N P T T M G I
 301 - GCGGGAGGTAATGTTACGCCGGGTTATAATATAAACGGTACCTTTGTAGCAACGCCAAT - 360
 - A G G N V T P G Y N I N G T F V S N A N
 361 - GTCATTCAGTGCTTAGGAACAACGATGTTGTTGGTATGAGGCAAATTTTCCGGAGGCA - 420
 - V N S V L R N N D V V G M R Q I F P E A
 421 - TCCAATAATCAAATGAATGGTTTGACAACTTGAGGAGAGCCGACAATATTCCAGATTCT - 480
 - S N N Q M N G L T N L R R A D N I P D S
 481 - ACTATTCACAGTTTAAAAACAAGAAAAATAACGTTAAACAATCACATCCCAGAAACAGTA - 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 - TTAAGGGGTGCCGGATATGTGACATTGTTGGAGTGGGTGTTTATTAGTTATTAATGTC - 660
 - L R G A G Y V T L F G V G V Y L V I N V
 661 - GCCGATCTCGTCCGGTCAATAGTTGACGCGATAAACCGCACTGGTGGTAGCTGGTATTTT - 720
 - A D L V G S I V D A I N R T G G S W Y F
 721 - AGAGGCAATAATGGGGGTGATAATTTTAAACAACATTCAATCGTGCATATTGCGCTACAGA - 780
 - R G N N G G D N F N N I Q S C I L R Y R
 781 - ACTTGCGCGTACCATTACAGATTTCAAGATTCCGGTATGTGTTCTAGATCCACAGAC - 840
 - T C G V P F T D I Q D S V C V L D P H D
 841 - GCTACAAACGTTGACCCACTAATGACTTTAGAAGAAGCAAGTGATTTTGAATAACTAT - 900
 - A T N V D P L M T L E E A S D F C N N Y
 901 - AATCATGCAACAGAA**CAAAGTGTTCGTCG**GGTTCAGATCCAAACGCAAATCCTACAAGT - 960
 - N H A T E Q S V C R G S D P N A N P T S
 961 - TTACAATACCTTGACATAAGTCTTCTTGCATCCAATCAAATATAGAATGTATCGAACCC - 1020
 - L Q Y L D I S L L A S N Q T I E C I E P
 1021 - TACGATTTTGGTGATTTAATAGGAGATTAGGATTAGATTGGTTATTAGGCGATAATGGT - 1080
 - Y D F G D L I G D L G L D W L L G D N G
 1081 - ATTGTCACCGCCAGCTCTAACAGTTTATTAAGTGTATCGGACAATTTTGTGACCATAATA - 1140
 - I V T A S S N S L L S V S D N F L T I I
 1141 - TTAGTGATAGGGGGAATTTAGTTTTCATTGGATTATAATCTACAAAGTAACT - 1200
 - L V I G G I L V L L F I G F I I Y K V T
 1201 - ATAACGAAAAGAAGTATATGAACACGAGTT**AATAAT**GAGCGTGAGTTAAAAATTAAATTT - 1260
 - I T K R S I *
 1261 - TTAATTTTAAAAATATTGCTATTTTATGTGCGCTCATTATTTGCTAAAAATACACAAA - 1320
 1321 - TTTGTATATCGATAAGCGCGAGTTAAACAGCAAAAATTTTAAATTTTAAATTTGCTGT - 1380
 1381 - TTTAACTCGCGCTCATTCTAGAAACAAACAAGCGAAGAGCGAGCGGTGTAACCCAAATT - 1440
 1441 - GAGATTAAAGCTCATTGTAAATGTGTATAATCAATAATATAGTAACATATTAAGTGATG - 1500
 1501 - TAAAATAAGTATT - 1513

Figure 2. Alignment of ChfuGV ODVP-6E/ODV-E56 with other baculoviral ODVP-6E/ODV-E56 proteins 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; SiNPV, AAL01703; and CnBV, AAK13276.

ChfuGV_	1	MA	SFF	GL	RR	TN	KV	YF	NT	NS	FL	SD	HAL	F	IR	NO	TF	VG	FN	LE	NP	TT	MG	IA	G	CN	VT	F																					
CpGV_	1	MS	FF	R	GL	RR	TN	KV	YN	LP	SG	FI	TD	HA	QL	IR	NO	TP	AG	FN	LE	NP	TT	MG	LA	N	GT	V	F																				
PxGV_	1	MS	LF	K	GL	RR	TN	KV	YA	NG	AG	FI	TD	HA	TF	VR	PN	TI	NG	FN	LS	AP	DV	VP	VA	T	G	TE	P																				
XcGV_	1	MS	TF	R	GL	RR	TN	KV	YN	AP	SG	FI	AD	HT	Q	L	I	K	NI	P	SG	FN	LD	NP	TT	IS	GP	N	NT	HI	P																		
HaNPV_	1	MS	FF	R	N	L	R	N	V	N	R	P	P	EP	MH	AS	FV	IT	AN	TS	VV	AN	AP	SG	FN	NV	F	TP	SS	QS	L	Q	N	NT	VL	G													
SeNPV_	1	---	M	F	R	P	L	S	V	N	T	F	FT	N	N	I	A	S	FT	N	Q	N	L	N	V	I	N	N	P	Q	G	R	N	V	I	S	N	E	R	T	V	D	I	C	N	G	V	R	P
OpSNPV_	1	MS	---	FF	T	N	L	R	R	V	N	K	V	Y	P	---	LS	AS	F	I	V	DN	RL	LL	MT	TP	AG	FT	N	V	L	N	V	P	ST	R	N	L	C	N	G	R	T	E	P				
CfMNPV_	1	MS	---	TF	F	T	N	L	R	R	V	N	K	V	Y	P	---	NO	AT	FL	I	DN	TR	LL	TT	TP	AG	FT	N	V	L	R	AP	ST	R	N	L	C	N	G	R	T	E	P					
EpNPV_	1	---	MT	---	FF	S	N	L	R	R	V	N	K	V	Y	P	---	NP	AT	F	AT	DN	AR	LL	TT	TP	AG	FT	N	V	L	N	V	P	ST	R	N	L	C	N	G	R	T	E	P				
BmNPV_	1	MS	---	FF	T	N	L	R	R	V	N	K	L	Y	P	---	NO	AS	F	LA	I	DN	TR	LL	T	ST	PA	AG	FT	N	V	L	N	AP	S	V	R	N	L	C	N	G	R	T	E	P			
AcMNPV_	1	MS	---	FF	S	N	L	R	A	V	N	K	L	Y	P	---	NO	AS	F	IT	I	DN	TR	LL	T	ST	PA	AG	FT	N	V	L	N	AP	S	V	R	N	L	C	N	G	R	T	E	P			
LdMNPV_	1	MS	---	FF	S	N	L	R	R	V	N	K	V	Y	P	---	MP	NQ	F	LA	V	DN	L	R	V	L	N	ST	EN	G	F	ON	V	F	S	AP	ST	R	I	Q	I	G	T	N	R	E	V	P	
SlNPV_	1	MS	---	FF	S	N	L	R	R	V	N	K	V	Y	P	---	NO	IS	F	AA	I	DN	V	T	LA	AS	T	EN	G	F	OH	V	F	S	AP	ST	R	I	Q	I	G	T	N	R	E	V	P		
CnBV_	1	MA	SK	L	F	G	P	L	R	V	M	G	H	V	VD	RS	AL	AM	DR	FR	TA	FR	AV	K	ND	N	P	L	R	I	F	D	L	E	---	IR	E	V	P	T	---	V	S	R	N				

ChfuGV_	55	G	Y	N	I	N	G	T	---	F	V	S	N	A	N	V	M	S	V	L	R	N	N	D	V	V	G	M	---	Q	I	F	P	E	---	A	S	N	N	Q	M	N	G	L	T	N	L	R	R	A	D	N	I	P	D	S				
CpGV_	54	G	Y	N	I	N	G	A	---	F	I	S	N	T	I	N	V	N	T	V	L	R	N	N	D	V	V	G	M	---	Q	L	F	P	D	---	A	S	N	N	Q	M	N	G	L	T	N	L	R	R	A	D	N	I	P	D	A			
PxGV_	53	G	Y	V	I	N	D	R	---	F	V	P	N	A	R	A	N	N	V	M	R	N	N	D	V	I	G	M	---	E	I	F	P	N	---	A	N	S	N	Q	I	N	G	L	S	Y	R	R	I	D	N	I	P	D	S					
XcGV_	56	G	Y	G	T	N	T	N	---	T	S	N	S	T	V	N	R	V	M	R	N	N	D	V	T	C	M	Q	---	Q	L	F	P	S	---	A	T	P	Q	I	N	S	L	G	L	P	Q	L	D	N	I	P	D	A						
HaNPV_	55	G	Y	C	T	P	A	G	---	F	V	S	T	P	Q	M	N	S	F	M	R	N	N	D	V	P	R	I	---	Q	V	F	P	N	---	A	N	T	N	Q	L	D	S	L	S	T	L	R	R	M	D	N	V	P	D	A				
SeNPV_	55	G	Y	A	L	P	N	I	Q	---	S	T	A	E	M	N	S	I	M	R	N	N	D	V	S	A	C	M	---	R	V	E	G	N	N	---	I	S	N	N	D	Y	I	G	L	S	Q	L	R	R	A	D	N	I	P	D	A			
OpSNPV_	55	G	Y	N	L	S	N	N	Q	---	F	V	S	A	G	D	I	N	R	I	T	R	S	N	D	V	P	R	I	---	G	V	F	Q	---	I	S	D	P	Q	I	G	S	I	N	Q	L	R	R	A	D	N	V	P	D	A				
CfMNPV_	56	G	Y	N	L	S	N	N	Q	---	F	V	S	A	G	D	I	N	R	I	T	R	G	N	D	V	P	R	I	---	N	V	F	Q	---	I	S	D	P	Q	I	G	S	I	N	Q	L	R	R	A	D	N	V	P	D	A				
EpNPV_	56	G	Y	N	L	S	N	N	Q	---	F	V	S	A	G	D	I	N	R	I	T	R	S	N	D	V	P	R	I	---	N	V	F	Q	---	I	S	D	P	Q	I	N	S	L	I	Q	L	R	R	A	D	N	V	P	D	A				
BmNPV_	55	G	Y	Q	L	S	N	N	R	---	F	V	S	T	S	D	I	N	R	I	T	R	N	N	D	V	P	N	I	---	N	V	F	Q	---	I	S	D	P	Q	I	N	S	L	I	Q	L	R	R	M	D	N	V	P	D	F				
AcMNPV_	55	G	Y	Q	L	S	N	N	Q	---	F	V	S	T	S	D	I	N	R	I	T	R	N	N	D	V	P	N	I	---	G	V	F	Q	---	I	S	D	P	Q	I	N	S	L	I	Q	L	R	R	V	D	N	V	P	D	F				
LdMNPV_	56	G	Y	N	L	P	N	N	R	---	F	V	S	S	D	I	N	R	V	M	R	N	N	D	T	A	C	M	---	N	I	F	T	S	---	A	D	N	S	Q	L	N	S	L	I	Q	L	R	R	V	D	N	I	P	D	A				
SlNPV_	56	G	Y	N	L	S	N	N	R	---	F	V	S	T	A	I	N	S	A	L	R	N	S	D	T	N	S	I	R	---	G	I	E	G	N	---	L	N	N	T	Q	L	N	G	I	T	Q	M	R	R	I	D	N	V	A	D	P			
CnBV_	58	G	D	L	Y	E	P	H	L	---	M	M	R	P	D	E	P	T	---	P	M	R	E	F	D	T	G	L	R	G	D	Q	---	V	L	N	R	L	R	M	T	S	S	E	R	A	C	I	N	N	I	R	---	G	N	S	I	P	D	N

ChfuGV_	108	T	I	H	S	L	K	T	R	K	N	N	V	K	Q	S	H	E	---	E	T	V	V	R	D	R	A	G	V	E	N	A	L	A	Q	N	P	R	L	G	D	Y	L	R	---	G	A	C	Y	---	V	T	L	F	C	V	G	V	Y	L		
CpGV_	107	T	L	H	G	L	D	T	R	K	N	G	V	K	T	S	H	E	---	E	T	A	V	R	D	R	V	G	V	E	N	A	L	A	Q	N	P	R	L	A	D	Y	L	R	---	G	A	C	Y	---	V	T	L	F	C	V	S	V	Y	L		
PxGV_	106	T	L	H	G	L	D	I	R	K	N	I	K	N	S	R	E	---	E	T	R	V	N	R	Q	G	V	E	A	L	I	K	N	P	R	L	R	D	Y	L	Y	---	G	A	G	R	---	I	T	L	F	C	V	G	S	Y	L					
XcGV_	110	T	L	N	S	L	S	T	K	K	S	S	V	R	A	N	E	---	E	T	A	T	R	T	R	S	G	V	G	V	L	D	R	H	P	R	L	N	Q	Y	L	K	---	T	A	G	I	---	V	G	I	S	C	A	T	I	Y	L				
HaNPV_	108	R	L	H	A	N	L	R	R	Q	S	V	K	T	N	E	F	---	S	T	N	T	R	T	A	G	V	T	N	V	L	N	Q	Q	P	R	L	S	R	Y	L	E	T	A	Q	A	G	Y	---	V	G	L	V	C	V	G	V	Y	L			
SeNPV_	110	N	H	S	Q	L	R	E	D	A	V	K	N	N	E	---	S	T	R	T	R	E	A	G	I	E	N	S	L	N	R	N	E	N	N	R	L	Q	G	L	N	A	G	Y	---	A	A	L	L	C	T	G	A	Y	L							
OpSNPV_	109	N	L	H	V	K	R	T	F	S	D	A	V	K	Q	S	F	---	E	T	N	V	R	E	A	G	V	D	R	A	L	Q	Q	N	P	R	L	N	T	Y	L	Q	G	A	A	C	V	---	G	V	L	L	A	G	C	A	Y	L				
CfMNPV_	110	G	L	H	V	K	R	T	F	S	D	A	V	K	Q	N	E	---	E	T	N	V	R	S	A	D	G	V	D	R	A	L	Q	Q	N	P	R	L	N	T	Y	L	Q	G	A	T	A	G	Y	---	G	V	L	L	A	G	C	A	Y	L		
EpNPV_	110	S	F	H	V	K	R	T	F	S	D	A	V	K	Q	N	E	---	E	T	N	V	R	T	---	G	V	D	R	A	L	Q	Q	N	P	R	L	N	T	Y	L	Q	G	A	T	A	G	Y	---	A	V	L	L	A	G	C	A	Y	L			
BmNPV_	109	H	Y	E	T	K	Q	T	R	S	N	A	V	K	Q	N	E	---	E	T	N	V	R	T	---	P	E	G	V	Q	N	A	L	Q	Q	N	P	R	L	H	N	H	M	R	T	L	N	V	A	G	Y	---	G	V	L	L	A	G	C	G	Y	L
AcMNPV_	109	N	Y	H	T	K	Q	T	R	S	N	A	V	K	Q	N	E	---	E	T	N	V	R	T	---	P	E	G	V	Q	N	A	L	Q	Q	N	P	R	L	H	S	Y	Q	S	L	N	V	G	C	T	---	G	V	L	L	A	T	C	G	Y	F	
LdMNPV_	110	G	L	H	A	A	F	L	R	E	Q	G	V	K	S	N	E	---	S	T	N	T	R	T	---	P	E	G	V	E	G	V	L	A	T	N	P	R	L	H	A	R	L	N	G	L	N	S	V	C	T	---	A	L	L	C	C	G	V	Y	L	
SlNPV_	110	T	I	F	N	---	R	D	E	Q	Q	V	R	N	E	---	D	S	A	T	R	T	---	P	E	G	T	O	N	F	M	I	N	Q	P	R	I	N	Y	L	D	N	L	N	K	A	G	H	---	V	G	L	Y	C	F	G	I	L				
CnBV_	117	L	Y	Q	A	V	R	A	D	E	A	A	T	R	V	R	D	E	N	L	A	N	A	R	T	E	Q	E	---	I	T	D	A	L	A	R	H	P	R	L	R	D	R	L	T	---	A	G	---	V	I	K	A	C	A	C	V	S	L	V		

ChfuGV_	164	V	I	N	V	A	D	L	V	G	S	I	V	D	A	---	N	R	T	G	G	S	Y	F	R	E	N	N	G	D	N	F	N	N	I	Q	S	C	L	L	R	Y	---	R	T	C	G	Y	P	---	F	T</
---------	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	-----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	-----	---	---	---	---	---	---	-----	---	-----

ChfuGV_ 222 V -LDFHDATNVDPMLTLEEASDFCNVNHATEQSVCRGSDENAFETSLQYLDISLLASN
 CpGV_ 221 E -LDFHDPNVDPPLSFDEARNFCNGSLAAEGSVCRGSDTNADPSLQYLDISLEPN
 PxGV_ 220 V -LDPLDPTNVDPILNLEEAOQLCTGYNKREKSVCRASDTTANFDSLQYLDISSLEAN
 XcGV_ 224 --FDFHDPTWRDPYLTSEQLATVCNNYLAAEETVCPASDFYAEYGRGYDYVSLLATN
 HaNPV_ 223 P -SDPLIDN-----INALQSVCHGNAEVERTVCRASDENADELSLQYVDISPLATG
 SeNPV_ 225 Q -NDPLIAD-----QSELNRICDGFYEVERTVCRASDENAYEFSPQYVDISLSTD
 OpSNPV_ 225 A -NDPLVSM-----TAQLQAICSGFNVEQEQTVCRQSDPAADFDSPQYVDVSDLLPG
 CfMNPV_ 226 N -HDPLIAD-----TAQLQAICSGFNVEQEQTVCRQSDPAADFDSPQYVDVSDLLPG
 EpNPV_ 226 A -NDPLIST-----SAQLQSTICAGFNVEIEQTVCRQSDPAADFDSPQYVDISDLPAG
 BmNPV_ 225 S -RDPLIAND----SPLLTNMCGGFNYETKTVCRGSDNFAAFNPSPOYVDISDLPAG
 AcMNPV_ 225 P -FDPLIPNN-----PPELTNMCGGFNYEVEKTVCRGSDPSADFDSPQYVDISDLPAG
 LdMNPV_ 225 Q -FDPLIPDD-----PAGQRAICNGFNVEREQSVCPASDFYADFDSPQYVDISDLPAG
 SlNPV_ 225 T -SDPLIQD-----QSTLQNLCLNNVNOEAEQSVCRASDENADPASLQYVDISDLIVN
 CnBV_ 229 GESGDPIYDS-----VGEAQATCSGNNKSVERSVCRQADVLAEPSQQFLDARTLPEN

ChfuGV_ 280 QTIECIEPYDFGDLIGDLGLDWLLGNGEVTASSNSLLSVSDNFTTILLVICGILVLLFI
 CpGV_ 279 QTVCQVEPYDFGDLIGDLGLDWLLGNGEVTASSNSLTSVSNNFSTTILLVICGILVLLTFI
 PxGV_ 278 QTIQCVEPYDFGDLIGDLGLDWALGEDGLITASSNSFTSVSNNFSTILLAIGGVLLLAFTI
 XcGV_ 281 EIIQCVEPYDMADLIADLGLDHLGLGNGILTNSSASSOSVSNNFSTILLIVCGVIVLLFV
 HaNPV_ 274 HTISCIEPYDFGDLIGDLGLDGLLGEGLNKSSTSFQKLPRIIVVLGIVLLIIFI
 SeNPV_ 276 QMIHCIEPYDFGDLIADLGLDGLLGENGSIQTSSNKSSTISLSELPAILMCAIILLIVLV
 OpSNPV_ 276 QTIMCIEPYSGLDGLIGDLGLDHLGLGEEGLVCKSSNSSDSVSNKLMPLIWLICAVLFLGLI
 CfMNPV_ 277 QTIMCIEPYNLGDGLIGDLGLDHLGLGEDGLVCKSSNSSDSVSNKLMPLIWLICAVLFLGLI
 EpNPV_ 277 QTIMCIEPYNLGDGLIGDLGLDHLGLDGLVCKSSNSSDSVSNKLMPLIWLICAVLFLGLI
 BmNPV_ 277 QTIMCIEPYSFSDLVGDGLGLDWLLGREGLVCKSSNSSDGIRNKIMPIIMMIGAVLFLGLI
 AcMNPV_ 277 QTLMCIEPYSFGDLVGDGLGLDWLLGDEGLVCKSSNSVSDSVSGKLMPIILLICAVLFLGLI
 LdMNPV_ 277 HTLMCIEPYNLGDGLIGDLGLDHLGLGEEGLVCKSSNSKDISSKLWPLLVAIGIFAVVLI
 SlNPV_ 276 TTITCIEPYDLADLIGDLGLDWLLNENGLVCKSKNSKSIQEALVPLFIATGVIFIVFTI
 CnBV_ 282 AHIYCFEPGTLGRLVADLGLADLVDAVGGSVSGSSGNSSGKSSGNPLILLISAFVLLIIII

ChfuGV_ 340 GFIIYKVTIT-----KRSI-----
 CpGV_ 339 GFVIKVVNE-----SSNNNTS-----
 PxGV_ 338 GFIIKVVNN-----KSGK-----
 XcGV_ 341 MYVVIRKMLE-----PKS-----
 HaNPV_ 334 GYIYIK-----RMLMQPPPPFANYNR-----
 SeNPV_ 336 GYFIKKNLS-----NRQTVQFQMPVETAVPTPLPIQVR-----
 OpSNPV_ 336 VYLIYRFLIK----GGG--SSTTN--APEVVI--PPPATTNLN--PQQQI
 CfMNPV_ 337 IYLIYRFVIK----GGAGAAGAAR--APEVIV--PPPPTQQTYNSTKQQI
 EpNPV_ 337 IYLIYRYIFK----SGSSGSAVAMPQPPFVIV--QTPTTRA-----
 BmNPV_ 337 LYFIYRYITK----GGGGGG--GSGGAPTEIVI--QHPASTTAPRR-----
 AcMNPV_ 337 FYFIYRYIMK----GGGGGGVGAATSPTEIVIS--QNPTPTTAPR-----
 LdMNPV_ 337 LVFVFKRLMN----KQTVDIRANA-----
 SlNPV_ 336 GFMIYKRITAPSVRIEGGGGGGGGGGAGVGVQNIR-----
 CnBV_ 342 FVVFVGYST-----RRNSDADRTI-----

Figure 3. Position of potential transmembrane domains in ChfuGV ODVP-6E/ODV-E56 and other baculoviral ODVP-6E/ODV-E56 proteins. The amino acid numbers are on the X-axis and probability of the existence of transmembrane (red), inside (blue) and outside (pink) domains are on Y-axis.

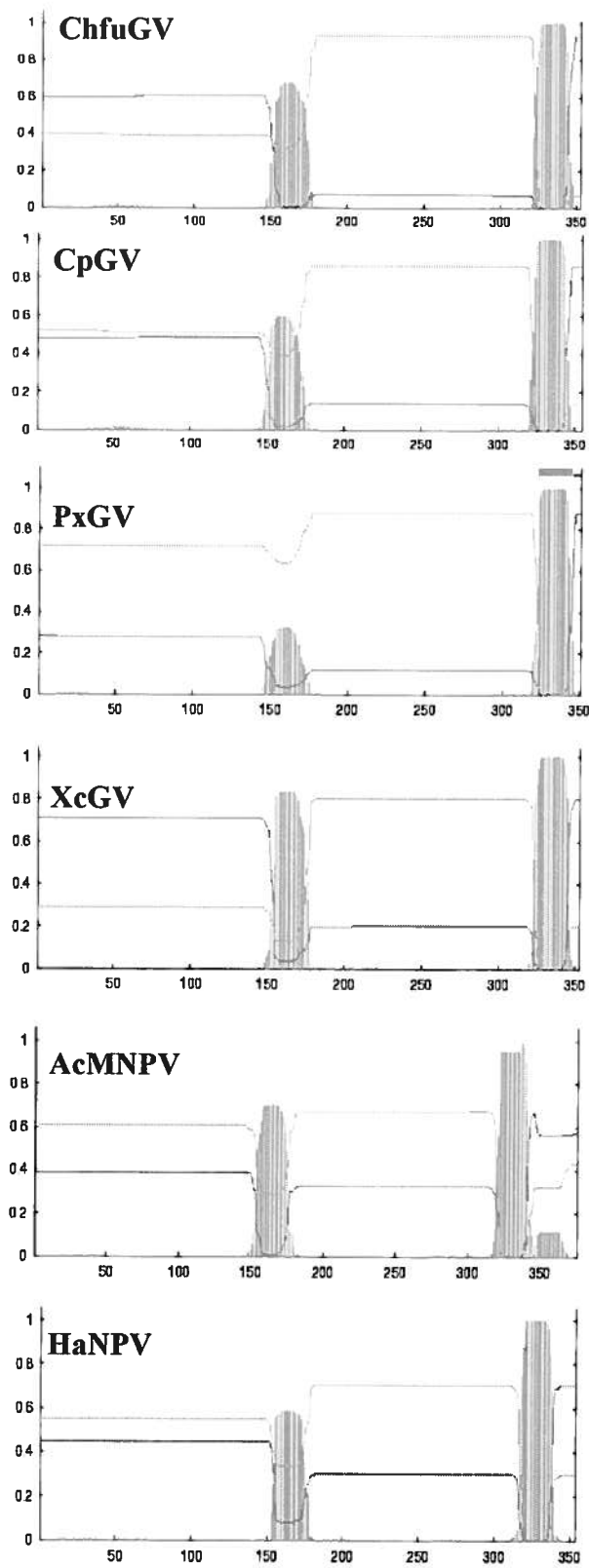


Figure 4. Comparison of the Kyte-Doolittle hydrophilicity plots of ODVP-6E/ODV-E56 from ChfuGV to CpGV, PxGV, XcGV, AcMNPV, and HaNPV ODVP-6E/ODV-E56 proteins. Above the axis (+) denotes hydrophilic regions and below the axis (-) indicates hydrophobic regions.

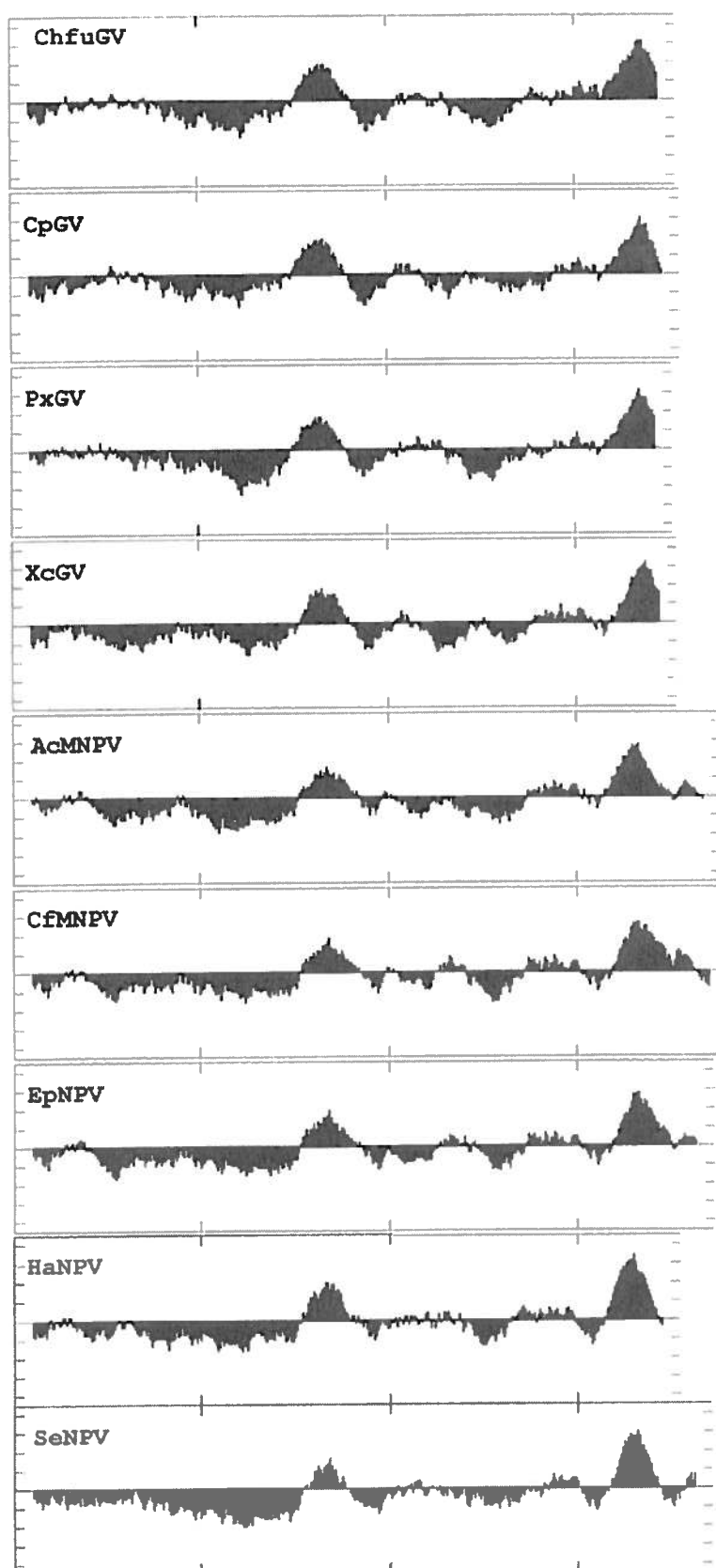


Figure 5. Phylogenetic analysis performed on baculoviral ODVP-6E/ODV-E56 proteins. The unrooted tree was produced by a Branch and Bound search using PAUP4.0b4a. Bootstrap values (100 replicates) are shown. Tree was constructed according to the following settings: 1- uninformative characters ignored. 2- Branch-and-bound search options: an initial upper bound: compute via stepwise; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e 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; SINPV, AAL01703; and CnBV, AAK13276.

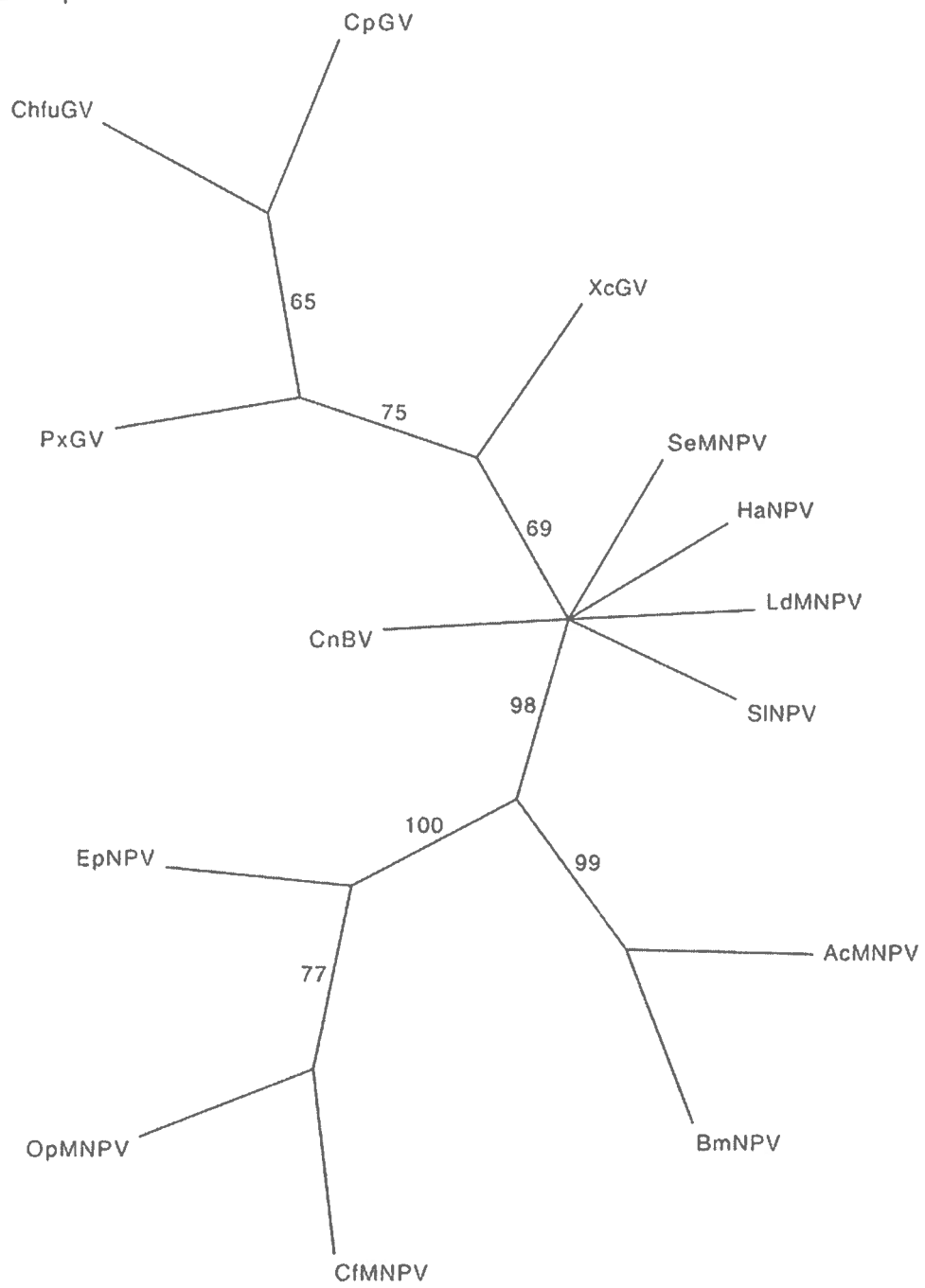


Table 1. List and the location of all putative *N*- glycosylation, *O*-glycosylation, *N*-myristoylation phosphorylation sites within the ChfuGV ODVP-6E/ODV-E56 protein.

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

[illegible]

3.5. Acknowledgements

The authors would like to thank Dr. Anne Bruneau, Department of Biology, University of Montreal for helping us in the phylogenetic studies. We also thank Dr. David Morse, Department of Biology, University of Montreal for his scientific discussions. This research was funded by the *Ministère des Ressources Naturelles du Québec* Grant number 0316-240S to CG.

CHAPTER 4 – Publication #3

Identification and Characterization of a Putative Baculoviral Transcriptional Factor (IE-1) from *Choristoneura fumiferana* Granulovirus

Authors: Kianoush Khajeh Rashidan¹, Nasha Nassoury², Abderrazzak Merzouki¹, and Claude Guertin^{1*}

1- Institut national de la recherche scientifique-Institut Armand-Frappier, 531 Blvd des Prairies, H7V 1B7, Laval-Canada

2- Université de Montréal, 4101 Sherbrooke east, H1X 2B2, Montreal-Canada

The nucleotide sequence data reported in this paper has been submitted to the GenBank nucleotide sequence database and has been assigned the accession number AY064489.

Published in Journal of Biochemistry and Molecular Biology in 2002. Number 35, pages 553-561.

Contribution of the authors

I was the principal investigator in this report while Dr. Nassoury's expertise in screening to locate the clone was used to complete this paper. Dr. Merzouki participated in scientific discussions about the paper and Dr. Guertin was the director of the project and the majority of the experiences have been performed in his laboratory.

Abstract in French :

Identification et caractérisation du facteur transcriptionnel IE-1 du granulovirus de *Choristoneura fumiferana*

Un gène codant pour une protéine baculovirale homologue à l'IE-1 a été identifié et séquencé dans le génome du *Choristoneura fumiferana* granulovirus (ChfuGV). Le gène possède un cadre de lecture ouvert (open reading frame, ORF) de 1278 nucléotides (nt) codant pour 426 acides aminés avec un poids moléculaire estimé de 50,3 kDa. Au niveau nucléotidique, quelques éléments régulateurs agissant en cis ont été détectés à l'intérieur de la région du promoteur du gène ie-1 du ChfuGV. Deux motifs CCAAT ont été détectés dans la région non-codante du gène (5'UTR), un a été localisé sur le brin opposé à -92 nt et l'autre à -420 nt à partir du triplet d'initiation. Deux motifs baculoviraux de promoteur tardif TAAG ont aussi été détectés à l'intérieur de la région du promoteur du gène ie-1 du ChfuGV. Un seul signal de polyadénylation AATAAA, a été localisé 18nt en aval du codon de terminaison traductionnelle putatif du gène ie-1 du ChfuGV. Au niveau protéine, les données de la séquence d'acides aminés du IE-1 dérivée de la séquence nucléotidique du ChfuGV ont été comparées à celles du *Cydia pomonella* granulovirus (CpGV), *Xestia c-nigrum* granulovirus (XcGV) et *Phutella xylostella* granulovirus (PxGV). Les régions C-terminales des séquences IE-1 granulovirales semblent être plus conservées comparativement aux régions N-terminales. Un domaine similaire au domaine basique ressemblant à l'hélice-boucle-hélice (basic helix-loop-helix-like, bHLH-like) chez les NPV a été détecté dans la région C-terminale de IE-1 du ChfuGV (résidus 387 à 414). Un arbre phylogénétique des IE-1 de baculovirus a été construit en utilisant une analyse de parcimonie maximale. Une estimation

phylogénétique démontre que la protéine IE-1 du ChfuGV est plus rapprochée à celle du CpGV.

4.1. Abstract

A gene encoding for a protein homologue to baculoviral IE-1 was identified and sequenced in the genome of *Choristoneura fumiferana* granulovirus (ChfuGV). The gene has a 1278 nucleotide (nt) open reading frame (ORF) encoding 426 amino acids with an estimated molecular weight of 50.33 kDa. At the nucleotide level, several *cis*-acting regulatory elements were detected within the promoter region of the ChfuGV *ie-1* gene of along with other studied granuloviruses (GVs). Two putative CCAAT elements were detected within the noncoding leader region of this gene, one located on the complementary strand at 92 and the other at 420nt upstream from the putative start codon. Two baculoviral late promoter motifs (TAAG) were also detected within the promoter region of *ie-1* gene of ChfuGV. A single polyadenylation signal AATAAA was located 18nt downstream of the putative translational stop codon of the ChfuGV *ie-1*. At the protein level, the amino acid sequence data derived from nucleotide sequence in ChfuGV IE-1 was compared to those of *Cydia pomonella* granulovirus (CpGV), *Xestia c-nigrum* granulovirus (XcGV) and *Plutella xylostella* granulovirus (PxGV). The C-terminal regions of granuloviral IE-1 sequences appeared to be more conserved compared to N-terminal regions. A domain similar to basic helix-loop-helix like (bHLH-like) domain in NPVs was detected at the C-terminal region of IE-1 from ChfuGV (residues 387 to 414). A phylogenetic tree for baculoviral IE-1 was constructed using maximum parsimony analysis. The phylogenetic estimation demonstrated that ChfuGV IE-1 is most closely related to that of CpGV.

Keywords: *Choristoneura fumiferana* granulovirus, *ie-1*, transcriptional factor, protein analysis, phylogeny

4.2. Introduction

The Baculoviridae are large and complex DNA viruses that primarily infect arthropods especially insects, belonging largely to the orders *Lepidoptera*, *Dieptera* and *Hymenoptera*. The Baculoviridae comprised of two genera: the nucleopolyhedroviruses (NPVs) and the granuloviruses (GVs). Granuloviruses have been well characterized at the cytological level and their application for pest control in agriculture and forestry has been fairly successful (Crook, 1991; Moscardi, 1999). The Spruce budworm, *Choristoneura fumiferana*, is the most devastating coniferous tree pest in eastern Canada and the United States. *Choristoneura fumiferana* granulovirus (ChfuGV) are being considered as an alternative biological insecticide due to their specificity for their insect host.

Baculovirus genes are expressed in a transcriptional cascade in which each successive phase is dependant on the expression of genes during the previous phase. Basically, there are presently four commonly established phases in this cascade; immediate early (IE), delayed early (DE), late (L) and very late (VL). During the early phases, the molecular environment and other aspects of intracellular environment of the cell are altered in preparation for the replication and expression of viral DNA. Early virus promoters are composed of multiple *cis*-acting regulatory components. These promoters require host RNA polymerase II for proper transcription (Guarino and Dong, 1994; Pullen and Friesen 1995a; Theilmann and Stewart, 1991; Rodems and Friesen, 1993; Nissen and Friesen, 1989). DE genes need the presence of regulatory proteins that are transcribed during the immediate early phase in order to be efficiently transcribed (Guarino and Summers, 1986). Baculoviral genes encoding *trans*-activating factors have been identified. These

regulatory proteins are able to detect enhancer elements located in several regions of baculoviral genome and contribute to the speed and facility with which the virus is able to direct the host cell metabolic machinery to transcribe particular genes necessary for viral replication.

Immediate early protein IE-1 is the principal transcriptional regulator known in *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) and *Orgyia pseudotsugata* MNPV (OpMNPV) (Hayakawa *et al.*, 2000). Transcription factors are commonly known as structurally complex proteins containing various functional domains associated with DNA binding, protein oligomerization, phosphorylation, activation, and other activities. Transcription factors fall into different families such as zinc fingers, helix-turn-helix (HTH), helix-loop-helix (HLH), or basic leucine zippers. These proteins share a short domain associated with DNA binding or oligomerization (Lewin, 1997; Altshul *et al.*, 1992). Baculovirus IE-1 protein is detected throughout infection (from IE to VL phases) and this involvement suggests that the presence of IE-1 is essential for a productive infection (Okano *et al.* 1999; Choi and Guarino, 1995; Ribeiro *et al.*, 1994). Presence of homologous regions (*hrs*) in *cis* position to the promoter of DE genes enhances the transactivation effect of IE-1 (Guarino and Summers, 1986; Choi and Guarino, 1995). Studies conducted on IE-1 from AcMNPV, the prototype subgroup I baculovirus, demonstrated that oligomerization of IE-1 is a key factor for IE-1 transcriptional stimulation (Olson *et al.*, 2001). These data also indicate that IE-1 needs to be oligomerized as a dimer in order to turn into an active protein capable of binding to the 28bp palindromic repeats of the *hrs*. This was revealed using *ie-1* defective mutants

of AcMNPV. These mutants have been shown to lose the oligomerization activity and hence suffer the loss of IE-1 *hr*-dependant transactivation (Olson *et al.*, 2001). The 28-mer palindrome is the minimal requirement for the enhancer activity of *cis* linked viral promoters (Guarino and Dong, 1994; Olson *et al.*, 2001; Leisy and Rohrmann, 2000; Massari and Murre, 2000; Rodems *et al.*, 1997; Kremer and Knebel-Morsdorf, 1998). Evidences demonstrated that it is essential for IE-1 dimers to interact through cooperative binding with both palindromic half sites of 28-mer palindrome in order to stimulate the *hr* enhancer activity (Leisy and Rohrmann, 2000; Rodems *et al.*, 1997; Kremer and Knebel-Morsdorf, 1998). The model presented for DNA binding by IE-1 proposed that interaction of IE-1 dimer across the 28-mer axis of symmetry makes a simultaneous contact with both half sites (Olson *et al.*, 2001). This contact is a vital parameter that can eventually lead to transcriptional enhancement.

To date, *ie-1* homologues have been reported from 10 NPVs and 3 GV; however, no significant information concerning the granuloviral *ie-1* genes is available in the literature. This paper is the first major report concentrating on the analysis of a granuloviral *ie-1* gene and its deduced amino acid sequence.

4.3. Materials and Methods

4.3.1. In vivo Production and Purification of Virus

Choristoneura fumiferana fourth-instar larvae were infected by ChfuGV using virus-contaminated artificial diet (Forté *et al.*, 1999). Progeny viruses were then extracted from infected larvae as previously described (Bah *et al.*, 1997).

4.3.2. DNA Extraction, Cloning and Sequencing

ChfuGV DNA was isolated from purified enveloped nucleocapsids of the virus. To isolate the enveloped nucleocapsids, occlusion bodies were solubilized under alkaline conditions (0.5M Sodium Carbonate pH 10.5 for 60 min on ice) and undissolved granules were separated by centrifugation (5,000 g for 10 min). Supernatant was then layered on a sucrose gradient (10% to 50%) and centrifuged at 25,000 g (Beckman SW-41) for 60 min at 4°C. Enveloped nucleocapsids were then collected and after being washed in distilled water, centrifuged at 100,000 g (Beckman SW-41) for 60 min at 4°C and resuspended in TE buffer (10mM-Tris-HCl, 1mM-EDTA, pH 7.4). The purity and integrity of enveloped nucleocapsids were examined by transmission electron microscopy. For DNA extraction, purified enveloped nucleocapsids of ChfuGV were digested in 250 µg / ml proteinase K in the presence of 0.5% SDS, for 1h at 37°C. Digested proteins were extracted once with saturated phenol by adding equal volume of phenol to preparation, twice with one volume of Phenol: Chloroform (1:1) and once with an equal volume of Chloroform: Isoamylalcohol (24:1). The DNA present in aqueous phase was precipitated with 2.5 volume of 100% ethanol in the presence of 0.1 volume sodium acetate (3 M, pH 5.2) for

3 h at 4°C and pelleted at 16,000 g for 30 min. The purified DNA was dried and dissolved in TE buffer. The viral genome was digested by *Bam*HI and the fragments were ligated into pSK-Bluescript (Stratagene). During sequence analysis of an 11kb *Bam*HI restriction fragment of ChfuGV genome, a homologue of baculoviral *ie-1* gene was detected. The sequencing was performed using specific primers in both directions by primer walking (Sanger *et al.* 1977).

4.3.3. Computer Analysis

The sequences were analyzed by Sequencher program version 4.0.5 (Gene Codes Corp.) and 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 by using the BLAST algorithm (Altschul *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). Amino acid composition and isoelectric points were calculated using Peptide Statistic (<http://web.umassmed.edu/cgi-bin/biobin/pepstats>) and ProtParam (<http://www.expasy.ch/tools/protparam.html>) Tools. Hydropathy plots were obtained using Kyte and Doolittle hydropathy values (-4.5 to 4.5) via Nixon web tools (<http://www.bmb.psu.edu/nixon/webtools.html>).

4.4. Results and Discussion

During sequencing of the ChfuGV genome library, a putative gene encoding a protein homologue to baculoviral IE-1 protein was identified. The *ie-1* gene from ChfuGV has a 1278 nucleotide open reading frame (ORF) encoding 426 amino acids with a predicted molecular mass of 50.33kDa (Fig.1). The amino acid composition in terms of non-polar (ACFGILMPVWY), polar (DEHKNQRST), basic (HRK), and acidic residues (DE) indicates a content of non-polar (48.4%), polar (51.6%), acidic (11%) and basic (16.7%). The protein is particularly rich in leucine (10.3%) and lysine (10.1%) residues. Table 1 presents the amino acid composition, predicted molecular mass and the isoelectric points of all known baculoviral IE-1 proteins.

The upstream region of *ie-1* gene from ChfuGV along with all other granuloviruses was studied with the aim of detecting all possible *cis*-acting control elements within this region. Former studies on the promoter region of *ie-1* gene from AcMNPV pointed out that these elements were located from 34 to 546 nt upstream from the start codon (Leisy and Rohrmann, 2000; Pullen and Friesen, 1995a). To identify all *cis*-acting DNA elements regulating the expression of ChfuGV *ie-1* gene along with other known granuloviral *ie-1* genes, we examined a region extending from nucleotide -1 to -800 within the noncoding leader region of all granuloviral *ie-1* genes. This analysis revealed the presence of several *cis*-acting regulatory elements within the promoter region of *ie-1* gene of ChfuGV and other studied granuloviruses.

When these regions in granuloviruses were compared to the same regions in NPVs, no significant preservation was observed between the studied sequences from the two genera. Furthermore, alignment of all granuloviral *ie-1* gene noncoding leader regions revealed that, although the similar motifs were identified in these regions, no significant level of conservation was observed amongst them (data not shown).

The tetranucleotide CAGT is the consensus baculovirus early transcriptional start site found in numerous baculovirus genes (Pullen and Friesen, 1995b; Blissard *et al.*, 1992). Current evidence suggests that the CAGT motif is an initiator element that has been shown to be required for proper *ie-1* transcription during AcMNPV infection. This motif plays a critical role in accelerating the expression of viral genes, such as *ie-1*, that are vital to launch a productive infection. These data suggest that in AcMNPV the CAGT motif affects the rate of *ie-1* transcription and indeed deletion of sequences that included the CAGT eliminated early *ie-1* transcription. It has been revealed that in AcMNPV the nucleotides surrounding the CAGT motif were adequate for appropriate transcription in a TATA-independent mode (Pullen and Friesen, 1995b). Although no CAGT motif was detected in *ie-1* promoter region of ChfuGV, we detected a similar motif (TCATT) centered at position -482. The motif (A/C/T)CA(G/T)T is a proposed consensus for arthropod transcriptional initiator elements. A large group of arthropod polymerase II promoters contains one or more copies of the pentanucleotide TCAGT or one of its cognates TCATT, ACAGT, and GCAGT. The high level of conservation of this motif at promoter region of arthropods implies the significance of these promoter elements (Cherbas and Cherbas, 1993). The presence of this motif in ChfuGV as well as CpGV

and XcGV suggests the importance of this motif as a transcriptional initiator element in granuloviruses.

The only TATA element in *ie-1* promoter region of ChfuGV is located 203nt upstream of start codon triplet. Potential TATA motifs were observed in *ie-1* promoter region of all studied GVs, however in only two of them (PxGV and CpGV) these elements located within a reasonable distance from the putative ATG translation initiation codon. A CCAAT element was located 92nt upstream of start triplet of *ie-1* gene in ChfuGV on the opposite strand. Another CCAAT element was also located at position –420nt. CCAAT elements were also detected in the promoter region of PxGV. Two late motifs (TAAG) were observed at positions –368 and –508 respectively from the putative start triplet. The late motifs were detected in all available GV promoter regions. A single polyadenylation signal AATAAA was located 18nt downstream of the putative translational stop codon of ChfuGV *ie-1* (Fig.1).

The multiple sequence alignment of ChfuGV and other granuloviral IE-1 proteins revealed that the C-terminal regions of these proteins are more conserved compared to the N-terminal regions (Fig.2). This finding was in agreement with the data from nucleopolyhedral IE-1 (Olson *et al.* 2001; Wang *et al.*, 2001). Available data on IE-1 proteins from NPVs propose that the C-terminal region of these proteins plays a critical role in DNA binding while the N-terminal regions of IE1 which is rich in acidic residues has been hypothesized to be an acidic activation domain (Olson *et al.* 2001; Leisy and Rohrmann, 2000; Rodems *et al.*, 1997; Theilmann, and Stewart, 1991; Forsythe *et al.*,

1998). Two stretches of amino acid sequences are well conserved in the C-terminal region of all granuloviral IE-1 proteins, those are $^{332}[R/H][Y/F]R[L/I]NCF[K/R]^{339}$, and $^{345}[L/V]W[I/V]NS[M/I]V^{351}$. Another well-conserved amino acid sequence $^{393}KL[V/I]IRY[I/V]L^{400}$ was detected in all GVs except PxGV. Although PxGV keep almost the same pattern, the second, fourth and the eighth residues were found to be shifted to F, M, and M respectively.

The deduced amino acid sequence of ChfuGV IE-1 was compared to other known baculoviral IE-1s. Results revealed a non-significant similarity between granuloviral and nucleopolyhedral IE-1 proteins (Table 2). Even the hydrophobicity plot analysis did not demonstrate any significant structural conservation between IE-1 proteins from GVs and NPVs (Fig.3). Despite the lack of significant homology between amino acid sequences of IE-1 proteins of GVs and NPVs, the ChfuGV IE-1 protein from shares a conserved domain (residues 387 to 414) with nucleopolyhedral IE-1s (Fig.4). This conserved domain was originally discovered within the C-terminal region of NPVs IE-1 proteins. Each domain contains two amphipathic α -helices preceded by a group of basic residues and forms a bHLH-like domain. The HLH protein family, of transcriptional factors, is characterized by highly conserved DNA binding, dimerization, and transactivation motifs and contains very important elements that have been identified in different living organisms ranging from yeast to human (Atchley *et al.*, 2000). Despite the high levels of similarity inside the conserved domain, HLH proteins usually exhibit considerable sequence divergence outside this region. It has been shown in AcMNPV IE-1 that the existence of the hydrophobic residues within the HLH-like domain in C-terminus is

essential for IE-1 oligomerization and contributes to IE-1 stability (Rodems *et al.*, 1997; Slack and Blissard, 1997). It seems that the hydrophobic surface of the amphipathic helices is directly involved in homophilic interaction between IE-1 monomers. The basic residues preceding the HLH-like domain of baculoviral IE-1 can be regarded as elements involved in DNA-binding activity. These residues are thought to interact through ionic interactions with the negatively charged phospho-sugar backbone during binding to DNA molecule.

A highly conserved stretch of five amino acids could be marked within this domain in most of the known baculoviral IE-1 proteins apart from CpGV, XcGV, and PxGV. The conserved sequence is [V/L][V/L]KL[V/I/L] (residues 391 to 395 in ChfuGV IE-1). Very similar sequences could be also detected in CpGV (VSKLV), XcGV (V-KLI) and PxGV (V-KFV) (Fig. 4). A multiple sequence alignment of deduced amino acid sequences of granuloviral and nucleopolyhedral IE-1 proteins revealed that, beside the basic helix-loop-helix like (bHLH-like) domain, no significant conservation could be detected amongst IE-1 proteins from the two genera of *Baculoviridae* family. IE-1 proteins from ChfuGV and CpGV showed a high degree of similarity and shared several distinct motifs within both C and N-terminals. Five such motifs were located within the N-terminal (⁶³N[K/H]F[R/K]TTT⁶⁹, ⁷¹HMF[I/V]C⁷⁵, ⁸²YVKNER[Y/F]⁸⁸, ¹⁰²GNA[F/Y]E[L/M]¹⁰⁷, and ¹¹³KF[F/Y]ITT[K/R]L[I/L]Q¹²²) while seven were found in C-terminal region (²²²[L/M]F[D/E][I/L][M/L]YKNG²³⁰, ³⁰³DV[D/E][D/E]F[L/M]R[I/L]S [L/V]³¹², ³¹⁷GDV[L/V]FNMK³²⁴, ³³¹QRYR[L/I]NCF[K/R][I/M]D³⁴¹, ³⁴⁵[L/V]W[I/V]NSMVY³⁵², ³⁸³M[L/M]SKLH³⁸⁸, and ³⁹³KL VIRY[I/V]LS[K/R]R⁴⁰³).

A phylogenetic analysis of IE-1 proteins is shown in Fig.5. Tree was produced via maximum parsimony to estimate the evolutionary relationship between baculoviral IE-1 proteins. These analyses show a clear division between the GV and NPV IE-1 proteins. Two main branches for NPVs were found in this tree, but due to the limitation in number of known granuloviral IE-1 proteins, the tree cannot present a complete picture of this genus. ChfuGV is positioned along side with CpGV in the same cluster, supported by a 99% bootstrap value.

This study has provided further information about the importance of the basic helix-loop-helix like domain in baculoviral IE-1 proteins. It also demonstrates that the presence of several potential *cis*-acting regulatory elements within the upstream region of the *ie-1* gene is a common property in all known baculoviral IE-1 proteins. It will be of interest to determine the cellular location of the *ie-1* gene product during the process of infection.

Figure 1. Nucleotide sequence and predicted amino acid sequence of *ie-1* gene from ChfuGV. 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.

1 - TGAAGGCACGCAACAAGTCGAATGGAACAAAACCGCTAACTAAGCACCTTACTTTATAGT - 60
 61 - ATTCGTTTTCATTTGCTTTAATTTTTTTGTATACCATAGCAGCGGGTAATTCACCTACCTT - 120
 121 - TTTTATACCAATGACGTACCAAAATCTTGGTATTGATTAGTAATTACTTTAACGACGT - 180
 181 - TAAGTTAATTTTATCTACCACGTCGGGTAGTAGACAAGGTGCGGAAACATGCTTATTA - 240
 241 - CGAATTTGTCGTGATGTGACTCGTTATTGATTGAAAAACGCAATTTATAGGATTGTTGA - 300
 301 - CGTCGAGACCGCTAACCGTTTTTGTTTTTTTATTGACTGGTTCCTTTTATAAATTTATAGT - 360
 361 - TAAACTGTGTAAGTGAATTTAGAGTGTGGAAAATTGAAATATTATATCTACGTCGTTGT - 420
 421 - TGGATTTGCTTTGCCTATTTGAAACATTGGAAATTTTTTTACAGGAATAACACTTATAT - 480
 481 - CGATATGGAAGAGGAGACGCCGTTTGAATTGTTGACCACAAACGATTTCGGAGTACGAAAG - 540
 541 - CGAAGAGAATGAAATCGGAAATGGACATATCTGAAGTTAAAACGGGTCGAGGTGCAGCCT - 600
 - M K S E M D I S E V K T G R G A A Y
 601 - ATTTGTTCAACCTTACGTCAATCCTGATAATCATGTCAATCCTTGATAATCATGTCAATC - 660
 - L F N L T S I L I I M S I L D N H V N P
 661 - CTGATAAAACTGAAAGCGTTCAAGAAACACAATTTCAAGATGAAGACGATAAATTAACAA - 720
 - D K T E S V Q E T Q F Q D E D D K L T I
 721 - TTTTTCGGTGGTTGAATCATTTTAGAACTACAACGTTTCACATGTTTATATGTAATTGGG - 780
 - F R W L N H F R T T T F H M F I C N W D
 781 - ACGGACTAAATATGTCAAAATGAACGGTATTGTGAAGATGTGTATTTAAGAAAGTACC - 840
 - G L K Y V K N E R Y C E D V Y L R K Y H
 841 - ATCAAGTCAATGGCAACGCTTATGAACTCCTGTTAAAGATTGTAAATTTTTCATTACTA - 900
 - Q V N G N A Y E L L L K D C K F F I T T
 901 - CAAAATTGATTAGCATGGCAAGATATTCTCATTTCCAAGCGAAAAAGTTATGAAGCATG - 960
 - K L I Q H G K I F S F P S E K V M K H V
 961 - TAGATTTAAAAATGTACACCGCTTTTACAGATTATTATGAAGGACGGCAACTCTTTAC - 1020
 - D L K I V T P L L Q I I M K D G N S L H
 1021 - ATCAACATATTATGGAAGCTTTGATTAACACTGTATTTATTGGGGTTTCTAATAACAGTG - 1080
 - Q H I M E A L I N T V F I G V S N N S V
 1081 - TTGGGGTGTAAAGCGACGAATGTAGCTGATAATTTCGGCAAGAGAATTAACAAGATTA - 1140
 - G V V K R R M L A D N S A K R I N K I T
 1141 - CCAACAGATATAGCTGGAATAGTGAAAAATAAAAAATTAGAAATCTGTGCGAAAGCAATG - 1200
 - N R Y S W N S E K I K I R N L C E S N D
 1201 - ATGTTCAATCACTGTTTGATATAATGTACAAAAACGGAGGCTATTACAAATTTACACAAG - 1260
 - V Q S L F D I M Y K N G G Y Y K F T Q D
 1261 - ACTTTACTTTGGATTACACTAAATTCATAGCCCAATTTAATTATAATGTGCAGCGATTAA - 1320
 - F T L D Y T K F I A Q F N Y N V Q R L K
 1321 - AAACAGATATAATAGAACATGTGGATAATATTAGTCAAACAAAAATGTGGTCAAAATTT - 1380
 - T D I I E H V D N I S Q T K M L V K N L
 1381 - TGGCGGCTTTGTTGTATAAACTGAAATCTAAGGGATATGCAAGAATTAAAGATATCA - 1440
 - A A L L Y K T E N L R D M Q E L K D I K
 1441 - AAGGTAATGAGCAGGACGTGGATGATTTTTTACGAATTTCTTTAAATTATCCAATGGGAG - 1500
 - G N E Q D V D D F L R I S L N Y P M G D
 1501 - ATGTGTTATTTAACATGAAACTAAAGGGACCAACAGTCAACGTTACCGGCTAAATTGCT - 1560
 - V L F N M K T K G T N S Q R Y R L N C F

1561 - TTAAATAGATCAAGTTTATTGTGGATTAATAGTATGGTTTATAAACTAGCAAGTTTG - 1620
 - K I D Q V Y L W I N S M V Y K T S K F D
 1621 - ATTTGGAATCTGTAATACTAAAATATAAACAGGAACACACTACATTATCAGTTTAAAT - 1680
 - L E S V I L K Y K T G T H Y I I S F K Y
 1681 - ATGTATTTAATACCATGTTAAGCAAATTACACTCTGAAGTGGTCAAATTGGTAATACGGT - 1740
 - V F N T M L S K L H S E V V K L V I R Y
 1741 - ATATTCTTTCTAAAAGAGAAATATAGTTTATTGGAAAATGATATCAAAATAACAACAAAT - 1800
 - I L S K R E Y S L L E N D I K I N N K L
 1801 - TACTTTACAAATGTTTAATATTTTCATAAATTTGGTTATGTAGAGAAATATAATGAGA - 1860
 - L Y K C L I F S *
 1861 - TTGAATTAAAACGTTAAAAGTTTTTAAAAATATTTTTTAAGATTTAATTCAAACTCATTA - 1920
 1921 - TAATTTTGTGTGTTTGACAAAAATTTTGTAAAACGTTGTGTAATATGAATCTGATTTA - 1980
 1981 - AAACGCTAAAAAATTTTAAAAGTTGTTTTTGGAAATTAAATCGCATTCAAATTTGTCAA - 2040

Figure 2. Alignment of the N-terminal (a) and C-terminal (b) of granuloviral IE-1 protein. 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. Accession numbers for the sequences used are as follow: CpGV AAK70674, XcGV CQ9PZ32; PxGV Q9DW20.

(a)

```

ChfuGV_ 1 -----MKSEMDISEVKTGRGAYLFNLTSLIIM
CpGV_ 1 -----MEVETTMVLEDOIMTTNDSENEVDCDEEEARGASHTFANQRFVKLV
XcGV_ 1 MDEFHTLSNIDSQMOLEDAHYEQNPITLSTDEENDDLEDSDECDSDTKLKEKGVSK
PxGV_ 1 -----MSDILTTIRDSEDEECDDIEEP-----

ChfuGV_ 30 SILDNVNPD-----TESVQETGFQDEDDKTIFRWLNHFPTTF
CpGV_ 48 NNP EQQPPPREQQQPSELNRIFLP-INYNKTTTNDDAEDCVSENYRWFNRKTTTC
XcGV_ 61 SSS FETMPDSLLPFNKRSFVLPEITNALSTSSSSKCHQWPAPDNTCPWLSMTTTN
PxGV_ 22 -----SPAKQPRDNNGAEDWILNKTTLI

ChfuGV_ 71 HMFICNWDGKYVKNERPCEDVYLRKYHQVNGNAYELLKDCKFITTKLIQGNFSFP
CpGV_ 107 HMFVCHSR-VEYVKNERFPAEVYMQYHELYGNAENTIKDAKFYITPRLQLLKVD--
XcGV_ 121 HMFVCHSD-VDYVKNIPFGSALYIDNYHTKYGNTRYEHIHS-----
PxGV_ 47 NMFVCHDS-FEYYSVRHPSSVYMMFYROYGSAYELHYNHIRVYMTENCLERNGVGKNR

ChfuGV_ 131 SEKVMKHVDIKI-----
CpGV_ -----
XcGV_ -----
PxGV_ 106 YPNESVQVNVVSEQVFLGLKVDVK

```


(b)

```

ChfuGV_ 1 -----VT-- PLLQIMKDGNSIHOHMEALINTVF
CpGV_ 1 -----MWHPAES IHHVDSVTSARILNLRRPDGYTVOQHTIEKLFYLL
PxGV_ 1 -----GVTERLTCAAIVLQFNLIIN
XcGV_ 1 KLYVTSTLLKINNSPNDLYENOMESIKDAAAI SYLLNLCTKNGSTVMTQILLFHSINL

ChfuGV_ 29 IGSNNVGVVVRPMLADNSAKPIKNTNPYSWSEKININLCESN-----
CpGV_ 45 IGLGPRDISAVMEPINNKATEPCALQSKYSWSEERKLNVEFRNSTVASVTANINGS
PxGV_ 22 EGLDNETINIVKNLINATIIINRDNKPKPPNEDTDPIAKLFES-----
XcGV_ 61 LSLSTEQIDSIHNTLP FVHRLAHMKNYSWKSELYKISKIKTIP-----

ChfuGV_ 76 -----DVQSLFDIMYKNGGYK-----FTDFTLDYTKFIAEFNYNVQRILKTD
CpGV_ 105 NSNTTCHTTNKSIEKIEELLYKNGNTPP-----ETEDKELAELEYRREHQNVCOLNKH
PxGV_ 69 -----ARLEYF-----VEDKEYTHQSKYIENMAIN-----N
XcGV_ 108 -----GQTHIQLLKASTYPRNKPLEARKQSYVHSPYVCEFYDAMQPTNIT

ChfuGV_ 119 IIEHVDNISITRMIVNNAAALYKHELRDMQELNDIKGNEQDVDDFLPISLNYPMGDV
CpGV_ 160 VTEYMDKVSNNVFLINDAMVLYKHEKQELONLKNEYGTEADVEEFMRISVAHPRCDV
PxGV_ 96 NINVQDNVSNITWYFKIVNVICPS-EEELNFKSFDCDYTTKKFIMSRMNKKLNV
XcGV_ 156 TAFTPDPTVTHVKWLESATVKSIVNPP-MSFRLKTFDSTPNDAEFLPASKSHPNVI

ChfuGV_ 179 FNMSTNCTNSQYPLNCFKIDQYWINSMVY---KISKFDLESVILNYKTGTHYIISFN
CpGV_ 220 FNMKVEDTNTQRYPLNCFRMDSVHVVNSMVYS--VQQFNLKKTIQPHWGTHHIFQFD
PxGV_ 155 LHTKVNIVNKLHYRLNCFKLCDDHVWINSIVLNKEESKKLDLENLIFREKNGSEYTIINLK
XcGV_ 215 LHTAPITGKEHERLNCFKLNKHVWINSMVYDKIQHNKVDLKQIIEVAVGEHHIISFQ

ChfuGV_ 236 YVENIMLSKLHSEVVKLVIRYILSPREYSLLENDIKI-NKLLYKCLIFS
CpGV_ 278 YMANSMLSKLHAEVSNLVIRYVLSRPSFDLLQNDCS--LKLSYKIVIE
PxGV_ 215 YMNQELRPVHMETVNFVVKYIERPDALLKNDIDC-SKLYKCIREF-
XcGV_ 275 YIHQCKLKQTHHETVKLIPYILEPPNFIKLAQVAFCPRKIKKDYIIF-

```

Figure 3. Hydrophobicity profiles using Kyte-Doolittle from C-terminal regions of all known granuloviral and two nucleopolyhedral IE-1 proteins. Above the axis (+) denotes hydrophilic regions and below the axis (-) indicates hydrophobic regions.

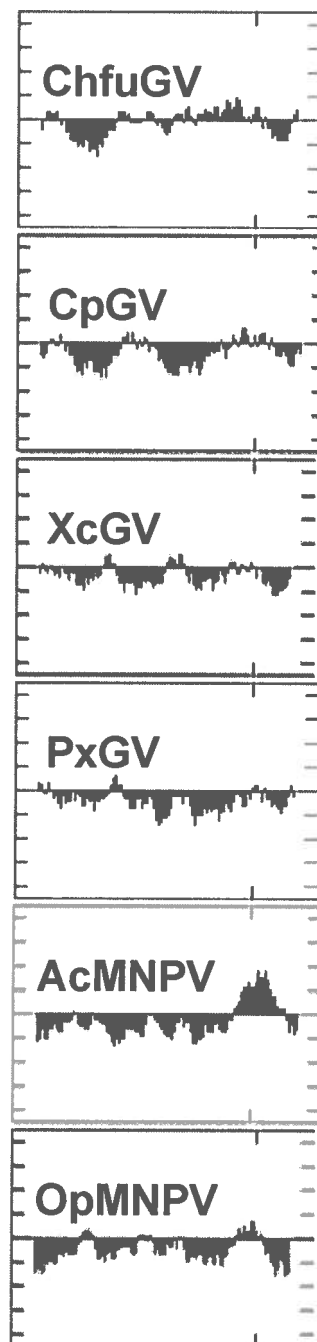


Figure 4. (a) HLH-like domain of IE-1. ChfuGV IE-1 residues 387 to 414 were aligned with equivalent IE-1 residues from NPVs and other GVs. Conserved hydrophobic (ϕ) residues are illustrated. (b) Basic residues (+) preceding the HLH-like domain of ChfuGV IE-1 are depicted.

a)

	Helix	Loop	Helix
	ϕ $\phi\phi$ $\phi\phi$ ϕ $\phi\phi$		ϕ ϕ ϕ
ChfuGV	LHSE-VVKLV-IRYIL	SKREYS	-LLEND---IK--I
CpGV	LHAE-VSKLV-IRYVL	SRRSFD	-LLQNDCKSLK--L
PXGV	VHMETV-KFV-MKYIM	ERRD--	-LA-----L---L
XcGV	MHIETV-KLI-IRYIL	ERRNFT	KLAK-D---V---A
AcMNPV	LHNN-LLKLLAL--IL	QGLVP-	-L--SD---AITFA
OPMNPV	WHNN-LLKLLQL--LL	QNL-IR	-L--ED---VQSYS
CfNPV	WHNN-LLKLLQL--LL	QNLRIK	-I--DD---VQQYS
LdMNPV	RHNG-LLKLVS---LY	TSAAVD	-LSV-----LVEIA
BmNPV	LHNN-LLKLL-AALIL	QGLVP-	-LS--D---AITFA

b)

	Basic	Helix	Loop	Helix	
	+ + + + +				
365	KYKTGTHYIISFKYVFNTMLSK	LHSEVVKLVIRYIL	SKREYS	LLENDIKINN	414

Figure 5. Phylogenetic analysis performed on baculoviral IE-1 proteins. The unrooted tree was produced by a Branch and Bond 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; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e addition sequences: furthest. Accession numbers for the sequences used are as follow: CpGV AAK70674, XcGV CQ9PZ32; PxGV Q9DW20; HaNPV AAK96269; SINPV AAL01702; HzNPV 010619; AcMNPV P11138; BmNPV P33245; EpNPV AAK85693; OpMNPV 22114; CfNPV P41716; SeNPV Q9J811; LdMNPV 036454; TnSNPV AAL14888.

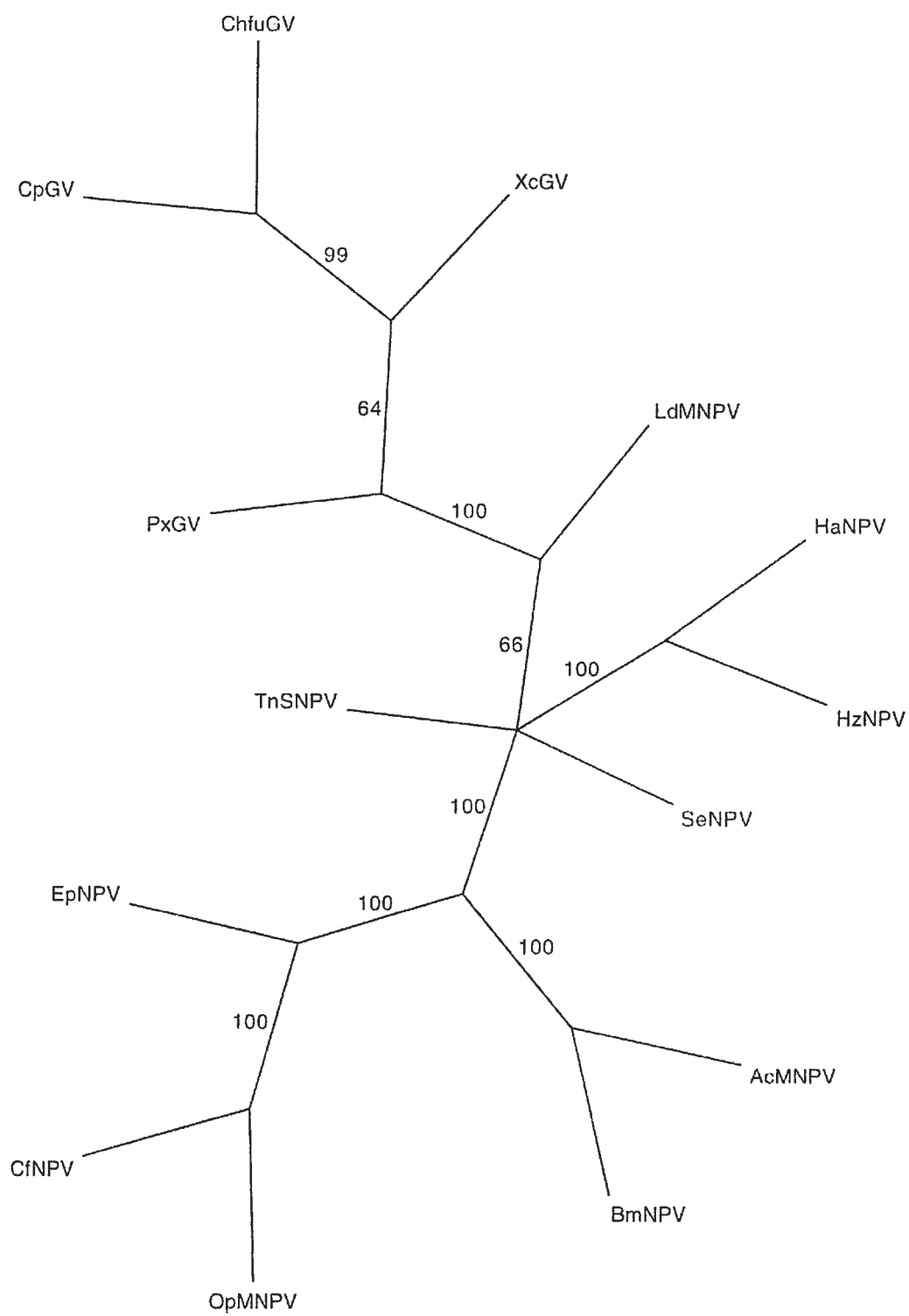


Table 1. Amino acid compositions (%), isoelectric points and molecular mass of ChfuGV and other baculoviral IE-1 proteins. Peptide Statistic and ProtParam tools were used for calculation.

	ChfuGV	CpGV	XcGV	PxGV	AcMNPV	OpMNPV	BmNPV
(%)							
Non polar	48.4	43.9	45.2	47.6	46.4	46.6	46.7
Polar	51.6	56.1	54.8	52.4	53.6	53.4	53.3
Acidic	11.0	13.1	10.7	13.7	11.5	12.3	11.8
Basic	16.7	17.4	16.9	17.0	12.9	14.6	12.3
(%)							
Ala	2.1	3.9	3.9	2.5	5.3	6.6	5.7
Arg	4.0	6.1	3.9	4.6	2.2	5.2	2.1
Asn	8.0	8.0	6.4	9.4	9.5	8.6	9.9
Asp	6.3	5.3	6.6	6.9	5.3	6.4	5.7
Cys	1.4	1.6	1.2	2.3	1.5	0.9	1.5
Gln	3.8	5.3	4.1	2.8	4.0	4.8	3.9
Glu	4.7	7.8	4.1	6.9	6.2	5.9	6.2
Gly	3.3	2.3	1.7	2.0	2.4	3.4	2.6
His	2.6	3.7	5.2	2.5	2.7	1.8	2.2
Ile	8.7	4.3	7.2	6.6	4.6	3.6	4.6
Leu	10.3	7.8	8.9	9.2	7.7	8.2	7.7
Lys	10.1	7.6	7.9	9.9	7.9	7.7	8.0
Met	3.5	4.1	3.1	3.8	1.9	3.4	2.1
Phe	5.4	4.3	4.3	4.1	5.8	4.8	5.7
Pro	0.9	2.0	4.3	2.5	3.1	3.4	3.3
Ser	6.3	6.1	8.7	5.3	9.3	7.5	8.9
Thr	5.9	6.1	7.9	4.1	6.5	5.5	6.3
Trp	0.9	1.0	1.2	0.8	0.2	0.4	0.2
Tyr	5.4	4.9	4.3	5.6	6.4	4.8	6.3
Val	6.3	7.6	5.0	8.1	7.4	7.1	7.2
Isoelectric Point	9.17	8.23	8.61	8.19	6.09	8.28	5.79
Molecular Mass(kDa)	50.33	57.92	56.42	46.60	66.88	64.31	66.92

	EpNPV	LdMNPV	SeMNPV	TnSNPV	HaNPV	CfNPV
(%)						
Non polar	46.03	50.53	44.26	41.68	45.84	44.46
Polar	53.96	49.47	55.74	58.32	54.16	55.54
Acidic	12.59	15.19	12.47	11.77	10.89	12.68
Basic	13.31	15.37	15.47	15.43	16.04	14.11
(%)						
Ala	4.9	8.8	4.1	4.1	3.8	3.9
Arg	4.1	6.2	3.5	3.0	4.7	4.6
Asn	10.6	4.9	8.7	13.1	9.7	8.2
Asp	5.8	6.2	7.1	7.3	4.7	5.9
Cys	0.9	2.5	0.8	1.2	0.6	0.9
Gln	3.8	3.9	3.5	3.7	3.9	5.2
Glu	6.8	9.0	5.3	4.5	6.2	6.8
Gly	2.2	3.0	3.1	2.0	2.4	2.9
His	2.2	1.4	2.9	1.9	3.0	2.0
Ile	4.9	4.6	7.6	7.7	6.4	5.2
Leu	8.8	8.0	6.4	6.2	7.9	7.5
Lys	7.0	7.8	9.1	10.6	8.3	7.5
Met	4.0	2.7	4.1	2.4	3.8	2.9
Phe	4.7	4.2	4.5	4.3	4.2	5.2
Pro	3.6	4.4	3.6	4.3	4.7	3.0
Ser	6.3	6.4	9.2	8.4	8.6	8.9
Thr	7.4	3.7	6.3	6.0	5.0	6.4
Trp	0.5	0.9	0.3	0.3	0.3	0.4
Tyr	5.0	4.6	4.1	3.2	4.2	5.2
Val	6.7	6.9	5.7	5.8	7.6	7.5
Isoelectric Point	5.92	5.85	7.64	8.92	9.17	6.46
Molecular Mass(kDa)	64.63	65.08	81.94	84.52	76.49	64.95

4.5. Acknowledgements

The authors would like to thank Dr. Anne Bruneau, Department of Biology, University of Montreal for helping us in the phylogenetic studies. This research was funded by the *Ministère des Ressources Naturelles du Québec* Grant number 0316-2405 to GC.

CHAPTER 5 – PUBLICATION #4

Identification, Characterization and Phylogenic Analysis of Conserved Genes within the p74 Gene Region of *Choristoneura fumiferana* Granulovirus Genome

Authors: Kianoush Khajeh Rashidan^{1*}, Nasha Nassoury² Paresa N. Giannopoulos¹, Yves Mauffette³ and Claude Guertin¹

1- Institut national de la recherche scientifique-Institut Armand-Frappier, 531 Blvd des Prairies, H7V 1B7, Laval, Canada

2- Université de Montréal, 4101 Sherbrooke east, H1X 2B2 Montreal, Canada

3- Université du Québec à Montréal, Succursale Centre Ville CP8888, H3C 3P8 Montreal, Canada.

The nucleotide sequence data reported in this paper has been submitted to the GenBank nucleotide sequence database and has been assigned the accession numbers ANN77187, ANN77188, ANN77189, ANN77190, ANN77191, ANN77192, ANN77193, ANN77194, ANN77195, ANN77196, ANN77197, ANN77198, ANN77199, ANN77200.

Published in Journal of Biochemistry and Molecular Biology in 2004. Number 37, pages 700-708.

Contribution of the authors

I was the principal investigator in this report while Dr. Nassoury's expertise in the screening used to locate the clone was used in this report. Mrs. Giannopoulos helped me by performing some of the PCR reactions. Dr. Mauffette was the co-director of this project with whom I had several scientific discussions prior the submission of this paper. Dr. Guertin was the director of the project and the majority of the experiences have been realized in his laboratory.

Abstract in French :

Identification, caractérisation et analyse phylogénétique des gènes conservés dans la région du gène *p74* du granulovirus de *Choristoneura fumiferana*

Les gènes localisés dans la région du gène *p74* de *Choristoneura fumiferana* granulovirus (ChfuGV) ont été identifiés par séquençage du fragment de restriction de 8,9 kb de la banque génomique *Bam*HI du ChfuGV. Le contenu global en GC de cette région du génome était 33,02%. Les ORF à l'intérieur de la région du gène *p74* et leurs orientations transcriptionnelles sont présentés. Cette région contient un nombre total de 15 ORF. Parmi ceux-ci, 8 ORF semblent être homologues aux ORF baculoviraux communs tels que Cf-i-*p* , Cf-vi, Cf-vii, Cf-viii (*ubiquitin*), Cf-xi (*pp31*), Cf-xii (*lef11*), Cf-xiii (*sod*) and Cf-xv-*p* (*p74*). Les ORF les plus importants localisés dans cette région du génome sont ceux des gènes *ubiquitine*, *lef-11*, *sod*, *p10* et *p74*. Les produit conceptuels d'ORF majeurs hautement conservés ont été utilisés dans cette étude pour la construction d'arbres phylogénétiques. Les résultats de cette étude supportent les résultats antérieurs pour la classification des granulovirus en deux groupes distincts : Groupe I : *Choristoneura fumiferana* granulovirus (ChfuGV), *Cydia pomonella* granulovirus (CpGV), *Phthorimaea operculella* granulovirus (PhopGV) et Groupe II : *Xestia c-nigrum* granulovirus (XcGV) et *Plutella xylostella* granulovirus (PxGV). Les protéines conservées du ChfuGV sont plus rapprochées à celles du CpGV et du PhopGV. Des études comparatives ont été réalisées sur l'arrangement des gènes à l'intérieur de cette région génomique, démontrant que les GV du groupe I maintiennent des arrangements de leurs gènes similaires.

5.1. Abstract

The genes located within the p74 gene region of *Choristoneura fumiferana* granulovirus (ChfuGV) were identified by sequencing an 8.9 kb *Bam*HI restriction fragment on ChfuGV genome. The global GC content of this region of the genome was 33.02%. The ORFs within p74 gene region and their transcriptional orientations is presented. This region contains a total number of 15 open reading frames (ORFs). Among those, 8 ORFs were found to be homologues to baculoviral ORFs: Cf-i-*p* , Cf-vi, Cf-vii, Cf-viii (*ubiquitin*), Cf-xi (*pp31*), Cf-xii (*lef-11*), Cf-xiii (*sod*) and Cf-xv-*p* (p74). To date no specific function has been assigned to ORFs: Cf-i, Cf-ii, Cf-iii, Cf-iv, Cf-v, Cf-vi, Cf-vii, Cf-ix and Cf-x. The most noticeable ORFs located in this region of ChfuGV genome were *ubiquitin*, *lef-11*, *sod*, fibrillin and p74. Phylogenetic trees (constructed using conceptual products of major conserved ORFs) and gene arrangement in this region were used in order to further study the classification of the members of granulovirus genus. Comparative studies demonstrated that ChfuGV along with *Cydia pomonella* granulovirus (CpGV), *Phthorimaea operculella* granulovirus (PhopGV), *Adoxophyes orana* granulovirus (AoGV) and *Cryptophlebia leucotreta* granulovirus (ClGV) share a high degree of amino acids sequence and gene arrangement preservation within the studied region. These results support our previous results to classify granuloviruses in 2 distinct groups: Group I: ChfuGV, CpGV, PhopGV and AoGV and Group II: *Xestia c-nigrum* granulovirus (XcGV) and *Plutella xylostella* granulovirus (PxGV). Phylogenetic and gene arrangement studies also placed ClGV as a novel member of Group I granuloviruses.

Key words: Baculovirus, *Choristoneura fumiferana* granulovirus, phylogeny, gene arrangement.

5.2. Introduction

Granuloviruses (GV) are members of the family *Baculoviridae* which also includes nucleopolyhedroviruses (NPV) (Miller, 2001). The GV genome is a double-stranded super-coiled DNA, usually greater than 100 kb, that is packaged within a single nucleocapsid (Lange and Jehle, 2003; Wormleaton *et al.*, 2003; Luque *et al.*, 2001; Hashimoto *et al.*, 2000; Hayakawa *et al.*, 1999). Generally, there is only one nucleocapsid per envelope and together they are occluded within a paracrystalline protein matrix called granulin. Like polyhedrin of NPVs, granulin protects the virions against environmental effects and is involved in the horizontal transmission of the virus (Miller, 2001).

Baculoviral gene expression happens generally in three steps, namely early, late and very late, controlled in a regulatory cascade manner (Hess and Falcon, 1997; Blissard, 1996). All genes are located randomly throughout the genome, even though their products fill distinct regulatory or structural functions (Lange and Jehle, 2003; Wormleaton *et al.*, 2003; Chen *et al.* 2002; Li *et al.*, 2002; Luque *et al.*, 2001; Pang *et al.*, 2001; Hashimoto *et al.*, 2000; Gomi *et al.*, 1999; Hayakawa *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Ahrens *et al.*, 1997; Ayres *et al.* 1994).

Choristoneura fumiferana granulovirus (ChfuGV) is pathogenic towards the spruce budworm (*Choristoneura fumiferana*, (Clem.), (Lepidoptera: Tortricidae). Spruce budworms are responsible for the defoliation of coniferous trees in eastern North American forests. ChfuGV has been considered as an alternative to chemical insecticides, for control of the spruce budworm populations.

Knowledge about the molecular characteristics of ChfuGV is limited and in order to enhance our understanding on the nature of ChfuGV, our efforts are being concentrated on molecular characterization of this virus. In our previous communication we reported the identification and characterization of the ORFs located within the *odvp-6e/odv-e56* gene region of ChfuGV (Rashidan *et al.*, 2004). In present paper we report the identification and characterization of the ORFs located within the p74 gene region of ChfuGV. The gene p74 is a highly conserved baculoviral gene coding for an envelope protein, the complete sequence and characterization of this gene and its product in ChfuGV has been reported recently (Rashidan *et al.*, 2003). ORFs located within the p74 gene region of ChfuGV are positioned downstream to those within the *odvp-6e/odv-e56* gene region.

5.3. Materials and Methods

5.3.1. In vivo Production and Purification of Virus and DNA Extraction

Choristoneura fumiferana fourth-instar larvae have been infected by ChfuGV using virus contaminated artificial diet (Forté *et al.*, 1999). The virus isolation from infected larvae was carried out as previously described (Bah *et al.*, 1997).

5.3.2. DNA extraction, cloning and sequencing

ChfuGV genomic DNA was extracted from purified enveloped nucleocapsids as we have already described (Rashidan *et al.*, 2002). An 8.9 kb *Bam*HI restriction fragment were cloned into a pBlueScript-SK⁺ cloning vector (Stratagen) and transformed in to *E.coli* XL-1 Blue (Stratagen) and sequenced using techniques explained previously (Rashidan *et al.*, 2002).

5.3.3. Computer Analysis

Sequence data assembly and analysis were performed with Sequencher software version 4.0.5 (Gene Codes Corp.) and MacVector program version 4.5.0 (Eastman Kodak). Nucleotide sequence and its predicted amino acid sequence were compared to homologues in GenBank/EMBL and SWISSPROT by using BLAST (Altshul *et al.*, 1990). CLUSTALX (1.81) (Thompson *et al.*, 1997) was used for multiple amino acid sequence alignments. BOXSHADE version 3.21 was used for similarity shading and scoring among the aligned sequences. The alignments were used then as the input to construct the phylogenetic tree with Branch and Bound search settings of PAUP 4.0b4a

employing Neighbor joining (NJ) method (Swofford, 2000). The reliability of the phylogenetic trees was evaluated by bootstrap analysis with 1000 bootstrap replicates.

5.3.4. Assignment of putative ORFs

The information obtained from sequencing the 8.9 kb genomic fragment of ChfuGV was used to detect homologue sequences in the Genbank database (whole database and restricted database containing only virus sequences) using BLASTN (Altshul *et al.*, 1990) and Fasta3 (Pearson and Lipman 1988) programs. Alternatively, putative ORFs were first translated from ChfuGV nucleotide sequences and amino acid sequences were compared with existing sequences in the Genbank database using BLASTP program (Altshul *et al.*, 1990). These data and the one obtained from the gene order from other granulovirus genomes were used to identify ChfuGV ORFs.

5.4. Results and Discussion

The open reading frames (ORFs) located within the p74 gene region of *Choristoneura fumiferana* granulovirus (ChfuGV) were identified by sequencing an 8.9 kb *Bam*HI restriction fragment on ChfuGV genome. The global GC content of this region of the genome was 33.02%. The studied region contains a total number of 15 ORFs and among those, 8 ORFs were found to be homologues to baculoviral ORFs such as

Cf-i-*p* (*p* for partial), Cf-vi, Cf-vii, Cf-viii (*ubiquitin*), Cf-xi (*pp31*), Cf-xii (*lef11*), Cf-xiii (*sod*) and Cf-xv-*p* (p74). To date no specific function has been assigned to ORFs Cf-i, Cf-ii, Cf-iii, Cf-iv, Cf-v, Cf-vi, Cf-vii, Cf-ix and Cf-x.

One of identified ORFs codes for a conceptual protein similar to ubiquitin. Ubiquitin is a globular protein found in all eukaryotic cells and whose sequence is extremely well conserved from simple organisms to more sophisticated organisms such as vertebrates (*ubique*, Latin for everywhere). A putative ORF (Cf-viii) of 288 nt located 7 ORFs upstream to p74 potentially encodes a baculoviral ubiquitin-like protein homologue. This is a highly conserved gene present in all known baculoviral genomes (Lange and Jehle, 2003; Wormleaton *et al.*, 2003; Chen *et al.* 2002; Li *et al.*, 2002; Luque *et al.*, 2001; Pang *et al.*, 2001; Hashimoto *et al.*, 2000; Gomi *et al.*, 1999; Hayakawa *et al.*, 1999; Ijkel *et al.*, 1999; Kuzio *et al.*, 1999; Ahrens *et al.*, 1997; Ayres *et al.*, 1994).

Ubiquitin plays essential roles in several cellular processes in eukaryote cells, such as ATP-dependent selective degradation of cellular proteins, maintenance of chromatin structure, regulation of gene expression, programmed cell death, stress response, survival

of protein-damaging stresses and ribosome biogenesis DNA repair. The conjugation of ubiquitin to other cellular proteins regulates a wide range of eukaryotic cell functions.

The key role of ubiquitin in these procedures is to provide a tag to mark proteins for degradation by the proteasome. The covalent attachment of ubiquitin to targets necessitates a series of enzymes which consist of a ubiquitin activating enzyme, conjugating enzymes, and usually a ubiquitin-protein ligase. Biochemically this process is characterized as a cascade of events in which ubiquitin is passed from an activating protein, E1, to the conjugating protein E2 and then finally to the targeted substrate, sometimes with the assistance of a ligating protein E3 (Pickart, 2004; 2001).

Targets of the ubiquitin system include proteins which are malfunctioning due to mutation or exposure of the cell to physical or chemical stresses, or naturally short-lived proteins such as cellular regulators. In the latter class, many oncoproteins, the tumor suppressor p53, mitotic cyclins and several cell surface receptors are turned over by the ubiquitin-dependent proteolysis system. Sorting signals establish the location and destiny of proteins. These signals can be detected inside the amino acid sequence of a protein or they can be added to the protein in posttranslational steps. In the light of recent findings, it is now apparent that the ubiquitin can operate as a regulated sorting signal at different levels of the endosomal and biosynthetic pathways.

Beside the chief function of ubiquitin which is to target proteins for degradation by the proteasome, other noncanonical functions have been also proposed. Among other functions, the role of ubiquitin in controlling protein trafficking within the cell has been

established. The role of ubiquitin in protein transport has been come to attention when it was revealed that ubiquitination is necessary for the endocytosis of a number of yeast plasma membrane proteins (Hicke, 2001; Pickart, 2001). In most species, there are many genes coding for ubiquitin. However they can be classified into two classes. The first class produces polyubiquitin molecules consisting of exact head to tail repeats of ubiquitin. The number of repeats is variable. There are a number of proteins which are evolutionary related to ubiquitin, including ubiquitin-like proteins from baculoviruses proteins (Archibald *et al.*, 2003; Hicke, 2001).

Another important ORF found in the p74 region is ORF Cf-xiii with the potential to code a 161 amino acid polypeptide homologue to baculoviral copper/zinc superoxide dismutase (SOD). This ORF is located two ORFs upstream to p74 on ChfuGV genome. Three evolutionarily distinct families of SODs are known, depending on the metal cofactor: Cu/Zn, Fe and Mn types. The Fe and Mn forms are alike in their primary, secondary and tertiary structures, but are different from the Cu/Zn form. Prokaryotes and protists contain Mn, Fe or both types, while most eukaryotic organisms utilize the Cu/Zn type (Van Camp *et al.*, 1990). ChfuGV SOD homologue along with other baculoviral SOD homologue proteins is Cu/Zn type as well. SOD is the best known and perhaps most important of the antioxidant enzymes. SOD is a ubiquitous metalloprotein that catalyzes the destruction (dismutation) of the $O_2^{\cdot -}$ free radical ($2O_2^{\cdot -} + 2H^+ \rightarrow O_2 + H_2O_2$). It converts the very harmful superoxide free radicals to the less active peroxide, which then can be further converted by other antioxidant enzymes into water. The action of SOD results in protection of the biological integrity of cells and tissues against damages caused

by oxygen-mediated free radicals by catalyzing the dismutation of superoxide into molecular oxygen and hydrogen peroxide (Petkau *et al.*, 1975; Malmström *et al.*, 1975; Schinina *et al.*, 1989). SOD is widespread in nature and present in all oxygen-metabolizing cells (Gregory *et al.* 1974). Superoxide dismutase extracted from Bovine erythrocytes consists of two subunits of identical molecular weight joined by a disulfide bond. The molecular weight is 32.5 kDa and there are two Cu^{2+} and two Zn^{2+} per molecule (Bannister *et al.* 1971; Keele *et al.* 1971). Genes coding SOD have been reported in several eukaryotic viruses as well, for example in vertebrate poxviruses a gene encodes homologues of cellular cupro-zinc superoxide dismutases (Cu-Zn SOD) has been reported. Sequencing of the *sod*-like genes from various poxviruses demonstrated that the protein is highly conserved in all viruses sequenced (Almazan *et al.*, 2001; Cao *et al.*, 2002).

The gene encoding SOD is almost present in all baculoviruses. The only exception is *Epiphyas postvittana* nucleopolyhedrovirus (EpNPV) which lacks a homologue of the superoxide dismutase (*sod*) gene (Hyink *et al.*, 2002). The role of SOD in baculovirus infection is yet to be established but considering the function of SOD as an antioxidant enzyme, one may speculate that baculoviral SOD might play a role in the anti-apoptotic process that takes place during the virus infection cycle. Oxidants may be essential biochemical intermediates in the progression of many forms of apoptosis induced by different stimuli. Exposure of various cell types to oxidative stress-causing agents can directly induce apoptosis which can be blocked by a wide range of antioxidants such as SOD (Atlante *et al.*, 2003; Hockenbery *et al.* 1993; Buttke and Sandstorm, 1994; Wolf *et al.*, 1994).

Another identified ORF (Cf-xi) potentially codes for protein known as baculoviral 39k/pp31 protein, a nuclear matrix associate protein required for late expression during the baculoviral infection cycle (Gong *et al.*, 1998). Baculoviral 39k/pp31 protein is thought to play a role either at transcriptional level or at that of mRNA stabilization. (Lu and Miller, 1995).

Another conserved baculoviral gene homologue present in this region of ChfuGV genome is a putative late expression factor (*lef-11*). The baculaviral *lef-11* gene has been shown to be a gene of a great importance for baculoviral late gene expression and is essential for viral DNA replication during the infection cycle to support optimal levels of transient expression from baculoviral late promoter (Lin *et al.*, 2001; Lin and Blissard 2002).

The ORF encoding the ChfuGV LEF-11 homologue is situated three ORFs upstream of ChfuGV p74 and has a potential to encode 92 amino acids. Phylogenetic analysis of LEF-11 proteins separated GVs from NPVs, and demonstrates that ChfuGV LEF-11 is most closely related to the *Phthorimaea operculella* granulovirus (PhopGV), *Cryptophlebia leucotreta* granulovirus (ClGV), and *Cydia pomonella* granulovirus (CpGV) homologues (Fig.2).

Sequence analysis of the p74 gene region also revealed the presence of an ORF (Cf-xiv) of 315 nucleotides with homology to baculoviral fibrillin proteins. This protein owed its name to its fibrillar structure (van Oers and Vlak, 1997). This gene has been located

adjacent to p74 gene which encodes a highly conserved viral envelope protein. The presumed ChfVGV fibrillin gene codes for a 104 amino acids protein with a molecular weight equal to 11811 Da. To date several fibrillin genes have been identified in different baculoviruses, such as *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) (Kuzio *et al.*, 1984), *Orgyia pseudotsugata* multicapsid nucleopolyhedrovirus (OpMNPV) (Leisy *et al.*, 1986), *Bombyx mori* nucleopolyhedrovirus (BmNPV) (Hu *et al.*, 1994), *Perina nuda* nucleopolyhedrovirus (PnNPV) (Genbank/EMBL accession no. AAC13879), *Spodoptera exigua* nucleopolyhedrovirus (SeNPV) (Zuidema *et al.*, 1993), *Choristoneura fumiferana* nucleopolyhedrovirus (CfNPV) (Wilson *et al.*, 1995), *Hyphantria cunea* nucleopolyhedrovirus (HcNPV) (Lee *et al.*, 1996a), *Lymantria dispar* nucleopolyhedrovirus (LdNPV) (Kuzio *et al.*, 1999), *Spodoptera litura* nucleopolyhedrovirus (SlNPV) (Faktor *et al.*, 1997; Wei *et al.*, 1998), *Mamestra configurata* nucleopolyhedrovirus (McNPV) (Li *et al.*, 1997; Li *et al.*, 2002), *Trichoplusia ni* nucleopolyhedrovirus (TnNPV) (Fielding and Davison, 2000; Fielding *et al.*, 2002), *Buzura suppressaria* nucleopolyhedrovirus (BsNPV) (van Oers *et al.*, 1998), *Helicoverpa armigera* nucleopolyhedrovirus (HaNPV) (Wang *et al.* 2001), *Helicoverpa zea* nucleopolyhedrovirus (HzNPV) (Chen *et al.*, 2002), *Anticarsia gemmatilis* nucleopolyhedrovirus (Razuck *et al.*, 2002), EpNPV (Genbank/EMBL accession no. AAK85684) and *Rachiplusia ou* nucleopolyhedrovirus (RoNPV) (Genbank/EMBL accession no. AAN28152).

The Baculoviral fibrillin gene product is generally produced throughout the late stages of the baculoviral infection cycle in substantial quantities along with the baculoviral occlusion matrix protein (granulin or polyhedrin). Studies have shown that fibrillin might be implicated in nuclear disintegration, occlusion body formation and release of the occlusion body (Rohrmann, 1986; Williams *et al.* 1989; van Oers *et al.*, 1993; Lee *et al.*, 1996b). It has also been shown that fibrillin is not a structural protein, although, it can be found in fibrillar structures within the cytoplasm of the infected insect cells (Vlak *et al.*, 1988; van Oers *et al.*, 1994). Several studies have shown that fibrillin genes in AcMNPV, OpMNPV, CfMNPV, SeNPV, and SINPV are located adjacent to p74 genes on the genome of these viruses (Ayers *et al.*, 1994; Leisly *et al.*, 1986; Ahrens *et al.*, 1997; Wilson *et al.*, 1995; Zuidema *et al.*, 1993; Faktor *et al.*, 1997).

Although the gene is conserved between ChfuGV and other baculoviruses, comparative sequence analysis of various fibrillin proteins shows significant deviation in the amino acid sequences (data not shown). A relatively higher level of similarity in their secondary structures, such as the existence of a coiled-coil region in the N-terminal, can be noticed in the majority of the fibrillin proteins. This similarity suggests a parallel function for baculoviral fibrillin protein homologues. Not only the gene homology but also the location of fibrillin genes may give information on the relatedness among baculoviral fibrillin homologues. Comparative studies revealed that the arrangement of p74 and fibrillin genes in ChfuGV was identical to those of AcMNPV, OpMNPV, SeNPV, TnNPV, CfNPV, McNPV, SINPV, EpNPV, and RoNPV (data not shown). The arrangement of p74 and fibrillin genes in LdNPV was different although they were

located within the same region of the genome and have maintained similar direction of transcription compared to the abovementioned baculoviruses.

ChfuGV fibrillin homologue contains a 12-fold EP (glutamic acid and proline) dipeptide repeat within an EP-rich region at its conceptual C-terminal. These EP repeats have been also observed in the C-terminal region of *Trypanosoma sp.* procyclins. Procyclins, the product of a small multigenic family, are glycosyl phosphatidylinositol-anchored proteins containing characteristic amino acid repeats at C-terminal (either EP or GPEET) (Ruepp *et al.*, 1999; Rangarajan *et al.*, 2000; Acosta-Serrano *et al.*, 2001; Butikofer *et al.*, 2002). It has been suggested that one function of the protease-resistant C-terminal is to protect the parasite surface from digestive enzymes in the insect gut. Compared to other fibrillin proteins, ChfuGV fibrillin along with SINPV fibrillin are the longest with 104 amino acids. A previous study has shown that SINPV fibrillin additional amino acid residues reside within a defined region. This region is well-known for its potential to take a coiled-coil configuration, thus SINPV fibrillin was predicted to form longer coiled-coil region (Faktor *et al.*, 1997). Analysis of the ChfuGV fibrillin protein sequence using diverse multiple alignment tools revealed that the extra amino acid residues of ChfuGV fibrillin protein was mostly concentrated inside the EP rich region (from 71 to 91) and at the terminal region of the protein, instead of being within the coiled-coil region as it was reported for SINPV fibrillin. The result of the phylogenetic analysis on fibrillin protein species showed that ChfuGV fibrillin shares a common ancestor with EpNPV, AgNPV PnNPV and OpMNPV fibrillin proteins (data not shown).

An ORF coding for a putative baculoviral envelope protein (p74) was located at the extremity of the clone (Cf-xv-p). This protein has been already identified in a number of other baculoviruses such as *Adoxophyes orana* granulovirus (AoGV) (Wormleaton *et al.*, 2003), CIGV (Lange and Jehle, 2003), PhopGV (Accession no. NC-004062), CpGV (Luque *et al.*, 2001), *Plutella xylostella* granulovirus (PxGV) (Hashimoto *et al.*, 2000), *Xestia c-nigrum* granulovirus (XcGV) (Hayakawa *et al.*, 1999), AcMNPV (Ayres *et al.*, 1994), OpMNPV (Ahrens *et al.*, 1997), CfNPV (Hill *et al.*, 1993), LdNPV (Kuzio *et al.*, 1999), BmNPV (Gomi *et al.*, 1999), SiNPV (Pang *et al.*, 2001), SeNPV (Ijkel *et al.*, 1999), HaNPV (Zhang and Wu 2001), McNPV (Li *et al.*, 2002), and *Culex nigripalpus* baculovirus (CnBV)(Afonso *et al.*, 2001). Baculoviral p74 proteins are considered to be a highly conserved ODV envelope protein that may play a major role in the initiation of viral infection (Slack *et al.*, 2001). It has been suggested that the N terminal of AcMNPV p74 is located outside of the ODV envelope while the C terminal acts as a transmembrane anchor (Faulkner *et al.*, 1997).

The results of a study on ChfuGV p74 protein support the idea that the N terminal of p74 protein is presumably a flexible structure with an elevated number of turns and surface loops, as multiple alignment of ChfuGV p74 protein with those of other BVs revealed that several glycine and proline residues were preserved in nearly all baculoviral p74 proteins (Rashidan *et al.*, 2003). Another remarkable aspect of ChfuGV p74 protein is a cluster of 8 cysteine residues, six of which were shown to be conserved in all known p74 homologues (Rashidan *et al.*, 2003). The high level of preservation of cysteine, glycine and proline residues in all baculoviral p74 proteins could mean that p74 protein is an

essential component in the life cycle of all baculoviruses and that mutations on these preserved residues might ultimately lead to the exclusion of the mutated virus from the evolutionary pathway.

Traditionally, the sequence of occlusion body matrix proteins granulins and polyhedrin is used to determine phylogenetic relationship in baculaviridae family. The soundness of this approach was initially questioned by Herniou and his associates (2001). In our latest paper (Rashidan *et al.*, 2004) we employed combined data from different genes and gene arrangement we demonstrated that GVs can be grouped in 2 distinct groups: Group I (ChfGV, CpGV, PhopGV and AoGV) and Group II (XcGV, PxGV, LoGV and TnGV). The results here confirm these results and the recent report on ClGV complete genome sequencing allowed us to add them to the group I.

ChfGV conserved proteins are most closely related to those of CpGV, PhopGV and AoGV. ChfGV along with CpGV, PhopGV, AoGV and ClGV share a high degree of gene order preservation within the studied genome region (Figure 1). GVs and NPVs demonstrated different patterns in their gene order in p74 regions (data not shown).

Figure 1. Schematic illustration of the gene structure and orientation of p74-containing region of various GVs (i- ChfuGV, ii- ClGV, iii-CpGV, iv-PhopGV, v-AoGV, vi-PxGV, vii-XcGV). The arrows represent the relative length and the direction of transcription of the ORFs.

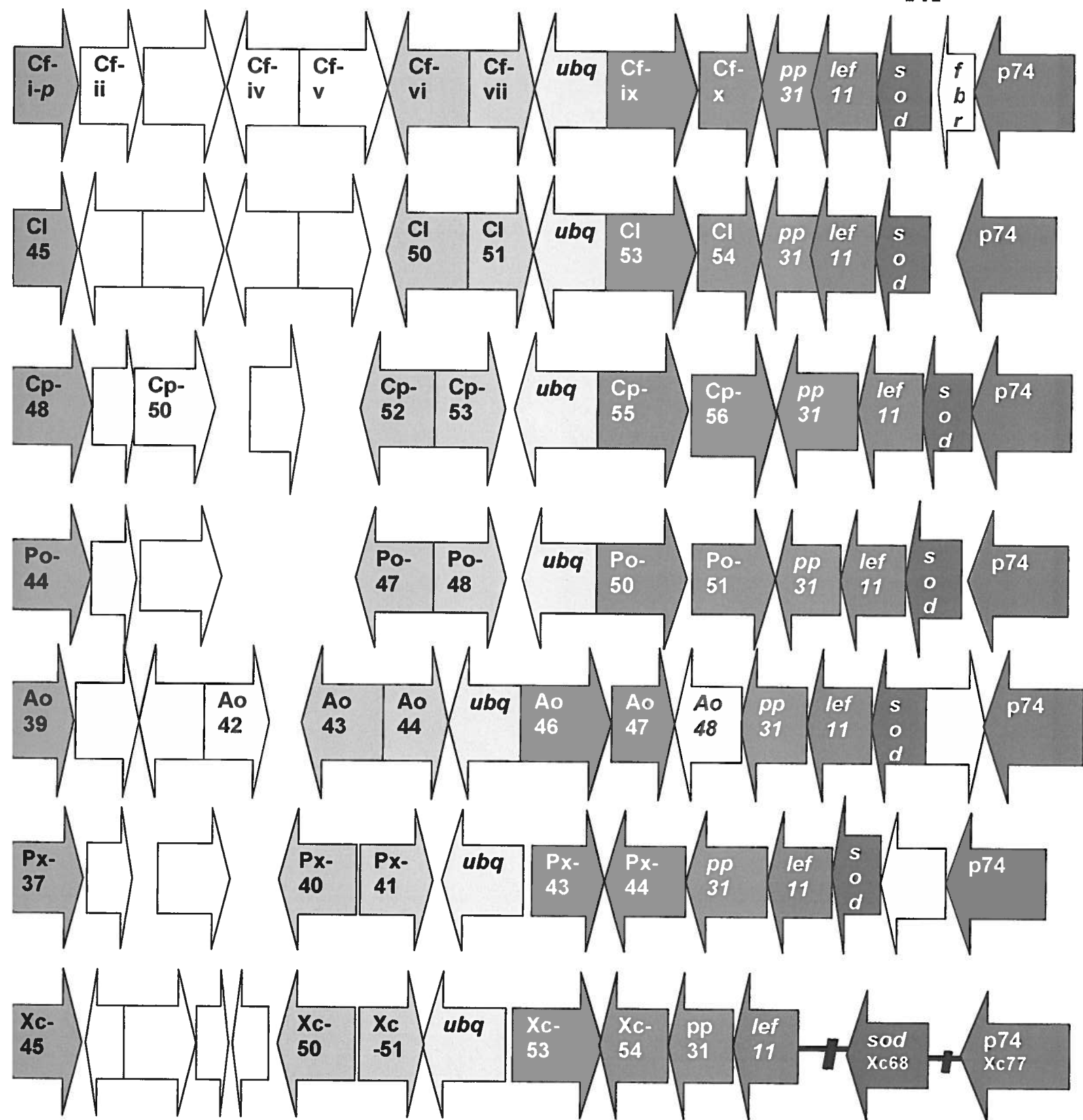
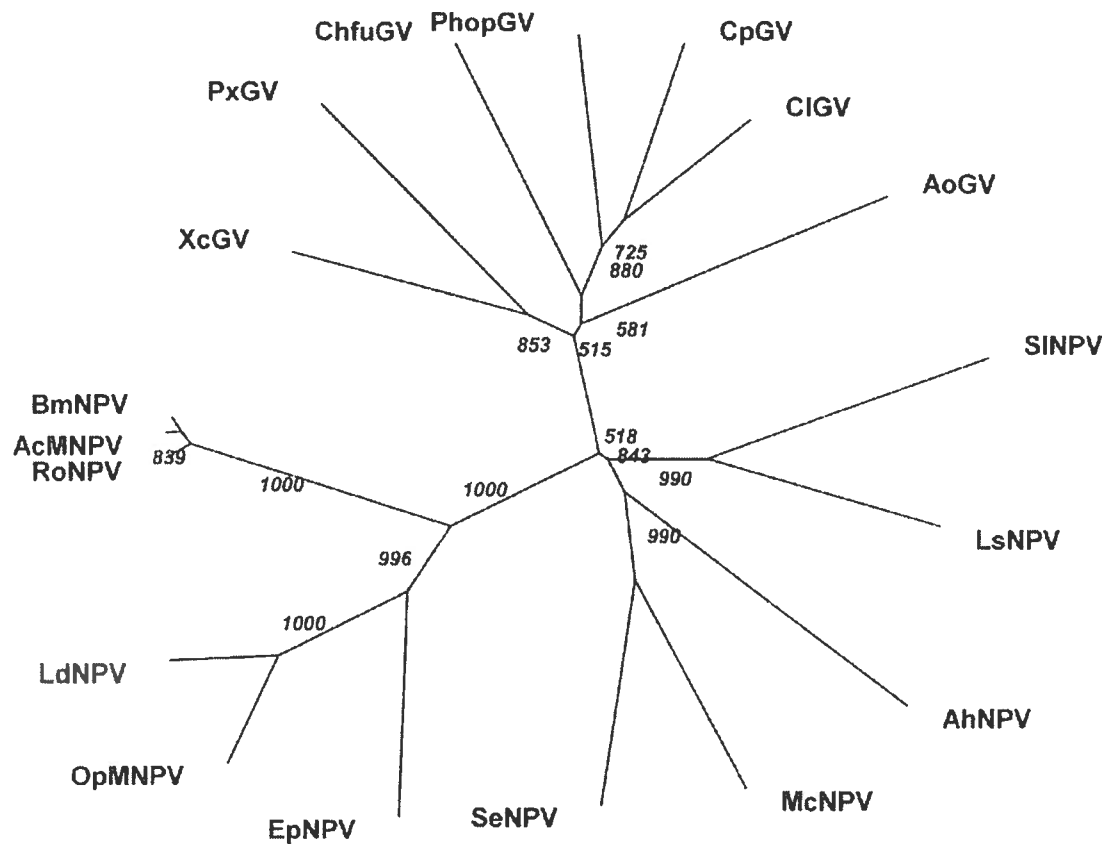


Figure 2. Phylogenetic analysis of the conceptual protein products of baculoviral late expression factor 11 (*lef-11*) genes. The unrooted tree was produced employing the Neighbor joining (NJ) method using PAUP4.0b4a. Bootstrap values (1000 replicates) are shown. Tree was constructed according to the following settings: 1- uninformative characters ignored. 2- Branch-and -bound search options: an initial upper bound: compute via stepwise; b keep minimal only; c collapse zero-length branches ON; MULPARS ON; e addition sequences: furthest. Accession numbers for the sequences used are as follow: ChfV AAN77198; CpV AAK70718; ClV AAQ21651; PhopV AAM70251; XcV AAF05170; AoV AAP85687; PxV AAG27344; McNPV AAM95134; SeNPV AAF33648; AhNPV AP006270; SiNPV AAL01720; RoNPV AAN28128; LsNPV AB009614; AcMNPV AAA66667; BmNPV AAC63712; LdNPV AAC70230; OpMNPV BAA02637; EpNPV AAK85585.



0.1

Table 1. Properties of ChfuGV ORFs detected in 8.9 kb *Bam*HI restriction fragment and their related ORFs on CpGV, PhopGV, AoGV ClGV, PxGV and XcGV genomes.

ChfuGV			CpGV			PhopGV			AoGV			CIGV		
ORF	length (nt)	D i r	ORF	length (nt)	D i r	ORF	length (nt)	D i r	ORF	length (nt)	D i r	ORF	length (nt)	D i r
Cf-i -p	-----	>	Cp48	1119	>	Po44	1101	>	Ao39	1128	>	CI45	1122	>
Cf-ii	303	>							Ao40	417	>			
			Cp49	390	<	Po45	204	<	Ao41	231	<	CI46	255	<
Cf-iii	1098	>	Cp50	2184	>	Po46	2433	>	Ao42	2628	>	CI47	2145	>
Cf-iv	735	<										CI48	477	<
Cf-v	621	>	Cp51	639	>							CI49	1746	>
Cf-vi	633	<	Cp52	1029	<	Po47	909	<	Ao43	609	<	CI50	621	<
Cf-vii	150	>	Cp53	147	>	Po48	144	>	Ao44	147	>	CI51	156	>
Cf-viii ubiquitin	288	<	Cp54 ubiquitin	285	<	Po49 ubiquitin	285	<	Ao45 ubiquitin	285	<	CI52 ubiquitin	231	<
Cf-ix	1032	>	Cp55	981	>	Po50	1020	>	Ao46	1050	>	CI53	1047	>
Cf-x	141	>	Cp56	210	>	Po51	147	>	Ao47	228	>	CI54	150	>
									Ao48	360	<			
Cf-xi <i>pp31</i>	756	<	Cp57 <i>pp31</i>	726	<	Po52 <i>pp31</i>	696	<	Ao49 <i>pp31</i>	837	<	CI55 <i>pp31</i>	678	<
Cf-xii <i>lef-11</i>	279	<	Cp58 <i>lef-11</i>	405	<	Po53 <i>lef-11</i>	267	<	Ao50 <i>lef-11</i>	285	<	CI56 <i>lef-11</i>	285	<
Cf-xiii <i>sod</i>	486	<	Cp59	399	<	Po54 <i>sod</i>	501	<	Ao51 <i>sod</i>	459	<	CI57 <i>sod</i>	426	<
Cf-xiv fibrillin	315	<							Ao52	171	>			
Cf-xv P74-p	-----	<	Cp60 p74	2067	<	Po55 p74	1977	<	Ao53 p74	1929	<	CI58 p74	2013	<
Most homologous ORFs			Cf-i, Cp48, Po44, Ao39, CI45, Px37, Xc45		Cf-vi, Cp52, Po47, Ao43, CI50, Px40, Xc49		Cf-vii, Cp53, Po48, Ao44, CI51, Px41, Xc51		Cf-viii, Cp54, Po49, Ao45, CI52, Px42, Xc52		Cf-xi, Cp57, Po52, Ao49, CI55, Px45, Xc55			

PxGV	XcGV
-------------	-------------

ORF	length (nt)	D i r	ORF	length (nt)	D i r
Px37	1107	>	Xc45	1167	>
			Xc46	246	<
Px38	462	>	Xc47	663	>
Px39	1131	>	Xc48	2520	>
Px40	621	<	Xc49	546	<
			Xc50	819	<
Px41	162	>	Xc51	162	>
Px42 ubiquitin	345	<	Xc52 ubiquitin	234	<
Px43	1245	>	Xc53	1062	>
Px44	393	<	Xc54	333	>
Px45 pp31	759	<	Xc55 pp31	888	<
Px46 lef-11	291	<	Xc56 lef-11	309	<
Px47 sod	462	<	Xc68 sod	462	<
Px48	156	<			
Px49 p74	1737	<	Xc77 P74	2133	<
Most homologous ORFs:		Cf-xii, Cp58, Po53, Ao50, Cl56, Px46, Xc56	Cf-xiii, Cp59, Po54, Ao51, Cl57, Px47, Xc68	Cf-xv, Cp60, Po55, Ao53, Cl58, Px49, Xc77	

5.5. Acknowledgements

The authors would like to thank Dr. Anne Bruneau, Department of Biology, University of Montreal for helping us in the phylogenetic studies. We also thank Dr. David Morse, Department of Biology, University of Montreal for his scientific discussions. This research was funded by the *Ministère des Ressources Naturelles du Québec* Grant number 0316-240S to CG.

CHAPTER 6 – PUBLICATION #5

***Choristoneura fumiferana* Granulovirus p74 Protein, a Highly Conserved Baculoviral Envelope Protein.**

Authors: Kianoush Khajeh Rashidan^{1*}, Nasha Nassoury², Samia Tazi¹, Paresa N. Giannopoulos¹, and Claude Guertin¹

1- Institut national de la recherche scientifique-Institut Armand-Frappier, 531 Blvd des Prairies, H7V 1B7, Laval-Canada

2- Université de Montréal, 4101 Sherbrooke east, H1X 2B2 Montreal-Canada

The nucleotide sequence data reported in this paper has been submitted to the GenBank nucleotide sequence database and has been assigned the accession number AAL13071.

Published in Journal of Biochemistry and Molecular Biology in 2003. Number 36, pages 475-487.

Contribution of the authors

I was the principal investigator in this report while Dr. Nassoury's expertise in screening to locate the clone was used to complete this report. Mrs. Giannopoulos and Mrs. Tazi helped me by performing some of the PCR reactions. Dr. Guertin was the director of the project and the majority of the experiences have been realized in his laboratory.

Identification et caractérisation de la protéine structurale p74 du granulovirus de *Choristoneura fumiferana*

Un gène codant pour une protéine homologue à la p74 baculovirale, une protéine structurale virale associée à l'enveloppe, a été identifié et séquencé sur le génome du *Choristoneura fumiferana* granulovirus (ChfuGV). Un segment du gène *p74* du ChfuGV a été localisé sur le fragment subgénomique *Bam*HI de 8,9 kb en utilisant différentes séries d'amorces dégénérées, lesquelles ont été dessinées en se basant sur les résultats du séquençage d'une protéine structurale majeure de 74 kDa associée au virus d'occlusion (occlusion-derived virus, ODV). Le gène possède un cadre de lecture ouvert (open reading frame, ORF) de 1992 nucléotides (nt) codant pour une protéine de 663 acides aminés dont le poids moléculaire prédit est de 74 812 Da. Des études comparatives ont révélé la présence de deux régions majeures conservées dans la protéine p74 du ChfuGV. Ces études montrent également que toutes les protéines p74 contiennent deux domaines transmembranaires putatifs à leurs segments C-terminaux. Au niveau nucléotidique, deux motifs de promoteur tardif (TAAG et GTAAG) ont été localisés en amont du premier ATG du gène *p74*. Le gène contenait un signal poly(A) canonical, AATAAA, dans la région 3' non-traduite (3'untranslated region, 3'UTR). Un arbre phylogénétique des p74 de baculovirus a été construit en utilisant une analyse de parcimonie maximale. L'estimation phylogénétique a démontré que la protéine p74 du ChfuGV semble d'avantage associée aux protéines p74 des granulovirus de *Cydia pomonella* (CpGV) et de *Phthorimaea operculella* (PhopGV).

6.1. Abstract

A gene encoding a homologue to baculoviral p74, an envelope-associated viral structural protein, has been identified and sequenced on the genome of *Choristoneura fumiferana* granulovirus (ChfuGV). A part of ChfuGV p74 gene was located on an 8.9kb *Bam*HI subgenomic fragment using different sets of degenerate primers, which were designed using the results of protein sequencing of a major 74 kDa structural protein associated with the occlusion-derived virus (ODV). The gene has a 1992 nucleotide (nt) open reading frame (ORF) encoding a protein with 663 amino acids with a predicted molecular mass of 74.8 kDa. Comparative studies have revealed the presence of two major conserved regions in ChfuGV p74 protein. These studies also show that all p74 proteins contain two putative transmembrane domains at their C-terminus segments. At the nucleotide sequence level two late promoter motifs (TAAG and GTAAG) were located upstream of the first ATG of the p74 gene. The gene contained a canonical poly (A) signal, AATAAA, at its 3' non-translated region. A phylogenetic tree for baculoviral p74 was constructed using maximum parsimony analysis. The phylogenetic estimation demonstrated that ChfuGV p74 is most closely related to those of *Cydia pomonella* granulovirus (CpGV), and *Phthorimaea operculella* granulovirus (PhopGV).

Key words: Baculovirus, ChfuGV, p74, envelope-associated protein, protein analysis, phylogenetic analysis.

6.2. Introduction

Granuloviruses (GVs) are members of the *Baculoviridae*, a family of large rod-shaped enveloped viruses with relatively large double-stranded DNA genomes. The *Choristoneura fumiferana* granulovirus (ChfuGV) virion is occluded in a highly organized proteinous matrix known as granulin. Infection cycle commences when *Choristoneura fumiferana* (spruce budworm) larvae ingest the occluded virions. The proteinous matrix of occlusion bodies dissociates under the alkaline pH condition in the midgut and releases the enveloped virions known as occlusion-derived virus (ODV).

Like other members of the family, ChfuGV produces two distinct forms of virion during a single infection cycle: occlusion-derived virus (ODV), and budded virus (BV). ODVs transmit infection solely from insect to insect, whereas the BVs spread infection from infected cells to healthy cells within the infected host (Keddie *et al.*, 1989). BVs acquire their envelopes at the plasma membrane of the infected cell by budding (Whitford *et al.*, 1989; Blissard and Rohrmann, 1989). In contrast, ODVs acquire their envelope within the nucleus of the infected cell (Braunagel and Summers, 1994). The attachment of ODVs to microvilli of insect midgut cells followed by the virus penetration are considered as vital steps for the initiation of infection cycle in the host. The protein(s) that could participate in these steps have not yet been identified but there is evidence suggesting that the ODV penetration is a non-endocytotic process and requires interaction of the virion envelope proteins with midgut cell surface proteins which eventually leads to membrane fusion (Horton and Burand, 1993).

The ChfúGV ODV like all other known baculoviruses ODVs, has a complex structure. Baculoviral ODV envelopes contain several proteins such as: VP17 (Funk and Consigli, 1993), ODV-E25 (Russell and Rohrmann, 1993), ODV-E35 (Braunagel *et al.*, 1996a), GP41 (Whitford and Faulkner, 1993), p74 (Kuzio *et al.*, 1989), ODV-E18 (Braunagel *et al.*, 1996a), ODV-E66 (Hong *et al.*, 1994), and ODVP-6E/ODV-E56 (Braunagel *et al.*, 1996b, Theilmann *et al.*, 1996; Rashidan *et al.*, 2002). One or more of these proteins could participate in adsorption, fusion and the penetration process. In *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) (prototype of the *Baculoviridae* family) p74 protein is exposed on the virion surface (Faulkner *et al.*, 1997) and has an essential role in the infectivity of the virus (Kuzio *et al.* 1989; Faulkner *et al.*, 1997). It has also been shown that although p74 null mutants fail to initiate infection in insect larvae when virus are fed orally, they are infectious once injected into the haemocoel of the larvae (Faulkner *et al.*, 1997). These observations suggest p74 protein is essential for initiation of infection and its presence is an indispensable element for attachment and /or fusion of the ODV envelope to the membrane of midgut epithelial cells. All this evidence places baculoviral p74 protein in a particular position for being considered as an ODV envelope protein that may play a major role in the initiation of viral infection. It has been suggested that the N terminal of AcMNPV p74 is located outside of the ODV envelope while the C terminal acts as a transmembrane anchor (Faulkner *et al.*, 1997). In AcMNPV the hydrophobic C-terminal region has been shown to play a significant role in protein's localization and its transmembrane anchoring character (Slack *et al.*, 2001).

In the present study we report the identification and characterization of a p74 homologue in ChfuGV. To date, this is the first report devoted to characterization of a granuloviral p74 protein, as all previous reports were concentrated on p74 protein homologues in nucleopolyhedrovirus (NPVs).

6.3. Materials and Methods

6.3.1. In vivo Production and Purification of Virus and DNA Extraction

Choristoneura fumiferana fourth-instar larvae were infected by ChfuGV using virus contaminated artificial diet (Forté *et al.*, 1999). The virus isolation from infected larvae followed by DNA extraction was carried out as previously described (Bah *et al.*, 1997).

6.3.2. SDS-PAGE and Protein microsequencing

Enveloped nucleocapsids of ChfuGV were purified described (Rashidan *et al.*, 2002). The purity and integrity of enveloped nucleocapsids were examined by transmission electron microscope. Purified enveloped nucleocapsids were disrupted in an equal volume of Laemmli sample buffer (Laemmli, 1970) in presence of β -mercaptoethanol (Bio Rad). The samples were then boiled for 5 min and clarified by centrifugation at 13,000 g for 5 min before electrophoresis on a 3% stacking/12.5% separating sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel (SDS-PAGE). Proteins were visualized using light Coomassie Brilliant Blue (Invitrogen) staining, and the molecular mass (*Mr*) of virion proteins was determined by comparison to those of standards (Bio Rad).

Gel purified ChfuGV structural proteins were used for protein sequencing. Major protein bands from ChfuGV enveloped nucleocapsids preparation, including a 74 kDa band, were excised from the gel and were subjected to protein sequencing analysis. Sequencing was performed at Harvard Microchemistry facility using either microcapillary reverse-

phase HPLC nano-electrospray tandem mass spectrometry (μ LC/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 degenerated primers (with the least redundancy) to be used in PCR reactions on ChfuGV genomic DNA in order to locate the gene coding for this protein on ChfuGV genome.

6.3.3. Cloning and DNA sequencing

Using different combinations of degenerate primers, PCR amplifications were performed on ChfuGV genomic DNA. A part of the p74 gene was amplified with the combination of the following pair of degenerate primers, p74FRIII: AC(A/G) CC(A/G) AA(T/C) GC(T/C) AA(T/C) AAA and p74-RVV: TTT (A/G)AT (T/C)AA (G/A)GC (C/T)GA. All PCR reactions were carried out using Pwo DNA polymerase (Roch) and *Taq* DNA polymerase (Amersham Pharmacia Biotech) in the buffer supplied by the manufacturer supplemented with 200 μ M of each dNTP and $MgCl_2$ to a final concentration of 2.5 mM. PCR reactions were subjected to heated for 10 min at 95° C prior to the start of the amplification (30 cycles of 95 °C, 30sec; 47 °C 25 sec; and 72 °C, 40 sec). The PCR amplicon was cloned in PCR2.1 (Invitrogen) cloning vector and sequenced, and the sequence was compared to homologues in GenBank/EMBL by using the BLAST algorithm (Altshul *et al.*, 1990). In order to locate p74 gene on ChfuGV genome, subgenomic fragments were generated by *Bam*HI endonuclease and fractionated in 1% agarose gel. Gel extracted fragments were then used as templates for PCR reactions using the same set of primers. A PCR reaction on an 8.9kb ChfuGV *Bam*HI fragment yielded an amplicon with the expected size. This 8.9kb restriction fragment was cloned in a

pBlueScript-SK⁺ cloning vector (Stratagene), and transformed into *E.coli* XL-1 Blue (Stratagen). All the manipulations including 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 a primer walking technique. Automatic sequencing was carried out using an Applied Biosystems automated DNA sequencer model 377XL (Applied Biosystems CA, USA). The first series of sequencing results revealed that the pbam8.9 clone only contained a segment of p74, and in order to attain the entire sequence of the p74 ORF, a primer was designed using the preliminary sequencing results. This primer was then employed in direct sequencing of ChfuGV genomic DNA. Primer walking on genomic DNA continued until the complete sequence of p74 ORF was obtained.

6.3.4. Computer Analysis

Sequence data assembly and analysis was performed with Sequencher software version 4.0.5 (Gene Codes Corp.) and 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 by using BLAST (Altshul *et al.*, 1990). CLUSTALX (1.81) (Thompson *et al.*, 1997) was used for multiple amino acid sequence alignments. BOXSHADE version 3.21 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. Amino acid composition was calculated using

Peptide Statistic (<http://web.umassmed.edu/cgi-bin/biobin/pepstats>), Pepinfo (www.ebi.ac.uk/service/pepinfo/1598281014310854.html) and ProtParam (<http://www.expasy.ch/tools/protparam.html>) Tools. Hydrophobic plots were obtained using Kyte and Doolittle hydrophobicity values (1982) via EMBOSS Pepwindow program (<http://www.ebi.ac.uk/emboss/pepinfo/>). Secondary structure prediction method used in this study was GORIV (Garnier *et al.*, 1996). Potential transmembrane regions were identified using programs TMHMMver. 2.0 (Moller *et al.*, 2001), Tmpred (Hofmann and Stoffel, 1993), and SOUSI program (<http://sosui.proteome.bio.tuat.ac.jp/sosui/menu0.html>). Potential glycosylation, and myristoylation sites were predicted using the programs NetOglyc (Hansen *et al.*, 1998) and Proscan (Bairoch *et al.*, 1997). The prediction of coiled-coil region was performed using COILS program version 2.1 (Lupas *et al.* 1991). Accession numbers for p74 protein sequences used in this study are as follow: *Cydia pomonella* granulovirus (CpGV), AAK70720; *Phthorimaea operculella* granulovirus (PhopGV), AAM70253; *Plutella xylostella* granulovirus (PxGV), AAG27347; *Xestia c-nigrum* granulovirus (XcGV), AAF05191; *Helicoverpa armigera* nucleopolyhedrovirus (HaNPV), AAK96273; *Spodoptera exigua* nucleopolyhedrovirus (SeNPV), AAF33660; *Orgyia pseudotsugata* multicapsid nucleopolyhedrovirus (OpMNPV), AAC59133; *Choristoneura fumiferana* multicapsid nucleopolyhedrovirus (CfMNPV), S29849; *Epiphyas postvittana* nucleopolyhedrovirus (EpNPV), AAK85685; *Bombyx mori* nucleopolyhedrovirus (BmNPV), AAC63805; *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV), AAA66768; *Lymantria dispar* multicapsid nucleopolyhedrovirus (LdMNPV), AAC70212; *Spodoptera litura* nucleopolyhedrovirus (SiNPV), AAL01707; *Rachiplusia ou*

nucleopolyhedrovirus (RoNPV), AAN28023; *Mamestra configurata* nucleopolyhedrovirus (McNPV), AAM09268; *Culex nigripalpus* baculovirus (CnBV), AAK94152; and *Heliothis zea* virus1 (Hz-1 virus) AAM45758.

6.4. Results and Discussion

In this report we have identified and characterized of an ORF from ChfuGV encoding a 74 kDa protein homologue to a highly conserved envelope protein termed as baculoviral p74. This protein has been already identified in a number of other baculoviruses such as CpGV (Luque *et al.*, 2001), PhopGV, PxGV (Hashimoto *et al.*, 2000), XcGV (Hayakawa *et al.*, 1999), AcMNPV (Ayres *et al.*, 1994), OpMNPV (Ahrens *et al.*, 1997), CfMNPV (Hill *et al.*, 1993), LdMNPV (Kuzio *et al.*, 1999), BmNPV (Gomi *et al.*, 1999), SINPV (Pang *et al.*, 2001), SeNPV (Ijkel *et al.*, 1999), HaNPV (Zhang and Wu 2001), EpNPV, RoNPV, McNPV (Li *et al.*, 2002), CnBV (Afonso *et al.*, 2001), and Hz-1 virus (Cheng *et al.*, 2002). In few other BVs only partial sequences of p74 protein are known, such as *Buzura suppressaria* single-nucleocapsid nucleopolyhedrovirus (BsSNPV) (Hu *et al.*, 1998), *Trichophusia ni* single capsid nucleopolyhedrovirus (TnSNPV) (Fielding *et al.*, 2002), and *Leucania separata* nucleopolyhedrovirus (LsNPV).

The results of protein sequencing of a major 74 kDa protein in ChfuGV enveloped nucleocapsid yielded several stretches of amino acids. These amino acid sequences were found to be very similar to a highly conserved ODV-specific envelope protein known as p74 after being compared to homologous proteins in GenBank/EMBL using the BLAST algorithm (Altshul *et al.*, 1990). The ORF coding for ChfuGV p74 protein was located on the genome of ChfuGV as described in materials and methods section. The strategy adopted for sequencing is shown in figure1. The sequence and the position of the primers used to complete the sequence of p74 are presented in figure 2. The ORF is 1992 nt potentially encoding 663 amino acids, with an estimated molecular mass of 74.8 kDa.

The 5' non-translated region (NTR) of ChfuGV p74 gene was analyzed in order to identify all possible *cis*-acting elements and possible transcription start sites. No obvious TATA box element was detected however two late promoter motifs (TAAG and GTAAG) were located respectively at 5 nt and 161 nt upstream of the first ATG of the p74 coding sequence. The TAAG motif is known to be a strong promoter of the late genes in baculoviruses. The gene contained a canonical poly (A) signal, AATAAA, at 3' NTR at the position +16 nt downstream from termination signal. The nucleotide sequence surrounding the predicted translational start codon (AGTATGG) obeyed Kozak's rule for efficient eukaryotic translation initiation with the presence of two purine bases at +4 and -3 positions (Kozak, 1986) (Fig.2).

The deduced amino acid composition in terms of non-polar (ACFGILMPVWY), polar (DEHKNQRST), basic (HRK), and acidic residues (DE) indicates a content of non-polar (54.75%), polar (45.25%), acidic (11.01%) and basic (9.50%). The protein is particularly rich in Leucine (9.20%), Serine (8.60%), and Alanine (7.54%) residues. The amino acid composition of ChfuGV p74 protein also revealed the presence of a high number of glycine and proline residues (26 and 29, respectively), which total 8.3% of ChfuGV p74 in. In all other p74 homologues studied here the matching high number of glycine and proline residues were also observed. The number of glycine and proline residues respectively are as follows: in GVs 26 to 35 and 27 to 33; in NPVs 22 to 34, and 25 to 32; in CnBV 45 and 39; and finally in Hz-1 virus 26 and 33. The N-terminal two-thirds of the ChfuGV p74 homologue contains most of these glycine and proline residues (14 and 18 residues respectively). Proline with its side-chain bound to the amide nitrogen of the

main chain, and glycine due to the absence of a side chain always result turns in the conformation of a protein. Both proline and glycine often are also associated with surface loops. The unhindered rotational freedom of glycine can result in an increased loop flexibility that leads to a smooth progress of conformational changes in the protein. This observation suggests that the N terminal of ChfuGV p74 protein can be characterized as a flexible structure with a high number of turns and surface loops. Multiple alignment of ChfuGV p74 protein with those of other BVs and Hz-1 virus revealed that several glycine and proline residues were conserved in almost all baculoviral p74 proteins (Fig.3a). This high conservation of proline and glycine residues may imply their significance in the creation of a functional conformation in baculoviral p74.

Another remarkable aspect of ChfuGV p74 protein was a clustered 8 cysteine residues. Multiple sequence alignment of ChfuGV p74 with those of other baculoviruses revealed that six of these cysteine residues are astonishingly well conserved in each and every one of p74 homologues studied here (Fig.3a and b). The seventh and the eighth cysteine residues, initially detected in ChfuGV p74 protein, were also detected in CpGV, PxGV, SINPV, HaNPV and Hz-1 virus. This level of preservation of cysteine residues, which are perhaps involved in creation of disulfide bridges, implies that presence and the arrangement of these residues are matter of a great importance for the function of the p74 protein in all baculoviruses as well as Hz-1 virus and without them the protein may not be able to maintain the proper folding for a functional protein. The high degree of conservation as well as the position preservation of cysteine, glycine and proline residues in all baculoviral and Hz-1 p74 proteins may imply this important fact that p74 protein

plays a crucial role in the life cycle of all baculoviruses and the removal of these conserved properties may eventually lead to the elimination of the virus in their evolutionary pathway.

The size of baculoviral p74 proteins varies between 578 aa in PxGV to 710 aa in XcGV. The deduced amino acid sequence of ChfuGV P74 was compared to other known baculoviral p74 proteins. These results revealed a considerable similarity between ChfuGV p74 sequence and other p74 homologues (Table 1). ChfuGV p74 has 63.7% amino acid sequence identity to its closest relatives being CpGV and PhopGV. Hz-1 virus p74 was the most distant, showing only 20.9% amino acid sequence identity.

The one-third C-terminal of ChfuGV p74 protein in ChfuGV is rich in non-polar residues (60.63%) and the predicted hydrophobic profile of ChfuGV p74 demonstrated the existence of two highly hydrophobic regions within this segment of ChfuGV p74 protein. Similar hydrophobic regions were also identified in the C-terminal regions of all other studied p74 homologue proteins (Fig.4a). Various computer assisted prediction models used in this study predict that ChfuGV has two membrane-spanning regions within its C-terminal region. These transmembrane domains were located between amino acid residues 600 to 620 and 633 to 656. Despite low sequence conservation at the C-terminal, almost all baculoviral p74 homologues (excluding PxGV) have conserved these two hydrophobic regions at approximately parallel positions (Fig.4b). The lack of transmembrane domains in PxGV, as Slack and his associates have also suggested, might be due to a recent evolutionary change in PxGV p74 (Slack *et al.*, 2001).

Beside these two regions an additional transmembrane region was positioned almost in the center of p74 proteins, in all NPVs, CnBV, Hz-1 virus as well as XcGV (Fig.4b). All computer assisted models used in this study predicted that ChfuGV p74 N-terminal is exposed outside of ODV envelope, while a part of the C-terminal is anchored within the envelope membrane through two highly hydrophobic regions. This arrangement places the last seven amino acid residues of ChfuGV p74 out side the envelope membrane (Fig.4c). Similar arrangement is shared by two other studied GVs (CpGV and PhopGV). In XcGV, the entire NPVs group, CnBV and Hz-1, the N-terminal of p74 proteins are exposed out side as well, although the presence of three transmembrane regions directs the C terminal end towards inside of the ODV envelope (Fig.4b).

ChfuGV p74 has two potential *N*- glycosylation N-{P}-[ST]-{P} sites located at residues 202, and 332. Glycosylation is important for proper folding or targeting of some polypeptides (Darvey, 1989). Two *N*-myristoylation (G-{EDRKHPFYW}-X-X-[STAGCN]-{P})) sites were also observed at residues 314 and 571 respectively.

A close look at the multiple alignments of p74 amino acid sequences revealed the existence of two major conserved regions (Fig.3a). The first region began from amino acid residue number 22 and is extended up to residue 302 (in ChfuGV p74 protein), and the second conserved region was located between residues 357 and 593. All conserved cysteine residues were found within the first region while the conserved proline and glycine residues reside within either one of the two conserved regions. Areas situated

outside of these two preserved boxes demonstrate high levels of sequence divergence. Several conserved motifs were detected within these two major conserved boxes in all baculoviruses. These motifs were ³²[F/L/M]P[H/N][I/V][L/F][I/V][D/N]Y³⁹, ⁴⁸[D/E][Y/F][Y/F][V/I]PP⁵³, ⁷¹GC[E/Q/D/V][A/S]M[S/T]C[F/Y]P[Y/F][T/H/R/N][E/A]TG[V/P/T][I/V]⁸⁶, ¹⁰⁷QPAC[F/Y][N/H]LD¹¹⁴, ¹⁵⁰N[A/S/T]PY[I/M/L/V]RT¹⁵⁶, ¹⁶³G[V/I]DDVP¹⁶⁸, ¹⁹¹AYCRRFGR¹⁹⁸, ²¹⁶FVL[G/E][E/D][T/S][I/V/L]²²², ⁴⁷³[I/L/F]WDP[F/L]G[Y/F][N/S]N[N/Q]MFP[R/K]⁴⁸⁵. ChfuGV, CpGV, and PhopGV p74 proteins shared several highly conserved stretches of amino acids. These highly conserved stretches were located at the following positions: residue 1 to 9, 91 to 132, 138 to 184, 205 to 228, 372 to 388, 391 to 407, 413 to 426, 430 to 513, 522 to 530, and finally 534 to 553. The longest was a stretch of amino acids covering 83 amino acid residues started from residue 430 in ChfuGV p74 protein. Interestingly, all abovementioned highly conserved regions were located on the external segment of the protein suggesting the importance of conformational preservation of this segment. This level of preservation can be usually seen in the envelope proteins that play a role in the process of virion attachment to the cell surface receptors.

Phylogenetic analysis of p74 proteins is shown in Fig.5. Tree was produced via maximum parsimony to estimate the evolutionary relationship between p74 proteins species in baculoviruses and Hz-1 virus. These analyses showed a clear division between the GV and NPV p74 proteins. ChfuGV is positioned along side with CpGV and PhopGV in the same cluster, supported by 98 bootstrap values.

The data presented in this paper offer a broad comparative study on an essential baculoviral envelope protein p74 that may play a significant role during the initiation phase of baculoviral infection and put more weight on the importance of this protein for the *Baculoviridae* family and related viruses.

Figure 1. Schematic presentation of the 8.9kb BamHI fragment from ChfuGV genomic DNA showing the position and direction of the putative ORFs p74, p10, and *sod*. Only 573nt of p74 *orf* (denoted as p74p, “p” for partial) is on this fragment. The strategy employed for sequencing of complete p74 *orf* is presented, arrows show the position of the primers used for sequencing and the direction of sequencing.

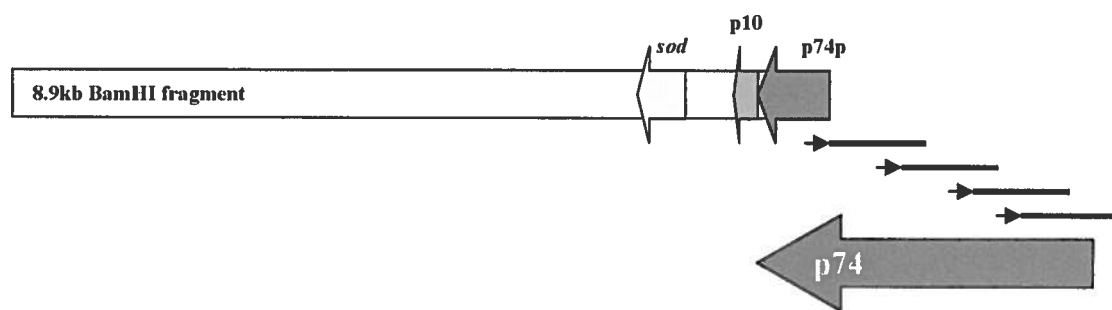


Figure 2. Figure shows nucleotide and protein sequence of ChfuGV p74. Two late promoter motifs and a putative polyadenylation signal for p74 gene are in bold. The positions of degenerated primers are in underlined bold, and the locations of primers used for sequencing are underlined. Asterisk (*) shows the stop codon.

1 - CTTGAATCGTAATCACAGAAGTCGTAGAAGGGCACCCAGAAATCCATCCCGAAGCCCATC - 60
 61 - CCGAAGCCCATCACCTCAAAGGCAGAGGTCCAGATCTCGTTCACGTTCAAGATCCATTGA - 120
 121 - TTATTGCCAGTAAGGAGATCTCGTTCACGTTCAAGATCCATTGATTATTCGCCAGTAAG - 180
 181 - GAGGTCGAGATCTAGGTCCCCAATAAATCGTAGGTCCAGATCAAGATCAATTCATCTGA - 240
 241 - CGACACTTTAATTTTAGATTAATAAATAGACGCAAATAATACAATTGTATATTTTATTTTC - 300
 301 - TAATATACTTCTCCTACAATATCGCTTACATTAAGTATGGCTACTGTAACACAATTAGAC - 360
 - M A T V T Q L D
 361 - ATATTAAATGCAGTGCAATATTTATCTAATCGTGAATCGCTATCTTTTATTTCAGATGG - 420
 - I L N A V Q Y L S N R E S L S F I S R W
 421 - CGCAGTGTTTTTCCTCACATATTGATAGATTATAGTATTAGGTGGGCTACTAACGATGAT - 480
 - R S V F P H I L I D Y S I R W A T N D D
 481 - TATTACGCTCCACCAATGCGACAACTAGCGCTATCGTGGTGGAAATAATATTTTCG - 540
 - Y Y V P P P M R Q T S A I V V E I I F S
 541 - AAAGAAGGTTGCGAAGCTATGTCCTGCTTTCCTTACACTGAAACAGGCGTAATCGATTTT - 600
 - K E G C E A M S C F P Y T E T G V I D F
 601 - ATGAAAAGTCTATCGGTGGTTACACTCAGACCTCGAACACTGCAGTGCAATATAATCAG - 660
 - M K S P I G G Y T Q T S N T A V Q Y N Q
 661 - CCGGCATGTTTTAATTTGGACTCTGCGTTGGCTGCACGCGATGGTAAAATACAATCAGTA - 720
 - P A C F N L D S A L A R D G K I Q S V
 721 - GAGCTGCGATATACATCGTCTAATAAATGTGAATGGTGGATAGTTTTACAAAAGCTTGG - 780
 - E L R Y T S S N K C V M V D S F T K A W
 781 - TTAAACGCGCCTTATATTGCGACAGACGCACACGTTGTGCGTGGTGTGACGATGTGCCA - 840
 - L N A P Y I R T D A H V V R G V D D V P
 841 - GGATTTGATGTGTCGTACGATGATGATCTGCTTTTCCTGAACGAATTAAGGCAAATTT - 900
 - G F D V S Y D D D P A F P E R I K G K F
 901 - AACACGGCCTATTGTGCGCGCTTTGGTCTGTTCTGAAATAAATAACAGTTGCTCTCAACCG - 960
 - N T A Y C R R F G R S E I N N S C S Q P
 961 - TGGTATGAAACATTTGTGTCGTTTGTATTAGGCGAATCTATATTGACCACCTTTAAATTG - 1020
 - W Y E T F V S F V L G E S I L T T F K L
 1021 - GCATCGACAAACGTTTTTGATGATTTACGTGATTTTGATTACAGTCGTCCATCTAACATT - 1080
 - A S T N V F D D L R D F D Y S R P S N I
 1081 - CTGCCTCTGCACCACCTCCTGAAGGAGACGAAATGTTGTTTAAATGGTTAAACACCAGA - 1140
 - L P P A P P P E G D E M L F K W L N T R
 1141 - GACACTACGGTGGATGTTGGAAGAGAAAACAATTTCTTGAACAATAAATTTGATATGGTA - 1200
 - D T T V D V G R E N N F L N N K F D M V
 1201 - TTGGGACAGAGCATAATTATGTGCGAAACGAAGGCTTTAGCACTGTCGCAACTCAAAAT - 1260
 - L G Q S I I Y V A N E G F S T V A T Q N
 1261 - ACAAATATGCACAGTGGGTTGATGGAAAATTTGCTGATGCGTCGTCGTAAGTGTTACAA - 1320
 - T N M H S G L M E N L L M R R R E V L Q
 1321 - AATAACATTAATTATAGTAACAAAAAGAACCTAATCAATTTAGTGGTTACGAAAATAGT - 1380
 - N N I N Y S N K K N P N Q F S G Y E N S
 1381 - TTCAATGTGAACCCCTTCAACAGACGAGTTAGAAATTATAATAATACAATTTTAGAAGAC - 1440
 - F N V N P S T D E L E I I I I Q F L E D
 1441 - CATGCTTTTATAATGAGTATATTAACGGACTTGGGTTTTAGTGTGTTAGAATCATCTTTG - 1500
 - H A F I M S I L T D L G F S V L E S S L
 1501 - AGTAGTATGTTGCAACAATTAACAAGGTGTTAATACCTTCGTTAAAAGGATGCTTAGT - 1560
 - S S M L Q Q L N K V L I P S L K R M L S
 1561 - TTGCAAAGTCGCCGAGTGACCGCTGCGCTATTAGGCGAAACATACAAAGCGCGCATGATA - 1620
 - L Q S R R V T A A L L G E T Y K A A M I
 1621 - AACGCGTTGAATCGTGCTTTTATCAGCACAATTTCTACTGTTGCCAAAGCTACCGCGCGA - 1680
 - N A L N R A F I S T I S T V A K A T A R
 1681 - ACTGTGAGAGCTGCCGCTTCAATAGCCAATTTGGCATTAAACATTTCTAACTATCGCCGAC - 1740
 - T V R A A A S I A N L A L T F L T I A D
 1741 - TTGGTGTGATGATATGGGATCCGTTCCGGTACAGCAATATGTTTCCGCGAGGTTATTTG - 1800
 - L V L M I W D P F G Y S N M F P R G Y L
 1801 - GACGATTTATCTTCTGCTTTTTTGTGCGCCTATTACGAATCTATTGATGCTCCTACACGG - 1860
 - D D L S S A F L S A Y Y E S I D A P T R
 1861 - GACATAATCGAGTTTAAACCTCAACATTTCTCAAACCTTGATCATTGACGAGGAAGAAGAG - 1920
 - D I I E F K P Q H F T S N L I I D E E E E
 1921 - TATTTTGTGTAAGGTATGTTACATTTAGCAGACTATCTAGCTGTCTTAGATGTAAACAGC - 1980

- Y F V E G M L H L A D Y L A V L D V N S
 1981 - AATGGACAAGTTATCGATTTACTGCGAGGTGTAGAAAGTGTACGAGGTTAACGACGAGGAA - 2040
 - N G Q V I D L L R G V E V Y E V N D E E
 2041 - ATAATAGGTGCTAGTTTGGCGTCTACCACATGGGCCTATTTTAAATGGTTTTGTGCTCGT - 2100
 - I I G A S L A S T T W A Y F K W F C A R
 2101 - CATGATGCATTAATAAAGACACCAAATGCTAATAAAATTTTAGTTGTGCCGTCAATCGTA - 2160
 - H D A L I K T P N A N K I L V V P S I V
 2161 - TTGTGTGTGGCCGGTCTAATCTATTCTTTAAAAATACCACAGCGTTTTACAAATAGAGCAA - 2220
 - L C V A G L I Y S L K Y H S V L Q I E Q
 2221 - CAAACAAATATTCATCTCTGCTTGTGCTGATCATTTTATTGTCCTTTTTATTATTATTC - 2280
 - Q T N I H L C L L L I I L L S F L L L F
 2281 - ACACCGTCGGTGCAATATTATTCAGCTTTGATCAAAACACAGATATGATTAAAAAATATAT - 2340
 - T P S V Q Y Y S A L I K H R Y D *
 2341 - ~~TATAAGCA~~TATAAATTAAAAAATGTCTAAACCGAGTATTTTAAACGCAAATCTTGACGCTG - 2400
 2401 - TGAATGTTGTTGACCAAAAAGTTGATGTTATAGACCAAAAAGTTGACAACCTGCAAACAC - 2460
 2461 - AATTGAACAATTTTCAATTGGATGACATTACTGCTCTTGTCAACGATGTCAACGAAAAAG - 2520
 2521 - TTACAAACATTCAAGACATGCTAACCGGCTTGAACCCGAACCCAGAACCCGAACCACTAAGGCGTT - 2580
 2581 - CCGAACCAGAACCCGAACCCAGAACCCGAACCCGAACCCGAACCACTAAGGCGTT - 2640
 2641 - CTGTTGTTAAATCTCACCATAAAAAAATAAGTGAATAGCTGTAATCAATTCACGCTGTT - 2700
 2701 - TCAAAAGGAAATCGATGTTATACGCCACATATGAATGCGAACAAAACTCAATTTTTTTT - 2760
 2761 - AATATTAATAATTTTACAAATTAAATCGTAGTCATTTTGAATATATTATTGTCAAATTA - 2820
 2821 - AAATGTACAAATGTTTGTATAAATTTGTACATTTTATTCACAAATAATAAGTGTGATAT - 2880

Figure 3. a) Amino acid alignments of deduced p74 protein of ChfuGV along with other p74 protein homologues studied here (see materials and methods for accession numbers). Conserved proline, glycine are depicted by (*), and conserved cysteine residues are illustrated by (+). Two highly conserved regions (C.R.) are shown within two boxes as C.R.I and C.R.II. b) Figure shows the position of conserved cysteine residues and their spacing patterns in baculoviral p74 protein homologues,

			*	*	*	*	*	*	*	+	*
ChfuGV	150	NAPYPTLAHVVRGVDDVP	GFDVSYDDPAFPERI	RCKENTAYCRRFGPSEINN	-----						
CpGV	150	NAPYPTSRHVVRGVDDVP	GFDVANDDPAFPERI	RCRFNEAYCRRFGPSENNN	-----						
PhopGV	148	NAPYPTSRHIVKGVDDVP	GFDVSYDDPHMFERIV	CKFNAYCRRFGRSERDN	-----						
PxGV	147	NSPYLRTENHILRGVDDVP	GFDVYDDDPAFPERI	RCKINSAYCRRFGPAGET	-----						
XcGV	149	NPYPTENPKSTGIDVDP	GFTRLSDNPVFPD	VDCIFNSAYCRRFGP	VVNANSSSA						
AcMNPV	151	NSPYLPTTEHTIMGVDDVP	AFNVRPDPDLFPERF	KCFNEAYCRRFGPELF	NG						
OpMNPV	151	NSPYLPTTEHTIMGVDDVP	AFNVRPDPDLFPERF	KCFNEAYCRRFGP	LI	NG					
CfMNPV	151	NSPYLPTTEHTIMGVDDVP	AFNVRPDPDLFPERF	KCFNEAYCRRFGP	LI	NG					
LdMNPV	150	NPYLPTTEHTVQGVDDVP	AFNVTPETG	AFPEMFCEFNEAYCRRFGPSLI	NG						
BmNPV	151	NSPYLPTTEHTIMGVDDVP	AFNVRPDPDLFPERF	KCFNEAYCRRFGPNLF	NG						
SlNPV	154	NSPYLPTTEHTIKGIDVDP	GFNVEYGS	DPHFPEIKCTFNEAYCRRFGPEM	NG						
SeNPV	150	NSPYPTTEHVIKIDVDP	AFNVEEAPDLLFPERF	VCKFNEAYCRRFGPTLV	NG						
HaNPV	176	NPYLPTDDHLIQGVDDVP	GFNVTNDTDQLFPERF	ECFFNEAYCRRFGPSLQPN	NG						
EpNPV	151	NSPYLPTTEHTIMGVDDVP	AFNVDPDPDLFPERF	KCFNEAYCRRFGP	LV	NG					
RoMNPV	151	NSPYLPTTEHTIMGVDDVP	AFNVRPDPDLFPERF	KCFNEAYCRRFGP	LF	NG					
McNPV	155	NSPYLPTTEHVIKIDVDP	AFNVAASNPLLP	PEQLGTENDAYCRRFGPSLQ	NG						
CnBV	151	ISPYLPTTEHILQCHDDVP	GFNVEPTLDDIFFETED	KFNRCYCDYEMPOLDPSC	-----						
Hz-1	169	EIPRYPESELYAPRLNDLDTG	FYEBSTDTYHFN	Y	-----						
			+	*	*	*	*	*	*	*	*
ChfuGV	204	-SCSQPWYETFSFVLGESILTFK	ASTNVFDOLPDEFDYS	PPSNLP	PPAPPPG-DEML						
CpGV	204	-ACQPWYEIFVSFVLGESIMTFK	ATTHVFDLPNFDTKPSA	LPDAPEP	EG-VAMV						
PhopGV	202	-GCSQPWYEIFVSFVLGESILTFK	SECHVFAELPDEFDYP	APSA	LPPEPSPG-LDML						
PxGV	201	-SCIRPWYETFTFVLGESEVTF	TKLITTVISLDEFDYKPSA	LPDAPPP	N-N--L						
XcGV	208	SGCQAQWYEMLVGFVLGDSVYAS	FKMLATGVVADMA	ASHNYDRPSS	LPQPPPPAG-RTML						
AcMNPV	205	-GCSFRWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSS	LPPEPBNVDS-NAIL						
OpMNPV	205	-GCSFRWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	AAPSP	LPPEPPADS-NAVL						
CfMNPV	205	-GCSFRWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	TAPSP	LPPEPPADS-NAVL						
LdMNPV	203	-GCSMQWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSP	LPPEPPADS-NAVL						
BmNPV	205	-GCSFRWWESLIGFVLGDTIL	TFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
SlNPV	208	-TCLIRWWESMIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
SeNPV	204	-GCSVNWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
HaNPV	231	-GCSLQWWESLIGFVLGDTIL	VSFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
EpNPV	205	-GCSFRWWESMIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
RoMNPV	205	-GCSFRWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
McNPV	209	-GCHLQWWESLIGFVLGDTIY	TFKMLANNIFSELDPDEFDY	APSS	LPPEPPADS-NAVL						
CnBV	206	-GCLQWWESLIGFVLGDSIYAT	FGFIRE	SG	QQEDFKRHRDLETAEV	NP-EQVR					
Hz-1	219	-TCYTKWLNKLANNIVGDAVIK	---	LINAGLAA	SNGG	TLPP	EKDP	PPAIDDKWT			
ChfuGV	262	FKVLNTPDTTDY	G-RENNFLNNKFD-M	LG-QS	-----						
CpGV	262	EEWYSVPDES	QSG-VEAGELNNIFP	MKSADEV	-----						
PhopGV	260	EAWYNISDTTDLD	RELKMVYNNEF-G	IASDER	-----						
PxGV	257	QEWLR	PDSAYDPD-EESKVL	-----	ASVKN	-----					
XcGV	267	DEWLAAPDDRFDDE	FEERFI	THDDYGDMA	PGTT	-----					
AcMNPV	263	AQWPSVPDNATDLE	FEKLENKNPT	NDLGMIVN	-----	GSP	QITYTAETG	FTPT			
OpMNPV	263	AQWPAVPDRADWD	FEKQFSEAPT	LOQLGDAN	-----	GVL	QISYTAETG	FTKTP			
CfMNPV	263	AQWRAVRDRAVDWN	FEKQFSEAPT	LOQLGLADN	-----	GGL	QIAYTAETG	FAATP			
LdMNPV	261	ARWPSVRDPRADE	FEKSELEFRT	ADLRIDAR	-----	T	KLVTAE	SGFATVG			
BmNPV	263	AQWPSVRDTTTTLE	FEKLETKRPT	NDLGMIVD	-----	GSP	QITYTAETG	FTKTP			
SlNPV	266	DEWPNVRDPADVD	FENKESTFKS	ISELDITQN	-----						
SeNPV	262	DEWPNVRDPEVDVD	FEIAENNYKT	ADLNDA	-----	LNKLV	TRAEV	GSPEY			
HaNPV	290	QEWPSQPRDKAPD	LELSELDYEQYSD	GLTAN	-----	T	VLE	VAENG	FRVNP		
EpNPV	263	AQWPSVRDPADVD	FEKLENQPT	LOQLGMIVN	-----	GVL	QISYTAETG	FTKTP			
RoMNPV	263	AQWPSVRDTATDLE	FEKLENKNPT	NDLGMIVN	-----	GSP	QITYTAETG	FTKTP			
McNPV	267	RDWQOTPDPAADLD	FEKLEVNYSYDDLD	HDG	-----	E	KIV	RAE	EGLSEVP		
CnBV	264	EAWPRVRDSTPHL	LEVNE	SDPTRDLDK	GLG	-----	LS	RA	VANR	GLFEF	
Hz-1	274	SGWPKNVDPNFK	PPIDVT	VSDEL	NRN	NHRSKRSPNT	IPMSDN	KIV	RFD	LAMKSKP	

ChfuGV_ 306 TQNTNMHSGIMENLLMRRREVLQNN-----INYSNKKNPNQFSGYENSFNVN
 CpGV_ 307 RGDR--RRGLMEEMMIYRNKVLKMHNYKTASTNSFISTNKINDTTNNLYKNKEKSSLNN
 PhopGV_ 305 LDEE--SRRLNENLLL-----ELFQRRQKAAARN
 PxGV_ 295 VFGVESAVKARRRLLGG-----GKAEPKRFYNTNLTTHKNN
 XcGV_ 313 SSSFSDTKRSLLRQINSARLAMLNGRERREN-----ITKSKIRVTDNPLNVSEVIADV
 AcMNPV_ 314 IAYNYRGNEP-----ARVEHFEALDRSI
 OpMNPV_ 314 IAYSARGAVP-----VARE-SRAADRAM
 CfMNPV_ 315 IAYSARVTPP-----AVRD-FGAPGRQL
 LdMNPV_ 309 MA-RDLKFRP-----ATAPPPAAFSAAA
 BmNPV_ 314 ITYNYRGTEP-----VHVEHFEALDRSI
 SlNPV_ 314 LPASKFASAY-----RIGDKTGDRKSFN
 SeNPV_ 311 MP-RTLNYRV-----ATDFVSHSVAAAA
 HaNPV_ 338 YRGTTDRWQR-----ETTTLYNDAKQTT
 EpNPV_ 314 ITYNARTAPP-----TVIK-NRVLNESV
 RoMNPV_ 314 IAYNYHGNEP-----ARVEHFEALDRSI
 McNPV_ 315 LPKRKLEFRI-----AVAPATSTLRDTN
 CnBV_ 314 RTDSPEREFP-----ETRRPTRHTN
 Hz-1_ 334 KPTKPSTSTTKNQDIEKLLN-----LVKSLNDTSGADRDLK

* *

ChfuGV_ 353 -----PSTDLEIIIIQFLEDHA---FIMSTLTDIGFSVLESLSMLQOINKVLIP
 CpGV_ 365 NLFNSTTTTNNLETIIVDFLNDHA---LIMSTLTDIGFSVLESTINNMLTOINKVLIPA
 PhopGV_ 332 -----LESKSLIEDIIVEFLEDHS---FIMSTLTDIGFNVLESTLNSMLTOINKVLIPA
 PxGV_ 331 -----FALEGLITDFLEIHE---LISSLLIDIGFSFLESTLTNMLQOINKVLIPA
 XcGV_ 367 N-----FSNDENLSMIIDFLEDHA---FIIGILTDLGFNILESQLNKLKETSQCLPI
 AcMNPV_ 337 S-----DQDLESIIITSFLEDYA---LVFGIATDIGFDMLMSCFKSMLKKINTSLIPA
 OpMNPV_ 336 S-----DDLELAIASFLLEYA---LVFGIATDIGFDMLLAFKAMLLKKINTALIPA
 CfMNPV_ 337 N-----DELELAIASFLLEYA---LVFGIATDIGFDMLLAFKAMLLKKINTALIPA
 LdMNPV_ 331 -----EALLEAIISQFLEDHS---LVLGILTDLAFDALLDQFKA LKKSTCLIPA
 BmNPV_ 337 S-----DQDLESIIITKFLEDYA---IIFGIATDAGFEMLMSCFKSMLKKINTTLIPA
 SlNPV_ 337 R-----DYDDLELAIASFFEDHA---MLFGIETDMGFALMSNEKSMMLKKINTTLIPK
 SeNPV_ 333 AA-----APTDELEAIISQFLEDNA---LILGIATSFSGFDYLFDAIKALKKINTSLIPL
 HaNPV_ 361 -----IDDQTLKDIIITQFLEDNA---LMAGIASFGFDFLFDVINDMLKPIINTQIPL
 EpNPV_ 336 T-----DELELAIASFLLEYA---LVFGIATDIGFDMLLAFKAMLLKKINTALIPA
 RoMNPV_ 337 S-----DQDLESIIITSFLEDYA---LVFGIATDIGFDMLMSCFKSMLKKINTSLIPA
 McNPV_ 338 -----SPYYTDLFVISOFLEDNA---LMGIAASFGFDFLFDNLKMLKPIINTLIPT
 CnBV_ 334 GGPIMRLEDPSLEDQIDFN NP---LISSLVSAGEDFINDGEALMKKAMVRYIPM
 Hz-1_ 370 NALASQTGKQTKFESIILSLNLTSAFASAVDVGVTOLASQLNKIMTVKAIQKIPPL

*

ChfuGV_ 403 LKRMLSLQSRVTAALLGETYKAAMINATNPAFISTISTVAKATAPTIVRAASLANALT
 CpGV_ 422 LRRMLQQSGRVTAALLGETYKAATIHANPFAITTVSTVAKATVETVKAANLANALT
 PhopGV_ 383 LTRMLSMOTRRVTAVLLGETYKASVHAMNPAFSTVSTVAKASVPAISAAASLANALT
 PxGV_ 380 LKVAASGITTFIVPRLGETYKAAVIQVLPNTLINTVSVVAATTPVTASLSVENIAQY
 XcGV_ 419 LNNLLSGGRTFTNPFARVYKATVINTVHPTLINTVALVSKMFPALKMAMSVINVL
 AcMNPV_ 386 LKMLSTTRRVTPVLLGETYKAALVHSLNVAIKTLTVTAKATPRIAICASSVIGIVLI
 OpMNPV_ 385 LKPMLVGTSQPVTVPLLGETYKAALVHSMNRIAKTLTTAAKATPRIAICASSVIGIVLI
 CfMNPV_ 386 LKPMLVSTSQPVTVPLLGETYKAALVHSMNRIAKTLTTAAKATPRIAICASSVIGIVLI
 LdMNPV_ 379 LKRALLSTSRVTSKMLGETKAAAMVHOFNRIAKTVSTVAKAMTKIGIKASVVGIVLI
 BmNPV_ 386 LKRMVSTTRPVTAIPLGETYKAAVHSENAIAIKTLTTAKATPRIAICASSVIGIVLI
 SlNPV_ 389 LRAALLSTTSRVTAPELSYVYNKAVINMSPIAIKTVTAMAKATPRIAICASSVIGIVLI
 SeNPV_ 386 LKQTLNTSRVTTPLLGETYKAAMTHAFNRIAKTISAAKAMTPRIAKASTIGITLI
 HaNPV_ 411 LQQLVLSGSPQFTTPLLGETYKAAVHSMNRIAKTVTAVAKANTHIAIKASVIGIVLI
 EpNPV_ 385 LKPMLVTTQRVTVPLLGETYKAALVHSLNRIAKTLTTAAKATPRIAICASSVIGIVLI
 RoMNPV_ 386 LKRMVSTTRRVTVPLLGETYKAALVHSLNVAIKTLTVTAKATPRIAICASSVIGIVLI
 McNPV_ 389 LKRLLYNTTEPVTVPPLGESYKAFVAFQFNRIAKTISAAKAMTPRIAKASVIGITLI
 CnBV_ 391 LQAAARFGEGTRIMVSEAFVLFVFSRINQALVLTGAAAKATAPFAMASSVIGIVLI
 Hz-1_ 430 LQKKAALADTSELVLCASRVVARDLVPVTVRSLGTCVYLTSLAMASTGVGVLL


```

ChfuGV_ 615 NLFFKIPQRFNTRAT-----NKYSSILADHFTVLEIIH
CpGV_   636 LYYLKNHTQLTLKSK-----FTMEILLVLVTCTLFIL
PhopGV_ 609 IYYTKNYTILSKSEK-----TOISTELLIIWFLILML
PxGV_   556 LRTSNGTAPATTRCS-----IDRNGSCI
XcGV_   642 VAVYKNINRKSLEPPQL-----VAFSILFLLIIGIWLMT
AcMNPV_ 596 FATNTVAFTAFVIHT-----ELIFFEFVIFLMTFYIHK
OpMNPV_ 595 FLTNTAVALAFAIAHK-----ELTFFVYFAIFLMAFYLVK
CfMNPV_ 596 FLTGTAVALAFAFLHK-----KLTFFVYFAIFLMAFYLVK
LdMNPV_ 585 WALAGVAVVIAASPTRPRPAGASTTLASLSIRAEEDLRRHASLFFVIFLITLYTVMV
BmNPV_  596 FATNTVAFTAFVIHK-----QLIFFEFVIFLMTFYIHK
SlNPV_  596 LTFLLVCAGIAWISGG-----GPVMSSIVIFILMIVEKLIN
SeNPV_  593 FIAGAVLAMMPR-----DTNVTALFIFLLIAYTLIV
HaNPV_  626 FGLSSLILMALVMITN-----DHNAVCLFVIVLLIILEVICR
EpNPV_  595 FFTNTGTIAAFIMYK-----ELTFFVYFAIFLMAFYLIK
RoMNPV_ 596 FATNTVAFTAFVIHT-----ELIFFEFVIFLMTFYIHK
McNPV_  596 FIMGALILMLMPQQ-----DTNIVALEFIFLLIAYLMID
CnBV_   609 LVALVAVLTGSQIFSTKAP-----DLATVILVILVAFVIVLQL
Hz-1_   638 LIVLRVYLLGVLMLLS-----VALVSISYANAFNVQWIT

```

```

ChfuGV_ 650 TVGAILFSFDQTQI-----
CpGV_   671 PSYQYTRIDANHQFFLE-----
PhopGV_ 644 PSYQYTLIQ--DFYKT-----
PxGV_   -----
XcGV_   679 ESEKYMMLLHHKTPAAPRKLAQNKPLVYKN-
AcMNPV_ 633 ESYEYKKTIDLLF-----
OpMNPV_ 632 EPYEYKKTIDLLF-----
CfMNPV_ 633 EPYEYKKTIDLLF-----
LdMNPV_ 645 ESYSYLLRLHREARVTASPLWYQNLDT----
BmNPV_  633 ESYGYKKTIDLLF-----
SlNPV_  634 APFMYVGVQRFAAG--DPLLWYINY-----
SeNPV_  627 DALSYVNVRRQTNYLQNR--WYDNLyse----
HaNPV_  663 SSIMYMGIRKHAQYATMP--WYHNLYT-----
EpNPV_  632 EPYEYKKTIDMLF-----
RoMNPV_ 633 ESYEYKKTIDLLF-----
McNPV_  631 ESHYMGIRAYTHRIQDR--WYDNLyte----
CnBV_   649 DRITPLARLAIVKHEENEKNRVGQRFAGLLRRA
Hz-1_   675 DRINPEKV-----

```

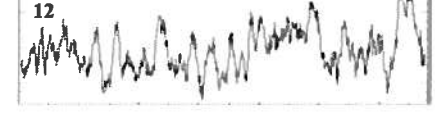
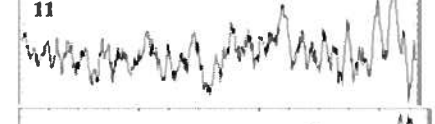
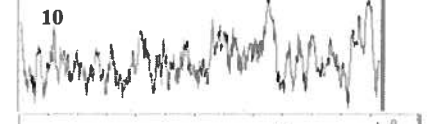
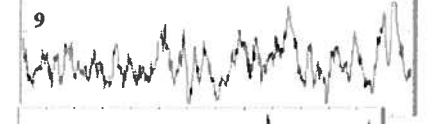
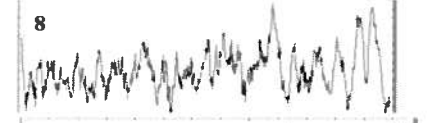
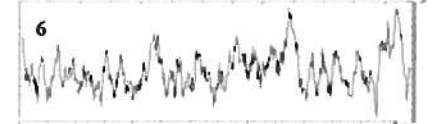
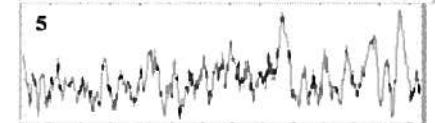
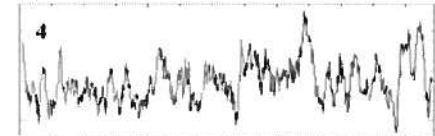
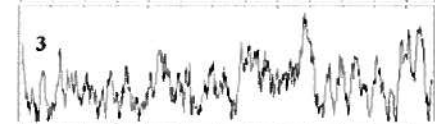
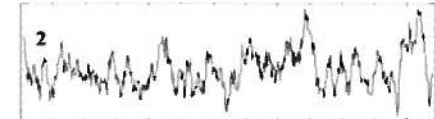
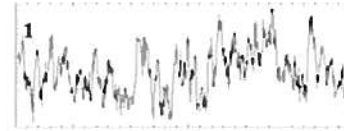
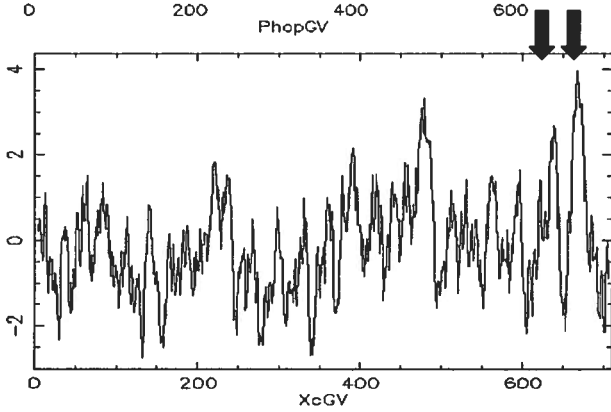
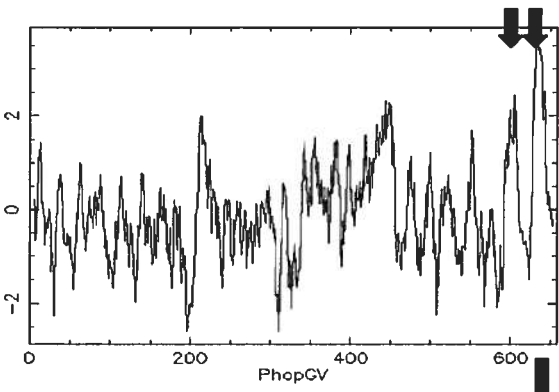
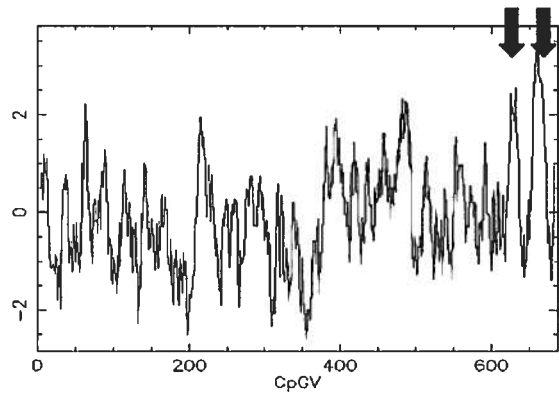
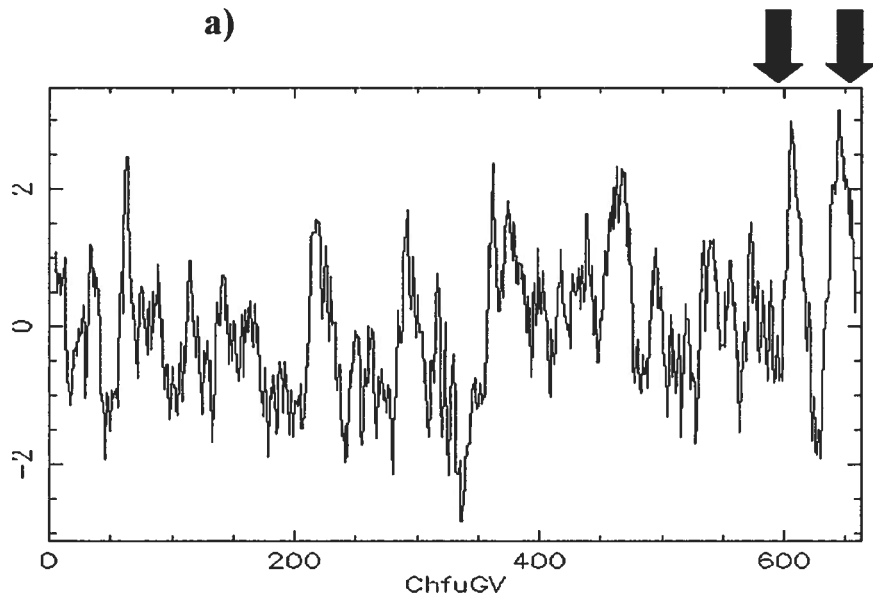
b)

		-S-S-	-S-S-	-S-S-		
ChfuGV	H2N.....-71	C-4-C-33	C-26-C-54	C-11-C-380	C-23-C-53-.....	COOH
CpGV	H2N.....-71	C-4-C-33	C-26-C-54	C-11-C-400	C-57-C-24-.....	COOH
PhopGV	H2N.....-70	C-4-C-33	C-25-C-54	C-11-C-363	C-91-.....	COOH
PxGV	H2N.....-70	C-4-C-32	C-25-C-54	C-11-C-366	C- 7-C-1-.....	COOH
XcGV	H2N-11-C-59	C-4-C-32	C-26-C-54	C-17-C-500	COOH
AcMNPV	H2N.....-71	C-4-C-31	C-29-C-54	C-11-C-439	COOH
OpMNPV	H2N.....-71	C-4-C-31	C-29-C-54	C-11-C-438	COOH
CfMNPV	H2N.....-71	C-4-C-31	C-29-C-54	C-11-C-439	COOH
LdMNPV	H2N.....-70	C-4-C-31	C-29-C-53	C-11-C-378	C-89-	COOH
BmNPV	H2N.....-71	C-4-C-31	C-29-C-54	C-11-C-439	COOH
SlNPV	H2N.....-74	C-4-C-31	C-29-C-54	C-11-C-378	C-14-C-54-.....	COOH
SeNPV	H2N.....-70	C-4-C-31	C-29-C-54	C-11-C-448	COOH
HaNPV	H2N.....-96	C-4-C-31	C-29-C-54	C-12-C-414	C-13-C-27-.....	COOH
EpNPV	H2N.....-71	C-4-C-31	C-29-C-54	C-11-C-438	COOH
RoNPV	H2N.....-71	C-4-C-31	C-29-C-54	C-11-C-439	COOH
McNPV	H2N-20-C-54	C-4-C-31	C-29-C-54	C-11-C-78	C-368-.....	COOH
CnBV	H2N.....-72	C-4-C-31	C-28-C-54	C-12-C-474	COOH
H _z -1	H2N.....-87	C-4-C-8	C-20-C-3	C-30-C-47	C-13-C-322-C-139-	COOH

ChfuGV	71.....	205
CpGV	71.....	205
PhopGV	70.....	203
PxGV	70.....	202
XcGV	71.....	203
AcMNPV	71.....	206
OpMNPV	71.....	206
CfMNPV	71.....	206
LdMNPV	70.....	204
BmNPV	71.....	206
SlNPV	74.....	209
SeNPV	70.....	205
HaNPV	96.....	232
EpNPV	71.....	206
RoNPV	71.....	206
McNPV	75.....	210
CnBV	72.....	205
H _z -1	87.....	158

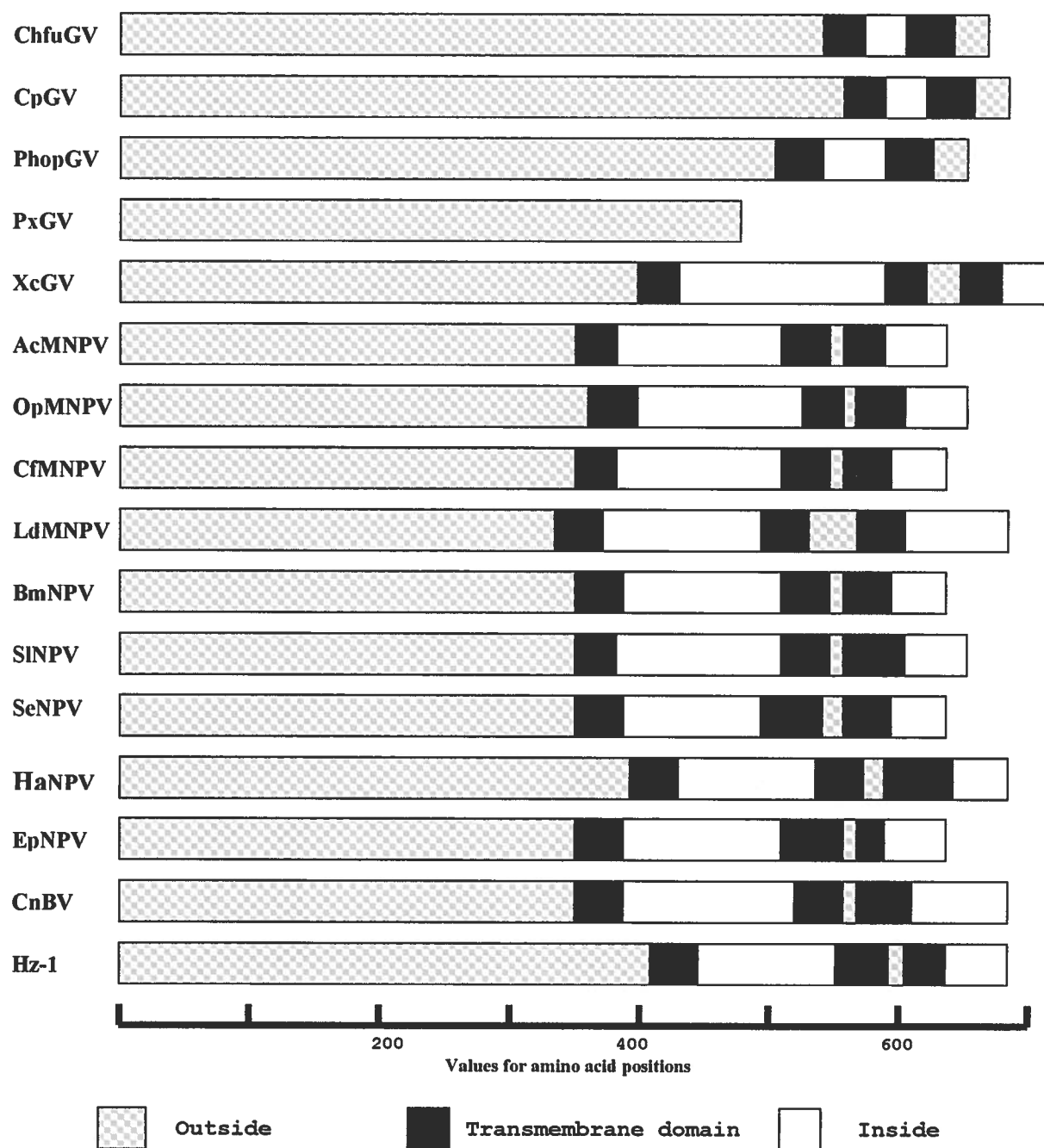
Figure 4 Structural analysis of p74. a) Predicted Kyte-Doolittle hydrophobic profiles of ChfuGV p74 and those of other GVs, NPVs, CnBV and Hz-1 virus. Arrows indicate two C-terminal hydrophobic domains. Profiles were generated using Pepwindow hydropathy analysis (7 residue average; Kyte and Doolittle, 1982). Positions of the amino acids are shown on the X-axis and hydrophobicity values are shown on the Y-axis. The list of baculoviral p74s in the right column is as follow: 1-PxGV; 2- AcMNPV; 3- OpMNPV; 4- CfMNPV; 5- LdMNPV; 6- BmNPV; 7- SINPV; 8- SeNPV; 9- HaNPV; 10- EpNPV; 11- CnBV; 12- Hz-1 virus. b) Demonstration of transmembrane, out side, and inside domains of p74 proteins. (c) Schematic presentation of ChfuGV p74. Computer analysis (TMHMM-ver. 2) of the amino acid sequence predict that the N terminal of the protein is located outside of the envelope followed by two transmembrane regions which results in an external C terminal. Each transmembrane domain spans about 20 amino acids and are shown as cylinders.

a)



┌ = 100 aa

b)



c)

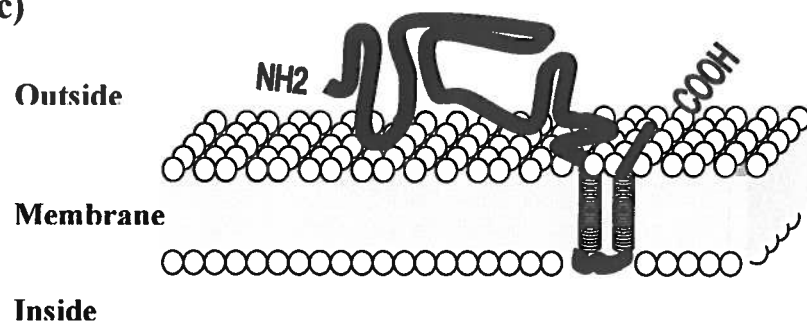


Figure 5. Phylogenetic analysis performed on baculoviral p74 proteins. The unrooted tree was produced by a Branch and Bond search using PAUP4.0b4a. Bootstrap values (100 replicates) are shown. Tree was constructed according to the following settings: 1- uninformative characters ignored. 2- Branch-and –bound search options: an initial upper bound: compute via stepwise; keep minimal only; collapse zero-length branches ON; MULPARS ON; addition sequences: furthest. Accession numbers for protein sequences used are shown in material and methods section.

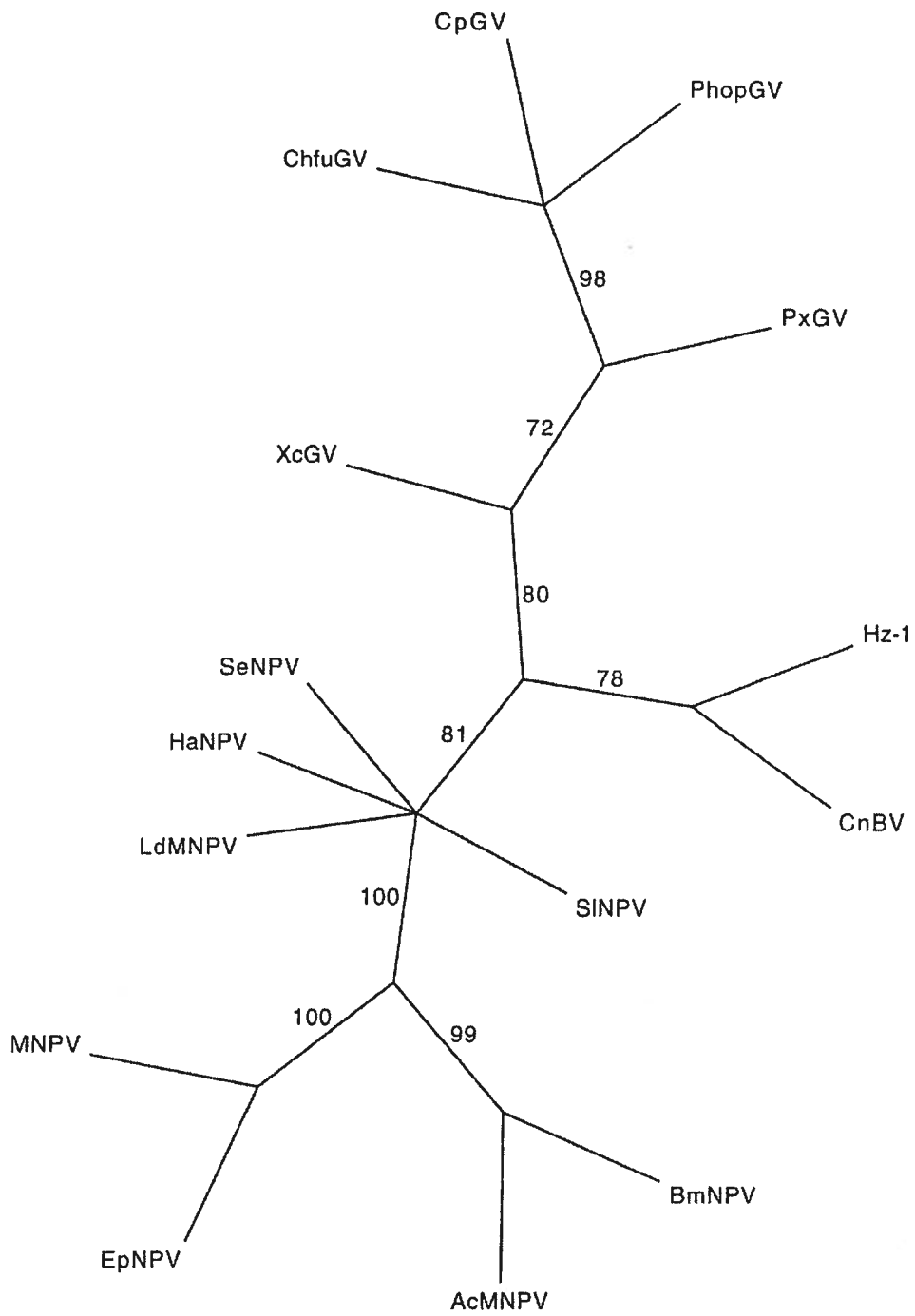


Table 1. Comparison of conceptual amino acid sequences of ChfuGV p74 with those of other baculoviruses studied here. Pair wise identity values (%) between taxa are shown.

[illegible]

6.5. Acknowledgements

The authors would like to thank Dr. Anne Bruneau, Department of Biology, University of Montreal for helping us in the phylogenetic studies. We also thank Dr. David Morse, Department of Biology, University of Montreal for his scientific discussions. This research was funded by the *Ministère des Ressources Naturelles du Québec* Grant number 0316-2405CG.

CHAPTER 7 – PUBLICATION #6

Transcription, Translation, and Immunolocalization of ODVP-6E/ODV-E56 and p74 Proteins: Two Highly Conserved ODV-associated Envelope Proteins of *Choristoneura fumiferana* Granulovirus.

Authors: Kianoush Khajeh Rashidan^{1*}, Nasha Nassoury², Paresa N. Giannopoulos¹ and Claude Guertin¹

1- Institut national de la recherche scientifique-Institut Armand-Frappier, 531 Blvd des Prairies, H7V 1B7, Laval, Canada

2- Université de Montréal, 4101 Sherbrooke east, H1X 2B2 Montreal, Canada

Published in Journal of Biochemistry and Molecular Biology in 2005. Number 38, pages 65-70.

Contribution of the authors

I was the principal investigator in this report while Dr. Nassoury's expertise in immunolocalization was used to complete this report. Mrs. Giannopoulos helped me by performing the larvae infection and larvae harvesting. Dr. Guertin was the director of the project and the majority of the experiences have been realized in his laboratory.

Abstract in french :**Transcription, traduction et immunolocalisation de l'ODVP-6E/ODV-E56 et p74: deux protéines hautement conservées chez le phénotype ODV du granulovirus de *Choristoneura fumiferana***

Deux phénotypes viraux sont associés à l'infection des larves de *Choristoneura fumiferana* par le granulovirus (ChfuGV). Il s'agit des deux types de virions enveloppés, ODV et BV. Deux des protéines majeures hautement conservées associées à l'enveloppe de l'ODV sont l'ODVP-6E/ODV-E56 et p74, lesquelles pourraient jouer un rôle dans l'attachement et/ou la pénétration du virus aux cellules de l'intestin moyen de l'insecte hôte. Cette étude présente la caractérisation transcriptionnelle et traductionnelle de l'ODVP-6E/ODV-E56 et p74 du ChfuGV ainsi que l'immunolocalisation de ces protéines dans le virion inclus du ChfuGV. Les résultats de cette étude révèlent que l'expression des gènes *odvp-6e/odv-e56* et p74 débute à 24 h. p.i. En utilisant des anticorps polyclonaux monospécifiques produits contre l'ODVP-6E/ODV-E56 et p74, nous avons démontré que ces protéines sont exprimées tardivement durant l'infection. L'immunolocalisation à l'or en utilisant l'antisérum contre l'ODVP-6E/ODV-E56 et p74 a mis en évidence que l'ODVP-6E/ODV-E56 et p74 sont associées à l'enveloppe de l'ODV du ChfuGV.

7.1. Abstract

Choristoneura fumiferana granulovirus (ChfuGV) infection results two types of enveloped virions, ODV and BV. Structural proteins ODVP-6E/ODV-E56 and p74 are two major conserved ODV-associated proteins that may be involved in the initiation of viral infection cycle in susceptible host insect larvae. This study presents the characterization of ChfuGV *odvp-6e/odv-e56* and p74 transcription and translation as well as immunolocalization of these proteins in the occluded ChfuGV virion. Our results revealed that the transcription of *odvp-6e/odv-e56* and p74 genes, both, start at 24 hours post infection (hr p.i.). Using monospecific polyclonal antibodies made against ODVP-6E/ODV-E56 and p74 we demonstrated that these proteins are both expressed late in infection (24hr. p.i.). Immunogold labeling using antisera against ODVP-6E/ODV-E56 and p74 proteins demonstrated that ODVP-6E/ODV-E56 and p74 proteins are both associated with the ODV envelope of ChfuGV.

Keywords: Baculovirus, *Choristoneura fumiferana* granulovirus, ODV envelope associated proteins, ODVP-6E/ODV-E56 and p74 proteins, transcription, translation and immunolocalization.

7.2. Introduction

The *Baculoviridae* are complex and diverse family of entomopathogens that infect especially insects belonging largely to the orders *Lepidoptera*, *Dieptera* and *Hymenoptera* (Crook, 1991). This family of viruses contains a double-stranded, supercoiled DNA genome which is packed in a rod-shaped enveloped nucleocapsid (Miller, 2001). Members of *Baculoviridae* family are comprised of two genera: the nucleopolyhedroviruses (NPVs) and the granuloviruses (GVs). Granuloviruses have been well characterized at the cytological level and their application for pest control in agriculture and forestry has been fairly successful (Crook, 1991; Moscardi, 1999). The spruce budworm, *Choristoneura fumiferana*, is the most devastating coniferous tree pest in eastern Canada and the United States. *C. fumiferana* granulovirus (ChfuGV) is being considered as an alternative biological insecticide due to the specificity for its insect host.

The Infection cycle initiates after *C. fumiferana* (spruce budworm) larvae ingest the occluded virions. The proteinous matrix of occlusion bodies dissociates under the alkaline pH condition in the midgut and liberates enveloped virions known as an occlusion-derived virus (ODV) (Miller, 2001). Throughout the baculoviral infection cycle, two virion phenotypes are produced: budded virus (BV) and occlusion-derived virus (ODV). These phenotypes can be distinguished by the origin of their envelopes: BVs obtain their envelopes during budding process through the plasma membrane of the infected cells (Blissard, 1996; Adams and Mc Clintock, 1991; Blissard and Rohrmann, 1989) and ODVs attain their envelope inside the nucleus of the infected cell from intracellular microvesicles within the invaginated inner nuclear membrane (Braunagel

and Summers, 1994; Hong *et al.*, 1994; Braunagel *et al.*, 1996a). The BV phenotype is responsible for the spread of infection from cell to cell inside the infected insect host body (Blissard, 1996; Tanada and Hess, 1991). The ODV phenotype is produced in late steps of infection within the nucleus of infected cells and is subsequently occluded in a proteinous matrix to form the occluded virions. ODVs transmit infection exclusively from insect to insect, while the BVs spread infection from infected to healthy cells inside the infected host (Blissard, 1996).

The attachment of ODVs to microvilli of insect midgut cells followed by the virus penetration is considered as vital steps for the initiation of infection cycle in the host. The protein(s) that could participate in these steps have not yet been identified but there is evidence suggesting that the ODV penetration is a non-endocytotic process and requires interaction of the virion envelope proteins with midgut cell surface proteins which eventually leads to membrane fusion (Horton and Burand, 1993).

In ChfugV, like other baculoviruses, the ODV phenotype has a complex structure. Baculoviral ODV envelopes contain a number of proteins such as: VP17 (Funk and Consigli, 1993), ODV-E25 (Russell and Rohrmann, 1993), ODV-E35 (Braunagel *et al.*, 1996b), GP41 (Whitford and Faulkner, 1993), p74 (Rashidan *et al.*, 2003; Kuzio *et al.*, 1989), ODV-E18 (Braunagel *et al.*, 1996b), ODV-E66 (Hong *et al.*, 1994), and ODVP-6E/ODV-E56 (Rashidan *et al.*, 2002; Braunagel *et al.*, 1996a, Theilmann *et al.*, 1996). One or more of these proteins could participate in the adsorption, fusion and penetration process.

Baculoviral ODVP-6E/ODV-E56 protein is an extremely conserved protein (Rashidan *et al.*, 2002); the elevated extent of primary and secondary structures preservation, even between two distantly related members of *Baculoviridae* family, points out the significance of ODVP-6E/ODV-E56 and suggest a probable major role for this protein during the infection cycle. The deduced amino acid composition of ChfuGV ODVP-6E/ODV-E56 indicates the presence of three potential *N*- glycosylation N-{P}-[ST]-{P} sites located at residues 59, 179, and 279. Two hydrophobic membrane-spanning regions and six conserved cysteine residues between two hydrophobic membrane-spanning regions has been identified in ChfuGV ODVP-6E/ODV-E56 and those of other baculoviruses (Rashidan *et al.*, 2002). These highly conserved cysteine residues have been speculated to be implicated in the creation of disulfide bridges that ultimately form the right folding of baculoviral ODVP-6E/ODV-E56 proteins (Rashidan *et al.*, 2002).

Baculoviral p74 protein is another major protein associated with the ODV envelope. Studies on this protein in *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) (prototype of the *Baculoviridae* family) revealed that p74 protein is exposed on the virion surface with an N-terminal located outside of the ODV envelope and a transmembrane C-terminal that acts as an anchor (Slack *et al.*, 2001, Faulkner *et al.*, 1997). It has been shown that the baculoviral p74 protein may play a crucial role in the virulence of the virus (Kuzio *et al.* 1989; Faulkner *et al.*, 1997). It has been shown that although p74 null mutants fail to initiate infection in insect larvae when virus are fed orally, they are infectious once injected into the haemocoel of the larvae (Faulkner *et al.*,

1997). These observations suggest p74 protein is essential for initiation of infection and its presence is an essential constituent for attachment and /or fusion of the ODV envelope to the membrane of midgut epithelial cells. Wu and his associates have demonstrated that baculovirus p74 gene is a species-specific gene (Wu *et al.*, 2003). They have shown that that the substitution of AcMNPV p74 knock out with a p74 gene from *Spodoptera litura* nucleopolyhedrovirus (SINPV), would inhibit recombinant AcMNPV from infecting the *Argyrogramma agnata* larvae (Wu *et al.*, 2003).

By nucleotide and amino acid sequence analyses it has been predicted that ChfuGV ODVP-6E/ODV-E56 and p74 proteins were ODV-envelope associated proteins that are expressed during the late stages of infection cycle (Rashidan *et al.*, 2002; 2003). Here, we confirm, by transcriptional and translational analyses, the late expression of these proteins. We also examine the validity of our computer assisted predictions which classified ChfuGV ODVP-6E and p74 as ODV associated envelope proteins by the means of Electron microscopy immunolocalization.

7.3. Materials and Methods

7.3.1. RNA extraction and analysis of Transcription

Total RNA was extracted from the fat body of the infected fourth instar larvae at different hours post infection (hr p.i.). Extraction was carried out using the TRIZOL reagent (Stratagene) according to the protocol provided by supplier. Extracted RNAs were quantified and stored under RNAGuard and ethanol at -70°C . RNA samples (12ug) were loaded on formaldehyde-agarose gels and electrophoresis was carried out. RNA was either visualized by ethidium bromide staining or transferred to Hybond N nylon membrane by capillary blotting as described in standard protocols (Sambrook *et al.*, 1989). Transcripts of *odvp-6e/odv-e56* and p74 genes were detected by Northern hybridization using ^{32}P [dATP] DNA probes labeled by random priming (PROMEGA).

The probe for *odvp-6e/odv-e56* was generated using a 758bp PCR fragment as template. The fragment was amplified using the following two *odvp-6e/odv-e56* internal primers: *FRe56*-TTT ACG GGT CTA AGA CGC AC, and *RVe56*- ACG ACA AAC ACT TTG TTC. Meanwhile the probe for p74 was generated using a 1161bp PCR fragment as template, amplified using the following two p74 internal primers: *FRp74*: CCT ATT GTC GCC GCT TTG GTC GTT CTG and *RVp74*: GT GGT AGA CGC CAA ACT AGC ACC.

Temporal transcriptional activity of the *odvp-6e/odv-e56* transcripts were examined by RT-PCR using total RNA isolated from ChfuGV-infected larvae collected at various times p.i. RT-PCR reactions were realized with an one-step Reverse Transcriptase-PCR (Invitrogen, Life technologies), as provided by manufacturer, using two internal specific primers for the *odvp-6e/odv-e56* gene which have been mentioned before (*FRe56*- and *RVe56*) . Using these primers a 758 bp fragment of the target gene was amplified, cloned into a *PCR-2.1* cloning vector (Invitrogen), and sequenced to confirm the identity of amplicon. The same process was repeated for p74 using p74-specific primers and probes.

7.3.2. Total protein extraction from infected *Choristoneura fumiferana* larvae

Total protein was extracted from the fat body of infected larvae at different hours post infection in order to monitor the pattern of ChfuGV ODVP-6E/ODV-E56 and p74 protein expression. Fourth instar *C. fumiferana* larvae were infected with 1 µl of purified ChfuGV (10^8 virus / ml) by force feeding. The infected larvae were then transferred to a normal insect diet and were used for time course expression studies. The fat body of the larvae was harvested at various times post infection and placed in 0.01 M Tris, pH 7.5 with 1% SDS and protease inhibitors (10 µg / ml each of aprotinin, leupetin, phenylmethylsulfonyl fluoride, *N*-tosyl-phenylalanyl-chloromethyl ketone, n-tosyl-lysyl chloromethyl ketone; Sigma). The sample was passed through a 27-gauge needle attached to a tuberculin syringe using firm and rapid pressure on the plunger and then was subjected to a freeze-thaw cycle. These two steps were repeated one more time to disrupt the cells.

7.3.3. SDS-PAGE and Western blot analysis

Protein samples were disrupted in an equal volume of Laemmli sample buffer (Laemmli, 1970) in presence of β -mercaptoethanol (Bio Rad). The samples were then boiled for 5 min and clarified by centrifugation at 13,000 g for 5 min before sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 10% acrylamide gel. The protein gels were then electrophoretically transferred to a nitrocellulose membrane (Millipore). The membrane was then examined for presence of proteins by staining the membrane with Ponceau S (Sigma). After washing with PBS buffer, membranes were blocked in PBS containing 3% BSA for 1 hr in room temperature followed by 2h incubation with anti-ODVP-6E/ODV-E56 (dilution 1:8000) or anti-p74 (dilution 1:5000). Membranes were then washed in PBS (pH 7.2), and reacted with the peroxidase-conjugated secondary antibody and visualized using a chemiluminescence substrate (Amersham).

7.3.4. Monospecific antisera preparation

Gel purified ChfuvGV structural proteins were used for production of monospecific polyclonal antibodies. For preparation of antisera purified enveloped nucleocapsids were subjected to electrophoresis on 10% SDS-PAGE as was explained before. Gels were then lightly (10 min) stained in 0.05 % Coomassie brilliant blue (Sigma) to locate the positions of the protein bands in the gels. After destaining, 39 and 74 kDa bands were excised from the gels and washed in deionized water three times each for 10 min followed by an over night wash. Following the washes, the preparation of proteins was injected to animals as described by Harlow and Lane (1988). Briefly, the gel pieces were extruded through the barrel of a 5ml syringe with an equal volume of deionized water to

another 5-ml syringe. Antisera were produced in 6 weeks, using male New Zealand rabbits by subcutaneous injection of each antigen at multiple sites along the back. ODVP-6E/ODV-E56 and p74 proteins were injected four times. The initial injection was performed with Freund's complete adjuvant and subsequent injections were done with incomplete Freund's adjuvant. A quantity equal to 1.5 mg of either ODVP-6E/ODV-E56 or p74 protein was used for immunization.

7.3.5. Immunogold localization of ODVP-6E and p74 proteins

Immunoelectron microscopy was carried out on complete occluded ChfV particles. Purified occluded virions were fixed in 1 ml 1 % glutaraldehyde / 1 % paraformaldehyde / phosphate buffer (0.028 M Na_2HPO_4 , 0.0387 M KH_2PO_4 , pH 7.2) for 1 h in room temperature. The sample was then centrifuged in 8000 g for 10 min and the pellet was resuspended in 150 μl in 0.25% agarose / phosphate buffer at 37 °C.

The agarose plug was cut into small blocks and incubated for 1 h at room temperature. Samples were washed twice in phosphate buffer, each for 10 min. The samples were dehydrated in 10, 30, 50, 70, 90 and 100 % ethanol for 15 min in each concentration. Following dehydration with ethanol, the virus were infiltrated in a 1: 1 mixture of L. R. White resin and 100% ethanol for 30 min, and then incubated in three changes of fresh LR White resin for 1 h. The samples then were placed in embedding capsules, which were filled with fresh L. R. White resin and incubated at 52-55°C for 12 hours. Thin sections were collected on formvar-coated nickel grids.

The grids were incubated in 0.1% trypsin in PBS buffer for 15 min in order to expose antigenic sites. After washing with distilled water, non-specific binding sites were

blocked with 3 % BSA in EM-immunogold (EMG) buffer (0.05 % Tween, 0.5 *M* NaCl, 0.01 *M* phosphate buffer, pH 7.2) for 60 min. The sections were then incubated for 90 min with the primary antibody diluted in EMG with 0.5 % BSA (1:1000 for both anti ODVP-6E/ODV-E56 and anti-p74). After two rounds of washing with EMG buffer with 0.5 % BSA, the grids were incubated with goat anti-rabbit antibody conjugated to 10 nm gold particles (Ted Pella Inc.) for 90 min. Grids were washed twice each in EMG buffer with 0.5 % BSA, PBS, and distilled water and stained for 15 min with 3 % uranylacetate and 5 min with lead citrate. TEM was performed using a Philips model 410 transmission electron microscope operating at 80 kV.

7.4. Results and Discussion

In a previous study we have reported the identification of genes encoding proteins homologue to ODVP-6E/ODV-E56 and p74 proteins in ChfuGV. In those reports we have analyzed the nucleotide sequences of *odvp-6e/odv-e56* and p74 genes and their deduced amino acid sequences (Rashidan *et al.*, 2002; 2003). The genes encoding for these proteins were respectively located in an 11-kb and 8.9-kb *Bam*HI restriction fragments from ChfuGV genomic DNA.

7.4.1. Temporal transcription and translation of ChfuGV *odvp-6e/odv-e56* gene

The open reading frame (ORF) is 1062 nt, potentially encoding 353 amino acids with an estimate molecular mass of 38.5 kDa. A late promoter motif (GTAAG) was located at – 14 nt upstream to the first ATG. The presence of a TAAG motif, which is a known strong late gene promoter in baculoviruses, signals the strong possibility that ChfuGV *odvp-6e/odv-e56* is expressed during the late phase of the infection cycle.

To examine the temporal nature of *odvp-6e/odv-e56* transcription, RT-PCR and Northern blot analyses were performed. RT-PCR analysis on total RNA extracted from infected larvae using the *odvp-6e/odv-e56* specific primers detected the presence of a transcript from 24 hr p.i (Fig.1a). The sequences of the RT-PCR amplicons showed that all amplicons corresponded to ChfuGV *odvp-6e/odv-e56* transcripts present in the infected larvae. To substantiate the results obtained by the RT-PCR, Northern blot analysis was performed using a ³²P-labeled probe. Results of Northern blot analysis revealed the presence of a prominent band detected at 1.2 kb in samples taken from larvae at 48 hr p.i

(Fig. 1b). Hybridization reactions using the ^{32}P -labeled probe on non-infected larvae RNA demonstrated no reaction (data not shown). Studies done on several baculoviral ODVP-6E/ODV-E56 proteins identified them as ODV associated envelope proteins encoded by baculovirus late genes and in all these genes TAAG motifs have been observed in the 5' Untranslated regions (UTRs) (Braunagel *et al.*, 1996a; Theilman *et al.*, 1996).

In order to characterize the expression pattern of ChfuGV ODVP-6E/ODV-E56 protein, an anti- ODVP-6E/ODV-E56 monospecific antiserum was produced using a SDS-PAGE purified 39 kDa protein. The specificity of this antiserum was examined performing Western blot analysis on ODV enveloped nucleocapsid proteins (data not shown). After confirming its specificity, anti- ODVP-6E/ODV-E56 antisera were used on total protein extracts of infected larvae collected at different times p.i. A band of 39 kDa was first detected at 24 hr p.i in infected larvae (Fig.2). Taken together, these data indicate that the ChfuGV ODVP-6E/ODV-E56 is a late protein and is translated from a 1.2 kb mRNA. The presence of ChfuGV ODVP-6E/ODV-E56 protein at the same time as its transcripts suggests that the expression of this protein is controlled at the transcriptional level.

7.4.2. Temporal transcription and translation of ChfuGV p74 gene

The gene has a 1992 nucleotide (nt) open reading frame (ORF) encoding a protein with 663 amino acids with a predicted molecular mass of 74.8 kDa. Two late promoter motifs (TAAG and GTAAG) were located upstream of the first ATG of p74 gene. As it mentioned before, the existence of a TAAG motif indicates the elevated likelihood that the gene is expressed during the late phase of the infection cycle.

To examine the temporal nature of p74 transcription, RT-PCR and Northern blot analyses were performed. RT-PCR analysis on total RNA extracted from infected larvae using the p74 specific primers detected the presence of a transcript from 24 hr p.i (Fig.1c). The sequences of RT-PCR amplicons showed that all amplicons corresponded to ChfuGV p74 transcripts present in the infected larvae. To substantiate the results obtained by the RT-PCR, Northern blot analysis was performed using a ^{32}P -labeled probe. Results of Northern blot analysis revealed the presence of a band at 2.0 kb in samples taken from larvae at 48 hr p.i (Fig.1d). Hybridization reactions using the ^{32}P -labeled probe on non-infected larvae RNA demonstrated no reaction (data not shown).

In an attempt to examine the expression pattern of ChfuGV p74 protein an anti-p74 monospecific antiserum was produced using a SDS-PAGE purified 74 kDa protein. The specificity of this antiserum was examined performing Western blot analysis on ODV enveloped nucleocapsid proteins. After confirming its specificity, anti- p74 antisera were used on total protein extracts of infected larvae collected at different times p.i. A band of 74 kDa was first detected at 24 hr p.i. in infected larvae (Fig.2b). Altogether, these information point to the fact that ChfuGV p74 is a late protein and is translated from a 2.0 kb mRNA. The presence of ChfuGV p74 protein at the same time as its transcripts, as we have also observed in the case of ChfuGV ODVP-6E/ODV-E56, suggests that the expression of ChfuGV p74 protein is possibly controlled at the transcriptional level.

7.4.3. Immunolocalization of ChfuGV ODVP-6E/ODV-E56 and p74 proteins

Analysis of ODVP-6E/ODV-E56 hydrophobicity plots suggested the presence of two hydrophobic domains within the C-terminal that were conserved in all baculoviral ODVP-6E/ODV-E56 proteins. These highly hydrophobic regions suggested the existence of two membrane-spanning regions in ChfuGV ODVP-6E/ODV-E56 protein. The same hydrophobic regions were also detected in all other known 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 (Rashidan *et al.*, 2002).

Studies on ChfuGV p74 using various computer assisted prediction tools suggested that ChfuGV p74 protein is a membrane protein with two membrane-spanning regions within its C-terminal region. These transmembrane domains were located between amino acid residues 600 to 620 and 633 to 656. In addition to these two regions, an additional transmembrane region was located almost in the center of ChfuGV p74 protein (Rashidan *et al.*, 2003).

To confirm our previous computer-assisted prediction of a membrane location for the ChfuGV ODVP-6E/ODV-E56 and p74 proteins, we performed an immunogold electron microscopy (IEM) analysis on occluded ChfuGV virion using either anti-ODVP-6E/ODV-E56 or anti-p74 monospecific antiserums. ODV was purified and consequently prepared for IEM. The results of IEM confirmed our prediction by localizing the ChfuGV ODVP-6E/ODV-E56 and p74 proteins on the envelope region of the occluded virion (Fig.3a, b). To evaluate the accuracy of our IEM method we performed a count of gold beads associated with three regions on individual virions to obtain the total number of

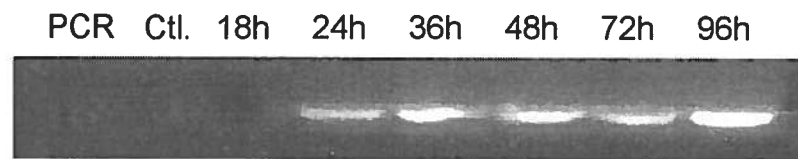
beads associated with capsids, envelopes and on occlusion bodies of 150 virions. The results in the case of ODVP-6E/ODV-E56 demonstrated that 78.68% of beads were associated with envelopes while 6.36% and 13.96% were associated with capsids and occlusion bodies respectively. In the case of this protein the overall average of label density on envelopes was roughly 6 and 13 fold greater than the occlusion bodies and the capsids, respectively. The results for p74 revealed almost the same proportions: 85.58% of beads were associated with envelopes, 5.65% associated with capsids and 8.77% were localized on occlusion bodies. In the case of p74 protein the overall average of label density on envelopes was roughly 10 and 17 fold greater than the occlusion bodies and the capsids, respectively. These observations are in complete agreement with the findings on other baculoviral ODVP-6E/ODV-E56 and p74 homologue proteins. Studies carried on *AcMNPV*, *Orgyia pseudotsugata* multinucleocapsid nucleopolyhedrovirus (OpMNPV), and *Cydia pomonella* granulovirus (CpGV) ODVP-6E/ODV-E56 protein homologues all characterized this protein as a late protein associated with the ODV phenotype enveloped nucleocapsid (Braunagel *et al.*, 1996a; Theilmann *et al.*, 1996). The result of immunoelectron microscopy demonstrated that ChfuvGV ODVP-6E/ODV-E56 and p74 proteins are indeed bona fide ODV envelope associated proteins. Results of immunolocalization with preimmune antisera demonstrated no unspecific reaction (data not shown).

Similar studies performed on other baculoviruses demonstrated the absence of ODVP-6E/ODV-E56 protein in BV phenotype. Presently, the absence of permissive cell lines for ChfuvGV denied us the chance to compare the BV and ODV phenotypes in regard to the

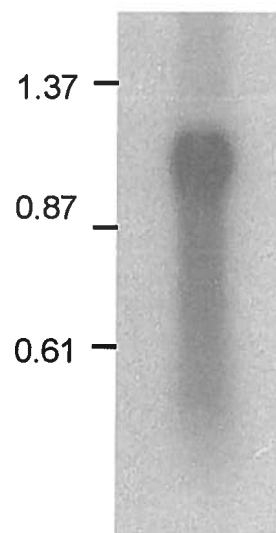
ODVP-6E/ODV-E56 and p74 proteins. This would be a very interesting subject to look at in future studies on ODVP-6E/ODV-E56 and p74 proteins. Another attractive study which may shed more light on characterization of ChfuGV ODVP-6E/ODV-E56 and p74 proteins is IEM studies on cells infected with ChfuGV.

Figure 1. (a) Analysis of *odvp-6e/odv-e56* gene transcripts using RT-PCR using *odvp-6e/odv-e56*-specific primer sets and total RNA purified from infected *Choristoneura fumiferana* larvae or uninfected larvae (Ctl.). The result of PCR when the reverse-transcription step is omitted is shown on the first lane. (b) Northern blot analysis on transcripts from infected larvae extracted 48 hr p.i. performed using a ^{32}P labeled DNA probe specific to *odvp-6e/odv-e56* revealed the presence of a prominent band detected at 1.2 kb. (c) Analysis of p74 gene transcripts using RT-PCR using p74-specific primer sets and total RNA purified from infected *Choristoneura fumiferana* larvae or uninfected larvae (Ctl.). The result of PCR when the reverse-transcription step is omitted is shown on the last lane. (d) Northern blot analysis on transcripts from infected larvae extracted 48 hr p.i. performed using a ^{32}P labeled DNA probe specific to p74. The Northern blot analysis revealed the presence of a smear extent from 1.7 to almost 2 kb.

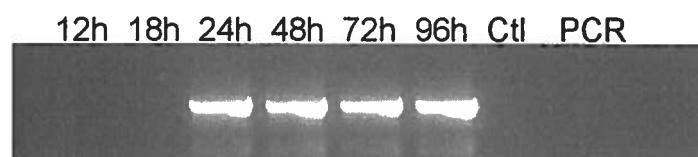
(a)



(b)



(c)



d)

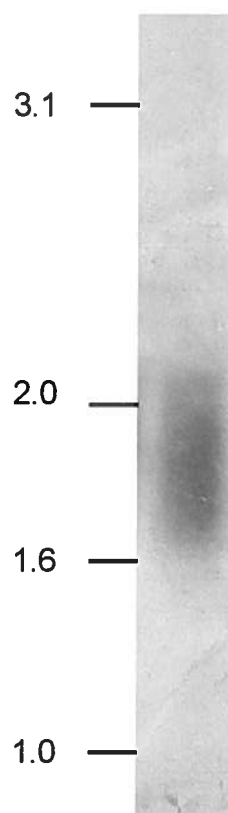


Figure 2. Preparation of uninfected (UL) or infected larvae with ChfuGV. Immunoblot analysis of ODVP-6E/ODV-E56 and p74 expression in ChfuGV infected fourth instar larvae. Larvae were harvested at different hours post infection (hr p.i.). Samples were analyzed by immunoblotting with the (a) anti-ODVP-6E/ODV-E56 antiserum. Band at 39kDa and (b) anti-p74 antiserum. Band at 74 kDa. Uninfected larvae (UL) were used as control.

(a)

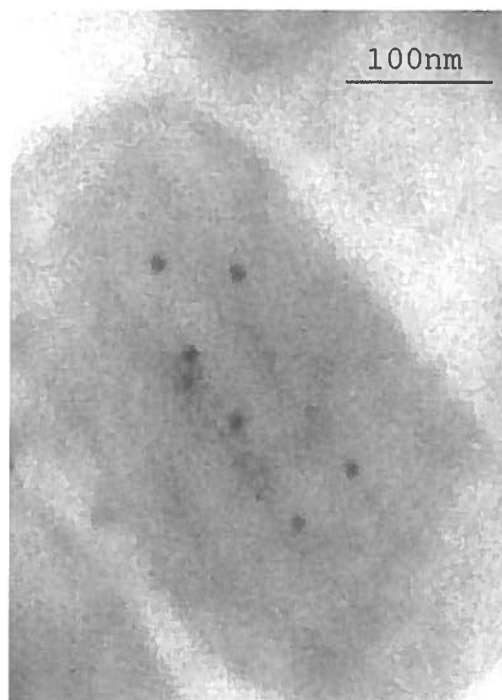
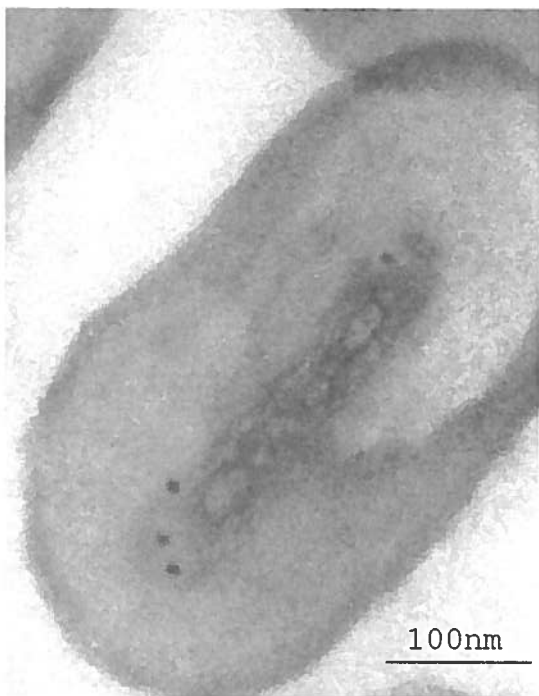


(b)

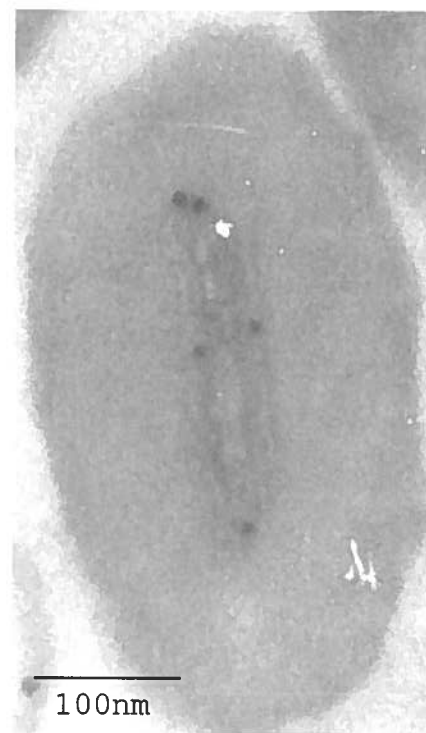
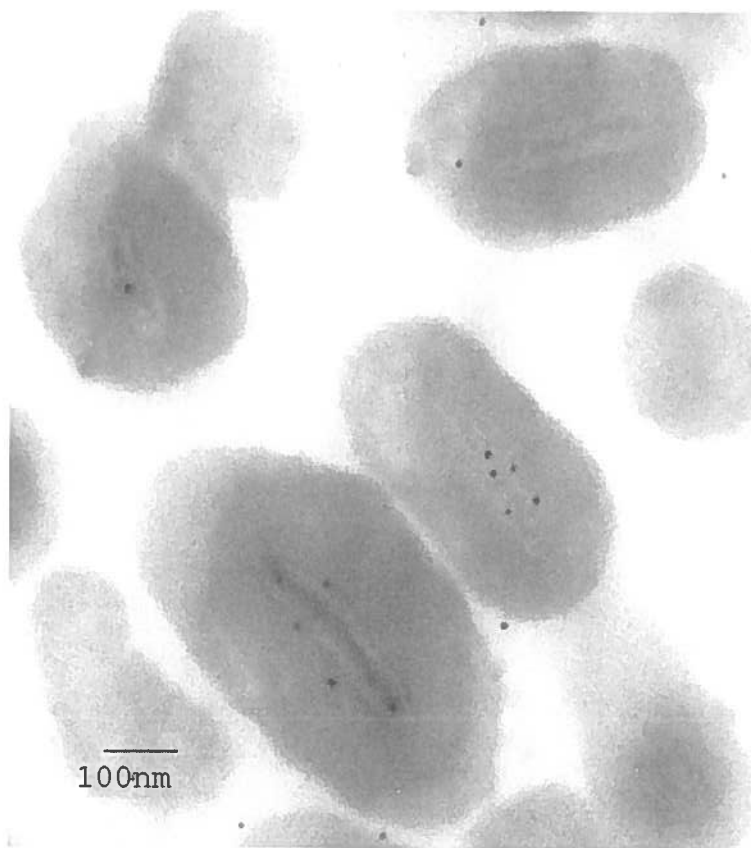


Figure 3. Immunoelectron microscopy of the ODV phenotype of ChfuGV. Thin sections of occluded virions were incubated with (a) antisera to ODVP-6E/ODV-E56 and (b) antisera to p74 protein, then the grids were incubated with goat anti-rabbit antibody conjugated to 10 nm gold particles and stained with 3 % uranyl acetate and 5 min with lead citrate. The grids then were viewed on a Philips model 410 transmission electron microscope operating at 80 kV.

(a)



(b)



7.5. Acknowledgements

The authors would like to thank Dr. David Morse, Department of Biology, University of Montreal for his scientific discussions. This research was funded by the *Ministère des Ressources Naturelles du Québec* Grant number 0316-240S to CG.

GENERAL DISCUSSION AND CONCLUSION

The main interest in ChfuGV studies derives from the possibility of using them as an alternative microbiological agent in the battle against the spruce budworm, a safe and cost effective alternative to protect the forests. To achieve this goal we require a comprehensive understanding of different aspects of the ChfuGV life cycle, starting from ecological features and continue all the way to their molecular characteristics. It may take us several years to reach that point but this is a journey that we can not afford not to start.

Existence of numerous important dissimilarities between the NPV genus and GV genus and the growing demand for safe and narrow range bio- insecticide agents are between the key motives for concentrating on learning more about the viruses belonging to this genus. No matter how many granuloviruses are characterized it does not close the way for other workers to investigate new GVs since each GV has its own insect target and any new finding on novel GVs are most welcomed by the workers with the interest to discover new bio-insecticide alternatives in order to be used against pest insects. The results of the present study can be judged as the groundwork for other researchers whose goal is either to further characterize this virus or to develop a bio-insecticide using ChfuGV as the active ingredient. The information existing in this work may be, also, considered very interesting for any worker in this field with an interest to develop genetically modified ChfuGV particles with more virulence and shorter infection cycle. The phylogenetic analyses presented here are not only important in molecular evolution and taxonomical viewpoint but also they might be considered as helpful tools to be used in any effort to create a new insecticide with ChfuGV as its active ingredient by

providing valuable comparative studies on ChfuGV and other GVs that are currently used as active ingredient in commercially available bioinsecticides.

The virus used in this study was an isolate of ChfuGV that was amplified, *in vivo*, in *Choristoneura fumiferana* larvae. As a starting point, we focused on characterization of ChfuGV's major structural proteins. To realize this, enveloped nucleocapsids were run on SDS-PAGE in order to isolate major structural proteins. Two major proteins (ODV-E56/ODVP-6E and p74) were isolated and sequenced. The amino acid sequences were then employed in the design of degenerate primers. Degenerate primers were used in order to determine the location of *odv-e56/odvp-6e* and p74 genes using ChfuGV genomic *Bam*HI restriction fragments as templates. Restriction fragments containing targeted genes were cloned in plasmid vectors and subsequently used as templates for nucleotide sequence determination.

An ORF coding for ODV-E56/ODVP-6E was located inside an 11 kb *Bam*HI restriction fragment. The gene *odv-e56/odvp-6e* is ORF number 14 on the ChfuGV genome (granulin is considered ORF 1). The ORF coding for ODV-E56/ODVP-6E is 1062 nt long coding for protein with an estimate molecular mass of 38.5 kDa. The size of baculoviral ChfuGV ODVP-6E/ODV-E56 protein is within the range known for other baculoviruses (the smallest at 351 aa is in PxGV and the longest at 379 aa is in CfMNPV). The localization and the orientation of the *odvp-6e/odv-e56* gene on the ChfuGV genome are similar to that described for CpGV. The 5' untranslated region (5'UTR) of ChfuGV *odvp-6e/odv-e56* analyses demonstrated that 5'UTR *odv-e56/odvp-*

6e in ChfuGV demonstrate a high degree of similarity to those of other baculoviral *odv-e56/odvp-6e*. A TATA box and a CCAAT element were identified, respectively, at 134 nt and 150 nt upstream from the start codon. A putative baculoviral late promoter motif (GTAAG) was located at 14 nt upstream to the first ATG. This motif has been reported in several baculoviral *odv-e56/odvp-6e* 5'UTRs. The late promoter motif TAAG is known to be a strong promoter of the late genes in *Baculoviridae*.

The presence of a putative late promoter motif upstream to ChfuGV *odv-e56/odvp-6e* gene indicated the fact that *odv-e56/odvp-6e* gene is a late gene. To substantiate this prediction we performed studies on expression of this gene. RT-PCR detected the transcripts 24 hr p.i. and immunoblots confirm the absence of the protein in the early stages of the infection. The first appearance of ChfuGV ODV-E56/ODVP-6E protein in infected larvae was detected at 24 hr p.i.

Analysis of ODVP-6E/ODV-E56 hydrophobicity plots suggested the presence of two hydrophobic domains within the C-terminal that were conserved in all baculoviral ODVP-6E/ODV-E56 proteins. These highly hydrophobic regions encompassing two membrane-spanning regions presented in ChfuGV and all other baculoviral ODVP-6E/ODV-E56 proteins. In addition to these two regions ChfuGV ODVP-6E/ODV-E56 protein shared almost similar hydrophobicity plots with others baculoviral ODVP-6E/ODV-E56 proteins.

A significant similarity was found between the sequence of ChfuGV ODVP-6E/ODV-E56 protein and other granuloviral and nucleopolyhedral ODVP-6E/ODV-E56 proteins. A high degree of conservation in amino acid sequences was observed within the two membrane-spanning regions of granuloviral ODVP-6E/ODV-E56 proteins. ChfuGV, CpGV, and PxGV shared a highly conserved stretch of 31 amino acids and a close look at this region in nucleopolyhedral ODVP-6E/ODV-E56 proteins showed that they maintained the same pattern although some residues were shifted. The existence of six conserved cysteine residues flanked by two hydrophobic membrane-spanning regions was observed by performing multiple alignments on deduced amino acid sequences of ChfuGV ODVP-6E/ODV-E56 and those of other baculoviruses. We propose that these highly conserved cysteine residues might be involved in the formation of disulfide bridges that maintain correct folding of baculoviral ODVP-6E/ODV-E56 proteins.

ChfuGV ODVP-6E/ODV-E56 protein seems to have no N-terminal signal sequence even when the ATG at the position 118 was considered as the presumed initiation codon (according to SignalP 3.0 and PSORT tools). Homology studies performed on ChfuGV ODVP-6E/ODV-E56 protein and other known baculoviral ODVP-6E/ODV-E56 proteins suggested that the ATG at the position 160 as the most probable start site for *odvp-6e/odv-e56*. Although both bioinformatics tools used in this study did not detect the existence of a signal peptide the presence of the putative motif LXL (alternate motif for AXA) immediately upstream and adjacent to the first methionine of ChfuGV ODVP-6E/ODV-E56 protein as well as the presence of a proline residue upstream to the LXL motif which usually form a turn prior to the signal peptide site (LXL) both strongly

suggest the possibility the existence of a signal peptide region upstream to the reported start site.

As mentioned previously, bioinformatics analysis of the ODVP-6E/ODV-E56 protein predicted the protein to be an envelope protein with two hydrophobic domains within the C-terminal. To confirm our computer-assisted prediction we performed an immunogold electron microscopy analysis on occluded ChfuGV virion using an anti-ODVP-6E/ODV-E56 monospecific antiserums. The results of immunogold electron microscopy confirmed our prediction by localizing the ChfuGV ODVP-6E/ODV-E56 protein on the envelope region of occluded virion. This was not surprising to us knowing that in at least three other baculoviruses, ODVP-6E/ODV-E56 proteins were localized on the ODV envelope (Braunagel *et al.*, 1996a; Theilmann *et al.*, 1996).

The region of ChfuGV genome in which we located *odvp-6e/odv-e56* was found to be composed of 14 ORFs including three ORFs unique to ChfuGV with no obvious homologues in other baculoviruses. The eleven ORFs with homologues to granuloviral ORFs included granulin, CfORF2, *pk-1*, *ie-1*, *odv-e18*, *p49* and *odvp-6e/odv-e56*. The conceptual products of seven major conserved ORFs, granulin, CfORF2, IE-1, ODV-E18, p49 and ODVP-6E/ODV-E56, were used in order to construct phylogenetic trees in this study. Our results classified the granuloviruses in two distinct groups, Group I containing (ChfuGV, CpGV, PhopGV and AoGV) and Group II containing (XcGV, PxGV and TnGV). ChfuGV conserved proteins are most closely related to those of CpGV, PhopGV and AoGV. Comparative studies carried out on gene arrangement inside

this region of genome showed that three GVs from group I maintain similar gene arrangements.

Other important ORFs have been also identified within the same *Bam*HI restriction fragment upstream to *odv-e56/odvp-6e*. An ORF with the potential of coding for a protein kinase homologue with the predicted molecular mass of 32 kDa was identified as ORF number 3 in this region of ChfuGV genome. At nucleotide level, an early element (CAGT) and a late consensus (TAAG) motif were detected within the 5'UTR of ChfuGV *pk-1* gene. Phylogenetic analysis of baculoviral protein kinase proteins separated GVs from NPVs, and showed that ChfuGV *pk-1* is most closely related to the PhopGV and CpGV.

An ORF with the potential of encoding a second ODV envelope associated protein was also identified in this region. ORF number 7 potentially codes for an ODV-E18 homologue. Baculoviral ODV-E18 is a conserved ODV-associated envelope protein. ORFs coding for this protein have been detected on the genome of several members of *Baculoviridae* family. ODV-E18 has been characterized as a late protein in AcMNPV associated with both the ODV envelope and capsid (Braunagel *et al.*, 1996). The presence of two baculoviral late promoter motifs (TAAG) within the 5'UTR of ChfuGV *odv-e18* gene indicates the possibility that this protein like other ODV-associated envelope proteins is expressed during the late phase of the viral infection cycle.

One of the most interesting ORFs detected inside the *Bam*HI restriction fragment upstream to *odv-e56/odvp-6e* was an ORF coding for a protein homologue to an essential baculoviral regulatory protein. The conceptual amino acid sequence of ORF 7 was found to be very similar to a highly conserved baculoviral IE-1 protein. Immediate early protein IE-1 is the principal transcriptional regulator known in several baculoviruses (Hayakawa *et al.*, 2000). Generally, transcription factors are identified as structurally complex proteins containing a range of functional components related with DNA binding, protein oligomerization, phosphorylation, activation, and other activities. Transcription factors can be classified into different groups such as zinc fingers, helix-turn-helix (HTH), helix-loop-helix (HLH), or basic leucine zippers. These proteins share a small domain associated with DNA binding or oligomerization (Lewin, 1997). Baculoviral IE-1 proteins are detected throughout infection (from immediate early to very late phases) and this suggests the significance of the presence of IE-1 for a productive infection (Okano *et al.* 1999; Choi and Guarino, 1995; Ribeiro *et al.*, 1994).

Comparative amino acid sequence studies on ChfVGV and other granuloviral IE-1 proteins revealed that the C-terminal regions of these proteins demonstrate a higher degree of preservation compared to the N-terminal regions. This observation was in agreement with the data from nucleopolyhedral IE-1 proteins (Olson *et al.* 2001; Wang *et al.*, 2001). Existing data on IE-1 proteins from NPVs suggest that the C-terminal regions of baculoviral IE-1 proteins complete a significant function in DNA binding while the N-terminal regions of IE-1 that is rich in acidic residues has been hypothesized to be an

acidic activation domain (Olson *et al.* 2001; Leisy and Rohrmann, 2000; Rodems *et al.*, 1997; Theilmann, and Stewart, 1991; Forsythe *et al.*, 1998).

Despite the lack of significant homology between amino acid sequences of IE-1 proteins of GVs and NPVs, the IE-1 protein from ChfugV shares a conserved domain with nucleopolyhedral IE-1 proteins. This conserved domain was originally reported within the C-terminal region of nucleopolyhedral IE-1 proteins. Each domain encloses two amphipathic alpha-helices preceded by a group of basic residues and forms a bHLH-like domain. The HLH protein family of transcriptional factors is characterized by highly conserved DNA binding, dimerization, and transactivation motifs, and contains very important elements that have been identified in different living organisms ranging from yeast to human (Atchley *et al.*, 2000). In spite of the high level of similarity inside the conserved domain, HLH proteins usually demonstrate significant sequence divergence outside this region. It has been shown in AcMNPV IE-1 that the existence of the hydrophobic residues within the HLH-like domain in C-terminus is essential for IE-1 oligomerization and contributes to IE-1 stability (Rodems *et al.*, 1997; Slack and Blissard, 1997). One can speculate that the hydrophobic surface of the amphipathic helices is directly involved in homophilic interaction between IE-1 monomers. The basic residues preceding the HLH-like domain of baculoviral IE-1 can be regarded as elements involved in DNA-binding activity. These residues are thought to interact via ionic interactions with the negatively charged phospho-sugar backbone during binding to DNA molecule. Besides the basic helix-loop-helix like (bHLH-like) domain, no noteworthy preserved domain could be detected among baculoviral IE-1 proteins. Determination of

the localization of IE-1 protein during different stages of infection will assist us to have a better understanding about the possible maturation process of this protein.

Another major structural protein characterized in the present study was p74. In other members of *Baculoviridae* family, p74 is known as a species specific (Wu *et al.*, 2003) major ODV-associated envelope protein which is exposed on the virion surface (Faulkner *et al.*, 1997) and has an essential role in the infectivity of the virus (Kuzio *et al.* 1989; Faulkner *et al.*, 1997). These observations suggest p74 protein is essential for initiation of infection and its presence is an indispensable element during the initiation step of viral infection cycle. The ORF coding for ChfuGV p74 protein is 1992 nt potentially encoding 663 amino acids, with an estimated molecular mass of 74.8 kDa. Two late promoter motifs (TAAG and GTAAG) were located respectively at 5 nt and 161 nt upstream of the first ATG of the p74 gene. As mentioned previously, the TAAG motif is a strong promoter of the late genes in baculoviruses. Similar to the case of ChfuGV ODVP-6E/ODV-E56 protein, the results on transcriptional and translational studies on ChfuGV p74 confirm that p74 is a late protein and its expression begins after 24 hr. p.i. in the infected larvae. This was in agreement with the existence of two late promoter motifs identified at nucleotide level in 5'UTR of ChfuGV p74 gene.

The high degree of conservation as well as the position preservation of cysteine, glycine and proline residues in all baculoviral p74 proteins may imply, also, that p74 protein plays a critical role in the life cycle of all baculoviruses and the removal of these conserved properties may eventually lead to the elimination of the virus in their

evolutionary pathway. The results on topology prediction presented on page 181 is produced by using TMHMM-VER.2 which predicts the localization of the N-terminal of p74 protein as an outside portion, but it is note worthy that if we consider the possibility of the existence of a peptide signal at N-terminal, this part of the protein can be considered as an inside portion of the protein.

Further studies on the nucleotide sequence of the region of ChfuGV genome which contains the gene that codes for p74 protein revealed the existence of a number of other important ORFs. We identified a total number of 15 ORFs and among those, 8 ORFs were found to be homologues to other known baculoviral ORFs such as Cf-i-*p* (*p* for partial), Cf-vi, Cf-vii, Cf-viii (*ubiquitin*), Cf-xi (*pp31*), Cf-xii (*lef11*), Cf-xiii (*sod*) and Cf-xv-*p* (p74). To date no specific function has been assigned to ORFs Cf-i, Cf-ii, Cf-iii, Cf-iv, Cf-v, Cf-vi, Cf-vii, Cf-ix and Cf-x.

One of identified ORFs codes for a conceptual protein similar to ubiquitin protein.

Ubiquitin is a globular protein found in all eukaryotic cells and whose sequence is extremely well conserved. Another important ORF detected in p74 region is an ORF with the potential to code a 161 amino acid polypeptide homologue to baculoviral copper/zinc superoxide dismutase (SOD). Genes coding SOD has been identified in several eukaryotic viruses such as poxviruses in which a gene encodes homologues of cellular cupro-zinc superoxide dismutases (Cu-Zn SOD) has been reported. Sequencing of the *sod*-like genes from various poxviruses demonstrated that the protein is highly conserved in all viruses sequenced (Almazan *et al.*, 2001; Cao *et al.*, 2002). The gene encode SOD

is almost present in all baculoviruses with only exception of *Epiphyas postvittana* nucleopolyhedrovirus (EpNPV) which lacks a homologue of the superoxide dismutase (*sod*) gene (Hyink *et al.*, 2002). Another ORF in this region of ChfuGV genome has the potential to code for a baculoviral protein known as 39k/pp31 protein, a nuclear matrix associate protein required for late expression during the baculoviral infection cycle (Gong *et al.*, 1998). Baculoviral 39k/pp31 protein is thought to play a role either at transcriptional level or at that of mRNA stabilization. (Lu and Miller, 1995). Another conserved baculoviral gene homologue present in this region of ChfuGV genome is a putative late expression factor (*lef-11*). Baculaviral *lef-11* gene has been known to be a gene of a great significance for baculoviral late gene expression and is crucial for viral DNA replication throughout the infection cycle to sustain optimal levels of transient expression from baculoviral late promoter (Lin *et al.*, 2001; Lin and Blissard 2002). Presence of an important ORF with homology to baculoviral fibrillin proteins in p74 region. This protein has a fibrillar structure (van Oers and Vlak, 1997). This gene has been located adjacent to p74 gene which encodes a highly conserved viral envelope protein. Predicted probability of coiled-coil structure formation in fibrillin and other similar proteins was performed using COILS program version 2.1, based on 21 residue window, as described by Lupus *et al.* (1991).

Granulin and polyhedrin are two major baculoviral proteins that are used to institute phylogenetic relation in baculaviridae family. The accuracy of this approach was originally questioned by Herniou and his associates (2001). In this study we employed combined data from different genes and gene arrangement. The results of phylogenetic

studies performed on several baculoviral conserved gene products demonstrated that GVs can be grouped in 2 distinct groups: Group I (ChfuGV, CpGV, PhopGV, ClGV and AoGV) and Group II (XcGV, PxGV, LoGV and TnGV). ChfuGV conserved proteins are most closely related to those of CpGV, PhopGV, AoGV and ClGV. ChfuGV along with CpGV, PhopGV, AoGV and ClGV share a high degree of gene order preservation within the studied genome region. In the two ChfuGV genome regions studied here GVs and NPVs demonstrated complete different patterns in their gene.

Several nucleopolyhedroviruses have been efficiently adapted to cell cultures, which has resulted in extensive progress in studies on their gene expression and function. On the other hand, efforts to adapt GVs to cell cultures have not been quite successful. The main cause of this limitation on GV studies. As a consequence, identification of GV genes sequences is principally essential in prediction of virus functions by the identification of homologues of NPV genes within GV genomes. During the past decade members of *Baculoviridae* were under extensive scientific investigations in order to shed more light on their genomic characterization. The fruit of these studies was complete or partial genome characterization of several nucleopolyhedrovirus and granulovirus. At the time that this study was initiated only three granuloviral complete or partial genome were existed in the Genbank. Today nine granuloviral complete or partial genomes (including ChfuGV) are recorded in the databank.

The findings in present study not only characterized two ChfuGV major ODV-associated proteins but also identified two previously unknown regions on ChfuGV genome. ORFs

identified in these two regions can be easily used to begin future studies on ChfuGV genes. The knowledge of the molecular aspect of ChfuGV is extremely limited, and prior to this study just two papers had been published on this subject matter. In fact, the present study can be considered the first extensive effort in molecular characterization of ChfuGV. When it comes to ChfuGV molecular characterization we are simply beginning our scientific journey while this study may only be able to put in place some of the first pieces of a colossal puzzle, however, it opens several doors for future workers who will ultimately complete the puzzle.

REFERENCES

- ACOSTA-SERRANO, A., E. VASSELLA, M. LINIGER, C. KUNZ RENGGLI, R. BRUN, I. RODITI and P.T. ENGLUND. 2001. The surface coat of procyclic *Trypanosoma brucei*: programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proc. Natl. Acad. Sci. U S A*; 98, 1513-1518.
- ADAMS, J. R. and J. T. MCCLINTOCK. 1991. Baculoviridae, Nuclear polyhedrosis viruses of insects. In *Atlas of Invertebrate Viruses*, pp. 87- 204. Edited by J. R. Adams and J. R. Bonami. Boca Raton. FL: CRC Press.
- ADAMS, J. R., R. H. GOODWIN, and T. A. WICOX. 1977. Electron microscopic investigations on invasion and replication of insect baculoviruses *in vivo* and *in vitro*. *Biol. Cell.*, 28:261-268.
- AFONSO, C.L., E.R. TULMAN, Z. LU, ,C.A. BALINSKY, B.A. MOSER, J.J. BECNEL, D.L. ROCK and G.F. KUTISH. 2001. Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. *J. Virol.* 75, 11157-11165.
- AHRENS, C. H., and G. F. ROHRMANN. 1995a. Identification of essential *trans*-acting regions required for DNA replication of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus; *lef-1* is an essential replication gene. *Virology* 207, 417-428.
- AHRENS, C. H., and G. F. ROHRMANN. 1995b. Replication of *Orgyia pseudotsugata* baculovirus DNA: *ie-1* and *lef-2* are essential and *ie-2* and *Op-iap* are simulatory genes. *Virology* 209, 572-576.
- AHRENS, C.H., R.L. RUSSELL, C.J. FUNK, J.T. EVANS, S.H. HARWOOD, and G.F. ROHRMANN. 1997. The sequence of the *Orgyia pseudotsugata* multinucleocapsid nuclear polyhedrosis virus genome *Virology* 229, 381-399.
- ALMAZAN, F., D.C. TSCHARKE and G.L. SMITH. 2001. The vaccinia virus superoxide dismutase-like protein (A45R) is a virion component that is nonessential for virus replication. *J. Virol.* 75, 7018-7029.
- ALTSHUL, S. F., W. GISH, W. MILLER, E. W. MEYERS and D.J. LIPMAN. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- ARCHIBALD, J.M., D. LONGET, J. PAWLOWSKI and P.J. KEELING. 2003. A novel polyubiquitin structure in cercozoa and foraminifera: evidence for a new eukaryotic supergroup. *Mol Biol Evol.* 20, 62-66.
- ATCHLEY, W. R., K. R. WOLLENBERG, W.M. FITCH, W. TERHALLE, A.W. DRESS. 2000. Correlations among amino acid sites in bHLH protein domains: an information theoretic analysis. *Mol. Biol. Evol.* 17, 164-178.

- ATLANTE, A., A. BOBBA, P. CALISSANO, S. PASSARELLA and E. MARRA. 2003. The apoptosis/necrosis transition in cerebellar granule cells depends on the mutual relationship of the antioxidant and the proteolytic systems which regulate ROS production and cytochrome C release en route to death. *J. Neurochem.* 84, 960-971.
- AYRES, M.D., S.C. HOWARD, J. KUZIO, M. LOPEZ-FERBER, and R.D. POSSEE. 1994. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* 202, 586-605.
- BAH, A., J. BERGERON, M. ARELLA, C.J. LUCAROTTI, and C. GUERTIN. 1997. Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. *Arch. Virol.* 142, 1577-1584.
- BAIROCH, A., P. BUCHER and K. HOFMANN. 1997. The PROSITE database, its status in 1997. *Nucleic Acids Res.* 25, 217-221.
- BANNISTER, J., W. BANNISTER and E. WOOD. 1971. Bovine Erythrocyte Cupro-zinc Protein. I. Isolation and General Characterization, *Eur. J. Biochem.* 18, 178.
- BARLEY, S. K., A. MILLER, K. A. HARRAP, and D. C. KELLY. 1982. Structure of the Baculovirus nucleocapsid. *Virology* 120, 433-438.
- BEATON, C. D., and B. K. FILSHIE. 1976. Comparative ultrastructural studies of insect granulosis and nuclear polyhedrosis viruses. *J. Gen. Virol.* 31, 151-161.
- BEEGLE, C. C., L. P. PEDIGO, F. L. POSTON, and J. D. STONE. 1973. Field effectiveness of the granulosis virus of the green clover worm as compared with *Bacillus thuringiensis* and selected chemical insecticides on soy bean. *J. Econ. Entomol.* 66, 1137-1138.
- BENZ, G., and R. WAGER. 1971. Autoradiographic studies on nucleic acid metabolism in granulosis-infected fat body of larvae of *Carpocapsa*. *J. Invertebr. Pathol.* 18, 70-80.
- BENZ, G. A., 1986. *The Biology of Baculoviruses*, vol. 1. pp. 1-35. eds Granados, R. R. and Federici, B. A.), CRC Press, Boca Raton.
- BERGERON, Y., H. MORIN, A. LEDUC and C. JOYAL. 1995. Balsam fir mortality following the last spruce budworm outbreak in northwestern quebec. *Can. J. For. Res.* 25, 1375-1384.
- BERGOIN, M., and S. DALES. 1971. Comparative observations on poxviruses of invertebrates and vertebrates, p. 169-205. In K. Maramorosch and E. Kurstak (ed.), *Comparative virology*, Academic press, Inc., New York.

BJORNSON, R. M. and G. F. ROHRMANN. 1992a Nucleotide sequence of p39-capsid gene region of the *Lymantria dispar* nuclear polyhedrosis virus. *J. Gen. Virol.* 73, 1505-1508.

BJORNSON R. M., and G. F. ROHRMANN. 1992b. Nucleotide Sequence of the polyhedron envelope protein gene region of the *Lymantria dispar* nuclear polyhedrosis virus. *J. Gen. Virol.* 73, 1499- 1504.

BLISSARD G. W. 1996. Baculovirus-insect cell interactions. 1996. *Cytotechnology* 20, 73-93.

BLISSARD, G., B. BLACK, N. CROOK, B. A. KEDDIE, R. POSSEE, G. ROHRMANN, D. THEILMANN, and L. E. VOLKMAN. 2000. Baculoviridae: taxonomic structure and properties of the family, p. 195-202. In M.H.V. van Regenmortel et al. (ed.), Virus taxonomy. The classification and nomenclature of viruses. The seventh report of the International Committee for the Taxonomy of Viruses. Academic Press, San Diego, Calif

BLISSARD, G. W., and G. F. ROHRMANN. 1989. Location, sequence, transcriptional mapping, and temporal expression of gp64 envelope glycoprotein gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 170, 537-555.

BLISSARD, G. W., and G. F. ROHRMANN. 1990. Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.*; 35, 127-55.

BLISSARD, G. W., R. L. QUANT-RUSSELL, G. F. ROHRMANN and G. S. BEAUDREAU . 1989. Nucleotide sequence, transcriptional mapping, and temporal expression of the gene encoding p39, a major structural protein of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*. *Virology*. 168:354-362.

BLISSARD, G. W., P. H. KOGAN, R. WEI and G. F. ROHRMANN. 1992. A synthetic early promoter from a baculovirus: roles of the TATA box and conserved start site CAGT sequence in basal levels of transcription. *Virology* 190, 783-93.

BRAUNAGEL, S. C., H. HE, P. RAMAMURTHY and M. D. SUMMERS. 1996a. 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., D.M. ELTON, H. MA and M.D. SUMMERS. 1996b. Identification and analysis of an *Autographa californica* nuclear polyhedrosis virus structural protein of the occlusion-derived virus envelope: ODV-E56. *Virology* 217, 97-110.

- BRAUNAGEL, S. C., and M.D. SUMMERS. 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.
- BULACH, D. M., C. A. KUMAR, A. ZAIA, B. LIANG, D. E. TRIBE. 1999. Group II nucleopolyhedrovirus subgroups revealed by phylogenetic analysis of polyhedrin and DNA polymerase gene sequences. *J. Invertebr. Pathol.* 73, 59-73.
- BUTTKE, T. M. and P. A. SANDSTROM. 1994. Oxidative stress as a mediator of apoptosis. *Immunol. Today* 15, 7-10.
- BUTIKOFER, P., E. VASSELLA, M. BOSCHUNG, C.K. RENGGLI, R. BRUN, T.W. PEARSON and I. RODITI. 2002. Glycosylphosphatidylinositol-anchored surface molecules of *Trypanosoma congolense* insect forms are developmentally regulated in the tsetse fly. *Mol. Biochem. Parasitol.* 119, 7-16.
- CAO, J.X., M.L. TEOH, M. MOON, G. MCFADDEN and D.H. EVANS. 2002. Leporipoxvirus Cu-Zn superoxide dismutase homologs inhibit cellular superoxide dismutase, but are not essential for virus replication or virulence. *Virology* 296, 125-135.
- CHEN, X., W.-J. ZHANG, J. WONG, G. CHUN, A. LU, B.F. MCCUTCHEN, J.K. PRESNAIL, R. HERRMANN, M. DOLAN, S. TINGEY, Z.-H. HU, and J.M. VLAK. 2002. Comparative analysis of the complete genome sequences of *Helicoverpa zea* and *Helicoverpa armigera* single-nucleocapsid nucleopolyhedroviruses. *J. Gen. Virol.* 83, 673-684.
- CHENG, C.H., S.M. LIU, T.Y. CHOW, Y.Y. HSIAO, D.P. WANG, J.J. HUANG, and H.H. CHEN. 2002. Analysis of the Complete Genome Sequence of the Hz-1 Virus Suggests that it is related to members of the Baculoviridae. *J. Virol.* 76, 9024-9034.
- CHERBAS, L. and P. CHERBAS. 1993. The arthropod initiator: the capsite consensus plays an important role in transcription. *Insect Biochem. Mol. Biol.* 23, 81-90.
- CHOI, J. and L. A. GUARINO. 1995. A temperature-sensitive IE1 protein of *Autographa californica* nuclear polyhedrosis virus has altered transactivation and DNA binding activities. *Virology* 209, 90-98.
- CLEM, R. J., M. FECHHEIMMER, and L. K. MILLER. 1991. Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* 254, 1388-1390.

CROOK, N. E. 1991. Baculoviridae: subgroup B: comparative aspects of granulosis viruses. In *Viruses of Invertebrates*. pp73-110. Edited by E. Kurstak. Newyork: Marcel Dekker.

DARVEY, J. 1989. Sorting out the secretory pathway. *Bioassays* 11, 185-187.

DAVID, W. A. L. 1969. The effect of ultraviolet radiation of known wavelengths on a granulosis virus of *Pieris brassicae*. *J. Invertebr. Pathol.* 14, 336-342.

DAVID, W. A. L., and B. O. C. GARDINER. 1967. The effect of heat, cold, and prolonged storage on a granulovirus of *Pieris brassicae*. *J. Invertebr. Pathol.* 8, 180-183.

DAVID, W. A. L., B. O. C. GARDINER, and M. WOOLNER. 1968. The effects of sunlight on a purified granulosis virus of *Pieris brassicae* applied to cabbage leaves. *J. Invertebr. Pathol.* 11, 496-501.

DOUGHERTY, E. M., K. P. GUTHRIE, and M. SHAPIRO. 1996. Optical brighteners provide Baculovirus activity enhancement and UV radiation protection. *Biological Control* 7, 71-74.

FAKTOR, O., M. TOISTER-ACHITUV, O. NAHUM and B. KAMENSKY. 1997. The p10 gene of *Spodoptera littoralis* nucleopolyhedrovirus: nucleotide sequence, transcriptional analysis and unique gene organization in the p10 locus. *J. Gen. Virol.* 78, 2119-2128.

FAULKNER, P., J. KUZIO, G. V. WILLIAMS and J. A. WILSON. 1997. Analysis of a PDV envelope protein of *Autographa californica* nucleopolyhedrovirus required for occlusion body infectivity in vivo. *J. Gen. Virol.* 78, 3091-3100.

FEDERICI, B. A. 1997. Baculoviral pathogenesis. In *The Baculoviruses*, pp 33-59. Edited by L. K. Miller. New York: Plenum Press.

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

FIELDING, B.C. and S. DAVISON. 2000. Identification and characterization of the *Trichoplusia ni* single capsid nuclear polyhedrosis virus p10 gene. *Virus Genes* 20, 189-192.

FIELDING, B.C. S. KHAN, W. WANG, C. KRUGER, R. ABRAHAMS and S.P. DAVISON. 2002. The genetic organization of a 2,966 basepair DNA fragment of a single capsid nucleopolyhedrovirus isolated from *Trichoplusia ni*. *Virus Genes* 25, 35-43.

- FORSYTHE, I. J., C. E. SHIPPAM, L. G. WILLIS, S. STEWART, T. GRIGLIATTI, and D. A. THEILMANN. 1998. Characterization of the acidic domain of the IE1 regulatory protein from *Orgyia pseudotsugata* multicapsid nucleopolyhedrovirus. *Virology* 252, 65-81.
- FORTÉ, A. J., C. GUERTIN and J. CABANA. 1999. Pathogenicity of a granulovirus towards *Choristoneura fumiferana*. *Canadian Entomologist* 131, 725-727.
- FUNK, C.J. and R.A. CONSIGLI. 1993. Temporal expression and immunogold localization of *Plodia interpunctella* granulosis virus structural proteins. *Virus Res.* 28, 57-66.
- Garnier, J., J-F. Gibrat and B. Robson. 1996. GOR secondary structure prediction method version IV *Methods in Enzymology* R.F. Doolittle Ed., vol 266, 540-553.
- GOMBERT, A. F., G. W. BLISSARD and, G. F. ROHRMANN. 1989. Characterization of the genetic organization of the Hind III M region of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata* reveals major differences among baculoviruses. *J. Gen. Virol.* 70, 1815-1828.
- GOMI, S., MAJIMA, K. and S. MAEDA. 1999. Sequence analysis of the genome of *Bombyx mori* nucleopolyhedrovirus *J. Gen. Virol.* 80, 1323-1337.
- GONG, M. and L. A. GUARINO. 1994. Expression of 39k promoter of *Autographa californica* nuclear polyhedrosis virus is increased by the apoptotic suppressor p35. *Virology* 204, 38-44.
- GONG, M., J. JIN and L.A. GUARINO. 1998. Mapping of ORF121, a factor that activates baculovirus early gene expression. *Virology* 244, 495-503.
- GREGORY, E., S. GOSCIN and I. FRIDOVICH. 1974. Superoxide Dismutase and Oxygen Toxicity in a Eukaryote, *J. Appl. Bacteriol.* 117, 456-462.
- GROSS, C. H., R. L. Q. RUSSELL, and G. F. ROHRMANN. 1994. The *Orgyia pseudotsugata* baculovirus p10 and polyhedron envelope protein genes: analysis of their relative expression levels and role in polyhedron structure. *J. Gen. Virol.* 75, 1115-1123.
- GUARINO, L. A., and SUMMERS, M. D. 1986. Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene. *J. Virol.* 57, 563-571.
- GUARINO, L. A. and W. DONG. 1994. Functional dissection of the *Autographa californica* nuclear polyhedrosis virus enhancer element hr5. *Virology* 200, 328-335.

- HANSEN, J.E., O. LUND, N. TOLSTRUP, A.A. GOOLEY, K.L. WILLIAMS and S. BRUNAK. 1998. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj. J.* 15, 115-130.
- HARLOW, E., and D. LANE. 1988. Antibodies: A laboratory manual. Cold Spring Harbor Press. N. Y.
- HARRAP, K. A., C. C. PAYNE, and J. S. ROBERTSON. 1977. The properties of three baculoviruses from closely related hosts. *Virology* 79, 14-31.
- HASHIMOTO, Y., T. HAYAKAWA, Y. UENO, T. FUJITA, Y. SANO, and T. MATSUMOTO. 2000. Sequence analysis of the *Plutella xylostella* granulovirus genome. *Virology* 275, 358-372.
- HAYAKAWA, T., R. KO, K. OKANO, S.I. SEONG, C. GOTO, and S. MAEDA. 1999. Sequence analysis of the *Xestia c-nigrum* granulovirus genome. *Virology* 262, 277-297.
- HAYAKAWA, T., G. F. ROHRMANN and Y. HASHIMOTO. 2000. Patterns of genome organization and content in lepidopteran baculoviruses. *Virology* 278, 1-12.
- HERNIOU, E. A., T. LUQUE, X. CHEN, J.M. VLAK, D. WINSTANLEY, J.S. CORY and D.R. O'REILLY. 2001. Use of whole genome sequence data to infer baculovirus phylogeny. *J. Virol.* 75, 8117-8126.
- HESS, R. T. and L. A. FALCON. 1987. Temporal events in the invasion of the codling moth, *Cydia pomonella*, by a granulosus virus: an electron microscope study. *J. Invertebr. Pathol.* 50, 85-92.
- HICKE, L. 2001. A new ticket for entry into budding vesicles: ubiquitin. *Cell* 106, 527-530.
- HILL, J.E., J. KUZIO, J.A. WILSON, E.A. MACKINNON and P. FAULKNER. 1993. Nucleotide sequence of the p74 gene of a baculovirus pathogenic to the spruce budworm, *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus. *Biochim. Biophys. Acta* 20, 1172, 187-189.
- HILL-PERKINS, M. S., and R. D. POSSEE. 1990. A baculovirus expression vector derived from the basic protein promoter of *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* 71, 971-976.

- HOCKENBERY, D.M., Z.N. OLTVAI, X.M. YIN, C.L. MILLIMAN and S.J. KORSMEYER. 1993. Bcl-2 functions in an antioxidant pathway to prevent apoptosis, *Cell* 75, 241-251.
- HOFMANN, K. and, W. STOFFEL .1993. TMbase - A database of membrane spanning proteins segments *Biol. Chem. Hoppe-Seyler* 374,166.
- HONG, T., S. C. BRAUNAGEL, and M. D. SUMMERS. 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.
- HORTON, H.M., and J.P. BURAND. 1993. Saturable attachment sites for polyhedron-derived baculovirus on insect cells and evidence for entry via direct membrane fusion. *J. Virol.* 67, 1860-1868.
- HU, Z.H., B.M. ARIF, F. JIN, J.W. MARTENS, X.W. CHEN, J.S. SUN, D. ZUIDEMA, R.W. GOLDBACH and J.M. VLAK. 1998. Distinct gene arrangement in the *Buzura suppressaria* single-nucleocapsid nucleopolyhedrovirus genome. *J. Gen. Virol.* 79, 2841-2851.
- HU, N.T., Y.F. LU, Y. HASHIMOTO, S. MAEDA and R.F. HOU. 1994. The p10 gene of natural isolates of *Bombyx mori* nuclear polyhedrosis virus encodes a truncated protein with an M(r) of 7700. *J. Gen. Virol.* 75, 2085-2088.
- HUH, N. E. and R. F. WEAVER. 1990. Identifying the RNA polymerases that synthesize specific transcripts of the *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* 71, 195-201.
- HUNTER, D. K., D. F. HOFFMAN and S. J. COLLIER. 1975. Observations on a granulosis virus of the potato tuberworm, *Phthorimaea operculella*. *J. Invertebr. Pathol.* 26, 397-400.
- HYINK, O., R.A. DELLOW, M.J. OLSEN, K.M. CARADOC-DAVIES, K. DRAKE, E.A. HERNIOU, J.S. CORY, D.R. O'REILLY and V.K. WARD. 2002. Whole genome analysis of the *Epiphyas postvittana* nucleopolyhedrovirus. *J. Gen. Virol.* 83, 957-971.
- IJKEL, W.F., E.A. VAN STRIEN, J.G. HELDENS, R. BROER, D. ZUIDEMA, W. GOLDBACH, R. and J.M. VLAK. 1999. Sequence and organization of the *Spodoptera exigua* multicapsid nucleopolyhedrovirus genome. *J. Gen. Virol.* 80, 3289-3304.
- JAKUES, R. P. 1973. Tests on microbial and chemical insecticides for control of *Tricoplusia ni* (Lepidoptera: Noctuidae) and *Pieris rapae* (Lepidoptera: Pieridae) on cabbage. *Can. Entomol.* 105, 21-27.

- KAWANISHI, C. Y., M. D. SUMMER, D. B. STOLTZ and H. J. ARNOTT. 1972. Entry of an insect virus in vivo by fusion of viral envelopes and microvillus membrane. *J. Invertebr. Pathol.* 20, 104-108.
- KEDDIE, A. and L. E. VOLKMAN. 1985. Infectivity difference between the two phenotypes of *Autographa californica* nuclear polyhedrosis virus: Importance of the gp64 envelope glycoprotein. *J. Gen. Virol.* 66, 1195-1200.
- KEDDIE, B.A., G.W. APONTE and L.E. VOLKMAN. 1989. The pathway of infection of *Autographa californica* nuclear polyhedrosis virus in an insect host. *Science* 243, 1728-1730.
- KEELE, B., J. MCCORD and I. FRIDOVICH. 1971. Further Characterization of Bovine Superoxide Dismutase and the Isolation from Bovine Heart. *J. Biol. Chem.* 246, 2875.
- KELLY, D. C. and X. WANG. 1981. The infectivity of nuclear polyhedrosis virus DNA. *Ann. Virol.* 132^E, 247-259.
- KELSEY, J. M. 1957. Virus sprays for control of *Pieris rapae* L. N. Z. J. *Sci. Technol. Sect. A* 38, 644-646.
- KOOL, M., C. AHRENS, R. W. GOLDBACH, G. F. ROHRMANN and J. M. VLAK. 1994a. Identification of genes involved in DNA replication of the *Autographa californica* baculovirus. *Proc. Nat. Acad. Sci. U.S.A.* 91, 11212-11216.
- KOOL, M., J. T. M. VOETEN and J. M. VLAK. 1994b. Functional mapping of regions of the *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* 75, 3345-3352.
- KOOL, M., J. T. M. VOETEN, R. W. GOLDBACH, J. TRAMPER and J. M. VLAK. 1993. Identification of seven putative origins of *Autographa californica* nuclear polyhedrosis virus DNA replication. *J. Gen. Virol.* 74, 2661-2668.
- KORNBERG, A. and T. A. BAKER. 1992. *DNA replication*, 2nd ed. New York: W. H. Freeman.
- KOVACS, G. R., J. CHOI, L. A. GUARINO and M. D. SUMMERS. 1992. Functional dissection of the *Autographa californica* nuclear polyhedrosis virus immediate-early 1 transcriptional regulatory protein. *J. Virol.* 66, 7429-7437.
- KOZAK, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulate translation by eukaryotic ribosomes. *Cell* 44, 282-292.

KOZLOV, E. A., T. L. LEVITINA, N. M. GUSAK and S. B. SEREBRYANI. 1981. Comparison of the amino acid sequences of inclusion body proteins of nuclear polyhedrosis viruses of *Bombyx mori*, *Porthesia dispar*, and *Galleria mellonema*. *Bioorganic. Khimia* 7, 1008-1015.

KREMER, A. and D.KNEBEL-MORSDORF.1998. The early baculovirus he65 promoter: On the mechanism of transcriptional activation by IE1. *Virology* 249, 336-351.

KUZIO, J., M.N. PEARSON, S.H. HARWOOD, C.J. FUNK, J.T. EVANS, J.M. SLAVICEK, and G.F. ROHRMANN. 1999. Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria dispar*. *Virology* 253:17-34.

KUZIO J., R. JAQUES and P. FAULKNER.1989. Identification of p74, a gene essential for virulence of baculovirus occlusion bodies. *Virology* 173, 759-763.

KYTE, A J. and R.F. T. DOOLITTLE. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105-132.

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

LANGE, M. and J.A. JEHLE. 2003. The genome of the *Cryptophlebia leucotreta* granulovirus. *Virology* 317, 220-236.

LEE, H.H., S. PARK and S.C. CHA. 1996a. Nucleotide sequence analyses of p10 gene and its promoter of *Hyphantria cunea* nuclear polyhedrosis virus strain HL-2. *Taehan Paerosu Hakhoe Chi* 26, 131-137.

LEE, S.Y., A. POLOUMIENKO, S. BELFRY, X. QU, W. CHEN, N. MACAFEE, B. MORIN, C. LUCAROTTI and M. KRAUSE. 1996b. A common pathway for p10 and calyx proteins in progressive stages of polyhedron envelope assembly in AcMNPV-infected *Spodoptera frugiperda* larvae. *Arch. Virol.* 141, 1247-1258.

LEISY, D.J., G.F. ROHRMANN, M. NESSON and G.S. BEAUDREAU.1986. Nucleotide sequencing and transcriptional mapping of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus p10 gene. *Virology* 153, 157-167.

LEISY, D. J. and G.F. ROHRMANN. 2000. The *Autographa californica* nucleopolyhedrovirus IE-1 protein complex has two modes of specific DNA binding. *Virology* 274, 196-202.

LEWIN, B. 1997. Genes VI. Oxford University Press, New York.

LI , S., M. ERLANDSON, D. MOODY and C. GILLOTT. 1997. A physical map of the *Mamestra configurata* nucleopolyhedrovirus genome and sequence analysis of the polyhedrin gene. *J. Gen. Virol.* 78, 265-271.

- LI, L., S. H. HARWOOD, and G. F. ROHRMANN. 1999. Identification of additional genes that influence baculovirus late gene expression. *Virology* 255, 9-19.
- LI, Q., C. DONLY, L. LI, L.G. WILLIS, D.A. THEILMANN, and M. ERLANDSON. 2002. Sequence and organization of the *Mamestra configurata* nucleopolyhedrovirus genome. *Virology* 294, 106-121.
- LIN, G. and G. W. BLISSARD. 2002. Analysis of an *Autographa californica* Nucleopolyhedrovirus *lef-11* Knockout: LEF-11 Is Essential for Viral DNA Replication. *J. Virol.* 76, 2770-2779.
- LIN, G., J.M. SLACK and G.W. BLISSARD. 2001. Expression and localization of LEF-11 in *Autographa californica* nucleopolyhedrovirus-infected Sf9 cells. *J. Gen. Virol.* 82, 2289-2294.
- LU, A. and E. B. CARSTENS. 1992. Nucleotide sequence and transcriptional analysis of the p80 gene of *Autographa californica* nuclear polyhedrosis virus: a homologue of the *Orgyia pseudotsugata* nuclear polyhedrosis virus capsid-associated gene. *Virology* 190, 201-209.
- LU, A., and L. K. MILLER. 1995. The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *J. Virol.* 69, 975-982.
- LUPAS, A., M. VAN DYKE and J. STOCK. 1991. Predicting Coiled Coils from Protein Sequences, *Science* 252, 1162-1164.
- LUQUE, T., R. FINCH, N. CROOK, D.R. O'REILLY, and D. WINSTANLEY. 2001. The complete sequence of the *Cydia pomonella* granulovirus genome. *J. Gen. Virol.* 82, 2531-2547.
- MASSARI, M. E. and MURRE, C.2000. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell Biol.* 20, 429-440.
- MCLACHLIN, J. R. and L. K. MILLER. 1994. Identification and characterization of *vlf-1*, a baculovirus gene involved in very late gene expression. *J. Virol.* 68, 7746-7756.
- MCINTOSH, A. H. 1975. In vitro specificity and mechanism of infection. pp. 63-72 in *Baculoviruses for insect pest control: Safety Considerations*. Edited by Summers, M. D., Engler, R., Falcon L.A., Vail P. American Society for Microbiology. Washington D.C. ISBN 0-914826-07-7.
- MILLER, L. K. Insect viruses. 2001. pp. 533-585. in B. N. Fields *et al.* (eds.) *Fields Virology* (vol.1., 4th edition). Lippincott-Raven publishers, New York.

- MITCHELL, A. C. MITTER, J.C. REGIER. 2000. More taxa or more characters revisited: combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). *Syst. Biol.* 49, 202-224.
- MOLLER, S., M.D.R. CRONING and R. APWEILER. 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646-653.
- MONSMA, S. A., A.G. OOMENS and G.W. BLISSARD. 1996. The GP64 envelope fusion protein is an essential baculovirus protein required for cell-to-cell transmission of infection. *J Virol.* 70, 4607-4616.
- MORIN, H., D. LAPRISE and Y. BERGERON. 1993. Chronology of spruce budworm outbreaks in the lake duparquet region, abitibi, Québec. *Can. J. For. Res.* 23, 1497-1506.
- MOSCARDI, F. 1999. Assessment of the application of baculoviruses for control of lepidoptera. *Annu. Rev. Entomology* 44, 227-289.
- MULLER, R., M. N. PEARSON, R. L. Q. RUSSELL, and G. F. ROHRMANN. 1990. A capsid-associated protein of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*: genetic location, sequence, transcriptional mapping, and immunocytochemical characterization. *Virology* 176, 133-144.
- NISSEN, M. S. and P.D. FRIESEN. 1989. Molecular analysis of the transcriptional regulatory region of an early baculovirus gene. *J. Virol.* 63, 493-503.
- OELLIG, C., B. HAPP, T. MULLER and W. DOERFLER. 1987. Overlapping sets of viral RNAs reflect the array of polypeptides in the EcoRI J and N fragments (map position 81.2 to 85.0) of the *Autographa californica* nuclear polyhedrose, virus *J. Virol.* 61, 3048-3057.
- OLSON, V. A., J. A. WETTER and P. D. FRIESEN. 2001. Oligomerization mediated by a helix-loop-helix-like domain of baculovirus IE1 is required for early promoter transactivation. *J. Virol.* 75, 6042-6451.
- OKANO, K., V. S. MIKHAILOV and S. MAEDA. 1999. Colocalization of baculovirus IE-1 and two DNA-binding proteins, DBP and LEF-3, to viral replication factories. *J. Virol.* 73, 110-119.
- OLSON, V. A., J. A. WETTER and P. D. FRIESEN. 2001. Oligomerization mediated by a helix-loop-helix-like domain of baculovirus IE1 is required for early promoter transactivation. *J. Virol.* 75, 6042-6451.

- OOMENS, A.G. and G.W. BLISSARD. 1999. Requirement for GP64 to drive efficient budding of *Autographa californica* multicapsid nucleopolyhedrovirus. *Virology* 254, 297-314.
- PANG, Y., J. YU, L. WANG, X. HU, W. BAO, G. LI, C. CHEN, H. HAN, S. HU, and H. YANG. 2001. Sequence analysis of the *Spodoptera litura* multicapsid nucleopolyhedrovirus genome. *Virology* 287, 391-404.
- PAILLOT, A. 1926. Sur une nouvelle maladie au noyau au grasserie des chentilles de *P. brassicae* et un nouveau groupe de microorganismes parasites. *Compt. Rend.* 182, 180-182.
- PASSARELLI, A. L. and L. K. MILLER. 1993a. Identification and characterization of *lef-1*, a baculovirus gene involved in late and very late gene expression. *J. Virol.* 67, 3481-3488.
- PASSARELLI, A. L. and L. K. MILLER. 1993b. Three baculovirus genes involved in late and very late gene expression: *ie-1*, *ie-n*, and *lef-2*. *J. Virol.* 67, 2149-2158.
- PEARSON, W. R. and D. J. LIPMAN. 1988. Improved Tools for Biological Sequence Comparison. *PNAS* 85, 2444- 2448.
- PEARSON, M. N., R. L. Q. RUSSELL, G. F. ROHRMANN, and G. S. BEAUDREAU. 1987. P39, a major baculovirus protein; Immunocytochemical characterization and genetic location. *Virology* 167, 407-413.
- PEI, Z., G. RESKE, Q. HUANG, B.D. HAMMOCK, Y. QI, N. CHEJANOVSKY. 2002. Characterization of the apoptosis suppressor protein P49 from the *Spodoptera littoralis* nucleopolyhedrovirus. *J. Biol. Chem.* 277, 48677-48684.
- PETKAU, A., W. CHELACK, S. PLESKACH, B. MEEKER and C. BRADY. 1975. Radioprotection of Mice by Superoxide Dismutase, *Biochem. Biophys. Res. Commun* 65, 886.
- PICKART, C.M. 2004. Back to the future with ubiquitin. *Cell* 116, 181-190.
- PICKART, C.M. 2001. Mechanisms underlying ubiquitination. *Ann. Rev. Biochem.* 70, 503-533.
- PINNOCK, D. E. and R. T. HESS. 1978. Morphological variation, in the cytopathology associated with granulosis virus in the fruit-tree leaf roller, *Archips argyrospilla*. *J. Ultrastructure. Res.* 18, 143-157.
- PRINGLE, C.R. 1999. Virus Taxonomy. *Arch. Virol.* 144, 421-429.

- PULLEN, S. S. and P.D. FRIESEN. 1995a. The CAGT motif functions as an initiator element during early transcription of the baculovirus transregulator IE-1. *J. Virol.* 69, 3575-83.
- PULLEN, S. S. and P.D. FRIESEN. 1995b. Early transcription of the *ie-1* transregulator gene of *Autographa californica* nuclear polyhedrosis virus is regulated by DNA sequences within its 5' noncoding leader region. *J. Virol.* 69, 156-165.
- QUANT, R. L., M. N. PEARSON, G. F. ROHRMANN and G. S. BEAUDEAU. 1984. Production of polyhedrin monoclonal antibodies for distinguishing two *Orgyia pseudotsugata* baculoviruses. *Appl. Envir. Microbio.* 48, 732-736.
- RAGHOW, R. and T. D. C. GRACE. 1974. Studies on a nuclear polyhedrosis virus in *Bombyx mori* cells in vitro. I. Manipulation kinetics and ultrastructural studies. *J. Ultrastruct. Res.* 47, 384-399.
- RANGARAJAN, D., T.I. HARVEY and J.D. BARRY. 2000. Characterization of the loci encoding the glutamic acid and alanine rich protein of *Trypanosoma congolense*. *Mol. Biochem. Parasitol.* 105, 281-290.
- RASHIDAN, K. K., N. NASSOURY, A. MERZOUKI and C. GUERTIN. 2002a. Identification and Characterization of a Putative Baculoviral Transcriptional Factor (IE-1) from *Choristoneura fumiferana* granulovirus. *J. Biochem. Mol. Biol.* 35, 553-561.
- RASHIDAN, K. K., N. NASSOURY, P. N. GIANNOPOULOS and C. GUERTIN. 2002b. Identification and characterization of a conserved baculoviral structural protein ODVP-6E/ODV-E56 from *Choristoneura fumiferana* granulovirus. *J. Biochem. Mol. Biol.* 35, 595-603.
- RASHIDAN, K.K., N. NASSOURY, S. TAZI, P.N. GIANNOPOULOS, and C. GUERTIN. 2003. *Choristoneura fumiferana* Granulovirus p74 protein, a highly conserved baculoviral envelope protein. *J. Biochem. Mol. Biol.* 36, 475-487.
- RASHIDAN, K. K., N. NASSOURY, P. N. GIANNOPOULOS, Y. MAUFFETTE and C. GUERTIN. 2004. Identification, Characterization and Phylogenic Analysis of Conserved Genes within the *odvp-6e/odv-e56* Gene Region of *Choristoneura fumiferana* Granulovirus. *J. Biochem. Mol. Biol.* 37, 206-212.
- RAPP, J. C., J. A. WILSON and L. K. MILLER. 1998. Nineteen baculovirus open reading frames, including *LEF-12*, support late gene expression. *J. Virol.* 72, 10197-10206.
- RAZUCK, F.B., B. RIBEIRO, J.H. VARGAS, J.L. WOLFF and B.M. RIBEIRO. 2002. Characterization of the p10 gene region of *Anticarsia gemmatilis* nucleopolyhedrovirus. *Virus Genes* 24, 243-247.

- RIBEIRO, B. M., K. HUTCHINSON, and L. K. MILLER. 1994. A mutant baculovirus with a temperature-sensitive IE-1 transregulatory protein. *Virology* 68, 1075-1084.
- ROBERTS, T. E. and P. FAULKNER. 1989. Fatty acid acylation of the 67K envelope glycoprotein of a baculovirus: *Autographa californica* nuclear polyhedrosis virus. *Virology* 172, 377-381.
- RODEMS, S. M. and P. D. FRIESEN. 1993. The hr5 transcriptional enhancer stimulates early expression from the *Autographa californica* nuclear polyhedrosis virus genome but is not required for virus replication. *J. Virol.* 67, 5776-5785.
- RODEMS, S.M., S. S. PULLEN, AND P. D. FRIESEN. 1997. DNA-dependent transregulation by IE1 of *Autographa californica* nuclear polyhedrosis virus: IE1 domains required for transactivation and DNA binding. *Virology* 71, 9270-9277.
- ROELVINK, P. W., M. M. M. VAN MEER, C. A. D. DEKORT, R. D. POSSEE, B. D. HAMMOCK and J. M. VLAK. 1992. Dissimilar expression of *Autographa californica* multiple polyhedrosis virus polyhedrin and p10 genes. *J. Gen. Virol.* 73, 1481-1489.
- ROHRMANN, G.F. 1986. Polyhedrin structure. *J. Gen. Virol.* 67,1499-1513.
- ROHRMANN, G. F. 1992. Baculovirus structural proteins, *J. Gen. Virol.* 73, 61-74.
- ROHRMANN, G. F., D. J. LEISY, K. -C. CHOW, G. D. PEARSON and G. S. BEAUDREAU. 1982. Identification, cloning and R-Loop mapping of polyhedrin gene from the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*. *Virology* 121, 51-60.
- ROHRMANN, G. F., M. N. PEARSON, T. J. BAILEY, R. R. BECKER and G. S. BEAUDREAU. 1981. N-terminal polyhedrin sequences and occluded baculovirus evolution. *J. Mol. Evol.* 17, 329-333.
- ROHRMANN, G. F., T. J. BAILEY, R. R. BECKER and G. S. BEAUDREAU. 1980. Comparison of the structure of C- and N- polyhedrins from two occluded viruses pathogenic for *Orgyia pseudotsugata*. *J. Virol.* 34, 360-365.
- RUEPP, S., U. KURATH, C.K. RENGGLI, R. BRUN and I. RODITI. 1999. Glutamic acid/alanine-rich protein from *Trypanosoma congolense* is the functional equivalent of 'EP' procyclin from *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 98, 151-156.
- RUSSELL, R. L Q., M. N. PEARSON and G. F. ROHRMANN. 1991. Immunoelectron microscopic exanuation of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus-infected *Lyrnantria dispar* cells: time course and localization of maior polyhedron - associated proteins. *J. Gen Virol.* 72, 275-283.

- RUSSELL, R. L. Q. and G. F. ROHRMANN. 1997. Characterization of P91, a protein associated with virions of an *Orgyia pseudotsugata* baculovirus. *Virology*. 233, 210-223.
- RUSSELL, R. L. Q. and G. F. ROHRMANN. 1993. A 25-kda protein is associated with the envelopes of occluded baculovirus virions. *Virology* 195, 532-540.
- SAMBROOK, J., E. F. FRITSCH and T. MANIATIS. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York.
- SANGER, F., S. NICHLEN and A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- SCHININA, M.E., D. BARRA, F. BOSSA, L. CALABRESE, L. MONTESANO, M.T. CARRI, P. MARIOTTINI, F. AMALDI and G. ROTILIO, (1989). Primary structure from amino acid and cDNA sequences of two Cu, Zn superoxide dismutase variants from *Xenopus laevis*. *Arch Biochem Biophys*. 272, 507-515.
- SCHONHERR, J. 1969. Field experiment on the biological control of *Choristoneura murinana* (Hbn.) by Bergoldiavirus. *Entomophaga* 14, 251-260.
- SHIKATA, E. 1977. Tabamovirus (tobacco mosaic virus) group, -p.237-255. In K Maramorosch (ed.), The atlas of insect and plant viruses. Academic Press, Inc. New York.
- SLACK, J. M. and G. W. BLISSARD. 1997. Identification of two independent transcriptional activation domains in the *Autographa californica* multicapsid nuclear polyhedrosis virus IE1 protein. *Virology* 71, 9579-9587.
- SLACK, J.M., E.M. DOUGHERTY and S.D. LAWRENCE. 2001. A study of the *Autographa californica* multiple nucleopolyhedrovirus ODV envelope protein p74 using a GFP tag. *J. Gen. Virol.* 82, 2279-2287.
- SUMMERS, M. D. Biophysical and biochemical properties of baculoviruses. 1975a. pp. 17-29 in *Baculoviruses for insect pest control: Safety Considerations*. Edited by Summers, M. D., Engler, R., Falcon L.A., Vail P. American Society for Microbiology. Washington D.C. ISBN 0-914826-07-7.
- SUMMERS, M. D. Comparison of in vivo and in vitro specificity of insect viruses. 1975b. pp. 77- 84 in *Baculoviruses for insect pest control: Safety Considerations*. Edited by Summers, M. D., Engler, R., Falcon L.A., Vail P. American Society for Microbiology. Washington D.C. ISBN 0-914826-07-7.
- SUMMERS, M. D. 1971. Electron microscopic observations on granulosis virus entry, uncoating and replication process during infection of the midgut cells of *Trichoplusia ni*. *J. Ultrastruct. Res.* 35, 606-625.

SUMMERS, M. D. 1969. Apparent in vivo pathway of granulosis virus invasion and infection. *J. Virol.* 4, 188-190.

SUMMER, M. D., and L. E. VOLKMAN. 1976. Comparison of biophysical and morphological properties of occluded and extracellular nonoccluded baculovirus from *in vivo* and in vitro host systems. *J. Virol.* 17, 962-972.

SWOFFORD, D. L. 2000. PAUP Version 4, Sinauer Associates, Sunderland, Massachusetts.

TANADA, Y, R. T. HESS and E. M. OMI. 1975. Invasion of a nuclear polyhedrosis virus in midgut of the armyworm, *Pseudaletia unipuncta*, and the enhancement of a synergistic enzyme. *J. Invertebr. Pathol.* 26, 99-104.

TANADA, Y. and R. T. HESS. 1976. Development of a nuclear polyhedrosis virus in midgut cells and penetration of the virus into the hemocoel of the armyworm, *Pseudaletia unipuncta*. *J. Invertebr. Pathol.* 28, 67-76.

TANADA, Y. and R. LEUTENEGGER. 1970. Multiplication of a granulosis virus in larval midgut cells of *Trichoplusia ni* and possible pathways of invasion into the hemocoel. *J. Ultrastruct. Res.* 30, 589-600.

TANADA, Y. and R. T. HESS. 1991. Baculoviridae, Granulosis viruses. In *Atlas of Invertebrate Viruses*, pp. 227- 257. Edited by J. R. Adams and J. R. Bonami. Boca Raton, FL: CRC Press.

THEILMANN, D. A. and S. STEWART. 1993. Analysis of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus trans-activators IE-1 and IE-2 using monoclonal antibodies. *J. Gen. Virol.* 74, 1819-1826.

THEILMANN, D. A. and S. STEWART. 1991. Identification and characterization of the IE-1 gene of *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 180, 492-508.

THEILMANN, D. A., J. K. CHANTLER, S. STEWART, H. T. M FLIPSEN, J. M. VLAK and N. E. CROOK. 1996. Characterization of a highly conserved baculovirus structural protein that is specific for occlusion-derived virions. *Virology* 218, 148-158.

THIEM, S. M. and L. K. MILLER. 1989a. Identification, sequence and transcriptional mapping of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* 63, 2008-2018.

THIEM, S. M. and L. K. MILLER. 1989b. A baculovirus gene with a novel transcription pattern encodes a polypeptide with a zinc finger and a leucine zipper. *J. Gen. Virol.* 63, 4489-4497.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN and D.G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24, 4876-4882.

TODD, J. W., A. L. PASSARELLI and L. K. MILLER. 1995. Eighteen baculovirus genes, including *lef-11*, *p35*, *39K*, and *p47*, support late gene expression. *J. Virol.* 69, 968-974.

TODD, J. W., A. L. PASSARELLI, A. LU and L. K. MILLER. 1996. Factors regulating baculovirus late and very late gene expression in transient-expression assays. *J. Virol.* 70, 2307-2317.

TWEETEN, K. A., L. A. BULLA JR., and R. A. CONSIGLI. 1980. Characterization of an extremely basic protein derived from granulosis virus nucleocapsids. *J. Virol.* 33, 866-892.

VAN CAMP, W., C. BOWLER, R. VILLARROEL, E.W. TSANG, M. VAN MONTAGU and D. INZE. 1990. Characterization of iron superoxide dismutase cDNAs from plants obtained by genetic complementation in *Escherichia coli*. *Proc Natl. Acad. Sci. USA* 87, 9903-9907.

VAN OERS, M. M., J. T. M. FLIPSEN, C. B. E. M. REUSKEN, E. L. SLIWINSKY, R. W. GOLDBACH and J. M. VLAK. 1993. Functional domains of the p10 protein of *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* 74, 563-574.

VAN OERS, M. M., J.T. FLIPSEN, C.B. REUSKEN and J.M. VLAK. 1994. Specificity of baculovirus p10 functions. *Virology* 200, 513-523.

VAN OERS, M.M. and J.M. VLAK. 1997. The baculovirus 10-kDa protein. *J. Invertebr. Pathol.* 70, 1-17.

VAN OERS, M.M., Z. HU, B.M. ARIF, E.A. VAN STRIEN, J.W. VAN LENT and J.M. VLAK. 1998. The single-nucleocapsid nucleopolyhedrovirus of *Buzura suppressaria* encodes a P10 protein. *J. Gen. Virol.* 79, 1553-1562.

VLAK, J. M. and G. F. ROHRMANN. 1985. The nature of polyhedrin. In *Viral Insecticides for Biological Control*. pp.489-544. Edited by K. Maramorosch & K. E. Sherman. Orlando: Academic Press.

VLAK, J. M., F. A. KLINKENBERG, K. J. ZAAL, M. USMANY, E.C. KLINGE-ROODE, J.B. GEERVLIET, J. ROOSIEN and J.W. VAN LENT. 1988. Functional studies on the p10 gene of *Autographa californica* nuclear polyhedrosis virus using a recombinant expressing a p10-beta-galactosidase fusion gene. *J. Gen. Virol.* 69, 765-776.

VOLKMAN, L.E. and P. A. GOLDSMITH. 1985. Mechanism of neutralization of budded *Autographa californica* nuclear polyhedrosis virus by a monoclonal antibody: inhibition of entry by adsorptive endocytosis. *Virology* 143, 183-195.

VOLKMAN, L. E. and M. D. SUMMERS. 1977. *Autographa californica* nuclear polyhedrosis virus: comparative infectivity of the occluded, alkali liberated, and nonoccluded forms. *J. Inverteb. Pathol.* 30, 102-103.

WANG, H., X. CHEN, J.M. VLAK and Z. HU. 2001. Sequence and transcriptional analysis of the p10 gene of *Helicoverpa armigera* single nucleocapsid polyhedrovirus. *Bing Du Xue Bao* 17, 81-86.

WANG, W., N. LEAT, B. FIELDING, S. DAVISON. 2001. Identification, sequence analysis, and phylogeny of the immediate early gene 1 of the *Trichoplusia ni* single nucleocapsid polyhedrosis virus. *Virus Genes* 23, 53-62.

WASHBURN, J.O., E.Y. CHAN, L.E. VOLKMAN, J.J. AUMILLER and D.L. JARVIS. 2003. Early synthesis of budded virus envelope fusion protein GP64 enhances *Autographa californica* multicapsid nucleopolyhedrovirus virulence in orally infected *Heliothis virescens*. *J Virol.* 77, 280-90.

WEI, Y.J., Q.X. LONG, S.W. CHEN, and X.Z. WANG. 1998. Nucleotide sequence and characterization of the p10 gene of *Spodoptera litura* nuclear polyhedrosis virus. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 30, 550-555.

WHITFORD, M., S. STEWART, J. KUZIO and P. FAULKNER. 1989. Identification and sequence analysis of a gene encoding gp67 an abundant envelope glycoprotein of the baculovirus, *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* 63, 1393-1399.

WHITFORD, M. and P. FAULKNER. 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, 2427-2435.

WHITT, M. A. and J. S. MANNING. 1988. A phosphorylated 34-kDa protein and a subpopulation of polyhedrin are thiol linked to the carbohydrate layer surrounding a baculovirus occlusion body. *Virology* 163, 33-42.

WILLIAMS, G.V., D.Z. ROHEL, J. KUZIO and P. FAULKNER. 1989. A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion/deletion mutants. *J. Gen. Virol.* 70, 187-202.

WILSON, M. E. 1988. A synthetic peptide to the to the predicted 6.9K translation product of the *HindIII*-H/*EcoRI*-D region of Ac MNPV genome induces antibodies to tyhe basic-DNA binding protein. *Virus Res.* 9, 21-31.

WILSON, M. E. and L. K. MILLER. 1986. Changes in the nucleoprotein complex of a baculovirus DNA during infection. *Virology* 151, 315-328.

WILSON, M. E., T. H. MALNPRIZE, P. D. FREISEN and L. K. MILLER. 1987. Location, transcription, and sequence of a baculovirus gene encoding for a small arginine-rich polypeptide. *J. Virol.* 61, 661-666.

WILSON, M. E. and R. A. CONSIGLI. 1985a. Characterization of a protein kinase activity associated with purified capsids of the granulosis virus infecting *Plodia interpunctella*. *Virology* 143, 516-525.

WILSON, M. E. and R. A. CONSIGLI. 1985b. Functions of a protein kinase activity associated with purified capsids of the granulosis virus infecting *Plodia interpunctella*. *Virology* 143, 526-532.

WILSON, J.A., J.E. HILL, J. KUZIO and P. FAULKNER. 1995. Characterization of the baculovirus *Choristoneura fumiferana* multicapsid nuclear polyhedrosis virus p10 gene indicates that the polypeptide contains a coiled-coil domain. *J. Gen. Virol.* 76, 2923-2932.

WOLGAMOT, G. M., G. H. CHRISTIAN, R. L. Q. RUSSEL and G. F. ROHRMANN. 1993. Immunocytochemical characterization of p24, a baculovirus capsid-associated protein. *J. Gen. Virol.* 74, 103-107.

WOLFE, J. T., D. ROSS and G. M. COHEN. 1994. A role for metals and free radicals in the induction of apoptosis in thymocytes. *FEBS Lett.* 352, 58-62.

WORMLEATON, S., J. KUZIO and D. WINSTANLEY. 2003. The complete sequence of the *Adoxophyes orana* granulovirus genome. *Virology* 311, 350-65.

WU, W.W., J.W. WANG, F. XIE, Q.X. LONG and X.Z. WANG. 2003. Baculovirus p74 gene is a species-specific gene. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 35, 834-840.

XEROS, N. 1956. The virogenic stroma in nuclear and cytoplasmic polyhedroses. *Nature (London)* 128, 412-413.

XING L., A. PANG, H. A. M. LAUZON, S. S. SOHI and B. M. ARIF. 1997. The gene encoding the capsid protein P82 of the *Choristoneura fumiferana* multicapsid nucleopolyhedrovirus: sequencing, transcription and characterization by immunoblot analysis. *J. Gen. Virol.* 78, 2665-2673.

YANG, C. L., D. A. STETLER and R. F. WEAVER. 1991. Structural comparison of the *Autographa californica* nuclear polyhedrosis virus- induced RNA polymerase and three nuclear RNA polymerases from the host, *Spodoptera frugiperda*. *Vir. Res.* 20, 251-264.

YANG, S. and L. K. MILLER. 1999. Activation of baculovirus very late promoters by interaction with very late factor 1. *J. Virol.* 73, 3404-3409.

YANG, S. and L. K. MILLER. 1998. Control of baculovirus *polyhedrin* gene expression by very late factor 1. *Virology* 248, 131-138.

ZANOTTO, P.M., B.D. KESSING, J.E. MARUNIAK. 1993. Phylogenetic interrelationships among baculoviruses: evolutionary rates and host associations. *J. Invertebr. Pathol.* 62, 147-164.

ZHANG, C.X. and J.C. WU. 2001. Genome Structure and the p10 Gene of the *Helicoverpa armigera* Nucleopolyhedrovirus. *Sheng wu hua hsueh yu sheng wu wu li hsueh pao* 33, 179-184.

ZOOG, S. J., J. J. SCHILLER, J. A. WETTER, N. CHEJANOVSKY and P. D. FRIESEN. 2002. Baculovirus apoptotic suppressor P49 is a substrate inhibitor of initiator caspases resistant to P35 in vivo. *EMBO* 21, 5130-5140.

ZUIDEMA, D., M. M. VAN OERS, E. A. VAN STRIEN, P. C. CABALLERO, E. J. KLOK, R. W. GOLDBACH and J. M. VLAK. 1993. Nucleotide sequence and transcriptional analysis of p10 gene of *Spodoptera exigua* nuclear polyhedrosis virus. *J. Gen. Virol.* 74, 1017-1024.

**APPENDIX A- LIST OF PUBLICATIONS AND COMMUNICATIONS DURING
MY Ph.D.**

Rashidan, K. K., N. Nassoury, P. N. Giannopoulos and C. Guertin. 2005. Transcription, Translation, and Immunolocalization of ODVP-6E/ODV-E56 and p74 Proteins: Two Highly Conserved ODV-associated Envelope Proteins of *Choristoneura fumiferana* Granulovirus. *J. Bioch. Mol. Biol.* 38, 65-70.

Rashidan, K. K., N. Nassoury, and C. Guertin. 2004. Identification, characterization and phylogenic analysis of conserved genes within the p74 gene region of *Choristoneura fumiferana* granulovirus. *J. Bioch. Mol. Biol.* 37, 700-708.

Rashidan, K. K., N. Nassoury P. N. Giannopoulos, Y. Mauffette and C. Guertin. 2004. Identification, characterization and phylogenic analysis of conserved genes within the *odvp-6e/odv-e56* gene region of *Choristoneura fumiferana* granulovirus. *J. Bioch. Mol. Biol.* 37, 206-212.

Rashidan, K. K., C. Guertin, and J. Cabana. 2004. Granulovirus. In Encyclopedia of Entomology, University of Florida, pp.991-997. Capinera, J. L. (Ed.) ISBN: 0-7923-8670-1.

Rashidan, K. K., N. Nassoury, S. Tazi, P. N Giannopoulos and C. Guertin. 2003. *Choristoneura fumiferana* granulovirus p74 protein, a highly conserved baculoviral envelope protein *J. Bioch. Mol. Biol.* 36, 475-487.

Rashidan, K. K., N. Nassoury, P. N Giannopoulos and C. Guertin. 2002. Identification and characterization of a conserved baculoviral structural protein ODVP-6E/ODV-E56 from *Choristoneura fumiferana* granulovirus. *J. Bioch. Mol. Biol.* 35, 595-603.

Rashidan, K. K., N. Nassoury, A. Merzouki and C. Guertin. 2002. Identification and Characterization of a Putative Baculoviral Transcriptional Factor (IE-1) from *Choristoneura fumiferana* granulovirus. *J. Bioch. Mol. Biol.* 35, 553-561.

Rashidan, K. K., N. Nassoury, A. Merzouki, J. Sawicki and C. Guertin. 2001. Identification and Characterization of a new 40 kDa protein in ChfuGV that may play a similar role as ODV-E56 in other baculoviruses. American Society for Virology- 20th Annual Meeting, July 20-25, 2001, Wisconsin, Madison.

Smati, R., C. Guertin, M. Henrichon, K. K. Rashidan, M. Arella, S. Slilaty and A. Merzouki. 2001. Random subcloning of *ChfuGV* baculovirus genome into gap-free sequence provider pTB cloning vector. American Society for Virology- 20th Annual Meeting, July 20-25, 2001, Wisconsin, Madison.

Rashidan, K. K., N. Nassoury, A. Merzouki, R. Smati, J. Sawicki and C. Guertin. 2001. Characterization of the *Choristoneura fumiferana* granulovirus ODVP-6E: A Baculovirus conserved envelope- associated protein. 51st Annual Meeting of Canadian

Society of Microbiologists, June 10-13, 2001, Waterloo, Ontario. *"Bioinformatics & Biotechnology in the Post-Genome Era"*.

Bourassa, N, K. K. Rashidan, J. Cabana and C. Guertin. 2001. Identification of microbial contaminants during *in vivo* production of ChfuGV in *Choristoneura fumiferana* larvae. 51st Annual Meeting of Canadian Society of Microbiologists, June 10-13, 2001, Waterloo, Ontario. *"Bioinformatics & Biotechnology in the Post-Genome Era"*.

Smati, R., C. Guertin, M. Henrichon, K. K. Rashidan, M. Arella, S. Slilaty and A. Merzouki. 2001. Random subcloning and sequencing of *Choristoneura fumiferana* granulovirus genome. 51st Annual Meeting of Canadian Society of Microbiologists, June 10-13, 2001, Waterloo, Ontario. *"Bioinformatics & Biotechnology in the Post-Genome Era"*.

Rashidan, K. K., C. Guertin and J. Cabana. 2001. Granulovirus; a promising alternative in microbiological control programs against spruce budworm. *Tordeuse des bourgeons de l'épinette: l'apprivoiser dans nos stratégies d'aménagement. 4th annual meeting March 27-29 Shawinigan, Quebec.*