

UNIVERSITÉ DU QUÉBEC

**THÈSE PRÉSENTÉE À L'INRS-INSTITUT ARMAND-FRAPPIER
COMME EXIGENCE PARTIELLE DU DOCTORAT EN
VIROLOGIE ET IMMUNOLOGIE**

PAR

ALIOU BAH

**CONTRIBUTION À LA CARACTÉRISATION MOLÉCULAIRE DU
GRANULOVIRUS DE LA TORDEUSE DES BOURGEONS
DE L'ÉPINETTE**

CENTRE DE MICROBIOLOGIE ET BIOTECHNOLOGIE

Janvier 2000

TABLE DES MATIÈRES

	Pages
LISTE DES TABLEAUX	V
LISTE DES FIGURES	VI
SOMMAIRE	VIII
INTRODUCTION	1
CHAPITRE 1	
REVUE BIBLIOGRAPHIQUE	6
1.1 Généralités	7
1.2 Baculovirus	7
Généralités	7
1.2.1 Classification et nomenclature des baculovirus	8
1.3 Granulovirus	9
1.3.1 Propriétés et caractéristiques	9
Structure du virion	9
Infection et réplication	11
1.3.2 Organisation des gènes sur le génome	14
1.3.3 Études rapportées sur le génome du GV	15
1.4 Variations génotypiques chez les baculovirus	17
1.4.1 Techniques utilisées pour détecter des mutations dans les régions variables d'ADN génomiques	17
Détection des mutations en gel de polyacrylamide en gradient dénaturant perpendiculaire	19
Électrophorèse des fragments d'ADN homoduplexes et hétéroduplexes en gel de polyacrylamide en gradient parallèle de dénaturant	19
Électrophorèse des fragments d'ADN homoduplexes et hétéroduplexes en gel de polyacrylamide en gradient constant de dénaturant	22
1.5 Conclusion	22

CHAPITRE 2

IDENTIFICATION ET ANALYSES DE LA SÉQUENCE DU GÈNE DE LA GRANULINE DU GRANULOVIRUS DE <i>Choristoneura fumiferana</i> , ChfuGV	25
Identification and sequence analyses of the granulin gene of <i>Choristoneura fumiferana</i> granulovirus	26
Introduction	29
Material and methods	30
Results	31
Discussion	34
Acknowledgements	35
References	35

CHAPITRE 3

LE GRANULOVIRUS DE <i>Choristoneura fumiferana</i> , ChfuGV: ANALYSES DE LA SÉQUENCE ET CARACTÉRISATION DE L'EXTRÉMITÉ 5' DU GÈNE ORF891	37
<i>Choristoneura fumiferana</i> granulovirus, ChfuGV: sequence analysis and 5' end characterization of ORF891	38
Introduction	41
Material and methods	42
Results	43
Discussion	45
Acknowledgements	48
References	48

CHAPITRE 4

LE GRANULOVIRUS DE <i>Choristoneura fumiferana</i> , ChfuGV: ANALYSE DE LA SÉQUENCE DE L'ORF381 ET DE LA	
---	--

TRANSCRIPTION DE SA RÉGION ADJACENTE, LE GÈNE DE LA GRANULINE	51
The Granulovirus of <i>Choristoneura fumiferana</i> , ChfuGV: sequence analysis of ORF381 and its adjacent granulin gene transcriptional region	52
Introduction	55
Material and methods	56
Results	59
Discussion	60
Acknowledgements	62
References	62
DISCUSSION GÉNÉRALE	78
CONCLUSION GÉNÉRALE	86
REMERCIEMENTS	90
RÉFÉRENCES	93
COMMUNICATIONS ET PUBLICATIONS	105

LISTE DES TABLEAUX

	Pages
CHAPITRE 2	
Tableau 1 Percent identity of the 5' and 3' flanking regions of the granulin gene of CfGV compared to ClGV, PbGV and TnGV	32
Tableau 2 Percent identity of nucleotide and amino acid sequences of the granulin gene of CfGV, ClGV, PbGV and TnGV	35
 CHAPITRE 4	
Tableau 1 Percent identity and similarity of sequences producing significant alignment with the N and C-terminal amino acid sequences of ChfuGV ORF381	65

LISTE DES FIGURES

		Pages
CHAPITRE 1		
Figure 1	Représentation schématique d'un gel de gradient perpendiculaire en conditions dénaturantes	20
Figure 2	Représentation schématique d'un gel de gradient parallèle ou constant en conditions dénaturantes	23
 CHAPITRE 2		
Figure 1	Schematic representation of clones pGV199 and pGV606	31
Figure 2	Multiple alignments of the 5' flanking sequence of the granulin ORF of CfGV, CIGV, CpGV, PbGV and TnGV	32
Figure 3	The nucleotide sequence of CfGV granulin gene coding region including its 5' and 3' flanking regions	33
Figure 4	Multiple alignment of CfGV, CIGV, PbGV and TnGV granulin amino acid sequences	34
 CHAPITRE 3		
Figure 1	Schematic representation of the ChfuGV ORF891 with a portion of the ChfuGV granulin and <i>egt</i> gene regions	42
Figure 2	Nucleotide sequence of ChfuGV ORF891	44
Figure 3	Multiple alignment of amino acid sequences of ChfuGV ORF891, CIGV ORF909 and CpGV ORF124R	45
Figure 4	Transcriptional analysis of ChfuGV ORF891	46
 CHAPITRE 4		
Figure 1	Schematic representation of the ChfuGV ORF381, location on clone pGV199	66
Figure 2	Nucleotide sequence of the ChfuGV ORF381 and its adjacent granulin gene transcriptional region	68
Figure 3	Schematic representation of the MNPV polyhedrin and GV granulin gene regions organization	70

Figure 4	Transcriptional analysis of ChfuGV granulin gene	72
Figure 5	Analysis by the DCode system of the 5' intergenic and the 3' flanking regions of the ChfuGV granulin gene in a perpendicular, parallel and constant denaturing gradient gels	74
Figure 6	Comparison of the restriction enzyme (REN) profiles of the original genome of Bonaventure ChfuGV isolate (BI), with each of the first to five serial passages (P1-P5) in spruce budworm and isolates 87-181 and 87-271 from New Brunswick (NBI)	76

INTRODUCTION

La tordeuse des bourgeons de l'épinette, *Choristoneura fumiferana* (Clem.), est un lépidoptère qu'on rencontre uniquement en Amérique du Nord. Les larves de cet insecte s'attaquent principalement aux conifères, notamment le sapin baumier (*Abies balsamea* (L.) Mill.) et l'épinette blanche (*Picea glauca* (Moench) Voss). Cet insecte est considéré comme le principal ravageur des forêts conifériennes du Canada. Afin de démontrer l'importance des impacts économiques de cet insecte, notons par exemple qu'au Québec, en 1985, plus de 110 millions de dollars ont été dépensés pour combattre l'épidémie causée par la tordeuse des bourgeons de l'épinette (Comtois, 1988). Selon le même auteur, pour la période allant de 1970 à 1990, les pertes de bois en essences de sapin et d'épinette ont été évaluées à plus de 135 millions de mètres cubes.

De nos jours, plusieurs stratégies d'interventions sont proposées afin de minimiser les impacts des insectes ravageurs. Dans le cas de la tordeuse des bourgeons de l'épinette, l'utilisation d'agents chimiques et biologiques demeure encore l'approche curative proposée lors de situations épidémiques. Cependant, on observe de plus en plus, des systèmes de lutte intégrée qui laissent place à de nouvelles pratiques, notamment au niveau de la sylviculture, du maintien des complexes prédateur-proie et de l'utilisation de phéromones sexuelles. C'est dans ce contexte que s'inscrivent les démarches visant le développement de nouvelles approches en lutte biologique.

Les approches biologiques utilisent les pathogènes naturels d'insectes tels, les bactéries, les virus, les champignons et les protozoaires comme agents de lutte biologique. Actuellement au Québec, seul l'emploi d'insecticides biologiques est accepté dans les programmes de lutte contre la tordeuse des bourgeons de l'épinette (Comtois, 1988). Parmi les agents pathogènes d'insectes, les baculovirus, qui regroupent les nucléopolyédrovirus (NPV) et les granulovirus (GV), sont de plus en plus considérés comme des candidats ayant un fort potentiel insecticide dans la lutte contre la tordeuse des bourgeons de l'épinette (Hagedorn *et al.*, 1989). En effet, les larves de la tordeuse des bourgeons de l'épinette semblent être susceptibles aux infections causées par le granulovirus de *C. fumiferana* (ChfuGV). C'est dans ce contexte que des travaux ont été entrepris afin d'étudier ce virus.

En général, les GV sont considérés comme des agents potentiels qu'on pourrait utiliser en lutte biologique contre certaines populations d'insectes ravageurs en

agriculture et en foresterie. Comme indiqué précédemment, les granulovirus (GV) sont des virus d'insectes qui appartiennent à la famille des *Baculoviridae*. Ils infectent en général les lépidoptères et plus particulièrement les *Noctuidae* et les *Tortricidae*. Les GV ont été répertoriés chez environ 150 espèces de lépidoptères (Adams et Bonami, 1991). Citons, cependant, que plusieurs insectes d'importance économique comme la tordeuse des bourgeons de l'épinette, *C. fumiferana* (Clem.), la carpocapse de la pomme, *Cydia pomonella* (L.), la fausse arpenreuse du chou, *Trichoplusia ni* (Hbn.), la piéride du chou, *Pieris brassicae* (L.) sont susceptibles au GV (Crook, 1981; Akiyoshi *et al.*, 1985; Arif *et al.*, 1986). Selon la convention établie par le comité international sur la taxonomie des virus (ICTV), les GV portent le nom de leurs hôtes naturels respectifs et c'est pourquoi le GV de la tordeuse des bourgeons de l'épinette se nomme ChfuGV (Murphy *et al.*, 1995).

Le génome des granulovirus est constitué d'un ADN double brin, circulaire et superenroulé avec une longueur moyenne de 115 à 135 kilopaires de bases (kpb) (Summers et Anderson, 1972; Tweeten *et al.*, 1977). Leur ADN viral est contenu dans une capsidie entourée par une nucléocapsidie de 40-100 sur 250-360 nm. La nucléocapsidie est recouverte par une enveloppe virale qui à son tour est contenue dans un corps d'inclusion ou granule (Adams et Bonami, 1991). Cette dernière possède une matrice constituée par une protéine paracristalline appelée granuline. Elle contient en général une nucléocapsidie par enveloppe (Adams et Bonami, 1991). La masse moléculaire de la granuline est estimée à environ 27,6 à 37,8 kDa dépendant des types de GV (Falcon et Hess, 1985). La granuline du GV comme le polyèdre du nucléopolyédrovirus (NPV) serait impliquée dans l'infection du virion. Elle se dissout au niveau de l'intestin moyen de la larve dans un milieu alcalin à pH supérieur à 10 et sous l'action de certaines protéases associées au corps d'inclusion (Adams et Bonami, 1991). La granuline représente plus de 18% des protéines totales du virion. Elle joue un rôle très important dans la protection du virus contre les rayons ultraviolets du soleil, la transmission horizontale du virus et serait aussi impliquée dans le pouvoir infectieux du virion (Blissard et Rohrmann, 1990; Jehle *et al.*, 1992).

Le gène de la granuline a été séquencé chez les granulovirus de *Trichoplusia ni*, TnGV; *Cydia pomonella*, CpGV; *Cryptophlebia leucotreta*, CrleGV et *Xestia c-nigrum*, XcGV (Akiyoshi *et al.*, 1985; Chakerian *et al.*, 1985; Jehle et Backhaus, 1994; Bah *et al.*,

1997; Goto *et al.*, 1998). Des études comparatives ont montré que la granuline est hautement conservée chez ces GV avec plus de 75% d'homologie en acides aminés et en séquences nucléotidiques. Cependant, l'arrangement des gènes dans la région de la granuline chez les GV n'est pas identique à celle observée chez la polyédrique des NPV (Jehle *et al.*, 1994; Crook *et al.*, 1997; Bah *et al.*, 1999; Goto *et al.*, 1998; Smith et Goodale, 1998).

L'étude et la caractérisation moléculaire du ChfuGV sont essentielles à la compréhension des rôles et des fonctions des gènes impliqués dans le processus d'infection. Elles sont les prémisses à l'utilisation du virus en milieu naturel sur de grandes surfaces. De plus, avec le développement de nouvelles approches en biologie moléculaire, elles devraient conduire à une augmentation du pouvoir infectieux et une amélioration de la production du virus.

La présente étude avait pour buts de tester la stabilité du génome du ChfuGV par comparaison des profils de restriction du génome de différents GV, de comprendre l'organisation génomique de la région du gène de la granuline du ChfuGV, d'analyser et de caractériser quelques gènes de cette région et enfin, d'étudier la variabilité génomique au niveau de cette région. Nos objectifs à atteindre étaient les suivants: 1) cloner, séquencer et caractériser le gène de la granuline du ChfuGV; 2) analyser et caractériser le gène situé immédiatement en amont du gène de la granuline du ChfuGV; 3) analyser et caractériser le gène situé immédiatement en aval du gène de la granuline; 4) vérifier la stabilité du virus en comparant les profils de restrictions du génome du virus original de laboratoire avec celui des virus issus de passages sériés chez la tordeuse des bourgeons de l'épinette et celui d'isolats viraux de différentes provenances géographiques; 5) étudier la variabilité génomique dans la région du gène de la granuline du ChfuGV.

Les travaux réalisés dans ce projet de doctorat sont présentés sous la forme de trois articles scientifiques dont deux sont déjà publiés et le troisième a été soumis pour publication. Dans le souci de simplifier la compréhension des différents résultats et afin de préserver l'intégrité des documents publiés dans des chapitres distincts, le lecteur retrouvera dans le chapitre 1 la revue de littérature et dans les chapitres 2, 3 et 4, les trois articles en question qui regroupent les différents objectifs avancés précédemment. Cependant, il est à souligner que les résultats de l'analyse du transcrit du gène de la

granuline du ChfuGV ont été insérés dans le troisième article du chapitre 4, vu que cette partie du travail a été réalisée après la publication du premier article. Une copie de chaque article a été insérée dans le chapitre correspondant.

Il est très important de souligner que les résultats présentés dans cette thèse sont innovateurs puisqu'il n'existe aucune étude antérieure portant sur la caractérisation moléculaire des gènes du granulovirus du ChfuGV. De façon générale, tous les résultats de recherches en laboratoire émanant de ce projet de doctorat sont l'oeuvre personnelle de l'étudiant. L'identification et la caractérisation du gène de la granuline du ChfuGV; l'identification et la caractérisation du gène ORF891 du ChfuGV; les analyses de séquences de l'ORF381 du ChfuGV; les études de variabilité génomique dans la région de la granuline et de la stabilité du génome du ChfuGV sont la somme du travail réalisé par ce dernier. Cependant, ce projet de doctorat doit sa réalisation par le dévouement et l'implication des coauteurs, particulièrement mon directeur et codirecteur de recherche, Claude Guertin et Max Arella. Je dois aussi souligner la contribution de Christopher J. Lucarotti qui s'est aussi bien impliqué dans ce travail. Nous lui devons les isolats viraux du ChfuGV en provenance du Nouveau Brunswick. Ses idées et sa connaissance du sujet ont été indispensables à la rédaction des trois articles. Jean Bergeron a aussi participé aux premiers stades de ce projet, dans l'amplification du cadre de lecture du gène de la granuline du ChfuGV. La contribution de ces différentes personnes à la réalisation des travaux se traduit par leur présence comme coauteurs des différentes publications.

CHAPITRE 1

Revue Bibliographique

1.1 Généralités

Les travaux sur les maladies d'insectes d'origines inconnues ont commencé à attirer l'attention des chercheurs bien avant le dix-neuvième siècle (Miller, 1996). En 1926, on rapportait la présence d'une maladie virale connue sous le nom de maladie du noyau ou grasserie chez les larves de *P. brassicae* et *Agrotis segetum* et grâce aux techniques de filtration des virus développées à l'époque (revue de Benz, 1986), il était désormais possible d'identifier l'agent causal de cette maladie larvaire, le virus de la granulose, connu maintenant sous le nom de granulovirus (Murphy *et al.*, 1995). En 1948, on arrivait à décrire les formes circulaires de ces particules virales contenues dans des corps d'inclusion ou granules chez des larves infectées de *Cacoecia murinana* et en 1949, il était déjà possible d'observer les virions et les corps d'inclusion de baculovirus chez des larves infectées (Adams et Bonami, 1991).

Depuis plusieurs années, les virus pathogènes d'insectes, particulièrement les baculovirus, retiennent l'attention des agriculteurs et des forestiers pour leur potentiel comme agents de régulation des populations d'insectes ravageurs.

1.2 Baculovirus

Généralités

L'intérêt pour l'utilisation des virus pathogènes d'insectes, principalement les baculovirus, comme pesticides biologiques a commencé vers 1949. Depuis, on assista à la découverte de nouveaux types de virus entomopathogènes tels, les iridovirus, les nodavirus et les polydnavirus (revue de Miller, 1996). Ce sont, cependant, les baculovirus qui ont attiré la plus grande attention chez les chercheurs, les producteurs et les utilisateurs en raison de leur grand potentiel comme insecticide biologique (Hagedorn *et al.*, 1989).

Comparés aux autres agents biologiques, les baculovirus ont un spectre d'hôtes plus étroit, on les retrouve principalement chez les arthropodes et plus particulièrement chez les insectes de l'ordre des lépidoptères et des hyménoptères (Ignoffo, 1968). Plusieurs des insectes nuisibles répertoriés susceptibles à ces virus sont reconnus pour leur grande importance économique (Adams et Bonami, 1991). Les baculovirus peuvent rester stables après de longues périodes de conservation (Jarvis et Garcia, 1994a). Ils ne semblent représenter aucun danger ni pour les autres invertébrés non ciblés ni pour les

vertébrés (Smirnoff, 1979; Hartig *et al.*, 1989; Hartig *et al.*, 1991; Adams et Bonami, 1991). Les baculovirus, notamment les nucléopolyédrovirus, constituent l'un des larges groupes de virus d'insectes les plus diversifiés et les plus étudiés (Wood et Granados, 1991). Dans un contexte écologique, l'utilisation des baculovirus en lutte biologique intégrée constitue une avenue prometteuse.

Aujourd'hui, en plus de leur potentiel comme agent de lutte biologique et grâce à l'avènement de la biologie moléculaire et de la biotechnologie, les baculovirus sont très utilisés dans les industries pharmaceutiques et dans les laboratoires, notamment comme vecteurs d'expression. Il est possible de cloner dans ces vecteurs des fragments d'ADN génomiques hétérologues de grandes tailles et produire des quantités importantes de protéines recombinantes (Kidd et Emery, 1993). De plus, il est possible d'obtenir à partir de manipulations génétiques, des baculovirus recombinants beaucoup plus efficaces que les souches originales, tant pour l'expression génomique que pour l'utilisation comme insecticide biologique (Wood et Granados, 1991).

1.2.1 Classification et nomenclature des baculovirus

Le nom des baculovirus dérive du latin *baculum*, qui veut dire bâton. Dans l'ancienne classification, la famille des *Baculoviridae* étaient regroupée en deux sous familles, les *Eubaculovirinae* recouverts d'un corps d'inclusion et les *Nudibaculovirinae*, caractérisés par la présence ou l'absence de corps d'inclusion (Adam et Bonami, 1991). Lors du sixième congrès du comité international sur la taxonomie des virus (ICTV), l'ancienne classification des *Baculoviridae* a été modifiée. Les *Baculoviridae* sont maintenant classés en deux groupes, les nucléopolyédrovirus (NPV) et les granulovirus (GV) (Murphy *et al.*, 1995). Le premier groupe de virus, les NPV, est le plus connu et le plus étudié. Le deuxième groupe, les granulovirus qui font l'objet de la présente étude, sont par contre beaucoup moins connus.

En général, on rencontre les baculovirus à la surface des plantes et dans le sol (Blissard et Rohrmann, 1990). Chez les NPV, la polyédrine qui représente le corps d'inclusion du virion, peut contenir une ou plusieurs nucléocapsides regroupées dans une même enveloppe. C'est pourquoi chez les différents membres du groupe, on retrouve des nucléopolyédrovirus avec une seule nucléocapside (SNPV) comme c'est le cas chez *Bombyx moris* (BmSNPV) et d'autres avec plusieurs nucléocapsides (MNPV) comme

chez *Autographa californica* (AcMNPV) et *Orgyia pseudotsugata* (OpMNPV). Par contre, chez les GV, dans la plupart des cas, le corps d'inclusion contient une seule nucléocapside par enveloppe, par exemple, le GV de *Choristoneura fumiferana* (ChfuGV) (Blissard et Rohrmann, 1990).

1.3 Granulovirus

Les granulovirus (GV) sont retrouvés principalement chez les lépidoptères, dans les familles des *Noctuidae* et des *Tortricidae* qui renferment bon nombre d'insectes d'importance économique. Aujourd'hui, ces virus sont répertoriés chez environ 150 espèces d'insectes. Il faut noter que les GV infectent seulement les membres de la famille-hôte (Adams et Bonami, 1991).

1.3.1 Propriétés et caractéristiques

Structure du virion

Chez les GV, l'ADN viral est contenu dans une capsidie entourée par une nucléocapsidie. La nucléocapsidie est recouverte par une enveloppe virale qui à son tour est contenue dans un corps d'inclusion ou granule (Adams et Bonami, 1991).

Les corps d'inclusions possèdent une forme ovocylindrique mesurant 120 à 350 nm par 300 à 500 nm avec une matrice constituée par une protéine paracristalline nommée granuline. Leurs formes et leurs dimensions varient considérablement en fonction de l'hôte. Cependant, il existe aussi des formes cubiques qui ressemblent à celles des polyèdres des NPV (Hunter *et al.*, 1972). Ils sont recouverts d'une fine couche composée majoritairement d'hydrates de carbones. Lors de l'infection primaire, une protéase endogène du corps d'inclusion, serait impliquée dans la dissolution de ce dernier. Les corps d'inclusion sont constitués de lipoprotéines considérées comme facteurs synergistiques, facilitant l'attachement de la molécule de l'enveloppe virale avec les sites des récepteurs situés à la surface des cellules de la membrane plasmique (Tweeten *et al.*, 1978; Langridge et Balter, 1981). En plus des lipoprotéines, elles contiendraient aussi des facteurs de protéines appelés VEF "virale enhancing factors", qui augmenteraient l'infection virale en dégradant la membrane péritrophique exposant ainsi les cellules épithéliales de l'intestin moyen de la larve aux virions (Hashimoto *et al.*, 1991).

La masse moléculaire de la granuline varie entre 27,6 et 37,8 kDa dépendant des types de GV (Falcon et Hess, 1985). La granuline des GV comme la polyédrine des NPV serait impliquée dans la protection, la transmission horizontale et le pouvoir infectieux du virion. Le corps d'inclusion qui contient en général une nucléocapside est recouverte d'une enveloppe (Blissard et Rohrmann, 1990; Adams et Bonami, 1991; Jehle *et al.*, 1992).

La nucléocapside est une structure cylindrique faite de protéines qui constituent la capsid et d'un ADN. Elle est munie d'une coiffe à ses deux extrémités et mesure 40-100 nm sur 250-360 nm. Comme chez les NPV, la nucléocapside peut acquérir une enveloppe par synthèse *de novo* dans le noyau par morphogénèse ou par synthèse *de novo* dans le cytoplasme en bourgeonnant à travers les cellules de la membrane plasmique. En général, chaque enveloppe contient une nucléocapside (Adams et Bonami, 1991).

La capsid est constituée d'une série de sous-unités d'anneaux alignés perpendiculairement à l'axe longitudinal (Tweeten *et al.*, 1977). Elle est composée de deux polypeptides principaux, un polypeptide structural majeur de haut poids moléculaire qui serait responsable de l'encapsulation de l'ADN viral et un autre très basique, riche en arginine et de faible poids moléculaire qui servirait de récepteur pour l'activation de la protéine kinase (Tweeten *et al.*, 1978; Burley *et al.*, 1982; Tweeten *et al.*, 1980). Cette activation a pour effet de phosphoryler la protéine basique, ce qui provoque en conséquence la libération de l'ADN à partir de la capsid dans le nucléoplasme.

Les GV ont un ADN double brin superenroulé avec une taille située entre 115 à 135 kilopaires de bases (kpb). Le génome du ChfugV mesure environ 107 kpb (Arif *et al.*, 1986). Cependant, il existe des exceptions où la taille de l'ADN génomique peut atteindre 178 kpb, comme c'est le cas chez le GV de *Xestia c-nigrum*, XcGV (Goto *et al.*, 1998). L'ADN viral, est replié sur lui même dans la capsid (Miller *et al.*, 1990; Summers et Anderson, 1972).

L'expression des gènes et les mécanismes moléculaires impliqués dans la réplication virale chez les GV sont encore mal connus, principalement à cause du manque de lignées cellulaires permissives à ces virus (Crook *et al.*, 1997; Winstanley et Crook, 1993; Dwyer et Granados, 1988). Cependant, la disponibilité de ces dernières dans le cas des NPV, a permis d'élaborer des hypothèses sur le mode de réplication des baculovirus.

Infection et réplication

Infection

Dépendant du type d'infection on distingue deux phénotypes de virions chez les granulovirus comme chez les nucléopolyédrovirus, ceux avec un corps d'inclusion ou granules (OV "occluded virus") et ceux issus du bourgeonnement cellulaire et qui deviennent des virus libres ou extracellulaires (NOV "non occluded virus") (Blissard et Rohrmann, 1990). Les virions avec corps d'inclusion acquièrent leur enveloppe à l'intérieur du noyau. Ils sont impliqués dans la transmission horizontale et sont responsables de l'infection primaire chez les larves d'insectes (Rohrmann, 1992). Les virus extracellulaires se distinguent des virus avec corps d'inclusion à partir des variations dans la composition des protéines membranaires.

Après ingestion, les GV infectent l'insecte à travers les cellules épithéliales de l'intestin. Chez les GV comme chez les NPV, les corps d'inclusion ou granules se dissolvent au niveau de l'intestin moyen de la larve dans un milieu alcalin à pH supérieur à 10 et probablement aussi sous l'action de certaines protéases associées au corps d'inclusion (Adams et Bonami, 1991). Les polyèdres qui constituent le corps d'inclusion chez les NPV présentent plusieurs caractères communs avec les granules des GV. En général, les corps d'inclusion sont résistants aux conditions atmosphériques et peuvent persister dans la nature d'où leur fonction de protection du virus.

Réplication

Chez les baculovirus, l'expression des gènes et la réplication de l'ADN génomique sont soumises à une cascade de régulation ou d'évènements où chaque phase successive est dépendante de la phase précédente (Dobos et Cochran, 1980). L'expression des gènes se divise en deux grandes étapes, la phase précoce qui précède la réplication virale et la phase tardive qui a lieu pendant ou après la réplication.

Apparemment, la réplication chez les GV commencerait à l'entrée du noyau, au niveau des pores ou dans le noyau. À la suite, le stroma virogène se forme et des éléments en forme de tubes apparaissent (possiblement la capsidie qui s'associe au stroma virogène), puis se transforment en nucléocapsides. Au fur et à mesure que la réplication progresse, la nucléocapsidie déjà synthétisée dans le noyau migre vers le cytoplasme où elle bourgeonnera à travers la membrane plasmique où il y aura formation de l'enveloppe

du virus (Granados, 1978). Chez les GV, la réplication virale provoque une désintégration précoce du noyau et se fait en même temps dans le noyau et dans le cytoplasme. Tandis que chez les NPV, la réplication virale se fait dans le noyau sans que ce dernier soit endommagé (Crook *et al.*, 1997).

Phase précoce

La phase précoce est subdivisée en deux, la phase précoce immédiate (IE) durant laquelle, la transcription virale peut avoir lieu dans une cellule non infectée sans l'aide d'aucune protéine virale et la phase tardive immédiate qui nécessite la présence des protéines virales. Les gènes précoces sont exprimés de façon séquentielle, par exemple les transcrits de certains gènes précoces apparaissent une heure suivant l'infection et déclinent rapidement à la sixième heure suivant l'infection, alors que les transcrits d'autres gènes précoces s'accumulent avec le temps (Blissard et Rohrmann, 1990). Chez les NPV, six heures après infection, le virus exprime les protéines nécessaires pour sa réplication comme, l'ADN polymérase (*dnapol*), l'ADN hélicase (*p143*) et une protéine homologue de l'antigène nucléaire pour la prolifération cellulaire (*etl*). Pendant cette période, d'autres protéines impliquées dans la régulation de la transcription des gènes précoces, *ie-1*, *ie-2*, *pe-38* sont exprimées ainsi que les facteurs d'expression des protéines tardives (*lefs*) (Blissard et Rohrmann, 1990; Friesen et Miller, 1986).

En général tous les gènes précoces, *IE* portent le motif du doigt de zinc, $CX_2C..CXHX_{2,3}CX_2C...CX_2C$ (Miller, 1996). La plupart des promoteurs des gènes précoces chez les baculovirus possède des boîtes TATA situées entre 25-30 paires de bases (pb) en amont du site d'initiation de la transcription, CAGT. Certains promoteurs précoces comme ceux des gènes *p35*, *gp64* et *p39* des NPV, peuvent chevaucher des promoteurs tardifs, ce qui fait que ces gènes peuvent avoir à la fois des transcrits précoces et tardifs (Toister-Achituv et Faktor, 1997).

Durant les phases précoces et suite à l'infection cellulaire, l'environnement intérieure et extérieure de la cellule est assujettie à plusieurs changements physiologiques. La structure cytosquelettique des cellules infectées change de façon dramatique, les parois cellulaires et le noyau s'hypertrophient. Ce phénomène pourrait être responsable de l'apoptose ou mort programmée de la cellule (Crook *et al.*, 1993).

Comme indiqué précédemment, après être synthétisées, les nucléocapsides peuvent rester dans le noyau où elles acquièrent une enveloppe *de novo* et ensuite un corps d'inclusion. En général les virions avec un corps d'inclusion peuvent être observés dans le noyau dans les 24 heures suivant l'infection (Blissard et Rohrmann, 1990). D'un autre côté, les nucléocapsides peuvent être transportées vers le cytoplasme où, après bourgeonnement à travers la membrane plasmique, elles acquièrent une nouvelle enveloppe. Les nouveaux virions issus de ce bourgeonnement sont de phénotypes NOV, responsables des infections secondaires. Les virions infectieux de phénotype OV pénètrent dans la cellule hôte par fusion de l'enveloppe virale avec les microvillosités des cellules épithéliales de l'intestin moyen et le cycle de réplication recommence.

Environ huit heures après l'infection, on peut déjà observer dans le noyau, des nucléocapsides issues de cette première série de réplication (Blissard et Rohrmann, 1990).

Des expériences effectuées *in vitro* ont montré que les gènes précoces des baculovirus sont transcrits à partir de l'ARN polymérase II de la cellule hôte, sensible à l'antibiotique alpha-amanitine et mieux, les promoteurs de ces gènes sont efficacement transcrits en présence d'un inhibiteur de la synthèse de protéine, la cycloheximide (Rice et Miller, 1986; Huh et Weaver, 1990a et 1990b).

Phase tardive

La phase tardive pourrait aussi être subdivisée en deux, la phase tardive simple et la phase très tardive. La phase tardive simple implique les gènes qui sont exprimés en même temps que se fait la synthèse de l'ADN viral. Ces gènes sont transcrits par l'ARN polymérase de la cellule hôte à partir de la séquence consensus, A/GTAAG des promoteurs tardifs des baculovirus (Morris et Miller, 1994). En général, la transcription commence à partir de cette séquence. Cependant on a aussi découvert qu'une nouvelle ARN polymérase exprimée par le virus serait responsable de la transcription des gènes tardifs (Fuchs *et al.*, 1983; Huh et Weaver, 1990a,b). La quantité d'ARN messager de ces gènes décline au fil du temps.

Chez les NPV, la phase tardive commence environ six heures après l'infection et continue jusqu'à 24 heures après l'infection. Durant cette phase, ce sont les protéines structurales de phénotypes NOV qui sont synthétisées en abondance, comme p6.9K, VP39 et gp64 (Blissard et Rohrmann, 1989; Jarvis *et al.*, 1994a); des protéines du stroma

virogène, 39K, sont aussi synthétisées. Les gènes *lef*, *dnapol* et *p143*, impliqués dans la réplication de l'ADN viral, et le gène *p47*, impliqué dans l'expression des gènes tardifs, font tous partie du groupe des gènes tardifs (Arhens et Rohrmann, 1996; Lu et Miller, 1994; Lu et Carsten, 1991). La phase très tardive de l'infection est caractérisée par une très forte expression des gènes tardifs comme le gène de la polyédrine, *polh* chez les NPV. Le taux d'ARN messagers de ces gènes reste élevé même tard après l'infection virale. D'ailleurs, la formation du corps d'inclusion des virions commence environ 18 à 24 heures suivant l'infection et peut durer environ 76 heures jusqu'à la lyse des cellules. On remarque aussi la présence d'un gène hyperexprimé, le gène *p10*, impliqué dans l'assemblage des microtubules (Vanoers *et al.*, 1992; Leisy *et al.*, 1986; Rohrmann, 1986).

1.3.2. Organisation des gènes sur le génome

Des études rapportées sur l'organisation génomique dans la région immédiate en amont du gène de la granuline montrent que l'arrangement des gènes dans cette région semble être conservé chez les GV (Jehle *et al.*, 1994; Crook *et al.*, 1997; Bah *et al.*, 1999), contrairement à la même région chez la polyédrine des NPV (Zanotto *et al.*, 1992). Cependant, la région en aval du gène de la granuline ne serait pas identique chez tous les GV (Bah *et al.*, soumis pour publication).

Chez le CpGV et le TnGV, on a identifié la présence de certaines régions du génome avec des séquences répétitives homologues appelées *ih*s (Hashimoto *et al.*, 1996). Chez les NPV, ces régions sont appelées *hr*s et elles ont été identifiées et séquencées chez plusieurs MNPV (Cochran et Faulkner, 1983; Arif et Doerfler, 1984; Pearson *et al.*, 1992; Pearson et Rohrmann, 1995). Ces régions seraient impliquées dans l'expression et faciliteraient la transcription des gènes précoces. Elles sont aussi considérées comme l'origine de la réplication de l'ADN chez bon nombre de baculovirus (Ayers *et al.*, 1994).

Les cadres de lecture (ORF) des baculovirus sont très rapprochés les uns des autres et se chevauchent avec des espacements de séquence d'ADN d'environ 2 à 200 pb à ce niveau (Miller, 1996). Selon Miller (1996), les régions promotrices et les signaux d'arrêt des ORF sont localisés en général au niveau des chevauchements. Très souvent, le codon d'arrêt de la traduction, UAA, chevauche le premier signal de polyadénylation, AAUAAA. Il est aussi très fréquent de voir que les transcrits d'un gène commencent à

partir du gène voisin immédiat et dans certains cas, atteignent ou dépassent même l'autre gène adjacent au gène voisin. Ce dernier cas est beaucoup plus accentué au niveau des gènes tardifs et très tardifs. Il existe aussi des situations où les gènes sont transcrits sous forme bicistronique ou multicistronique, mais on pense qu'il n'y a pas de translation dans les régions terminales de ces ORF. Deux extrémités 5' ou 3' peuvent se chevaucher sur le même gène tandis que leurs ARN sont transcrits à partir du brin d'ADN complémentaire. C'est pourquoi, le génome d'un baculovirus qui mesure environ 133 kpb pourrait facilement contenir 150 gènes de longueur moyenne (Miller, 1996).

Chez les NPV, les gènes qui ont des fonctions similaires ou qui se ressemblent sont dispersés de façon aléatoire sur le génome (Kool et Vlak, 1993; Kool *et al.*, 1993). Il s'agirait surtout des gènes qui codent pour les protéines de structure du virion et des gènes tardifs codant pour des facteurs d'expression. Toutefois, il est évident que la dispersion de différents transcrits régulés de façon temporelle à travers le génome, pourrait avoir une signification fonctionnelle dans certains cas. Par exemple, le transcrit du gène de la polyédrine se fait dans le sens opposé aux gènes tardifs qu'il chevauche et l'activation de son promoteur a une régulation négative sur le niveau des transcrits de ces gènes. Kool et collaborateurs (1993) ont aussi remarqué que les gènes précoces, tardifs et très tardifs sont dispersés à travers tout le génome. Ces auteurs précisent cependant, qu'il existe des exceptions, particulièrement dans le cas de certains gènes qui joueraient un rôle dans la régulation du groupe des gènes précoces *ie* et du gène *pe-38*. Ces gènes sont regroupés ensemble dans des compartiments spécifiques du génome et comptent pour environ 5% du génome total.

1.3.3. Études rapportées sur le génome du GV

Comme il a déjà été souligné précédemment, le manque de lignées cellulaires susceptibles aux GV constitue un sérieux handicap quant à l'avancement des connaissances de la biologie moléculaire de ce virus. À ce jour, il existe une seule lignée cellulaire fiable susceptible au CpGV (Crook *et al.*, 1997). Malgré ces difficultés, certaines études ont pu être effectuées sur quelques gènes de GV à savoir, le gène de la granuline (Akiyoshi *et al.*, 1985; Chakerian *et al.*, 1985; Jehle et Backhaus, 1994; Goto *et al.*, 1998), le gène inhibiteur de l'apoptose, *iap* (Crook *et al.*, 1993), les gènes régulateurs

d'ADN (Jehle et Backhaus, 1994; Bah *et al.*, 1999; Crook *et al.*, 1997), le gène de l'ecdysteroid UDP-glucosyl transferase, *egt* (O'Reilly et Miller, 1989a).

Dans l'introduction, il a été mentionné que la granuline est hautement conservée chez les GV avec plus de 75% d'homologie en acides aminés et en séquences nucléotidiques. De ce fait, la granuline et la polyédrine sont souvent utilisées pour établir les liens phylogéniques chez les baculovirus dans le but de bien comprendre leur évolution (Zanotto *et al.*, 1993). Le gène de la granuline est aussi utilisé comme point de départ, pour établir une carte de restriction physique.

D'autres gènes ont fait l'objet d'études. Il s'agit du gène qui code pour le facteur de croissance de l'infection virale "VEF" chez le TnGV (Hashimoto *et al.*, 1991) et le GV de *Pseudaletia unipuncta*, PuGV (Roelvink *et al.*, 1995); le gène *iap* du GV de *Cydia pomonella*, qui bloque l'apoptose ou la mort programmée de la cellule (Crook *et al.*, 1993); le gène *egt* du GV de *Lacanobia oleotreta* (O'Reilly et Miller, 1989a), qui code pour une hormone, l'ecdysteroid UDP-glucosyl transferase (EGT), capable de bloquer la mue chez l'insecte, permettant au virus de se répliquer de façon plus abondante. La délétion de ce gène entraînerait aussi la mort prématurée de la larve (O'Reilly et Miller, 1989b).

Les études des profils de restriction de certains GV ont montré que dans la plupart des cas, les génomes de GV partagent une forte homologie, comme c'est le cas chez le CpGV, le XcGV et le ChfuGV; le GV de *C. occidentalis*, ChooGV et le GV de *C. retiniana*, ChreGV (Crook *et al.*, 1985; Arif *et al.*, 1986; Goto *et al.*, 1992).

Quelques cartes de restrictions physiques ont été établies pour les génomes des GV de *P. rapae* (Dwyer et Granados, 1987), *X. c-nigrum*, XcGV (Goto *et al.*, 1992), de *C. pomonella* (Crook *et al.*, 1985) et de *T. ni* (Hashimoto *et al.*, 1996).

Des études de comparaison des profils de restriction de différents génomes de GV ont aussi été rapportées. Pour vérifier la stabilité d'un virus ou chez un groupe de virus phylogéniquement proches, on peut, après digestion enzymatique des génomes, comparer leurs profils de restriction. En utilisant cette méthode, on a pu comparer des génomes entiers de baculovirus qui, selon les cas, sont phylogéniquement proches ou distants. Par exemple, les profils de restrictions génomiques des GV de *C. fumiferana*, *C. occidentalis*, et *C. retiniana*, (Arif *et al.*, 1986) sont très semblables avec 97,7%

d'homologie. Cette homologie est d'autant plus importante si l'on considère que ces hôtes sont phylogénétiquement très proches. Les profils de restriction de CpGV en provenance du Mexique, de l'Angleterre et de la Russie, présentent aussi un fort pourcentage d'homologie (Crook *et al.*, 1985). Les génomes de GV de *Autographa gamma* (L.), *Hydraecia amurensis*, *Celaena leucostigma*, *Aletia pallens*, *Pseudaletia separata* et *Xestia c-nigrum* (L.) sont aussi très homologues (Goto *et al.*, 1992). Selon cet auteur, ces GV peuvent utiliser plusieurs hôtes alternatifs et seraient des variants d'un même ancêtre. En général, chez les MNPV, un isolat peut avoir plusieurs variants génotypiques comme c'est le cas chez *Heliothis zea*, *Mamestra brassicae* et *Autographa californica*, dont on a pu identifier plusieurs de ces variants (Hughes *et al.*, 1993). La stabilité observée chez les GV constitue un critère important dans leur choix comme outil de lutte biologique.

1.4 Variations génotypiques chez les baculovirus

Les variations génotypiques chez les baculovirus peuvent être perçues comme des changements dans la séquence nucléotidique du génome virale suite à des passages sériés du virus en cultures cellulaires ou chez son hôte naturel ou alternatif. Guettig et McCarthy (1982), pensent que les variations génotypiques chez les NPV pourraient être influencées par l'appartenance géographique du virus ou bien par la nature de l'insecte hôte. La possibilité de recombinaison entre un virus chez son hôte naturel avec un autre virus latent (Hughes *et al.*, 1993) serait une des raisons des variations génotypiques chez les baculovirus. Smith et Crook (1993) ont rapporté que la présence du GV de *Artogeia rapae*, ArGV, chez les larves de *Pieris brassicae* pourrait déclencher l'activité pathogène d'un autre GV en latence chez celles-ci. Chez les NPV, on a observé que dans plusieurs cas, le profil de restriction génomique d'un virus après passage chez un hôte alternatif ne ressemble plus à celui du virus initial. Ceci laisse croire à la possibilité d'une recombinaison du virus hétérologue avec le virus homologue de l'hôte (Kumar et Miller, 1987).

1.4.1 Techniques utilisées pour détecter des mutations dans les régions variables d'ADN génomiques

Pour augmenter les chances de détection des variations génomiques comme les insertions, délétions ou substitutions de simple bases de nucléotides, plusieurs techniques de biologie moléculaire ont été mises au point parmi lesquelles, le RFLP "Restriction

fragment length polymorphism" qui peut donner des informations sur la nature et les variations entre les séquences de fragments d'ADN générés à l'aide d'enzymes de restriction spécifiques; le RAPD "Random amplified polymorphic DNA", qui utilise des amorces aléatoires pour amplifier par PCR les régions polymorphiques de grands génomes. Cette technique est souvent utilisée pour développer des cartes génétiques et des marqueurs moléculaires dans les populations d'insectes (Hoy, 1994); le PCR-SSCP "Polymerase chain reaction-single strand conformation polymorphism", permet de distinguer des fragments d'ADN mutés et non mutés à partir de leur profil de migration sur gel de polyacrylamide en conditions dénaturantes (Sheffield *et al.*, 1992; Sheffield *et al.*, 1990; Oritta *et al.*, 1989); le CFLScan "Cleavage fragment length polymorphism scanning" est une des techniques de clivages chimiques. Quand une molécule d'ADN double brin est portée à des températures élevées, elle se dissocie en deux brins complémentaires. Lorsque la température diminue progressivement, les brins adoptent des structures secondaires qui peuvent varier selon que l'ADN soit de type sauvage ou muté. Par clivage enzymatique on peut donc distinguer l'ADN mutant du non mutant (Myers *et al.*, 1985).

L'une des techniques les plus récentes utilisée pour détecter le polymorphisme ou les variations génomiques est celle du "DCode system" développée par Bio-Rad (1996). Par rapport aux autres techniques, elle est efficace, fiable, sensible, simple d'application, rapide et les résultats obtenus sont faciles à reproduire. C'est un instrument de diagnostic qui permet d'utiliser plusieurs dizaines d'échantillons à la fois. Le "DCode system" permet de détecter des mutations sur des fragments d'ADN de 100 à 1000 pb, la présence de nouveaux gènes et des cas de polymorphisme. Il peut aussi servir à sélectionner des marqueurs génétiques. Le "DCode system" a été utilisé avec succès pour détecter des mutations génétiques responsables de plusieurs maladies, entre autre, le cancer, la dystrophie musculaire, le SIDA, la fibrose kystique.

L'utilisation du "DCode system" se résume en trois parties: a) la détection de la mutation en gel de polyacrylamide en gradient dénaturant perpendiculaire; b) l'électrophorèse des fragments d'ADN homoduplexes et hétéroduplexes en gel de polyacrylamide en gradient parallèle de dénaturant; c) l'électrophorèse des fragments

d'ADN homoduplexes et hétéroduplexes en gel de polyacrylamide en gradient constant de dénaturant.

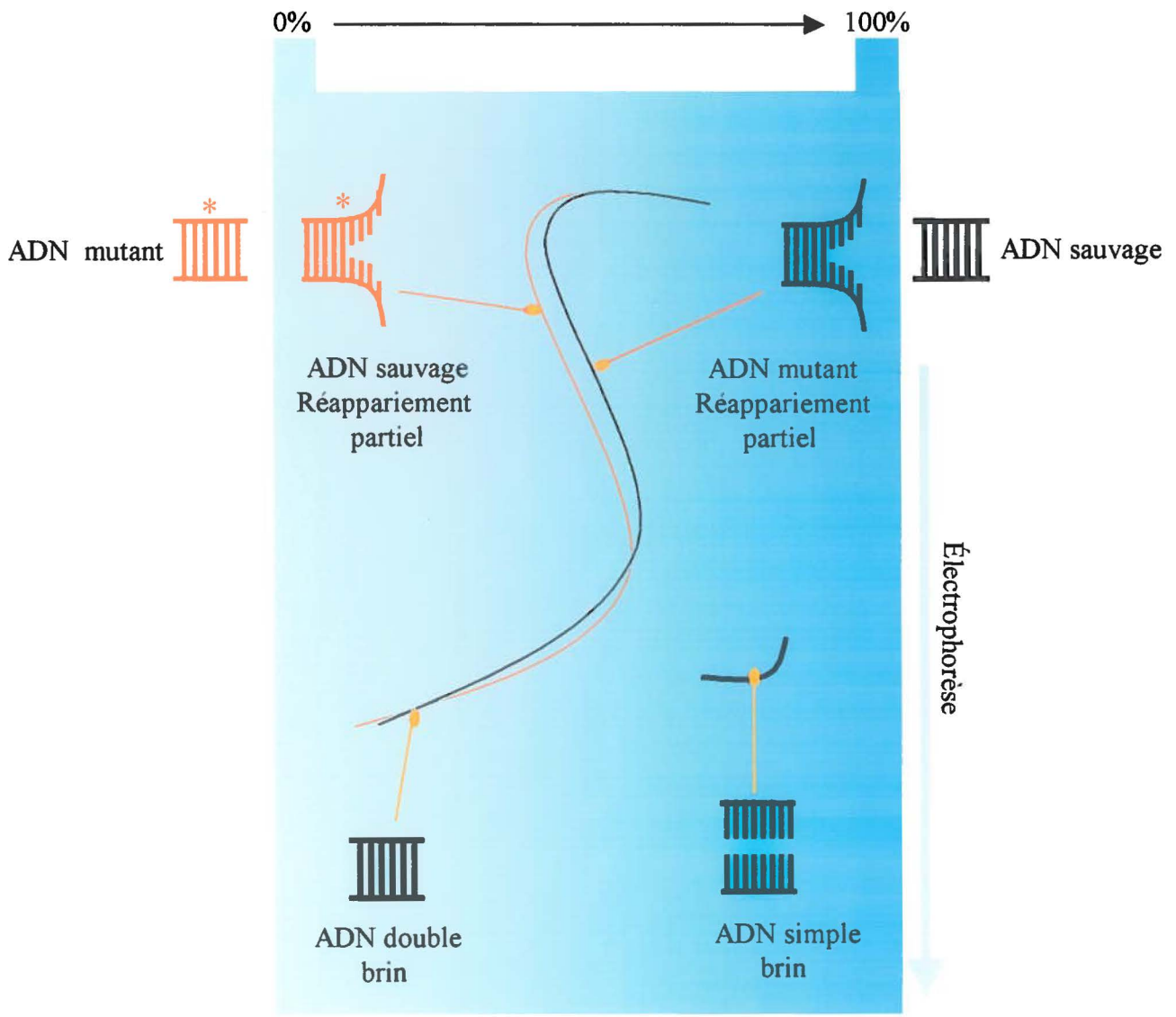
Détection des mutations en gel de polyacrylamide en gradient dénaturant perpendiculaire

Après avoir isolé l'ADN de type sauvage et celui du mutant, les régions qui sont identifiées variables sont d'abord amplifiées par PCR et les deux amplicons sont ensuite mélangés pour former des homoduplexes et des hétéroduplexes. Lorsqu'on fait migrer les amplicons homoduplexes et hétéroduplexes sur gel de polyacrylamide en sens perpendiculaire au gradient de dénaturant, dépendant du degré de dissociation de chaque brin d'ADN et de leur conformation, on peut observer une différence dans leur profil de migration. Les fragments d'ADN, en fonction de leur taille, migrent dans le gel dénaturant jusqu'à leur point de dénaturation où ils se dissocient, ralentissent leur mouvement et finissent par s'immobiliser. Dépendant de l'emplacement des deux extrémités du fragment d'ADN dans le gradient de dénaturant, celles qui se trouvent dans la région du gel où la concentration de dénaturant est la plus élevée se dissocient et migrent moins vite par rapport à l'extrémité opposée. C'est pourquoi à la fin de la migration d'un fragment d'ADN hétéroduplexe, on peut observer une bande d'ADN ressemblant à une lettre S avec une séparation nette des deux brins au milieu du S, le premier migrant plus vite que le second (Fig. 1).

Électrophorèse des fragments d'ADN homoduplexes et hétéroduplexes en gel de polyacrylamide en gradient parallèle de dénaturant

Une fois les mutations détectées, on peut faire migrer les fragments d'ADN homoduplexes et hétéroduplexes sur gradient parallèle de dénaturant où cette fois-ci, la direction de l'électrophorèse est parallèle au gradient de dénaturant. Sachant qu'en conditions dénaturantes la vitesse de migration de l'ADN est fonction de son point de dénaturation, une fois que les fragments hétéroduplexes et homoduplexes atteignent leur niveau maximal de dénaturation, ils se stabilisent à ce point. C'est pourquoi, à ce niveau, on peut nettement distinguer l'ADN hétéroduplexe qui dans les conditions dénaturantes migre plus lentement par rapport à l'ADN homoduplexe (Fig. 2).

Figure 1: Représentation schématique d'un gel de gradient perpendiculaire en conditions dénaturantes. À des concentrations faibles de dénaturants, le fragment d'ADN demeure double brin. Au fur et à mesure que la concentration de dénaturants augmente, le fragment d'ADN commence à se dissocier. À des concentrations plus élevées de dénaturants, le fragment d'ADN peut complètement se dissocier en deux brins séparés.



Électrophorèse des fragments d'ADN homoduplexes et hétéroduplexes en gel de polyacrylamide en gradient constant de dénaturant

Pour faciliter le diagnostic des régions variables à analyser chez plusieurs échantillons, après un gradient parallèle de dénaturant, on peut déterminer exactement le pourcentage de dénaturant au niveau duquel les fragments d'ADN homoduplexes et hétéroduplexes se dénaturent. Dans un gel de polyacrylamide en gradient constant, l'ADN se dénature dès que la migration commence et dépendant de la dimension du fragment et de la conformation de l'ADN, la migration continue jusqu'à un point où elle finit par se stabiliser.

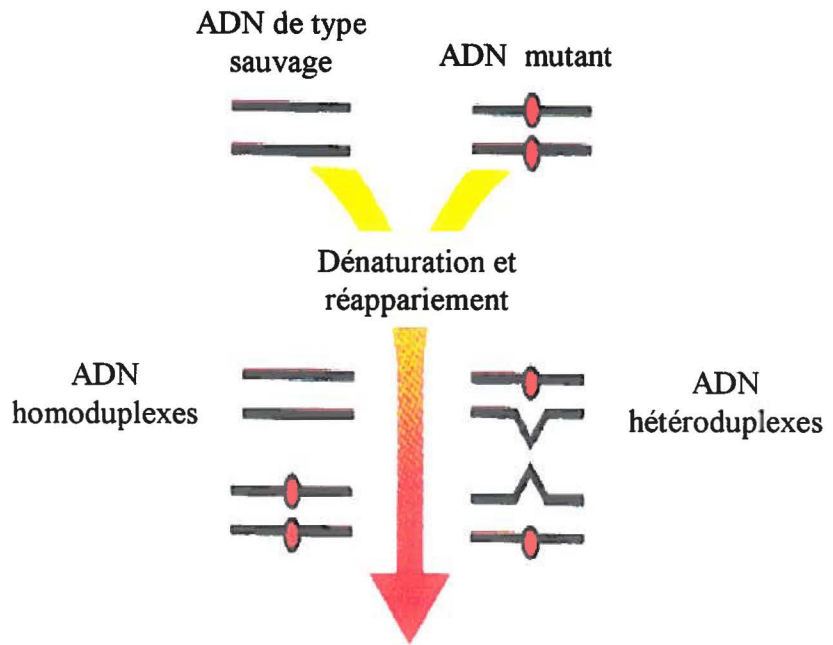
1.5 Conclusion

Les granulovirus sont des virus d'insectes qui sont considérés en général comme des agents potentiels pouvant être utilisés en lutte biologique contre certains insectes défoliateurs d'importance économique. La connaissance de la biologie moléculaire de ces virus est essentielle à la compréhension des processus et mécanismes associés à leur pathologie chez les insectes. Parmi les gènes de GV étudiés, c'est le gène de la granuline qui a retenu le plus d'attention. Le gène de la granuline est très conservé chez les GV et il joue un rôle très important dans la protection et le pouvoir infectieux du virion.

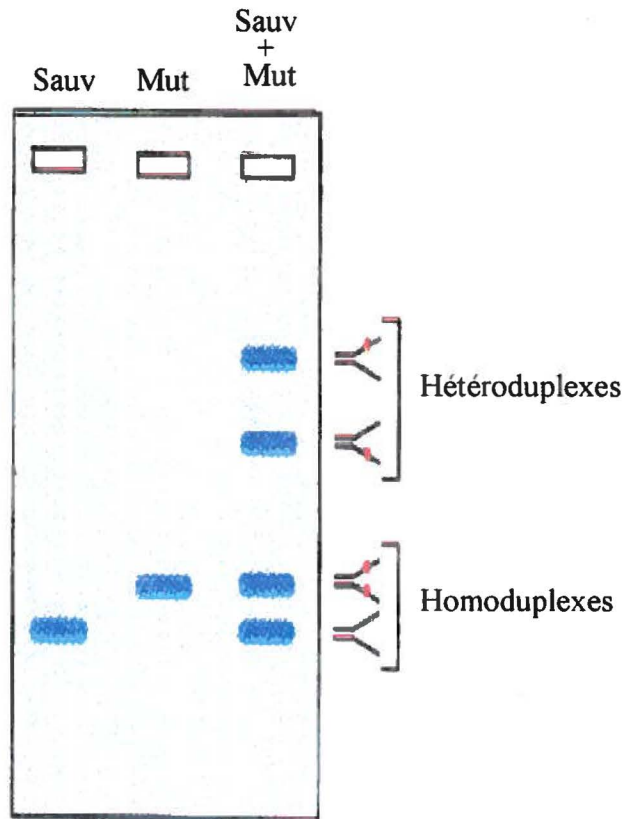
L'arrangement des gènes en amont du gène de la granuline chez les GV semble être identique. Mais avant ces travaux, aucune étude n'avait encore été rapportée sur l'organisation des gènes en aval de la granuline chez les GV.

Le but de cette recherche est d'étudier de façon plus précise le gène de la granuline du GV de *C. fumiferana*, ChfuGV et ses régions avoisinantes. Les résultats obtenus nous permettront d'améliorer notre compréhension sur l'organisation génomique de la région du gène de la granuline des GV et d'étendre nos connaissances sur la biologie moléculaire de ce virus.

Figure 2: Représentation schématique d'un gel de gradient parallèle ou constant en conditions dénaturantes. Après dénaturation, des fragments d'ADN sauvage et mutant sont mélangés ensemble pour générer quatre fragments dont deux hétéroduplexes et deux homoduplexes. Les propriétés de dissociation des fragments d'ADN hétéroduplexes sont modifiées de sorte qu'ils se dissocient plus rapidement que les fragments d'ADN homoduplexes dans un gradient de dénaturant de faible concentration.



Gel dénaturant de gradient parallèle ou constant



CHAPITRE 2

IDENTIFICATION ET ANALYSES DE LA SÉQUENCE DU GÈNE DE LA GRANULINE DU GRANULOVIRUS DE *Choristoneura fumiferana*, ChfuGV

**Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana*
granulovirus**

RÉSUMÉ

Le gène de la granuline du granulovirus de *Choristoneura fumiferana*, ChfuGV a été séquencé. Il code pour une protéine de 248 acides aminés avec une masse moléculaire estimée à 29,299 kDa. Les homologies de séquences entre le gène de la granuline du ChfuGV et celui des granulovirus de *Trichoplusia ni* (Hbn.), *Pieris brassicae* (L.) et *Cryptophlebia leucotreta* se situent entre 76,7 à 80,5% pour les nucléotides et 84,2 à 88,3% pour les acides aminés. La structure secondaire de la granuline du ChfuGV a été estimée y compris ses régions hydrophiliques (polaires) et hydrophobiques (basiques). Les résultats observés nous montrent que ces structures sont similaires à celles d'autres granulines. Le motif ATAAG, caractéristique des promoteurs très tardifs des baculovirus, a été localisé dans la région promotrice putative du gène de la granuline du ChfuGV.

ABSTRACT

The nucleotide sequence of the granulin gene of *Choristoneura fumiferana* granulovirus (ChfuGV) was determined. The gene encodes a protein of 248 amino acids with a predicted Mr of 29.299 kDa. The granulin genes of *Trichoplusia ni* (Hbn.), *Pieris brassicae* (L.) and *Cryptophlebia leucotreta* granuloviruses showed homologies ranging from 76.7 to 80.5% for nucleotide sequences and 84.2 to 88.3% for amino acid sequences when compared to ChfuGV. The secondary structure of CfGV granulin protein, including the hydrophilic (polar) and hydrophobic (basic) regions, was predicted and found to be similar to other granulins. A very late baculovirus promoter motif, ATAAG, was found within the putative promoter region of the ChfuGV granulin gene.

Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus*

A. Bah¹, J. Bergeron¹, M. Arella¹, C. J. Lucarotti², and C. Guertin¹

¹Institut Armand-Frappier, Centre de Recherche en Virologie, Laval, Québec, Canada

²Canadian Forest Service-Atlantic Forestry Centre, Fredericton,
New Brunswick, Canada

Accepted March 18, 1997

Summary. The nucleotide sequence of the granulin gene of *Choristoneura fumiferana* granulovirus (CfGV) was determined. The gene encodes a protein of 248 amino acids with a predicted Mr of 29.299 kDa. The granulin genes of *Trichoplusia ni*, *Pieris brassicae* and *Cryptophlebia leucotreta* granuloviruses showed homologies ranging from 76.7–80.5 % for nucleotide sequences and 84.2–88.3 % for amino acid sequences when compared to CfGV. The secondary structure of CfGV granulin protein, including the hydrophilic (polar) and hydrophobic (basic) regions, was predicted and found to be similar to other granulins. A very late baculovirus promoter motif, ATAAG, was found within the putative promoter region of the CfGV granulin gene.

Introduction

The spruce budworm, *Choristoneura fumiferana* (Clem.) (Lepidoptera: Tortricidae), feeds on coniferous trees in eastern North America causing serious defoliation resulting in significant economic damage. This insect is susceptible to a granulovirus (GV) [3–5] and GVs, generally, are being considered as agents for potential use in the biological control and regulation of pest insect populations in agriculture and forestry.

The GVs are members of genus *Granulovirus* in the family *Baculoviridae* [17]. The rod-shaped virions of GVs are occluded and enveloped and contain a circular double stranded DNA that is usually larger than 100 kilobase pairs (kb) [1]. The most abundant protein of the encapsulating occlusion body is granulin which is produced in the late stages of viral infection. Granulin plays the same role as polyhedrin in nuclear polyhedrosis viruses (subgroup A) in protecting

*The sequence data described in this paper have been deposited with the GenBank database and assigned the accession number U87621.

viral DNA from the deleterious effects of ultraviolet radiation [14]. Arif et al. [3] have reported on the genomic restriction profiles and granulin proteins of GV's isolated from three *Choristoneura* species. However, the CfGV granulin gene sequence and the sequences regulating the transcription elements of this gene are not known. Here, we report on the nucleotide and amino acid sequence analyses of the CfGV granulin gene and its putative promoter region.

Materials and methods

Virus purification

The virus was amplified by feeding fourth-instar spruce budworm larvae CfGV. Moribund and dead larvae were harvested and homogenized in sodium dodecyl sulphate (SDS), final concentration of 0.3%. The homogenate was filtered through three layers of cheesecloth and viral occlusion bodies (OBs) were pelleted at 12 000 g for 30 min at 15 °C in a Sorvall SS-34 rotor. The pellet was washed twice with distilled water and resuspended to a final volume of 5 ml.

Extraction and purification of DNA

Viral OBs were further purified on a 30–80% (v/v) glycerol gradient (0.1% SDS) [8]. Proteinase K was added to purified OBs to a final concentration of 2 mg/ml and the mixture was incubated for 20 min at 37 °C. Stock solution of 10% sarkosyl was added to a final concentration of 2%. The suspension was incubated at 56 °C for 2 h then at 37 °C for 3–5 h [15]. DNA was extracted in an equal volume of phenol/chloroform followed by chloroform rinses to remove traces of phenol [20]. DNA was purified using cesium chloride-ethidium bromide gradient centrifugation [6]. The DNA was washed several times with butanol to remove the ethidium bromide [20] and dialyzed against several changes of 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 at 4 °C [21].

Amplification conditions and sequencing

Total protein was extracted from purified OBs and was subjected to SDS-polyacrylamide gel electrophoresis. A Mr 30 kDa protein corresponding to the expected size of the granulin was recovered following protein transfer, by electrophoresis, to PVDF membrane. From this, the N-terminal sequence of the protein was determined by automated Edman degradation [10]. The resulting amino acid sequence, GYNKSLXYSXXDGTSSVIDN, was used to deduce a primer. CfGVFW (5'-ATGGGATA(T/A)GC(G/T)GG(T/A)GT, specific to the 5'-end of the CfGV granulin gene. A second, putative reverse consensus primer, GVRE (5'-TTAATA(T/A)GC(G/T)GG(T/A)GT), was generated on the basis of published TnGV [2] and PbGV [7] granulin gene sequences. DNA was amplified by PCR in a reaction volume of 50 µl containing 5 µl of 10×PCR buffer (Pharmacia), 2 µl of 10 mM dNTP, 2 µl of each primer, 25 ng of CfGV genomic DNA, 0.5 µl of Tween-20 and 0.5 unit of Taq DNA polymerase (Bio/Can Scientific). Amplification of the granulin gene was performed in a Perkin Elmer Cetus DNA Thermal Cycler (model 9600) using a 35-cycle program, each cycle consisting of 1 min at 94 °C, 1 min at 45 °C, 1.5 min at 72 °C and a final extension cycle of 10 min at 72 °C. Amplification of the promoter region consisted of a 3-cycle program, of 1 min at 95 °C, 1 min at 40 °C, 1 min at 72 °C followed by 35-cycle program of 1 min at 94 °C, 1 min at 45 °C and 0.5 min at 72 °C. The PCR products were cloned in the plasmid pCRII (Invitrogen) and sequenced using an A.L.F. DNA sequencer (Pharmacia). CfGV sequences were analyzed and compared to equivalent

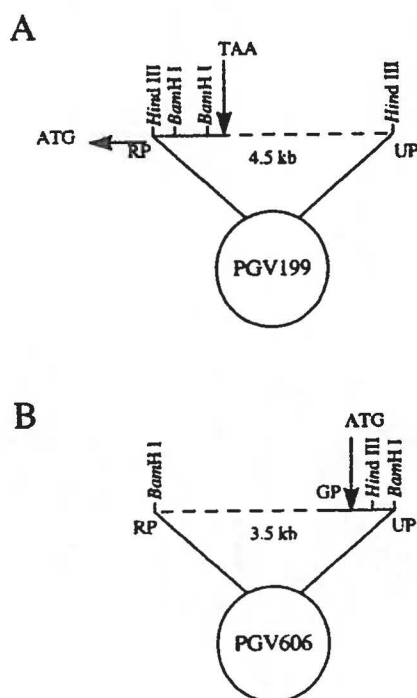


Fig. 1. Schematic representation of clones pGV199 and pGV606. **A** pGV199 contains a portion of 647 bp of the CfGV granulin gene coding region begin from the *Hind* III site to the TAA stop codon. Dashes represent the granulin gene downstream region. **B** pGV606 contains a second portion of 80 bp of the CfGV granulin gene coding region between the *Hind* III site and the ATG start codon. The putative promoter region is indicated by *GP* (granulin gene promoter). Dots represent the granulin gene upstream region. *UP* and *RP* indicate the universal and reverse primers respectively

sequences of CIGV [13], PbGV [7] and TnGV [2] using PCGENE 6.7 (Intelligenetics Inc.) and Mac Vector 4.5.0 (Eastman Kodak) softwares.

In order to localize the granulin gene and putative promoter region on the CfGV genome, a CfGV*Hind* III genomic library was constructed from purified viral DNA by shotgun cloning CfGV*Hind* III fragments into the cloning vector pUC19. A clone designated pGV199 hybridized strongly with a labeled PCR generated granulin gene probe. The 4.5 kb insert of this clone was sequenced as described above. One of the termini of this insert contained a 667 bp fragment of the CfGV granulin gene coding sequence. This sequence begins from a unique *Hind* III restriction site (located 80 bp downstream from the ATG start codon) to the TAA stop codon. Two *Bam*H I restriction sites were also identified within the 4.5 kb fragment. These were located at 190 bp and 476 bp downstream from the ATG start codon of the granulin gene (Fig. 1). In order to identify the granulin promoter region, another genomic library was constructed using the restriction enzyme *Bam*H I. An 80 bp fragment representing the deduced 5' end of the CfGV granulin gene was amplified with specific forward and reverse PCR-primers and was used as a probe to identify clone pGV606 that contained a 3.5 kb insert (Fig. 1). Sequencing the two termini of this insert enabled us to locate the putative granulin gene promoter region. This insert also contained the 80 bp sequence of the granulin gene 5' region. This sequence complemented the first 667 bp of the sequence found in clone pGV199.

Results

Analyses of the region upstream of the granulin gene open reading frame (ORF) indicated the presence of a consensus baculovirus late promoter ATAAG at position -28. Nucleotides -18 to 31 upstream of the ATG start codon were

```

          -30      -20      -10
CfGV TTTATAAGGAATTTAATTTAACAAGAACA--ATG
C1GV .....T..T..C..G.T-ATG
CpGV .....T..T..C..C.TCATG
PbGV .....T..C..A.T-ATG
TnGV .....TC.A..C.--ATG

```

Fig. 2. Multiple alignments of the 5' flanking sequence of the granulin ORF of CfGV, C1GV [13], CpGV [13], PbGV [7] and TnGV [2]. The invariable 14 nucleotides between the five GVs are double underlined. Conserved 6 nucleotides between CfGV, C1GV, CpGV and PbGV are in bold letters. Letters single underlined show the -3 position of each sequence. Dashes were introduced to optimize sequence alignment. Dots indicate the conserved nucleotides

Table 1. Percent identity of the 5' and 3' flanking regions of the granulin gene of CfGV compared to C1GV, PbGV and TnGV

Homology (%)	<i>Choristoneura fumiferana</i>	
	5' end	3' end
<i>Cryptophlebia leucotreta</i>	59.4	44.6
<i>Pieris brassicae</i>	66.7	46.4
<i>Trichoplusia ni</i>	53.6	42.0

perfectly homologous between the five GVs (Fig. 2). Nucleotides -12 to -17 were also found to be well conserved between CfGV, C1GV, CpGV and PbGV. For 138 bp upstream of the ATG translation initiation start site, the granulin gene sequence homology of CfGV compared to C1GV [13], PbGV [7] and TnGV [2] ranges between 53.6 and 66.7 % while for the 119 bp downstream of the TAA stop codons CfGV comparative homologies range between 42 and 46.4 % homology (Table 1). The 3' flanking region of the CfGV granulin gene contained some repetitive sequence motifs, ATTTTA and TTCTT, downstream of the putative stop codon (Fig. 3). Two stop codons, located 37 (TGA) and 208 (TAA) nucleotides downstream from the CfGV granulin gene TAA stop codon (Fig. 3), were also found (Fig. 3).

The ORF of the CfGV granulin gene is 747 bp long and codes for a predicted Mr 29.3 kDa protein. This size protein was confirmed by SDS-PAGE analysis of mature granulin and is the same predicted Mr as was obtained by Arif et al. [3]. Alignment of amino acid sequences of CfGV, C1GV, PbGV and TnGV granulin genes shows that regions from amino acids 55 to 80, 83 to 124, 126 to 144 and 197 to 248 are highly conserved (Fig. 4). However, the similarity decreases at amino acid positions 22 to 50, especially between CfGV and C1GV. Proline, cysteine and arginine residues are all conserved in the four granulin sequences. Basic (lysine, arginine) and acidic (asparagine, glycine) amino acid residues represent 12% and 13.7%, respectively, of the 248 amino acids in the granulin proteins of CfGV, PbGV and TnGV and 11.2% and 12% in C1GV granulin.

CfGV granulin gene

1581

-95

TACTGTTACAACAACTGTATGGAAGACGCAACCGAAAAGGTTATATGTTGTGCCAAC 60

-73 -50 -39 -28

ATAATGTTATAAACCCTGGATACTAGGAGAAACGAAAATGAAACGATTTTATAAGBAATT 120

TTAATTTAAACAAGAACAATGGGATATAACAAAGCATTAAAGATACAGTCGTCACGACGGC 180

M G Y N K A L R Y S R H D G

Hind III

ACCTCATGCGTTATTGACAACCACCACCTCAAAGCTTAGGTTCTGTGCTGAATGATGTT 240

T S C V I D N H H L K S L G S V L N D V

CGACACAAAAAGACCGCATACGTGAGGCTGAATACGAACCCATTCTAGATATTGCCAAT 300

R H K K D R I R E A E Y E P I L D I A N

*Bam*H III

CAGTACATGCTGACCGAGGATCCTTTTCGTGGACCCGGTAAGAACGTACGAATTACCCTT 360

Q Y M V T E D P F R G P G K N V R I T L

TTCAAGGAAATTAGGCGCGTTCAACCAGACACCATGAAATGGTATGCAACTGGAGCGGA 420

F K E I R R V Q P D T M K L V C W W S G

AAGGAATTCCTTCGCGAAAACATGGACACGCTTCATTCTGAAGAGTTTCCCATCACTACT 480

K E F L R E T W T R F I S E E F P I T T

GACCAAGAAATCATGGATTGTGGTTGAAATCAACTCAGACCAATGCAACCCAATCGC 540

D Q E I M D L W F E I Q L R P M Q P N R

TGTTACAAATTTACGATGCAATACGCGCTTGGTGTAAACCCGAATACGTTGCACATGAT 600

C Y K F T M Q Y A L G A N P E Y V A H D

*Bam*H I

GTAATCCGTC AACAGGATCCTTACTACGTTGGTCTCAACAACATGGAACGCATCAACTTG 660

V I R Q Q D P Y Y V G P N N M E R I M L

TCCAAGAAAGGATTCGCATTCGCGCTCACGTGCTTCAATCCGCTTACAATGACAACTTT 720

S K K G F A F P L T C L Q S V Y N D N F

GAGAGATTCCTFCGACGACATTCCTTTGGCCGATTTCCACCGTCTTTAGTATATGTGGGA 780

E R F F D D I L W P Y F H R P L V Y V G

ACCACCTCCGCGAAAATCGAAGAAATATGATCGAGGTGTCGCTATTGTTAAAAATAAA 840

T T S A E I E E I M I E V S L L F K I K

GAATTCGCACCAGATGTACCGCTTTTACTGGTCCCGCATATTAATCCTGTTTTCAGCG 900

E F A P D V P L F T G P A Y *

ACATTTTCGTCATAATTTTATTGTAGTATCAACAATTTTAGAATTTAAGTCTTCCATTTTA 960

TCACAAACTATTACGTTGATGTTATTTGGACCGCGCTTCGATTTGGGTAACGAAGAAGTA 1020

ATTGCAACGGAAGAATTTTTCGCAACATCAACCTTATTTGTTGGTGAATTCGAAAATTTTA 1080

ATAATCATATTATAAACGCTGTCCAAAACCTCAGTCAACTCTCCACCGTTAACGTTAAC 1140

ACGTTGTCGTCAGTTTCGGTGTATAATTTCTTTTAAAGCTATTTATTCACATTTCAATAAC 1200

TTTATTTGGGTCACGAACGTTTCAGCATTGTAATTTTACTTTTAA 1244

Fig. 3. The nucleotide sequence of CfGV granulin gene coding region including its 5' and 3' flanking regions. In the 5' flanking region, tandem repeat sequences are underscored with arrows or bold underlines. The baculovirus consensus late promoter is enclosed in the box. N-glycosylation sites are represented by amino acids in bold letters. In the 3' flanking region repetitive sequence patterns are underlined

Discussion

The first 14 nucleotides of the granulin promoter region, which contain the putative very late baculovirus promoter motif, ATAAG [18, 19], are identical in the five GVs studied to date [2, 7, 13] (Fig. 2). The granulin initiation codon, ATG, is in accordance with Kozak's [16] rule (A position -3; G position +4), and is located in the same position, only a few nucleotides from the ATAAG motif, in each GV. In *Autographa californica* (AcMNPV) and other MNPVs, the ATAAG motif is known to be associated with polyhedrin and p10 gene transcription promoter sites [11] and these late genes are transcribed by an α -amanitin-resistant RNA polymerase [9]. While this motif is speculated to serve a similar function in GV granulin late genes, it has yet to be demonstrated that this is the case. In the 3' flanking region of the TnGV granulin gene sequence, a potential polyadenylation signal (AATAAA) was found at 7 nucleotides beyond the termination codon [2]. As is the case with the ClGV granulin gene sequence [13], no similar polyadenylation signal sequence was found in the 3' flanking region of the CfGV granulin gene. Homologies between 76.7 and 88.3% were found when CfGV granulin nucleotide and amino acid sequences were compared to those of ClGV, PbGV and TnGV (Table 2). These results indicate that granulin is highly conserved between different GV species.

Multiple alignments of granulin amino acids of CfGV, ClGV, PbGV and TnGV showed that arginine, cysteine and proline residues are conserved in all cases (Fig. 4). This conservation is likely due to the role that these amino acids may play in granulin secondary structure. The presence of proline and arginine

CfGV	MGYNKALRYS	RHDGTSCVID	NHHLKSI	GEV	LMDVRHKKDR	IREAEYEPIL	50
ClGVS.....	..D.....	.H.....	EPC	YTMSDV..IA	SVKR.Y...I	50
PbGV	...RA....	K.E.....	.Q.Y....	A..K..K...DR	L....I..VL		50
TnGVS..._	..N..._	.K...T..	E..G..R..EEL	I...QFD..K		50
CfGV	DIANQYMVTE	DPFRGPGKNV	RITLPKEIRR	VQPDTMRLVC	NWSGKEFLRE		100
ClGV	...D.....H.....		100
PbGV	...D.....C..E..L...		100
TnGV	...N.....	K.....	IQ.....		100
CfGV	TWTRFISEEF	PITTDQEIMD	LWPEIQLRPM	QPNRCYKFTM	QYALGANPEY		150
ClGV	I...L....	H.....	...C.H...	150
PbGV	M.....	...Q..N	M...I.V...D.H...	150
TnGVL..._	..._A.N...	150
CfGV	VAHDVIRQQD	PYYVGPNNME	RINLSKKGFA	FPLTCLQSVY	NDNFERFFDD		200
ClGV	I.....	K.....	...M...	200
PbGV	.P...A...	...I...I...	...K...-	...M.....T..E...	199
TnGV_H...	..._D.R...	...R.L...	..._...I...	...E...E...Q		200
CfGV	ILWPYFHRPL	VYVGTSAEI	BEIMIEVSLI	FKIKEFAPDV	PLFTGPAY		248
ClGV		248
PbGVI...S.T	...LL..F.Y.....		247
TnGVV...A...		248

Fig. 4. Multiple alignment of CfGV, ClGV [13], PbGV [7] and TnGV [2] granulin amino acid sequences. Dots indicate homology of the three GV sequences with CfGV. The boxed sequences represent regions of high variability. Three conserved prolines, arginines and cysteines are underlined

Table 2. Percent identity of nucleotide and amino acid sequences of the granulin gene of CfGV, CIGV, PbGV and TnGV

Homology (%)	Sequence	<i>Choristoneura fumiferana</i>
<i>Cryptophlebia leucotreta</i>	Nucleotide	80.5
	Amino acid	88.3
<i>Pieris brassicae</i>	Nucleotide	77.6
	Amino acid	84.2
<i>Trichoplusia ni</i>	Nucleotide	76.7
	Amino acid	87.9

at conserved positions (Fig. 4) suggests similar flexibility in the granulins of CfGV, CIGV, PbGV and TnGV. The CfGV granulin amino acid sequence RHKK is highly basic and is located at position 35 through 38; the corresponding regions in granulins of CIGV, PbGV and TnGV are DVKK, KHKK and RKK, respectively (Fig. 4). Jarvis et al. [12] have identified a nuclear localization signal, KRKK in AcMNPV polyhedrin located between amino acids 32 and 35. It is plausible that this sequence may serve a similar function in GVs.

Comparative analyses of the CfGV granulin gene putative promoter region with other GVs indicate that this region is highly conserved and may be an important control element in the regulation of the granulin gene.

Acknowledgements

This project was supported by the Ministère des Ressources naturelles du Québec grant 0908-240S. CJL received research support through the Canadian Forest Service S&T Opportunities Fund. We thank Joanne Roger and Lise Forget for their technical assistance and Daniel Belkosky for the reviewing early drafts of the manuscript.

References

1. Adams JR, Bonami JR (1991) Atlas of invertebrate viruses, CRC Press, Boca Raton, pp 227-257
2. Akiyoshi D, Chakerian R, Rohrmann GF, Nesson MH, Beaudreau GS (1985) Cloning and sequencing of the granulin gene from *Trichoplusia ni* granulosis virus. *J Virol* 141: 328-332
3. Arif BM, Guangyu Z, Jamieson P (1986) A comparison of the three granulosis viruses from *Choristoneura* spp. *J Invertebr Pathol* 48: 180-186
4. Bird FT (1963) On the development of granulosis viruses. *J Insect Pathol* 5: 368-376
5. Bird FT (1976) Effects of mixed infections of two strains of granulosis virus of the spruce budworm, *Choristoneura fumiferana* (Lepidoptera: Tortricidae), on the formation of viral inclusion bodies. *Can Ent* 108: 865-871
6. Burand JP, Summers MD, Smith GE (1980) Transfection with baculovirus DNA. *J Virol* 101: 286-290.
7. Chakerian R, Rohrmann GF, Nesson MN, Leisy DJ, Beaudreau GS (1985) The nucleotide sequence of the *Pieris brassicae* granulosis virus granulin gene. *J Gen Virol* 66: 1263-1269

8. Crook NE (1981) A comparison of the Granulosis viruses from *Pieris brassicae* and *Pieris rapae*. *J Virol* 115: 173–181
9. Fuchs YL, Woods MS, Weaver RF (1983) Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *J Virol* 48: 641–646
10. Hewick RM, Hunkapiller MW, Hood LE, Dreyer WS (1981) A gas-liquid solid phase peptide and protein sequencer. *J Biol Chem* 256: 7990–7997
11. Hooft Van Iddekinge BJL, Smith GE, Summers MD (1983) Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus. *J Virol* 131: 561–565
12. Jarvis D, Bohlmeier DA, Garcia A (1991) Requirements for nuclear localization and supramolecular assembly of a baculovirus polyhedrin protein. *J Virol* 185: 795–810
13. Jehle JA, Backhaus H (1994) The granulin gene region of *Cryptophlebia leucotreta* granulosis virus: sequence analysis and phylogenetic considerations. *J Gen Virol* 75: 3667–3671
14. Jehle JA, Backhaus H, Fritsch E, Huber J (1992) Physical map of the *Cryptophlebia leucotreta* granulosis virus genome and its relationship to the genome of *Cydia pomonella* granulosis virus. *J Gen Virol* 73: 1621–1626
15. Kaupp WJ, Ebling PM (1993) Horseradish peroxidase-labelled probes and enhanced chemiluminescence to detect baculoviruses in gypsy moth and eastern spruce budworm larvae. *J Virol Methods* 44: 89–98
16. Kozak M (1986) Point mutations define a sequence flanking the ATG initiation codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283–292
17. Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds) (1995) *Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses*. Springer, Wien New York, pp 104–113 (*Arch Virol* [Suppl] 10)
18. Rohrmann GF (1986) Polyhedrin structure. *J Gen Virol* 67: 1499–1513
19. Rohrmann GF (1992) Baculovirus structural proteins. *J Gen Virol* 73: 749–761
20. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
21. Smith IL, Crook NE (1988) In vivo isolation of baculovirus genotypes. *J Virol* 166: 240–244

Authors' address: Dr. C. Guertin, Institut Armand-Frappier, Centre de recherche en Virologie, Laval, Québec, H7V 1B7, Canada.

Received August 15, 1996

CHAPITRE 3

LE GRANULOVIRUS DE *Choristoneura fumiferana*, ChfuGV: ANALYSES DE LA SÉQUENCE ET CARACTÉRISATION DE L'EXTRÉMITÉ 5' DU GÈNE ORF891

***Choristoneura fumiferana* granulovirus, ChfuGV: sequence analysis and 5' end
characterization of ORF891**

RÉSUMÉ

Le cadre de lecture, ORF891, a été identifié immédiatement en amont du gène de la granuline du granulovirus de *Choristoneura fumiferana*, ChfuGV. L'ORF891 code pour une protéine d'une masse moléculaire de 34,6 kDa. Le point d'initiation du transcrit de l'ARN messager de ce gène a été identifié et analysé. La région promotrice de l'ORF891 contient le motif, TAAG, caractéristique des promoteurs très tardifs des baculovirus et deux autres séquences, TATA reconnues comme étant des promoteurs précoces. Ces deux motifs de promoteurs sont similaires à ceux retrouvés dans la région promotrice du gène ORF909 du granulovirus de *Cryptophlebia leucotreta* (CIGV). Les comparaisons de séquences entre l'ORF891 du ChfuGV et les gènes ORF909 de CIGV et ORF124R du granulovirus de *Cydia pomonella* granulovirus (CpGV), ont montré des homologies respectives de 60,9 et 63,9% en nucléotides et de 46,3 et 49,3% en acides aminés. Les homologies en séquences nucléotidiques entre l'ORF891 et le gène ME53 des nucléopolyédrovirus de *Autographa californica* (AcMNPV) étaient de 68,2%. Mais ces deux gènes ne présentaient aucune homologie en acides aminés. Deux motifs de doigts de Zinc ont aussi été identifiés dans l'ORF891.

ABSTRACT

A gene located immediately upstream of the granulin gene of *Choristoneura fumiferana* (ChfuGV) granulovirus was identified, sequenced and named ORF891. The determined, putative open reading frame (ORF) of 891 bp encodes an estimated 34.6 Kda protein. The 5' end transcript of the gene was mapped and analysed. A putative promoter region organization of ChfuGV ORF891 contains a consensus late baculovirus promoter element, TAAG, and two putative early TATA boxes similar to the promoters of ORF909 of *Cryptophlebia leucotreta* granulovirus (CIGV). Sequence comparisons of ChfuGV ORF891 with CIGV ORF909 and *Cydia pomonella* granulovirus (CpGV) ORF124R showed respective homologies of 60.9 and 63.9% for nucleotides and 46.3 and 49.3% for amino acids. Homology of ChfuGV ORF891 with ME53 ORF of *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) was 68.2% for nucleotides but a total lack of homology for amino acid sequences. Two zinc finger motifs are also associated with ChfuGV ORF891.

***Choristoneura fumiferana* granulovirus: sequence analysis
and 5' characterization of ORF891***A. Bah¹, C. J. Lucarotti², M. Arella¹, and C. Guertin¹¹INRS-Institut Armand-Frappier, Centre de Recherche en Microbiologie-Biotechnologie,
Laval, Québec, Canada²Canadian Forest Service-Atlantic, Fredericton, New Brunswick, Canada

Accepted October 15, 1998

Summary. A gene located immediately upstream of the granulin gene of *Choristoneura fumiferana* (ChfuGV) granulovirus was identified, sequenced and named ORF891. The determined, putative open reading frame (ORF) of 891 bp encodes an estimated 34.6 kDa protein. The 5' end transcript of the gene was mapped and analysed. A putative promoter region organization of ChfuGV ORF891 contains a consensus late baculovirus promoter element, TAAG, and two putative early TATA boxes similar to the promoters of ORF909 of *Cryptophlebia leucotreta* granulovirus (CIGV). Sequence comparisons of ChfuGV ORF891 with CIGV ORF909 and *Cydia pomonella* granulovirus (CpGV) ORF124R showed respective homologies of 60.9 and 63.9% for nucleotides and 46.3 and 49.3% for amino acids. Homology of ChfuGV ORF891 with ME53 ORF of *Autographa californica* multicapsid nucleopolyhedrovirus (AcMNPV) was 68.2% for nucleotides but a total lack of homology for amino acid sequences. Two zinc finger motifs are also associated with ChfuGV ORF891.

Introduction

Granuloviruses (GV) are members of the family *Baculoviridae* [18]. They infect a specific range of host insects, mostly within the Lepidoptera or Hymenoptera [1]. Granuloviruses contain a circular double stranded DNA, greater than 100 kb, which is included within a nucleocapsid. Each nucleocapsid is surrounded by a single envelope and is embedded in a granulin occlusion body. Granulin protein, like polyhedrin of nucleopolyhedroviruses (NPV) [21], protects the viral DNA from adverse environmental conditions [12].

*The sequence data described in this paper have been deposited with the GenBank database and assigned the accession number AF058690.

Epidemic populations of spruce budworm (*Choristoneura fumiferana* (Clem.) Lepidoptera: Tortricidae) larvae cause serious defoliation in North American white spruce and balsam fir forests. These larvae are susceptible to a GV (ChfuGV) which has potential for use in integrated pest management programs directed against this pest. To this end, ChfuGV will have to be characterized but, compared to NPVs, little is known about GV genomes. The nucleotide and predicted protein sequences of the ChfuGV granulin gene have recently been reported [2]. To date, however, there is no report on the region upstream of the ChfuGV granulin gene, or of any DNA regulatory element related to this gene. Here, we report on the results of studies on sequence and transcriptional analyses on the ChfuGV ORF891 gene. The ORF891 is a homology of previously identified ORFs of CIGV ORF909 [11] and CpGV ORF124R [5], with unknown function. Its location adjacent to the granulin gene suggests a similar genome organization in this region among granuloviruses.

Materials and methods

Virus purification and DNA extraction

Purification of ChfuGV occlusion bodies was performed on a 30–80% (v/v) gradient of sucrose (0.1% SDS) and DNA was extracted and purified using cesium chloride-ethidium bromide gradient centrifugation as previously described [2].

Cloning and sequencing

To localize ChfuGV ORF891 gene, clone pGV606 from a ChfuGV/*Bam*H I library was used [2]. pGV606 was digested with *Hind* III to obtain two fragments of 1 Kb and 1.5 Kb (Fig. 1). These two fragments were cloned into pUC18 to generate pGV750 and pGV749, respectively. ChfuGV ORF891 sequence was determined by sequencing the two termini of both strands of pGV750, using universal (UP) and reverse (RP) primers of pUC18. By sequencing the two termini of both strands of pGV749, we localized the putative promoter

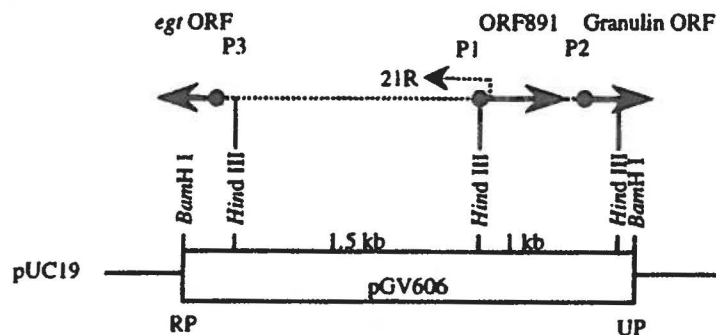


Fig. 1. Schematic representation of the ChfuGV ORF891 with a portion of the ChfuGV granulin and *egt* gene regions. Coding regions are represented by arrows. P1, P2 and P3 begin at the oval end of the arrows and indicate, respectively, the promoter regions of the ORF891, the granulin and the *egt* genes. UP and RP represent the universal and reverse primers. The position of the reverse primer, 21R, specific to the 5'-end of ORF891 is indicated by a dashed arrow

region at the RP terminus. To confirm that the identified region belonged to ChfuGV ORF891 gene, primer 21R (5'-GCATTATCCAAACCG-3') specific to pGV750 RP terminus was labelled with fluorescein-dATP nucleotides by an internal-labelling reaction (Pharmacia) and the sequencing was performed on clone pGV606 by T7 DNA polymerase using an A.L.F. DNA sequencer (Pharmacia). The ORF891 sequence was analyzed and compared to other sequences upstream from granulin genes of CIGV ORF909 [11], CpGV ORF124R [5] and *Autographa californica* multicapsid nucleopolyhedrovirus (Ac MNPV) ME53 gene [13] using PCGENE 6.7 (Intelligenetics Inc.) and Mac Vector 4.5.0 (Eastman Kodak).

Transcriptional mapping

Total RNA was extracted from spruce budworm larvae at 48 h post infection with ChfuGV. Briefly, infected larvae were homogenized in mRNA extraction buffer (Pharmacia Biotech) [19]. The homogenate was filtered through three layers of cheesecloth and pelleted by centrifugation at 2 500 g for 5 min in a microcentrifuge 851 rotor RCF. The supernatant was recovered and loaded onto an oligo(dT)-Cellulose Spun Column (Pharmacia Biotech) and the messenger RNA was isolated as described by the supplier (Pharmacia Biotech) [19]. The size of the transcript was evaluated against a dephosphorylated X174 *Hinf* I marker DNA (Promega) and a known sequence from M13 mp 18 vector (Pharmacia Biotech).

Primer extension was performed using a 29-mer oligonucleotide, 27R: 5-AATTGTGTCCTCAAATCTATCATTTTGAA-3 radiolabelled at its 5' end with 50 μ Ci of 32 P-ATP (ICN; 3000 Ci/mmol) and T4 polynucleotide kinase. During the extension reaction, 10 pmol of labelled 27R primer was annealed to 1 μ g of complementary mRNA and the transcription reaction was performed with 1 μ l of AMV reverse transcriptase (Promega) [23].

Results

Sequence identification in the region upstream of the granulin gene showed an ORF of 891 bp designated as ChfuGV ORF891. This ORF encodes 296 amino acids with a predicted Mr of 34,639 kDa (Fig. 2). Nucleotide sequence comparison of ChfuGV ORF891 with CIGV ORF909 [11], CpGV ORF124R [5] and ME53 gene (ORF 139) of AcMNPV [13] showed homologies of 63.9, 60.9 and 68.2%, respectively. Protein relatedness was 46.3, 49.3 and 3.4% respectively.


Two putative zinc finger motifs were identified in ChfuGV ORF891, one at the N amino terminus with the motif, CX₂CX₃₄CX₂C and a second at the carboxy terminus with the motif, CX₂CX₁₄CX₂C (Fig. 2). Multiple alignment of amino acid sequences of ChfuGV ORF891, CIGV ORF909 and CpGV ORF124R indicates that regions 8–12, 73–77, 89–94, 99–104, 113–116 and 254–259 are well conserved (Fig. 3). Two potential glycosylation sites were found on the sequence at amino acid positions 46 and 116 (Fig. 2).


Amino acid sequence analyses of ChfuGV ORF891, CIGV ORF909, CpGV ORF124R and AcMNPV ME53 ORF indicate that the four sequences share aspartic acid compositions of approximately 5.2% of the total amino acids present. Basic hydrophilic residue lysine, represents of about 9.1% of total amino acids. Of the total amino acids for the three GVs, neutral amino acid residues for threonine and cysteine accounted for approximately 4.7 and 4.1% respectively and hydrophobic residues for leucine about 9.3%.

In the putative promoter region of ChfuGV ORF 891, two TATA boxes at nucleotide (nt) positions –47 and –81 were identified, each located at 28 nt from

-229
 Y A A F V S Q H S L A P T P F V C L I N S S C
 AACTTTATCGACGTACGCTGCAAACACAGACTGATGACTCAACGCCGGTGTGGAAACACACACAAAATATTTGACGAACA 80

-116
 V F C P A L V V L I L Q V F M < Start *egf* l.p.
 AACAAAACACGGCCAGGACGACGAGTATTAGTTGCACGAACATGATAACGAATGAGGTGATAGTTAAGCGCAACGTTT 160

e.p. e.p. 
 ATATAGCCAACCTAACCTGTTTAAATGCAGAAAATAATATAAAATGTTTGTTCGCAGATGATTTGGGTCATTTTCAAA 240
 ATGATAGATTTGAGGACACAATTTTGTCAAAGAGATTCAGAGGTTATCAACGCTTAAATAGGAATTGCTCAAAGCT Hind III 320
 M I D L R T Q F L S K E I Q E V J N A L I G I A Q K L

21R 
 TATTCACGGTTTGGATAATGCAGTGTGCTTCAATTGCAAGACTATATTTAAACAAAACAAAACGTCAAAAAACTAATT 400
 I H G L D N A V

Hinc II
 TTATTTTATTTGTCATTACCGACTACCGCAATGTTGACGACGAAGTTTAAATTTTGTGTTGAAGTCTGCAAAACT 480
 AGCGACAATGTTATGGATGTTATTTAGCTCTATCCAAACGCTGTCGTTCATTCGTAATAAAAAGTTAATGATTCGGGAGT 560
 S D N V M D V I E L Y P T L S L H S V K K L M Y S G V

AAATAAAAAATTTTAAATTTAAATTTAACGACTCCACAATCGTTCAATATAAAAAGTATGTAATGTAGAATCGTGAACA 640
 I K K P F F K F D S T I V Q Y K K Y V I V E S S N S

GTGTTCTCAACAAATACTTGGTGATAAAAAGTATAACGAAGAAAATAATAAAAATTCGTTATTTGTTGTAATGAAACAGAA 720
 V L Q Q I L G D K K Y N E E I I K I R I V R N E T E

ATTGTTGCAGAAGATAATATTAATAGTTTGCGGGTAGATTATGTTGAAAAGTACAAATTTTGATTGCGCGTTAATGCTAAA 800
 I V A E D N I N S L R V D Y G B K Y N F D S P L M L N

TAAAAAGTTTGAAGATGTCACAATCGATGCAAAAAGTACAGATACTATTTAGAGGTGATTACAAAGAATATGAAA 880
 K K V L E D V N N R C K T D R Y Y L E V Y Y K B Y E K

AGTATGCCCTTTGCGCATTAATATAACCGACTTTATCAAAAATGAAATGCCTATATTGTGTAACAAAATATGCGCTGAA 960
 Y A P F A I N Y N R L Y Q N E

ACAGGTAATCCTATATTTGATTTGATGTTGCGGACCCACAAACCCCTAATTTTAAAGCACAAAATGATGATTC 1040
 G P T N P N Y P L K A Q N D D S

CCTTTGGCAATCAAATTAATGTTACAAACTGTTATGGAAAAGAGCGCAACCGAAAAGGTTATATGTTGCCAACATA 1120
 L L A I K L L L Q Q T V L E K T Q P K R L Y V C Q H N

ATGTTATATAACGTGGATACTAGGAGAAACGAAAATGAAACGATTTTATAAGGAATTTAATTTAAACAAGAACAATGGG 1200
 V I * < end ORF891 Start *granulin* >M G

ATATACAAAGCATTAAAGATACAGTCGTACGACGGCACCTCATGCGTTATTGACAAC 1258
 Y N K A L R Y S R H D G T S C V I D N

Fig. 2. Nucleotide sequence of ChfuGV ORF891. Early and late promoters are in bold letters and indicated by e.p. and l.p., respectively. Actual mRNA start site is overscored with an arrow. DNA sequence regulator element, GATA are represented by underlined letters. The position of the primer 21R is indicated with a dashed arrow. Zinc finger motifs at amino and carboxy termini are shaded letters. Pairs of two conserved cysteines are bold underlined letter. N glycosylation sites are in bold letters. Repetitive sequence patterns are underlined

CAGA motif (-25, -53) (Fig. 2). The late promoter motif TAAG (-94) was found to be 69 nt from first CAGA (-25) and 48 nt from the second CAGA (-53).

Transcription experiments indicated that the mRNA of ChfuGV ORF891 gene starts at a position 26 nt from the translation initiation codon ATG and is located at 18 nt from the first TATA box (Fig. 4).

Comparisons of the 119 nt upstream of the ATG initiation codon of ChfuGV ORF891 with the corresponding regions in CIGV ORF909, CpGV ORF124R

Choristoneura fumiferana granulovirus ORF891

741

Consensus	M.DPRTQ...IQEV.NAL VGLAQLH.G LD...CF.CN .QPKQ...N	50
CpGV ORF124R	.Q...TE..MF...S...TTP.VA...Q...H--IK.	48
CIGV ORF909	.H...L.S...N..N. INRGF.V...L...--VN.	48
ChFuGV ORF891	.I.L...I...I.I...IM. .NAV..N.K TI...NKS	50
Consensus	.NPPFIVIT NYLNE.DDTL K...NQ. SRGVMDI...K...K...K...	100
CpGV ORF124R	T...V..K...M...H...C	98
CIGV ORF909	N...V...S...N...H...V...Y	98
ChFuGV ORF891	KTN...D.R.VD.EVF...CKT .DN...V...HS	100
Consensus	...NGVLRK FFFNP.D.TT VL...AIVD S.DSVL.Q.L .EK.YNDEIQ	150
CpGV ORF124R	...V.N...V...E.M.Q...R...	148
CIGV ORF909	...N...L...R.TN...N.TL.D..K.VF R..NHH...	148
ChFuGV ORF891	...S..IK...K.N.S.I...V..E .SN...Q.I GD.K..E..I	150
Consensus	MVRLVRNENE FVAENIA.L R.EYG..YNF DRPL.LN.QL I..VNNHCKL	200
CpGV ORF124R	...N...SV.H..M...RGH...Q..A...AA..K..Q.	198
CIGV ORF909	...I...D...Q...I...TR.T...LY.G...NS...	198
ChFuGV ORF891	KI.I...T.I...D..NS..VD..EK...S..M..KCV LED...R..T	200
Consensus	.KY...YK EYKYSF.V CYN...QK...VY...GKI...TGHPIF.CS.	250
CpGV ORF124R	A..H...V...VKED...MPD...V.S..Y	248
CIGV ORF909	N...P...S..F...M..F.IDP...LKN...N..H	248
ChFuGV ORF891	DR...A..AI N..RLY.N..L..VN.CAE...N...D..I	250
Consensus	CG...K.HTMM.PFW T..YDYNKVY WK.LK.KGLL .CD.M.YGV	300
CpGV ORF124R	...T.R...F...YS...T.R...R..V.L...	298
CIGV ORF909	...L...K...I...HT...S..H..M.K..L.V...	298
ChFuGV ORF891	...L.AQND-SLL AKLLLLQOTV LEXTQP.R.Y V.Q-----	292
Consensus	.RR.I	305
CpGV ORF124R	A..SV	303
CIGV ORF909	T..N.	303
ChFuGV ORF891	-HNV.	296

Fig. 3. Multiple alignment of amino acid sequences of ChFuGV ORF891, CIGV ORF909 [11] and CpGV ORF124R [5]. Conserved regions are hatched. The pairs of conserved cysteines are underlined bold

and AcMNPV ME53 ORF showed homologies of 54, 52 and 49%, respectively. No canonical poly (A) signal, AATAAA, was identified within an appropriate distance from the stop codon at the 3'-end flanking region of ChFuGV ORF891.

Discussion

Compared to the genes immediately upstream of the granulin genes of other GVs ChFuGV ORF891 is the shortest by 19 nucleotides [5, 11]. Despite these deletions, the two zinc finger motif domains at the amino and carboxy termini are conserved between ChFuGV ORF891, CIGV ORF909 and CpGV ORF124R.

Comparisons of the 119 nucleotides immediately upstream of ChFuGV ORF891 with the corresponding regions in CIGV ORF909 and CpGV ORF124R indicated homologies of 54 and 52%, respectively. Multiple sequence alignment of the 119 nucleotides of the three GVs indicated no conserved region. These results show that among GV species, the putative promoter region of gene located immediately upstream of the granulin gene is not conserved.

Amino acid multiple alignment sequences of ChFuGV ORF891, CIGV ORF909 and CpGV ORF124R indicates that the pairs of cysteine residues at

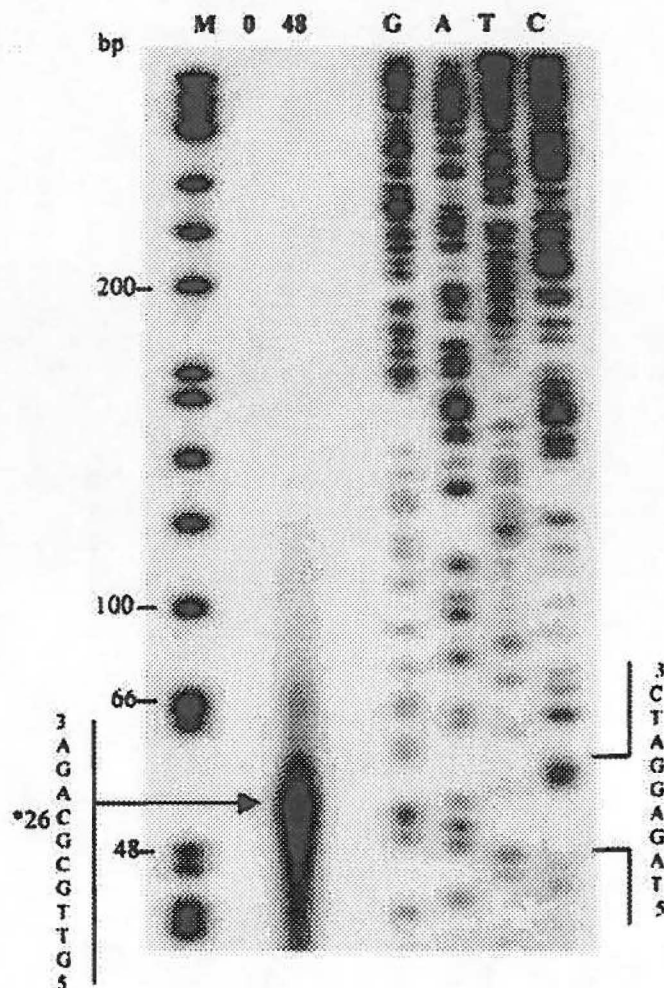


Fig. 4. Transcriptional analysis of ChfuGV ORF891. The transcript detection of the Chfu GV891 gene was performed by primer extension analysis with a specific ^{32}P -end-labelled, 29-mer oligonucleotide complementary to the mRNA. The primer was annealed to mRNA from uninfected (lane 0) and 48 h post infected larvae (lane 48) and extended by reverse transcription. The size of the extension product was estimated first with a DNA ladder X174 *Hinf* I (Promega) and determined by comparison with a sequence ladder (lanes G, A, T, C) (lettering at right) from plasmid M13mp/18 (Pharmacia Biotech). The transcript is shown by an arrow. The ChfuGV ORF891 promoter region is indicated by the lettering at left and the specific start site position, 26 nt from the translation initiation codon, is indicated by an asterisk

positions 36, 39, 74, 77, 230, 233, 248 and 251 are perfectly conserved (Fig. 3). In zinc finger motifs, cysteine is known to play an important role in protein binding by co-ordinating the single atom of zinc [17]. A single replacement of cysteine by a serine in the kruppel protein leads to a consequence similar to the complete deletion of the gene [20]. Certain DNA regulator sequences found in baculoviruses

[3, 5, 27], yeast [8], *Drosophila* [4] and mammalian [22] genes indicate that the two pairs of cysteine are invariably conserved. The zinc finger motif at the ChfuGV ORF891 carboxy terminus (Fig. 2) follows the same rule, $CX_2CX_{14}CX_2C$, as that found in OpMNPV p39 gene [3]. In the case of CIGV ORF909 and CpGV ORF124R, the rule is $CX_2CX_{13}CX_2C$ because the amino acid residues between the two invariant pair of cysteines may vary [28]. ChfuGV ORF891 presents exactly the same characteristic with two zinc finger motifs identified on the amino and carboxy termini of the gene (Fig. 2). Analysis of ChfuGV ORF891 indicates that the sequence may be a member of DNA regulator protein family since it contains all the necessary elements described above.

The putative promoter region of ChfuGV ORF891 contained two potential TATA boxes (-47, -81) that may be recognized by the host RNA polymerase II [10, 17]. Both are located 28 nt from CAGA motif (-25, -53) (Fig. 2). In most baculovirus early promoters, TATA boxes are separated from their mRNA start sites by 21 to 30 nt [11, 13]. The distance of 28 nt in ChfuGV ORF891 promoter is the same as that found in CIGV ORF909 and AcMNPV ME53 ORF promoters. A late promoter motif, TAAG (-95), is located 42 nt from sequence CAGA (-42). The TAAG motif of the CIGV ORF909 promoter is 48 nt from its potential mRNA start site [11]. The presence of two different promoter elements, an early TATA box and late TAAG motif upstream of ChfuGV ORF891, CIGV ORF909 and AcMNPV ME53 ORF indicates that these genes probably have a similar temporal regulation. However, no late promoter element was found in the homologous promoter region of CpGV ORF124R [5]. Despite promoter elements similarity and significant nucleotide sequence homology between ORFs of ChfuGV ORF891 and AcMNPV ME53 genes, there is a total lack of amino acid sequence homology. Previous studies [5, 11] indicated a low homology when amino acid sequence of AcMNPV ME53 ORF coding region was compared to the homologous coding regions of CIGV ORF909 and CpGV ORF124R genes. These observations indicate that the functional roles of ORFs adjacent to the granulin gene of the three GVs and AcMNPV ME53 ORF are probably different.

A sequence GATA was found in the ChfuGV ORF891 promoter region. Previous studies [15] have shown that a GATA sequence could be a promoter regulatory element and is recognized by a number of vertebrate, invertebrate and fungal proteins [6, 7, 9, 16, 25, 26, 28] all of which contain highly conserved zinc finger motifs. Whether the GATA motif serves a similar function in ChfuGV ORF891 gene has to be demonstrated.

Analyses of primer extension experiments indicated that the mRNA transcript of ChfuGV ORF891 starts at nucleotide position -26 (Fig. 4). The distance of 18 nt between the first TATA box (-47) and the beginning of the transcript is the same as has been suggested for the putative promoter of CIGV ORF909 [11]. Moreover, the location of a second translation initiation codon, ATG, downstream from the TATA box (-47), is in agreement with the Kozak consensus rule [14]. These results demonstrate that the TATA box (nt -47) is one of the early promoters of the ChfuGV ORF891 gene. Analysis of the transcript of ChfuGV ORF891

indicated that the mRNA start site motif, CAGA, differs from the consensus sequence, CAGT [11], by a change of the last nucleotide from T to A.

In a recent paper, Smith and Goodale [24] reported that the granulin and ecdysteroid UDP-glucosyltransferase (*egt*) genes of *Lacanobia oleracea* (Lo)GV are separated by ~8 000 base pairs (bp). In ChfuGV, CIGV [11] and CpGV [5] the granulin and *egt* genes are separated by ORFs of 891, 909 and 911 bp, respectively. The deduced amino acid sequence from 115 nt (-116 to -229) of the ChfuGV ORF891 sequence (Fig. 2) showed 48% homology with the translation product of the first 115 nt of the 2 231 bp sequence of CIGV [11] and 61% homology with the first 38 amino acids encoded by the LoGV *egt* gene [24]. These results show that the gene located immediately upstream of the ChfuGV ORF891 could be the *egt* gene, since the gene organization of the granulin gene upstream region of ChfuGV, CIGV and CpGV is almost identical.

The results of our study indicate that the organization of the region immediately upstream of the granulin gene of ChfuGV, CIGV and CpGV is probably identical. This is not the case for the region upstream of the granulin gene of LoGV [24] and the polyhedrin genes of MNPVs [29]. ChfuGV ORF891, CIGV ORF909 and CpGV ORF124R are homologous genes with two conserved zinc finger motifs. Because of their primary polypeptide structure, the three GV genes may belong to a DNA regulator family.

Acknowledgements

This project was supported in part by the Quebec Ministry of Natural Resources grant 0908-240S. CJL received research support through the Canadian Forest Service S & T Opportunities Fund. We thank Lise Forget and Louise Paris for their technical assistance.

References

1. Adams JR, Bonami JR (1991) Atlas of invertebrate viruses. CRC Press, Boca Raton, pp 227-257
2. Bah A, Bergeron J, Arella M, Lucarotti CJ, Guertin C (1997) Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. Arch Virol 142: 1 577-1 584
3. Blissard GW, Quant RRL, Rohrmann GF, Baeudreau GS (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
4. Brunk BP, Martin EC, Adler PN (1991) Drosophila genes *Posterior Sex Combs* and *Suppressor two* of *zeste* encode proteins with homology to the murine *bmi-1* oncogene. Nature 353: 351-353
5. Crook EN, James JD, Smith IRL, Winstanley D (1997) Comprehensive physical map of the *Cydia pomonella* granulovirus genome and sequence analysis of the granulin gene region. J Gen Virol 78: 965-974
6. Evans RM, Hollenberg SM (1988) Zinc fingers: guilt by association. Cell 52: 1-3
7. Evan T, Felsenfeld G (1989) The erythroid-specific transcription of factor Eryf1: a new finger protein. Cell 58: 877-885
8. Freemont PS, Hanson IM, Trawsdale J (1991) A novel cysteine-rich sequence motif. Cell 64: 483-484

9. Fu YH, Marzluf GA (1990) *nit-2*, the major nitrogen regulatory gene of *Neurospora crassa*, encodes a protein a putative zinc finger DNA-binding domain. *Mol Cell Biol* 10: 1 056–1 065
10. Fuchs LY, Woods MS, Weaver RF (1983) Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *J Virol* 67: 3 771–3 776
11. Jehle JA, Bachaus H (1994) The granulin gene region of *Cryptophlebia leucotreta* granulosis virus: sequence analysis and phylogenetic considerations. *J Gen Virol* 75: 3 667–3 671
12. Jehle JA, Bachaus H, Fritsch E, Hubert J (1992) Physical map of the granulin gene region of *Cryptophlebia leucotreta* granulosis virus genome and its relationship to the genome of *Cydia pomonella* granulosis virus. *J Gen Virol* 73: 1 621–1 626
13. Knebel-Mörsdorf D, Kremer A, Jahnel F (1993) Baculovirus gene ME53, which contains a putative zinc finger motif, is one the major early-transcribed genes. *J Virol* 67: 735–758
14. Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44: 283–292
15. Krappa R, Behn-Krappa JF, Doerfler W, Knebel-Mörsdorf D (1992) Differential factor binding at the promoter of early baculovirus gene PE38 during viral infection: GATA motif is recognized by an insect protein. *J Virol* 66: 3 494–3 503
16. Kudla B, Caddick MX, Langdon T, Martinez-Rossi NM, Bennett CF, Sibley S, Davies RW, Arst Jr HN (1990) The regulatory gene are a mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutation affecting specificity of gene activation alter a loop residue of a putative zinc finger. *EMBO J* 9: 1 355–1 364
17. Latchman DS (1990) Eukaryotic transcription factors. *Biochem J* 270: 281–289
18. Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds) (1995) *Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses*. Springer, Wien New York, pp 104–113 (*Arch Virol* [Suppl] 10)
19. Pharmacia Biotech Inc. (1994) QuickPrep mRNA Purification Kit: Instructions. XY-025-00-07
20. Redemann N, Gaul U, Jackle H (1988) *Nature* 332: 90–92
21. Rohrmann GF (1986) Polyhedrin structure. *J Gen Virol* 67: 1 499–1 513
22. Salvato MS, Shimomaye EM (1989) The completed sequence of lymphocytic choriomeningitis virus reveals a unique RNA structure and a gene for a zinc finger protein. *J Virol* 173: 1–10
23. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, New York
24. Smith I, Goodale C (1998) Sequence and *in vivo* transcription of *Lacnobia oleracea* granulovirus *egt*. *J Gen Virol* 79: 405–413
25. Skeiky YAW, Iatrou K (1991) Synergistic interactions of silkmouth chorion-promoter binding factors. *Mol Cell Biol* 11: 1 954–1 964
26. Spieth J, Shim YH, Lea K, Conrad R, Blumental T (1991) *elt-1*, an embryonically expressed *Caenorhabditis elegans* gene homologous to the GATA transcription factor family. *Mol Cell Biol* 11: 4 651–4 659
27. Thiem SM, Miller LK (1989) A baculovirus gene with a novel transcription pattern encodes a polypeptide with a zinc finger and a leucine zipper. *J Virol* 63: 4 489–4 497
28. Wirth UV, Fraefel C, Vogt B, Vlček C, Paces V, Schwyzer M (1992) Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1 are 3' coterminal and encode a putative zinc finger transactivator protein. *J Virol* 66: 2 763–2 772

746 A. Bah et al.: *Choristoneura fumiferana* granulovirus ORF891

29. Zanotto PM DE A, Sampao MJA, Johnson DW, Rocha TL, Maruniak JE (1992) The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *J Gen Virol* 73: 1 049–1 059

Authors' address: Dr. C. Guertin, INRS-Institut Armand-Frappier, Centre de Recherche en Microbiologie-Biotechnologie, Laval, Québec, Canada H7V 1B7.

Received July 9, 1998

CHAPITRE 4

LE GRANULOVIRUS DE *Choristoneura fumiferana*, ChfuGV: ANALYSE DE LA SÉQUENCE DE L'ORF381 ET DE LA TRANSCRIPTION DE SA RÉGION ADJACENTE, LE GÈNE DE LA GRANULINE

**The granulovirus of *Choristoneura fumiferana*, ChfuGV: sequence analysis of ORF381
and its adjacent granulin gene transcriptional region**

TITLE OF PAPER: *Choristoneura fumiferana* granulovirus: sequence analysis of ORF381 and its adjacent granulin gene transcriptional region

RUNNING TITLE: *Choristoneura fumiferana* granulovirus granulin gene and ORF381

AUTHORS: Aliou Bah, Christopher J. Lucarotti¹, Max Arella and Claude Guertin^{2,*}

ADDRESS: INRS-Institut Armand-Frappier
Centre de Recherche en Microbiologie et Biotechnologie
Laval, Québec, Canada H7V 1B7

¹Canadian Forest Service-Atlantic Forestry Center
P. O. Box 4000
Fredericton, New Brunswick
Canada, E3B 5P7

²Corresponding author

Telephone: (450) 687-5010
Facsimile: (450) 686-5533
e-mail: claudio.guertin@inrs-iaf.quebec.ca

The sequence data described in this paper have been deposited with the GenBank database and assigned the accession number AF058690.

RÉSUMÉ

Le cadre de lecture ORF381 a été localisé immédiatement en aval du gène de la granuline du granulovirus de *Choristoneura fumiferana*, ChfuGV. L'ORF381 code pour une protéine de 126 acides aminés ayant une masse moléculaire estimée à environ 14,2 kDa. La région promotrice du gène contient deux boîtes TATA et un motif, ATAAG, caractéristique des promoteurs très tardifs des baculovirus. Des identités significatives ont été retrouvées en comparant les portions N-terminales des séquences d'acides aminés de l'ORF381 avec les régions N-terminales des gènes "VEF" (facteur d'accroissement de l'infection virale) du granulovirus de *Trichoplusia ni* (L.), TnGV et des gènes de la chitinase et de la cathepsine du granulovirus de *Cydia pommonella* (L), CpGV. Une comparaison de la séquence en acides aminés de la partie C-terminale de l'ORF381 avec celle de la protéine de fusion du virus de la maladie de New Castle a révélé aussi une forte identité. Le transcrit de l'ARN messenger de l'extrémité 5' du gène de la granuline adjacent à l'ORF381, a aussi été identifié. Il débute à partir de la première lettre, ATAAG, de la séquence consensus des promoteurs tardifs des baculovirus.

Abstract

A putative open reading frame (ORF) of 381 nucleotides (ORF381) was identified immediately downstream of the granulin gene of the *Choristoneura fumiferana* granulovirus (ChfuGV). ORF381 encodes 126 amino acids with an estimate molecular weight of 14.2 kDa. The putative promoter region contains two TATA boxes and a baculovirus late promoter motif, ATAAG. Significant identities were found when a part of ChfuGV ORF381 amino acid N-terminus sequence was compared to the amino acid N-terminus sequence of the *Trichoplusia ni* granulovirus (TnGV) viral enhancing factor (VEF), *Cydia pomonella* granulovirus (CpGV) chitinase and cathepsin genes. A portion of the amino acid C-terminus of ORF381 was identical to the C-terminus portion of the F protein of the Newcastle disease virus. The mRNA transcript of the 5' end of the ORF381 adjacent region, the granulin gene, was determined. The transcript begins at the first adenine of the baculovirus very late promoter consensus sequence ATAAG.

Keywords: *Choristoneura fumiferana* granulovirus, ORF381, granulin gene, open reading frame.

1. Introduction

Granuloviruses (GV) are members of the family *Baculoviridae* which also includes nucleopolyhedroviruses (NPV) (Murphy et al., 1995). The GV genome is a double-stranded super-coiled DNA, usually greater than 100 kb, that is packaged within a single nucleocapsid. 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 (Jehle *et al.*, 1992; Blissard and Rohrmann, 1990). The genes that encode both granulin and polyhedrin are highly conserved and are often used to help establish phylogenetic relatedness among baculoviruses (Zanotto et al., 1993).

For many years, various GVs have been viewed with interest for use in controlling economically important insect pests in agriculture and forestry (Adams and Bonami, 1991). The spruce budworm larva, *Choristoneura fumiferana*, is considered the most important pest of coniferous trees in North America (Comtois, 1988) and is susceptible to a granulovirus (ChfuGV). To date, two genes of the ChfuGV genome have

been identified and sequenced, the granulin gene (Bah *et al.*, 1997) and a DNA regulatory family gene, ORF891, located immediately upstream to the granulin gene (Bah *et al.*, 1999). In addition to these, we have isolated and sequenced a gene (ORF381) located immediately downstream of the granulin gene. In contrast to NPVs, mRNA transcripts have not been mapped for any GV granulin gene, all studies have been based on analyses and comparisons of the putative promoter regions of GV granulin genes (Jehle and Backaus, 1994). Herein, we provide such a map for ChfuGV granulin gene. While granulin gene sequences are highly conserved among GV species, the same can not be said for the 5' intergenic and 3' flanking regions of this gene. Using a polymorphic detection technique, the Dcode System (Bio-Rad), mutations were identified in these regions between different isolates of ChfuGV.

2. Materials and methods

2.1. Virus isolates

Our laboratory strain of ChfuGV was isolated from infected spruce budworm larvae collected near Bonaventure, Québec, Canada. It was designated as the Bonaventure isolate and was the principle strain used in all analyses. Other ChfuGV isolates used were obtained from infected spruce budworm larvae collected at the Acadia Experimental Forest Station near Fredericton, New Brunswick, Canada. These latter isolates were designated by the year and the number of the larva from which they were isolated: (1985) 85-1089, (1987) 87-126, 87-181 and 87-271. Each ChfuGV isolate was amplified by serial passages in spruce budworm larvae in the laboratory (Bah *et al.* 1997).

2.2. Virus purification, DNA extraction and restriction enzyme digestion

ChfuGV occlusion bodies were isolated on a 30-80% (v/v) sucrose gradient (0.1% SDS) and genomic DNA was phenol/chloroform extracted and purified by cesium chloride-ethidium bromide gradient centrifugation as described previously (Bah *et al.*, 1997).

2.3. PCR amplification and DNA sequencing

The 5' intergenic and 3' flanking regions of the ChfuGV granulin gene were amplified with specific pairs of primers: intergenic region - 8F: 5-CATTTGCGGACCCACAAACC-3 and 2R: 5-CATCATTCAGCACAGAACC-3; flanking region - 5F: 5-GTACCGCTTTTTACTGGTCCCG-3 and 13R: 5-

GCCCCGCCCCCGCCGCCCTGCCCCGCGCCCCGCGCCGCCCGCACGGTGGGAAGAG TTGACTGA-3. For optimal resolution in perpendicular gel electrophoresis, a GC clamp (Sheffield et al., 1989) was attached to the primer 13R.

DNA was amplified by PCR in a reaction volume of 100 μ L containing 10 μ L 10X PCR buffer, 2 μ L 10mM dNTP, 2 μ L of each primer, 0.25 ng of ChfuGV genomic DNA, 0.5 μ L Tween-20 and 1 unit of Pwo DNA polymerase (Boehringer Mannheim). Amplification was performed in a Perkin Elmer Turbo-Thermal Cycler (model 9600). Samples were denatured for 7 min at 94°C followed by a 25-cycle program, each cycle consisting of 45 sec at 94°C, 45 sec at 55°C and 45 sec at 72°C. Amplification was completed with an extension cycle of 10 min at 72°C.

A portion of the 3' flanking region of the ChfuGV granulin gene (Bah et al., 1997) was extended using a specific internal primer 12F: 5-TCAGTCAACTCTTCCACCGT-3, labelled with fluorescein-dATP nucleotides using an internal-labelling reaction (Pharmacia) and the sequencing was performed on clone pGV199 (Fig. 1). The sequenced downstream region was analysed using PCGENE 6.7 (Intelligenetics Inc.) and MacVector 4.5.0 (Eastman Kodak).

2.4 Detection of sequence variability in the ChfuGV granulin gene intergenic and flanking regions

Sequence variations were identified using the components of the DCode Universal Mutation Detection System (Bio-Rad). This system involves the formation of heteroduplex (annealing between wild-type and mutant DNA or RNA) molecules which are characterized by mismatched nucleotides at the sites of alteration. The system can identify single base variations in segments of DNA using differences in melting behavior of DNA fragments which may be identical in sequence except for a single base pair. In denaturing gradient gel electrophoresis (DGGE), DNA fragments migrate according to size until they reach the gradient where the molecule begins to denature. When the DNA denatures, it opens in a domain which is anchored by a higher melting domain or domains that have not melted. At this point, the migration rate slows down. When the fragment completely denatures, then migration again becomes a function of size.

Two types of DGGE gels may be run to detect mutations in DNA. The first is a perpendicular gradient, in which the gradient is perpendicular to the electric field and

uses a broad denaturing gradient range, 0-100% or 0-70%. From this gel, one can determine the concentration of denaturants in which the wild type and mutant can be separated. The second is parallel DGGE, where the denaturing gradient is parallel to the electric field, and the range of denaturants is narrowed to allow better separation of fragments. A third component of the Dcode System is constant gradient gel electrophoresis where there is no gradient. Instead, the optimum denaturing concentration, which allowed separation in perpendicular DGGE, is used.

Variability in ChfuGV granulin gene intergenic and flanking region sequences were detected using the Dcode Universal Mutation Detection System (Bio-Rad).

Fragments of 306 and 275 bp were amplified by PCR in the respective 5' intergenic and 3' flanking regions of the ChfuGV granulin gene. Fifty μL of the amplified PCR product from the upstream or downstream region of the granulin gene region of two different ChfuGV isolates were pooled in a microfuge tube with 100 μL 2X loading buffer. Samples were denatured for 5 min at 95°C, incubated for 2 h at 60°C and chilled on ice. The mixture was loaded for perpendicular DGGE on an 8% acrylamide perpendicular gradient gel (37.5: 1 acrylamide: bis) with an increasing denaturant gradient of 0-100% where 100% denaturant was 7 M urea/40% formamide. Parallel DGGE gels were also run where the denaturing gradient range was 0-40% maximum denaturant. All DGGE runs were performed in 1X TAE buffer at 130 V at a constant temperature of 60°C for 1.5 h. After electrophoresis, gels were stained in a 50 $\mu\text{g}/\text{mL}$ ethidium bromide solution for 5 min and destained for 10 min in 1X TAE buffer. Visualization of gels was performed on a UV transilluminator (Bio-Rad). Once differences had been established using DGGE, individual mutations were identified using CDGE with 20% denaturant.

2.6. Transcriptional mapping

Total RNA was extracted from spruce budworm larvae at 48 h post infection with ChfuGV. Messenger RNA was isolated as described previously (Bah et al., 1999). Primer extension was performed using a 35-mer oligonucleotide, 26R: 5-TGACGACTGTATCTTAATGCTTTGTTATATCCCAT-3 radiolabeled at its 5' end with 50 μCi of ^{32}P -ATP (ICN; 3000 Ci/mmol) and T4 polynucleotide kinase. During the extension reaction, 10 pmol of labelled 35R primer was annealed to 1 μg of

complementary mRNA and the transcription reaction was performed with 1 μ L of AMV reverse transcriptase (Promega) (Sambrook and Magnatis, 1989).

3. Results

3.1. Analysis of the *ChfuGV OFR381* gene

The published portion of the 3' flanking region of the *ChfuGV* granulin gene (Bah et al., 1997) was further sequenced and analysed. The sequenced region harbours a small open reading frame (ORF) of 381 nucleotides located in the first complementary frame with a stop codon overlapping with the granulin gene TAA codon (Fig. 1). This ORF was called ORF381 and codes for 126 amino acids with an estimated molecular weight of 14.2 kDa. The putative promoter region of the ORF381 indicated the presence of two putative early TATA boxes located at -40 and -84 nucleotides (nt) from the start codon and a putative late promoter motif ATAAG located at -73 nt from the ATG initiation codon. A putative mRNA start motif, CAGT, was located at position -32. Two N-glycosylation sites were identified at amino acid sequence positions, 26 and 76 (Fig. 2).

Alignments of a portions of the N-terminus amino acid sequence of ORF381 and viral enhancing factor (VEF) gene of the granulovirus of *Trichoplusia ni* (TnGV) (Hashimoto et al., 1992) indicated an identity of 58% and a similarity of 74% (Fig. 3). Only 14.3% identity was found when the complete amino acid sequences of the two genes were aligned. Alignment of a portion of the N-terminus sequence of the *ChfuGV* ORF381 and the *CpGV* cathepsin gene (Kang et al., 1998) showed an identity of 40% and similarity of 58% (Fig. 3). An identity of 34% was found between a part of the C-terminus amino acid sequence of the *ChfuGV* ORF381 and the virus fusion F protein of the Newcastle disease virus (Toyoda et al., 1987).

3.2. Transcriptinal analysis of the *ChfuGV granulin* gene

Transcriptional mapping analyses of the *ChfuGV* granulin gene showed that the transcript of the gene begins at the first letter of the baculovirus consensus sequence promoter motif, TAAG (Fig. 4).

3.3. Dcode analysis of the *ChfuGV granulin* gene 5' intergenic and 3' flanking regions

Perpendicular DGGE of the 306 bp fragment, amplified from the intergenic 5' region of the granulin gene of the Bonaventure *ChfuGV*, yielded a single band indicating a homoduplex (Fig. 5A). Heteroduplexes were obtained when the amplified DNA

fragments of the Bonaventure ChfuGV and ChfuGV 87-181 were mixed (Fig. 5B). Similarly, mixtures of the 275 bp fragment amplified from the ChfuGV granulin gene 3' flanking regions of these two strains yielded heteroduplexes (Fig. 5B). Parallel DGGE of the 306 bp DNA fragment obtained from the Bonaventure ChfuGV and from ChfuGV isolates 85-1089, 87-126, 87-181 and 87-271 showed them all to be heteroduplexes. Heteroduplexes were also obtained when the 306 bp DNA fragment obtained from the Bonaventure ChfuGV was mixed with each of the amplified 306 bp fragments from ChfuGV isolates 85-1089, 87-126, 87-181 and 87-271 (Fig. 5C). Similarly, CDGE showed heteroduplex DNA for the Bonaventure strain 306 bp fragment and heteroduplexes when it was mixed with each of the 306 bp fragments from isolates 85-1089, 87-126, 87-181 and 87-271 (Fig. 5D).

4. Discussion

The sequence analysis of ORF381 gene indicates that the ATG initiation codon, located at the beginning of sequence respects Kozak's rule (Kozak, 1982) with a purine at position -3 and G at position +4, immediately following an ATG triplet.

The genes located immediately upstream and downstream of the polyhedrin genes in several NPVs, ORF9 of *Autographa californica* MNPV (Zanotto *et al.*, 1992), ORF7 of *Rachiplusia ou* MNPV (Harrison and Bonning, GeneBank: AF068270), ORF1629 of *Bombyx moris* NPV (Maeda, Genbank: L33180) and ORF2 of *Orgyia pseudotsugata* MNPV (Arhens *et al.*, 1997) are capsid-associated proteins. VEF represents a capsule protein that has been considered as a synergistic factor for virus infection by enhancing attachment of the viral envelope receptors in the plasma membranes of insect host midgut cells (Hashimoto *et al.*, 1992). Cathepsin is a lysosomal enzyme involved in the degradation of the insect peritrophic membrane, it also promotes liquefaction of the host in the late stages of infection (Hawtin *et al.*, 1997). F-protein is a glycoprotein involved in virus envelope fusion with host cell receptor proteins (Toyoda *et al.*, 1987). Even though no firm conclusions can be made regarding function of ORF381, the location of capsid-associated genes close to polyhedrin genes in NPVs and amino acid comparative analyses with TnGV VEF, PuGV cathepsin and Newcastle disease F-protein suggest a capsid-associated protein or an envelope protein function for ORF381.

Transcriptional studies of the promoter regions of NPV polyhedrin genes indicate that these regions contain an invariable sequence consensus promoter motif, ATAAG where transcription of polyhedrin mRNA begins on the first adenine of the motif (Rohrmann, 1986). The mRNA transcript analysis of the ChfuGV granulin gene demonstrates that transcription of this gene also begins at the first adenine of the same motif.

Heteroduplex DNA fragments obtained in perpendicular DGGE from mixtures of the amplified 5' intergenic and 3' flanking regions of the Bonaventure ChfuGV granulin gene with those of the other isolates indicated the presence of at least one mutational difference in each of these respective fragments between the Bonaventure isolate and each of the isolates from New Brunswick.

In parallel and constant denaturing gradient gels, all analysed homoduplex and heteroduplex DNA fragments of the granulin gene 5' intergenic region of the ChfuGV isolates showed two separate bands (Fig. 5 C, D). The presence of two separate bands of the homoduplex control fragment of ChfuGV in parallel and constant gradient gels was not expected. There are two possible explanations for this. It is possible to detect a mutation running just homoduplex DNA if there is a mutation especially if the mutation involves a base change of from A to G. Alternatively, it is well known that latent viruses can occur in baculoviruses and recombine with an heterologous virus or integrate to the host genome (Smith et Crook, 1988; Crook, 1985; Crook, 1981). These mutations were not detected using genomic DNA restriction profiles for isolates Bonaventure, 87-181 and 87-271 (Fig. 6).

In conclusion, there is no intergenic region between the ChfuGV granulin gene and the ORF381. Furthermore, no similarity was found when the amino acid N-terminus portion of the ChfuGV ORF381 was compared with the part of the amino acid sequence of the translated complementary strand of the granulin gene 3' flanking region of CrleGV (Jehle and Backhaus, 1994), PbGV (Chakerian et al. 1985) and *Xestia C-nigrum* GV (Goto et al., 1998;). Previous reports (Bah et al., 1999; Jehle and Backhaus, 1994) indicated that the organization in GV granulin gene upstream region is very similar. Our analyses indicate that this is not the case in the downstream region where the N-terminus of the ChfuGV ORF381 was dissimilar to other GV granulin gene 3' flanking regions.

Acknowledgements

This project was supported in part by the Quebec Ministry of Natural Resources grant 0908-240S. CJL received research support through the Canadian Forest Service S & T Opportunities Fund. We thank Lise Forget and Lise Paris for their technical assistance.

References

- Adams, J. R. and Bonami, J. R., 1991. Atlas of invertebrate viruses, CRC Press, Boca Raton, pp. 227-257.
- Bah, A., Bergeron, J., Arella, M., Lucarotti, C. J. and Guertin, C., 1997. Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. Arch. Virol. 142, 1577-1584.
- Bah, A., Lucarotti, C. J., Arella, M. and Guertin, C., 1999. *Choristoneura fumiferana* Granulovirus: Sequence Analysis and 5' Characterization of ORF891. Arch. Virol., 144, 737-746.
- Blissard, G. W. and Rohrmann, F., 1990. Baculovirus diversity and molecular biology. Annu. Rev. Entomol. 35, 127-155.
- Comtois, B., 1988. Notions d'entomologie forestière. Modulo éditeur, pp. 123. ISBN 2-89113-207-6.
- Chakerian, R., Rohrmann, G. F., Nesson, M. H., Leisy, D. J. and Beaudreau, G. S., 1985. The nucleotide sequence of the *Pieris brassicae* granulosis virus granulin gene. J. Gen. Virol. 66, 1263-1269.
- Crook, E. N., James, J. D., Smith, I. R. L. and Winstanley, D., 1997. Comprehensive physical map of the *Cydia pomonella* granulovirus genome and sequence analysis of the granulin gene region. J. Gen. Virol. 78, 965-974.
- Crook, N. E., 1981. A Comparison of the granulosis viruses from *Pieris brassicae* and *Pieris rapae*. Virology 115, 173-181.
- Crook, N. E., Spencer, R. A., Payne, C. C. et Leisy, D. J., 1985. Variation in *Cydia pomonella* granulosis virus isolates and physical maps of the DNA from three variants. J. Gen. Virol. 66, 2423-2430.
- Goto, C., Hayakawa, T. et Maeda, S. 1998. Genome organization of *Xestia c-nigrum* granulovirus. *Virus Genes* 16 (2), 199-210.

- Hashimoto, Y., Corsaro, B. G. and Granados, R., R., 1991. Location and nucleotide sequence of the gene encoding the viral enhancing factor of the *Trichoplusia ni* granulosis virus. *J. Gen. virol.* 72, 2645-2651.
- Hawtin, R. E., Zarkowska, T., Arnold, K., Thomas, C. J., Gooday, G. W., King, L. A., Kuzio, J. A. et Possee, R. D., 1997. Liquefaction of *Autographa californica* polyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. *Virology* 238, 243-253.
- Jehle, J. A. and Backhaus, H., 1994. The granulin gene region of *Cryptophlebia leucotreta* granulosis virus: sequence analysis and phylotgenetic considerations. *J. Gen. Virol.* 75, 3667-3671.
- Jehle, J. A., Backhaus, H., Fritsch, E., Hubert, J., 1992. Physical map of the granulin gene region of *Cryptophlebia leucotreta* granulosis virus genome and its relationship to the genome of *Cydia pomonella* granulosis virus. *J. Gen. Virol.* 73, 1621-1626.
- Kang, W., Tristem, M., Maeda, S., Crook, N. E. et O'Reilly, D. R., 1998. Identification and characterization of the *Cydia pomonella* granulovirus cathepsin and chitinase genes. *J. Gen. Virol.* 79 (9), 2283-92.
- Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44, 283-292.
- Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A., Summers, M. D., 1995. *Virus Taxonomy. Classification and Nomenclature of Viruses. Sixth Report of the International Committee on Taxonomy of Viruses.* Springer, Wien New York, pp 104-113 (Arch. Virol. [Suppl] 10).
- Pharmacia Biotech Inc., 1994. QuickPrep mRNA Purification Kit: Instructions. XY-025-00-07.
- Rohrmann, G. F., 1986. Polyhedrin structure. *J. Gen. Virol.* 67, 1499-1513.
- Sambrook, J., Fritsch, E. F., and Maniatis, T., 1989. *Molecular cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory.
- Smith, I. R. L. et Crook, N. E., 1988. *In vivo* isolation of baculovirus genotypes. *Virology* 166, 240-244.

- Toyoda, T., Sakaguchi, T., Imai, K., Inocencio, N. M., Gotho, B., Hamaguchi, M. et Nagai, Y., 1987. Structural comparison of the cleavage-activation site of the fusion glycoprotein between virulent and avirulent strains of Newcastle disease virus. *Virology* 158, 242-247.
- Zanotto, P. M. de A., Sampaio, M. J. A., Johnson, D. W., Rocha, T. L. and Maruniak, J. E., 1992. The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *J. Gen. Virol.* 73, 1049-1056.

Table 1. Percent identity and similarity of sequences producing significant alignment with the N and C -terminal aminoacid sequences of ChfuGV ORF381.

Sequences		ChfuGV ORF381	
		Identity	Similarity
Cathepsin gene of CpGV (9/22)		40%	58%
ChfuGV	9 SKNYNADVRD'TIKLMKWNN SL K 30 +K Y +D IKL + N+LK	9/22	13/22
CpGV	40 NKTYVSDEERA IKLEN FKNNLK 61		
Viral enhancing factor (VEF) of TnGV		58%	74%
ChfuGV	18 DTIKLMKWNN SL 29 +TI + WNN SL	7/12	9/12
TnGV	257 NTILIEIWNN SL 268		
The gene F protein of the Newcastle disease virus (NDV)		34%	54%
ChfuGV	75 KNSVITSL SNNI --DKMEDNSKI IKLNVA 122 ++SV+T+LS NNI DK+E+NSK++KL++A	17/50	27/50
NDV	451 QDSVVTNLS SNNI SNDKLEENSKLVKLHIA 500		

Fig. 1. Schematic representation of the ChfuGV ORF381, location on clone pGV199. The arrows indicate the ChfuGV ORF381 and a portion of the published granulin genes (Bah et al., 1997). The oval end of the arrow indicates the promoter region, P4, of the ORF381. The dashed arrow represents the primer 12F, used for the internal sequencing of ORF381.

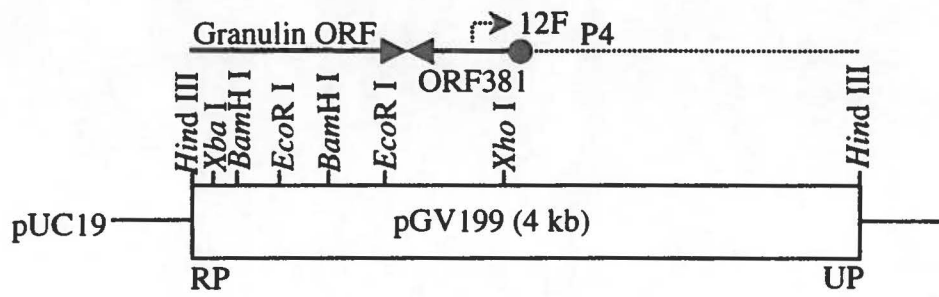
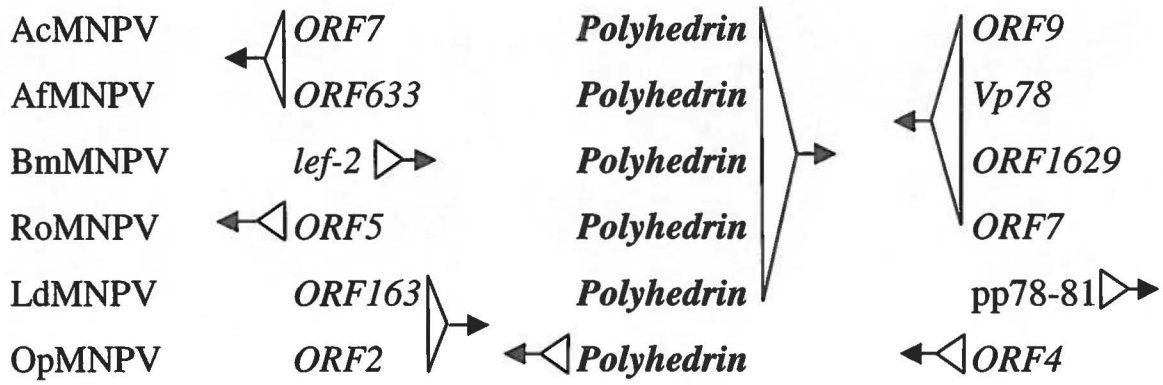


Fig. 2. Nucleotide sequence of the ChfuGV ORF381 and its adjacent granulin gene transcriptional region. The ORF381 putative early and late promoter elements are dashed and bold boxed respectively. A potential baculovirus mRNA initiation site is double underlined. Primer positions and direction of the elongation are showed by arrows. The ORF381 N-glycosilation sites are in bold and are underlined. Repetitive sequence patterns are single underlined. The granulin gene mRNA transcriptional start site is indicated by the oval arrow. The analysed granulin gene 5' and 3' flanking regions by the DCode system for mutation detection are dashed underlined.

↗ 8F
AACAGGTAATCCTATATTTGATTGTAGCATTTGCGGACCCACAAACCCTAATTATTTTTTAAAAGCACAAAATGATGATT 80
CCCTTTTGGCAATCAAATTAAGTGTACAACTGTATTGGAAAAGACGCAACCAGAAAAGGTTATATGTGTGCCAACAT 160
AATGTTATATAACGTGGATACTAGGAGAAAACGAAAAAGAAACGATTTTATAAGGAATTTAATTTAAACAAGAACAATGG 240
↖ 26R ↗ Hind III
GATATAACAAAGCATTAAGATACAGTCGTCACGACGGCACCTCATGCGTTATTGACAACCACCACCTCAAAGCTTAGGT 320
↖ 2R
TCTGTGCTGAATGATGTTTCGACACAAAAAGACCGCATACGTGAGGCTGAATACGAACCCATTCTAGATATTGCCAATCA 400
BamH III
GTACATGGTGACCGAGGATCCTTTTCGTGGACCCGGTAAGAACGTACGAATTACCCTTTTCAAGGAAATTAGGCGCGTTC 480
AACCAGACACCATGAAATTTGGTATGCAACTGGAGCGGAAAGGAAATCCCTTCGCGAAACATGGACACGCTTCATTTCTGAA 560
GAGTTTCCCATCACTACTGACCAAGAAATCATGGATTTGTGGTTTGAAATTCAACTCAGACCAATGCAACCCAATCGCTG 640
BamH I
TTACAAATTTACGATGCAATACGCGCTTGGTGC TAACCCGAAATACGTTGCACATGATGTAATCCGTCAACAGGATCCTT 720
ACTACGTTGGTCC TAACAACATGGAACGCATCAACTTGTCCAAGAAAGGATTTCGCATTCCCCTCAGGTGCTTCAATCC 800
GTCTACAAATGACAAC TTTGAGAGATTCTTCGACGACATTCTTTGGCCGTA TTTCCACCGTCTTTAGTATATGTGGGAAC 880
↗ 5F
CACTTCCGCCGAAATCGAAGAAATTATGATCGAGGTGTCGCTATTGTTTTAAATTAAGAATTCGCACCAGATGTACCGC 960
TTTTTACTGGTCCCGCATATTAATCC TGT TTTTCAGCGACATTTTCGTC TAATTTTATTTGAGTATCAACAATTTTAGAA 1040
* D Q K E A V N E D L K I Q T D V I K S
TTTAAGTCTTCCATTTTATCACAACTATTACGTTGATGTTATTGGACGCCGGCTTCGATTTGGGTAACGAAGAAGTAAT 1120
N L D E M K D C V I V N I N N S A P K S K P L S S T
TGCAACGGAAGAATTTTTTGCAACATCAACCTTATTGTTGGTGATTTGCAAATTTTAATAATCATATTATAAACGCTGT 1200
I A V S S N K A V D V K N T P S K C I K I I M N Y V S
↗ F12 ↖ 13R
CCAAAACTCAGTCAACTCTTCCACCGTTAACGTTAACACGTTGTCGTCAGTTTCCGTTTAAATTTTCTTTTTAAGCTA 1280
D L F E T L E E V T L T L V N D D T E T N L K R K L S
TTATTCATTTCAATTAAC TTTATTGTGTCACGAACGTCAGCATTTGTAATTTTACTTTTAAATGAATCCTAACAAATCCAT 1360
N N W K M L K I T D R V D A N Y N K S K I F G L L D M
ATCTCGAGTGTATTGGAATTA AAAAACACTGAAAT TATA AAAAATCTAGGCGAAAATGACGAAGACT CTTAT GATAACG 1440
TATA TTTATGCAAAAAAAAAGGTGACCC 1428

Fig. 3. Schematic representation of the MNPV polyhedrin (A) and GV granulin (B) gene regions organization. A. Polyhedrin gene region of AcMNPV (Zanotto *et al.*, 1992), AfMNPV (Federici, GenBank:AFU64896), BmMNPV (Maeda, Genbank: L33180), LdMNPV (Kuzio *et al.*,1999), OpMNPV (Arhens *et al.*, 1997) and RoMNPV (Harrison and Bonning, GeneBank: AF068270). The genes coding directions are indicated by the arrows.

A



B

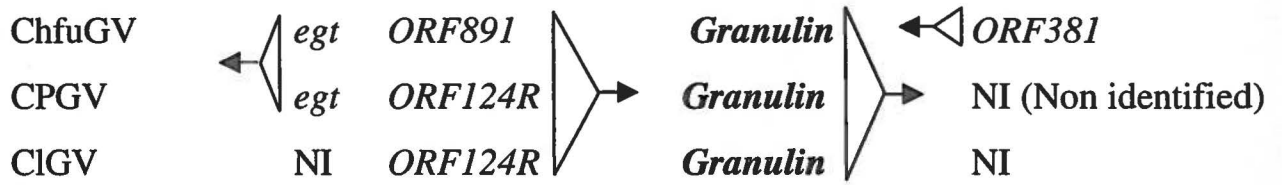


Fig. 4. Transcriptional analysis of ChfuGV granulin gene. The transcript detection of the ChfuGV granulin gene was performed by primer extension analysis with a specific ^{32}P -end-labelled, 29-mer oligonucleotide complementary to the mRNA. The primer was annealed to mRNA from uninfected (lane 0) and 48 h post infected larvae (lane 48) and extended by reverse transcription. The size of the extension product was estimated first with a DNA ladder X174 *Hinf* I (Promega) and determined by comparison with a sequence ladder (lanes G, A, T, C) (lettering at right) from plasmid M13mp/18 (Pharmacia Biotech). The transcript is showed by arrow. The promoter region of the ChfuGV granulin gene is indicated by the lettering at left and the specific start site position, 28 nt from the translation initiation codon, is indicated by an asterisk.

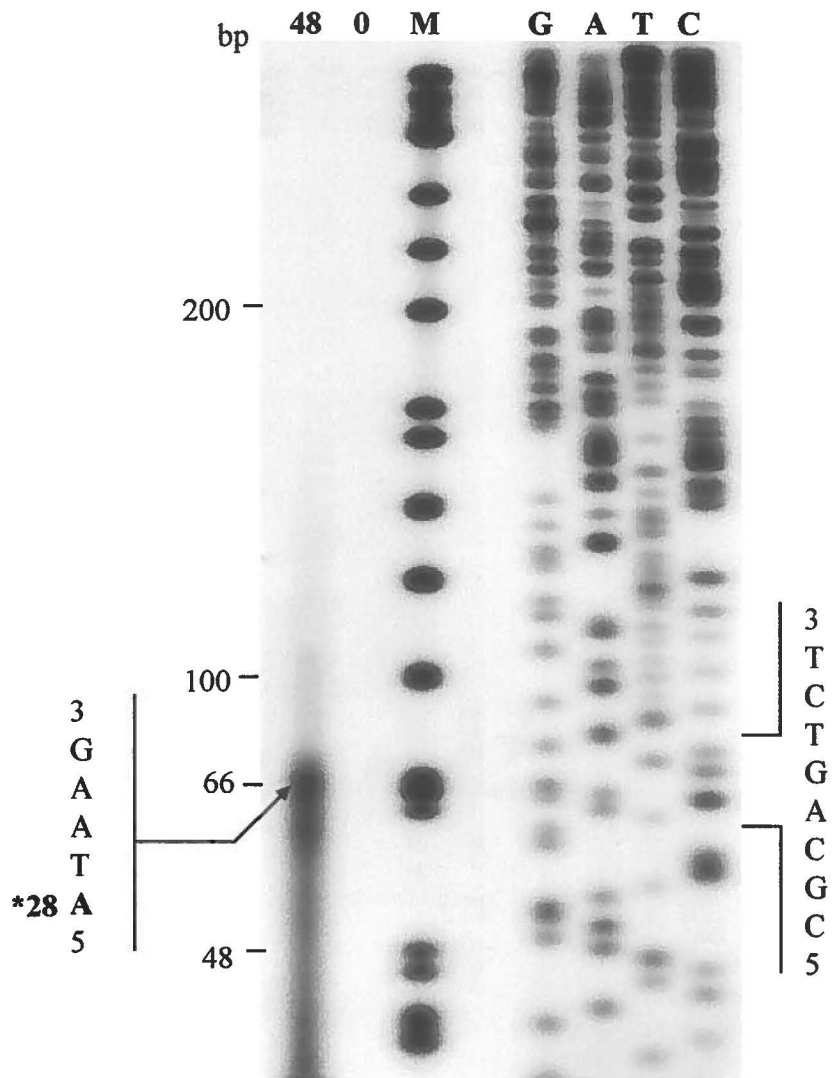


Fig. 5. Analysis by the DCode system of the 5' intergenic and the 3' flanking regions of the ChfuGV granulin gene in a perpendicular (A, B), parallel (C) and constant (D) denaturing gradient gels. A. The single S curve represents an homoduplex DNA fragment from ChfuGV Bonaventure isolate (BI). B. The double S curves represent heteroduplex DNA fragments obtained by mixing amplicons of Bonaventure and 87-181 ChfuGV isolates from New Brunswick (NBI). C and D. Parallel DGGE and constant DGGE of heteroduplex DNA fragments obtained from the mixture of ChfuGV Bonaventure isolate with each of New Brunswick 85-1089, 87-126, 87-181 and 87-271.

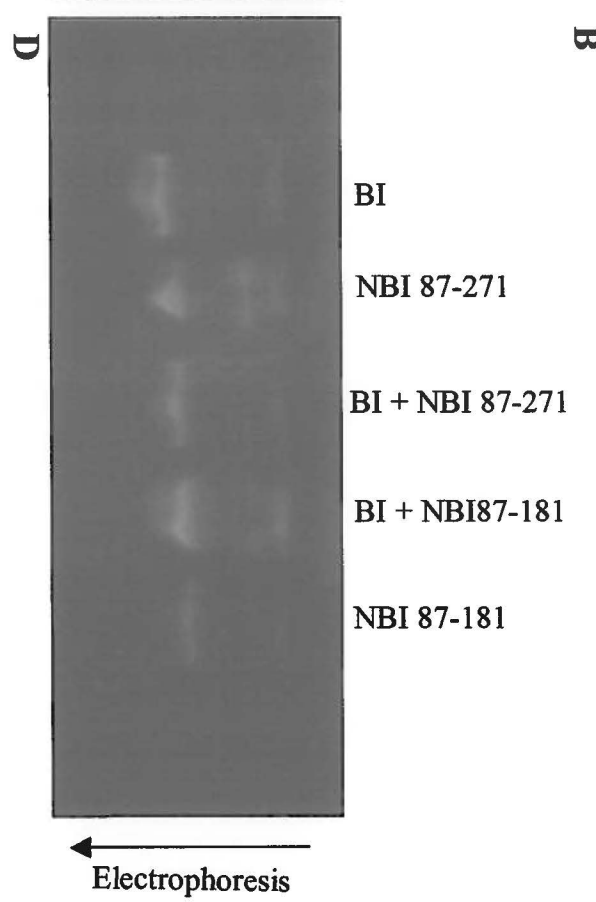
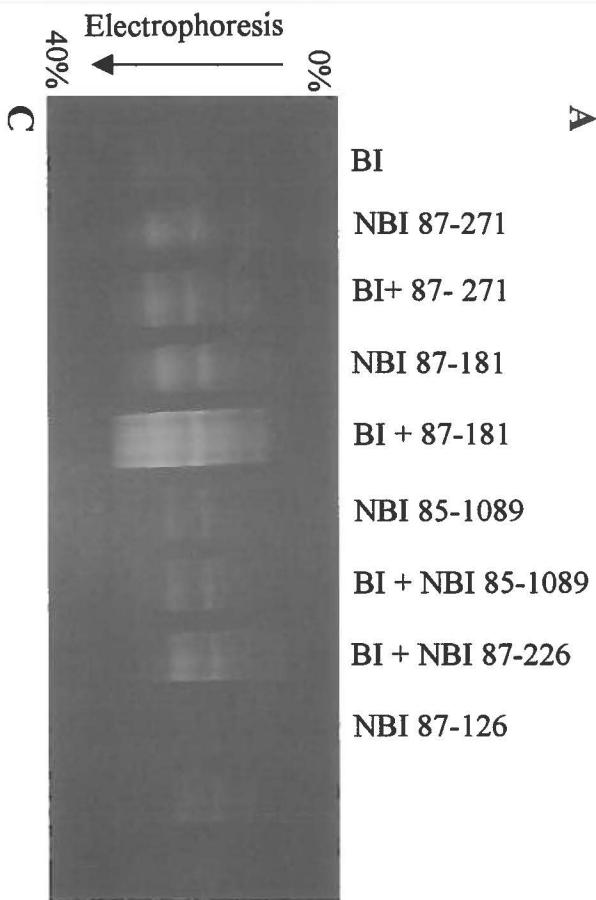
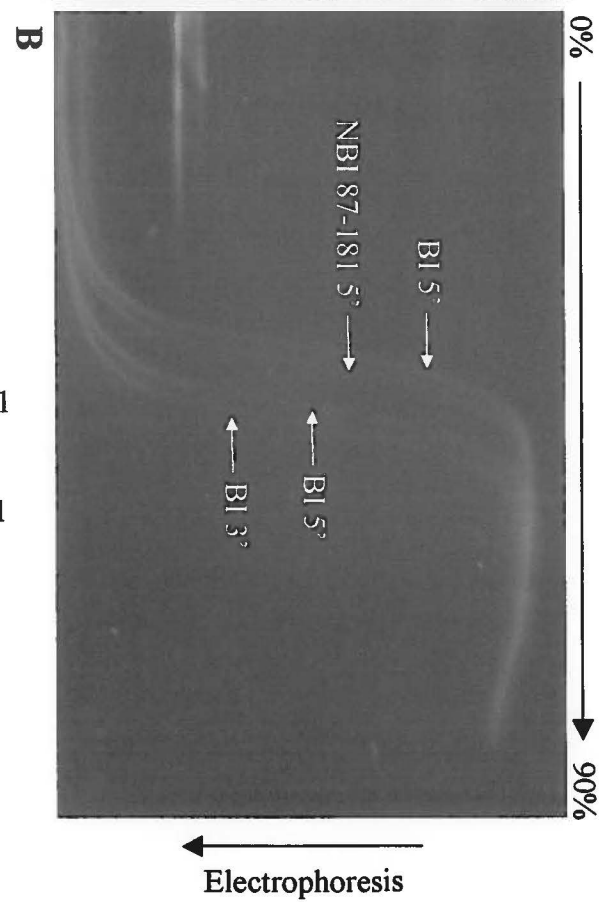
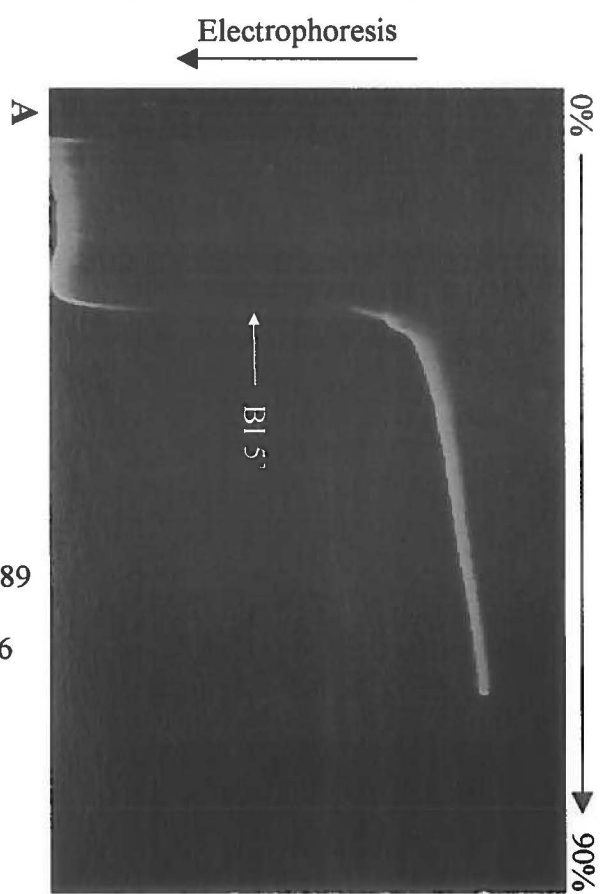
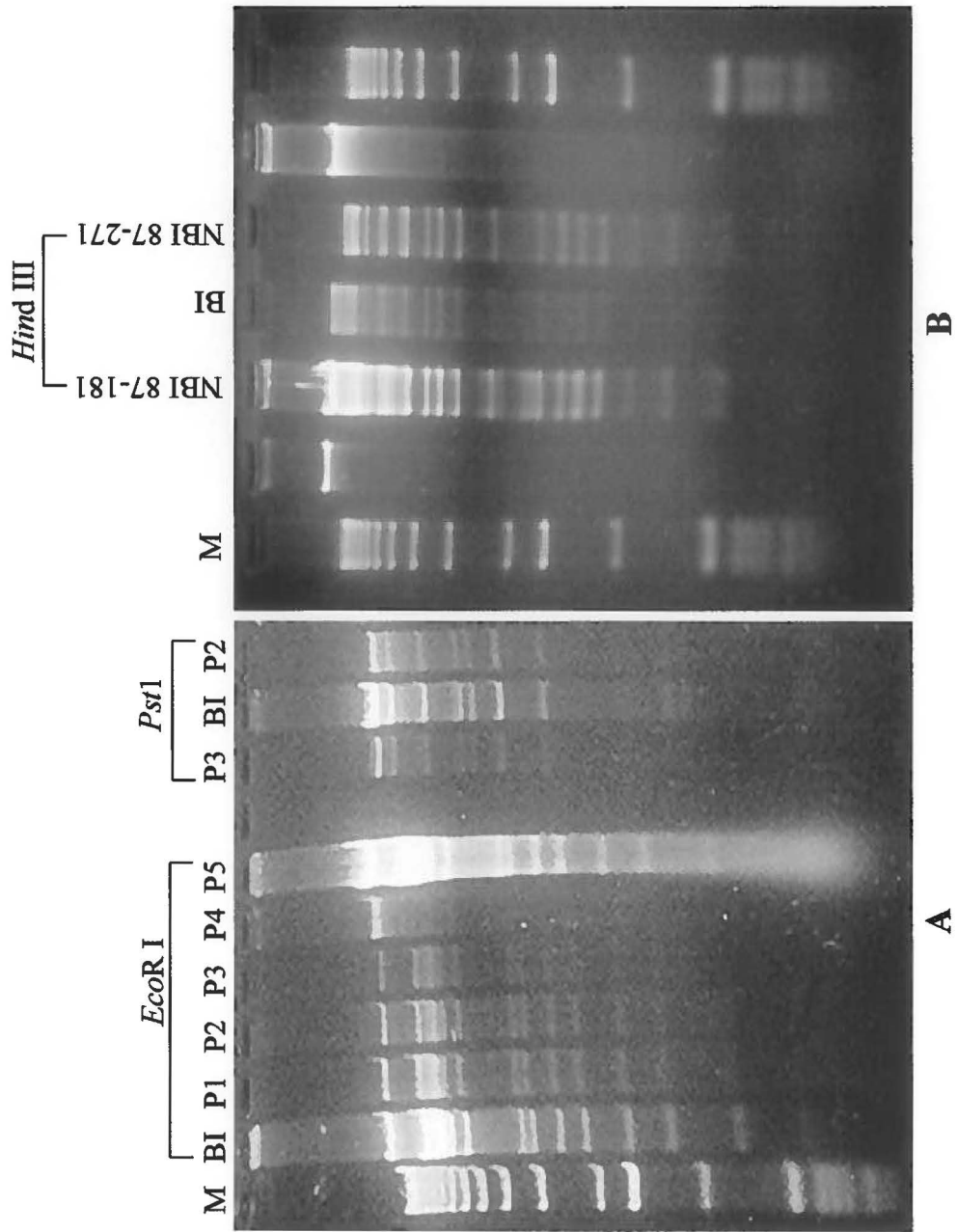


Fig. 6. Comparison of the restriction enzyme (REN) profiles of the original genome of Bonaventure ChfuGV isolate (BI), with each of the first to five serial passages (P1-P5) in spruce budworm and isolates 87-181 and 87-271 from New Brunswick (NBI). A DNA molecular marker (M) was also included.



DISCUSSION GÉNÉRALE

Identification et analyses de séquence du gène de la granuline

Chez les granulovirus, la granuline constitue la protéine la plus abondante produite par le virus. La granuline assure la protection des virions contre les aléas naturels comme la polyédricité chez les nucléopolyédrovirus (Jehle *et al.*, 1992). Arif et collaborateurs (1986) ont rapporté des études sur les profils de restriction génomiques et sur l'étude de la protéine des trois isolats de GV isolés chez le genre *Choristoneura*. Toutefois, aucune étude n'avait encore été rapportée sur le gène de la granuline du ChfuGV y compris les éléments de régulation de la région promotrice de ce gène.

Le gène de la granuline du ChfuGV a été séquencé et analysé. La constitution d'une banque génomique ChfuGV/*Hind* III et ChfuGV/*Bam*H I, nous a permis de localiser les portions N et C terminales du gène de la granuline sur les clones pGV606 et pGV199 respectivement (Chapitre II, Fig. 1, A et B). Dans la revue bibliographique, il a été mentionné que le gène de la granuline était très conservé parmi les GV, avec une forte homologie de plus de 75 et 80% en séquences nucléotidiques et en acides aminés respectivement (Jehle et Backhaus, 1994). Les résultats des analyses de séquences du gène de la granuline du ChfuGV confirment ces résultats. Ce gène présente une homologie en séquences nucléotidiques et en acides aminés se situant entre 76,7% et 88,3% avec les gènes de la granuline du CrleGV, PbGV et TnGV (Chapitre 2, Tableau 1). La région promotrice du gène de la granuline du ChfuGV a aussi été caractérisée. Le motif ATAAG du promoteur du gène est identique à la séquence consensus des promoteurs très tardifs des baculovirus (Rohrmann, 1986). Le codon d'initiation du gène est conforme à la règle de Kozak avec un A à la position -3 et un G à la position +4. Les 14 premiers nucléotides de la région promotrice situés immédiatement avant le codon d'initiation du gène de la granuline du ChfuGV, CrleGV, PbGV, CpGV et TnGV sont entièrement identiques (Chapitre 2, Fig. 2). Cette région hautement conservée parmi les GV, pourrait jouer un rôle très important dans la régulation de l'expression du gène et être utilisée aussi comme élément de diagnostic pour la détection des GV (Bah *et al.*, 1997).

Les alignements multiples des séquences en acides aminés des gènes de la granuline du ChfuGV, CrleGV, PbGV et TnGV indiquent que les résidus d'arginine, de proline et de cystéine restent bien conservés entre les quatre GV. La conservation des positions d'arginine et de prolines suggère une même flexibilité de la granuline chez les

quatre GV. À la lumière des études rapportées sur la polyédrine du nucléopolyédrovirus de AcMNPV (Rohrmann, 1986), la séquence RHKK de la granuline du ChfuGV pourrait servir de signal de localisation nucléaire (Jarvis *et al.*, 1991) (Chapitre 2, Fig. 4).

Des études comparatives en séquences nucléotidiques des extrémités 5' et 3' du gène de la granuline du ChfuGV avec celles des autres gènes de la granuline du CrleGV, PbGV et TnGV indiquent que les extrémités 5' sont beaucoup plus conservées que les extrémités 3' (Chapitre 2, Tableau 1).

Le granulovirus de *Choristoneura fumiferana*, ChfuGV: analyse de séquences et caractérisation de l'extrémité 5' du gène ORF891

Le gène ORF891 a été identifié immédiatement en amont du gène de la granuline chez le ChfuGV (Chapitre 3, Fig. 1). Comparé aux autres gènes homologues, l'ORF909 du CrleGV et l'ORF124R du CpGV, l'ORF891 représente le gène le plus court avec une délétion de 19 nucléotides. Malgré cette délétion, les deux motifs de doigt de zinc aux deux extrémités N amino et C carboxy terminales du gène sont restés bien conservés (Bah *et al.*, 1999).

Des comparaisons de séquences de 119 nucléotides localisées immédiatement en amont du cadre de lecture du gène ORF891 avec les régions homologues des gènes ORF909 et ORF124R ont montré des homologies respectives de 54 et 52%, respectivement. Mais les alignements multiples de séquences de ces mêmes régions indiquent qu'elles ne sont pas conservées.

Les alignements de séquences en acides aminés des gènes ORF891, ORF909 et ORF124R indiquent que les résidus de cystéines sont bien conservés à travers les trois séquences (Chapitre 3, Fig. 3). Il a été rapporté que les résidus de cystéines jouent un rôle très important dans l'interaction de la protéine au niveau du motif du doigt de zinc, plus précisément par coordination d'un simple atome de zinc. Parmi les séquences des gènes régulateurs d'ADN, la disposition des deux paires de cystéines au niveau du motif du doigt de zinc, CX₂CX₁₄CX₂C, reste invariable et conservée. On retrouve ces mêmes caractéristiques chez le gène ORF891 (Chapitre 3, Fig. 2) avec le motif, CX₂CX₁₄CX₂C ainsi que chez l'ORF909 du CrleGV et l'ORF124R du CpGV ayant pour motif, CX₂CX₁₃CX₂C. Il a été rapporté que les gènes ORF909 et ORF124R font partie des gènes régulateurs d'ADN (Crook *et al.*, 1997; Jehle et Backhaus, 1994). À la lumière de

ces remarques, le gène ORF891 ferait partie de la famille des gènes régulateurs de la transcription de l'ADN vu qu'il possède toutes les caractéristiques de ces derniers.

La région promotrice du gène ORF891 du ChfuGV contient deux boîtes TATA (-47, -81) pouvant être reconnues par l'ARN polymérase de l'hôte. Ces deux boîtes sont localisées chacune à 28 nucléotides du motif CAGA (-25, -53), site potentiel d'initiation de l'ARN messager (Chapitre 3, Fig. 2). Cette distance est la même retrouvée entre les promoteurs précoces de l'ORF909 du CrleGV, l'ORF ME53 de AcMNPV et leur site de transcription de l'ARN messager (Jehle et Backhaus, 1994; Knebel-Mörsdorf *et al.*, 1993).

Le motif TAAG des promoteurs tardifs des baculovirus, a été localisé à la position -95, à partir du site d'initiation ATG du gène ORF891. La présence d'un promoteur précoce, TATA et d'un autre tardif, TAAG, au niveau des gènes ORF891, ORF909 et ME53, laisse suggérer que ces gènes auraient une régulation temporelle identique. Cependant, aucun élément de promoteur tardif n'a été identifié dans la région promotrice du gène ORF124R du CpGV (Crook *et al.*, 1997). Il n'existe pas une très forte homologie en acides aminés entre le cadre de lecture du gène ORF ME53 et celui des gènes ORF909 et ORF124R, cette homologie est encore moins significative avec l'ORF891 du ChfuGV. Ces observations nous indiquent que les gènes de GV situés immédiatement en amont du gène de la granuline pourraient avoir une fonction différente de celui de l'ORF ME53.

Des expériences d'extension d'oligonucléotides indiquent que le transcrit de l'ARN messager de l'ORF891 du ChfuGV débute à partir du nucléotide situé à la position -26 (Chapitre 3, Fig. 2). La distance entre la boîte TATA (-47) et le début du transcrit est la même que celle observée au niveau de la région promotrice de l'ORF909 du CrleGV. En plus, la localisation d'un second codon d'initiation en aval du promoteur, TATA (-47) est conforme à la règle de Kozak (Kozak, 1986). Ces résultats nous indiquent que la boîte TATA, position -47, pourrait être un des promoteurs précoces du gène ORF891. Les analyses de transcrit de l'ORF891 indiquent que le motif du site d'initiation du transcrit de ce gène, CAGA, diffère de celui du motif consensus, CAGT par un changement du T à la place du A (Bah *et al.*, 1999).

Récemment, Smith et Goodale (1998) ont rapporté que les gènes de la granuline et de l'ecdystéroïde UDP-glucosyltransferase (*egt*) du granulovirus de *Lacanobia oleracea* (LoGV) étaient séparés par des nucléotides d'environ 8000 pb. Chez le ChfuGV, le

CrleGV et le CpGV, les gènes de la granuline et de l'*egt* sont séparés par des cadres de lectures de 891, 909 et 911 pb respectivement. Les acides aminés obtenus à partir de la traduction de 115 nucléotides (-116 à -229) en aval du codon d'arrêt de l'ORF891 (Chapitre 3, Fig. 2) présentent une homologie de 48% avec les acides aminés obtenus à partir de la traduction des 115 nucléotides des 2231 pb du CrleGV. Cette homologie est de 61% avec la portion de séquences homologue de LoGV. Ces résultats nous montrent que la séquence localisée immédiatement en amont du gène ORF891 du ChfuGV pourrait être le gène *egt*, vu que l'organisation génomique des gènes en amont du gène de la granuline du ChfuGV, CrleGV et CpGV est presque identique.

Les résultats de ce travail nous indiquent que l'organisation génomique de la région située immédiatement en amont du gène de la granuline du ChfuGV, CrleGV et CpGV serait probablement identique. Bien que ceci ne soit pas le cas des gènes de la granuline de LoGV (Smith et Goodale, 1998) et de la polyédrique chez les MNPV (Zanotto *et al.*, 1992). Les gènes ORF891 du ChfuGV, ORF909 du CrleGV et ORF124R du CpGV sont des gènes homologues possédant chacun deux motifs de doigt de zinc bien conservés. En se fiant à la structure primaire des polypeptides de ces gènes, on peut conclure qu'ils appartiendraient tous à la famille des gènes régulateurs de la transcription d'ADN.

Le granulovirus de *Choristoneura fumiferana*, ChfuGV: Analyses des séquences du gène ORF381 et de la transcription du gène adjacent, le gène de la granuline

Le gène ORF381 du ChfuGV a été séquencé et analysé. Le codon d'initiation de la traduction du gène, ATG, est conforme à la règle de Kozak avec une purine à la position -3 et un G à la position +4 (Chapitre 4, Fig. 2).

Les alignements de séquences du gène ORF381 du ChfuGV indiquent qu'une portion de séquence d'acides aminés du bout N-terminal est identique à 58% et similaire à 74% à celle du gène du facteur d'accroissement de l'infection virale ou "VEF" du TnGV et du CpGV. Une identité de 40% avec une similarité de 58% ont été obtenues en comparant la portion en acides aminés N-terminale de l'ORF381 et celle du gène de la cathepsine du PuGV (Chapitre 4, Tableau 1). Le VEF est une protéine du corps d'inclusion du virus qui est considéré comme un facteur synergistique pouvant augmenter l'infection virale par attachement de l'enveloppe du virus aux récepteurs cellulaires du plasma membranaire (Hashimoto *et al.*, 1991; Tanada *et al.*, 1980). La cathepsine est une

lysosome impliquée dans la dégradation de la membrane de l'insecte et elle provoque la liquéfaction de son hôte durant les phases tardives de l'infection (Hawtin *et al.*, 1997).

D'autre part, la comparaison de la portion d'acides aminés C-terminale de l'ORF381 avec la partie correspondante de la protéine de fusion F du virus de la maladie de New Castle a indiqué une identité de 34% (Chapitre 4, Tableau 1). La protéine F est une glycoprotéine responsable de la fusion de l'enveloppe virale avec les cellules réceptrices de l'hôte (Pritzer *et al.*, 1990; Toyoda *et al.*, 1987). Il a été rapporté que les gènes localisés dans les régions immédiates en amont et en aval du gène de la polyédrine soit, l'ORF9 de *Autographa californica* (AcMNPV) (Zanotto *et al.*, 1992) et ses homologues comme, l'ORF7 de *Rachiplusia* ou MNPV (RoMNPV) (Harrison et Bonning, GeneBank: AF068270), l'ORF1629 de *Bombyx moris* NPV (BmNPV) (Maeda, Genbank: L33180) et l'ORF2 de *Orgyia pseudotsugata* MNPV (OpMNPV) (Arhens *et al.*, 1997) sont tous des gènes qui codent pour des protéines associées à la capsid du virus. Pour le moment, il est impossible d'affirmer avec certitude le rôle et la fonction du gène ORF381 du ChfuGV. Toutefois, les analyses de comparaisons en acides aminés avec les gènes du VEF du TnGV, de la cathepsine du PuGV et de la protéine F du virus de la maladie de New Castle, en plus des observations sur la disposition des gènes associés à la capsid autour de la polyédrine des MNPV (Chapitre 4, Fig. 3) nous laissent croire que l'ORF381 pourrait être impliqué dans l'infection du virus.

Les études de transcription des régions promotrices des gènes de la polyédrine chez les MNPV ont indiqué que les promoteurs de ces gènes comprenaient un motif de séquence invariable, ATAAG, très conservé chez les baculovirus. Les transcrits de l'ARN messager du gène de la granuline du ChfuGV commencent à partir de la première lettre de la séquence consensus, ATAAG qui se trouve être aussi très conservé chez les granulovirus (Chapitre 4, Fig. 4). Jusqu'à présent, aucune étude n'avait encore été rapportée sur le transcrit d'un gène de la granuline de GV. Nos résultats indiquent clairement que les transcrits de l'ARN messager de la polyédrine et de la granuline sont identiques et commencent tous les deux à partir de la première lettre du motif, ATAAG.

L'électrophorèse en conditions dénaturantes sur gradient perpendiculaire de gel de polyacrylamide des fragments d'ADN hétéroduplexes, amplifiés à partir des régions 3' de l'ORF381 et 5' intergénique du gène de la granuline du ChfuGV de Bonaventure et les

autres isolats de ChfuGV du Nouveau Brunswick, montrent deux bandes distinctes en forme de S (Chapitre 4, Fig. 5B). Ceci indique la présence d'au moins une mutation dans ces régions analysées. Par contre, le fragment contrôle homoduplexe amplifié à partir de la région intergénique 5' du gène de la granuline du ChfuGV de Bonaventure montrait seulement une simple bande en S, indiquant l'absence de mutations (Chapitre 4, Fig. 5A).

Les électrophorèses en conditions dénaturantes sur gradients parallèles et constants de gel polyacrylamide des fragments d'ADN homoduplexes et hétéroduplexes, amplifiés à partir de l'extrémité intergénique 5' du gène de la granuline du ChfuGV de Bonaventure et des isolats de ChfuGV du Nouveau Brunswick ont montré deux bandes distinctes dans chaque cas (Chapitre 4, Fig. 5C, D). Ceci confirme bien la présence d'une mutation au niveau de ces deux régions de la granuline de ChfuGV. Par contre, on ne s'attendait pas à la présence de deux bandes d'ADN distinctes au niveau des fragments homoduplexes (Chapitre 4, Fig. 5C, D). Ceci pourrait s'expliquer par deux raisons. Dans certains cas, il est possible de détecter une mutation dans un fragment homoduplexe si celle-ci implique un simple changement de bases d'un A par un G; parce que la température de dissociation du complexe G-C est plus élevée que celui du complexe A-T. D'autre part, des études préalables ont démontré que chez les baculovirus, il pourrait exister des virus latents pouvant se recombiner avec des virus hétérologues ou bien s'intégrer dans le génome de la cellule hôte (Crook, 1981; Crook, 1985; Smith et Crook, 1988), ce qui pourrait probablement expliquer la présence des deux bandes à partir des fragments homoduplexes utilisés comme témoins (Chapitre 4, Fig. 5C, D). Ainsi les extrémités 5' du gène de la granuline et 3' de l'ORF381 du ChfuGV pourraient contenir des séquences mutées, même si les profils de restrictions enzymatiques du génome au complet n'ont indiqué aucune différence significative, en comparant le ChfuGV du laboratoire avec les virus issus de passages sériés chez la tordeuse des bourgeons de l'épinette (Chapitre 4, Fig. 6A) et aussi avec les isolats, 87-181 et 87-271 du Nouveau Brunswick (Chapitre 4, Fig. 6B).

Dans une publication récente (Bah *et al.*, 1999), il avait été souligné que l'organisation des gènes situés en amont de la granuline était identique chez les GV. Tel n'est pas le cas dans la région en aval de la granuline. Non seulement il n'y a pas de région intergénique entre le gène de la granuline et l'ORF381 (article III, Bah *et al.*,

soumis) comme c'est le cas avec l'ORF891 (Bah *et al.*, 1999), mais en plus il n'existe aucune similarité entre les acides aminés N-terminal de l'ORF381 et ceux des régions homologues N-terminal des autres gènes de CrleGV, CpGV et *Xestia C-nigrum* (XcGV) (Jehle et Backhaus, 1994; Crook *et al.*, 1997; Goto *et al.*, 1998).

CONCLUSION GÉNÉRALE

Ces études de doctorat ont permis d'apporter une contribution significative à la caractérisation moléculaire du granulovirus de la tordeuse des bourgeons de l'épinette, ChfuGV. Grâce à ce projet, le gène de la granuline qui code pour la protéine virale la plus abondante et l'une des plus importantes impliquée dans l'infection du virus, a été cloné séquencé et caractérisé. Des études comparatives de ce gène incluant la région promotrice entre le ChfuGV et d'autres granulovirus ont montré que le gène était très bien conservé chez les GV et qu'il présentait aussi de fortes homologues de séquence avec les gènes de la polyédrine des nucléopolyédrovirus (MNPV). De plus, la caractérisation de la région promotrice du gène de la granuline du ChfuGV a montré que le site d'initiation du transcrit de l'ARN messager de ce gène était identique à celui de la polyédrine des MNPV. Ainsi, comme le gène de la polyédrine des MNPV, le gène de la granuline du ChfuGV serait exprimé de façon très tardive par un promoteur fort ayant pour motif la séquence consensus, ATAAG, des baculovirus.

Cette étude a aussi permis de localiser deux autres gènes dans les voisinages immédiats du gène de la granuline, l'ORF891 en amont et l'ORF391 en aval. Il est à noter qu'en amont de l'ORF891, une portion de séquence d'un quatrième gène qu'on pourrait identifier comme étant celui de l'ecdysone glucosyle transférase (*egt*) a aussi été identifié.

L'ORF891 a été séquencé et caractérisé. Il appartiendrait à la famille des gènes régulateurs d'ADN. Ce gène montre une identité significative en comparaison avec les ORF909 et ORF124R du CrleGV et CpGV respectivement. Malgré une délétion de 19 nucléotides par rapport aux gènes homologues, les deux motifs de doigt de zinc sont restés bien conservés dans les extrémités N et C terminales de l'ORF891. Le site d'initiation du transcrit de l'ARN messager du gène a été identifié. La présence et l'emplacement des promoteurs précoces, TATA, et tardifs, TAAG de ce gène sont identiques à ceux des autres gènes homologues, les ORF909 et ORF124R de GV et ME53 de MNPV, ce qui suggère la possibilité d'une même régulation temporelle. Par ailleurs, la région promotrice d'un autre gène, l'*egt* dont la portion N-terminale a été identifiée, se chevauche avec celle de l'ORF891.

Quant à l'ORF381, il est contigu au gène de la granuline sans aucun espace intergénique. Si on s'en tient aux études comparatives, les portions N-terminales de ce

gène sont très identiques à celles des gènes responsables de l'accroissement du facteur d'infection du virus ou "VEF" et celles des gènes de la cathepsine et la chitinase qui provoquent la liquéfaction de la larve. L'autre portion C terminale de l'ORF381 est très identique à celle de la protéine de fusion F du virus de la maladie du New Castle. Ce qui nous laisse supposer que l'ORF381 pourrait être impliqué dans l'infection du virus.

Cette étude nous a permis de comprendre de façon très claire, l'organisation génomique de la région du gène de la granuline du ChfuGV. C'est ainsi que chez la plupart des GV étudiés, les gènes localisés immédiatement en amont du gène de la granuline sont organisés de façon très similaire. Les gènes régulateurs de la transcription de l'ADN et ceux de l'*egt* se partagent la même région promotrice et pourraient être probablement régulés à partir des boîtes TATA de manière identique. Les régions 3' des gènes régulateurs d'ADN se chevauchent avec la région promotrice du gène de la granuline, ce dernier étant lui même contigu à l'ORF381. Cependant, l'arrangement génomique en aval de la granuline chez les GV pourrait ne pas se ressembler. Des études comparatives entre les portions de séquences nucléotidiques et en acides aminés des régions N terminales en amont du gène de la granuline chez le ChfuGV, CrleGV, CpGV et XcGV n'ont montré aucune homologie significative.

Par l'utilisation d'un système de détection rapide des mutations avec le "DCode system", des mutations potentielles ont été localisées dans les extrémités 5' intergéniques du gène de la granuline et 3' de l'ORF381. Les résultats de ces analyses laissent supposer la présence d'un gène hétérologue d'un autre granulovirus chez la tordeuse des bourgeons de l'épinette.

Les résultats de ces travaux de doctorat sont certainement la prémisse à plusieurs autres projets. Par exemple, ils permettent d'envisager l'expression des gènes séquencés afin d'étudier leurs fonctions biologiques *in vivo* ou *in vitro*, via les protéines exprimées. On peut penser aussi à la caractérisation et à l'expression du gène de l'*egt*. Des expériences fort intéressantes pourront être conduites avec ce gène. La simple délétion de l'*egt* pourrait accroître l'infection virale et entraîner une forte mortalité larvaire. La surexpression de ce même gène conduirait à une augmentation de la répllication virale, ce qui serait idéale en bioindustrie. Encore plus, le promoteur du gène pourrait être utilisé dans les vecteurs d'expression, vu que l'*egt* est non essentiel à la répllication du virus. Le

promoteur du gène de la granuline est considéré comme un promoteur très fort, comparable à celui de la polyédrique et du gène p10 des MNPV, il serait possible d'utiliser aussi ce promoteur dans les vecteurs d'expression pour la production de protéines recombinantes ou comme marqueur génétique dans les tests de diagnostics pour l'identification des granulovirus. Cependant, il serait bien de noter que la mise en place d'une lignée cellulaire permettant la réplication du virus est essentielle à plusieurs de ces travaux.

REMERCIEMENTS

Je dédie cette thèse de doctorat à ma très chère soeur bien aimée, *Aïssata Bah* dite *Thilo*, décédée au moment même où j'achevais ce travail. Que Allaah* ait pitié de son âme.

Merci à mon père Amadou Bah (que Allaah* ait pitié de son âme), à ma très chère mère, *Cotiary Ly*, à mes frères et soeurs et à mon épouse bien aimée, *Fatima Azzouz*.

Merci à mon directeur de recherche, *Claude Guertin* qui, je dois le souligner, m'a beaucoup appris sur la patience en milieu de recherche. Ses idées scientifiques, ses suggestions, ses corrections, la qualité de ses critiques, ses remarques pertinentes, ont été incontournables tout au long de ce travail. Son support matériel et moral, sa disponibilité et sa largesse d'esprit ont contribué à m'assurer et à me rendre beaucoup plus confiant pour pouvoir affronter les défis de la recherche durant mes quatre années de doctorat dans son laboratoire.

Merci à mon codirecteur, *Max Arella*, qui a contribué de façon très remarquable à l'avancement de ce travail. Ses conseils en biologie moléculaire ont été très appréciés. Ses critiques ont permis de donner une forme beaucoup plus élaborée dans la conception des trois manuscrits. Sa gentillesse et sa disponibilité m'ont été bien utiles tout au long du projet.

Merci à *Christopher J. Lucarotti*, pour la qualité de ses corrections durant la rédaction des trois manuscrits. Je le remercie aussi pour sa gentillesse et sa bonne volonté de collaboration sans oublier *Jean Bergeron*, pour son implication durant les premiers stades du projet.

Merci à l'INRS-Institut Armand-Frappier et au Ministère des Ressources Naturelles du Québec, pour toutes les bourses d'études et le matériel scientifique dont j'ai bénéficiés depuis ma maîtrise.

Merci à tous ceux et celles avec qui j'ai collaboré, depuis ma maîtrise et à tout le personnel d'entretien de jour et de soir de l'INRS-Institut Armand-Frappier, pour m'avoir signifié leur gentillesse et leur sympathie.

Savez-vous, la science est belle! Elle est belle parce que personne ne peut se l'approprier, belle parce qu'on doit l'apprendre, la partager et l'utiliser pour le bien être de l'univers. Cependant, elle reste la propriété exclusive de Allaah, l'omniscient, le digne de louanges et de remerciements; celui qui a créé et agencé harmonieusement l'univers.*

Allaah (Dieu).*

RÉFÉRENCES

- ADAMS, J. R. et BONAMI, J. R. 1991. Atlas of invertebrate viruses, CRC Press Inc. pp. 31-40, 229.
- AKIYOSHI, D., CHAKERIAN, R., ROHRMANN, G. F., NESSON, M. H. et BEAUDREAU, G. S. 1985. Cloning and sequencing of the granulin gene from *Trichoplusia ni* granulosis virus. *J. Virol.* **141**: 328-332.
- ARHENS, C. H. et ROHRMANN, G. F. 1996. The DNA polymerase and helicase genes of baculovirus of *Orgyia pseudotsugata*. *J. Gen. Virol.* **77**: 825-837.
- ARHENS, C. H., RUSSEL, R. L., FUNK, C. J., EVAN, J. T., HARWOOD, S. H. et ROHRMANN, G. F. 1997. The sequence of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus genome. *Virology* **229**: 381-99.
- ARIF, B. et DOERFLER, W. 1984. Identification and localization of reiterated sequences in the *Choristoneura fumiferana* MNPV genome. *EMBO J.* **3**: 525-529.
- ARIF, B. M., GUANGYU, Z. et JAMIESON, P. 1986. A comparison of three granulosis viruses isolated from *Choristoneura Spp.* *J. Invert. Pathol.* **48**: 180-186.
- AYERS, M. D., HOWARD, S. C., KUZIO, J., LOPEZ-FERBER, M. et POSSEE, R. D. 1994. The complete sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**: 586-605.
- BAH, A., LUCAROTTI, C. J., ARELLA, M. et GUERTIN, C. 1997. Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. *Arch. Virol.* **142**: 1577-1584.
- BAH, A., LUCAROTTI, C. J., ARELLA, M. et GUERTIN, C. 1999. *Choristoneura fumiferana* granulovirus: sequence analysis and 5' characterization of the ORF891. *Arch. of Virol.* **144**:1-10.
- BENZ, G. A. 1986. Historical perspectives. In: Granados RR, Federici BA, eds. *The biology of baculoviruses*. Boca Raton, FL: CRC Press, pp.: 1-35.
- BIRD, F. T. 1963. On the development of granulosis viruses. *J. Insect. Pathol.* **5**: 368-376.
- BIRD, F. T. 1976. Effects of mixed infections of two strains of granulosis virus of the spruce budworm, *Choristoneura fumiferana* (Lepidoptera: Tortricidae), on the formation of viral inclusion bodies. *Can. Ent.* **108**: 865-871.
- BIO-RAD. 1996. Catalog Numbers 170-9080, 1790-9104.

- BLISSARD, G. W. et RHORMANN, F. 1989. Location, sequence, transcriptional mapping, and temporal expression of the gp64 envelope glycoprotein gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* **170**:537-555.
- BLISSARD, G. W. et ROHRMANN, F. 1990. Baculovirus diversity and molecular biology. *Annu. Rev. Entomol.* **35**: 127-155.
- BRUNK, B. P., MARTIN, E. C., ADLER, P. N. 1991. *Drosophila* genes *Posterior Sex Combs* and *Suppressor two* of *zeste* encode proteins with homology to the murine *bmi-1* oncogene. *Nature (London)* **353**: 351-353.
- BURAND, J. P., SUMMERS, M. D. et SMITH, G. E. 1980. Transfection with baculovirus DNA. *J. Virol.* **101**: 286-290.
- BURLEY, S. K., MILLER, A., HARRAP, K. A. et KELLY, D. C. 1982. Structure of the baculovirus nucleocapsid. *Virology* **120**: 433.
- CHAKERIAN, R., ROHRMANN, G. F., NESSON, M. H., LEISY, D. J. et BEAUDREAU, G. S. 1985. The nucleotide sequence of the *Pieris brassicae* granulosis virus granulin gene. *J. Gen. Virol.* **66**:1263-1269.
- COCHRAN, M. A. et FAULKNER, P. 1983. Location of homologous DNA sequences interspersed at five region in the baculovirus AcMNPV genome. *J. Virol.* **45**: 961-970.
- COMTOIS, B. 1988. Notions d'entomologie forestière. Modulo éditeur, pp. 123. ISBN 2-89113-207-6.
- CROOK, N. E. 1981. A comparison of the granulosis viruses from *Pieris brassicae* and *Pieris rapae*. *Virology* **115**:173-181.
- CROOK, N. E., SPENCER, R. A., PAYNE, C. C. et LEISY, D. J. 1985. Variation in *Cydia pomonella* granulosis virus isolates and physical maps of the DNA from three variants. *J. Gen. Virol.* **66**:2423-2430.
- CROOK, N. E., CLEM, R. J. et MILLER, L. K. 1993. An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J. Virol.* **67**: 2168-2174.
- CROOK N. E., JAMES, J. D., SMITH, I. R. L. et WINSTANLEY, D. 1997. Comprehensive physical map of the *Cydia pomonella* granulovirus genome and sequence analysis of the granulin gene region. *J. Gen. Virol.* **78**: 965-974.

- DOBOS, P. et COCHRAN. 1980. Protein synthesis in cells infected by *Autographa californica* nuclear polyhedrosis virus (AcNPV): the effect of cytosine arabinoside. *Virology* **103**: 446-64.
- DWYER, K. G. et GRANADOS, R. R. 1987. A Physical Map of the *Pieris rapae* granulosis virus genome. *J. Gen. Virol.* **68**:1471-1476.
- DWYER, K. G. et GRANADOS, R. R. 1988. Mapping *Pieris brassicae* granulosis virus transcripts and their *in vitro* translation products. *J. Virol.* **62**: 1535-1542.
- EVANS, R. M., HOLLENBERG, S. M. 1988. Zinc fingers: guilt by association. *Cell* **52**: 1-3.
- EVAN, T. et FELSENFELD, G. 1989. The erythroid-specific transcription of factor Eryf1: a new finger protein. *Cell* **58**: 877-885.
- FREEMONT, P. S., HANSON, I. M. et TRAWSDALE, J. 1991. A novel cysteine-rich sequence motif. *Cell* **64**: 483-484.
- FU, Y. H. et MARZLUF, G. A. 1990. *nit-2*, the major nitrogen regulatory gene of *Neurospora crassa*, encodes a protein a putative zinc finger DNA-binding domain. *Mol. and Cel. Biol.* **10**: 1056-1065.
- FALCON, L. A. et HESS, R. T. 1985. Electron microscopic observations of multiple occluded virions in the granulosis virus of the Codling moth, *Cydia pomonella*. *J. Invert. Pathol.* **45**: 356.
- FRIESEN, P. D. et MILLER, L. K. 1986. The regulation of baculovirus gene expression in Doerfler W, Boehm P ed. *The molecular virology of baculoviruses*. Berlin: Springer-Verlag, 1986: 301-50.
- FUCHS, L. Y., WOODS M. S. et WEAVER, R. F. 1983. Viral transcription during *Autographa californica* nuclear polyhedrosis virus infection: a Novel RNA polymerase induced in infected *Spodoptera frugiperda* cells. *J. Virol.* **48**: 641-646.
- GOTO, C., MINOBE, Y. et LIZUKA, T. 1992. Restriction endonuclease analysis and mapping of the genomes of granulosis viruses isolated from *Xestia c-nigrum* and five other noctuid species. *J. Gen. Virol.* **73**: 1491-1497.
- GOTO, C., HAYAKAWA, T. et MAEDA, S. 1998. Genome organization of *Xestia c-nigrum* granulovirus. *Virus Genes* **16** (2): 199-210.

- GRANADOS, R. R. 1978. Early events in the infection of *Heliothis zea* midgut cells by a baculovirus, *Virology* **90**:178.
- GUETTIG, R. R. et McCARTHY, W. J. 1982. Genotypic variation among wild isolates of *Heliothis* spp. nuclear polyhedrosis viruses from different geographical regions. *Virology* **117**: 245-252.
- HAGEDORN, H. H., HILDEBRAND, J. G., KIDWELL, M. G. et LAW, J. H. 1989. Molecular insect science. Proceedings of the International Symposium on Molecular Insect Science, October 22-27, in Tucson, Arizona. Plenum Press, New York 199, pp 113.
- HARTIG, P. C., CARDON, M. C. et KAWANISHI, C. Y. 1991. Insect virus: assays for viral replication and persistence in mammalian cells. *J. Virol. Meth.* **31**: 335-344.
- HARTIG, P. C., CHAPMAN M, A., HATCH G, G. et KAWANISHI, C. Y. 1989. Insect virus: assays for toxic effects and transformation potential in mammalian cells. *Appl. Envir. Microbiol.* **55**:1916-1920.
- HASHIMOTO, Y., CORSARO, B. G. et GRANADOS, R. R. 1991. Location and nucleotide sequence of the gene encoding the viral enhancing factor of the *Trichoplusia ni* granulosis virus. *J. Gen. virol.* **72**:2645-2651.
- HASHIMOTO, Y., HAYASHI, K., OKUNO, Y., HAYAKAWA, T., SAIMOTO, A., GRANADOS, R. R. et MATSMOTO, T. 1996. Physical mapping and identification of interspersed homologous sequences in the *Trichoplusia ni* granulosis virus genome. *J. Gen. Virol.* **77**: 555-563.
- HAWTIN, R. E., ZARKOWSKA, T., ARNOLD, K., THOMAS, C. J., GOODAY, G. W., KING, L. A., KUZIO, J. A., POSSEE, R. D. 1997. Liquefaction of *Autographa californica* polyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. *Virology* **238**: 243-253.
- HEWICK, R. M., HUNKAPILLER, M. W., HOOD, L. E. et DREYER, W. S. 1981. A gas-liquid solid phase peptide and protein sequenator. *J. Biol. Chem.* **256**: 7990-7997.
- HOOFT VAN IDEKINGE, B. J. L., SMITH, G. E. et SUMMERS, M. D. 1983. Nucleotide sequence of the polyhedrin gene of *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **131**: 561-565.
- HOY, M. A. 1994. Insect molecular genetic an introduction to principles and applications. *Academic Press Limited*: 338.

- HUGHES, D. S., POSSEE, R. D. et KING, L. A. 1993. Activation and detection of a latent baculovirus resembling *Mamestra brassicae* nuclear polyhedrosis virus in *M. brassicae* insects. *Virology* **194**: 608-615.
- HUH, N. E. et WEAVER, R. F. 1990a. Categorizing some early and late transcripts directed by the *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* **71**: 2195-2200.
- HUH, N. E. et WEAVER, R. F. 1990b. Identifying the RNA polymerase that synthesizes specific transcripts of the *Autographa californica* nuclear polyhedrosis virus. *J. Gen. Virol.* **71**: 195-202.
- HUNTER, D. K., DEXEL, T. D. et HOFFMANN, D. F. 1972. On the granulovirus of the indian meal moth, *Plodia interpunctella*. *J. Invert. Pathol.* **20**: 361-363.
- IGNOFFO, C. M. 1968. Specificity of insect viruses. *Bulletin of the entomological society of America* **14**: 265-276.
- JARVIS, D. L., BOHLMAYER, D. A. et GARCIA, Jr. A. 1991. Requirement for nuclear localization and supramolecular assembly of a baculovirus polyhedrin protein. *Virology* **185**:795-810.
- JARVIS, D. L. et GARCIA, A. Jr. 1994a. Long-term stability of baculoviruses stored under various conditions. *Biotechniques* **16**: 508-513.
- JARVIS, D. L. et GARCIA, A. Jr. 1994b. Biosynthesis and processing of the *Autographa californica* nuclear polyhedrosis virus gp64 protein. *Virology* **205**: 300-313.
- JEHLE, J. A., BACKHAUS, H. FRITSCH, E. et HUBER, JÜRIG. 1992. Physical map of the *Cryptophlebia leucotreta* granulosis virus genome and its relationship to the genome of *Cydia pomonella* granulosis virus. *J. Gen. Virol.* **73**:1621-1626.
- JEHLE, J. A. et BACKHAUS, H. 1994. The granulin gene region of *Cryptophlebia leucotreta* granulosis virus: sequence analysis and phylogenetic considerations. *J. Gen. Virol.* **75**: 3667-3671.
- KANG, W., TRISTEM, M., MAEDA, S., CROOK, N. E. et O'REILLY, D. R. 1998. Identification and characterization of the *Cydia pomonella* granulovirus cathepsin and chitinase genes. *J. Gen. Virol.* **79**: 2283-92.

- KAUPP, W. J. et EBLING, P. M. 1993. Horseradish peroxidase-labelled probes and enhanced chemiluminescence to detect baculoviruses in gypsy moth and eastern spruce budworm larvae. *J. Virol. Meth.* **44**: 89-98.
- KIDD, M. et EMERY, V. C. 1993. The use of baculoviruses as expression vectors. *Applied biochemistry and biotechnology* **42**: 137-159.
- KNEBEL-MÖRSDORF, D., KREMER, A. et JAHNEL, F. 1993. Baculovirus gene ME53, which contains a putative zinc finger motif, is one of the major early-transcribed genes. *J. Virol.* **67**: 753-758.
- KOOL, M. et VLAK, J. M. 1993. The structural and functional organization of the *Autographa californica* nuclear polyhedrosis virus genome. *Arch. Virol.* **130**: 1-16.
- KOOL, M., VOETEN, J. T. M., GOLDBACH, R. W., TRAMPER, J. et VLAK, J. M. 1993. Identification of seven putative origins of *Autographa californica* nuclear polyhedrosis virus DNA replication. *J. Gen. Virol.* **74**: 2661-2668.
- KOZAK, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**: 283-292.
- KRAPPA, R., BEHN-KRAPPA, J. F., Doerfler, W. et KNEBEL-MÖRSDORF, D. 1992. Differential factor binding at the promoter of early baculovirus gene PE38 during viral infection: GATA motif is recognized by an insect protein. *J. Virol.* **66**: 3494-3503.
- KUDLA, B., CADDICK, M. X., LANGDON, T., MARTINEZ-ROSSI, N. M., BENNETT, C. F., SIBLEY, S., DAVIES, R. W. et ARST, Jr. H. N. 1990. The regulatory gene are a mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutation affecting specificity of gene activation alter a loop residue of a putative zinc finger. *EMBO J.* **9**: 1355-1364.
- KUMAR, S. et MILLER, L. K. 1987. Effects of serial passage of *Autographa californica* nuclear polyhedrosis virus in cell culture. *Virus Research* **7**: 335-350.
- LANGRIDGE, W. H. R. et BALTER, K. 1981. Protease activity associated with the capsule protein of *Estigmene acraea* granulosis virus. *Virology* **114**: 595.
- LATCHMAN, D. S. 1990. Eukaryotic transcription factors. *Bioch. J.* **270**: 281-289.

- LEISY, D. J., ROHRMANN, G. F., NESSON, M., BEAUDREAU, G. S. 1986. Nucleotide sequence and transcriptional mapping of the p10 gene region of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* **153**: 157-67.
- LU, A. et CARSTENS, E. B. 1991. Nucleotide sequence of a gene essential for viral DNA replication in the baculovirus *Autographa californica* nuclear polyhedrosis virus. *Virology* **181**: 336-347.
- LU, A. et MILLER, L. K. 1994. Identification of three late expression factor genes within the 33.8- to 43.4-map-unit region of *Autographa californica* nuclear polyhedrosis virus. *J. Virol.* **68**: 6710-6718.
- MILLER, R., PEARSON, M. N., RUSSEL, R. L. Q. et ROHRMANN, G. F. 1990. A capsid-associated protein of the multicapsid nuclear polyhedrosis virus of *Orgyia pseudotsugata*: genetic location, sequence, transcriptional mapping, and immunocytochemical characterization. *Virology* 133-144.
- MILLER, L. K. 1996. Insect viruses. In *Fields Virology* third edition. Lippincott-Raven Publishers, Philadelphia, pp.:533-556.
- MORIS, T. D. et MILLER, L. K. 1994. Mutational analysis of a baculovirus major late promoter. *Gene* **140**: 147-153.
- MURPHY, F. A., FAUQUET, C. M., BISHOP, D. H. L., GHABRIAL, S. A., JARVIS, A. W., MARTELLI, G. P., MAYO, M. A. et SUMMERS, M. D. 1995. Virus Taxonomy. Six Report of the international Committee on Taxonomy of Viruses. *Arch. Virol. Suppl.* 10, Springer-Verlag, Wien.
- MYERS, R., FISCHER, S. G., LERMAN, L. S. et MANIATIS, T. 1985. Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucl. Acid research* **13**:3131-3145.
- MYERS, R., M, LUMELSKY, N, LERMAN, L. S. et MANIATIS, T. 1985. Detection of single base substitutions in total genomic DNA. *Nature* **317**:495-497.
- ORITTA, M., SUZUKI, Y., SEKIYA, T. et HAYASHI, K. 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**: 874-879.
- O'REILLY, D. R. et MILLER, L. K. 1989a. A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. *Science* **245**: 1110-1112.

- O'REILLY, D. R. et MILLER, L. K. 1989b. Regulation of expression of a baculovirus *egt* gene. *J. Virol.* **64**: 1321-1328.
- O'REILLY, D. R. et MILLER, L. K. 1992. Improvement of a baculovirus pesticide by deletion of the *egt* gene. *Bio/Technology* **9**: 1086-1089.
- PHARMACIA BIOTECH Inc. 1994. QuickPrep mRNA Purification Kit: Instructions. XY-025-00-07.
- PEARSON, M. N., BJORNSON, R. M., PEARSON, G. D. et ROHRMANN, G. F. 1992. The *Autographa californica* baculovirus genome: evidence for multiple replication origin. *Science* **257**: 1382-1384.
- PEARSON, M. N. et ROHRMANN, G. F. 1995. *Lymantria dispar* nucleopolyhedrosis virus homologous regions: characterization of their ability to function as replication origins. *J. Virol.* **69**: 213-221.
- REDEMANN, N., GAUL, U. et JACKLE, H. 1988. *Nature (London)* **332**: 90-92.
- RICE, W. C. et MILLER, L. K. 1986. Baculovirus transcription in the presence of inhibitors and in non permissive *Drosophila* cells. *Virus Res.* **6**: 155-172.
- ROHRMANN, G. F. 1986. Polyhedrin structure. *J. Virology* **67**: 1499-1513.
- ROHRMANN, G. F. 1992. Baculovirus structural proteins. *J. Gen. Virol.* **73**: 749-761.
- ROELVINK, P. W., CORSARO, B. G., GRANADOS, B. G. 1995. Characterization of the *Helicoverpa armigera* and *Pseudaletia unipuncta* granulovirus enhancing genes. *J. Gen. Virol.* **76**: 2693-705.
- SAMBROOK, J., FRITSCH, E. F. et MANIATIS, T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- SALVATO, M. S., SHIMOMAYE, E. M. 1989. The completed sequence of lymphocytic choriomeningitis virus reveals a unique RNA structure and a gene for a zinc finger protein. *J. Virol.* **173**: 1-10.
- SHEFFIELD, V. C., COX, D. R., LERMAN, L. S. et MYERS, R. M. 1989. Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA by the polymerase chain reaction results in improved detection of single-base changes. *Prot. Natl. Acad. Sci. USA* **86**: 232-236.

- SHEFFIELD, V. C., COX, D. R. et MYERS, R. M. 1990. Identifying DNA polymorphisms by denaturing gradient gel electrophoresis. PCR protocol in Academic Press, Inc:206-218.
- SHEFFIELD, V. C., BECK, J. S., KWITEK, A. E., SANDSTROM, D. W. et STONE, E. M. 1992. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitution. *Genomics* **16**:325-332.
- SMITH, I. R. L. et CROOK, N. E. 1988. *In vivo* isolation of baculovirus genotypes. *Virology* **166**:240-244.
- SMITH, I. R. L. et CROOK, N. E. 1993. Characterization of new baculovirus genotypes arising from inoculation of *Pieris brassicae* with granulosis viruses. *J. Gen. Virol.* **74**: 415-424.
- SMITH, I. et GOODALE, C. 1998. Sequence and *in vivo* transcription of *Lacanobia oleracea* granulovirus egt. *J. Gen. Virol.* **79**: 405-413.
- SMIRNOFF, W. A. 1979. Safety tests for the baculovirus of *Thymelicus lineola* (Lepidoptera: Hesperidae). *Can. Ent.* **111**: 459-464.
- SKEIKY, Y. A. W., IATROU, K. 1991. Synergistic interactions of silkmouth chorion-promoter binding factors. *Mol. and Cel. Biol.* **11**: 1954-1964.
- SPIETH, J., SHIM, Y. H., LEA, K., CONRAD, R. et BLUMENTAL, T. 1991. *elt-1*, an embryonically expressed *Caenorhabditis elegans* gene homologous to the GATA transcription factor family. *Mol. and Cel. Biol.* **11**: 4651-4659.
- SUMMERS, M. D. et ANDERSON, D. L. 1972. Granulosis virus deoxyribonucleic acid: a closed, double-stranded molecule. *J. Virol.* **9**:710-713.
- TANADA, Y., INOUE, H., HESS, R. T et O. M. I., E. M. 1980. Site of action of a synergistic factor of a granulosis virus of the armyworm, *Pseudaletia unipuncta*. *J. Invert. Pathol.* **34**: 249-255.
- THIEM, S. M. et MILLER, L. K. 1989. A baculovirus gene with a novel transcription pattern encodes a polypeptide with a zinc finger and a leucine zipper. *J. Virol.* **63**: 4489-4497.

- TOISTER-ACHITUV, M. et FACTOR, O. 1997. Transcriptional analysis and promoter activity of the *Spodoptera littoralis* multicapsid nucleopolyhedrosis ecdysteroid UDP-glucosyltransferase gene. *J. Gen. Virol.* **78**: 487-491.
- TOYODA, T., SAKAGUCHI, T., IMAI, K., INOCENCIO, N. M., GOTHO, B., HAMAGUCHI, M. et NAGAI, Y. 1987. Structural comparison of the cleavage-activation site of the fusion glycoprotein between virulent and avirulent strains of Newcastle disease virus. *Virology* **158**: 242-247.
- TWEETEN, K. A., BULLA, L. A. et CONSIGLI, R. A. 1977. Supercoiled circular DNA of an insect granulosis virus. *Proc. Natl. Acad. Sci. USA.* **74**:3574-3578.
- TWEETEN, K. A., BULLA, L. A. Jr. et CONSIGLI, R. A. 1978. Characterization of an alkaline protease associated with a granulosis virus of *Plodia interpunctella*. *J. Virol.* **26**: 702.
- TWEETEN, K. A., BULLA, L. A. Jr. et CONSIGLI, R. A. 1980. Characterization of an extremely basic protein derived from granulosis virus nucleocapsids. *J. Virol.* **33**: 866.
- VANOERS, M. M., MALARME, D., JORE, J. M. P. et VLAK, J. M. 1992. Expression of the *Autographa californica* nuclear polyhedrosis virus p10 gene: effect of polyhedrin gene expression. **123**: 1-11.
- WINSTANLEY, D. et CROOK, N. E. 1993. Replication of *Cydia pomonella* granulosis virus in cells cultures. *J. Gen. Virol.* **74**:1599-1609.
- WIRTH, U. V., FRAEFEL, C., VOGT, B., VLCEK, C., PACES, V. et SCHWYZER, M. 1992. Immediate-Early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1 are 3' coterminal and encode a putative zinc finger transactivator protein. *J. Virol.* **66**: 2763-2772.
- WOOD, H. A. et GRANADOS, R. R. 1991. Genetically engineered baculoviruses as agents for pest control. *Annu. Rev. Microbiol.* **45**: 69-87.
- ZANOTTO, P. M. de A., SAMPAIO, M. J. A., JOHNSON, D. W., ROCHA, T. L. et MARUNIAK, J. E. 1992. The *Anticarsia gemmatalis* nuclear polyhedrosis virus polyhedrin gene region: sequence analysis, gene product and structural comparisons. *J. Gen. Virol.* **73**: 1049-1056.

ZANOTTO, P. M. de A., KESSING, B. D. et MARUNIAK, J. E. 1993. Phylogenetic interrelationships among baculoviruses: Evolutionary rates and host associations. *J. Invert. Pathol.* **62**:147-164.

COMMUNICATIONS ET PUBLICATIONS

Communications

BAH, A., GUERTIN, C. et ARELLA, M. 1999. Analysis and characterization of the ORF891 and ORF381 of the ChfuGV granulin gene region. Regroupement des biologistes Marocains du Canada à Montréal.

BAH, A., GUERTIN, C. et ARELLA, M. 1998. The granulin gene sequence of *Choristoneura fumiferana* granulovirus: The 5' end transcriptional analysis. Société Canadienne des Microbiologistes, Guelph (Ontario).

BAH, A., LUCAROTTI, C. J., ARELLA, M. et GUERTIN, C. 1998. Comparison of different isolates of ChfuGV collected in different geographical areas: variation in susceptibility of spruce budworm larvae. Congrès des entomologistes du Canada à Québec (Québec).

BAH, A., BERGERON, J., ARELLA, M. et GUERTIN, C. 1995. Analyses of the granulin gene sequence of *Choristoneura fumiferana* granulosis virus. Société Canadienne des Microbiologistes, Kingston (Ontario).

BAH, A., BERGERON, J., DICKO, O., CABANA, J., ARELLA, M. et GUERTIN, C. 1995. Potentiel of the granulovirus of the spruce budworm as a biological control agent (Congrès de la tordeuse des bourgeons de l'épinette à Québec).

BAH, A., BERGERON, J., ARELLA, M. et GUERTIN, C. 1995. Molecular characterization of the granulin gene of *Choristoneura fumiferana* granulovirus. Congrès des entomologistes du Québec (Trois-Rivières).

Publications

BAH, A., LUCAROTTI, C. J., ARELLA, M. et GUERTIN, C. *Choristoneura fumiferana* granulovirus: sequence analysis of ORF381 and its adjacent granulin gene transcriptional region. *Virus Research* (soumis pour publication).

BAH, A., LUCAROTTI, C. J., ARELLA, M. et GUERTIN, C. 1999. *Choristoneura fumiferana* granulovirus: sequence analysis and 5' characterization of the ORF891. *Arch. of Virol.* **144**:1-10.

BAH, A., LUCAROTTI, C. J., ARELLA, M. et GUERTIN, C. 1997. Identification and sequence analyses of the granulin gene of *Choristoneura fumiferana* granulovirus. *Arch. Virol.* **142**: 1577-1584.

BAH, A., BERGERON, J., DICKO, O., CABANA, J., ARELLA, M. et GUERTIN, C. 1995. Potentiel of the granulovirus of the spruce budworm as a biological control agent dans: comptes rendus du séminaire sur la TBE pp.: 133-135. Québec.