



Diversité génétique et fonctionnelle des molécules homologues de PA1b chez *Medicago truncatula* Gaertn. ainsi qu'au sein de légumineuses originaires du Liban

Lamis Karaki

► **To cite this version:**

Lamis Karaki. Diversité génétique et fonctionnelle des molécules homologues de PA1b chez *Medicago truncatula* Gaertn. ainsi qu'au sein de légumineuses originaires du Liban. Sciences agricoles. INSA de Lyon; Université Libanaise, 2013. Français. NNT : 2013ISAL0156 . tel-02003474

HAL Id: tel-02003474

<https://theses.hal.science/tel-02003474>

Submitted on 1 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre 2013ISAL0156
Année 2013

Thèse en Cotutelle

Diversité génétique et fonctionnelle des molécules homologues de PA1b chez *Medicago truncatula* Gaertn. ainsi qu'au sein de légumineuses originaires du Liban

Présentée devant
L'Institut national des sciences appliquées de Lyon
L'Université Libanaise

Pour obtenir
Le grade de docteur

Formation doctorale
Biologie Evolutive, Biologie des Populations, Ecophysiologie

Écoles doctorales
Evolution Ecosystèmes Microbiologie Modélisation (E2M2)
Ecole Doctorale des Sciences et Technologie (EDST)

Par
Lamis KARAKI

Soutenue le 12/12/2013 (sous réserve d'acceptation) devant la
Commission d'examen

Jury MM.

Abou Merhi, R.	Professeure (UL, Liban)	Examinatrice
Bou Dagher-Kharrat, M.	Professeure associée (USJ, Liban)	Rapportrice
Burstin, J.	Directrice de Recherches (INRA, Dijon)	Rapportrice
Lopez-Ferber, M.	Professeur (Mines, Alès)	Rapporteur
Rahbé, Y.	Directeur de recherche INRA (INSA, Lyon)	Directeur
Rizk, F.	Professeure associée (UL, Liban)	Co-directrice
Royer, C.	Chargée de Recherches (INRA, Lyon)	Co-directrice
Safi, S.	Professeur, HDR (UL, Liban)	Directeur

Laboratoire de recherche : Biologie Fonctionnelle Insectes et Interactions (BF2I)

Diversité génétique et fonctionnelle des molécules homologues de PA1b chez *Medicago truncatula* Gaertn. ainsi qu'au sein de légumineuses originaires du Liban

Résumé

Les peptides Albumines 1 b sont des membres de la famille structurale des knottines et présentent un potentiel intéressant en tant que composés insecticides. À ce jour, leur diversité parmi les Fabacées a été essentiellement étudiée en utilisant des approches biochimiques et moléculaires. Les ressources bioinformatiques (le séquençage complet du génome, les bases de données transcriptomiques (EST...), les atlas d'expression...) de l'espèce modèle des légumineuses, *Medicago truncatula* Gaertn. (*Mtr*), nous a permis de développer une approche génomique de cette biodiversité. Deux objectifs principaux nous ont guidé à : 1) déchiffrer l'histoire évolutive de la famille A1 dans cette espèce et 2) explorer la biodiversité naturelle afin de découvrir de nouvelles molécules bioactives. L'exploration du génome de *Mtr* a révélé une remarquable expansion, à travers des duplications en tandem, des loci A1 qui retiennent presque toutes la même structure génique canonique (2 exons et 1 intron). L'analyse phylogénétique nous a permis de comprendre l'évolution des gènes A1 intraspécifique et l'analyse de leur expression (EST, puces à ADN) a révélé la distribution de la famille des gènes A1 dans les organes de la plante (tissus) : Cette dernière s'est révélée bien plus diverse que celle connue chez les autres espèces examinées de légumineuses, où la famille était jusque-là principalement graine-spécifique. Selon plusieurs critères, certains peptides ont été sélectionnés puis chimiquement synthétisés et repliés *in vitro* et testés pour leur activité biologique. Parmi eux, un peptide, nommé AG41, isoforme MtrA1013 et issu de l'EST orpheline TA24778_3880, a révélé un pouvoir insecticide élevé et inattendu. L'analyse à grande échelle en présence d'homologues d'A1 des légumineuses a montré l'ancestralité de la fonction insecticide et l'âge de la famille est estimé à plus de 58 million d'années.

Notre étude s'est aussi orientée vers l'analyse de cette famille peptidique chez des légumineuses originaires du Liban. Cette approche par biologie moléculaire nous a permis de caractériser 9 nouveaux gènes chez 6 espèces de Papilionoideae. L'étude plus approfondie de ces gènes au niveau structural et fonctionnel est envisagée.

Afin de relier les variations de structure et d'activité, un système d'expression hétérologue (baculovirus/cellules d'insectes Sf9) a été mis au point. Le peptide recombinant de référence PA1b (Pea Albumin 1 sub-unit b), même exprimé en faible quantité, présente une activité biologique et une masse bien conforme ainsi qu'une structure bien repliée. Ce système a

permis, de même, de produire la proprotéine PA1, forme intermédiaire entre la préproprotéine et le peptide PA1b mature. Cette proprotéine, identifiée pour la première fois, ne présente aucune toxicité envers les cellules Sf9.

Mots-Clés: Albumine 1, A1b, Fabacées, peptide de défense, résistance des plantes, expression hétérologue, knottine, *Medicago truncatula*

Genetic and functional diversity of homologous of PA1b in *Medicago truncatula* Gaertn. and within legumes from Lebanon

Abstract

Albumin 1 b peptides are members of the knottin structural family and display an interesting potential as insecticidal compounds. To date their diversity among Fabaceae was essentially investigated using biochemical and molecular approaches. The bioinformatic resources (full-genome sequencing, EST database, gene expression atlas...) of the Legume model, *Medicago truncatula* Gaertn. (*Mtr*), prompted us to develop a large-scale approach in two ways: 1) to decipher the evolutionary history of A1 family in this species and 2) to explore the natural biodiversity to uncover new bioactive molecules. Exploring *Mtr* genome revealed a remarkable expansion, through tandem duplications, of A1 loci that retain nearly all the primary structure (2 exons and 1 intron) . Phylogenetic analysis has allowed us to understand the evolution of intraspecific A1 genes and the analysis of their expression (EST, microarrays), and revealed the distribution of the A1 gene family in plant organs (tissue): the latter proved to be much more diverse than that seen in other examined legumes species, where the family until then was mainly seed-specific. Selected upon several criteria some peptides were chemically synthesized, folded *in vitro* and assayed for their biological activity. Among them one peptide, named AG41: isoform *Mtr*A1013 (orphan EST : TA24778_3880), revealed a high and unexpected insecticidal power. The large-scale analysis in the presence of legumes A1 homologous showed the ancestry of the insecticidal function and the age of this family is estimated to be more than 58 million years. Our study is also directed towards the analysis of this family of peptide in legumes from Lebanon. This approach based on molecular biology has allowed us to characterize nine new genes in six species of Papilionoideae. The further study of these genes at the structural and functional level is considered.

To link changes in structure and activity, a heterologous expression system (baculovirus / insect Sf9 cells) was developed. The reference recombinant peptide PA1b (Pea Albumin 1 sub-unit b), even expressed in small

quantities, was biologically active and harbouring the expected mass as well as a well-folded structure. This system has enabled also to produce the proprotein PA1, intermediate form between the preproprotein and the mature peptide PA1b. This proprotein, identified for the first time, has no toxicity towards Sf9 cells.

Key-words: Albumin 1, A1b, Fabaceae, defense peptide, plant resistance, heterologous expression, knottin, *Medicago truncatula*

INSA Direction de la Recherche - Ecoles Doctorales – Quinquennal 2011-2015

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	CHIMIE DE LYON http://www.edchimie-lyon.fr Sec :Renée EL MELHEM Bat Blaise Pascal 3° etage Insa : R. GOURDON	M. Jean Marc LANCELIN Université de Lyon – Collège Doctoral Bât ESCPE 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04.72.43 13 95 directeur@edchimie-lyon.fr
E.E.A.	ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE http://edeea.ec-lyon.fr Secrétariat : M.C. HAVGOUDOUKIAN eea@ec-lyon.fr	M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04.72.18 60.97 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr
E2M2	EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION http://e2m2.universite-lyon.fr Insa : H. CHARLES	Mme Gudrun BORNETTE CNRS UMR 5023 LEHNA Université Claude Bernard Lyon 1 Bât Forel 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cédex Tél : 06.07.53.89.13 e2m2@univ-lyon1.fr
EDISS	INTERDISCIPLINAIRE SCIENCES-SANTE http://www.ediss-lyon.fr Sec : Insa : M. LAGARDE	Mme Emmanuelle CANET-SOULAS INSERM U1060, CarMeN lab, Univ. Lyon 1 Bâtiment IMBL 11 avenue Jean Capelle INSA de Lyon 696621 Villeurbanne Tél : 04.72.68.49.09 Fax :04 72 68 49 16 Emmanuelle.canet@univ-lyon1.fr
INFOMATHS	INFORMATIQUE ET MATHEMATIQUES http://infomaths.univ-lyon1.fr Sec :Renée EL MELHEM Bat Blaise Pascal 3° etage infomaths@univ-lyon1.fr	Mme Sylvie CALABRETTO LIRIS – INSA de Lyon Bat Blaise Pascal 7 avenue Jean Capelle 69622 VILLEURBANNE Cedex Tél : 04.72. 43. 80. 46 Fax 04 72 43 16 87 Sylvie.calabretto@insa-lyon.fr
Matériaux	MATERIAUX DE LYON http://ed34.universite-lyon.fr Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry Ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72.43 83 18 Fax 04 72 43 85 28 Jean-yves.buffiere@insa-lyon.fr
MEGA	MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE http://mega.universite-lyon.fr Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry mega@insa-lyon.fr	M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél :04.72 .43.71.70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr
ScSo	ScSo* http://recherche.univ-lyon2.fr/scso/ Sec : Viviane POLSINELLI Brigitte DUBOIS Insa : J.Y. TOUSSAINT	M. OBADIA Lionel Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Tél : 04.78.77.23.86 Fax : 04.37.28.04.48 Lionel.Obadia@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Dédicaces

A ma très chère et adorée grand-mère

A mes très chers parents et sœur

Remerciements

Ce travail de thèse, réalisé en cotutelle entre l'Institut National des Sciences Appliquées de Lyon (INSA de Lyon) et l'Université Libanaise, a bénéficié du soutien de l'Agence Universitaire de la Francophonie (AUF), du Centre national de la recherche scientifique libanais (CNRS-L) ainsi que du Projet CEDRE, que je tiens à remercier en premier lieu.

J'exprime toute ma gratitude à mes deux directeurs de thèse, Yvan Rahbé et Samir Safi. Merci à vous deux de m'avoir fait confiance pendant la thèse.

Merci Yvan pour m'avoir épaulé, conseillé avec beaucoup de disponibilité et de gentillesse et pour être présent pour toutes discussions, questions, conseils et critiques nombreuses et constructives. J'ai été impressionnée par ton ouverture d'esprit, tes connaissances ainsi que par ta puissance de travail et de raisonnement.

Merci Samir, tu es un homme d'une grande culture. Merci pour les encouragements et la confiance que tu m'as accordé.

Je remercie vivement mes deux co-directrices de thèse, Corinne Royer et Francine Rizk, pour avoir assuré la codirection de ce travail, et m'avoir apporté la rigueur scientifique nécessaire à son bon déroulement, je tiens également à les remercier de leur gentillesse et de leur grande disponibilité. Le partage de l'ensemble de leurs connaissances m'a toujours permis de travailler dans les meilleures conditions.

Merci Corinne pour m'avoir contaminée avec ta passion pour la recherche et soutenue au cours de ces derniers trois ans. Je te remercie pour toute ta gentillesse et ton amitié, pour ta disponibilité et ton aide à n'importe quelle heure, chaque jour de la semaine voire les week-ends. Merci pour m'avoir fait voir le vrai travail en équipe. Je n'aurais jamais trouvé meilleur encadrant de thèse que toi...

Merci Francine pour m'avoir encadrée durant ces années avec beaucoup de patience, de disponibilité et de professionnalisme, et pour m'avoir transmise ta passion pour l'enseignement et la recherche...Ce fut très sincèrement un réel plaisir de travailler à tes côtés pendant ces années.

Je remercie Madame la Doyenne Zainab Saad pour m'avoir donné l'opportunité d'effectuer une thèse dans le cadre d'une Cotutelle. Merci pour votre gentillesse.

Je tiens également à remercier les membres du jury, Mesdames Judith Burstin, Magda Bou Dagher-Kharrat et Raghida Abou Merhi ainsi que Messieurs Miguel Lopez-Ferber et Christophe Douady pour avoir accepté d'évaluer ce travail. Merci pour vos remarques et questions constructives qui m'ouvrent de nouvelles pistes pour compléter et poursuivre ce travail.

Je remercie aussi l'ensemble de l'équipe Entomotoxines pour avoir fait de ces trois ans une expérience inoubliable dans une ambiance toujours stimulante et joyeuse.

Merci Frédéric Gressent pour ta disponibilité, ton aide, tes remarques et suggestions et surtout ta patience.

Merci Pedro Da Silva pour plein de raisons, mais surtout pour ton énergie, ton aide, ton enthousiasme, ton amitié et pour tous les rires que nous avons eus.

Merci Vanessa Eyraud pour ta présence tout au long de ma thèse, pour ta gentillesse, ta disponibilité et ta bonne humeur. Tu es devenue plus qu'une amie avec qui j'ai beaucoup échangé lors du travail que l'on a mené en parallèle.

Merci Isabelle Rahioui pour ton aide lors des manipulations (*le Western n'a plus de secret pour toi !*) et pour tout le soutien que tu m'as apporté. Merci pour cette complicité et pour ton amitié.

Merci Catherine Sivignon pour ton aide inestimable et ton optimisme contagieux. C'est une petite phrase toute simple pour te dire que « *Si la Vie est un Jardin, les Amis en sont les Fleurs* »☺.

Je tiens également à remercier tous les membres du laboratoire Biologie Fonctionnelle Insectes et Interactions (BF2I) pour leur aide ainsi que pour leur accueil chaleureux qui m'a permis de m'y sentir comme chez moi pendant ces années. Merci à Monsieur le Professeur Abdelaziz Heddi, directeur du laboratoire, pour ses encouragements incessants et son soutien qui m'ont été très précieux. Merci Gérard, Hubert, Stéfano, Carole, Gabrielle, Patrice, Séverine, Marjolaine, Karen, Agnès, Elyane, Heidi, Catherine, Lionel, Alain, Aurélien et Florent et tous les autres, pour votre sympathique compagnie durant mes séjours à Lyon. Un grand merci Federica, pour ta gentillesse, tes encouragements et ton appui sans faille.

Je remercie tous les membres de la Faculté des Sciences II à l'Université Libanaise : enseignants, chercheurs, techniciens et personnels administratifs avec qui j'ai eu le plaisir de travailler. Merci également à

l'équipe ER030 de m'avoir accueilli durant mes trois séjours à l'Université Libanaise. Merci aux docteurs Aline, Fadia, Myrna et Rania.

Un grand merci à l'ensemble du personnel du laboratoire Baculovirus et Thérapie du CNRS à Saint Christol-les-Ales, pour son accueil et la sympathie qu'ils m'ont témoigné lors de mon séjour chez eux, et très particulièrement à sa directrice Mme. Martine Cerutti pour sa disponibilité et sa collaboration et la participation à la réussite de la production du peptide dans le système baculovirus.

Je tiens à remercier également Christophe Chouabe surtout pour avoir partagé avec moi ses connaissances et son expérience en électrophysiologie. Merci tout simplement pour ton soutien.

Je remercie aussi Céline Brochier-Armanet pour l'aide qu'elle m'a apportée pour toute la partie phylogénétique réalisée pendant ces trois ans de thèse.

Je tiens également à remercier les membres de mon comité de thèse, qui s'est réuni à deux reprises et m'a permis de poursuivre le travail dans la sérénité, en apportant un regard critique sur l'avancement des travaux, et en contribuant à apaiser les crises d'anxiété que tout doctorant subit au cours de sa thèse. J'exprime donc toute ma gratitude à Nathalie Chantret, Florence Piola, Karine Galardo, Mylene Ogliastro et Pascal Marget pour leurs critiques et apports constructifs au cours des comités et réunions consacrés à cette thèse. Je remercie sincèrement Nathalie Chantret pour l'aide qu'elle nous a apportée pour réaliser des manipulations bio-informatiques notamment l'utilisation du programme PAML.

Je remercie également l'ICARDA ainsi que l'Université Saint-Joseph au Liban pour la contribution à la réalisation de mon projet Plantes Libanaises et la fourniture du matériel nécessaire au bon déroulement du projet.

Les remerciements ne seraient rien sans une mention particulière pour l'amitié. Merci à toi, Khaled, d'avoir été là, encore et toujours, et d'avoir supporté stoïquement tous les aléas de la thèse...Merci Lamice Habib pour les bons moments passés ensemble à Lyon qui ont rendu cette belle aventure encore plus agréable. Merci Anji et Nada pour votre bonne humeur, votre disponibilité et soutien moral.

Je n'oublie pas, bien-sûr, ma famille qui m'a toujours permis de faire ce que je voulais et m'a toujours donné les moyens de le faire, qui

m'a soutenue dans les bons comme dans les mauvais moments de la thèse et qui a toujours su être là pour moi. Je profite donc de ce moment pour remercier mon père, Mohamad, ma mère, Inaya et ma soeur, Rawa: «Merci d'être tels que vous êtes» ! Je remercie aussi ma famille «étendue» et tous ceux qui ont pensé à moi très fort ce jour-là et qui se reconnaîtront.

Sommaire

DÉDICACES.....	3
REMERCIEMENTS	4
LISTE DES FIGURES	13
LISTE DES TABLEAUX	17
LISTE DES ABRÉVIATIONS.....	19
LISTE DES ACTIVITÉS EFFECTUÉES AU COURS DE LA THÈSE.....	22
1. <i>Liste des publications</i>	23
a. Article paru dans une revue à comité de lecture	23
b. Article soumis.....	23
c. Article en préparation	23
d. Déclaration d'invention	24
2. <i>Participations à des congrès.....</i>	24
a. Présentations orales	24
b. Posters	25
3. <i>Encadrement de stagiaire</i>	26
4. <i>Formations.....</i>	26
AVANT-PROPOS	27
CHAPITRE 1 INTRODUCTION BIBLIOGRAPHIQUE.....	30
I. PREMIÈRE PARTIE LES KNOTTINES VÉGÉTALES	31
1. <i>Introduction.....</i>	31
2. <i>Diversité structurale.....</i>	32
a. Découverte du premier peptide knottine	32
b. Définition de la terminologie "knottine"	33
c. Divers représentants des knottines	34
3. <i>Biosynthèse.....</i>	37
a. Knottines à chaîne ouverte	37
a.1 Les inhibiteurs de la carboxypeptidase	37
a.2 La famille des Albumines 1	38
b. Knottines à chaîne cyclique	38
4. <i>Activités biologiques.....</i>	39
a. Activité anti-VIH et hémolytique	39
b. Activité antitumorale.....	40
c. Activité antimicrobienne.....	40
d. Activité antihelminthique	41

e.	Activité molluscicide	42
f.	Activité insecticide	42
5.	<i>Bioingénierie et Production hétérologue des knottines</i>	43
6.	<i>Conclusion</i>	49
II.	DEUXIÈME PARTIE PA1B, LA PREMIÈRE KNOTTINE À ACTIVITÉ INSECTICIDE	50
1.	<i>Intérêt potentiel de PA1b pour la lutte contre les ravageurs</i>	50
2.	<i>Structure et propriétés de PA1b</i>	52
3.	<i>Spectre d'hôte de PA1b</i>	55
4.	<i>Mécanisme d'action de PA1b chez les insectes</i>	56
CHAPITRE 2 MEDICAGO TRUNCATULA, ESPÈCE MODÈLE DE LA FAMILLE DES ALBUMINES 1?		58
1.	<i>Introduction Medicago truncatula – Plante modèle des légumineuses</i>	59
a.	Les légumineuses	59
b.	Medicago truncatula	60
c.	Définition de l'approche : à la découverte de la famille des A1b chez Medicago truncatula	62
2.	<i>Résultats</i>	64
	Article Genome-wide analysis identifies gain and loss/change of function within the small multigenic insecticidal Albumin 1 family in Medicago truncatula	64
CHAPITRE 3 LA FAMILLE DES A1B AU SEIN DES LÉGUMINEUSES ORIGINAIRES DU LIBAN		66
1.	<i>Introduction</i>	67
2.	<i>Matériels et Méthodes</i>	69
a.	Choix des espèces testées	69
b.	Extraction d'ADN	70
c.	Recherche de fragments de gènes homologues par PCR	70
d.	Electrophorèse sur gel d'agarose	72
e.	Clonage	72
e.1	Préparation des plasmides portant l'insert	72
e.2	Transformation bactérienne	73
f.	Minipréparation d'ADN plasmidique	73
g.	Séquençage	73
h.	Analyses des séquences	74
3.	<i>Résultats</i>	75
	<i>Recherche des gènes homologues au gène codant pour PA1b dans le génome des plantes libanaises</i>	75
a.	Amplification de fragments de gènes homologues à PA1b (choix des amorces)	75
a.1	Approche par PCR classique	76

b.	Analyse des séquences	79
b.1	Variabilité des séquences	79
b.2	Phylogénie moléculaire.....	81
4.	<i>Discussion</i>	82
CHAPITRE 4 EXPRESSION HÉTÉROLOGUE DE PA1B		85
1.	<i>Introduction</i>	86
a.	Le système baculovirus-cellules d'insectes.....	87
a.1	Le virus	87
a.2	Les cellules.....	90
a.3	Le système d'expression	91
b.	Etat de l'art de la production de PA1b en système d'expression hétérologue	93
2.	<i>Matériels et Méthodes</i>	95
a.	<i>Matériels</i>	95
a.1	Plasmides	95
a.2	Souche bactérienne	96
a.3	Cellules et virus	96
a.4	Anticorps polyclonaux anti-PA1b.....	96
b.	<i>Méthodes</i>	96
b.1	Clonage moléculaire et analyse d'ADN	96
b.1.1	Electrophorèse de l'ADN en gel d'agarose	96
b.1.2	Purification des fragments d'ADN	97
b.1.3	Transformation bactérienne	97
b.1.4	Extraction de l'ADN génomique viral des cellules infectées	97
b.1.5	Analyse par PCR.....	97
b.2	Analyses d'ARN.....	98
b.3	Entretien des cellules d'insectes	98
b.4	Construction des baculovirus recombinants	98
b.5	Production des peptides recombinants	99
b.6	Purification des peptides recombinants	101
b.7	Western-Blot.....	101
b.8	Caractérisation des protéines par spectrométrie de masse	102
b.9	Caractérisation de l'activité des protéines.....	102
3.	<i>Résultats</i>	103
a.	Insertion de l'ADN complémentaire de PA1 et PA1b dans le baculovirus.....	103
b.	Cinétique d'expression des ARN messagers.....	108
c.	Caractérisation du peptide recombinant	109
c.1	HPLC.....	109
c.2	Spectrométrie de masse (MALDI TOF MS)	111

c.3 Western Blot.....	113
d. Caractérisation de l'activité des protéines.....	115
4. <i>Discussion</i>	116
DISCUSSIONS GÉNÉRALES ET PERSPECTIVES.....	121
a. Evolution de la famille des A1bs au sein de l'espèce modèle des Légumineuses <i>Medicago truncatula</i>	123
a.1 L'émergence de la famille multigénique Albumines 1	125
a.2 <i>Medicago truncatula</i> : modèle d'interaction racine-insecte.....	126
b. Ancestralité des Albumines 1.....	126
c. A la découverte de nouvelles molécules homologues de PA1b chez les Légumineuses originaires du Liban.....	127
d. Expression hétérologue de PA1b dans le système baculovirus/ cellules d'insectes.....	128
d.1 Peptide signal d'adressage à caractéristique universelle	128
d.2 Le rôle du peptide PA1a (et de la proprotéine PA1)	128
d.3 Importance du Baculovirus recombinant PA1 dans la lutte biologique.....	129
RÉFÉRENCES BIBLIOGRAPHIQUES.....	132
ANNEXES	143
Annexe 1 : Liste totale des espèces endémiques sélectionnées du patrimoine végétale Libanais	144
Annexe 2 Choix des amorces selon les différents clades phylogénétiques des A1bs de <i>Medicago truncatula</i> (génome version 3.5).....	145
Annexe 3: Résultats du séquençage des homologues de PA1b	149
Annexe 4 : Etape de production de la protéine recombinante par le système baculovirus/cellules d'insectes	153
Annexe 5 Sélection des baculovirus recombinants par la technique des plages de lyse	154
Annexe 6 Le Réactif CellTiter-Blue®	155
Annexe 7 Arbre phylogénétique des légumineuses [S LOUIS '07]	155

Liste des Figures

FIGURE 1 : DIAGRAMMES SCHÉMATIQUES DES DEUX CLASSES DES MOLÉCULES À NŒUD DE CYSTINE. LES BRINS B SONT REPRÉSENTÉS PAR DES FLÈCHES. LES RÉSIDUS DE CYSTÉINE SONT ÉTIQUETÉS I-VI DANS L'ORDRE DE LA RÉGION N VERS LA C-TERMINALE ET LES LIAISONS DISULFURES SONT REPRÉSENTÉES PAR DES LIGNES HACHURÉES. LE PONT DISULFURE PÉNÉTRANT LA BOUCLE DES FACTEURS DE CROISSANCE À NŒUD DE CYSTINE EST LE CYS (I-IV) ALORS QU'IL EST LE CYS (III-VI) CHEZ LES INHIBITEURS À NŒUD DE CYSTINE ET LES CYCLOTIDES [D. J. CRAIK '01B].....	33
FIGURE 2 : SCHÉMA REPRÉSENTATIF DE LA STRUCTURE DES KNOTTINES.....	35
FIGURE 3 : DE GAUCHE À DROITE, FLEURS D' <i>OLDENLANDIA AFFINIS</i> (RUBIACEAE), DE <i>VIOLA ODORATA</i> (VIOLEACEAE) ET DE <i>MOMORDICA COCHINCHINENSIS</i> (CUCURBITACEAE), TROIS PLANTES PRODUCTRICES DE CYCLOTIDES.....	36
FIGURE 4 : BIOSYNTHÈSE DES INHIBITEURS DE LA CARBOXYPEPTIDASE.	37
FIGURE 5 : BIOSYNTHÈSE DE LA FAMILLE DES ALBUMINES 1	38
FIGURE 6 : BIOSYNTHÈSE DE LA KALATA B1 [N. L. DALY '09]	39
FIGURE 7 : REPRÉSENTATION DES RÉSIDUS HYDROPHOBES DE A) KALATA B1 ET B) KALATA B7. LES RÉSIDUS NON-HYDROPHOBES SONT REPRÉSENTÉS EN VERT ET LES RÉSIDUS FORMANT UNE RÉGION SUSCEPTIBLE DE S'INSÉRER DANS LA MEMBRANE CELLULAIRE SONT EXPLICITÉS [N. L. DALY '11B].	41
FIGURE 8 : DIAGRAMME REPRÉSENTANT LA DIVERSITÉ DANS LA SÉQUENCE DES CYCLOTIDES [N. L. DALY '11A]	44
FIGURE 9 : MÉTHODES DE PRODUCTION BIOTECHNOLOGIQUE DE PEPTIDES À PONTS DISULFURE.	46
FIGURE 10 : STRUCTURE DU GÈNE PA1 ET MATURATION DES PEPTIDES PA1B ET PA1A.	53
FIGURE 11 : SÉQUENCE PRIMAIRE DE PA1B.	53
FIGURE 12 : ALIGNEMENT PROTÉIQUE DE LA PRÉPROTÉINE PA1 DE DIFFÉRENTES LÉGUMINEUSES.....	54
FIGURE 13 : STRUCTURE TRIDIMENSIONNELLE DES PA1B [JOUVENAL '03].....	55
FIGURE 14 : SCHÉMA HYPOTHÉTIQUE DE LA V-ATPASE DE <i>MANDUCA SEXTA</i>	57
FIGURE 15 : PROTOCOLE D'AMPLIFICATION D'ADN PAR PCR AVEC DES AMORCES DÉGÉNÉRÉES (CONCENTRATIONS FINALES).....	71
FIGURE 16 : FRAGMENTS D'ADN AMPLIFIÉS PAR PCR AVEC DIFFÉRENTS COUPLES D'AMORCES « PA1 » SUR ADN GÉNOMIQUE DES ESPÈCES VÉGÉTALES. M : MARQUEUR DE POIDS MOLÉCULAIRE ; 1 : <i>LABLAB PURPUREUS</i> ; 2 : <i>ASTRAGALUS EHDENENSIS</i> ; 3 -4-5 : <i>MEDICAGO ROTATA</i> ; 6 : <i>A. ZACHLENSIS</i> ; 7 : <i>LABLAB PURPUREUS</i> ; 8 : <i>A. ANGULOSUS</i> ; 9 : <i>A. ZACHLENSIS</i> ; 10 : <i>A. TRICHOPTERUS</i> ; 11 : <i>A. ANGULOSUS</i> ; 12 : <i>A. ZACHLENSIS</i> . LE GEL A : COUPLE FA-3/R2 ; B : FB-9/R2 (POUR L'ÉCHANTILLON 3) ET FC+9/R2 (POUR L'ÉCHANTILLON 4) ; C : FCN-10/RNC (POUR L'ÉCHANTILLON 5) ; D ET E : FE8G+30/R2 ; F ET G : FA+9/R2 ; H : -1/R2 ; I ET J : +26/R2. LES BANDES CLONÉES QUI ONT PRÉSENTÉ UNE HOMOLOGIE À PA1B SONT ENCADRÉES EN ROUGE.	78

FIGURE 17 : ALIGNEMENT PROTÉIQUE DES SÉQUENCES CARACTÉRISÉES DANS NOTRE TRAVAIL ET DES SÉQUENCES ISSUES DE LA LITTÉRATURE.	80
FIGURE 18 : ARBRE DE PHYLOGÉNIE MOLÉCULAIRE (SÉQUENCES "INCOMPLÈTES", BIONJ, MAXIMUM DE VRAISEMBLANCE) RÉALISÉ SUR LES SÉQUENCES PROTÉIQUES DES A1B.....	81
FIGURE 19 : REPRÉSENTATION DES DEUX PHÉNOTYPES DU BACULOVIRUS.....	88
FIGURE 20 : CYCLE RÉPLICATIF DU BACULOVIRUS [DEVAUCHELLE '04]	89
FIGURE 21 : DIAGRAMME RELATIF AUX PRINCIPAUX ÉVÉNEMENTS TRANSCRIPTIONNELS ET RÉPLICATIFS DU CYCLE VIRAL. CE SCHÉMA A POUR AMBITION DE DONNER UN CERTAIN NOMBRE DE REPÈRES EN CE QUI CONCERNE LES ÉVÈNEMENTS DE TRANSCRIPTION ET RÉPLICATION VIRALE SANS TENIR COMPTE DE L'ASPECT QUANTITATIF [KING '92, ROHRMANN '92].....	90
FIGURE 22 : CONSTRUCTION D'UN BACULOVIRUS RECOMBINANT. DANS CET EXEMPLE, LE GÈNE CIBLE EST CELUI DE LA POLYÉDRINE.....	92
FIGURE 23 : CARTE DU VECTEUR PGEM-T EASY AVEC CDS AJ574795 EN ANTISENS (T7 FORWARD, SP6 REVERSE ET CDS EN ANTI-SENS -> BORDÉ PAR DEUX SITES : KPN I ET BAMHI).	104
FIGURE 24 : A- EN GRISÉ LOCALISATION DES AMORCES R1-PA1, F1-PA1 ET F4-PA1 AU NIVEAU DE LA SÉQUENCE NUCLÉOTIDIQUE DE PA1 (PRÉSENTÉE EN SENS). LA RÉGION DE L'AMORCE F1-PA1, NE S'HYBRIDANT PAS À LA SÉQUENCE DE PA1, S'HYBRIDE AVEC LA SÉQUENCE DU VECTEUR DANS LEQUEL LA SÉQUENCE DE PA1 EST INSÉRÉE. B- SÉQUENCE PROTÉIQUE DE PA1 ; EN JAUNE : SÉQUENCE PROTÉIQUE DE PA1B ET EN VERT : SÉQUENCE PROTÉIQUE DE PA1A.	105
FIGURE 25 : IDENTIFICATION DES POLYÈDRES EN MICROSCOPIE PHOTONIQUE À FAIBLE GROSSISSEMENT SUITE À L'INFECTION PAR LES VIRIONS RECOMBINANTS.	107
FIGURE 26 : CINÉTIQUE D'EXPRESSION DES ARNs DANS LES CELLULES Sf9 INFECTÉES PAR LES VIRIONS RECOMBINANTS PA1B ET PA1.	108
FIGURE 27 : COMPARAISON DES CHROMATOGRAMMES D'HPLC DES PEPTIDES EXTRAITS À PARTIR DES CULOTS CELLULAIRES.....	110
FIGURE 28 : CHROMATOGRAMME D'HPLC: SÉPARATION DE PICS D'ISOFORMES DE PA1B DE LA FRACTION ETOH60 DE LA FARINE DU POIS- <i>PISUM SATIVUM</i> : VARIÉTÉ FRISSON.....	110
FIGURE 29 : COMPARAISON DES CHROMATOGRAMMES D'HPLC DES PEPTIDES EXTRAITS À PARTIR DES MILIEUX DE CULTURE DES CELLULES Sf9	111
FIGURE 30 : ANALYSE SPECTRALE (MALDITOF-MS) OBTENUE POUR LA FORME DU PEPTIDE RECOMBINANT TROUVÉE DANS LES CELLULES Sf9 INFECTÉES PAR LES VIRIONS RECOMBINANTS PA1 (RÉSULTAT CORRESPONDANT AU PIC A).	112
FIGURE 31 : ANALYSE SPECTRALE (MALDITOF-MS) OBTENUE POUR LA FORME DU PEPTIDE RECOMBINANT TROUVÉE DANS LE MILIEU DE CULTURE DES CELLULES Sf9 INFECTÉES PAR LES VIRIONS RECOMBINANTS PA1 (RÉSULTAT CORRESPONDANT AU PIC C).....	113
FIGURE 32 : WESTERN BLOT DES PURIFICATIONS D'HPLC, AVEC POUR ANTICORPS ANTI-PA1B:	114

FIGURE 33 : DIFFÉRENTES COURBES DE VIABILITÉ CELLULAIRE.	115
FIGURE 34: REPRÉSENTATION SCHÉMATIQUE DE LA PRODUCTION DE PA1B DANS LES CELLULES Sf9 INFECTÉES PAR LES VIRIONS RECOMBINANTS PA1.	117
FIGURE 35 : CONVERSION DE LA RÉSAZURINE EN RÉSORUFINE PAR DES CELLULES MÉTABOLIQUEMENT ACTIVES ENTRAÎNANT LA GÉNÉRATION D'UN PRODUIT FLUORESCENT. LA FLUORESCENCE PRODUITE EST PROPORTIONNELLE AU NOMBRE DE CELLULES VIABLES (CELLTITER-BLUE®CELL VIABILITY ASSAY- INSTRUCTIONS FOR USE OF PRODUCTS G8080, G8081 AND G8082-PROMEGA).	155

Liste des Tableaux

TABLEAU 1 : LISTE DE QUELQUES FAMILLES DE KNOTTINES VÉGÉTALES (HTTP://KNOTTIN.CBS.CNRS.FR/KNOTTINS.PHP)	34
TABLEAU 2 : ACTIVITÉ ANTIMICROBIENNE DE QUELQUES KNOTTINES.	41
TABLEAU 3 : FRÉQUENCE TOTALE DE DÉTECTION DES ESPÈCES D'INSECTES AU SEIN DES 190 LOTS DE GRAINS EXTRAITS DE SILOS DE BLÉS DANS LE CADRE DE L'ENQUÊTE ÉCOPROTECT GRAIN AUX PRINTEMPS 2010 PUIS 2011 [DECOIN '12]	51
TABLEAU 4 : LISTE DES ESPÈCES SÉLECTIONNÉES, ORIGINE GÉOGRAPHIQUE ET FOURNISSEUR.	69
TABLEAU 5 : SÉQUENCES DES AMORCES DÉGÉNÉRÉES UTILISÉES EN PCR	70
TABLEAU 6 : DIFFÉRENTS COUPLES D'AMORCES UTILISÉS AINSI QUE LA TAILLE APPROXIMATIVE MINIMALE DES « BANDES » ATTENDUES.	76
TABLEAU 7 : COMPOSITION DES AMORCES UTILISÉES EN PCR.	101
TABLEAU 8 : TITRAGE DES STOCKS VIRALX.	106

Liste des abréviations

Abréviation	Définition
#	Nombre
µg	Microgramme
µM	Micromolaire
AcMNPV	<i>Autographa californica</i> multicapsid nucleopolyhedrovirus
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ARN	Acide ribonucléique
BET	Bromure d'éthidium
BV	Budded virus
CCK	Cyclic cystine-knot
CDS	Coding DNA sequence
cfu	Colony-forming unit
CL ₅₀	Concentration létale provoquant la mort de 50% des individus
CMI	Concentration minimale inhibitrice
Cys	Cystéine
Da	Dalton
dNTP	Mélange des quatre désoxyribonucleotides : dATP (désoxy adénine triphosphate), dCTP (désoxy cytosine triphosphate), dGTP (désoxy guanine triphosphate), dTTP (désoxy thymine triphosphate)
DTT	Dithiothréitol
EDTA	Acide diaminotétracarboxylique (Ethylenediaminetetraacetic acid)
HPLC	Chromatographie liquide haute performance (High performance liquid chromatography)
ICK	Inhibitor cystine-knot
K _i	Constante d'inhibition
ml	Millilitre
<i>Mtr</i>	<i>Medicago truncatula</i>
ODV	Occlusion-derived virus
PA1a	Pea Albumin 1 subunit a
PA1b	Pea Albumin 1 subunit b
PCR	Polymerase chain reaction
pfu	Plaques-forming unit
PH	Polyédrine
pH	Potentiel hydrogène
RMN	Résonance magnétique nucléaire
<i>Sf9</i>	<i>Spodoptera frugiperda</i> 9
SVF	Sérum du veau fœtal

TAE	Tris-acétate
UV	Ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (C ₁₄ H ₁₅ BrClNO ₆)

Liste des activités effectuées au cours de la thèse

1. Liste des publications

a. Article paru dans une revue à comité de lecture

Gressent F., Da Silva P., Eyraud V., **Karaki L.**, Royer C. (2011) **Pea Albumin 1 subunit b (PA1b), a promising bioinsecticide of plant origin.** *Toxins* (Basel) 3:1502-17. DOI: 10.3390/toxins3121502.

Eyraud V., **Karaki L.**, Rahioui I., Sivignon C., Da Silva P., Rahbe Y., Royer C., Gressent F. (2013) **Expression and biological activity of the cystine knot bioinsecticide PA1b (Pea Albumin 1 subunit b).** *Plos One*.

b. Article soumis

Karaki L., Da Silva P., Rizk F., Chouabe C., Chantret N., Eyraud V., Gressent, F., Sivignon C., Rahioui I., Brochier-Armanet, C., Rahbé Y. and Royer C., 2013. **Genome-wide analysis identifies gain and loss/change of function within the small multigenic insecticidal Albumin 1 family in *Medicago truncatula*.** *Genome Biology*.

c. Article en préparation

Karaki L., Da Silva P., Rizk F., Eyraud V., Gressent, F., Sivignon C., Rahioui I., Cerutti M., Rahbé Y. and Royer C., 2013. **Expression, purification and biological activity of a cystine knot plant toxin in baculovirus/ Sf9 cells system.**

d. Déclaration d'invention

Da Silva P., Gressent F., **Karaki L.**, Rahbe Y., Royer C., 2013.
Titre de l'invention : L'invention réside en un nouveau variant de PA1b (Pea Albumin 1b), une entomotoxine végétale active par ingestion. Ce nouveau variant, AG41, est près de 10 fois plus actif et plus affiné à son récepteur que la molécule originale de pois. Il peut ainsi apporter une solution technique à une lacune non résolue par le PA1b déjà connu. DI RV 13-0040.

2. Participations à des congrès

a. Présentations orales

Karaki L., Eyraud V., Sivignon C., Rahioui I., Da Silva P., Gressent F., Rahbe Y. and Royer C., 2013. **Legume entomotoxic type 1 albumins: Promising candidates for the protection of stored cereals.** Oral presentation presented at the 9th IPSP Conference, 1-4 July 2013, Bordeaux, Talence.

Karaki L., Da Silva P., Eyraud V., Gressent F., Chantret N., Rizk F., Rahbe Y. and Royer C., 2012. **The characterization of the insecticidal small multigenic albumin1 subunit b family in *Medicago truncatula*.** Oral presentation presented at the Natural Products and Biocontrol Congress, 19-21 September 2012, Perpignan.

Eyraud V., **Karaki L.**, Sivignon C., Rahioui I., Da Silva P., Rahbe Y., Royer C. and Gressent F., 2012. **Heterologous production of PA1b, a plant biopesticide.** Oral presentation presented at the Natural Products and Biocontrol Congress, 19-21 September 2012, Perpignan.

Karaki L., Da Silva P., Eyraud V., Gressent F., Chantret N., Brochier-Armanet C., Chouabe C., Rizk F., Rahbe Y. and Royer C., 2013. **Caractérisation de la famille insecticide multigénique albumine 1 sous unité b (A1b) chez *Medicago truncatula*.** Journée des doctorants de

l'école E2M2 : Doc E2M2 Lyon, 19 Mars 2013, Villeurbanne. (A obtenu le prix de la meilleure présentation).

Eyraud V., **Karaki L.**, Rahioui R., Sivignon C., Da Silva P., Rahbe Y., Royer C., Gressent F., 2013. **Synthèse de nouvelles isoformes de PA1b, un pesticide naturel, par expression hétérologue**. CBI, 17ème colloque de biologie de l'insecte, 7-9 Octobre 2013, Montpellier.

b. Posters

Karaki L., Da Silva P., Eyraud V., Gressent F., Chantret N., Brochier C., Rizk F., Rahbe Y. and Royer C., 2012. **Identification of Homologous Genes to PA1b, A Cysteine-Rich Plant Peptide, In *Medicago truncatula***. Poster presented at the third international symposium on antimicrobial peptides. June, 13-15, 2012, Lille (Villeneuve d'ascq), France.

Karaki L., Da Silva P., Eyraud V., Gressent F., Chantret N., Brochier C., Rizk F., Rahbe Y. and Royer C., 2012. Identification of Homologous Genes to PA1b, A Cysteine-Rich Plant **Peptide, In *Medicago truncatula***. Poster presented at Journées Scientifiques à l'Ecole Doctorale de Sciences et Technologie-2011/2012, Hadath, Liban

Karaki L., Da Silva P., Gressent F., Eyraud V., Cerutti M., Rahioui I., Sivignon C., Rizk F., Rahbe Y. and Royer C., 2013. **Heterologous expression of plant toxin PA1b in the Sf9cell/ Baculovirus system**. Poster presented at Journées Scientifiques à l'Ecole Doctorale de Sciences et Technologie-2012/2013, Hadath, Liban.

Karaki L., Da Silva P., Eyraud V., Gressent F., Chantret N., Brochier C., Rizk F., Rahbe Y. and Royer C., 2012. **Identification of Homologous Genes to PA1b, A Cysteine-Rich Plant Peptide, in *Medicago truncatula***. Poster presented at Cellular and molecular biology meeting 2012, Lebanese university, Hadath, Liban.

Eyraud V., **Karaki L.**, Rahioui I., Sivignon C., Rahbe Y., Royer C. and Gressent F. **Production of the peptidic entomotoxin PA1b extracted from pea seeds in *Nicotiana benthamiana***. Antimicrobial Peptide Symposium, Lille, France, 13-15 juin 2012

3. Encadrement de stagiaire

Juin 2012 - Stage de première année IUT (Génie Biologique-Villeurbanne) d'Amandine Demolder

Octobre-Décembre 2012 - Participation à l'initiation de stage d'Imad Alam et Jean Pierre Karam, Master 2 – Université Libanaise Faculté de Science 2.

Février 2013 - Encadrement de deux lycéennes (Séverine Gubian et Véronique Mastroianni) dans le cadre de la formation doctorale : Module accueillir un jeune dans son laboratoire

4. Formations

Les bases de la culture cellulaire : Stage de formation réalisé par Fc-3Bio (Lyon) sous la direction de M. Jean-Francois Prost.

Analyse de séquences génomiques et phylogénie : Stage de formation réalisé à l'UMR 5558 LBBE (UCBL 1) sous la direction de Mme. Céline Brochier-Armanet.

Module accueillir un jeune dans son laboratoire dans le cadre de la formation doctorale (encadrement de deux lycéennes).

Entre innovation et propriété intellectuelle : quelle stratégie adopter pour protéger ses travaux : logiciels, BDD, brevet ? Stage de formation réalisé par Urlist de Lyon sous la direction de M. Samuel Le Cacheux.

Transformer son résultat de recherche innovant en brevet : comment le rédiger et assurer sa protection intellectuelle? Stage de formation réalisé Urlist de Lyon sous la direction de M. Mathieu Maillard.

EndNote : initiation au logiciel de gestion de références bibliographiques : Stage de formation réalisé Urlist de Lyon sous la direction de Mme. Frédérique Cohen ADAD.

Avant-propos

Les plantes représentent une source riche en nutriments pour de nombreux organismes tels que les bactéries, les champignons, les protistes, les insectes et les vertébrés. Malgré l'absence d'un système immunitaire comparable à ceux des animaux, et durant leur longue évolution et confrontation avec les insectes, les plantes ont développé des arsenaux insecticides divers pour contrôler ces agresseurs majeurs ; cette panoplie métabolique est dominée par des substances dites secondaires, à la chimie et la biosynthèse souvent complexes et botaniquement spécialisées. A côté de cette « coévolution chimique », des métabolites d'origine biosynthétique plus simple ont également été utilisés, comme certains polypeptides toxiques.

Les êtres humains dépendent presque exclusivement de plantes pour se nourrir, et les végétaux fournissent aussi de nombreux produits non alimentaires importants comprenant bois, colorants, textiles, médicaments, cosmétiques, savons, caoutchouc, plastiques, encres et les produits chimiques industriels... Comprendre comment les plantes se défendent contre les agents pathogènes et les herbivores est essentielle pour protéger notre approvisionnement alimentaire et par conséquent développer des espèces de plantes hautement résistantes aux maladies. La disponibilité des méthodologies recombinantes a permis de tester l'importance potentielle de nombreux polypeptides dans les fonctions de défense des plantes, notamment contre les insectes.

Au cours du XX^e siècle, des substances naturelles dites peptides « cystine-knot » ont été identifiées [D. C. REES '80]. Ces peptides ont fait l'objet d'une attention grandissante depuis la dernière décennie en raison de leur caractéristique structurale et de leur activité biologique et donc de leur potentiel d'utilisation dans de nombreux domaines industriels (phyto-sanitaire, pharmaceutique...). Les cystine-knot jouent le rôle de substances de défense (peptides antimicrobiens) ou d'attaque (toxines). Leur structure inédite en nœud de cystine leur confère une stabilité accrue contre les dégradations chimiques, protéolytiques ou physiques. Ceci renforce l'activité des peptides en augmentant leur durée de vie.

Chez les eucaryotes, des cystine-knot de structure compacte participent à l'immunité innée. Ils sont impliqués dans des phénomènes de défense de l'organisme dans un environnement riche en micro-organismes. Par exemple, les défensines de mammifères [GANZ '03] sont synthétisées par des cellules en contact permanent avec des micro-organismes comme

les neutrophiles du sang, les cellules de Paneth de l'épithélium intestinal, de l'épiderme ou de la trachée. Les cyclotides [GRUBER '08, N. L. DALY '09] sont des molécules de défense des plantes produits par les organes aériens des plantes. La famille des conotoxines [TERLAU '04], produites par des mollusques du genre *Conus*, se distingue des autres structures peptidiques compactes eucaryotes par le fait que ces molécules servent à la nutrition en paralysant les proies de l'organisme producteur.

L'avancée technologique que les méthodes spectroscopiques ont connue ces dernières années a fourni aux chercheurs de nouveaux outils pour la compréhension de la structure, des propriétés et des fonctions de ces molécules si particulières. La spectrométrie de masse (MS), la résonance magnétique nucléaire (RMN) ou encore le dichroïsme circulaire (CD) sont autant de méthodes de précision qui, utilisées conjointement, participent à cette dynamique.

Chez une famille de plantes dicotylédones appelée couramment Légumineuses, des peptides cystine-knot appelés « albumines de type 1 » ont été identifiés. Ces peptides offrent un exemple remarquable de stabilité d'une structure peptidique. Ces albumines présentent une activité entomotoxique vis-à-vis de certains ravageurs comme les charançons des céréales ou le puceron du pois [GRESSENT '07].

C'est dans ce dernier contexte impliquant la protection des cultures contre les pathogènes, spécifiquement les insectes que s'inscrit ce travail de thèse.

En tant que laboratoire de l'INRA s'inscrivant dans les problématiques du département Santé des Plantes et Environnement, l'équipe EntomToxine de l'unité Biologie Fonctionnelle Insectes et Interaction (UMR 203 INSA/ INRA, BF2I) de Lyon s'intéresse particulièrement à la recherche de nouvelles molécules à activité insecticide chez les plantes, à la connaissance de leurs mécanismes d'action ainsi qu'à la compréhension des relations structure-fonction de ces peptides. Il s'intéresse particulièrement à une famille de molécule, les PA1b (pour Pea Albumin 1 sous-unité b), pour laquelle ils ont été les premiers à décrire une activité entomotoxique [DELOBEL '98].

Mon travail de thèse a porté sur l'étude de la diversité génétique et fonctionnelle des homologues naturels des albumines 1 chez l'espèce modèle des légumineuses *Medicago truncatula* ainsi qu'au sein de quelques légumineuses originaires du Liban.

Le **chapitre 1** consiste en une étude bibliographique portant sur A) : les knottines de structure compacte, avec une présentation des principaux représentants de ces peptides chez les plantes, suivie B) : d'une partie plus détaillée dédiée aux peptides entomotoxiques albumines 1.

Le **chapitre 2** concerne la recherche des homologues des albumines 1 au sein de l'espèce modèle des Légumineuses, *Medicago truncatula* selon une approche *in silico* basée sur un criblage des bases de données disponibles suivie par une étude phylogénétique. Ce chapitre permettra de mieux comprendre les forces évolutives s'exerçant sur les gènes homologues aux albumines 1, les fonctions des gènes, ainsi que leur régulation spatio-temporelle ou en lien avec des facteurs environnementaux (stress hydrique...) dans la plante. Ces résultats sont présentés sous forme d'un article soumis en 2013 dans la revue Genome Biology.

Le **chapitre 3** rassemble les résultats de la recherche des homologues des albumines 1 au sein de quelques espèces de légumineuses originaires du Liban par la voie plus classique de la biologie moléculaire (clonage par homologie).

Le **chapitre 4** décrit la mise en place d'un système de production hétérologue de l'albumine 1b du pois (PA1b) : Baculovirus/ cellules d'insecte Sf9. Il inclura les résultats des tests biologiques réalisés avec le peptide, produit d'expression.

Enfin, une **Discussion** générale de ces résultats permet d'apporter quelques conclusions sur l'étude de ces peptides originaux et d'ouvrir sur de nouvelles perspectives tant dans le domaine fondamental (évolution, relation structure-fonction) que dans le domaine appliqué (lutte contre les ravageurs phytophages ou les insectes vecteurs de maladies humaines ou animales).

Chapitre 1

Introduction bibliographique

I. Première partie

Les knottines végétales

1. Introduction

Les protéines se trouvent dans toutes les cellules vivantes et possèdent une variété d'activités biochimiques. Elles fonctionnent comme des enzymes, des hormones, des antibiotiques et des récepteurs. Elles composent une partie majeure de l'être vivant. Dans ses « **Dialectique de la nature** » Friedrich Engels a écrit que *La vie est le mode d'existence des corps albuminoïdes* [JEAN-MARIE TREMBLAY '68]. On prend ici le corps albuminoïde au sens de la chimie moderne, qui rassemble sous ce nom tous les corps composés de façon analogue à l'albumine ordinaire et appelés aussi substances protéiques. Par conséquent, les scientifiques ont été très intéressés par les peptides et les protéines et ses différentes conformations et fonctions. Le principal objectif dans ce domaine est d'étudier la relation entre la structure et l'activité des peptides biologiquement actifs et d'établir leurs mécanismes moléculaires.

Malgré la présence de plusieurs autres objectifs, la question initiale est la détermination de la structure des peptides ou des protéines. Dans la nature, les polypeptides adoptent un espace conformationnel spécifique déterminée par la séquence d'acides aminés. Sur la base des connaissances des propriétés physiques des acides aminés et leur comportement en milieu aqueux et chimique, il est possible de prédire l'architecture du peptide en se basant sur une structure primaire connue ou de la déterminer expérimentalement. Au cours des cinquante dernières années, une variété de méthodes physiques, chimiques, informatiques et biologiques a été créée pour l'étude de la structure tridimensionnelle de différents peptides. De plus, il a été constaté que l'existence et la typologie de la fonction biologique des peptides sont intimement liées à la conformation, suggérant que l'étude de ces différentes activités pourrait également contribuer à la connaissance de la structure spatiale des protéines. Au cours du XXe siècle, des substances naturelles dites des peptides « knottines » ont été identifiées [D. C. REES '80]. Ces peptides ont fait l'objet d'une attention grandissante depuis la dernière décennie en raison de leur caractéristique structurale et de leur activité biologique et donc de leur potentiel d'utilisation dans de nombreux domaines.

2. Diversité structurale

Les peptides « knottines » constituent une classe structurale de peptides riches en ponts disulfures aux propriétés fascinantes. Les membres de cette classe de molécules sont connus comme des mini protéines car elles montrent toutes les propriétés des protéines de grande taille, à l'exception de la taille. Ils sont petits, compacts et très stables (quelques-uns des membres de cette classe de molécules conservent leurs fonctions biologiques dans l'eau bouillante). Ces peptides possèdent tous de 26 à 48 résidus dont 6 cystéines formant trois ponts disulfures participant à la création d'un nœud de cystine [LAURENT CHICHE '10].

a. Découverte du premier peptide knottine

C'est en 1982, que Rees et Lipscomb ont découvert un nouveau type de repliement protéique en déterminant la structure tridimensionnelle d'un inhibiteur de la protéase chez la pomme de terre [D. C. REES '82]. Cette mini-protéine agit comme un inhibiteur de la carboxypeptidase (PCI, *potato carboxypeptidase inhibitor*) [D. C. REES '80] et peut former des complexes avec plusieurs métallo-carboxypeptidases, en les inhibant d'une manière fortement concurrentielle avec une constante d'inhibition (K_i) dans une gamme nanomolaire [MICHAEL HASS '81]. Cette nouvelle structure caractérisée montre la présence de trois ponts disulfures formant un nœud compact de cystine. Cette molécule est constituée de 39 acides aminés (masse moléculaire : 4295 Da), formant un noyau globuleux de 27 résidus et montrant une absence de la structure secondaire régulière à l'exception d'une courte hélice de 5 résidus et d'un petit feuillet-beta. De son noyau central hydrophobe émerge une queue C-terminale (résidus 35-39) qui sera insérée dans le site actif de la carboxypeptidase et modifie le site de liaison primaire de l'enzyme.

Ces ponts disulfures assemblés dans une structure particulière (formation d'un nœud), partagée par quelques autres peptides bioactifs, est appelée la structure « knottine » [D. LE-NGUYEN '90].

b. Définition de la terminologie "knottine"

Les ponts disulfures entre les paires de résidus cystéine dans certaines protéines forment une signature fonctionnelle unique et sont considérés comme les principaux déterminants de la stabilité de la protéine et de son repliement. Le motif cystine-knot, officiellement identifié comme un motif structural il y a environ 20 ans [MURRAY-RUST '93], est un arrangement de ponts disulfures présent dans les peptides et les protéines provenant d'une grande variété d'espèces. Cet arrangement noué des ponts disulfures est généralement associé à des structures en feuillet β dans les protéines dans lesquelles ils se produisent. Bien que le motif ait été initialement considéré comme caractéristique de certains facteurs de croissance, il est vite apparu que la « cystine-knot » est aussi assez fréquent dans une variété de petits peptides, en particulier les petites toxines riches en cystéine. Toutefois, lorsque les propriétés spatiales du nœud dans ces petits peptides ont été prises en compte, le motif ne pouvait pas se superposer directement avec celles des facteurs de croissance. Cela a conduit à la classification des protéines contenant des nœuds de cystine en deux groupes (Figure 1) [PALLAGHY '94]: Facteur de croissance à nœuds de cystine (GFCK) et les Knottines. Un autre nom a été ensuite proposé pour les knottines : les inhibiteurs cystine-knot (ICK) incluant les cyclotides à chaîne cyclique (cycliques cystine-knot) (CCK).

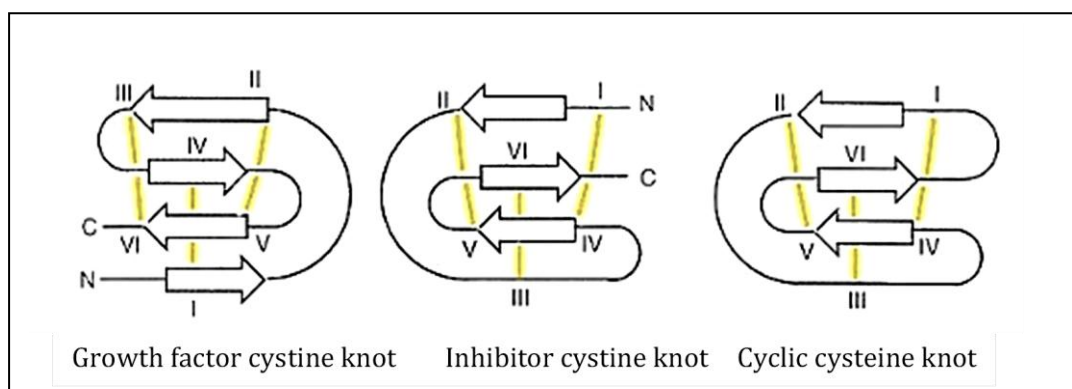


Figure 1 : Diagrammes schématisques des deux classes des molécules à nœud de cystine. Les brins β sont représentés par des flèches. Les résidus de cystéine sont étiquetés I-VI dans l'ordre de la région N vers la C-terminale et les liaisons disulfures sont représentées par des lignes hachurées. Le pont disulfure pénétrant la boucle des facteurs de croissance à nœud de cystine est le Cys (I-IV) alors qu'il est le Cys (III-VI) chez les inhibiteurs à nœud de cystine et les cyclotides [D. J. CRAIK '01b].

La connectivité des ponts disulfure dans toutes ces molécules à nœud de cystine est identique. Généralement six résidus cystéine, étiquet-

tées dans l'ordre depuis l'extrémité N-terminale à l'extrémité C-terminale, sont impliqués dans la formation du nœud. Le nœud est un anneau intégré formé par deux ponts disulfure et leurs segments de squelette, qui est traversé par un troisième pont disulfure. Les trois ponts disulfures sont formés entre CysI et CysIV, CysII et CysV et CysIII et CysVI. La principale caractéristique de ces deux familles, c'est que le pont disulfure pénétrant dans l'anneau chez les GFCK est Cys (I-IV), tandis que pour la famille des knottines est Cys (III-VI). La caractéristique unique qui distingue les CCK des ICK est la nature cyclique du squelette de la protéine [IYER '11].

Le motif « knottine » agit en tant qu'une matrice structurale pour la totalité de la protéine ou pour un domaine précis de la protéine conférant ainsi une structure compacte et remarquablement stable en face des pH extrêmes, des dénaturations thermiques et chimiques ainsi qu'en face d'une attaque protéolytique [IRELAND '06, WERLE '07].

c. Divers représentants des knottines

Les inhibiteurs « cystine-knot » sont présents dans de nombreuses espèces. Cette matrice a été découverte en 1982 avec la découverte du PCI, un inhibiteur de carboxypeptidase de pomme de terre (Rees et Lipscomb, 1982). Elle a été depuis observée dans un grand nombre de familles non apparentées. Ces knottines proviennent d'une variété d'espèces, y compris les champignons, les plantes, les mollusques marins, les insectes et les araignées. En se limitant au règne végétal, nous pouvons citer quelques familles de knottines extraites à partir des plantes (Tableau 1).

Tableau 1 : Liste de quelques familles de knottines végétales (<http://knottin.cbs.cnrs.fr/Knottins.php>)

Famille	Séquence	
	#	Organisme
Knottine d'algue	3	1
Knottine du maïs	8	1
Cyclotides	192	25
Antimicrobiens de plante	16	7
Défensines de plante	4	1
Toxines de plante	28	17

Les inhibiteurs « cystine-knot » sont de faible poids moléculaire et à ce jour, n'ont pas la flexibilité nécessaire pour engendrer des change-

ments conformationnels lors de leurs interactions avec leurs protéines cibles [PALLAGHY '94].

Ces ICKs sont composées d'un feuillet β à 3 brins antiparallèles stabilisés par le nœud de cystine comme présentée dans la Figure 2.

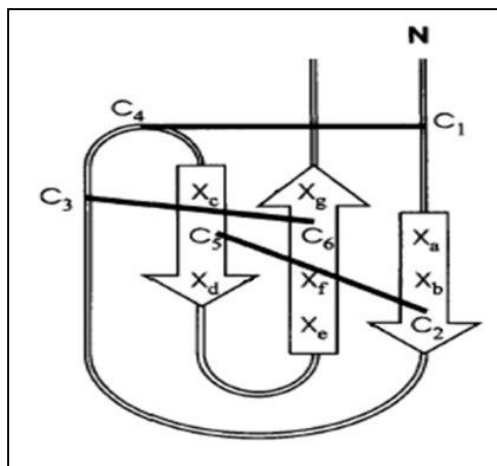


Figure 2 : Schéma représentatif de la structure des knottines.

C représente un résidu cystine et X représente n'importe quel autre résidu (Adapté par [PALLAGHY '94]).

Les signatures très conservées des membres de la famille des knottines sont : 1) l'espace entre les résidus des cystéines adjacentes et 2) le mode de connectivité des ponts de disulfures. Les boucles du squelette, cependant, sont variables selon la longueur et la séquence d'acides aminés.

Parmi les ICKs végétales, nous pouvons citer de plus les inhibiteurs de trypsine découvertes en 1989 chez des courges comme le CMTI-I de *Cucurbita maxi* et l'EETI d'*Ecballium elaterium* [BODE '89, L. CHICHE '89, HEITZ '89]. PA1b [JOUVENSAL '03] et la léginsuline [YAMAZAKI '03], deux homologues trouvés dans les graines du pois et du soja, respectivement, rejoignent en 2003 le club des knottines.

De plus, dans cette famille de knottines se trouvent les cyclotides (CCKs), encore appelés « cyclo-peptides », une grande famille de molécules présentes exclusivement dans les plantes [GORANSSON '12], dont la particularité est de présenter une cyclisation en tête-à-queue [D. J. CRAIK '06, D. J. CRAIK '07]. Ce sont des peptides dont la taille est comprise entre 28 et 37 acides aminés. Ils présentent un squelette peptidique cyclisé par une liaison peptidique entre les extrémités N-terminale et C-terminale, et possèdent, de même, six résidus cystéines engagés dans des ponts disulfure, qui leur confèrent une structure nouée dont le motif structural appartenant à la famille des knottines est appelé «cyclic cystine knot » (CCK) [D. J.

CRAIK '99, DAVIDJ CRAIK '01a]. En 1973, l'équipe de Gran isola un peptide de 29 résidus à partir de la plante *Oldenlandia affinis* utilisée en médecine traditionnelle chez des tribus congolaises [**GRAIN '73, GRAN '73**]. Cette molécule, fut nommée kalata B1, en référence au nom que les indigènes donnaient à leurs pratiques médicales, « kalata-kalata ». C'est seulement en 1995 que la structure tridimensionnelle de ce peptide révéla les particularités structurales (cyclisation) de cette nouvelle famille de composés [**SAETHER '95**]. Le premier cyclotide venait d'être découvert. Depuis, les mises en évidence de motif CCK se sont accentuées et le nombre de cyclotides qui existeraient dans la biodiversité est évalué aujourd'hui à environ 50 000 [**GRUBER '08**], principalement dans les familles des rubiacées, des violacées et des cucurbitacées (Figure 3).

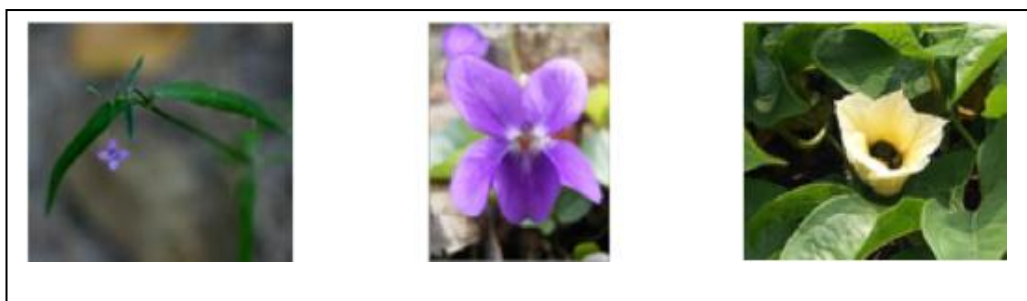


Figure 3 : De gauche à droite, fleurs d'*Oldenlandia affinis* (Rubiaceae), de *Viola odorata* (Violaceae) et de *Momordica cochinchinensis* (Cucurbitaceae), trois plantes productrices de cyclotides.

En 2006, Jason Mulvenna a montré la présence des cyclotides dans une variété importante socialement et économiquement de plantes cultivées au sein des Poacées, dont le riz, le maïs et le blé [**MULVENNA '06**]. Cette découverte est potentiellement d'une grande importance car les cyclotides ont été trouvés dans les monocotylédones, ainsi que dans les deux lignées d'astéridées et de rosidées des dicotylédones vraies [**D. C. DALY '01**], ce qui suggère que les cyclotides peuvent être dérivés d'un gène ancestral qui existait avant la divergence des lignées monocotylédones et dicotylédones il y a ~150 millions années.

Les cyclotides ont été classifiés en trois catégories appelées Möbius, bracelet et inhibiteur de trypsine selon l'agencement de leur ponts disulfure et la présence ainsi que la position d'une hélice α dans la séquence.

Les knottines, que ce soit à chaîne ouverte ou circulaire (CCKs), sont remarquablement stables face aux pH extrêmes, à la dénaturation thermique et chimique et aux attaques protéolytiques [**IRELAND '06, WERLE '07**]. C'est probablement le résultat de l'énorme rigidité conforma-

tionnelle qui est introduite grâce aux liaisons disulfures du noyau noué [COLGRAVE '04, KRATZNER '05].

3. Biosynthèse

La synthèse des knottines se réalise de différentes façons selon la structure du gène et la nature de la molécule. Malgré cette différence dans la synthèse de ces différentes molécules, le résultat final est le même : formation d'un nœud de cystine grâce aux trois ponts disulfures liants 6 cysteines en 1-4, 2-5, 3-6. Dans la suite de cette partie, nous allons présenter quelques exemples de synthèse de différentes knottines ou ICKs, que ce soit à chaîne ouverte ou à chaîne cyclique (CCK).

a. Knottines à chaîne ouverte

a.1 Les inhibiteurs de la carboxypeptidase

La synthèse des inhibiteurs de la carboxypeptidase est composée de trois étapes distinctes [HABIB '07]. Le peptide est d'abord synthétisé sous la forme d'un pré-peptide comportant une séquence signal dans sa partie N-terminale, la séquence de l'inhibiteur mature et une petite région propeptide dans sa partie C-terminale. Ce peptide subit un clivage protéolytique et des modifications post-traductionnelles pour donner le peptide mature et replié (Figure 4).

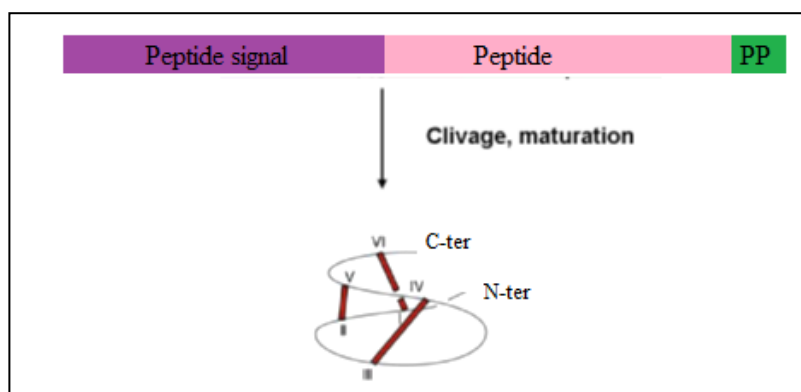


Figure 4 : Biosynthèse des inhibiteurs de la carboxypeptidase.

PP=propeptide

a.2 La famille des Albumines 1

Les albumines 1 sont des knottines entomotoxiques actives contre des insectes de différents ordres, notamment les charançons des céréales et la teigne de la farine [GRESSENT '03]. Cette famille a été identifiée dans la graine de pois. Le PA1b, citée ci-dessus, appartient à cette famille. Tous les homologues de PA1b sont synthétisés d'une manière semblable. Le peptide est d'abord synthétisé sous la forme d'un pré-propeptide comportant une séquence signal dans sa partie N-terminale, la séquence de la toxine PA1b, une petite région propeptide suivie d'un autre peptide PA1a et une petite région propeptide dans sa partie C-terminale. Ce peptide subit un clivage protéolytique et des modifications post-traductionnelles pour donner le peptide mature et replié (Figure 5). La biosynthèse des albumines 1 sera détaillée dans la deuxième partie du chapitre 1.

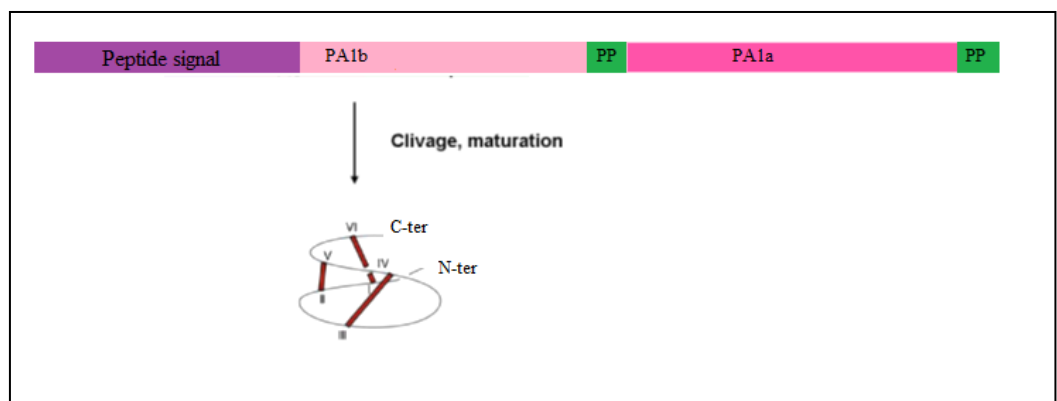


Figure 5 : Biosynthèse de la famille des Albumines 1

b. Knottines à chaîne cyclique

Les cyclotides sont synthétisés sous forme d'un pré-propeptide constitué par quatre segments peptidiques (Figure 6). La région N-terminale détermine une zone d'interaction avec le réticulum endoplasmique (ER), la région suivante est une région dite « pro- » puis on trouve une région N-terminale (NTR) suivie de la séquence peptidique du cyclotide proprement dit (ici B1 pour la kalata B1). On sait que le peptide subit une ou plusieurs étapes de clivage puis de maturation (cyclisation et formation des ponts disulfure) mais l'ordre de ces étapes ainsi que les enzymes mises en jeu sont mal connus [GRUBER '08].

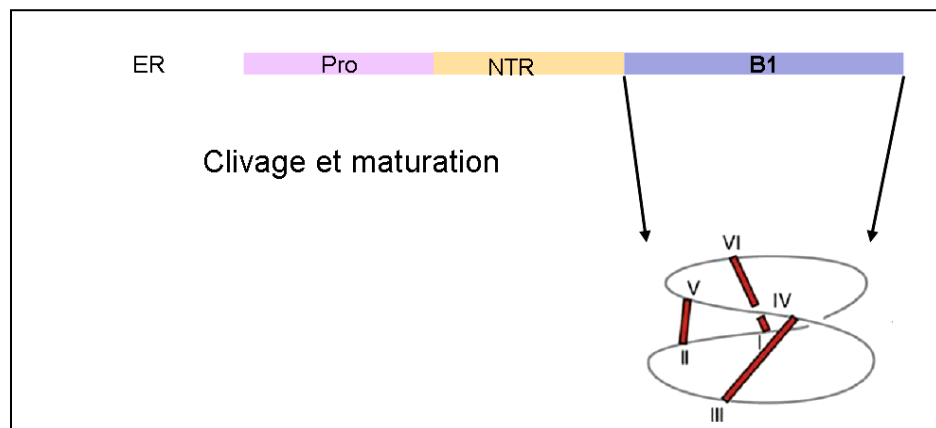


Figure 6 : Biosynthèse de la kalata B1 [N. L. DALY '09]

4. Activités biologiques

L'activité biologique des knottines s'est avérée être intéressante et montre une forte diversité [KOLMAR '08]. Ces fonctions biologiques variées sont conférées par différentes parties de la molécule, parmi lesquelles les boucles qui participent à une ou plusieurs fonctions.

La majorité de ces fonctions est exercée par les cyclotides végétaux. Cela est dû d'une part à leur structure tridimensionnelle à caractère amphiphile capable de traverser des membranes cellulaires et exerçant une fonction déterminée et d'autre part aux liaisons amides covalentes formées entre les extrémités N- et C-terminales et s'ajoutant aux trois ponts disulfures interconnectés rendant la structure encore plus compacte et rigide [D. J. CRAIK '10, N. L. DALY '11a].

a. Activité anti-VIH et hémolytique

La plupart des cyclotides possèdent une activité anti-VIH et hémolytique, avec une concentration inhibitrice médiane (CI_{50}) de l'ordre de 100 nM [GUSTAFSON '00, GUSTAFSON '04], sans doute en ciblant la membrane plasmique des cellules infectées par le virus. Malheureusement, il apparaît

que la spécificité de ces peptides pour les cellules infectées par rapport aux cellules saines n'est pas assez importante pour qu'ils fassent encore l'objet d'une recherche clinique poussée [IRELAND '08].

b. Activité antitumorale

Les plantes, par exemple, les familles des Violacées, Rubiacées, et cucurbitacées et des plantes marines, sont une importante ressource de peptides ayant le potentiel d'être des agents anticancéreux. Des études ont montré que ces peptides présentent des effets inhibiteurs marqués sur la prolifération de diverses lignées cellulaires tumorales de vertébrés. Les cyclotides Varv A et Varv F isolés à partir de *Viola arvensis*, et cycloviolacin O2 isolés à partir de *V. odorata* présentent de fortes activités cytotoxiques. Cycloviolacin O2 est le plus efficace contre toutes les lignées cellulaires (CI₅₀ 0,1- 0,3 µM), suivie par Varv A (IC₅₀ de 2,7-6,35µM) et Varv F (IC₅₀ de 2,6-7,4µM), respectivement [XUDONG MAA '06]. De la même manière, l'inhibiteur de la carboxypeptidase des pommes de terre PCI inhibe la croissance des cellules tumorales. C'est un antagoniste de l'EGF humain [BLANCO-APARICIO '98].

c. Activité antimicrobienne

De nombreux cyclotides présentent une activité antimicrobienne importante avec des concentrations minimales inhibitrices (CMI) comprises entre 0,2 et 50 µM et dirigées contre de nombreuses souches bactériennes et fongiques, notamment *Escherichia coli*, *Staphylococcus aureus* et *Candida kefyr* [TAM '99]. Ces cyclotides s'insèrent dans la membrane cellulaire grâce à leurs résidus hydrophobes plutôt que grâce à leur structure tridimensionnelle qui n'apparaît pas jouer un rôle très important dans l'orientation du peptide lors de son insertion dans la membrane [SHENKAREV '08] (Figure 7).

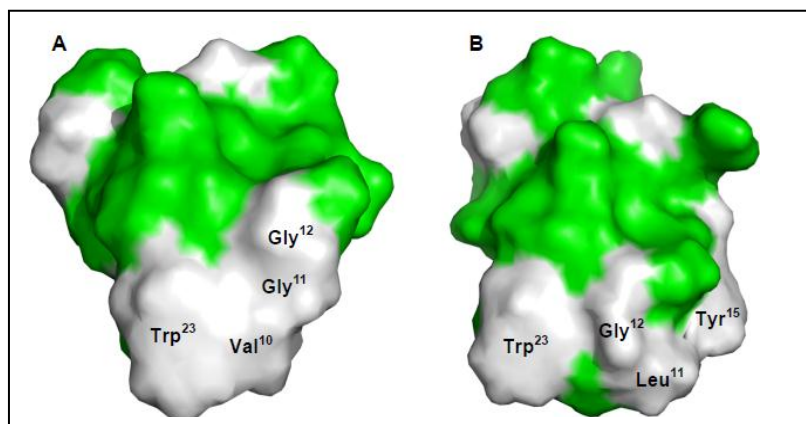


Figure 7 : Représentation des résidus hydrophobes de A) kalata B1 et B) kalata B7. Les résidus non-hydrophobes sont représentés en vert et les résidus formant une région susceptible de s'insérer dans la membrane cellulaire sont explicités [N. L. DALY '11b].

Pareillement aux cyclotides, la PAFP-S (Protéine antifongique de *Phytolacca*) est une knottine antifongique très basique de 38 acides aminés isolée à partir des graines de *Phytolacca americana* [SHAO '99]. PAFP-S présente un large spectre d'activité antifongique (Tableau 2).

Tableau 2 : Activité antimicrobienne de quelques knottines.

Kalata, circuline A et B et cyclopsychotride, ont été isolés à partir des plantes de café [TAM '99]

Source	Knottine	Inhibition de la croissance de		
		Gram +	Gram -	Champignon
Plante	PAFPS			+
Plante	Kalata B1 (cyclotide)	+		
Plante	Circuline A (cyclotide)	+		
Plante	Circuline B (cyclotide)	+	+	
Plante	Cyclopsychotride (cyclotide)	+	+	+

d. Activité antihelminthique

Plusieurs cyclotides (Kalata et Cycloviolacin) ont présenté une activité significative dans l'inhibition du développement des larves et la mobilité des adultes d'*Haemonchus contortus* et des larves de *Trichostron-*

gylus colubriiformis, deux importants nématodes gastro-intestinaux parasites des moutons. L'activité ne présente aucune sélectivité chirale et semble être liée à des résidus de base regroupés suggérant que l'interaction avec la membrane plutôt qu'avec un récepteur est impliquée dans l'activité antihelminthique des cyclotides [COLGRAVE '08b]. Les cyclotides sont aussi actifs contre des nématodes parasites du mouton [COLGRAVE '08a] et de l'homme [COLGRAVE '09].

e. Activité molluscicide

Les extraits purs et bruts des cyclotides d'*O. affinis* et de *Viola odorata* ont montré des activités molluscicides contre le plus important ravageur des productions de riz dans les zones humides, l'escargot doré (Golden apple snails GAS), *Pomacea canaliculata*. La LC_{50} déterminée au bout de 24 h pour Kalata B1, B2 et le métaldéhyde, un molluscicide synthétique connu, est de 90, 54 et 105 μ M, respectivement. À 65 μ M, les cyclotides pures de *V. odorata*, cycloviolacin O1 et Kalata S, causent la mortalité de 100% de GAS, alors que Kalata B1, B2 et le métaldéhyde causent 78, 68 et 60% de mortalité, respectivement. Kalata B7 était faiblement active, tandis que Kalata B8 ne présente aucune activité molluscicide [PLAN '08]. L'activité de ces cyclotides semble être liée à un dommage au niveau des mucocytes de l'appareil digestif menant d'abord à une sécrétion excessive de mucus suivie par des changements au niveau du métabolisme énergétique.

f. Activité insecticide

En réponse à l'attaque des insectes, les plantes accumulent un ensemble de protéines de défense y compris des knottines ayant une activité nocive pour les insectes. Ces mécanismes de défense inhérents confèrent un certain degré de résistance naturelle à la plante. Les cyclotides étant des molécules synthétisées par des plantes, on s'attend à dégager des activités insecticides contre les organismes dangereux pour les végétaux. Les cyclotides sont ainsi actifs contre des larves de lépidoptères [JENNINGS '01]. Des études par microscopie optique et électronique ont montré que les pep-

tides provoquaient la rupture des membranes cellulaires de l'intestin des larves par un mécanisme similaire à celui responsable de l'activité antibactérienne [BARBETA '08]

Parmi ces knottines nous trouvons aussi des inhibiteurs de protéases [ABDEEN '05] ainsi que des albumines de type 1 végétales [DELOBEL '98]. Ces knottines, actives par ingestion, sont capables de perturber les cellules épithéliales de l'intestin moyen des lépidoptères et des coléoptères et par conséquent de nuire à l'absorption des nutriments [ABDEEN '05, XINGWEI HOU '06].

5. Bioingénierie et Production hétérologue des knottines

La tolérance de certaines positions de la séquence des knottines aux substitutions a suggéré une application en bioingénierie pour le « design » de peptides bioactifs stabilisés par le motif structural « cystine-knot ». Ainsi, le squelette peptidique des knottines a pu être utilisé afin de synthétiser des peptides possédant une activité biologique différente de celle des knottines. Le but de ces études est de stabiliser un motif bioactif en utilisant la matrice structurale des knottines très résistantes aux dégradations diverses [COLGRAVE '04].

En effet, le motif ICK ou CCK, semble être une excellente matrice compacte et globulaire pour la présentation d'une variété de groupes fonctionnels, et dans la nature il a été utilisé pour produire une gamme de polypeptides avec des cibles biologiques diverses. Par exemple, MCoTI-II est le cyclotide inhibiteur des trypsines le plus exprimé et a été isolé à partir des graines de *Momordica cochinchinensis*, un membre de la famille végétale des cucurbitacées. Greenwood et ses collaborateurs ont montré que des cellules de cancer de seins incubées avec un MCoTI-II marqué internalisent les cyclotides par le biais d'endocytose dans les 15 minutes d'exposition. Cela fournit la première preuve qu'au moins une sous-famille de cyclotides est capable de pénétrer des cellules par les voies d'endocytose. Une fois englobés dans des vésicules au sein du cytoplasme par macropinocytose, MCoTI-II marqué n'a pas été transféré vers les lysosomes pour la dégradation. La pénétration cellulaire de MCoTI-II ouvre la possibilité pour le greffage d'épitopes peptidiques bioactifs sur la structure de MCoTI-II afin d'être livré à des cibles intracellulaires pour des applications pharmaceutiques [GREENWOOD '07]. De plus, une équipe a réussi à greffer sur le squelette de la kalata B1 la région linéaire d'une protéine possédant une ac-

tivité antagoniste au facteur de croissance des vaisseaux sanguins endothéliaux A (VGEF-A), ouvrant la voie à une nouvelle branche de synthèse de molécules bioactives [GUNASEKERA '08]. La Figure 8 montre qu'outre les six cystéines du squelette peptidique des cyclotides, les boucles 1 et 4 sont hautement conservées par leur nombre d'acides aminés ainsi que par la nature des résidus pour la boucle 1, ce qui laisse penser qu'elles sont nécessaires à l'acquisition de la structure CCK. De plus, certaines positions dans la séquence peptidique des cyclotides semblent être plus tolérantes que d'autres à la nature des acides aminés. Ceci laisse supposer que ces peptides peuvent être utilisés dans des expériences de chimérisation de peptides ou de modifications séquentielles afin de leur faire acquérir une autre activité biologique.

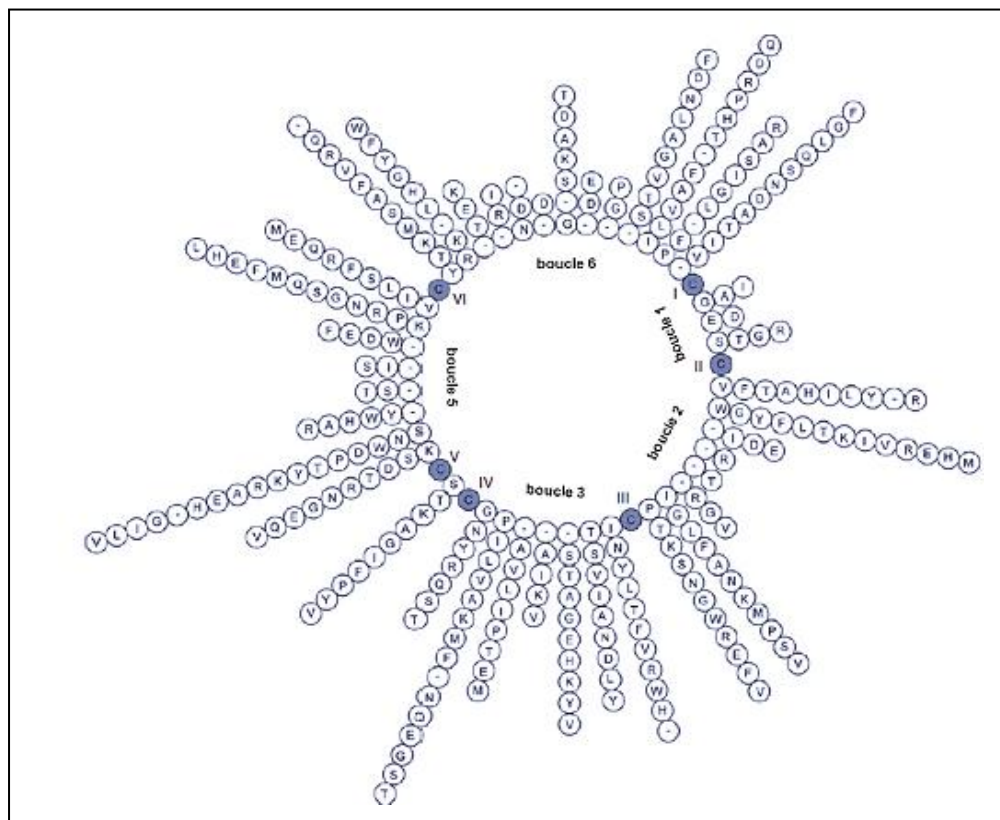


Figure 8 : Diagramme représentant la diversité dans la séquence des cyclotides [N. L. DALY '11a]

En résumé, la matrice des knottines semble être une charpente inédite pour la création de petites protéines repliées avec des activités de liaison. Cela est très prometteur pour la conception de petites protéines avec des fonctions de liaison complètement nouvelles. Depuis que ces knottines sont considérées comme des structures intéressantes dans la conception des médicaments, il a été nécessaire que ces protéines soient obtenues facilement. En tant que petites molécules/protéines, différentes knottines

ont été obtenues soit par une extraction directe à partir des ressources naturelles, soit par synthèse chimique ou soit par une expression dans des systèmes de recombinaison hétérologues [NORTON '98].

À partir d'une ressource naturelle végétale, les knottines ne peuvent être obtenues qu'en faibles quantités limitant leur disponibilité pour des applications que ce soit dans le domaine de la recherche fondamentale ou appliquée (industrie phytosanitaire, pharmaceutique, médicale...). Pour obtenir de plus grandes quantités de ces peptides, deux approches sont possibles: synthèse chimique et production recombinante dans des systèmes d'expression hétérologues (Figure 9). En raison des modifications post-traductionnelles de nombreuses knottines décrites ci-dessus [CRAIG '99, BUCZEK '05], la synthèse chimique par synthèse peptidique en phase solide (SPPS) sur un support de résine [MERRIFIELD '63] a été la méthode de choix pour produire des knottines en grandes quantités. Une ligature chimique des fragments de peptides purifiés peut être utilisée pour produire des plus longs polypeptides contenant jusqu'à plusieurs centaines d'acides aminés [DAWSON '00]. En 1989, Le-Nugyen a montré la possibilité de produire un inhibiteur de la trypsine isolé à partir d'*Ecballium elaterium* de la famille des cucurbitacées par la méthode SPPS [DUNG LE-NGUYEN '89].

Le coût des produits chimiques ainsi que des acides aminés protégés destinés à synthétiser ces peptides est assez élevé. Une approche alternative et rentable produisant des quantités plus importantes est l'utilisation de systèmes d'expression bactériens et eucaryotes. Le facteur limitant de ces systèmes de production est le mécanisme de modification post-traductionnelle qui ne permet pas la production de ces peptides avec les mêmes modifications post-traductionnelles trouvées dans de nombreuses petites knottines [D. J. CRAIK '99, BUCZEK '05].

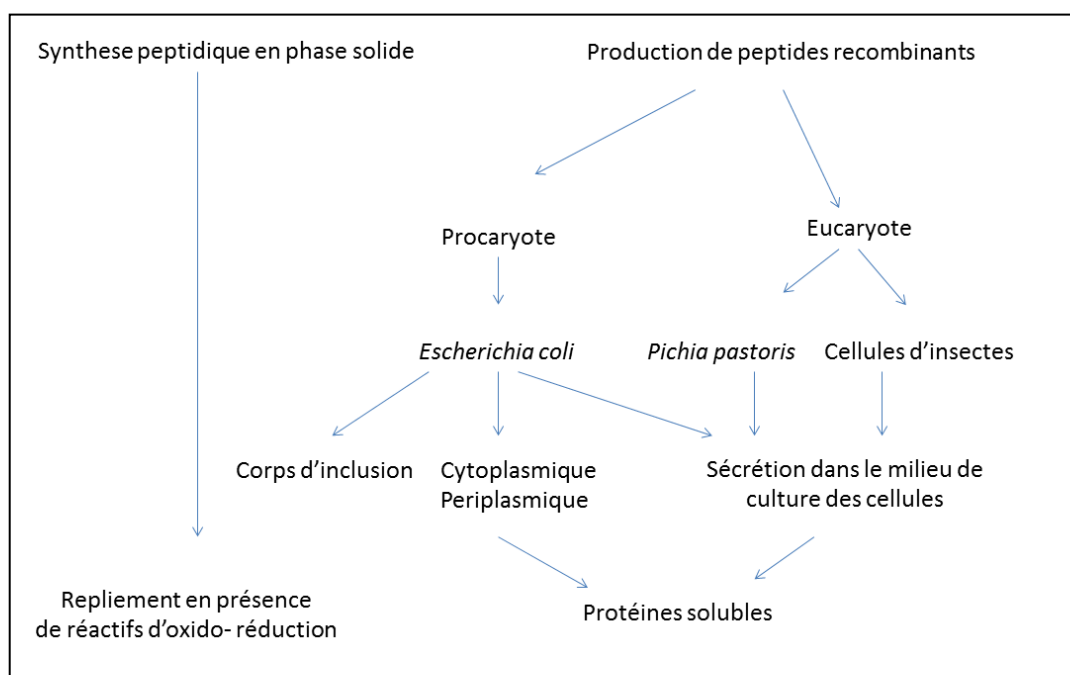


Figure 9 : Méthodes de production biotechnologique de peptides à ponts disulfure.

La synthèse peptidique en phase solide nécessite un repliement ultérieur du peptide, en présence d'oxygène. L'addition de thiols à faible poids moléculaire à l'état oxydé et réduit dans la solution peut augmenter le rendement du peptide correctement replié. La production de peptide recombinant peut entraîner des protéines insolubles (corps d'inclusion) ou des protéines solubles [BECKER '08]

L'Escherichia coli est l'hôte bactérien le mieux caractérisé et le plus utilisé pour la production de protéines recombinantes [BANEYX '99, PINES '99]. La production est rapide car le temps de doublement de la bactérie est de l'ordre d'une demi-heure. Il est peu coûteux du fait de la simplicité du milieu de culture. En général, la surexpression des petits peptides directement dans *E. coli* est difficile, car ils sont soit rapidement dégradés par les protéases cellulaires soit ils s'accumulent pour former des agrégats insolubles, formant ainsi des corps d'inclusion [GEORGIU '96]. En outre, le cytoplasme d'*E. coli* est un environnement réducteur qui empêche la formation de ponts disulfure [PRINZ '97], ce qui accroît la tendance à former des corps d'inclusion sur la surexpression. Pour atteindre un niveau élevé d'expression de ces petits peptides, les exprimer en fusion avec des protéines bien exprimées a été l'approche la plus réussie à ce jour. L'utilisation de ces protéines porteuses fortement solubles comme, par exemple, la protéine liant le maltose [KAPUST '99], la thiorédoxine [LAVALLIE '93], ou le glutathion-S-transférase [NYGREN '94] peut même conduire à l'ex-

pression des protéines de fusion solubles. Grâce à une chromatographie d'affinité ciblant ces protéines porteuses, ces protéines de fusion solubles peuvent être facilement purifiées [NYGREN '94, ESPOSITO '06, PETI '07]. L'utilisation additionnelle de souches bactériennes déficientes en protéase [MAURIZI '92, GOTTESMAN '96] peut aider à éviter la dégradation protéolytique non spécifique. Ensuite, le peptide doit être libéré par clivage chimique ou protéolytique dans un site de clivage créé dans la séquence protéique entre la protéine porteuse et la séquence du peptide d'intérêt [ESPOSITO '06]. Lorsque l'expression en fusion avec des protéines solubles ne peut être réalisée, la protéine doit être solubilisée à partir des corps d'inclusion en utilisant des réactifs dénaturants tels que l'urée et le chlorhydrate de guanidine, puis repliée par dilution dans ou par dialyse contre un tampon ne contenant pas d'agent dénaturant [FAHNERT '04].

En résumé, la production de protéines à ponts disulfure dans *E. coli* nécessite souvent un repliement de ces protéines à partir des corps d'inclusion ou la sécrétion dans l'espace périplasmique. Ces deux approches ont tendance à diminuer le rendement final.

Une alternative à la production dans *E. coli* est la production de ces protéines dans des systèmes d'expression eucaryotes. La levure méthylotrophe, *Pichia pastoris*, fournit un environnement de repliement intracellulaire similaire aux cellules de mammifères et la sécrétion dans le milieu environnant peut être facilement accomplie [WHITE '94]. En utilisant ce système d'expression, un certain nombre de protéines riches en ponts disulfures ont été produites (deux défensines végétales : AX2 et Psd1 extraite à partir des betteraves sucrières et du pois, respectivement) [KRISTENSEN '99, CABRAL '03]. Le niveau de sécrétion par cellule est relativement faible dans *P. pastoris*. Par contre, la fermentation à haute densité est à l'origine d'une production en grandes quantités de protéines sécrétées. Comme *E. coli*, *P. pastoris* ne nécessite aucun équipement spécial pour sa manipulation et les milieux de culture ne sont pas plus chers que ceux pour *E. coli*. Ainsi, la production de protéines à ponts disulfure dans *P. pastoris* est devenue une alternative intéressante à la production dans *E. coli* [WHITE '94]. Les knottines ont ainsi été produites avec succès dans le système levure (<http://knottin.cbs.cnrs.fr>).

Les cellules d'insectes ont également été utilisées pour la production de protéines à ponts disulfure via le système baculovirus [VOGEL '04, GALESI '07]. Une knottine neurotoxique, HWTX-I, isolée à partir du venin d'araignée chinoise (*Ornithoctonus huwena*), a été produite par le système baculovirus tout en conservant une activité similaire à celle de la molécule

native [Ji '05]. Les cellules S2 provenant de *Drosophila melanogaster* et transfectées d'une façon stable [GALESI '07] sont un autre système d'expression de cellules d'insectes. Une toxine d'araignée, PcTx1, a été exprimée avec succès dans les cellules S2. Cette knottine, exprimée pour la première fois dans ce système recombinant, conserve les mêmes propriétés structurales et fonctionnelles du peptide natif [ESCOUBAS '03]. Ces méthodologies ont permis la production de peptides bien conformés ayant retenu leur activité biologique. Un frein à leur développement est les coûts élevés des milieux de culture des cellules d'insectes, cependant elles restent intéressantes à explorer.

6. Conclusion

Les structures peptidiques compactes se retrouvent dans la grande diversité de l'arbre du vivant. On les retrouve ainsi produites par les procaryotes (bactéries) et par les eucaryotes (cyclotides des plantes, défensines circulaires des mammifères, conotoxines des mollusques). Leurs activités biologiques variées et généralement cruciales pour l'organisme producteur sont alors renforcées par l'extrême résistance de ces structures aux dégradations diverses, notamment chimiques et protéolytiques.

Un de ces peptides knottines, PA1b (Pea Albumin 1 subunit b), est un peptide entomotoxique actif contre des insectes de différents ordres, notamment les charançons des céréales et la teigne de la farine. Il a été identifié dans la graine de pois. Outre sa forte toxicité pour ces insectes, cette protéine de la famille des albumines contient 16 % de cystéines et représente la principale réserve en acides aminés soufrés des graines de pois. Au cours de ces dix dernières années, son étude a permis la connaissance à la fois de sa biosynthèse et de son mécanisme d'action. La présence d'un tel composé toxique a été initialement suspectée par les travaux de Delobel *et al* [DELOBEL '98]. La purification de la molécule a conduit à l'identification d'un peptide dont la structure primaire et la structure du gène avaient été élucidées, en 1986, par Higgins *et al.*, mais sans aucune démonstration de ses effets toxiques [HIGGINS '86]. PA1b est l'un des rares peptides entomotoxiques actifs par voie orale actuellement connus. La deuxième partie de cette introduction sera dédiée à PA1b dont l'activité entomotoxique a été brevetée en 1998 par l'INSA de Lyon et l'INRA et qui depuis est l'objet principal des recherches effectuées dans l'équipe « Entomotoxines ».

II. Deuxième partie

PA1b, la première knottine à activité insecticide

1. Intérêt potentiel de PA1b pour la lutte contre les ravageurs

Les insectes sont à l'origine de la plupart des dommages subis dans les réserves de denrées stockées, bien que dans certains pays les rongeurs se montrent encore plus dangereux. Les insectes nuisibles qui sévissent dans les entrepôts de stockage ont un taux de reproduction élevé et se développent rapidement, ce qui les met en mesure de provoquer à court terme de très graves dégâts à partir d'une population originelle modeste [DAVID REES '04]. Les **coléoptères (Coleoptera)** constituent de loin le groupe le plus important au sein des insectes ravageurs des stocks. Ils sont suivis par les **teignes (Lepidoptera)**. On trouve ensuite une série constituée d'autres groupes, tels que les **psocques (Psocoptera)** qui, si elles sont pratiquement incapables de provoquer des dommages au niveau des stocks, peuvent néanmoins poser des problèmes d'hygiène. Parmi les espèces d'insectes présents, la plus souvent trouvée dans les lots de stockage est le charançon du riz *Sitophilus oryzae*. Il est déniché dans 23 % des échantillons de blés tamisés en 2011 contre 24 % de ceux de 2010 (Tableau 3). Une enquête ancienne menée en 1977 et 1978 a montré que l'espèce était déjà présente dans 18 puis 37 % des échantillons. Pour sa part le charançon du blé *Sitophilus granarius* est présent cinq fois moins souvent que *S. oryzae* en 2011 comme en 2010 (respectivement 4 % et 5 % des échantillons) [DECOIN '12].

D'un point de vue économique, et sur le plan de la protection des ressources naturelles, il est beaucoup plus raisonnable de protéger les récoltes contre les pertes que d'investir pour une augmentation continue de la production agricole. Ces pertes causées par les insectes ravageurs des denrées stockées posent un problème non seulement européen [ADLER '13] mais aussi mondial [GUÈYE '11, TOGOLA '13] (<https://colloque.inra.fr/iobc-ips-2013conference>). Les céréales et leurs sous-produits sont les plus touchés par ces ravageurs : gâteaux, biscuits, pâtes, pain, crackers, riz, farine, ravioli et bien autres produits. De plus, la protection de ces denrées stockées a un enjeu économique : un lot de grains victime d'une infestation n'est plus viable, car refusé systématiquement par

les groupes agroalimentaires [SAVOLDELLI '13]. Si aucune protection n'est faite, après sept mois de stockage, la perte des denrées peut être totale [AKOU-EDI '84, AGOUNKÉ '94, BELL '94]. Les méthodes utilisées pour limiter les pertes dans les stocks sont généralement les insecticides chimiques qui peuvent induire une intoxication chronique des consommateurs, une résistance chez les ravageurs et avoir un impact négatif sur l'environnement (Attia, 1984; Marion-Poll, 2002). Parallèlement, le contexte législatif étant plus strict, les dernières directives européennes imposent une sévère réduction de l'usage de pesticides et d'insecticides après la récolte. En effet, des résidus risqueraient d'imprégner les fibres des céréales et de se retrouver finalement dans l'assiette du consommateur [ERVEL '07].

Tableau 3 : Fréquence totale de détection des espèces d'insectes au sein des 190 lots de grains extraits de silos de blés dans le cadre de l'enquête Ecoprotect grain aux printemps 2010 puis 2011 [DECOIN '12]

Espèce	Fréquence 2010 (%)	Fréquence 2011 (%)	Commentaires
Charançon du riz <i>Sitophilus oryzae</i>	24	23	Fréquence proche de celles détectées en 1977 et 1978
Tribolium roux <i>Tribolium castaneum</i>	13	11	Espèces de ravageurs « secondaires» (= sans « formes cachées » internes au grain), pourtant trouvées dans certains des 2 ^{es} tamisages
Petit silvain plat <i>Cryptolestes ferrugineus</i>	11	14	
Silvain <i>Oryzaephilus surinamensis</i>	8	15	
Capucin des grains <i>Rhyzoperta dominica</i>	7	8	Recrudescence (non décelé en 1977-78)
Charançon des grains <i>Sitophilus granarius</i>	5	4	Régression (trouvé dans env. 20 % des lots en 1977-78)
Mycetophagidae	4	5	Mangent davantage les moisissures que les grains
Tribolium sombre <i>Tribolium confusum</i>	2	1	Espèce trouvée seulement au 1 ^{er} tamisage
Parasites et prédateurs	5	4	Pas ravageurs... mais pas acceptés par les acheteurs
Toutes espèces	27	40	Beaucoup d'échantillons avec plusieurs espèces
N.B. Les trois espèces citées en gras sont des ravageurs primaires			

Par conséquent, trouver une alternative pour contrôler efficacement ce ravageur majeur du blé est devenu impératif. Plusieurs travaux ont rapporté l'effet insecticide de la farine de pois sur *Sitophilus oryzae* [DELOBEL '98, X. HOU '03, PRETHEEP-KUMAR '04]. Cette toxicité est principalement corrélée avec l'activité biologique d'une albumine majeure isolée à partir de grains de pois, PA1b (Pea Albumin 1 subunit b) [DELOBEL '98]. La forte activité entomotoxique de PA1b est aussi couplée à une forte stabilité de cette protéine. Elle est capable de conserver son activité biologique pendant plusieurs années dans des graines sèches [J. PETIT '05]. La forte toxicité de la protéine PA1b envers les charançons des grains, sa stabilité lors du stockage et de la dessiccation des graines, associées au fait de la coexistence de plusieurs isoformes de la protéine PA1b chez le pois comme chez d'autres fabacées [S LOUIS '07], font de ce peptide un candidat idéal considéré comme une nouvelle classe d'entomotoxines végétales.

2. Structure et propriétés de PA1b

Le peptide PA1b est le produit de l'expression du gène *PA1*. Ce gène est formé par deux exons séparés par un intron localisé dans la séquence qui code pour le peptide signal. Le gène *PA1* est transcrit en un seul ARNm codant pour une préproprotéine PA1 de 13,9 kDa (130 acides aminés). La préproprotéine est composée d'un peptide signal, de la protéine PA1b, de son propeptide, de la protéine PA1a et de son propeptide. La séquence signal de cette protéine est d'abord clivée par des endoprotéases, pour donner une proprotéine. Cette dernière comprend deux peptides, PA1b et PA1a avec leurs propeptides. Après clivages des propeptides, les deux peptides matures PA1b (3.8 kDa, 37 aa) et PA1a (6 kDa, 53 aa) sont libérés [EALING '94] (Figure 10).

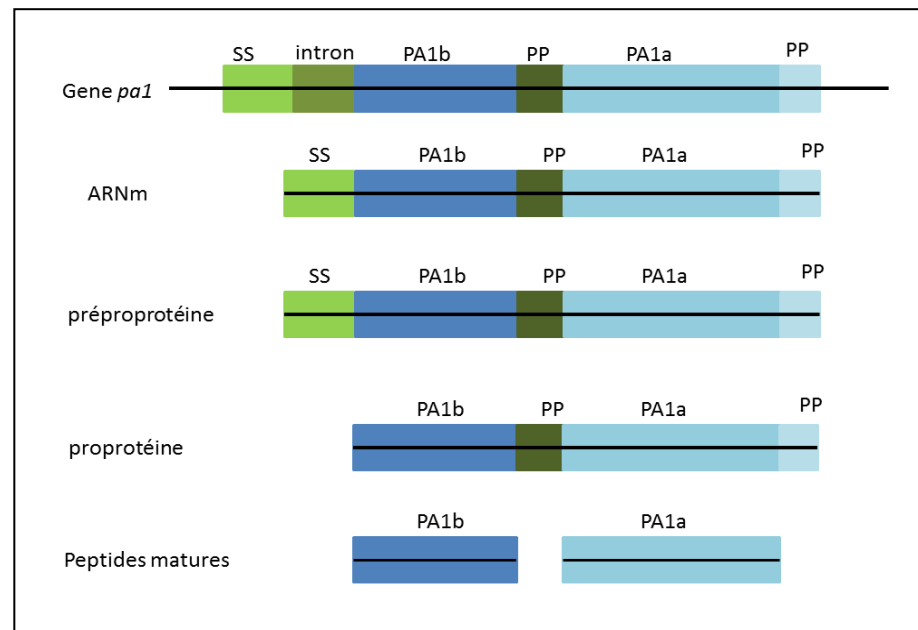


Figure 10 : Structure du gène *pa1* et maturation des peptides PA1b et PA1a.

SS=Séquence Signal ; PP=Propeptides

Le gène *PA1* est représenté en de nombreux exemplaires chez le pois (au moins sept gènes connus). Les travaux de thèse de Sandrine Louis au laboratoire Biologie Fonctionnelle Insectes et Interaction (BF2I) ont permis de déterminer plusieurs autres gènes *PA1* chez d'autres légumineuses comme le haricot (*Phaseolus vulgaris*), ou le soja (*Glycine max*) [SANDRINE LOUIS '04b]. Actuellement, l'existence de PA1b est attestée chez de nombreuses Légumineuses par des preuves biologiques, biochimiques ou de biologie moléculaire, et est principalement présente dans cette famille de plante. Cependant, il est à noter que PA1b a été trouvé dans une plante parasite *Phelipanche aegyptiaca* de la famille des Orobanchacées suite à un phénomène de transfert de gène horizontal [Y. ZHANG '13]. PA1b est un peptide de 37 acides aminés comportant six cysteines impliquées dans trois ponts disulfure intramoléculaires, dont la séquence est donnée dans la Figure 11.

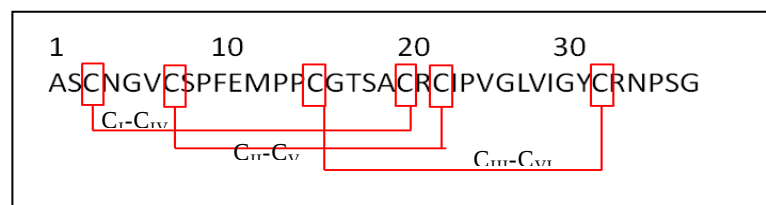


Figure 11 : Séquence primaire de PA1b.

Les résidus cysteines sont surlignés en rouge et leur appariement est indiqué en dessous.

L'alignement des séquences de PA1 (obtenues le plus souvent par la biologie moléculaire) chez différentes Légumineuses (Figure 12) montre une conservation assez importante de la préproprotéine, en particulier dans le peptide PA1a.

Pour le peptide PA1b, les acides aminés de structure sont très fortement conservés, c'est-à-dire principalement les six cysteines et les cinq proline. Deux boucles sont également fortement conservées : une première boucle ou la séquence « CSPFE » est retrouvée dans quasiment toutes les séquences (acides aminés 7-11), et la zone des acides aminés 23-29 où la conservation n'est pas stricte, cette boucle conserve toujours un caractère très fortement hydrophobe. Enfin le motif CRC (acides aminés 20-22), avec l'arginine encadrée par les deux cysteines est également très conservé.

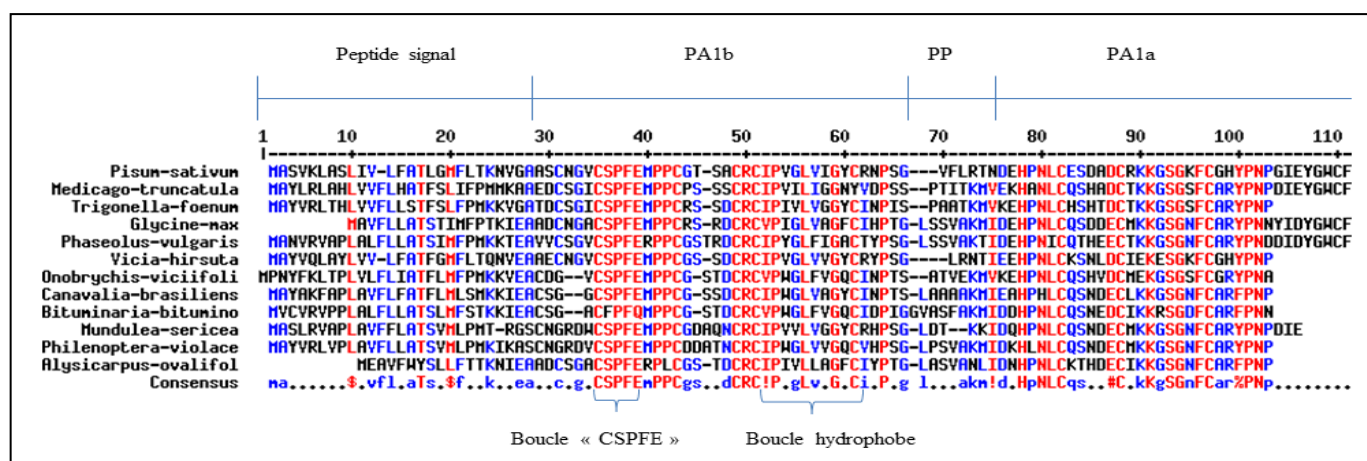


Figure 12 : Alignement protéique de la préproprotéine PA1 de différentes légumineuses.

La fin du peptide PA1a ainsi que le propeptide terminal ne sont pas représentés par manque de données dans cette zone. Les alignements permettent de caractériser deux boucles, l'une conservée nommée « CSPFE », et l'autre dont le caractère hydrophobe est conservé.

La structure tridimensionnelle du peptide, ainsi que l'appariement des ponts disulfures ont ensuite été établis par RMN et modélisation moléculaire au Centre de Biophysique Moléculaire d'Orléans [JOUVENSALE '03]. PA1b est un peptide compact, qui comporte trois feuillets antiparallèle et un appariement des cysteines en 1-4, 2-5 et 3-6. Cet appariement, ainsi que le troisième pont disulfure qui passe dans les deux premiers en une sorte de nœud, marque l'appartenance de PA1b à la famille des *cystine-knot inhibitors* (ICK) ou knottines (Figure 13). Ce repliement confère à la structure une grande compacité et une très grande stabilité : PA1b conserve son activité après autoclavage, digestion par des protéases (PA1b est insensible à

une digestion par la trypsine, la chymotrypsine [HANCOCK '94], la papaine, la protéinase K mais il s'est avéré que ce peptide n'est sensible qu'à une seule protéase complexe de source microbienne, la pronase E [GRESSENT '03]) ou extraction dans un grand nombre de solvants organiques. Seule l'ouverture des ponts disulfure par un chauffage à 100°C en présence de DTT aboutit à la perte totale de l'activité. Si PA1b ne présentait pas d'homologie au niveau de sa séquence, en revanche au niveau de sa structure tertiaire il présente de fortes similarités structurales avec une toxine d'araignée, les deux protéines ayant pour cible des insectes (peptide ACTX-Hi : OB4219) [JOUVENAL '03].

Enfin, concernant PA1a, aucune fonction n'a pas pu être associée à ce peptide jusqu'à présent.

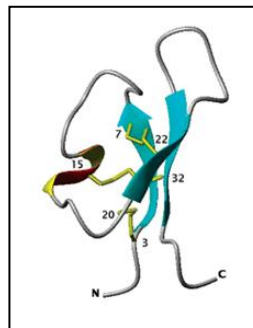


Figure 13 : Structure tridimensionnelle des PA1b [JOUVENAL '03]

3. Spectre d'hôte de PA1b

En tant que peptide entomotoxique naturel actif *per os*, PA1b présente, dans son potentiel d'utilisation en tant que biopesticide, *a priori* de nombreux avantages par rapport aux pesticides chimiques pour la préservation de la qualité des sols et des eaux après traitement, et, en général, pour le maintien des biotopes environnants.

Au sein de la famille des insectes, les travaux de Frédéric Gressent ont permis de dégager la sensibilité relative de certains Coléoptères à la toxine (*Sitophilus sp*, *Harmonia axyridis*), avec cependant quelques exceptions comme le *Tribolium castaneum* qui reste insensible à des doses de PA1b 20 fois supérieures à la dose habituellement utilisée pour tuer 100% des charançons [GRESSENT '07]. Si tous les Diptères ne sont pas sensibles à la toxine (*D. melanogaster*), il est cependant intéressant de constater que les moustiques, au travers des tests effectués sur des larves de *Culex pi-*

piens et d'*Aedes aegyptii* présentent une sensibilité aigue (100% de mortalité en deux jours) à des doses relativement faibles de PA1b (200 $\mu\text{g}.\text{ml}^{-1}$). Les Hémiptères, et plus précisément les pucerons, sont plus ou moins sensibles en fonction de l'espèce ; l'une (*Acyrtosiphon pisum*) est sensible, tandis que les deux autres espèces, *Aphis gossypii* et *Myzus persicae*, sont insensibles. Enfin, les Lépidoptères semblent les moins sensibles à la toxine ; *Mamestra brassicae*, *Spodoptera frugiperda*, *Ostrinia nubilalis* sont insensibles à la toxine, même si *Ephestia khuniella* et les cellules Sf9 (originaires de *Spodoptera frugiperda*), sensibles à la toxine, demeurent une exception

Parmi les insectes dont l'infestation est potentiellement maitrizable par l'utilisation de PA1b, certains ont un impact économique ou sanitaire très important. Par exemple, les charançons des céréales sont responsables de pertes céréalieres approchant les 20% dans le monde (et jusqu'à 40% dans les pays tropicaux) [NGAMO '07, DECOIN '12], les moustiques sont les premiers vecteurs de maladies humaines et des mammifères dans le monde [GUILLAUMOT '05] et les pucerons sont des vecteurs de virus des plantes [LECOQ '96].

4. Mécanisme d'action de PA1b chez les insectes

Des études électrophysiologiques récentes menées sur des cellules Sf9 ont montré que la liaison de la toxine sur la membrane cellulaire entraîne une dépolarisation de la membrane. Une étude pharmacologique a démontré que PA1b était capable de reproduire l'effet de la bafilomycine, une toxine bactérienne connue pour être un inhibiteur puissant d'une pompe à proton membranaire, la V-ATPase (pour Vacuolar ATPase) [CHOUABE '11]. Ces travaux ont permis de suggérer que PA1b agissait en bloquant la V-ATPase membranaire des cellules Sf9, qui est donc son récepteur. La V-ATPase est une pompe à protons utilisant l'ATP, caractérisée d'abord dans les membranes vacuolaires, mais présente aussi au niveau des membranes plasmalemmiques (Figure 14). La V-ATPase est un complexe protéique multimérique, hautement conservé parmi les espèces vivantes, des bactéries à l'homme. Chez les insectes, la protéine de *Manduca sexta* (Lépidoptères) a été étudiée de près en ce qui concerne sa structure, sa fonction et sa régulation. Il s'agit d'une protéine essentielle pour les insectes, en particulier dans l'intestin, puisque son action fournit l'énergie nécessaire à l'absorption des nutriments. La V-ATPase de l'insecte est composé de 14 sous-unités organisées en deux complexes. Le complexe V0 de la membrane a quatre

sous-unités, et le complexe V1 est cytosolique et supporte la fonction d'ATPase. Les tests biochimiques réalisés sur V-ATPase purifiée de *M. sexta* ont montré que PA1b inhibe l'enzyme mais sans aucune interférence avec l'unité V1. Par conséquent, les auteurs ont conclu que PA1b agit en se liant et en inhibant le complexe V0 de la V-ATPase chez les insectes [GRESSENT '11].

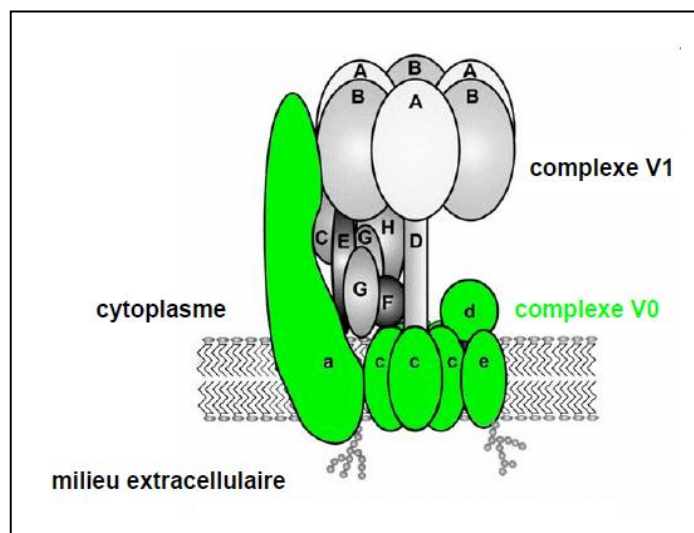


Figure 14 : Schéma hypothétique de la V-ATPase de *Manduca sexta*.

Le complexe V1 (gris) porte l'activité ATPasique. Le complexe V0 (vert) comprend quatre types de sous-unités, dont les sous-unités a et c qui forment le canal proton.

Chapitre 2

Medicago truncatula, espèce modèle de la famille des Albumines 1?

1. Introduction

***Medicago truncatula* – Plante modèle des légumineuses**

a. Les légumineuses

Les légumineuses (Fabacées) sont une grande et diverse famille qui comprend des plantes herbacées annuelles jusqu'à des arbres pérennes. Ces plantes sont des composantes essentielles dans l'écosystème terrestre dû à leur capacité à fixer l'azote atmosphérique dans les nodules symbiotiques, et elles sont donc d'excellents colonisateurs des environnements pauvres en azote.

Depuis le début de la civilisation, les graines et gousses (fruits spécifiques des légumineuses) ont joué un rôle très important dans l'alimentation humaine, car elles constituent une source exceptionnelle de protéines (légumineuses protéagineuses : le pois, la féverole, le haricot, le pois chiche, les lentilles) et de lipides (légumineuses oléo-protéagineuses : le soja et l'arachide). Les légumineuses sont aussi une source importante de fourrage pour les animaux (luzerne, trèfle).

Les légumineuses avec plus de 18,000 espèces sont la troisième plus grande famille de plantes supérieures et la deuxième plus importante pour les pâturages d'intérêt agricoles, après les poacées (autrefois appelées graminées) qui incluent la canne à sucre et les céréales tels que le maïs, riz, blé, orge, avoine, seigle et millet [FARIA '89, YOUNG '03].

La capacité de différencier, sur leurs racines, des nodules hébergeant des bactéries symbiotiques capables de fixer l'azote a été démontrée pour 88% des espèces examinées jusqu'à ce jour. Le partenaire bactérien de cette symbiose, appartenant à la famille des Rhizobiacées réduit l'azote à l'intérieur des nodules grâce à la nitrogénase. L'activité de cette enzyme est inhibée par l'oxygène d'où la présence dans les nodules fonctionnels de la leghémoglobine régulant les taux d'oxygène libre. L'intérêt agronomique des légumineuses à son origine dans cette fixation symbiotique de l'azote qui leur permet de produire en abondance des protéines en l'absence de fertilisation azotée.

Dans une politique désireuse de promouvoir l'agriculture durable et de réduire les nitrates, la culture de ces plantes trouve un regain justifié.

b. *Medicago truncatula*

Parmi les légumineuses *Medicago truncatula*, proche de la luzerne cultivée (*Medicago sativa*) présente un intérêt particulier. En plus de sa parenté avec la luzerne, qui est la légumineuse fourragère économiquement la plus importante au monde, *M. truncatula* a un intérêt agronomique propre, puisqu'elle est utilisée comme plante fourragère dans un grand nombre de régions dans le monde en incluant la zone méditerranéenne, l'Afrique du Sud et l'Australie [CRAWFORD '89]. Il y a 20 ans des chercheurs ont commencé à proposer *M. truncatula* comme plante modèle des légumineuses [BARKER '90]. *M. truncatula* est une espèce sauvage annuelle du pourtour méditerranéen également cultivée comme légumineuse fourragère.

Des populations et lignées collectionnées autour de la Méditerranée représentent un réservoir de ressources génétiques d'adaptation à des conditions environnementales variées. Des caractéristiques comme son petit génome (~500 Mbp - [BENNETT '97]), la diploïdie ($2n = 16$), l'autogamie et un temps de régénération court (environ trois mois dans des conditions les plus favorables) ont contribué à en faire la plante modèle des légumineuses, pour laquelle les ressources ont été développées plus vite et à plus grande échelle (<http://www.medicago.org/genome>) que pour l'autre plante modèle, le lotier (*Lotus japonicus*). Le séquençage complet de son génome est publié en 2011 [YOUNG '11]. Jusqu'aux années 1990s, il n'existait que deux plantes modèles. La première plante modèle *Arabidopsis thaliana* (L.) de la famille des crucifères a permis des progrès considérables dans la connaissance des bases moléculaires de la biologie des plantes. Le séquençage de son génome [ARABIDOPSIS-GENOME-INITIATIVE '00], a permis d'avoir des outils de génétique et de génomique disponibles pour accélérer la recherche. *Arabidopsis thaliana* a été choisie grâce à sa petite taille, son petit génome nucléaire, son court temps de génération et sa grande production de graines. De plus elle pouvait être transformée facilement. La deuxième plante modèle fut le riz (*Oryza sativa*) pour les monocotylédones (céréales) qui sont le groupe des plantes cultivées économiquement le plus important.

Cependant, ces deux espèces ne sont pas suffisantes pour représenter toute la diversité biologique du monde végétale [ADAM '00]. Il fallait surtout trouver une plante représentative d'un des taxons végétaux les plus importants : les légumineuses. Ni *Arabidopsis* ni le riz ne font des symbioses fixant l'azote atmosphérique. De plus, contrairement à la majorité des espèces végétales, *Arabidopsis* est incapable d'établir une symbiose mycorhizienne.

M. truncatula comme plante modèle permet d'apporter un outil de recherche pour mieux comprendre la biologie des légumineuses et en particulier la génétique moléculaire de deux associations plantes-microorganismes importantes : la symbiose *Rhizobium* légumineuses [D. COOK '95] et la symbiose avec des mycorhizes arbusculaires vésiculaires [MARIA J. HARRISON '94]. Elle permet également de comparer les mécanismes qui lui permettent de former les relations symbiotiques avec ceux qui la défendent contre des agents pathogènes.

La luzerne, le pois et le soja figurent parmi les légumineuses les mieux caractérisées génétiquement avec de nombreux marqueurs génétiques, des cartes génétiques bien développées et d'autres d'outils génomiques. Cependant, ces espèces ne pouvaient pas être utilisées pour faire des études de génétique moléculaire à cause de leurs caractéristiques défavorables comme un génome de grande taille, une ploïdie complexe, de l'allogamie et des difficultés à être transformées. Pour cette raison *M. truncatula* et *Lotus japonicus* ont été proposées comme plantes modèles et elles sont actuellement utilisées par de nombreux laboratoires dans le monde [BARKER '90, D. R. COOK '97, M.J. HARRISON '00, FRUGOLI '01, UDVARDI '01].

M. truncatula et *Lotus japonicus* ont des caractéristiques génétiques similaires mais leurs nodules se développent de façon différente. De ces deux légumineuses, *M. truncatula* est la plus proche phylogénétiquement des légumineuses cultivées en Europe, puisque *M. truncatula* appartient au groupe des Galégoïdes renfermant les Trifoliées (luzernes, trèfles), Viciées (pois, féveroles, lentilles, vesces) et Cicérées (pois chiche).

Beaucoup d'espèces de la famille des légumineuses sont difficiles à régénérer, ce qui constitue un handicap pour des études fonctionnelles par transformation génétique. Barker et ses collaborateurs [BARKER '90] ont réalisé des études de transformation de *M. truncatula* Gaertn cv Jemalong avec *Agrobacterium tumefaciens*, et sa capacité de régénération a été un argument en faveur du rôle de *M. truncatula* comme plante modèle. Les travaux de Nolan et ses collaborateurs [NOLAN '89] ont permis d'augmenter la capacité de régénération de *M. truncatula* cv. Jemalong par des cycles de culture *in vitro* et de régénération, donnant la lignée Jemalong 2HA [ROSE '99]. La lignée de référence A17 est également obtenue à partir de la variété Jemalong.

Ainsi, *M. truncatula* est maintenant bien établie comme plante modèle pour les légumineuses, permettant l'étude des symbioses de fixation d'azote et de mycorhizes, mais aussi des stress biotiques, de la biologie de la plante et de la réponse aux agents pathogènes.

c. Définition de l'approche : à la découverte de la famille des A1b chez *Medicago truncatula*

L'intérêt de PA1b en tant que toxine insecticide d'origine végétale incite à l'étude de la variabilité naturelle des homologues des Albumines 1 chez l'espèce modèle *Medicago truncatula* dans l'objectif de répondre à deux aspects : le premier fondamental consiste à retracer l'histoire évolutive des homologues d'Albumines 1 dans le génome de *Medicago truncatula* afin de reconstruire les parentés au sein des A1b.

Et le deuxième plus appliqué concerne l'étude de la relation structure-activité dans l'espoir de trouver des séquences originales plus toxiques ou à spectre d'action modifié.

Toutefois, la diversité des peptides PA1b au sein de la même espèce a d'abord été suggérée par les travaux de Higgins *et al.*, qui a permis d'identifier quatre gènes fonctionnels présents dans le génome de pois et exprimés dans les cotylédons de pois [HIGGINS '86]. À ce jour, sept isoformes peptidiques de PA1b ont été isolés et caractérisés biochimiquement dans le petit pois [HIGGINS '86, WATANABE '94, DELOBEL '98, SANDRINE LOUIS '04b, TAYLOR '04], ce qui indique que ces peptides appartiennent à une famille multigénique dont les membres ont divergé légèrement. Presque toutes les formes auraient conservé leurs propriétés insecticides puisque les variations d'acides aminés ne sont pas situées à l'intérieur des positions clés nécessaires au maintien de l'activité toxique.

La recherche et l'identification des homologues de PA1b dans les graines des Fabacées ont été réalisées par l'élaboration d'une approche combinant trois méthodes complémentaires: (1) un niveau moléculaire qui consiste à cloner les séquences génomiques homologues à PA1, (2) une méthode biologique qui se base sur une réponse à des essais biologiques spécifique aux PA1b et défini par une toxicité différentielle entre des souches sensibles et résistantes du charançon du riz, *S. oryzae*, le test étant d'abord effectué sur les graines extraites avec différents solvants; (3) un niveau biochimique qui montre la capacité de liaison de l'extrait membranaire de l'intestin des insectes sensibles à PA1b du pois radioactif. De cette façon, les nouveaux gènes albumine 1, et leurs produits associés, ont été caractérisés dans le soja *Glycine max* et le haricot *Phaseolus vulgaris* alors que dans la luzerne *Medicago truncatula* aucun peptide n'a été détecté biochimiquement, malgré la présence d'une forte activité insecticide et de gènes homologues [SANDRINE LOUIS '04b].

La pertinence de l'approche décrite ci-dessus conduit à une étude plus large sur 88 espèces supplémentaires dispersées parmi les trois sous-familles (Caesalpinioideae, Mimosoideae et Papilionoideae) des Fabacées [S LOUIS '07]. 19 gènes PA1-like ont été caractérisés. Ainsi, à ce jour, les peptides homologues a PA1b dans les plantes semblent être présents majoritairement chez les légumineuses, ce qui suggère fortement que cette famille peptidique est une importante ligne de défense des graines contre les insectes de la famille des Fabacées [SANDRINE LOUIS '04b, S LOUIS '07].

L'exploration des génomes est apparu comme une ouverture vers la découverte de nouveaux représentants de cette famille des albumines 1 dont l'étude ouvre de nombreuses perspectives sur la connaissance de leurs activités biologiques et la possibilité de déviation ou de diversification vers une néo-fonctionnalisation. Le génome de *Medicago truncatula*, publié en 2011, a donné la possibilité d'étudier cette famille de protéines dans cette espèce modèle. L'étude de leurs diversités apparaît intéressante pour plusieurs raisons. Elle pourrait permettre de déterminer l'importance écologique de ces peptides (rôle de défense et autre rôle ?) et également d'obtenir des indications sur l'évolution de cette famille peptidique.

Ainsi le **chapitre 2** présentera une première approche «*in silico*» mise en place afin d'étudier l'évolution des homologues de PA1b chez *Medicago truncatula*. Plus précisément, une étude plus approfondie des bases de données génomiques a été réalisée afin d'identifier des homologues de PA1b en utilisant des outils de recherche de similarité de séquence (BLASTP, profils HMM, etc.). Ensuite, une analyse phylogénétique a été conduite en utilisant des méthodes de reconstruction d'arbres les plus efficaces (maximum de vraisemblance et approche bayésienne). Ces travaux ont permis d'identifier l'origine de PA1b pendant la diversification des légumineuses et d'identifier les événements qui ont façonné l'histoire évolutive des PA1b (duplications de gènes, pertes, etc.) dans les différentes lignées. Les pressions sélectives qui agissent sur les gènes Albumines 1 et les résidus d'acides aminés clés potentiels ont été étudiées en comparant les taux de substitution synonyme et non synonymes dans les séquences nucléiques correspondantes à l'aide de différentes approches telles que PAML. Quelques séquences choisies ont été synthétisées chimiquement suivies d'un repliement *in vitro*, et l'activité biologique de certaines variantes A1b a été étudiée à la fois en mesurant leur affinité de liaison au récepteur et leur toxicité contre les cellules d'insectes Sf9. La recherche de variants naturels de PA1b pourrait aider à souligner de potentiels acides aminés importants pour l'activité. Enfin, il sera peut-être

possible de déterminer des homologues de CL₅₀ inférieure à celle de PA1b, ou contournant la résistance des insectes insensibles à la toxine de pois.

2. Résultats

Ces travaux ont fait l'objet d'un article soumis dans *Genome biology*.

Article

Genome-wide analysis identifies gain and loss/change of function within the small multigenic insecticidal Albumin 1 family in *Medicago truncatula*

Chapitre 3

La famille des A1b au sein des Légumineuses originaires du Liban

1. Introduction

Dans le cadre de la bourse doctorale qui m'a été attribuée par l'Agence Universitaire de la Francophonie (AUF), le Conseil National de la Recherche Scientifique Libanais (CNRS-L) et le Projet CEDRE, des collaborations avec l'université Saint-Joseph (USJ) au Liban ainsi qu'avec le Centre international des recherches agricoles dans les régions sèches (pour *International Center for Agricultural Research in the Dry Areas* « ICARDA ») ont été établies. Un des centres d'intérêt de l'université Saint Joseph concerne d'une part la connaissance de la diversité et de l'histoire évolutive de la flore libanaise et d'autre part l'organisation internationale à Beyrouth « ICARDA » maintient des ressources génétiques des Légumineuses. Ce projet a occasionné 3 séjours de plusieurs mois (1 séjour par an) au Liban.

L'objectif de ce projet consistait 1) à explorer la diversité de la famille des albumines 1 dans l'attente d'un élargissement du panel de la biodiversité et 2) à faire émerger de nouvelles molécules homologues de PA1b aux propriétés originales (variabilité de cibles insectes, niveau de toxicité moléculaire, toxicité envers des souches résistantes de charançons, propriété structurales inattendues ...), mais également 3) à évaluer les propriétés insecticides des espèces de légumineuses d'intérêt agronomique local/régional ou endémiques du Liban (*Pisum*, *Lens*, *Cicer*, *Vicia*, *Medicago* sp....) : ces propriétés pouvant concerner la variabilité d'expression des toxines A1b (tissu spécificité), comme la présence de molécules/fractions toxiques soit insoupçonnées, soit non-identifiées dans les espèces ciblées.

Ainsi cette étude répondra d'une part à une volonté de l'équipe française « Entomotoxines » d'explorer le volet plante d'un projet important sur le mode d'action de l'entomotoxine de légumineuses, avec une démarche partenariale méditerranéenne, et d'autre part à celle de l'équipe libanaise qui étudie la gestion et la santé des cultures et des plantes en développant des stratégies pour la conservation de la biodiversité végétale et pour améliorer la protection des ressources végétales vis-à-vis des pathogènes.

Plus précisément, d'une part, l'équipe « Entomotoxines » canalise principalement ses recherches sur l'identification et la caractérisation du mode d'action de molécules naturelles à activité insecticide utilisables dans le cadre d'une stratégie de développement durable. Ses activités se focalisent autour du peptide naturel entomotoxique (PA1b) d'origine végétale [GRESSENT '07] découvert dans le laboratoire BF2i. D'autre part, dans le cadre du développement et de l'étude de biopesticides, l'équipe libanaise

travaille dans le champ de l'interaction hôte-pathogène et mène des activités de recherche d'agents biologiques (eg agents viraux, densovirus, spécifiques des insectes) utilisables en lutte contre des ravageurs de culture majeurs, comme les lépidoptères, dans différentes régions agricoles au Liban.

Pour répondre à cet objectif, une stratégie de travail a été mise en place.

Il s'agit d'une part de sélectionner une source de variabilité naturelle du trait recherché (ressources génétiques d'intérêt pour l'équipe libanaise), d'autre part de choisir une stratégie de mise en évidence des gènes ou des traits recherchés (phénotypes biologiques), pour laquelle l'équipe « Entomotoxines » a une expérience de plusieurs années.

Le choix des plantes testées fera l'objet d'une attention particulière. Par référence aux acquis du laboratoire, ce choix se limitera aux légumineuses (Fabaceae), seule famille de plantes où des peptides A1b-like ont été détectés [S LOUIS '07]. En complément de légumineuses d'intérêt, dont l'espèce modèle de référence *Medicago truncatula* (groupe des Galégoïdes), la sélection privilégiera des espèces sauvages ou ancestrales originaires du Liban, centre d'origine de la domestication de ces plantes, ou des espèces d'intérêt pour l'agriculture régionale. A ce titre les contacts ont été pris et concrétisés avec l'ICARDA, organisme international qui maintient des ressources génétiques de légumineuses et l'université Saint-Joseph qui s'intéresse à la diversité des plantes libanaises. En explorant la littérature, des espèces potentiellement intéressantes avaient été identifiées [TOHMÉ '07] et des contacts ont été établis avec les auteurs (M. et Mme Tohmé). Des prélèvements sur le terrain, en réalisant des groupes de recherche locale, ont également été réalisés.

La recherche dans les plantes sélectionnées, des gènes homologues à celui codant pour la toxine peptidique PA1b sera réalisée par une approche moléculaire sur ADN génomique suivie par une PCR et un clonage par homologie.

2. Matériels et Méthodes

a. Choix des espèces testées

Les critères appliqués lors de la sélection des espèces étudiées sont triples :

La position phylogénétique : les espèces choisies sont réparties au sein de l'arbre phylogénétique des Légumineuses établie par Wojciechowski [WOJCIECHOWSKI '04].

L'originalité : seules les espèces originaires voire endémiques du Liban ont été sélectionnées.

La disponibilité : les espèces dont nous pouvions nous procurer des graines ont pu être étudiées. Les fournisseurs sont multiples.

La liste des espèces choisies et leur provenance sont indiquées dans le Tableau 4.

Tableau 4 : Liste des espèces sélectionnées, origine géographique et fournisseur.

Sous-famille	Tribu	Genre	Espèce	Fournisseur
Caesalpinioideae	Cercideae	Cercis	<i>siliquastrum</i>	USJ
Papilionoideae	Genisteae	Spartium	<i>Junceum</i>	L.KARAKI
Papilionoideae	Phaseoleae	Lablab	<i>purpureus</i>	L. KARAKI
Papilionoideae	Trifolieae	Medicago	<i>rotata</i>	ICARDA/ L.KARAKI
Papilionoideae	Trifolieae	Medicago	<i>radiata</i>	ICARDA
Papilionoideae	Trifolieae	Ononis	<i>adenotricha</i>	USJ
Papilionoideae	Loteae	Lotus	<i>angustissimus</i>	USJ
		Astragalus	<i>trichopterus</i>	USJ
		Astragalus	<i>ehdenensis</i>	USJ
		Astragalus	<i>angulosus</i>	USJ
		Astragalus	<i>cephalotes</i>	USJ
		Astragalus	<i>zachlensis</i>	USJ
		Astragalus	<i>berytheus</i>	USJ

La liste complète des espèces sélectionnées (annexe 1) n'a pas pu être traitée par l'approche moléculaire pour des raisons de disponibilité de la matière végétale. De plus, les Papilionoideae apparaissent bien plus re-

présentées que les deux autres sous-familles, les Caesalpinioideae et Mimosoideae. Ceci est dû au fait que les Papilionoideae sont les plus nombreuses des Fabacées.

b. Extraction d'ADN

L'ADN génomique des espèces récupérées a été extrait à partir des graines à l'aide du kit DNeasy Plant Mini Kit (QIAGEN). Les graines sont broyées dans l'azote liquide avec un pilon et un mortier puis mises en suspension dans une solution de lyse fournie avec le kit d'extraction. L'extraction de l'ADN génomique est poursuivie selon les indications mentionnées dans le protocole.

c. Recherche de fragments de gènes homologues par PCR

La PCR (Polymerase Chain Reaction) est une méthode d'amplification de l'ADN utilisant les propriétés de réplication de l'ADN. Différentes amorces ont été utilisées pour ce travail :

17 amorces dégénérées (Tableau 5) (voir la détermination de ces amorces dans la partie résultats).

Tableau 5 : Séquences des amorces dégénérées utilisées en PCR

Nom de l'amorce	Séquence de l'amorce
FA-3	5'CAG-TAA-TGT-TSC-CRA-YGR-AGA-A3'
FA+9	5'CCR-AYG-RAG-AAB-GTA-GAA-GCA3'
FB-9	5'CAT-TTA-CAG-TAA-TGT-TYC-CRA-TG3'
FB+9	5'CCR-ATG-AAR-AAA-RTA-GAA-GGA-G3'
FC-9	5'CAT-TTK-CAG-TCA-TAT-KYC-CGA-T3'
FC+9	5'CCG-ATG-AAG-AAG-GTA-GAA-GCA3'
FCN-10	5'CCA-TTT-GCA-GTA-ATY-GTC-CAG-A3'
FCN+18	5'AAC-GTT-GAA-GMA-GGG-CAR-TGT3'
FEAC+9	5'TGA-WGA-AGM-WGG-TAR-KAG-CAG3'
FE8g+30	5'AGA-WTG-TTY-GGG-TRY-TTG-TTC3'

-1	5' AAT-GRS-TTM-BGY-DAR-RST-YRY-WHC-T3'
+26	5' TGG-CTC-THT-TCY-TGB-TTR-CCA-C3'
+24	5' TTT-KGY-TVY-NKT-CWT-KYT-YRC3'
+35	5' TCW-TKY-TYR-CNV-H AW-THT-GTV-HGT-AT3'
RC4	5' AAT-CTC-TAC-AGG-CAT-CTT-CAA-G3'
RCN	5' ATT-CTT-AAG-TGG-CAA-TTT-CCA-AA3'
R2	5' AAR-CAC-CAN-CCR-TRN-TCD-AT3'

Ces amorces sont dites dégénérées car certaines de leur positions ne sont pas pleinement déterminées, ainsi R représente A ou G ; Y, C ou T ; M, A ou C ; K, G ou T ; S, C ou G ; W, A ou T ; H, A ou C ou T ; B, C ou G ou T ; V A ou C ou G ; D, A ou G ou T et N s'apparie indifféremment à n'importe laquelle des quatre bases. Plus le degré de dégénérescence est grand, moins l'oligonucléotide est spécifique et plus il doit être concentré lors de la PCR. Les réactions de PCR ont été effectuées dans un thermocycleur (Eppendorf MasterCycler) dans un volume réactionnel de 25 µl sur des quantités de matrice variant de 10 à 50 ng, avec 2 unités de Taq polymérase et son tampon 10 X (Uptitherm DNA polymerase d'Optima INTERCHIM) selon les schémas de la Figure 15.

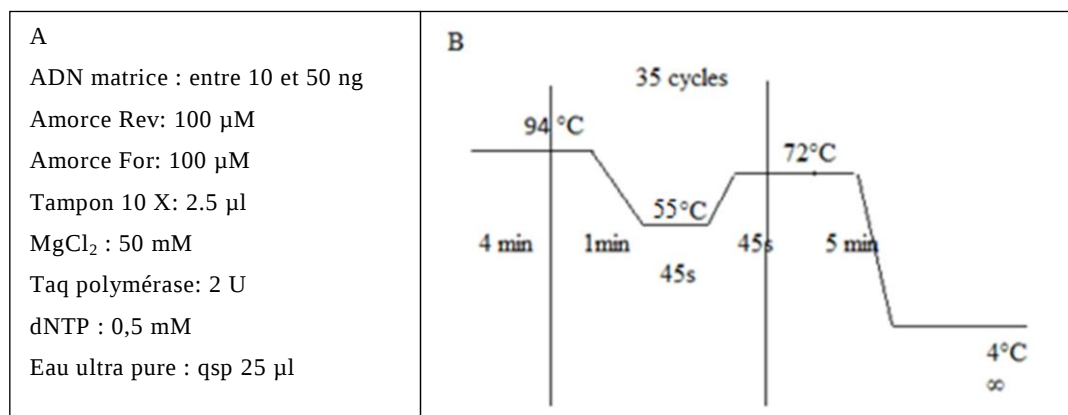


Figure 15 : Protocole d'amplification d'ADN par PCR avec des amorces dégénérées (concentrations finales).

A : mélange réactionnel, B : cycle de réaction.

d. Electrophorèse sur gel d'agarose

Pour s'assurer de la qualité de l'amplification, une électrophorèse des produits de PCR a été réalisée sur gel d'agarose à 1%, sous une tension de 100 V. Le tampon de migration utilisé est le TAE (Tris-HCl 1,6 mM pH 8 ; EDTA 4 μ M ; acétate de sodium 1,6 mM). La migration est suivie grâce à une solution de dépôt (Tampon de charge) ajoutée à la solution d'ADN à environ 20 % du volume. Un marqueur de poids moléculaire allant de 80 bp à 10000 bp (Euromedex Mass DNA ladder Mix) a été déposé dans le gel à côté de l'ADN et permet d'évaluer la taille des fragments amplifiés. Après migration, les fragments d'ADN ont été observés sous UV grâce au bromure d'éthidium (BET) ajouté au gel d'agarose.

e. Clonage

Le clonage se fait en plusieurs étapes. Après avoir purifié les fragments d'ADN, ceux-ci sont insérés dans un vecteur plasmidique. Le plasmide est alors introduit par choc thermique dans des bactéries compétentes.

e.1 Préparation des plasmides portant l'insert

Les fragments d'ADN amplifiés à liguer ont tout d'abord été purifiés. Ils ont été déposés sur gel d'agarose 1%. Après migration, les bandes visualisées sous UV ont été découpées. La purification de l'ADN a alors été effectuée à l'aide du kit NucleoSpin® Gel and PCR Clean-up de Macherey Nagel selon les recommandations du fournisseur. Avant la ligation, l'ADN purifié, a été dosé par un spectrophotomètre Nanodrop ND 1000. La ligation a été effectuée grâce au kit TOPO TA cloning (Invitrogen). Les plasmides sont fournis linéarisés avec une simple thymidine (T) à chaque extrémité du plasmide. Les fragments d'ADN purifiés possèdent une simple deoxyadenosine (A) à chaque extrémité grâce à l'activité terminale transférase non-dépendante de la matrice de la *Taq* polymérase. Le produit de PCR purifié a été placé en contact du plasmide et soumis à la réaction de

ligation (ligation entre les extrémités T et A) selon les recommandations du fournisseur (Topo TA Cloning, Invitrogen).

e.2 Transformation bactérienne

La bactérie employée ici est *Escherichia coli* : Souche TOP10 (Invitrogen) chimiquement compétente avec une efficacité de transformation de 1×10^9 cfu/ μ g de plasmide d'ADN utilisé.

La transformation par choc thermique nécessite 50 μ l de bactéries chimiquement compétentes et environ 50ng d'ADN. Le mélange est laissé 30 minutes sur la glace, puis transféré 30 secondes à 42°C puis incubé 2 minutes dans la glace. 250 μ l d'un milieu S.O.C (2% Tryptone, 0.5% Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) est ajouté pour une incubation de 1h à 37°C. Les bactéries sont ensuite étalées sur boîte de pétri contenant du LB-agar, de l'antibiotique (ampicilline) et du X-gal (40 mg/ml, Euromedex) pour la sélection des clones. Cette étape permet de différencier les clones transformés avec un plasmide portant un insert (blancs car lac⁻), des clones non-transformés (bleus car lac⁺).

f. Minipréparation d'ADN plasmidique

Les bactéries blanches sélectionnées seront mis en culture dans un milieu LB Broth dans lequel s'ajoute l'ampicilline. Après incubation, pendant une nuit à 37°C, l'ADN plasmidique sera extrait à partir des bactéries en suspension via le kit Nucleospin Plasmid- Plasmid DNA purification (Macherey Nagel), en suivant les protocoles indiqués par les fournisseurs.

g. Séquençage

Le séquençage de l'ADN plasmidique est réalisé par Biofidal (Villeurbanne). Et les résultats seront étudiés via le programme ApE : Plasmid Editor ainsi que par le programme SignalP 4.1 [PETERSEN '11] qui permet de prédire un site de clivage d'un peptide signal dans la séquence protéique trouvée.

h. Analyses des séquences

Les séquences protéiques complètes et incomplètes spécifiant des homologues de PA1b ont été alignées en utilisant le logiciel SeaView [GOUY '10].

3. Résultats

Recherche des gènes homologues au gène codant pour PA1b dans le génome des plantes libanaises

L'ADN génomique obtenu a été aliquoté et conservé à -20°C ou utilisé directement pour la recherche des gènes homologues à celui de PA1b.

a. Amplification de fragments de gènes homologues à PA1b (choix des amorces)

Le gène codant pour PA1b est constitué en réalité de trois parties (exon1-intron-exon2), codant pour une préproprotéine. La séquence signal, rapidement clivée lors de la maturation protéique, est suivie de la séquence codant pour PA1b puis de celle codant pour PA1a, une autre albumine de 6 kDa, riche en cystéine mais ne présentant pas d'activité entomotoxique et de fonction encore inconnue [DELOBEL '98]. L'alignement des séquences des homologues de PA1 auparavant connues chez les légumineuses a permis d'observer différentes régions bien conservées. Cinq amorces ont été choisies en se basant sur l'alignement des séquences de gènes des homologues de PA1b provenant de différentes espèces de légumineuses. Ces cinq amorces sont les suivantes : +1, +26, +24, +35 (Forward ; incrémentation à partir de l'ATG de la séquence génique) et R2 (Reverse). Toutes les autres amorces citées dans le tableau 2 de la partie « Matériels et Méthodes » ont été choisies en se basant sur l'alignement des séquences de gènes homologues de PA1b chez *Medicago truncatula*. La diversité de séquences observée chez l'espèce modèle des légumineuses dans le chapitre 2, nous a emmené à dessiner des amorces dégénérées représentant chaque clade phylogénétique des Albumines 1 chez *Medicago truncatula*. Les trois amorces reverse dégénérées ont été définies dans la séquence codant pour PA1a, mieux conservée que PA1b, ce qui doit permettre d'obtenir les séquences C-terminales des homologues de PA1b (voir localisation des amorces en Annexe 2).

a.1 Approche par PCR classique

Les différentes paires formées par ces 17 amorces ont été utilisées (Tableau 6). Après avoir déterminé le nombre de cycles (35), la température d'appariement et la concentration en amorces permettant une amplification correcte, les fragments obtenus ont été analysés.

Tableau 6 : Différents couples d'amorces utilisés ainsi que la taille approximative minimale des « bandes » attendues.

Couple d'amorce (pour plus de précisions voir annexe 2)	Taille approximative minimale de la bande attendue (pb)
FA-3/R2	286
FA+9/R2	274
FB-9/R2	295
FB+9/R2	277
FC-9/RC4	413
FC+9/RC4	395
FCN-10/RCN	232
FCN+18/RCN	204
FEAC+9/R2	265
FE8g+30/R2	237
-1/R2	398
+26/R2	371
+24/R2	304
+35/R2	293

Les 10 premiers couples d'amorces représentés dans le tableau 3 permettront une amplification qui débutera au début de la séquence de l'exon 2. Alors que les 4 derniers couples d'amorces permettront une amplification localisée au sein de l'exon 1. Le résultat de ces 4 derniers couples montrera ainsi une amplification de la séquence de l'intron en plus des séquences des exons.

Des résultats positifs ont été obtenus avec les paires FA-3/R2, FA+9/R2, FB-9/R2, FCN-10/RCN, FE8g+30/R2, -1/R2 et +26/R2. L'expérience a montré que certains couples d'amorces utilisés présentent une aspécificité se traduisant par une hybridation de chaque amorce au bon endroit dans les 2 sens de la séquence nucléique mais le résultat de l'amplification ne donne pas un homologue de PA1b ou une seule amorce du couple s'hybride deux fois dans les 2 sens de la séquence nucléique, n'engendrant pas de séquence homologue à PA1b.

Sur les 14 ADNs purifiés (provenant des 13 espèces listées dans tableau 4 ; 2 échantillons pour une même espèce *Medicago rotata*), 8 ont montré des amplifications avec plusieurs jeux d'amorces utilisés. Les séquences correspondantes ont été clonées et séquencées, mais s'avèrent résulter d'une amplification non spécifique. Concernant les 6 ADNs restants correspondants à 6 espèces différentes (*Lablab purpureus*, *Astragalus ehdenensis*, *Medicago rotata*, *A. zachlensis*, *A. angulosus* et *A. trichopterus*) et pour lesquels des produits de PCR de taille inférieure (rare) ou supérieure à celle attendue ont été obtenus (Figure 16), un clonage et un séquençage ont été réalisés révélant à chaque fois des séquences codantes homologues à celle de PA1b. Les amplifications obtenues sont ainsi correctes (spécifiques) (Annexe 3).

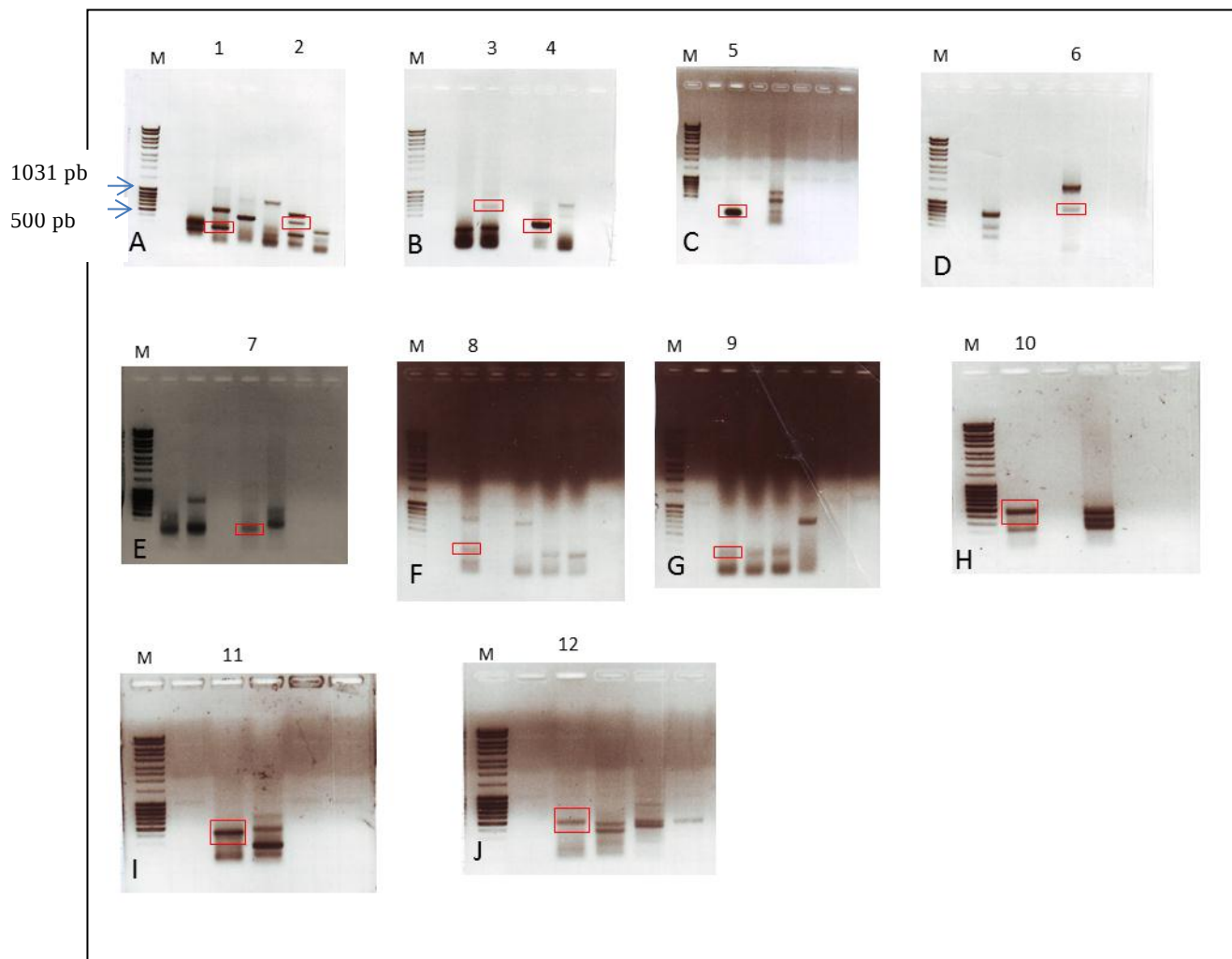


Figure 16 : Fragments d'ADN amplifiés par PCR avec différents couples d'amorces « PA1 » sur ADN génomique des espèces végétales. M : marqueur de poids moléculaire ; 1 : *Lablab purpureus* ; 2 : *Astragalus ehdenensis* ; 3 -4-5 : *Medicago rotata* ; 6 : *A. zachlensis* ; 7 : *Lablab purpureus* ; 8 : *A. angulosus* ; 9 : *A. zachlensis* ; 10 : *A. trichopterus* ; 11 : *A. angulosus* ; 12 : *A. zachlensis*. Le gel A : couple FA-3/R2 ; B : FB-9/R2 (pour l'échantillon 3) et FC+9/R2 (pour l'échantillon 4) ; C : FCN-10/RNC (pour l'échantillon 5) ; D et E : FE8g+30/R2 ; F et G : FA+9/R2 ; H : -1/R2 ; I et J : +26/R2. Les bandes clonées qui ont présenté une homologie à PA1b sont encadrées en rouge.

b. Analyse des séquences

b.1 Variabilité des séquences

L'alignement protéique de toutes les séquences obtenues est présenté dans la Figure 17. Cet alignement permet de distinguer les zones variables et les zones conservées entre A1bs, précisant et renforçant les conclusions établies à la fin du chapitre 2 sur la diversité des homologues de PA1b inter et intra-espèce. On constate en effet que, lors de l'élargissement de l'étude à différentes espèces de légumineuses, les acides aminés de structure sont conservés. Si nous considérons la numérotation des résidus d'après la séquence du PA1b du pois **P62931** (ALB1F_PEA), il s'agit en particulier des cysteines, des prolines, de l'arginine 21, de la leucine 27 et de la glycine 30 (avec de rares exceptions).

Différentes variations apparaissent au niveau de PA1b, en nombre et nature de résidus. Le peptide intermédiaire putatif encore appelé propeptide (séparant PA1b de PA1a) varie également en longueur par rapport à la séquence du pois. On peut également observer la présence d'un résidu cystéine supplémentaire dans les séquences 3_MEDRO et 5_MEDRO. En comparant les séquences obtenues, on remarque également la variabilité de la séquence au niveau du peptide signal. A noter, certains couples d'amorces ne permettent pas l'amplification de la totalité de la séquence du peptide signal. Au contraire, on constate que les parties de PA1a amplifiées par les amorces reverses restent fortement conservées.

De plus, la séquence 4_MEDRO possède un codon stop juste après la séquence de PA1b. Alors que 6_ASTZA et 9_ASTZA possèdent un codon stop au milieu de la séquence de PA1a. Parmi les séquences trouvées, celle de 4_MEDRO est la plus diversifiée. A l'exception des 6 cysteines, on trouve une très grande variabilité de la nature des résidus d'acides aminés.

En comparant les séquences obtenues, on remarque que les séquences 3_MEDRO et 5_MEDRO, provenant de la même espèce (*Medicago rotata*) mais en utilisant 2 couples d'amorces différentes, ont des séquences identiques. De même, les séquences 8_ASTAN et 11_ASTAN (*Astragalus angulosus*) sont identiques et 6_ASTZA et 9_ASTZA. Pour la suite de cette étude, nous allons conserver les 3 séquences 3_MEDRO, 11_ASTAN et 9_ASTZA et nous allons éliminer les 3 autres.

Les séquences 9_ASTZA, 10_ASTTR et 11_ASTAN ont été analysées par le programme SignalP 4.1 et ont montré un site de clivage du peptide signal entre les positions 27 et 28 (VEG-TD) pour 9_ASTZA, entre

20 et 21 (VEA-AD) pour 10_ASTTR et entre 17 et 18 (TFG-FA) pour 11_ASTAN.

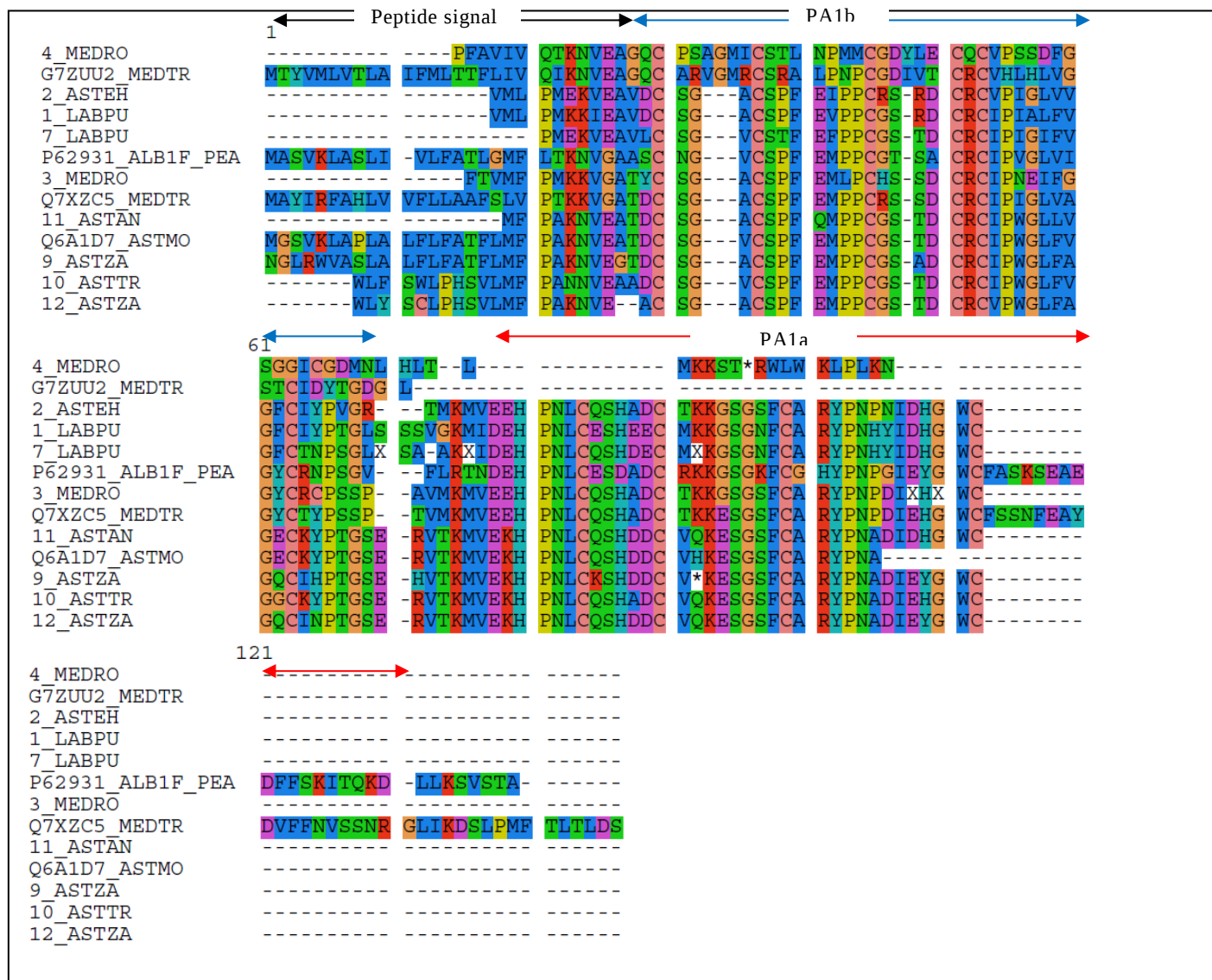


Figure 17 : Alignement protéique des séquences caractérisées dans notre travail et des séquences issues de la littérature.

Les espèces sont : MEDRO : Medicago rotata, LABPU : Lablab purpureus, ASTZA : Astragalus zachlensis, ASTAN : Astragalus angulosus, ASTEH : Astragalus ehdenensis, ASTTR : Astragalus trichopterus, MEDTR : Medicago truncatula, ALB1F_Pea : Pisum sativum, ASTMO : Astragalus monspessulanus.

Les séquences extraites de la littérature sont : Q6A1D7_ASTMO, P62931_ALB1F_PEA, Q7XZC5_MEDTR et G7ZUU2_MEDTR.

Les numéros 1 à 12 correspondent aux « bandes » définies dans la Figure 16.

b.2 Phylogénie moléculaire

Afin de localiser les séquences trouvées parmi les séquences A1b-like déjà connues, une analyse phylogénétique a été réalisée sur les 13 séquences protéiques incomplètes résultant de l'amplification par les différents couples d'amorces et quelques séquences déjà publiées (une séquence de *Pisum sativum*, une d'*Astragalus monspessulanus* et deux de *Medicago truncatula*). Le nombre de sites informatifs est limité par la petite taille des séquences. Ils ont été analysés en utilisant l'algorithme BioNJ [GASCUEL '97] et la méthode de maximum de vraisemblance. L'arbre consensus obtenu (non raciné), présenté dans la Figure 18, montre la séparation en trois groupes des séquences : le premier groupe englobe les séquences les plus proches du PA1b toxique du pois (Q7XZC5_MEDTR, 3_MEDRO, 2_ASTEH, 1_LABPU et 7_LABPU), le deuxième groupe regroupe des séquences homologues de PA1b provenant des astragales et le troisième groupe montre la présence de la séquence « noduline » (forme courte de PA1) en plus de celle de 4_MEDRO.

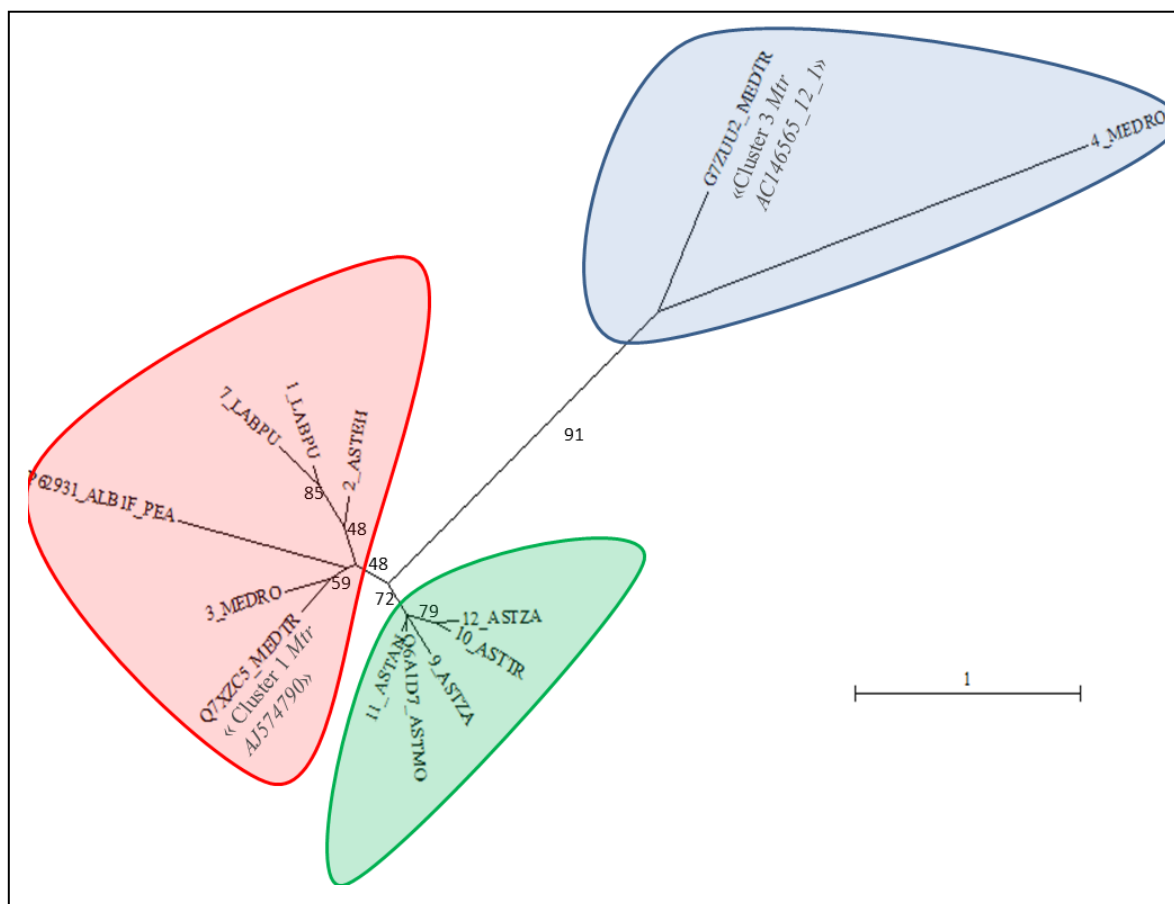


Figure 18 : Arbre de phylogénie moléculaire (séquences "incomplètes", BioNJ, Maximum de vraisemblance) réalisé sur les séquences protéiques des A1b.

Les valeurs de bootstrap sont indiquées. Les noms des séquences sont légendés Figure 17.

Les séquences d'Astragales, à l'exception de 2_ASTEH, se regroupent conformément à leur taxon. En comparant la séquence 4_MEDRO à celle du G7ZUU2_MEDTR (forme courte de PA1 identifiée chez *Medicago truncatula*), on trouve un pourcentage d'homologie égale à 33.3%.

Toutefois, le nombre limité de ces séquences apporte peu d'information supplémentaire à l'étude de l'évolution de la famille peptidique des Albumines 1. Même si l'originalité en termes de séquence peptidique peut paraître limitée, elles restent néanmoins intéressantes à exploiter notamment pour l'étude relation structure-fonction.

4. Discussion

Par notre approche de PCR avec amorces dégénérées, 6 nouvelles espèces, toutes originaires du Liban, ont révélé un produit spécifiant un homologue des A1bs. Pour certaines d'entre elles (*Lablab purpureus*, *Medicago rotata* et *Astragalus zachlensis*) deux séquences homologues différentes ont pu être caractérisées. Les séquences obtenues sont relativement variables en nature et nombre d'acides aminés. En totalité, pour chacune des 9 séquences, la structure du potentiel produit du gène en préproprotéine apparaît conservée, même si la maturation exacte de ces protéines reste inconnue.

Ainsi parmi les 13 plantes testées (14 ADNs extraits : deux différents lot de graines de *Medicago rotata* ont été testés), représentatives de la famille des fabacées et des deux sous-familles Papilionoideae et Caesalpinioideae et des genres : *Cercis*, *Spartium*, *Lablab*, *Medicago*, *Ononis*, *Lotus* et *Astragalus*, nous avons trouvé 9 formes de PA1-like. Toutes ces formes identifiées sont originales. Parmi les différentes espèces testées, nous avons 4 espèces non répertoriées (non trouvées dans les bases de données disponibles de NCBI). Toutes ces espèces appartiennent au genre *Astragalus* : *Astragalus ehdenensis*, *A.zachlensis*, *A.trichopterus* et *A.angulosus*.

Les résultats obtenus nous permettent d'émettre quelques hypothèses quant à l'évolution des A1b au sein des Légumineuses originaires du Liban.

L'analyse des relations phylogénétiques entre homologues de PA1b, bien que limitée par la faible longueur des séquences et leur faible nombre, semble montrer une évolution récente de cette famille peptidique. Pour pouvoir interpréter l'évolution de ces A1b multigéniques, il conviendrait de déterminer les éventuels orthologues et paralogues. Le regroupement des séquences d'une même espèce impliquerait une évolution essentiellement post-spéciation donc récente, alors que la présence d'homologies

plus importantes entre séquences d'espèces différentes plutôt qu'entre séquences d'une même espèce indiquerait une duplication avant spéciation et donc une histoire évolutive plus ancienne.

Enfin, avant toute conclusion, il ne faut pas négliger le fait que peu d'Astragales ont été testées. Or, le genre *Astragalus* comporte plusieurs centaines d'espèces au Proche-Orient, généralement liées à des habitats variés où prédominent des conditions de stress pouvant impliquer des adaptations particulières. Elles peuvent également être définies comme des espèces clés de certains écosystèmes fragiles, en facilitant la résilience après perturbation. Il conviendrait donc d'élargir l'échantillonnage des espèces dans ce genre d'*Astragalus*. Nous avons en notre possession 12 autres ADNs provenant de 12 espèces différentes d'Astragales originaires du Liban qui n'ont pas encore été testées. Ces espèces pourront être explorées dans le but de déterminer les différents paralogues des Albumines 1 au sein des différentes espèces d'*Astragalus* et de comparer ses paralogues selon la position phylogénétique des Astragales du Liban [ABDEL-SAMAD '13]. Cette étude pourra ainsi alimenter l'histoire évolutive des Albumines 1 au sein de ce genre et expliquer la diversité ou la conservation de cette famille au sein des Astragales en fonction de leur distribution bio-géographique.

D'autre part, de nouvelles perspectives ont été proposées pour la suite de ce travail. La PCR, malgré la dégénérescence des amorces, est sensible aux variations de séquences et pourrait ne pas avoir détectée des séquences divergentes mais néanmoins homologues chez les espèces *Cercis siliquastrum*, *Ononis adenotricha*, *Lotus angustissimus*, *Medicago radiata*, *Astragalus cephalotes*, *A. berytheus* et *Spartium junceum*. Pour essayer de contourner cette sensibilité, il est préférable de choisir de nouveaux jeux d'amorces dont le pourcentage de dégénérescence est faible. En fonction des moyens financiers disponibles une alternative serait de pouvoir investir dans le séquençage du génome d'espèces cibles d'intérêt.

L'étude plus approfondie des séquences trouvées consisterait, en se basant sur les résultats obtenus du travail de mutagenèse dirigée réalisé par Pedro Da Silva au sein de l'équipe « Entomotoxines » [DA SILVA '10]) et sur les résultats du travail chez *Mtr*, à développer une stratégie similaire à celle qui s'est avérée fructueuse dans le chapitre 2 à savoir : synthèse chimique de quelques-unes des formes les plus pertinentes suivie d'un repliement *in vitro* et exploration de leur activité biologique *in vivo* sur insectes en élargissant le spectre d'hôte (bactérie, champignon...) au cas où une divergence vers une néo-fonctionnalisation ait eu lieu au sein des homologues des A1bs.

De plus, ce projet est focalisé sur la relation entre la diversité des plantes cultivées et les processus de domestication et d'adaptation à l'envi-

ronnement agricole. L'objet principal de recherche associé est l'analyse de l'histoire évolutive et des modalités de la domestication et de l'adaptation des plantes cultivées méditerranéennes et tropicales afin de mieux exploiter et conserver leur diversité génétique. Il est connu que des composés du système de défense chez certaines plantes (cas des arcelines chez *Phaseolus vulgaris* : [OSBORNI '88, CARDONA '90]) sont perdus pendant le procédé de sélection imposé par la domestication. A ce titre, il serait très intéressant de comparer les patterns d'expression des A1-like entre les espèces sauvages (*Mtr*) et domestiquées (*Medicago sativa*). L'ICARDA possède des ressources génétiques de Médicago d'intérêt pour aborder cette problématique.

Toutes ces molécules homologues de PA1b et appartenant à la classe des « knottines » connues pour la diversité de leurs fonctions (antibactérienne...) permettent d'ouvrir des perspectives non seulement dans le domaine agronomique (lutte contre les ravageurs de cultures) mais aussi dans le domaine de la santé (pharmacologie) (voir le lien : <http://knottin.cbs.cnrs.fr>).

Finalement, les légumineuses représentent l'un des taxons végétaux les plus importants tant du point de vue de la biologie et de l'écologie fondamentale que du point de vue agronomique et environnemental. La recherche de molécules naturelles entomotoxiques pourra apporter, selon les choix effectués, des alternatives innovantes en lutte biologique ou lutte variétale. Ce travail participe ainsi à la connaissance et à la valorisation du patrimoine végétal du Liban. En ce sens il répond à une détermination nationale libanaise de préservation et de conservation des ressources génétiques sauvages (INRA de Tel Amara) et plus spécifiquement de réhabilitation des légumineuses (<http://www.fao.org/ag/agp/AGPC/doc/Counprof/lebanon/lebanon.html>) avec des collectes à partir de 2003 au Lebanese Agricultural Research Institute (LARI) (<http://www.tela-botanica.org/actu/article3214.html>).

Chapitre 4

Expression hétérologue de PA1b

1. Introduction

La troisième partie de ce travail de thèse avait pour objectif la mise au point d'un système d'expression et de purification de PA1b recombinant avec pour but ultime l'étude des relations structure - activité de ce peptide, ceci afin de relier les variabilités de séquence et de toxicité observée. La structure de PA1b, petit peptide compact et amphiphile, nécessite des modifications post-traductionnelles spécifiques (structuré par 3 ponts disulfures). Le choix du système d'expression est crucial pour obtenir un produit bien structuré et donc actif.

Un système d'expression résulte de la combinaison d'un vecteur d'expression et d'une cellule hôte. Le vecteur d'expression peut être un plasmide ou un virus. Son rôle est d'introduire un gène étranger dans la cellule hôte pour qu'il puisse y être exprimé. Pour cela, le vecteur doit contenir au moins un promoteur reconnu par la cellule hôte, des signaux de polyadénylation et dans certains cas un signal de sécrétion. Procaryote ou eucaryote, la cellule hôte doit permettre la production de la protéine recombinante en grande quantité, grâce à des systèmes de culture performants. Elle doit être capable de réaliser toutes les modifications post-traductionnelles permettant d'obtenir une protéine identique à la protéine naturelle tout en conservant la même activité biologique. Enfin, le milieu de culture doit être le moins onéreux possible [DODET '90].

Cependant, il n'existe pas de système d'expression universel utilisé, chacun possédant un certain nombre de qualités et d'inconvénients. La sélection d'un type cellulaire plutôt que d'un autre pour l'expression d'une protéine hétérologue dépendra de plusieurs critères, souvent fonction de la complexité de la protéine et de son utilisation.

Il existe de nombreux systèmes eucaryotes ou procaryotes : Bactéries (*E. coli* et *Bacillus subtilis*), levures (*S. cerevisiae* et *Pichia pastoris*), cellules d'insectes, cellules de mammifères en culture, animaux transgéniques (vache, mouton, chèvre), plantes transgéniques et ovocytes de xénope.

Nous allons présenter, dans cette partie introductive, le système baculovirus-cellules d'insectes utilisé en vue d'exprimer PA1b.

a. Le système *baculovirus-cellules d'insectes*

a.1 Le virus

Les baculovirus constituent un grand groupe de virus capable d'infecter les Arthropodes comme les lépidoptères, les diptères, les hyménoptères, les coléoptères mais aussi certains crustacés. Plus de 600 espèces d'insectes sont infectées par le baculovirus.

Son nom latin *baculum* traduit sa morphologie en petit bâtonnet. C'est un virus enveloppé de 450 nm de long et d'un diamètre moyen de 50 nm. Il appartient à la famille des *baculoviridae* qui est divisée en 2 groupes:

Les *Eubaculovirinae* avec 2 genres :

1. le genre NPV pour Nuclear Polyhedrosis Virus où les virus sont inclus dans des structures protéiques de 1 à 5 µm appelés polyèdres. Ces polyèdres contiennent plusieurs particules virales regroupées dans une membrane (MNPV) ou une seule particule par membrane (SNPV).

2. le genre GV pour Granulosis Virus où une seule nucléocapside est enfermée dans un corps d'inclusion appelé granule.

Les *Nudibaculoviridae* qui sont dépourvus de corps d'inclusion.

Un des virus probablement le plus étudié et le mieux caractérisé est celui qui infecte naturellement les larves d'*Autographa californica* appelé AcMNPV. Il se multiplie dans plus de 30 espèces de lépidoptères [VAIL '71]. Comme tous les Baculovirus, il existe dans la nature sous deux formes distinctes (Figure 19) :

Une forme incluse appelée ODV (occlusion-derived virus) qui permet la protection du virus lors de sa dissémination dans l'environnement (UV, chaleur). Ce corps d'inclusion paracristallin (polyèdre) est formé d'une protéine virale de 33 kDa appelée polyédrine (PH), elle-même protégée par une « enveloppe » constituée de la protéine pp34.

La forme libre appelée BV (budded virus) permet une infection secondaire (de cellule à cellule).

Lors du bourgeonnement des virions à la surface de la cellule, ils acquièrent une enveloppe d'origine cellulaire dans laquelle une glycoprotéine virale, la gp67 s'est installée. Cette protéine forme le péplome du virus, son rôle est essentiel pour l'infection secondaire.

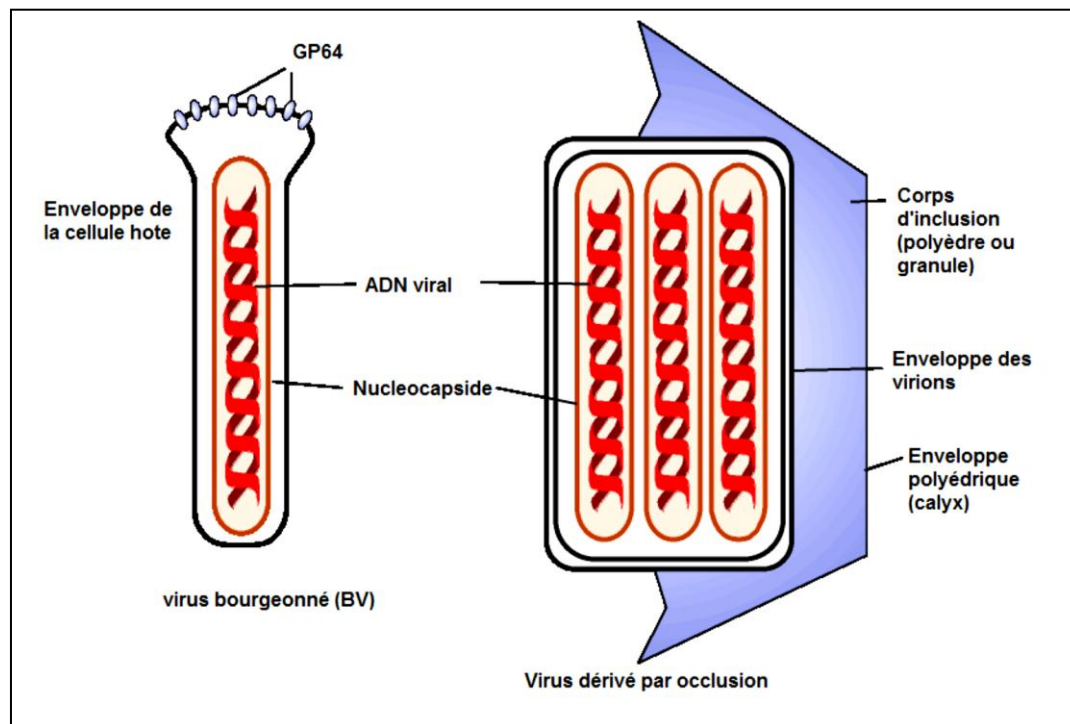


Figure 19 : Représentation des deux phénotypes du Baculovirus.

Du fait de la présence de deux types de virions, on distingue deux voies dans le cycle réplcatif (Figure 20). Dès l'ingestion d'un polyèdre, la polyédrine est dissoute par le suc intestinal très alcalin de la larve. Les virions ainsi libérés pénètrent dans la cellule intestinale grâce à la fusion de leurs membranes avec la membrane plasmique. Les nucléocapsides progressent dans le cytoplasme jusqu'au noyau où est libéré l'ADN viral. Le cycle de réplication du virus est alors amorcé, et les premiers virions apparaissent entre 15 et 17 heures post-infection (p.i). Une infection secondaire de tous les tissus de la larve est alors possible. *In vivo*, les polyèdres apparaissent dans les cellules environ 24 heures après l'infection.

Le génome d'AcMNPV est constitué d'une molécule d'ADN circulaire double brin de 133 894 pb [AYRES '94]. Une orientation a été proposée par Vlák en 1982 et permet de se repérer sur cet ADN [VLAK '82].

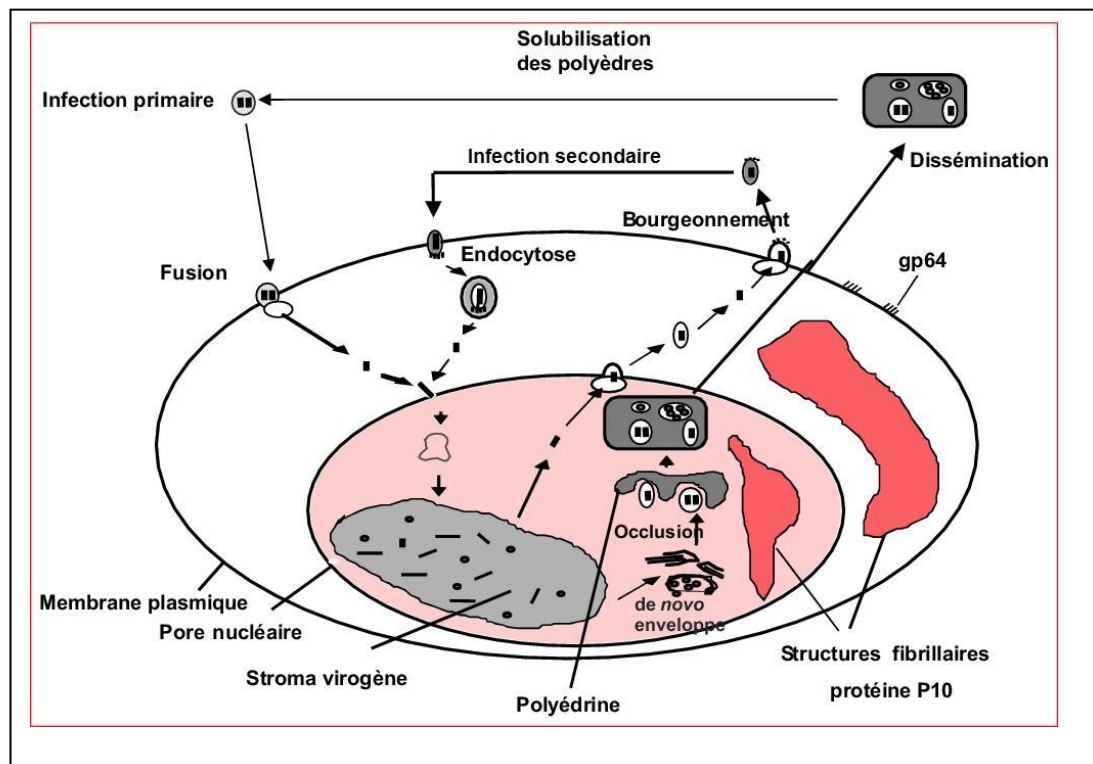


Figure 20 : Cycle réplcatif du Baculovirus [DEVAUCHELLE '04]

L'expression des gènes viraux peut être divisée en 4 phases : α , β , γ et δ (Figure 21). La phase α , phase très précoce où les gènes sont transcrits dès le début de l'infection et jusqu'à 4 heures. Le produit de certains gènes α permet l'activation de la phase précoce β entre 5 et 8 heures. Ces deux phases sont antérieures à la réplcation de l'ADN qui a lieu à partir de 8 heures p.i. Les gènes exprimés en phase tardive γ , entre 8 et 18 heures p.i. codent la plupart des protéines de structure.

Alors que la transcription des gènes des phases α et β est sous le contrôle de l'ARN polymérase cellulaire de type II (sensible à l' α amanitine), la transcription des gènes γ et δ est initiée à partir d'une séquence $^A/_T\text{TAAG}^T/_A\text{A}^T/_A$ (Boîte de Rohrmann) [ROHRMANN '86, RANKIN '88] contrôlée par une ARN polymérase, α amanitine résistante, viro-codée. Deux gènes δ sont surexprimés très tardivement. Il s'agit des gènes codant la polyédrine et la protéine P10. Alors que le rôle de la polyédrine est bien connu, celui de P10 demeure incertain mais elle contribuerait probablement à la lyse cellulaire [DEVAUCHELLE '03].

a.2 Les cellules

Les premières publications concernant l'établissement de lignées cellulaires d'insectes stables datent des années 70. Outre leur intérêt pour la virologie, elles ont surtout suscité un vif intérêt dans les études d'ontogenèse tel que la différenciation cellulaire et la régulation hormonale.

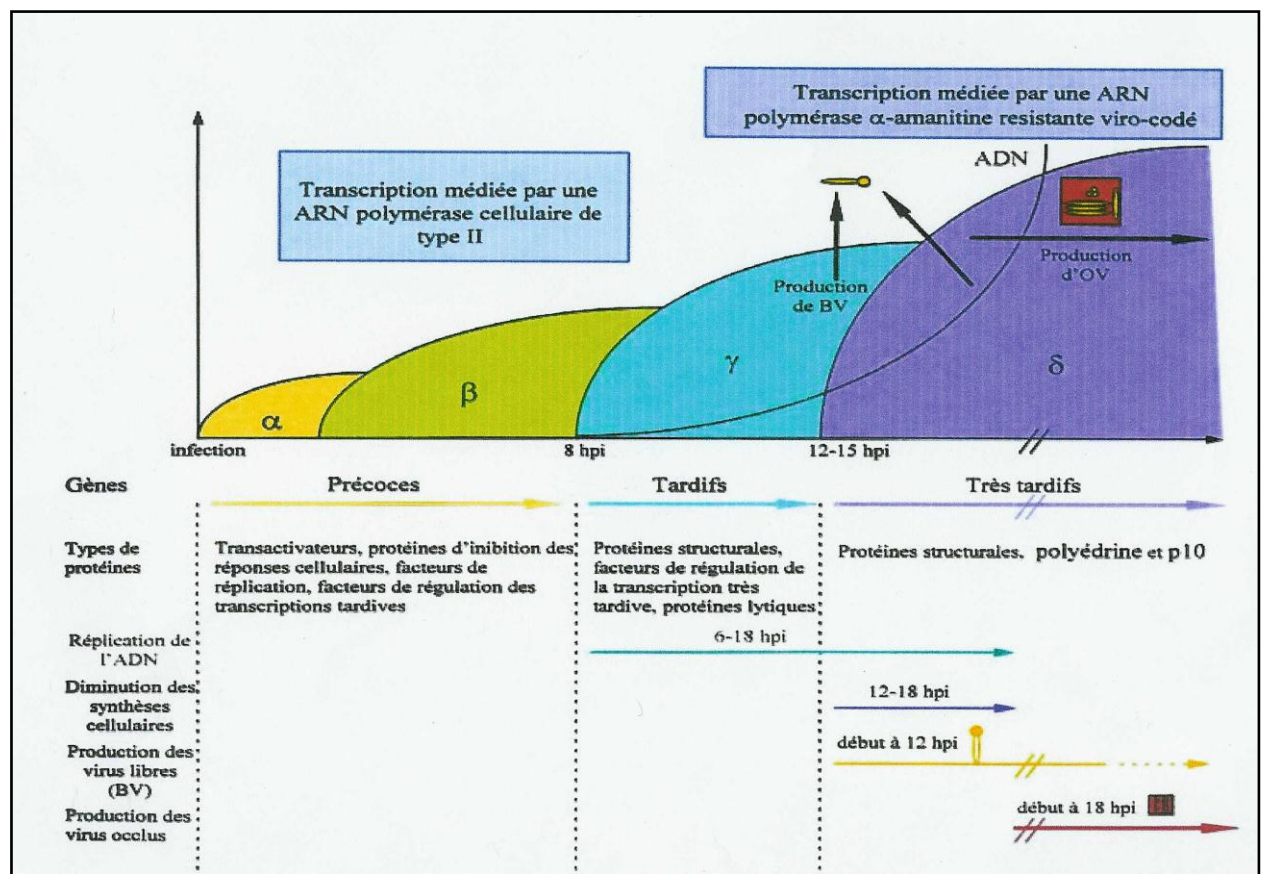


Figure 21 : Diagramme relatif aux principaux événements transcriptionnels et réplcatifs du cycle viral. Ce schéma a pour ambition de donner un certain nombre de repères en ce qui concerne les événements de transcription et réplcation virale sans tenir compte de l'aspect quantitatif [KING '92, ROHRMANN '92].

Bon nombre de lignées sont disponibles pour l'étude de la multiplication du baculovirus. La plus utilisée à l'heure actuelle est la lignée Sf9 (ATCC, CRL-1711), dérivée du clone cellulaire IPLB-Sf21AE [VAUGHN

'77] issu du tissu ovarien du lépidoptère *Spodoptera frugiperda*. La lignée BTI-TN5B14 (Invitrogen) dérivée de *Trichoplusia ni* et la lignée SPCMb-92-C6 (Station pathologie comparée, St Christol lez Ales) dérivée de *Mamestra brassicae* sont aussi très utilisées [DEVAUCHELLE '93]. Plus récemment, la lignée Ea4 (Novagen) dérivée d'*Estigmena acrea*, appartenant à l'ordre des lépidoptères et à la famille des Arctiidae, semble prometteuse car contrairement aux autres elle aurait une capacité de glycosylation plus complète.

Les cellules ont une température optimale de croissance de 28°C. Les milieux de culture sont simples et donc d'un coût faible, d'autant plus que l'on peut utiliser pour leur croissance des milieux dépourvus de tout élément d'origine animale, notamment le sérum de veau fœtal, source potentielle de contamination.

a.3 Le système d'expression

Parmi les systèmes d'expression eucaryote, le système Baculovirus-cellules d'insectes est un des plus performants pour la production de protéines complexes. Sa mise en œuvre est simple, rapide et peu coûteuse, le taux de production est élevé. Il présente aussi l'avantage d'être inoffensif puisque :

- Le baculovirus ne peut se multiplier en cellules de vertébrés,
- Aucun virus de vertébrés ne peut se multiplier sur les lignées de Lépidoptères utilisées,

- A ce jour, aucun prion pathogène n'a été décelé chez les insectes. Son seul inconvénient (ou avantage) est d'être un système lytique.

Ce système est basé sur l'utilisation du baculovirus comme vecteur pour l'expression en cellules d'insectes de gènes hétérologues sous contrôle de promoteurs viraux très tardifs. Le promoteur de la polyédrine, ainsi que celui de la protéine P10 sont en effet extrêmement actifs lors de la phase tardive de l'infection et ces protéines sont inutiles pour la réplication *in vitro*. Ces promoteurs sont donc utilisés pour l'expression de protéines hétérologues.

Ce système fût utilisé pour la première fois pour la production d'interféron β humain [SMITH '83]. Depuis, bon nombre de gènes ont été exprimés dans ce système, le plus souvent sous contrôle du promoteur polyédrine.

La Figure 22 présente les différentes étapes nécessaires à l'obtention d'un virus recombinant.

La première étape consiste en l'identification d'un gène viral cible. Celui-ci doit être non essentiel à la réplication virale *in vitro*. Une large région du chromosome viral incluant ce gène cible (1000 à 2000pb de part et d'autre du gène cible) est alors clonée dans le plasmide bactérien pUC. La seconde étape consiste en l'insertion du gène étranger. Dans l'exemple ci-dessous, le gène étranger est placé en aval du promoteur PH (vecteur de transfert chargé). Puis, les cellules sont cotransfectées avec de l'ADN du virus dit « sauvage » infectieux et de l'ADN du vecteur de transfert chargé (étape 3). Il peut alors se produire une recombinaison entre les régions homologues des deux ADN. Ceci conduit à l'obtention d'un virus recombinant présentant le gène étranger à la place du gène PH (étape 4). L'utilisation de la polyédrine comme marqueur est bien commode. Son phénotype particulier, la présence de polyèdres dans les cellules infectées, est très facile à identifier au microscope photonique et permet une sélection rapide des clones viraux recombinants.

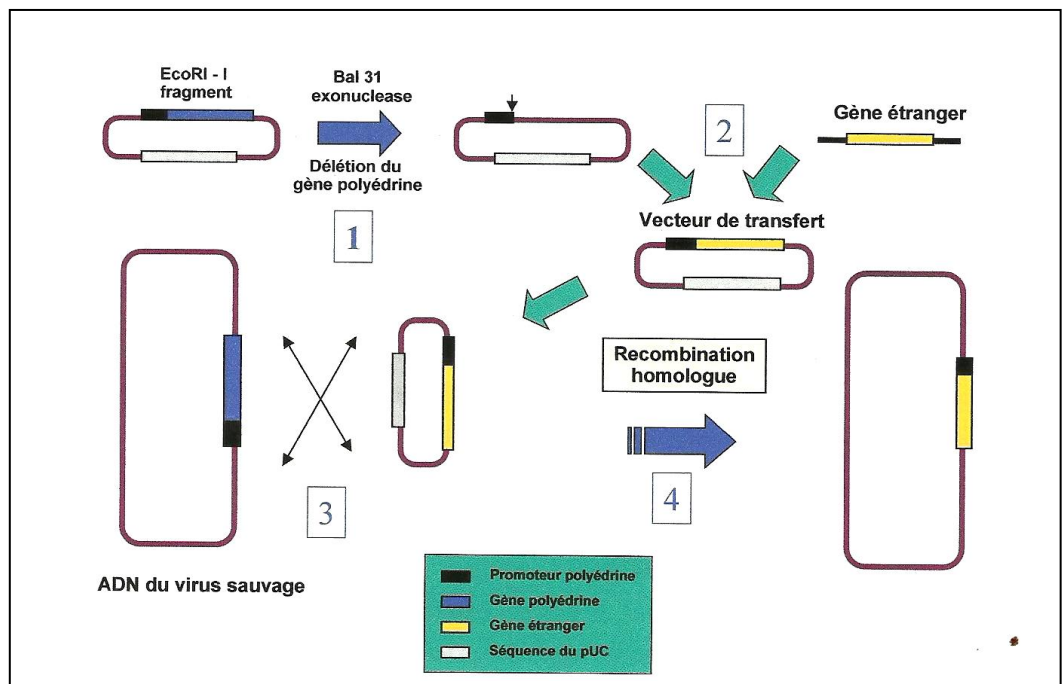


Figure 22 : Construction d'un baculovirus recombinant. Dans cet exemple, le gène cible est celui de la polyédrine.

- (1) Clonage du gène cible et de sa région puis délétion de la séquence codante.
- (2) insertion d'un gène étranger en aval du promoteur. Obtention du vecteur de transfert chargé.
- (3) Cotransfection de cellules Sf9 avec l'ADN viral et l'ADN du vecteur de transfert.
- (4) Obtention de virus recombinants.

b. Etat de l'art de la production de PA1b en système d'expression hétérologue

Quel que soit le système de production de toxine, celui-ci (ou ceux-ci) doit répondre à deux besoins : 1) produire en quantités suffisantes la toxine originale de pois pour la réalisation des différentes recherches aussi bien dans le laboratoire BF2I que dans les laboratoires partenaires du projet, et 2) également permettre la synthèse, en quantité moins importantes, d'isoformes et/ou de mutants purs de PA1.

A l'heure actuelle, la production en grosse quantité du peptide naturel passe par sa purification (extraction par des solvants, puis par HPLC) à partir des graines de pois. Ce travail est très demandeur en temps et en matériel. Les rendements sont de l'ordre de 600 mg de peptide pur par kg de graines.

La limitation de ce système de production est également que le pois produit au moins six isoformes, dont la séparation est très compliquée en HPLC, voire souvent impossible. Si un mélange d'isoforme convient en général à la plupart des tests biologiques (les plus demandeurs en quantité), les projets d'optimisation de la toxine ainsi que les projets de production des homologues de PA1b demandent de pouvoir produire des isoformes pures.

Ces synthèses se faisaient uniquement par voie chimique avec repliement de la structure tertiaire *in vitro* [DA SILVA '09]. Cette technique est relativement longue, coûteuse et les rendements sont parfois faibles. Afin d'optimiser la production de PA1b, en terme de coûts et de temps, et pour répondre à l'ensemble des besoins de l'équipe « Entomotoxines », il était de plus en plus nécessaire de pouvoir disposer d'un système de production hétérologue de la toxine. Des essais de production avaient été réalisés dans le passé dans *E.coli* (INRA Nantes et Lyon - [SANDRINE LOUIS '04a]) et dans *Pichia pastoris* (CNRS Lille), sans succès dans les deux cas. Le gène de PA1b était faiblement exprimé, et aucune activité toxique n'avait pu être détectée à cause d'un repliement incorrect du peptide. A l'heure actuelle le seul système efficace d'expression du peptide mature est le riz. Cette production a été réalisée par Julie Petit à CIRAD de Montpellier [JULIE PETIT '06]. Des lignées transformées ont été obtenues, et il a été montré que les grains de riz de ces plantes contiennent un PA1b natif

(analyse en spectrométrie de masse, par l'affinité au site d'action et par immuno-localisation). Le riz fut ainsi le premier système d'expression hétérologue à se révéler adéquat aussi bien pour la production que pour le repliement correct du peptide. Les plantes de 3^{ème} génération avaient produit des grains capables de provoquer une mortalité importante lors de tests biologiques sur charançons, mais pas totale. Cependant, à partir de la 5^{ème} génération, une baisse très nette de l'activité insecticide du grain est observée, qui n'est plus capable de tuer les insectes adultes. D'autre part, la production de PA1b n'a pas été envisagée dans les cellules de mammifères ainsi que par transgénèse animale essentiellement pour la lourdeur de leur manipulation. Après ces premiers essais de production de PA1b, Vanessa Eyraud (Doctorante au laboratoire BF2I) a entrepris l'expression transitoire de la toxine dans le tabac. Ces essais se sont avérés fructueux permettant de produire la toxine bien repliée et conservant sa toxicité [EYRAUD '13].

En parallèle, nous avons enclenché la production de PA1b dans le système de production hétérologue : baculovirus/ cellules d'insectes Sf9. Les cellules d'insectes (couplé au baculovirus) ont été choisies puisqu'elles sont *a priori* adaptées à la production de peptides soufrés et capable de maturer les peptides végétaux et de les produire en grande quantité [VETSCH '00]. Les protéines exprimées dans les cellules d'insectes présentent, dans la grande majorité des cas, des activités biologiques semblables à celles des protéines natives. De plus, les cellules d'insectes possèdent les équipements enzymatiques leur permettant d'effectuer les modifications post-traductionnelles semblables à celles des Vertébrés. De plus, ce système présente un avantage par rapport aux autres systèmes, puisqu'il possède la possibilité d'être manipulé dans un espace confiné de plus en adéquation avec les attentes industrielles. Le problème qui peut se poser avec le choix de ce système, c'est la possibilité de produire une toxine d'insecte dans une cellule d'insecte. Mais nous avons abordé ce problème en produisant notre toxine sous le contrôle d'un promoteur hypertardif. De plus ce choix risqué de système de production a été fait en parallèle avec celui du tabac pour lequel on ne savait pas initialement son efficacité pour produire la toxine PA1b.

2. Matériels et Méthodes

Le but de ce travail est d'obtenir des baculovirus recombinants exprimant le gène étranger *PA1*, qui code pour le peptide entomotoxique PA1b. Nous réaliserons la transfection puis la sélection des baculovirus recombinants. D'autre part, les cellules *Sf9* seront infectées par le baculovirus recombinant *AcMNPV-PA1* déjà sélectionné afin de vérifier la production de la protéine recombinante PA1b.

Afin de contrôler toutes les étapes du système, les manipulations suivantes vont être réalisées:

Infection des cellules *Sf9* par un baculovirus *AcMNPV* sauvage produisant des inclusions de polyédrine (Témoin).

Co-transfection des cellules *Sf9* avec l'ADN d'*AcMNPV* sauvage muté et du vecteur de transfert contenant le gène *PA1*.

Sélection des baculovirus recombinants par la méthode des plages de lyse.

Production de la protéine recombinante PA1b en cellules *Sf9* par infection des cellules avec le baculovirus recombinant *AcMNPV-PA1*.

a. Matériels

a.1 Plasmides

pGEM-T easy avec CDS AJ574795 en antisens : plasmide contenant la séquence nucléotidique CDS du gène homologue de *PA1* dont l'identifiant Genbank est AJ574795. Cette séquence est insérée entre deux sites de restriction KpnI et PstI (construit et fourni par Julie Petit, CIRAD Montpellier)

pGMAC116T : plasmide contenant un insert viral, cloné dans un pUC (construit et fourni par Martine Cerutti, Laboratoire baculovirus et thérapie, St-christol-les-Ales).

a.2 Souche bactérienne

Souche TOP10 d'*E.coli* (Invitrogen) : souche bactérienne chimiquement compétente avec une efficacité de transformation de 1×10^9 cfu/ μ g de plasmide d'ADN utilisé.

a.3 Cellules et virus

Les cellules Sf9 de *Spodoptera frugiperda*, obtenues à partir de l'American Type Culture Collection (ATCC) et récupérées du laboratoire de baculovirus et thérapie, St-christol-les-Alès, ont été cultivées à 27° C en monocouche en milieu Insect-Xpress (Lonza) complété avec 5% du sérum de veau fœtal (GIBCO) et 0.1% d'une solution antibiotique-gentamicine (SIGMA). La culture des cellules Sf9 a été réalisée sur boîte de Pétri de 25 et 75 cm² selon les méthodes de culture standard pour ces cellules. Le stock AcMNPV utilisé dans cette étude a été obtenu du laboratoire baculovirus et thérapie, St-christol-les-Alès.

a.4 Anticorps polyclonaux anti-PA1b

Les anticorps polyclonaux anti-PA1b sont fournis par Julie petit, CIRAD Montpellier.

b. Méthodes

b.1 Clonage moléculaire et analyse d'ADN

b.1.1 Electrophorèse de l'ADN en gel d'agarose

L'analyse des ADN cellulaires et plasmidiques et de leurs produits de digestion ou d'amplification se fait par électrophorèse en gel d'agarose

de 1% dans du tampon TAE 1x (Tris/acétate 20 mM, pH 98.0, EDTA 0.5 mM). Les gels sont ensuite colorés dans une solution de bromure d'éthidium (BET) à 500 µg/ml. Les profils de migration sont visualisés sous UV à 312 nm. La taille des fragments est estimée grâce à la co-migration dans le gel d'un marqueur de poids moléculaire, Mass DNA Ladder (Euromedex) : les fragments visualisés vont de 80 à 10000 pb.

b.1.2 Purification des fragments d'ADN

Après restriction des plasmides ou suite à une amplification par PCR, les ADN sont déposés sur gel d'agarose (1%) et migrent en fonction de leur taille. Les séquences sont excisées du gel et purifiées.

Après séparation électrophorétique, la bande d'agarose contenant le fragment d'ADN voulu est découpée puis placée dans un tube de 1.5 ml. La purification est faite à l'aide du kit Nucleospin Extract II, Macherey Nagel, selon le protocole indiqué par le fournisseur.

b.1.3 Transformation bactérienne

La transformation par choc thermique nécessite 50µl de bactéries chimiquement compétente et environ 50ng d'ADN. Le mélange est laissé 30 minutes sur la glace, puis transféré 30 secondes à 42°C puis incubé 2 minutes dans la glace. 250µl d'un milieu S.O.C (2% Tryptone, 0.5% Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) est ajouté pour une incubation de 1h à 37°C. Les bactéries sont ensuite étalées sur boîte de pétri contenant du LB-agar, de l'antibiotique (ampicilline) et du X-gal (40 mg/ml, Euromedex) pour la sélection des clones.

b.1.4 Extraction de l'ADN génomique viral des cellules infectées

Les cellules Sf9 infectées sont mises en suspension grâce à un grattoir puis centrifugées à 3000g pendant 10 minutes. L'ADN génomique total est extrait du culot à l'aide du kit Qiaprep spin miniprep (Qiagen).

b.1.5 Analyse par PCR

Les réactions de polymérisation en chaîne (PCR) servant à amplifier des fragments d'ADN destinés au clonage ou à l'analyse de séquences d'ADN, sont effectuées à l'aide de polymérases (Invitrogen), en suivant les

protocoles indiqués par le fournisseur. Les séquences des amorces utilisées pour les amplifications d'ADN sont citées dans le Tableau 7.

b.2 Analyses d'ARN

L'ARN total des cellules Sf9 infectées est extrait avec le kit Nucleospin RNA XS (Macherey Nagel) selon le protocole fourni. Pour analyser l'expression des transcrits d'ARN viraux dans les cellules infectées par les virions recombinants, une retro-transcription de l'ARN messager est réalisée avec le kit SuperScript II Reverse Transcriptase (Invitrogen) à l'aide d'amorces oligo dT. Une PCR est par la suite effectuée pour amplifier les séquences des transcrits *PA1* et *PA1b*.

b.3 Entretien des cellules d'insectes

Les cellules sont répliquées tous les 3 jours dans des flacons de culture Falcon de 25 cm². Les cellules sont mises en suspension dans 5ml de milieu en secouant les bouteilles vigoureusement et 1/10^{ème} de cette suspension sert à ensemer une boîte contenant 4,5ml de milieu neuf. Les cellules sont placées dans une étuve à 27°C.

b.4 Construction des baculovirus recombinants

(Annexe 4)

Nous avons dessiné trois amorces spécifiques (Tableau 7) qui permettent d'amplifier la séquence du CDS d'AJ574795 (en antisens) présent dans le vecteur pGEM-T easy et codant pour PA1 (peptide signal+PA1b+PP+PA1a+PP) et PA1b (peptide signal+PA1b). L'amorce R1PA1 (anti-sens) comporte un site de restriction Bgl II ainsi qu'une séquence Kozak. Cette dernière, dont la séquence est formée d'au moins six nucléotides (GCCACC), se trouve localisée juste avant le codon d'initiation ATG. Le rôle de cette séquence est d'augmenter l'efficacité de la traduction à partir de l'AUG qui la suit. L'amorce F1PA1 et F4PA1 (sens) comportent un site de restriction Kpn I et amplifie la séquence *PA1* et *PA1b* respecti-

vement. La PCR est effectuée sur 25 µL de milieu réactionnel, 10 µM des amorces R1PA1, F1PA1 et F4PA1, 50 mM de dNTPs, 50 mM de MgCl₂, le tampon *Taq* ADN polymérase et 2U de *Taq* ADN polymérase à haute-fidélité (Invitrogen). Les réactions d'amplification sont effectuées à l'aide d'un thermocycleur du type MasterCycler (Eppendorf). L'amplification est réalisée sur 35 cycles (45 sec à 95°C, 45 sec à 62°C, 1 min 30sec à 72°C). Les produits d'amplification ont ensuite été clonés dans le plasmide Topo-TA (pCR 2.1-TOPO) selon le protocole défini par « Invitrogen ».

Les plasmides sont purifiés sur colonne (Nucleospin Plasmid, Macherey-Nagel) et séquencés. Les plasmides qui contiennent les séquences correspondantes aux CDS codant PA1 et PA1b ont été digérés par les enzymes de restriction Bgl II et KpnI (10 U/µL, Roche). Les produits de digestion obtenus sont ensuite ligaturés avec le plasmide d'expression pGmAc116T préalablement digéré par BglII et KpnI et transformés les *E. coli* TOP10 [FREDERICK M. AUSUBEL '03]. De même, Les plasmides sont purifiés sur colonne (Nucleospin Plasmid, Macherey-Nagel) et séquencés (couple d'amorces utilisé : pGmAc116TF2- pGmAc116TR2).

b.5 Production des peptides recombinants

Les plasmides recombinants ont été par la suite utilisés pour co-transfecter des cellules d'expression Sf9 en même temps que l'ADN viral dit « sauvage muté » ou BacMid I par lipofection. Les liposomes sont obtenus en faisant interagir (i) du DOTAP (Roche) (ii) de l'ADN de virus sauvage muté- BacMid I (iii) de l'ADN du vecteur de transfert baculovirus « chargé » avec le gène étranger à exprimer. Les ADN viraux et plasmidiques ont été préalablement stérilisés une nuit minimum dans l'alcool à 70%. Les cellules Sf9 (2x10⁶) ont étéensemencées dans des boîtes de pétri de 25 cm² et sont incubées en présence du mélange DOTAP-ADN pendant 4 heures. Le milieu de transfection a été ensuite remplacé par du milieu de culture. Nous avons poursuivi l'incubation pendant 4 jours au cours desquels apparaissent les signes d'infection.

Il faut signaler que nous avons utilisé un système dérivé du système commercial, système mis au point par Martine Cerutti du laboratoire Baculovirus et Thérapie à St-christol-les-Alès et qui contrairement au système commercial engendre lors de l'infection la formation de polyèdres. Comme mentionné dans l'introduction paragraphe a.3, l'utilisation de la polyédrine comme marqueur est bien commode. Son phénotype particulier

dans les cellules infectées, facilite l'identification au microscope photomicroscopique et permet une sélection rapide des clones viraux recombinants.

Des baculovirus sauvages dit « mutés » développés à St-christol sont présentés sous forme d'ADN viral linéarisé avec une délétion létale (BacMid I) facilitant ainsi la sélection de recombinants. L'ADN du baculovirus AcMNPV a été modifié de telle sorte qu'il ne code plus pour un virus viable ; la co-transfection entre ce virus muté et l'ADN de transfert restitue la délétion létale par recombinaison homologue. De fait, seuls les baculovirus recombinés produisent des virions viables, cette fréquence de recombinaison excédant 99% (par rapport à une fréquence de 0,5 à 1,5 % avec un ADN sauvage). De plus, cet ADN viral muté dérive du baculovirus sauvage AcMNPV par délétion du locus polyédrine et réinsertion de la séquence codant la polyédrine dans le locus du gène p10. Il possède ainsi un phénotype avec formation de polyèdres.

La sélection des baculovirus recombinants s'effectue par la technique des plages de lyse en culture cellulaire en colorant les cellules au rouge neutre afin de rendre les plages visibles. Les virions recombinés présentent des plages avec des polyèdres (OB+ : Occlusion Body +) dus à la présence de la polyédrine. Ces plages de virions recombinés sont alors récupérées, et amplifiées.

Afin de s'assurer que les virions recombinants choisis ont bien intégré les plasmides recombinants *PA1* et *PA1b* résultant de la cotransfection, deux amorces spécifiques For116 et Bac116 ont été dessinées (Tableau I) permettant d'amplifier les CDS *PA1* et *PA1b*. La PCR est effectuée dans les mêmes conditions citées auparavant, additionné des ADN génomiques obtenus suite à l'extraction de l'ADN viral (voire paragraphe 2.2.1.4). De même, afin de contrôler la transcription des ADN viraux en ARN, une PCR est effectuée sur 25 µL de milieu réactionnel, additionné des ADNc obtenus de la transcription reverse d'ARN messagers d'extraits cellulaires infectés par les virions recombinants (voire paragraphe b.2). Pour réaliser cette PCR, nous avons utilisé deux couples d'amorces : Bac116-F1PA1 et bac116-F4PA1 (Tableau 7).

Suite à la détermination des virions recombinants et la réalisation de trois amplifications successives des clones choisis, nous avons déterminé le titre viral des virions recombinant PA1 et PA1b par la technique des plages de lyse. Les plages ont été observées au microscope inversé après coloration au rouge neutre (Annexe 5). Une fois le titre est déterminé, la production de la protéine est réalisée dans des flacons de culture Falcon de 25 ou 75 cm².

Tableau 7 : Composition des amorces utilisées en PCR.

Nom de l'amorce	Séquence nucléotidique des amorces
F1PA1	5'-GGGAATTCGATTGGTACCTTAA-3'
F4PA1	5'-ATTGGTACCTTATCCAGATGGATTTCTGCAGTAA-3'
R1PA1	5'-CAAGATCTGCCACCATGGCTTCCGTTAAACTCGCTT-3'
For116	5'-GCCATTGTAATGAGACGCAC-3'
Bac116	5'- CTCCTTTACTTTGAACACCAG-3'
pGmAc116TF2	5'-TGAGACGCACAACTAATATCAC-3'
pGmAc116TR2	5'-GTGCGTCTGGTGCAAAC-3'

b.6 Purification des peptides recombinants

Au 4ème jour p.i, les cellules *Sf9* infectées par les virions recombinants ont été mises en suspension grâce à un grattoir. Le tout a été centrifugé à 3000g pendant 10 minutes. Le culot cellulaire ainsi que le surnageant ont été séparé dans deux tubes différents et étudié séparément. Chacun de ces derniers a été mélangé avec de l'éthanol 60 en conservant une agitation pendant 2 heures. Cette étape a été suivie par une centrifugation de 20 min à 10000g à 4°C. Le surnageant de chaque tube a été ensuite étudié par chromatographie liquide à haute pression (HPLC) et injecté sur une colonne C18 préalablement équilibrée à 20% de tampon B (Acetonitrile-TFA 0.8%) – 80% de tampon A (TFA 0.1%). Les échantillons prélevés en sortie de la purification ont été analysés par électrophorèse, western Blot ainsi que par spectrométrie de masse.

b.7 Western-Blot

Les protéines ont été séparées par électrophorèse en gel dénaturant (NuPAGE 10% Bis-Tris Gel) et électrotransférées sur une membrane de nitrocellulose (Novex iBlot Gel Transfer stacks nitrocellulose, regular) en utilisant le système iBlot® 7-minute Blotting (Invitrogen) et en poursuivant le protocole fourni. Le marqueur de taille utilisé est le Novex Sharp Pre-Stained protein standard (INVITROGEN) dont la gamme de taille varie entre 3.5 et 260 kDa. La mise en contact de la membrane avec le premier anticorps s'effectue au 1/500ème dans du tampon TBS, pendant 1 heure. L'anticorps secondaire, couplé à la phosphatase alcaline, a été additionné

au 1/10000ème dans du TBS durant 1h sous agitation. Les membranes ont ensuite été lavées. Les protéines marquées ont alors été révélées par incubation dans un tampon salin (0,1 M Tris ; 0,1 M NaCl ; 0,05 M MgCl₂) en présence de BCIP (5- Bromo-4-Chloro-IndolylPhosphate, substrat de l'activité enzymatique de la phosphatase) et de NBT (NitroBlue Tetrazolium) à l'obscurité pendant 10 min. La réaction a été arrêtée par immersion dans de l'eau distillée.

b.8 Caractérisation des protéines par spectrométrie de masse

Après purification, les masses et la pureté des protéines ont été caractérisées par spectrométrie de masse couplant une source d'ionisation laser assistée par une matrice (MALDI, *Matrix-Assisted Laser Desorption/Ionisation*) et un analyseur à temps de vol (TOF, *time-of-flight mass spectrometry*) (Autoflex, Bruker Daltonics). Les spectres sont enregistrés entre 0 et 20000 M/z.

b.9 Caractérisation de l'activité des protéines.

Les échantillons de masses attendues ont été testés pour leur activité biologique sur les cellules Sf9, en incorporant la protéine dans le milieu de culture des cellules (15000 cellules/puit)ensemencées dans une plaque à 96-puits et en réalisant une cinétique de viabilité cellulaire pendant 5 heures. Le peptide PA1b extrait à partir des graines de pois a été utilisé en tant que molécule toxique de référence. Pour réaliser ces tests de viabilité cellulaire, un réactif fluorométrique CellTiter-Blue® est utilisé (Promega). Le dosage est basé sur la capacité des cellules vivantes à convertir un colorant redox (résazurine) en un produit final fluorescent (résorufine). Les cellules viables conservent la capacité de réduire la résazurine en résorufine. Les cellules non viables perdent rapidement leur capacité métabolique, ne réduisent pas le colorant indicateur, et donc ne génèrent pas un signal fluorescent (annexe 6). La valeur de la dose létale (LC₅₀) a été calculée par le biais du programme jmp. Cette valeur représentera la dose de PA1b causant la mort de 50 % des cellules Sf9 [JMP '89-'].

3. Résultats

Ces travaux ont été possibles grâce à une collaboration avec le Dr M. Cerutti (UPS3044 CNRS de Saint Christol Lez Alès).

a. Insertion de l'ADN complémentaire de PA1 et PA1b dans le baculovirus

La séquence de l'ADNc codant le peptide PA1 correspond à la séquence génique AJ574795 dans laquelle la séquence de l'intron a été supprimée (La séquence étant présente dans le vecteur pGEM-T easy (Figure 23)). Cette séquence codante est composée de 393 paires de bases (Figure 24, A), à partir du codon d'initiation de la transcription ATG jusqu'au codon d'arrêt de la traduction. La protéine correspondante est longue de 130 résidus (Figure 24, B). Dans cette séquence d'acides aminés, un peptide signal est composé de 26 acides aminés (Met1-Ala26). Deux peptides matures sont présents : PA1b est composé de 37 acides aminés et PA1a de 53 acides aminés. Nous avons focalisé notre attention sur le peptide PA1b, PA1a étant encore de fonction inconnue. PA1b débute par une Ala27 et se termine par une Gly63. Ce peptide est riche en résidus proline, 5 résidus qui représentent à eux seuls 13.5% du total de la séquence aminée. Il est également riche en résidus cystéine, 6 résidus qui représentent 16.21 % de la séquence, et qui sont répartis dans toute la séquence de la molécule (Figure 24, B).

Les deux couples d'amorces F1PA1-R1PA1 et F4PA1-R1PA1 ont été utilisés pour amplifier les séquences de PA1 et de PA1b (voire localisation des amorces Figure 24). Le résultat de ces deux amplifications a été inséré dans 2 vecteurs de transfert pGmAc116T en aval d'un promoteur viral tardif fort par l'intermédiaire d'un premier vecteur de transfert le TOPO TA.

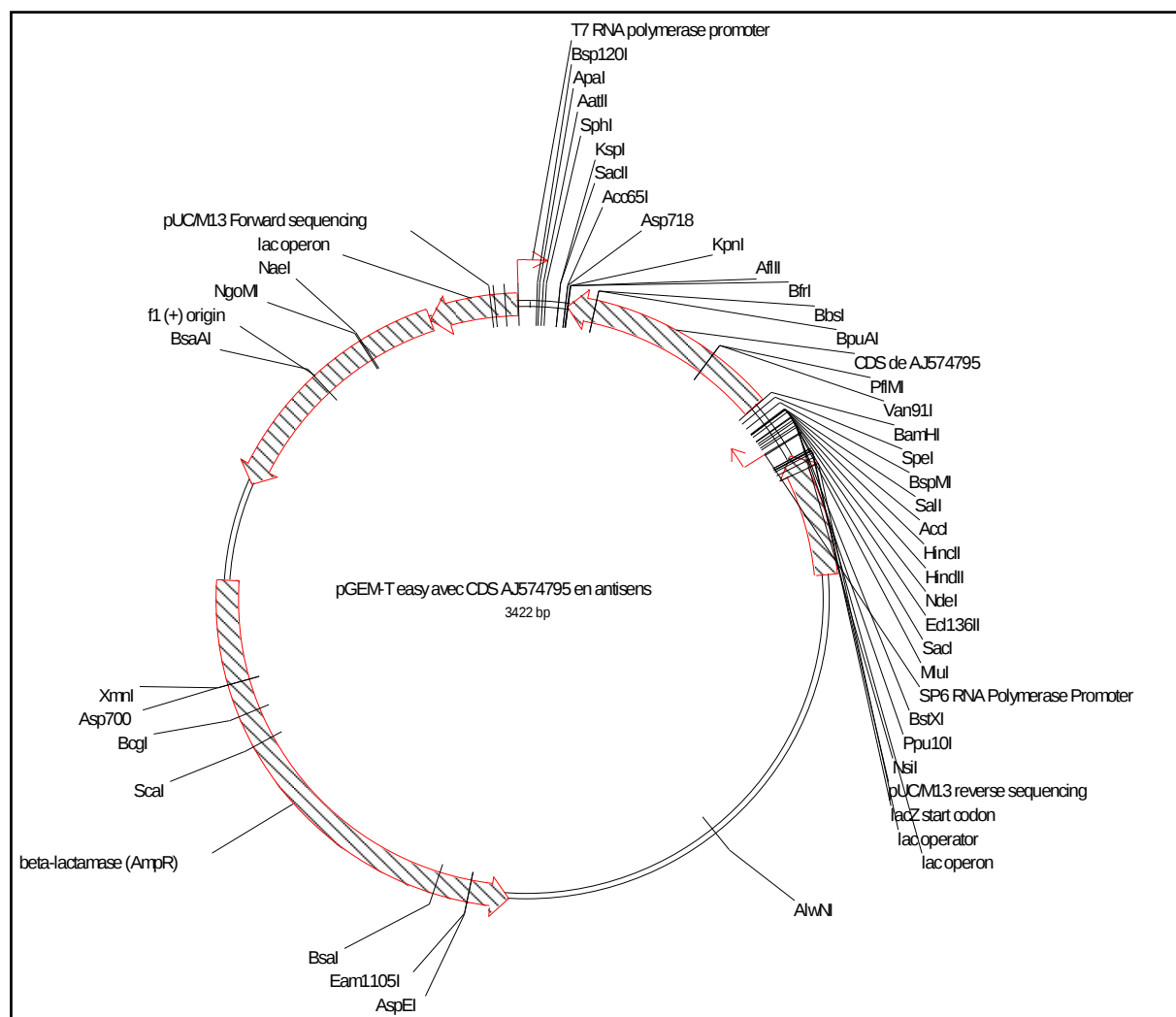
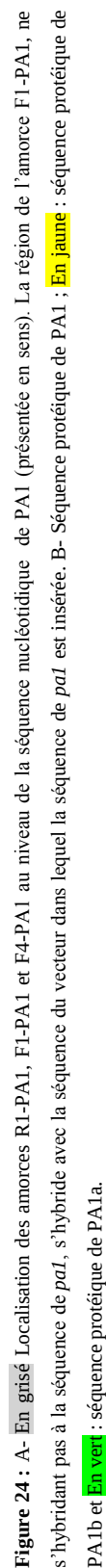


Figure 23 : Carte du vecteur pGEM-T easy avec CDS AJ574795 en antisens (T7 forward, SP6 reverse et CDS en anti-sens -> bordé par deux sites : KpnI et BamHI).



Des virus recombinants ont été obtenus après recombinaison homologue entre le vecteur de transfert chargé décrit ci-dessus et un ADN viral défectif non infectieux le BacMid I (Annexe 4). Cet ADN viral présente en effet une délétion dans un gène essentiel et ne peut donc pas générer de virus. Lors de la recombinaison homologue, l'ADNc de *PA1* et *PA1b* est intégré dans l'ADN viral et le gène essentiel réparé, ce qui conduit à la production de particules virales infectieuses. Un gène rapporteur, le gène codant la protéine polyédrine, a été inséré dans le BacMid I pour vérifier la présence de virus recombinants. En effet, au cours de la réplication des virus recombinants des polyèdres sont formés dans le noyau des cellules infectées. Ces polyèdres sont très faciles à identifier en microscopie photonique à faible grossissement (Figure 25). En parallèle, nous avons amplifié du virus sauvage wild type (wt) AcMNPV afin de l'utiliser en tant que témoin. Le stock viral a été titré par la méthode des plages de lyse et le résultat est présenté dans le Tableau 8.

Tableau 8 : Titrage des stocks viraux.

Les plages ont été observées au microscope inversé après coloration au rouge neutre. L'unité de mesure étant la pfu/ml ; pfu pour particule forming unit.

Virions	Titre viral (pfu/ml)
PA1	2×10^7
PA1b	4×10^6
AcMNPV – virus sauvage wt	1×10^7

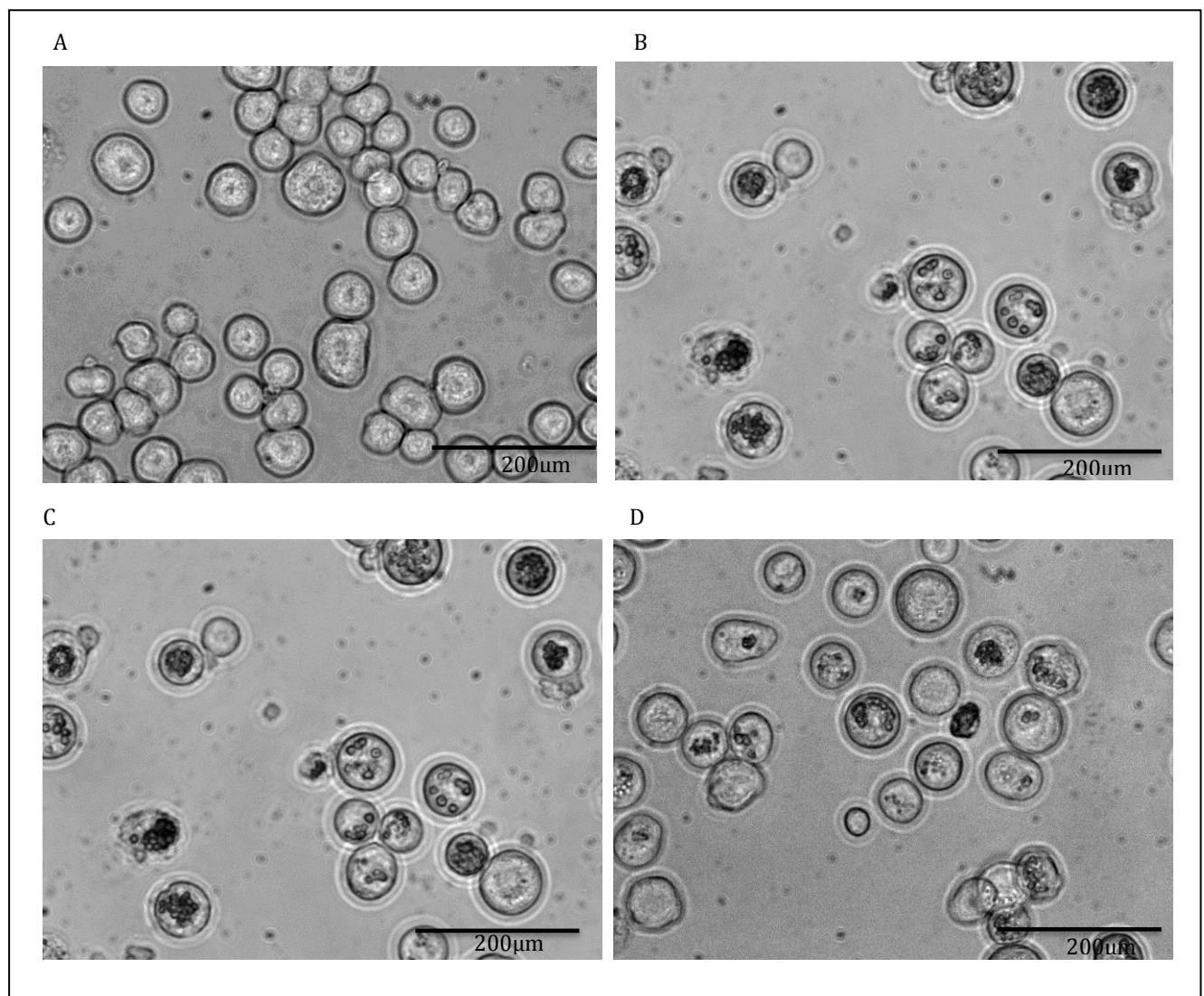


Figure 25 : Identification des polyèdres en microscopie photonique à faible grossissement suite à l'infection par les virions recombinants.

A- Image des cellules *Sf9* saines après 72 heures (Témoin).

B- Image des cellules *Sf9* infectées par le virus sauvage AcMNPV wt après 72 heures post-infection (Témoin).

C- Image des cellules *Sf9* infectées par les virions recombinants contenant l'ADN complémentaire de pa1 après 72 heures post-infection.

D- Image des cellules *Sf9* infectées par les virions recombinants contenant l'ADN complémentaire de pa1b après 72 heures post-infection.

b. Cinétique d'expression des ARN messagers

Pour analyser l'expression des transcrits d'ARN viraux dans les cellules infectées par les virions recombinants, une retro-transcription de l'ARN messager est réalisée suivie d'une PCR. Les résultats montrent bien l'expression des ARNs *PA1* et *PA1b* à partir de 48 heures post-infection avec une révélation des bandes observées à une taille de 399 et 200 pb, respectivement. D'autre part et logiquement, les cellules *Sf9* non infectées (saines) et les cellules infectées par le virus sauvage *AcMNPV* ne présentent aucune trace d'expression d'ARN homologues à *PA1* ou *PA1b* (Figure 26).

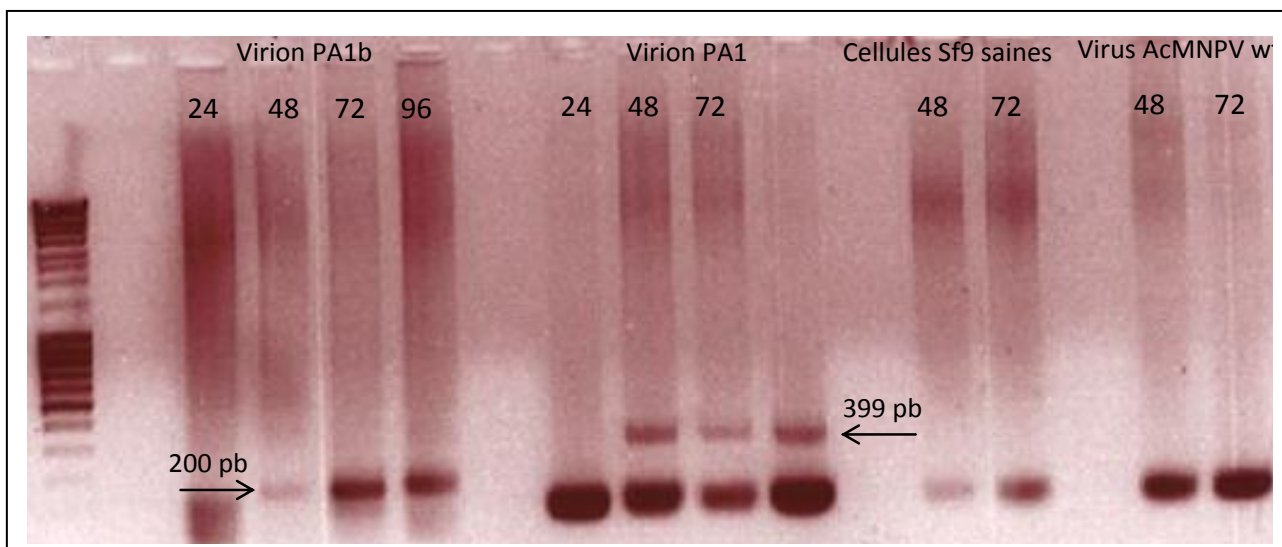


Figure 26 : Cinétique d'expression des ARNs dans les cellules *Sf9* infectées par les virions recombinants *PA1b* et *PA1*.

Les cellules *Sf9* non infectées (saines) et les cellules *Sf9* infectées par le virus sauvage *AcMNPV* (wt) représentent les contrôles négatifs de cette manipulation. M : marqueur de poids moléculaire allant de 100 à 10 000 pb.

c. Caractérisation du peptide recombinant

La purification du peptide a été réalisée à partir du milieu de culture et des cellules *Sf9* infectées par les virions recombinants. L'utilisation de deux vecteurs d'expression PA1 et PA1b a été réalisée dans le but de déterminer le rôle de PA1a dans la production de PA1b.

Nous avons envisagé une extraction à l'éthanol 60% de la protéine produite, PA1b étant soluble dans l'éthanol tout en conservant sa structure tridimensionnelle. La fraction soluble à l'éthanol 60% a été injectée en HPLC. Les pics observés dans la zone de rétention de PA1b natif ont été analysés en spectrométrie de masse, en électrophorèse et western blot, et leur activité biologique a été évaluée par des tests de viabilité cellulaire.

c.1 HPLC

Les peptides, provenant de la 4ème amplification virale et extraits à l'éthanol 60% dans le milieu de culture des cellules infectées et dans les cellules *Sf9*, ont été analysés par chromatographie liquide à haute performance (HPLC). Les profils obtenus sont représentés sur Figure 27- Figure 29.

a. Les deux pics supplémentaires (A et B) détectés dans le cas des cellules *Sf9* infectées par le virion recombinant PA1 et PA1b (Temps de rétention est de 20.35 minutes et 20.42 minutes respectivement) (Figure 27) ont le même temps de rétention qu'un standard PA1b extrait à partir de la farine du pois (Figure 28).

b. Un pic supplémentaire (C) a été détecté dans le cas du milieu de culture des cellules *Sf9* infectées par le virion recombinant PA1 (Temps de rétention est de 18.11 minutes) (Figure 29).

L'aire des pics a été quantifiée et les résultats montrent la production de 1.9 µg de PA1b en utilisant le virion recombinant PA1 (12×10^6 cellules *Sf9* dans un flacon de 75 cm²) alors que l'utilisation du virion recombinant PA1b, dans les mêmes conditions, permet la production de 94ng de PA1b.

D'autre part, l'aire du pic C a été quantifiée et le résultat montre la production de 14 µg de la proprotéine PA1 (forme non clivée : du PA1b au propeptide de la protéine PA1a) dans 15 ml de milieu de culture des cellules infectées par le virion recombinant PA1.

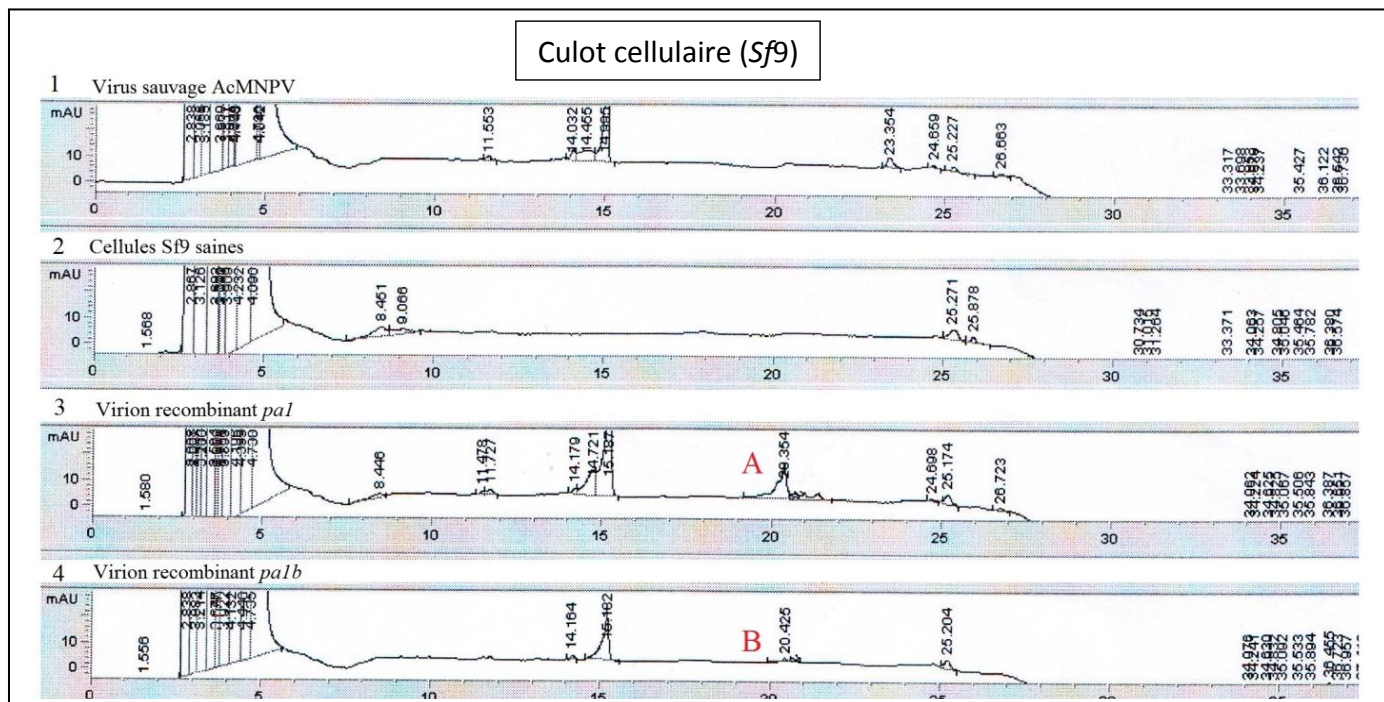


Figure 27 : Comparaison des chromatogrammes d'HPLC des peptides extraits à partir des culots cellulaires

- (1) infectées par le virion sauvage AcMNPV
- (2) non-infectées (saines)
- (3) infectées par le virion recombinant *pa1*
- (4) infectées par le virion recombinant *pa1b*

La quantité injectée représente le 1/5ème du volume totale.

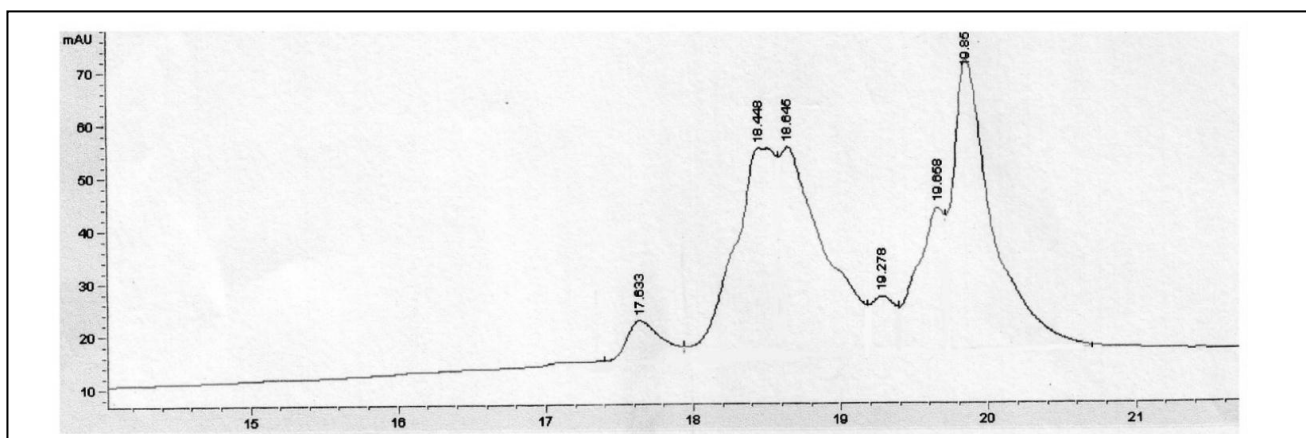


Figure 28 : Chromatogramme d'HPLC: séparation de pics d'isoformes de PA1b de la fraction EtOH60 de la farine du pois- *Pisum sativum* : variété Frisson.

résidus cystéines du fait de leur implication dans la formation de trois ponts disulfures.

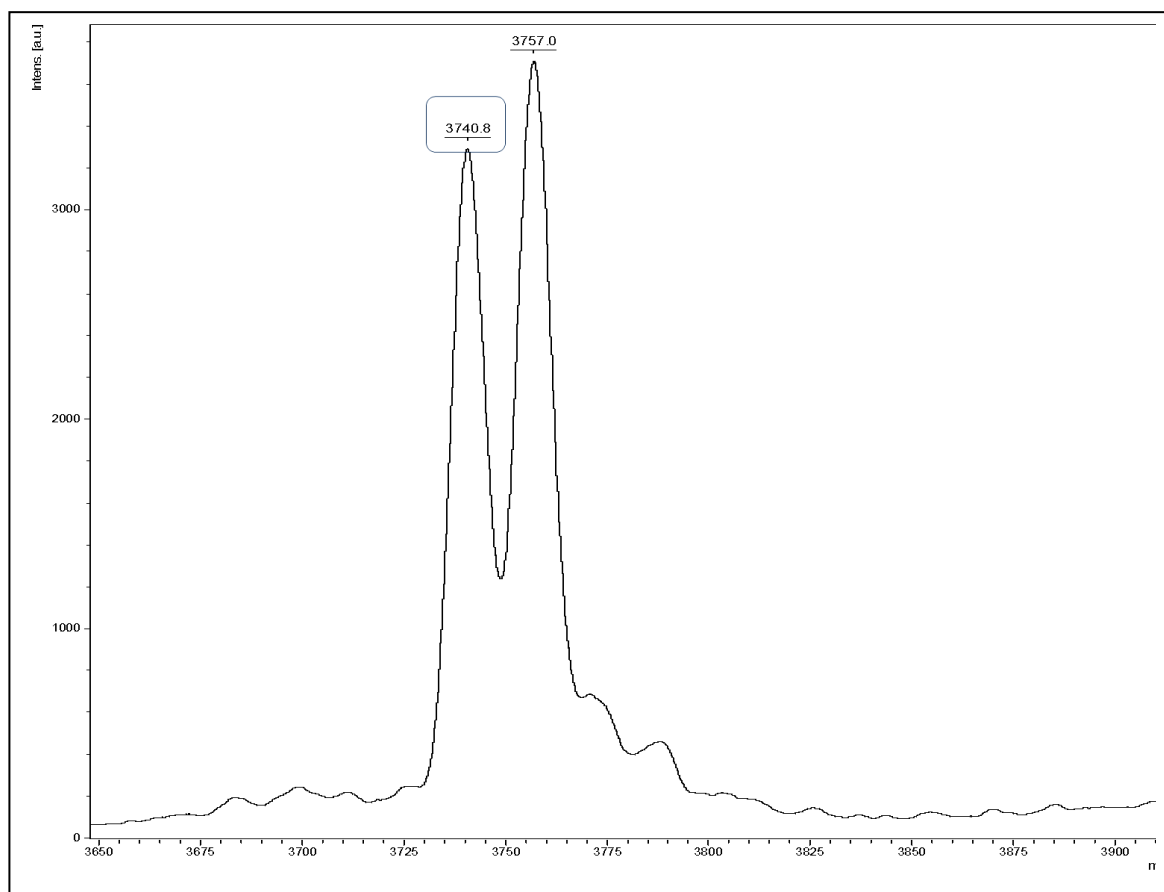


Figure 30 : Analyse spectrale (MALDI-TOF-MS) obtenue pour la forme du peptide recombinant trouvée dans les cellules *Sf9* infectées par les virions recombinants *PA1* (Résultat correspondant au pic A).

*Limite d'acceptabilité et intervalle de tolérance : 300- 1000 ppm en fonction de la taille des protéines et 50-300 ppm en fonction de la taille des peptides.

- b. L'analyse de la masse molaire du pic C a été mesurée à 11085.5* Da (Figure 31). Si l'on compare cette masse par rapport à la masse théorique du propeptide *PA1* (*PA1b*+*PP*+*PA1a*+*PP*) 11099 Daltons, on constate une perte de masse de 13.5* Daltons. Cette différence pourrait correspondre à la perte de 10 hydrogènes provenant de 10 résidus cystéines (six résidus de cystéines dans la séquence de *PA1b*

et quatre dans PA1a) du fait de leur implication dans la formation de cinq ponts disulfures.

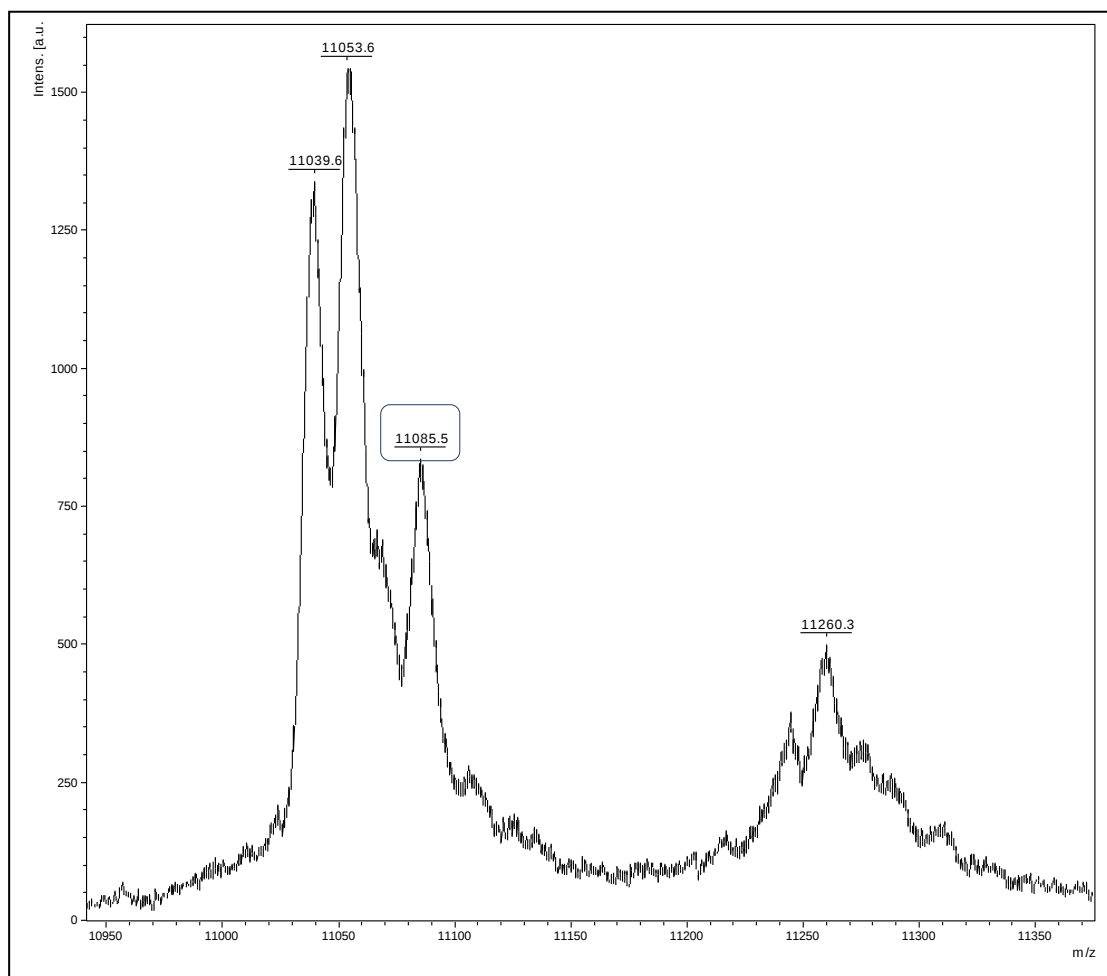


Figure 31 : Analyse spectrale (MALDI-TOF-MS) obtenue pour la forme du peptide recombinant trouvée dans le milieu de culture des cellules *Sf9* infectées par les virions recombinants *PA1* (Résultat correspondant au pic C).

*Limite d'acceptabilité et intervalle de tolérance : 300- 1000 ppm en fonction de la taille des protéines et 50-300 ppm en fonction de la taille des peptides.

c.3 Western Blot

Après avoir caractérisé l'expression des deux construits PA1 et PA1b dans l'ensemble du système d'expression (cellules et milieu de culture) par HPLC et spectrométrie de masse, une détection des peptides produits grâce à l'utilisation d'anticorps Anti-PA1b a été réalisée sur une membrane de nitrocellulose. Les albumines de pois, extraites à partir de fa-

rine de la variété Frisson, ont été déposées sur gel de polyacrylamide dénaturant et transférées sur une membrane de nitrocellulose. En parallèle, la récolte des pics d'HPLC (provenant de la 4ème amplification virale) a été déposée sur le même gel et transférées sur la membrane. Cette dernière a ensuite été incubée avec des anticorps anti-PA1b afin de détecter les protéines dans les extraits issus des milieux de culture et des cellules infectées. En effet, le profil protéique obtenu à partir des cellules *Sf9* infectées était similaire pour les extraits protéiques issus de la farine du pois (Figure 32). Ceci suggère que les anticorps anti-PA1b reconnaissent des protéines produites dans les cellules *Sf9* infectées. D'autre part, une bande de taille de 11000 Da a été détectée dans le milieu de culture des cellules *Sf9* infectées par le virion recombinant *PA1* et présentant un épitope similaire à celui de PA1b. Cette bande pouvait néanmoins contenir la forme non clivée de la proprotéine PA1 (du PA1b au propeptide de la protéine PA1a, 11 kDa). De plus, la récolte des pics d'HPLC du milieu de culture cellulaire d'une 5ème amplification virale a permis la détection simultanée du PA1b ainsi que du PA1.

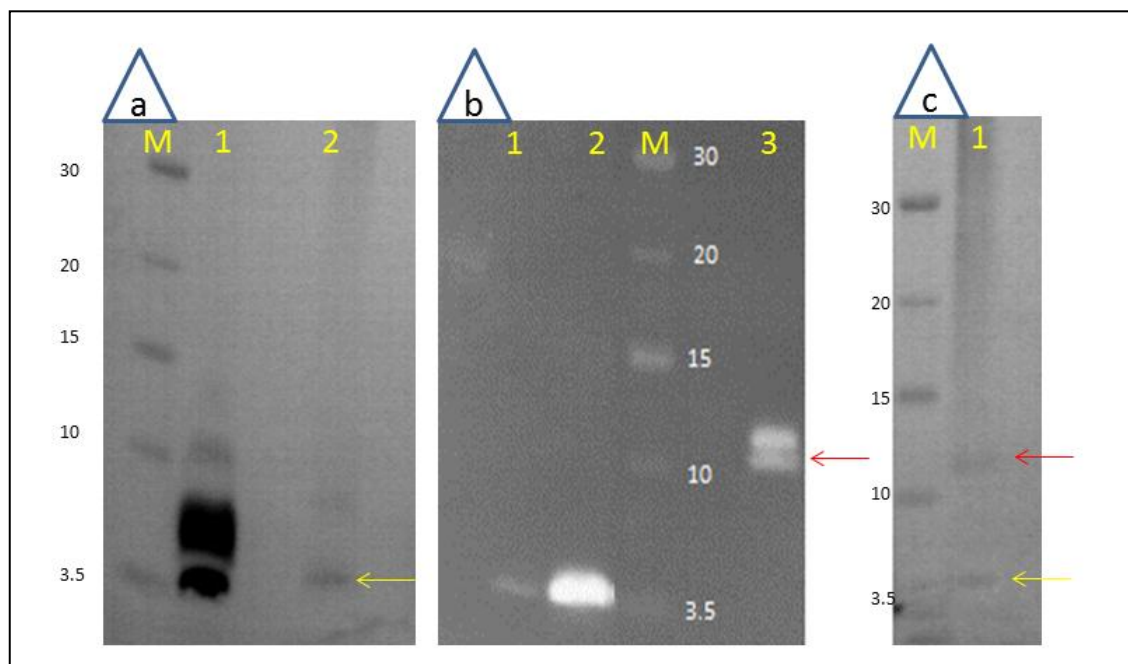


Figure 32 : Western Blot des purifications d'HPLC, avec pour anticorps anti-PA1b:

a) peptide PA1b produit (Amplification 4) : le puits N°1 correspond au PA1b extrait à partir du pois dont la concentration est 500 ng. Le puits N°2 correspond au pic A suggérant être le peptide PA1b d'une masse égale à 3700 Da.

b) proprotéine PA1 (Amplification 4) : le puits N°1 et 2 correspondent au PA1b extrait à partir du pois dont la concentration est 50 et 100ng respectivement. Le puits N°3 correspond au pic C suggérant être la proprotéine PA1 d'une masse égale à 11000 Da.

c) peptide PA1b + proprotéine PA1 (Amplification 5) : le puits N°1 correspond aux pics A et C trouvés simultanément dans le milieu de culture des cellules *Sf9* infectées par le virion *PA1*.

M étant l'échelle de marqueur de taille moléculaire (DNA-ladder).

d. Caractérisation de l'activité des protéines

Les échantillons de masses attendues ont été testés pour leur activité biologique sur cellules d'insectes *Sf9* saines sensibles à la toxine, par incorporation dans le milieu de culture l'échantillon récupéré par HPLC. Cette mesure de viabilité cellulaire consiste à exposer des cellules à la résazurine (couleur bleue). En présence d'agents réducteurs produits par l'activité cellulaire normale, particulièrement la production d'énergie par les mitochondries, la résazurine sera réduite en résorufine (un dérivé rosé), facilement quantifiable par spectrophotométrie. Les résultats obtenus par le test *in vitro* sur les cellules montrent une toxicité du pic A et B (donc le peptide PA1b) mais nous ne détectons pas d'activité pour le pic C (proprotéine PA1). La valeur de la constante LC_{50} est égale à 158 nM +/- 50 nM après 1 heure d'incubation avec le mélange des isoformes de PA1b extrait à partir de la farine du pois, alors que la valeur est de 389 nM +/- 108 nM après 1 heure d'incubation avec le peptide PA1b (pic A) produit par les cellules *Sf9* infectées par le virion recombinant *pa1* (Figure 33). Cette différence de valeur peut être expliquée par la forte toxicité du mélange d'isoforme de PA1b par rapport à une seule forme produite par le système baculovirus.

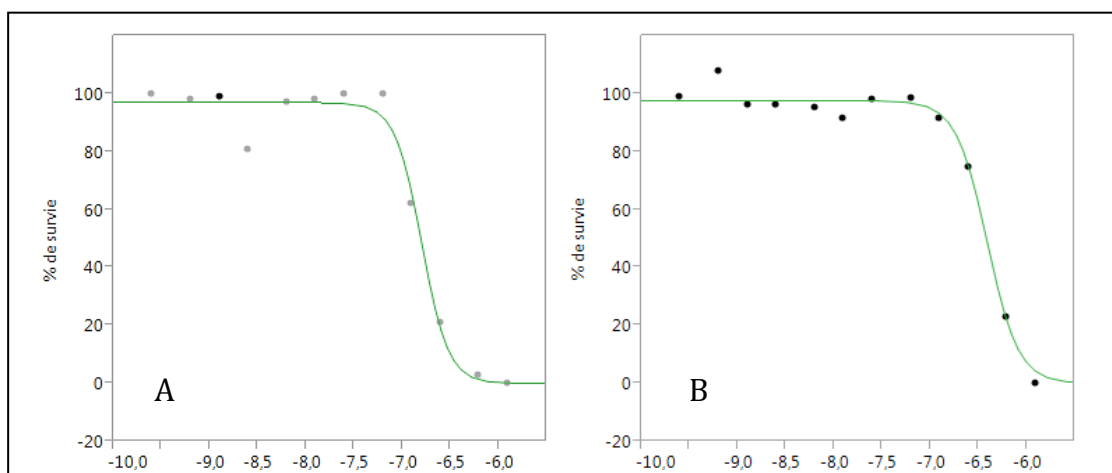


Figure 33 : Différentes courbes de viabilité cellulaire.

A-Courbes de viabilité cellulaire des cellules incubées en présence des isoformes de PA1b de la farine du pois.

B-Courbes de viabilité cellulaire des cellules incubées en présence de PA1b produit par le système baculovirus (virion recombinant *PA1*).

4. Discussion

L'objectif de la mise au point d'un système d'expression hétérologue de PA1b était la production de mutants pour une étude des relations structure-fonction. Toutefois, les difficultés rencontrées lors de cette mise au point et le temps limité imparti n'ont pas permis d'aborder la production des mutants. En effet, il convenait de vérifier avant tout que le peptide non muté était bien produit sous forme active dans le système choisi et développé, ce qui s'avère être le cas. Le système choisi, a priori adapté pour un peptide à ponts disulfures comme PA1b, s'avère capable de produire un peptide actif. Toutefois, nous avons pu montrer que l'expression d'un ADN complémentaire codant PA1 ou spécifiant PA1b dans les cellules d'insectes est possible tout en produisant un peptide conservant l'activité biologique toxique de base envers les cellules d'insectes *Sf9*. Le rendement d'expression obtenu est faible limitant ainsi la réalisation des tests d'activité *in vivo* sur les charançons des céréales (*Sitophilus oryzae*). Ce système, en présence de l'ADNc du PA1 complet, est donc capable de sécréter deux formes de peptides recombinants :

La première forme révélée est celle du PA1b (peptide identique à l'albumine-1E chaîne b dont le numéro d'accès Uniprot est P62930, résultant de la traduction du gène AJ475795). Il apparaît que ce peptide était bien produit, bien qu'en faible quantité, mais présentait une activité biologique. Le PA1b, extrait à partir des cellules *Sf9* infectées et non pas du milieu de culture, peut être interprété par le fait que le peptide est produit et sécrété à l'extérieur de la cellule, puis refixé sur son récepteur V-ATPase se trouvant à la surface de la membrane cellulaire. Ainsi, ce récepteur est capable de piéger les peptides sécrétés dans le milieu extérieur (Figure 34). Ainsi, pour confirmer cette hypothèse, nous avons réalisé des amplifications supérieures visant à produire du PA1b libre en quantité supérieure à celle fixée par les récepteurs. Par ce moyen, nous avons pu récupérer du PA1b libre dans le milieu de culture en plus de ceux attachés aux récepteurs (Figure 32). Ces résultats justifient la saturation de tous les sites des cellules *Sf9* et l'accumulation du PA1b libre dans le milieu de culture.

La deuxième forme révélée est celle de la proprotéine PA1 (du PA1b au propeptide du PA1a). Cette molécule intermédiaire entre la forme pré-proprotéine PA1 (séquence allant du peptide signal jusqu'à le propeptide du PA1a) et la forme toxique PA1b, ne présente pas de toxicité sur les cellules d'insectes *Sf9*. En présence de confirmation par spectrométrie de masse de la nature et de l'état de la proprotéine produite, cette protéine présente bien

ses 5 ponts disulfures, elle est bien structurée mais inactive. Afin de déterminer la structure tridimensionnelle de cette proprotéine, jamais extraite auparavant à partir de la farine du pois, une production à grande échelle est envisagée et la caractérisation de sa structure par RMN est proposée.

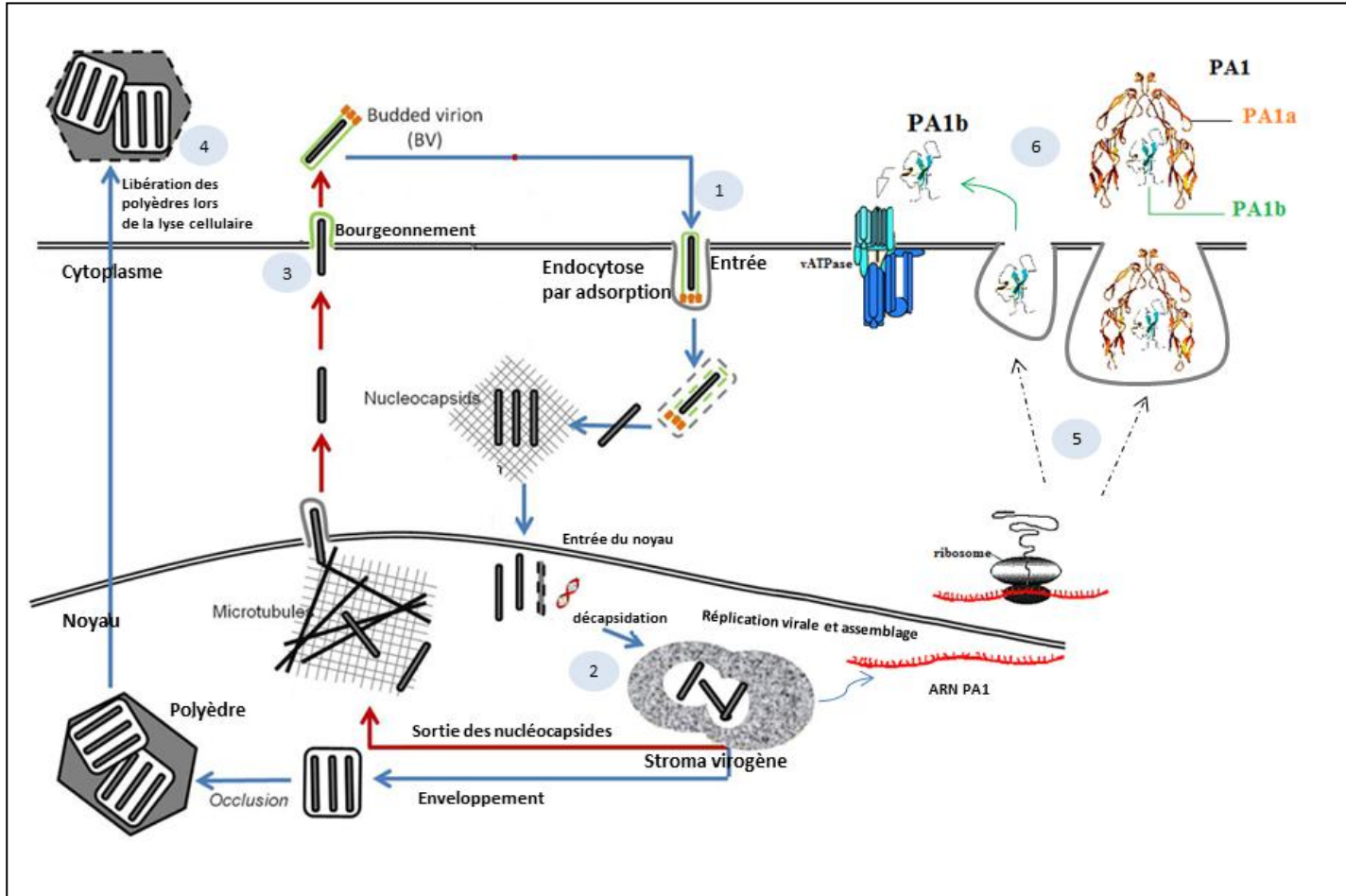


Figure 34: Représentation schématique de la production de PA1b dans les cellules *Sf9* infectées par les virions recombinants PA1.

(1) Les virions recombinants PA1 (BV) pénètrent dans la cellule principalement par un mécanisme d'endocytose et sont transportés par des vésicules jusqu'au noyau de la cellule.

(2) Dans le noyau, la répllication du virus conduit d'une part à la transcription de la séquence d'ARN d'intérêt (PA1) et d'autre part à l'assemblage des virions recombinants (virus de bourgeonnement) et à la formation des polyèdres.

Les nouveaux virus sont libérés, soit par bourgeonnement : BV (3), soit par lyse de la cellule : polyèdres (4).

(5) En outre, l'ARN PA1 sera transloqué dans le cytoplasme où il sera traduit. Les flèches en tiré représentent les différentes étapes de modifications post-traductionnelle permettant à la molécule produite d'être déversée dans le milieu externe via des vésicules de sécrétions.

(6) Les deux formes, le peptide PA1b et la proprotéine PA1, seront libérées dans le milieu de culture des cellules *Sf9*.

PA1b, dont le récepteur (V-ATPase) est localisé sur la membrane externe de la cellule *Sf9* productrice, sera à son tour fixé sur sa protéine cible alors que la forme complète PA1 (PA1b+PA1a) restera libre dans le milieu de culture. Une forte production de PA1b conduit à la saturation de tous les sites et ainsi PA1b se retrouve dans le milieu de culture.

Ps : la figure proposée par [MONTEIRO '12] a été modifiée selon les résultats obtenues dans ce chapitre.

La co-expression d'une protéine chaperonne des ponts disulfures, facilitant le repliement du peptide, a permis dans de nombreux cas d'améliorer le rendement de production de molécules actives [MASKOS '96, NISHIHARA '98, SCHAFFNER '01]. Dans ce sens, l'expression d'une séquence spécifiant tout PA1 a été envisagée. En effet, PA1b est suivi d'un polypeptide plus long, PA1a, dont la fonction est pour l'instant inconnue. La conservation de la structure de l'ADN complémentaire (peptide signal-PA1b- peptide intermédiaire -PA1a- peptide terminal) dans différentes espèces (cf. deuxième et troisième chapitre) témoigne de l'importance de celle-ci. Un des rôles possible de PA1a pourrait être d'aider à la conformation de PA1b. Nos résultats ont bien révélé l'importance de la présence de PA1a. Une production 20 fois supérieure de PA1b a été décelée en utilisant un virion recombinant contenant la séquence complète de PA1 par rapport à un virion PA1b.

D'autre part, la production de PA1b Chez *E.coli* [SANDRINE LOUIS '04a] a posé des problèmes très importants de rendement de synthèse et de repliement. Toutefois, des essais de production de PA1b recombinant par *P. pastoris* avaient été réalisés, dans l'unité BF2I, sans succès ([ZHANG '01], données non publiées). Le problème rencontré est celui de la formation d'une structure active de PA1b. Le peptide était apparemment bien produit, bien qu'en faible quantité, mais ne présentait pas d'activité biologique. Comme le gène s'était avéré correctement transcrit, deux hypothèses avaient été évoquées pour expliquer cet échec : soit un mauvais processus de repliement et/ou d'oxydation du peptide recombinant chez la levure, soit une interférence de la forte hydrophobicité de PA1b avec le système de sécrétion. En absence de confirmation par spectrométrie de masse de la nature et de l'état du peptide produit, cet échec n'a pas pu être totalement interprété. Il apparaît donc que dans les deux cas, le peptide est mal structuré et donc inactif.

Par ailleurs, l'équipe du Dr Hirano travaille sur le rôle de la léginsuline du soja (homologue de PA1b) dans la plante, et a montré sa liaison à un récepteur protéique des membranes plasmiques et parois cellulaires de soja. Cette équipe a réussi, d'une part, à produire dans *E. coli* la léginsuline sous forme active (ie liant le récepteur végétal) [HANADA '03], et d'autre part, à la synthétiser chimiquement [YAMAZAKI '03]. La production dans *E. coli* a nécessité l'utilisation d'une souche bactérienne nommée BL21trxB(DE3) caractérisée par une simple mutation (gène de la thioredoxine réductase) et une fusion avec la thioredoxine.

D'autre part, la structuration de PA1b après synthèse chimique a montré des résultats positifs vis-à-vis du rendement de synthèse et de re-

pliement [DA SILVA '09]. Cette synthèse s'est avérée satisfaisante en termes de reproductibilité et le peptide produit possédait une activité biologique conforme en test de compétition de liaison avec ^{125}I -PA1b ainsi qu'en test *in vivo* sur des charançons et *in vitro* sur les cellules d'insectes Sf9. De plus cette production par synthèse chimique a ouvert la possibilité d'étudier la relation existante entre la structure et la fonction de PA1b en réalisant une étude par mutagenèse dirigée (Alanine-scanning). Cette étude a permis l'identification de quelques acides aminés ainsi que de boucles hydrophobes nécessaire à la conservation de l'activité biologique du peptide [DA SILVA '10].

En outre, et au cours de la réalisation de notre production de peptides recombinants dans les cellules d'insectes, Vanessa Eyraud (Docteurante au laboratoire BF2I) a réussi à produire PA1b, bien structuré et actif, dans les plantes transgéniques de tabac. Le principe consiste à insérer le gène de PA1b dans la bactérie *Agrobacterium tumefaciens*, puis une culture bactérienne est co-injectée avec un second vecteur, contenant un suppresseur de silencing, sous vide sur une feuille de tabac (*Nicotiana benthamiana*). Cette plante va alors surproduire la protéine d'intérêt qui sera ensuite purifiée à partir des feuilles. L'analyse HPLC des extraits a montré que le tabac contenait un produit au même temps de rétention que le peptide PA1b du pois. L'analyse en spectrométrie de masse, puis en RMN, a confirmé que le tabac produit du PA1b avec une structure 3D identique à celle du peptide natif, et avec un rendement assez important. Enfin, des tests biologiques sur charançons et sur cellules Sf9 ont confirmé que le peptide était actif aux mêmes concentrations que le peptide natif. L'analyse en masse révèle cependant la présence de plusieurs formes du peptide. Il existe également, en plus de la forme native, des formes du peptides avec des coupures protéolytiques légèrement différentes de ce qui est observé chez le pois, avec un peptide où la glycine terminale manque, et une forme où l'alanine initiale manque [EYRAUD '13]. Egalement, différentes espèces de tabac ont été testées pour leur aptitudes à produire la toxine (*N. tabacum*, *N. glauca*, *N. suaveolens* et *N. rustica*), mais les résultats ont montré que *N. benthamiana* demeurerait le meilleur système.

Une étude semblable à celle proposée par Vanessa Eyraud a été réalisée en 2008 par Gillon et ses collaborateurs. Cette étude consistait à produire une knottine végétale (le cyclotide OaK1 : *O. affinis* Kalata B1) dans des plantes transgéniques : *Arabidopsis thaliana*, *Nicotiana tabacum* ainsi que *Nicotiana benthamiana*. L'expression de ce cyclotide dans ces plantes transgéniques a permis d'une part de déterminer le système le plus adapté (*N. benthamiana*) pour la production d'un peptide à nœud de cystine cyclique et d'autre part de caractériser les acides aminés nécessaires pour

cette cyclisation. Ces plantes utilisées possèdent des mécanismes de coupures enzymatiques différentes de celle trouvées chez la plante d'origine (*O. affinis*) permettant de plus l'obtention de forme linéaire de la séquence du cyclotide non caractérisée auparavant chez *O. affinis* [GILLON '08].

Finalement, pour essayer de mieux comprendre le rôle du peptide signal dans la production et l'accumulation du peptide dans le milieu de culture, nous avons envisagé la production de PA1b avec son peptide signal d'origine. Ce peptide signal végétal, inséré dans le génome du baculovirus suivi de la protéine majeure, a permis l'adressage du peptide produit. Le PA1b étant excrété dans le milieu de culture montre la capacité du système baculovirus à produire un peptide végétal en utilisant un peptide signal de la plante. D'autres essais de production d'un peptide végétal dans le système baculovirus ont montré la capacité de ce système à bien produire des molécules de plantes mais en fusion avec un peptide signal viral. Cette production illustre le travail de Michael Vetsch en 1999 où il a réussi à produire une prosystemine des tomates dans des cellules d'insectes *Sf21* après infection par un baculovirus recombinant contenant le gène de la protéine en fusion avec le peptide signal d'une protéine virale, la GP67 [VETSCH '00].

Discussion générale et perspectives

Ce travail de thèse, à l'image de poupées gigognes, comprenait plusieurs problématiques imbriquées les unes dans les autres, avec des objectifs somme toute assez différents : l'un à visée d'histoire évolutive d'une famille de peptides spécifique des légumineuses, l'autre à visée plutôt patrimoniale, le troisième enfin à objectif méthodologique, qui est inclus dans une stratégie d'équipe pour comprendre l'intégralité de l'information portée par les peptides PA1.

Ainsi, mon travail durant ces trois ans a consisté d'abord à rechercher des homologues de PA1b au sein de l'espèce modèle des légumineuses *Medicago truncatula* par une approche « *in silico* » complémentée par l'analyse de la toxicité des séquences choisies et synthétisées chimiquement envers deux systèmes modèles biologiques adoptés au laboratoire BF2I : i) *in vitro* : les cellules d'insectes Sf9 dont nous connaissons la sensibilité envers PA1b et ii) *in vivo* : *Sitophilus oryzae*, un charançon des céréales et ses souches sensibles et résistantes à la toxine PA1b. La méthodologie a été mise en place en premier lieu sur la version 3.5 du génome de *Medicago truncatula* publié en 2011 [YOUNG '11] et a été actualisée en fin de thèse pour la toute nouvelle version du génome de *Medicago* (version 4 - tout juste publiée en Septembre 2013). Cette dernière version du génome nous a permis de combler tous les manques trouvés dans la version antérieure et nous a aidés à déterminer les correspondances entre les séquences de gènes et les séquences d'ESTs auparavant classées orphelines dans la version 3.5.

Dans un autre temps, nous nous sommes focalisés sur la recherche d'homologues de PA1b au sein de différentes espèces de Légumineuses originaires du Liban, par voie de la biologie moléculaire (gènes homologues - clonage par homologie).

Finalement, afin de relier les variations de séquences aux variations de toxicité obtenues par ces approches, un système d'expression hétérologue a été mis au point afin de produire, dans le futur, certains variants naturels et des mutants choisis de PA1b.

Pour résumer et conclure ce travail, je vais extraire un certain nombre de faits marquants qui le caractérisent à mes yeux :

*a. Evolution de la famille des A1bs au sein de l'espèce modèle des Légumineuses *Medicago truncatula*.*

Comprendre les forces qui façonnent la diversité moléculaire au sein des espèces est un enjeu majeur de la biologie évolutive mais également pour la gestion et l'utilisation des ressources génétiques et la biologie de la conservation. Les duplications de gènes apparaissent, par ailleurs, comme un des principaux moteurs de l'évolution des génomes, donc des espèces. En effet, elles semblent être un tremplin pour l'acquisition de nouvelles fonctions et la génération de nouvelles adaptations [FORCE '99, MICHAEL LYNCH '00a, MOORE '05]. Elles contribuent significativement à l'augmentation de la complexité du fonctionnement des organismes [FREELING '06]. Dans notre étude, nous avons profité du séquençage complet du génome de *Medicago truncatula* pour identifier tous les paralogues des Albumines 1 au sein de cette espèce modèle des Légumineuses. Les résultats révèlent la présence de **53 séquences homologues** à la séquence de PA1 de l'espèce *Pisum sativum*. Ces séquences ont été distribuées sur 87.5% des chromosomes de *Medicago* avec une forte représentation au niveau du chromosome 3. Parmi toutes ces séquences, 43% présentent une expression diversifiée au niveau tissulaire et 57% semblent avoir perdu toute trace d'expression. Nous avons constaté une distribution tissulaire large des Albumines 1, inconnue à ce jour chez d'autres légumineuses. Ces résultats ont souligné le fait que, contrairement à nos connaissances antérieures, la graine n'est pas l'organe privilégié d'expression des A1-like chez *Medicago truncatula*. Les racines et les feuilles paraissent avoir un taux plus élevé d'expression des A1-like. Ces résultats nous ont ensuite poussés à rechercher les forces évolutives qui façonnent l'évolution des gènes dupliqués d'Albumines 1, celles qui s'exercent sur la redondance fonctionnelle initiale et mènent à l'acquisition de nouvelles fonctions.

Dans ce contexte, comprendre les mécanismes de l'adaptation au sein de l'espèce modèle des Légumineuses *Medicago truncatula* est un enjeu de taille. C'est pourquoi nous avons sélectionné quelques séquences d'Albumines 1 afin de tester leur fonction biologique de base : la fonction insecticide. Au sein des séquences sélectionnées, quatre conservent une activité insecticide contre d'une part un insecte séminivore (le charançon du riz *Sitophilus oryzae*) et d'autre part sur des cellules d'insectes modèles (la lignée Sf9). Les trois séquences testées restantes perdent complètement cette activité d'origine. En analysant les divergences entre les co-

pies de gènes d'Albumines 1, nous avons remarqué le dynamisme d'évolution au sein de ces gènes. Ces gènes d'Albumines 1, qui se retrouvent répartis dans le génome de Medicago peuvent subir différents processus évolutifs, comme proposés initialement par Ohno [OHNO '70] puis développés par différents groupes [FORCE '99, MICHAEL LYNCH '00a, MOORE '05]. Dans un grand nombre de cas, des copies des gènes subissent un processus de silencing, une délétion totale ou accumulent des mutations et dégénèrent en passant par l'état de pseudogène c'est-à-dire une copie de gène qui n'est plus exprimée ou dont la protéine n'est plus fonctionnelle. Par la suite ces pseudogènes peuvent être délétés du génome ou alors diverger de telle manière qu'aucune similarité de séquence avec la copie originelle ne peut plus être mise en évidence. Par contre, si les pseudogènes ont été formés récemment, ils pourront être identifiés grâce à leur similarité de séquence. Citons par exemple le cas de la séquence Medtr3g467760.1 présentant un codon stop au niveau de la deuxième cystéine de la séquence de PA1b. Cette séquence non exprimée est une des copies inactivée par l'accumulation de mutations. On parle ainsi de 'pseudogénisation'.

Dans d'autres cas, des copies de gènes se maintiennent et restent fonctionnelles. La présence de deux copies de gènes peut parfois être un avantage simplement parce que cette redondance permet d'augmenter les quantités de protéines ou d'ARN synthétisées. Ainsi la duplication génique peut donc s'expliquer par le besoin de maintenir une certaine quantité de produit des gènes dans la cellule (Citons l'exemple des deux gènes Medtr3g067540.1 et Medtr3g067555).

Bien qu'une quantité plus importante de produit des gènes peut être avantageuse, deux gènes ayant la même fonction ne sont, en général, pas stables au sein d'un génome [NOWAK '97]. En théorie, deux copies identiques peuvent être maintenues si la fonction correspondante diffère par certains points [NOWAK '97]. Le processus de 'subfonctionnalisation', encore appelé «partage de fonctions ancestrales», permet justement aux deux copies d'acquérir une différence de fonction ou d'expression assurant leur maintien. La subfonctionnalisation peut avoir lieu lors d'une duplication d'un gène codant deux domaines fonctionnels. Chaque copie va conserver un des deux domaines fonctionnels par accumulation de mutations dans la séquence codant l'autre domaine fonctionnel. Ainsi, la fonction ancestrale sera partagée entre les deux copies. La subfonctionnalisation peut également correspondre à la division de l'expression génique après une duplication c'est-à-dire que les deux copies seront exprimées dans des conditions différentes. Dans les deux cas, les deux gènes se com-

plémentent fonctionnellement et seront conservés dans le génome [MOORE '05, RIJPKEMA '07].

Enfin, la ‘néofonctionnalisation’ est une voie possible d’évolution de ces gènes, évoquée pour qualifier l’acquisition de nouvelles fonctions (Citons l’exemple d’AC146565_12.1 qui a perdu sa fonction d’origine entomotoxique envers les charançons et les cellules Sf9).

Toutes ces hypothèses peuvent être vérifiées en synthétisant chimiquement ces 53 PA1-like ou en les exprimant dans un système hétérologue, en testant ensuite leur activité biologique, par exemple en élargissant le spectre d’hôte (bactérie, champignon, levure ...).

a.1 L’émergence de la famille multigénique Albumines 1

Les duplications des gènes d’Albumines 1 entraînent une redondance de l’information génétique et sont une composante principale de l’évolution du génome et de l’apparition d’une **famille multigénique**. Sous la pression évolutive, la redondance initiale a tendance à diminuer au cours du temps, entraînant une modification du génome sur le plan structural et sur le plan fonctionnel [WENDEL '00]. L’acquisition progressive de différences, comme l’accumulation aléatoire des mutations dans leur séquence et la modification de leur structure ont pour conséquence leur évolution fonctionnelle. Leur évolution peut concerner la séquence codante, les parties non codantes ou les régions régulatrices du gène dupliqué [M. LYNCH '00b, M. LYNCH '01].

Les albumines 1 sont un bon exemple d’enrichissement d’une famille à partir d’un élément initial que nous avons montré être les toxines insecticides (qui ne semblent donc pas être une fonction dérivée récente de la diversification). A partir de ce patron structural, une diversification progressive ayant d’abord entraîné une diminution/extinction de la fonction ancestrale (Cluster 1/ chromosome 8), puis une expression diversifiée hors de son champ d’origine (graines), ayant permis la naissance d’une famille entière néofonctionnalisée (nodulines A1b). Enfin, plus récemment, est apparue une extension encore mal comprise sur le chromosome 3, de fonction a priori également non insecticide. L’auto-stop génomique de la famille est donc actuellement relativement lisible, et nous semble un exemple assez unique de petites molécules ayant subi plusieurs cycles de néofonctionnalisation chez les légumineuses.

a.2 *Medicago truncatula*: modèle d'interaction racine-insecte

D'après les données disponibles, on peut constater que *M. truncatula* a émergé en tant qu'une importante plante modèle des légumineuses facilitant les progrès dans le domaine de la symbiose et ouvrant de nouveaux domaines de recherches concernant le stress biotique, les réponses des plantes aux ravageurs et aux agents pathogènes et le développement des plantes. La famille des Albumines 1 se trouvant, pour la première fois, fortement exprimées dans les racines de *Mtr*, contrairement aux autres espèces de légumineuses, crée ainsi une **plateforme** pour la poursuite des recherches dans ce domaine, plus précisément dans le domaine de **l'interaction insecte /racine** (encore peu explorée).

L'ensemble de ces résultats expérimentaux et théoriques nous ont permis d'avancer dans notre compréhension de l'évolution des Albumines 1 au sein du génome de *Medicago truncatula* et de l'adaptation des ces gènes au sein de l'espèce modèle des légumineuses par l'intermédiaire de la duplication.

b. Ancestralité des Albumines 1.

La famille des Albumines 1 étant spécifique des légumineuses, elle a probablement son origine au sein de cette famille botanique dans une lignée de peptides que l'on pourrait reconnaître par ailleurs (famille peptidique déjà connue). Cette hypothèse de base nous a motivé à rechercher d'autres annotations possibles dans les A1 caractérisées. L'étude réalisée à cet effet n'a pas abouti à la détermination d'une famille reliée aux albumines 1 hors cette famille peptidique, qui reste donc un groupe relativement isolé dans les grandes banques de familles protéiques, qu'elles soient structurales (SCOP « **S**tructural **C**lassification **O**f **P**roteins », CATH représente quatre niveaux de classification « **C**lass, **A**rchitecture, **T**opology & **H**omologous superfamily») ou phylogénétiques (InterPro, PFAM).

C'est pourquoi une étude plus approfondie sur les A1b-like au niveau des génomes des Légumineuses complètement séquencées et d'une séquence ciblée récemment découverte chez *Styphnolobium japonicum*, a été réalisée et nous a permis de suivre l'évolution des A1b au sein des Légumineuses. L'ancestralité de ces peptides (au niveau structural et fonc-

tionnel) date d'au moins **58 million d'années**. La conservation de la fonction insecticide a été retrouvée chez *Styphnolobium japonicum* avec une activité quatre fois plus élevée que l'albumine PA1b du pois. La conservation de la fonction insecticide témoigne de son importance au niveau de la plante hôte, et dans les mécanismes de défense (**AG41** présente un facteur 10 supplémentaire de performance par rapport à la molécule de référence PA1b toxique du pois, et est en cours de protection à cet égard).

c. A la découverte de nouvelles molécules homologues de PA1b chez les Légumineuses originaires du Liban.

Lors de notre travail de thèse, la caractérisation de nouveaux gènes par PCR génomique avec amorces dégénérées, nous a permis de montrer que les homologues de PA1b constituent une véritable famille peptidique répartie au sein de la famille des Légumineuses originaires du Liban. Des gènes homologues, provenant de 6 espèces et représentant le clade de Papilionoideae (Annexe 7), ont en effet été caractérisés au cours de ce travail. La limitation de la caractérisation de ces gènes est très probablement due à une limitation technique, les amorces dégénérées utilisées pour rechercher les gènes homologues ayant été déterminées d'après les séquences des paralogues de PA1 chez *Medicago truncatula* d'une part, et d'autre part de quelques orthologues de PA1b provenant de quelques espèces différentes au sein des Légumineuses. Les gènes des A1b des autres espèces de Fabaceae sont peut-être trop variants pour être reconnus par nos amorces.

La diversité des Albumines 1 au sein de ces espèces originaires du Liban s'appuie sur le fait que la dynamique de la diversité génétique de ces espèces est dépendante non seulement des traits d'histoire de vie des espèces mais aussi des différents contextes culturels, sociaux et agronomiques qui interagissent. L'étude plus approfondie de ces Albumines 1 (structure comme fonction) contribuera éventuellement à une meilleure compréhension des processus évolutifs et biogéographiques à l'origine de l'endémisme au Liban. Nous pouvons tenter une hiérarchisation des priorités en la matière, en choisissant d'abord des groupes dont l'endémisme reflète plutôt une « diversité fondatrice » (espèces proches des aires de diversification reconnues par ailleurs pour les groupes botaniques considérés). Par ailleurs, le choix d'espèces proches de groupes cultivés pourra également enrichir l'échantillonnage taxonomique autour de questionnements liés à la domestication (genre *Pisum* pour le moyen orient, genre *Phaseolus* pour l'Amérique du Sud, etc...).

d. Expression hétérologue de PA1b dans le système baculovirus/ cellules d'insectes.

Nous avons également montré que la structure de PA1b, exposant ses acides aminés hydrophobes, possédant une proline cis et trois ponts disulfures imbriqués en cystine-knot [JOUVENSAL '03], peut être exprimée dans un système hétérologue baculovirus/ cellules d'insecte tout en conservant son activité biologique. Cette production ayant été accomplie, elle nous a permis de bénéficier de notions nouvelles sur le système utilisé et le peptide produit et de proposer une représentation schématique de la production de PA1b au sein du système :

d.1 Peptide signal d'adressage à caractéristique universelle

Le peptide PA1b et la proprotéine PA1 (forme intermédiaire entre la préproprotéine et le peptide mature PA1b) produits par les cellules Sf9 infectées par le virion recombinant se retrouve à la surface externe de la cellule et dans le milieu de culture, respectivement. Ces observations confirment bien l'universalité du mécanisme de reconnaissance de ce peptide signal. Ce dernier est un peptide de plante qui a pu être reconnu par le système cellulaire d'insectes et qui, en conservant sa fonctionnalité, a permis l'adressage de PA1b et PA1 vers le milieu extracellulaire.

Ce système, en ayant la capacité d'adresser le peptide produit vers le milieu externe, incorpora aussi une nouvelle possibilité d'appréhender le potentiel de liaison du peptide produit (homologue de PA1b) à sa cible cellulaire (V-ATPase). En effet, si l'homologue de PA1b produit par ce système se retrouve dans le milieu de culture, cela signifie l'absence de liaison de cet homologue à la V-ATPase cellulaire. De plus, ce système présente un outil de piégeage de PA1b au niveau des cellules Sf9 facilitant ainsi sa purification (PA1b étant soluble dans l'éthanol). Ainsi, une production de PA1b taguée n'est plus exigée.

d.2 Le rôle du peptide PA1a (et de la proprotéine PA1)

Grâce aux données des tests biologiques et des séquences peptidiques obtenues lors de notre travail, nous avons pu émettre des hypothèses quant au rôle du peptide PA1a dans la synthèse du peptide mature PA1b. La présence de PA1a auprès de PA1b favorise sa production, et

améliore le rendement d'une valeur de 20 fois environ par rapport à un virion recombinant ne contenant que la séquence de PA1b (sans PA1a).

Dans le cas du pois, PA1b est légèrement basique (Pi 7,8) et PA1a, bien qu'acide (Pi 4,9) n'aurait donc pas un rôle de neutralisateur comme dans le cas des thionines qui sont également exprimées sous forme de préproprotéine avec un peptide C-terminal acide qui neutraliserait la basicité du domaine thionine lors de la synthèse protéique [BOHLMANN '91]. Si aucune interaction entre A1a et A1b n'a été démontrée jusqu'à présent (ni vis-à-vis de l'activité entomotoxique de PA1b, ni vis-à-vis de son rôle putatif d'hormone végétale), la conservation très forte des A1a témoigne de l'existence d'une forte sélection purifiante, et donc de leur importance fonctionnelle. Il est possible que PA1a intervienne lors du repliement de PA1b, en tant que chaperonne. L'équipe « Entomotoxine » travaille actuellement sur cette hypothèse [EYRAUD '13].

Par ailleurs, **cette étude nous a permis de détecter, pour la première fois dans un système de production hétérologue, la proprotéine PA1** (non décelée dans aucun autre système d'expression déjà explorés : *E.coli*, *P. pastoris*, *O.sativa*, *N.benthamiana*...) et nous a permis de purifier cette molécule et de tester son activité biologique. En 1986, les travaux de Higgins ont conduit à la découverte de cette proprotéine dans les cotylédons du pois [HIGGINS '86]. Cette proprotéine a montré une absence d'activité insecticide envers les cellules Sf9 sensibles à PA1b et une absence de fixation de cette proprotéine au récepteur de PA1b. Ces observations sont en faveur du **rôle de PA1a en tant que protéine chaperonne** qui favoriserait le repliement de PA1b.

Ce système de production hétérologue est capable de produire PA1b mais en faible quantité, limitant ainsi son utilisation dans des tests biologiques *in vivo*, nécessitant une grande quantité de la toxine. Deux autres systèmes de production de PA1b ont été mis en place : la synthèse chimique et le système tabac. L'adoption d'un des systèmes sera faite selon les différents objectifs envisagés.

d.3 Importance du Baculovirus recombinant PA1 dans la lutte biologique.

Les Baculovirus ont présenté depuis longtemps un grand intérêt dans la lutte contre les ravageurs des plantes. Ils causent souvent des épidémies, exterminant presque des populations entières de chenilles et des insectes, généralement à la fin de la saison de croissance. Ces maladies sont relativement spécifiques à un nombre limité d'espèces d'hôtes poten-

tiels [WALTER '03]. Ces virus peuvent aussi rapidement causer la mortalité dans les 4 à 6 jours qui suivent l'infection selon la dose, la concentration et l'isolat du virus.

Les Baculovirus sont des virus exclusivement pathogènes d'invertébrés, en particulier les insectes, mais non infectieux pour les vertébrés. Les différents tests exécutés par Possée *et al* [POSSEE '93] ont attesté la non toxicité des Baculovirus pour les mammifères et même les formes recombinées. Le mode d'action hautement spécifique et unique, la spécificité de l'hôte, font des Baculovirus une forme de lutte séduisante pour les agriculteurs, et cadre bien avec les programmes de lutte intégrée. Ils constituent des agents relativement sûrs du point de vue de la santé des vertébrés et entraînent des impacts environnementaux très négligeables.

Les caractéristiques principales des biopesticides à base de Baculovirus sont : la spécificité, la haute virulence, la rapidité d'action et le niveau raisonnable de persistance dans l'environnement [WALTER '03]. Bien que des problèmes liés au développement de ces biopesticides soient présents tels qu'un spectre d'action étroit, une lenteur de leur action par rapport aux insecticides chimiques, la persistance limitée au champ, la sensibilité au rayonnement solaire et au pH plus élevé du feuillage, le **baculovirus recombinant PA1** pourrait apporter de nouvelles perspectives dans le domaine de la **lutte biologique** en ouvrant la possibilité de générer un **biopesticide** présentant une activité insecticide plus poussée couplant l'action virale de base du baculovirus et l'activité insecticide de la toxine du pois PA1b.

La lutte biologique, qui préconise l'emploi d'agents très spécifiques, ne peut être considérée comme une panacée universelle susceptible de protéger toute culture contre tout ennemi. C'est pourquoi aujourd'hui, la mise en œuvre de ce concept s'inscrit systématiquement dans le cadre de la lutte intégrée. C'est dans ce contexte que le présent travail effectué dans le laboratoire BF2I trouve sa pertinence.

Ainsi les polyèdres contenant les virions recombinants pourront délivrer le gène du PA1b aux insectes, créant ainsi une sorte de vecteur transportant PA1b à sa cible interne. Citons par exemple le cas de *Spodoptera frugiperda* insensible à l'action de PA1b par ingestion alors que les cellules Sf9 provenant de cet insecte présentent une forte sensibilité vis-à-vis de la toxine PA1b. Le nouveau virion recombinant PA1 pourra ainsi faciliter la traversée de la barrière intestinale (résistant au peptide PA1b) du gène PA1 et la délivrance aux cellules non digestives sensibles.

De façon globale, l'analyse des résultats soulève des questions qu'il serait souhaitable d'étudier dans l'avenir. Il s'agit, après avoir démontré la valeur ajoutée de cette nouvelle combinaison toxine-baculovirus

versus baculovirus sauvage (test de comparaison en laboratoire sur un échantillonnage conséquent), de :

- reprendre les mêmes investigations en milieu réel et d'en déduire les impacts réels dans une zone écologiquement contrôlable.
- étudier la rentabilité de cette méthode de lutte.
- comparer l'efficacité du baculovirus sauvage au champ et des baculovirus recombinant dans le milieu réel.
- vérifier l'absence de nocivité du biopesticide généré sur les vertébrés.

Ainsi cette étude apporte un nouvel outil via ce nouveau virus mixte qui consiste en un apport de connaissance au niveau fondamental et appliqué.

En guise de conclusion et de point final, nos investigations montrent clairement que la famille des A1 possède encore beaucoup de bio-activités originales à révéler. Seule une petite partie de l'immense biodiversité moléculaire qu'elles offrent a été étudiée expérimentalement. C'est une illustration, dans le domaine de la biodiversité, de l'intérêt (« structural scaffolds ») porté par les chimistes et biochimistes à la famille des knottines, dont les A1b font partie (<http://knottin.cbs.cnrs.fr>). Nos connaissances actuelles suggèrent que la famille des A1 a évolué vers plusieurs objectifs. Une vision plus large est clairement nécessaire et seules des études futures à grande échelle nous dira si cette famille est aussi prometteuse que l'équipe Entomotox le pense : au delà du stockage de l'azote ou du soufre, elle offre un modèle passionnant d'étude de ce qui pourrait être nommé l'« interactome » des légumineuses, et qui reflète la richesse écologique que cette famille végétale a su investir dans ses précieuses graines ou dans ses racines.

Références bibliographiques

- [ABDEEN '05] A. Abdeen *et al.*, Multiple insect resistance in transgenic tomato plants over-expressing two families of plant proteinase inhibitors. *Plant molecular biology*, 2005, **57**, 2, pp. 189.
- [ABDEL-SAMAD '13] F. Abdel-Samad, Caractérisation écogéographique et génétique du genre *Astragalus* du Liban : approches de conservation biogéographique. These en cotutelle-Universite Aix Marseille et Universite St Joseph, 2013.
- [ADAM '00] D. Adam, Arabidopsis thaliana genome. Now for the hard ones. *Nature*, 2000, **408**, 6814, pp. 792.
- [ADLER '13] Adler, N'Domo, paper presented at the 9th Conference on Integrated Protection of Stored Products IPSP 2013, Agora - University Bordeaux 1 TALENCE, FRANCE, 1 - 4 July, 2013 2013.
- [AGOUNKÉ '94] B. Agounké, Les règles de l'art: protection des denrées stockées combinant le fractionnement des récoltes et l'application d'insecticides. *GTZ, Eschborn, Allemagne.*, 1994.
- [AKOU-EDI '84] D. Akou-Edi, paper presented at the Proceedings of the 2nd Neem conference, Rauischolzhausen, 1984.
- [ARABIDOPSIS-GENOME-INITIATIVE '00] Arabidopsis-Genome-Initiative, Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, 2000, **408**, 6814, pp. 796.
- [AYRES '94] M. D. Ayres, S. C. Howard, J. Kuzio, M. Lopez-Ferber, R. D. Possee, The complete DNA sequence of Autographa californica nuclear polyhedrosis virus. *Virology*, 1994, **202**, 2, pp. 586.
- [BANEYX '99] F. Baneyx, Recombinant protein expression in Escherichia coli. *Current opinion in biotechnology*, 1999, **10**, 5, pp. 411.
- [BARBETA '08] B. L. Barbeta, A. T. Marshall, A. D. Gillon, D. J. Craik, M. A. Anderson, Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 4, pp. 1221.
- [BARKER '90] D. Barker *et al.*, Medicago truncatula, a model plant for studying the molecular genetics of the Rhizobium-legume symbiosis. *Plant Mol Biol Rep*, 1990, **8**, 1, pp. 40.
- [BECKER '08] S. Becker, H. Terlau, Toxins from cone snails: properties, applications and biotechnological production. *Applied microbiology and biotechnology*, 2008, **79**, 1, pp. 1.
- [BELL '94] A. Bell. *GTZ, Eschborn, Allemange*, 1994.
- [BENNETT '97] M. D. Bennett, I. J. Leitch, Nuclear DNA Amounts in Angiosperms-583 New Estimates. *Annals of Botany*, 1997, **80**, pp. 169.
- [BLANCO-APARICIO '98] C. Blanco-Aparicio *et al.*, Potato carboxypeptidase inhibitor, a T-knot protein, is an epidermal growth factor antagonist that inhibits tumor cell growth. *The Journal of biological chemistry*, 1998, **273**, 20, pp. 12370.
- [BODE '89] W. Bode, H. J. Greyling, R. Huber, J. Otlewski, T. Wilusz, The refined 2.0 Å X-ray crystal structure of the complex formed between bovine β -trypsin and CMTI-I, a trypsin inhibitor from squash seeds (*Cucurbita maxima*) Topological similarity of the squash seed inhibitors with the carboxypeptidase A inhibitor from potatoes. *FEBS Letters*, 1989, **242**, 2, pp. 285.
- [BOHLMANN '91] H. Bohlmann, K. Apel, Thionins. *Annual Review of Plant Physiology and Plant Molecular Biology*, 1991, **42**, 1, pp. 227.

- [BUCZEK '05] O. Buczek, G. Bulaj, B. M. Olivera, Conotoxins and the posttranslational modification of secreted gene products. *Cellular and molecular life sciences : CMLS*, 2005, **62**, 24, pp. 3067.
- [CABRAL '03] K. M. Cabral, M. S. Almeida, A. P. Valente, F. C. Almeida, E. Kurtenbach, Production of the active antifungal *Pisum sativum* defensin 1 (Psd1) in *Pichia pastoris*: overcoming the inefficiency of the STE13 protease. *Protein Expr Purif*, 2003, **31**, 1, pp. 115.
- [CARDONA '90] C. Cardona, J. Kornegay, C. E. Posso, F. Morales, H. Ramirez, Comparative value of four arcelin variants in the development of dry bean lines resistant to the Mexican bean weevil. *Entomologia Experimentalis et Applicata*, 1990, **56**, 2, pp. 197.
- [CHICHE '89] L. Chiche *et al.*, Use of restrained molecular dynamics in water to determine three-dimensional protein structure: prediction of the three-dimensional structure of Ecballium elaterium trypsin inhibitor II. *Proteins*, 1989, **6**, 4, pp. 405.
- [CHICHE '10] L. Chiche, J.-C. Gelly, J. Gracy, A. Heitz. 2010.
- [CHOUABE '11] C. Chouabe *et al.*, New Mode of Action for a Knottin Protein Bioinsecticide: PEA ALBUMIN 1 SUBUNIT b (PA1b) IS THE FIRST PEPTIDIC INHIBITOR OF V-ATPase*. *The Journal of biological chemistry*, 2011, **286**, 42, pp. 36291.
- [COLGRAVE '04] M. L. Colgrave, D. J. Craik, Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. *Biochemistry*, 2004, **43**, 20, pp. 5965.
- [COLGRAVE '08a] M. L. Colgrave *et al.*, Cyclotides: natural, circular plant peptides that possess significant activity against gastrointestinal nematode parasites of sheep. *Biochemistry*, 2008a, **47**, 20, pp. 5581.
- [COLGRAVE '08b] M. L. Colgrave, A. C. Kotze, D. C. Ireland, C. K. Wang, D. J. Craik, The anthelmintic activity of the cyclotides: natural variants with enhanced activity. *Chembiochem : a European journal of chemical biology*, 2008b, **9**, 12, pp. 1939.
- [COLGRAVE '09] M. L. Colgrave *et al.*, Anthelmintic activity of cyclotides: In vitro studies with canine and human hookworms. *Acta tropica*, 2009, **109**, 2, pp. 163.
- [COOK '95] D. Cook *et al.*, Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. *The Plant cell*, 1995, **7**, 1, pp. 43.
- [COOK '97] D. R. Cook, K. VandenBosch, F. J. de Bruijn, T. Huguet, Model Legumes Get the Nod. *The Plant cell*, 1997, **9**, 3, pp. 275.
- [CRAIG '99] A. G. Craig, P. Bandyopadhyay, B. M. Olivera, Post-translationally modified neuropeptides from *Conus* venoms. *European journal of biochemistry / FEBS*, 1999, **264**, 2, pp. 271.
- [CRAIK '01a] D. Craik *et al.*, Discovery and structures of the cyclotides: novel macrocyclic peptides from plants. *Lett Pept Sci*, 2001a, **8**, 3-5, pp. 119.
- [CRAIK '10] D. J. Craik, Discovery and applications of the plant cyclotides. *Toxicon : official journal of the International Society on Toxinology*, 2010, **56**, 7, pp. 1092.
- [CRAIK '07] D. J. Craik, M. Cemazar, N. L. Daly, The chemistry and biology of cyclotides. *Current opinion in drug discovery & development*, 2007, **10**, 2, pp. 176.
- [CRAIK '06] D. J. Craik, M. Cemazar, C. K. Wang, N. L. Daly, The cyclotide family of circular miniproteins: nature's combinatorial peptide template. *Biopolymers*, 2006, **84**, 3, pp. 250.

- [CRAIK '99] D. J. Craik, N. L. Daly, T. Bond, C. Waine, Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *Journal of molecular biology*, 1999, **294**, 5, pp. 1327.
- [CRAIK '01b] D. J. Craik, N. L. Daly, C. Waine, The cystine knot motif in toxins and implications for drug design. *Toxicon : official journal of the International Society on Toxinology*, 2001b, **39**, 1, pp. 43.
- [CRAWFORD '89] E. J. Crawford, A.W.H. Lake, K.G. Boyce., Breeding annual Medicago species for semiarid conditions in southern Australia. *Advances in Agronomy*, 1989, **42**, pp. 399.
- [DA SILVA '10] P. Da Silva *et al.*, Molecular requirements for the insecticidal activity of the plant peptide pea albumin 1 subunit b (PA1b). *The Journal of biological chemistry*, 2010, **285**, 43, pp. 32689.
- [DA SILVA '09] P. Da Silva *et al.*, A folded and functional synthetic PA1b: an interlocked entomotoxic miniprotein. *Biopolymers*, 2009, **92**, 5, pp. 436.
- [DALY '01] D. C. Daly, K. M. Cameron, D. W. Stevenson, Plant systematics in the age of genomics. *Plant physiology*, 2001, **127**, 4, pp. 1328.
- [DALY '11a] N. L. Daly, D. J. Craik, Bioactive cystine knot proteins. *Curr Opin Chem Biol*, 2011a, **15**, 3, pp. 362.
- [DALY '09] N. L. Daly, K. J. Rosengren, D. J. Craik, Discovery, structure and biological activities of cyclotides. *Advanced drug delivery reviews*, 2009, **61**, 11, pp. 918.
- [DALY '11b] N. L. Daly, K. J. Rosengren, S. T. Henriques, D. J. Craik, NMR and protein structure in drug design: application to cyclotides and conotoxins. *European biophysics journal : EBJ*, 2011b, **40**, 4, pp. 359.
- [DAWSON '00] P. E. Dawson, S. B. Kent, Synthesis of native proteins by chemical ligation. *Annual review of biochemistry*, 2000, **69**, pp. 923.
- [DECOIN '12] d. a. L. Decoin, fleurat-Lessard, Insectes des stocks de blé, l'enquête et ses moralités. *Phytoma*, 2012, **656**.
- [DELOBEL '98] B. G. Delobel, A.M.; Gueguen, J.; Ferrasson, E.; Mbaiguinam, M. France, 1998, vol. 98/05877
- [DEVAUCHELLE '04] Devauchelle, Un virus d'insecte au service des hommes. *Virologie*, 2004, **8**, 3, pp. 171.
- [DEVAUCHELLE '93] Devauchelle, Cerutti, *Les Baculovirus d'insectes vecteurs d'expression de genes e;trangers*. Société française de biochimie et biologie moléculaire, Châtenay-Malabry, FRANCE, 1993.
- [DEVAUCHELLE '03] Devauchelle, Cerutti, in *XIVème Colloque de Physiologie de l'Insecte*. Amiens, 14-16 avril 2003, 2003.
- [DODET '90] Dodet, *Production de proteines recombinantes : quel systeme choisir?*, Lavoisier, Cachan, FRANCE, 1990, pp. 10.
- [EALING '94] P. M. Ealing, K. R. Hancock, D. W. White, Expression of the pea albumin 1 gene in transgenic white clover and tobacco. *Transgenic research*, 1994, **3**, 6, pp. 344.
- [ERVEL '07] F. Ervel. ZDNet, 2007.
- [ESCOUBAS '03] P. Escoubas, C. Bernard, G. Lambeau, M. Lazdunski, H. Darbon, Recombinant production and solution structure of PcTx1, the specific peptide inhibitor of ASIC1a proton-gated cation channels. *Protein science : a publication of the Protein Society*, 2003, **12**, 7, pp. 1332.

- [ESPOSITO '06] D. Esposito, D. K. Chatterjee, Enhancement of soluble protein expression through the use of fusion tags. *Current opinion in biotechnology*, 2006, **17**, 4, pp. 353.
- [EYRAUD '13] V. Eyraud *et al.*, Expression and biological activity of the cystine knot bioinsecticide PA1b (Pea Albumin 1 subunit b). *Plos One.*, 2013.
- [FAHNERT '04] B. Fahnert, H. Lilie, P. Neubauer, Inclusion bodies: formation and utilisation. *Advances in biochemical engineering/biotechnology*, 2004, **89**, pp. 93.
- [FARIA '89] S. D. Faria, G. P. Lewis, J. I. Sprent, J. M. Sutherland, Occurrence of nodulation in the leguminosae. *New Phytologist*, 1989, **111**, pp. 607.
- [FORCE '99] A. Force *et al.*, Preservation of Duplicate Genes by Complementary, Degenerative Mutations. *Genetics*, 1999, **151**, 4, pp. 1531.
- [FREDERICK M. AUSUBEL '03] Frederick M. Ausubel *et al.*, *Molecular Biology - Current Protocols*. John Wiley & Sons, Inc., 2003.
- [FREELING '06] M. Freeling, B. C. Thomas, Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome research*, 2006, **16**, 7, pp. 805.
- [FRUGOLI '01] J. Frugoli, J. Harris, Medicago truncatula on the Move! *The Plant cell*, 2001, **13**, 3, pp. 458.
- [GALES '07] A. L. Galesi, C. A. Pereira, A. M. Moraes, Culture of transgenic *Drosophila melanogaster* Schneider 2 cells in serum-free media based on TC100 basal medium. *Biotechnology journal*, 2007, **2**, 11, pp. 1399.
- [GANZ '03] T. Ganz, Defensins: antimicrobial peptides of innate immunity. *Nature reviews. Immunology*, 2003, **3**, 9, pp. 710.
- [GASCUEL '97] O. Gascuel, BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol*, 1997, **14**, 7, pp. 685.
- [GEORGIOU '96] G. Georgiou, P. Valax, Expression of correctly folded proteins in *Escherichia coli*. *Current opinion in biotechnology*, 1996, **7**, 2, pp. 190.
- [GILLON '08] A. D. Gillon *et al.*, Biosynthesis of circular proteins in plants. *The Plant journal : for cell and molecular biology*, 2008, **53**, 3, pp. 505.
- [GORANSSON '12] U. Goransson, R. Burman, S. Gunasekera, A. A. Stromstedt, K. J. Rosengren, Circular proteins from plants and fungi. *The Journal of biological chemistry*, 2012, **287**, 32, pp. 27001.
- [GOTTESMAN '96] S. Gottesman, Proteases and their targets in *Escherichia coli*. *Annual review of genetics*, 1996, **30**, pp. 465.
- [GOUY '10] M. Gouy, S. Guindon, O. Gascuel, SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, 2010, **27**, 2, pp. 221.
- [GRAIN '73] L. Grain, Isolation of oxytocic peptides from *Oldenlandia affinis* by solvent extraction of tetraphenylborate complexes and chromatography on sephadex LH-20. *Lloydia*, 1973, **36**, 2, pp. 207.
- [GRAN '73] L. Gran, On the effect of a polypeptide isolated from "Kalata-Kalata" (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. *Acta pharmacologica et toxicologica*, 1973, **33**, 5, pp. 400.
- [GREENWOOD '07] K. P. Greenwood, N. L. Daly, D. L. Brown, J. L. Stow, D. J. Craik, The cyclic cystine knot miniprotein MCoTI-II is internalized into cells by macropinocytosis. *The international journal of biochemistry & cell biology*, 2007, **39**, 12, pp. 2252.

- [GRESENT '11] F. Gressent, P. Da Silva, V. Eyraud, L. Karaki, C. Royer, Pea Albumin 1 subunit b (PA1b), a promising bioinsecticide of plant origin. *Toxins*, 2011, **3**, 12, pp. 1502.
- [GRESENT '07] F. Gressent *et al.*, Biological Activity and Binding Site Characteristics of the PA1b Entomotoxin on Insects from Different Orders. *Journal of Insect Science*, 2007, **7**.
- [GRESENT '03] F. Gressent, I. Rahioui, Y. Rahbe, Characterization of a high-affinity binding site for the pea albumin 1b entomotoxin in the weevil *Sitophilus*. *European journal of biochemistry / FEBS*, 2003, **270**, 11, pp. 2429.
- [GRUBER '08] C. W. Gruber *et al.*, Distribution and evolution of circular miniproteins in flowering plants. *The Plant cell*, 2008, **20**, 9, pp. 2471.
- [GUÈYE '11] M. T. Guèye, D. Seck, J.-P. Wathélet, G. Lognay, Lutte contre les ravageurs des stocks de céréales et de légumineuses au Sénégal et en Afrique occidentale : synthèse bibliographique. *Biotechnol. Agron. Soc. Environ.*, 2011, **15**, 1, pp. 183.
- [GUILLAUMOT '05] L. Guillaumot. Institut Pasteur de nouvelle-Calédonie, 2005.
- [GUNASEKERA '08] S. Gunasekera *et al.*, Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. *Journal of medicinal chemistry*, 2008, **51**, 24, pp. 7697.
- [GUSTAFSON '04] K. R. Gustafson, T. C. McKee, H. R. Bokesch, Anti-HIV cyclotides. *Current protein & peptide science*, 2004, **5**, 5, pp. 331.
- [GUSTAFSON '00] K. R. Gustafson *et al.*, New circulin macrocyclic polypeptides from *Chassalia parvifolia*. *Journal of natural products*, 2000, **63**, 2, pp. 176.
- [HABIB '07] H. Habib, K. M. Fazili, Plant protease inhibitors: a defense strategy in plants. *Biotechnology and Molecular Biology Review*, 2007, **2**, 3, pp. 068.
- [HANADA '03] K. Hanada, Y. Nishiuchi, H. Hirano, Amino acid residues on the surface of soybean 4-kDa peptide involved in the interaction with its binding protein. *European journal of biochemistry / FEBS*, 2003, **270**, 12, pp. 2583.
- [HANCOCK '94] K. R. Hancock, P. M. Ealing, D. W. White, Identification of sulphur-rich proteins which resist rumen degradation and are hydrolysed rapidly by intestinal proteases. *The British journal of nutrition*, 1994, **72**, 6, pp. 855.
- [HARRISON '00] M. J. Harrison, Molecular Genetics of Model Legumes. *Trends Plant Sci.*, 2000, **5**, 414–415.
- [HARRISON '94] M. J. Harrison, R. A. Dixon, Spatial patterns of expression of flavonoid/isoflavonoid pathway genes during interactions between roots of *Medicago truncatula* and the mycorrhizal fungus *Glomus versiforme*. *The Plant Journal*, 1994, **6**, 1, pp. 9.
- [HEITZ '89] A. Heitz, L. Chiche, D. Le-Nguyen, B. Castro, ¹H ²D NMR and distance geometry study of the folding of *Ecballium elaterium* trypsin inhibitor, a member of the squash inhibitors family. *Biochemistry*, 1989, **28**, 6, pp. 2392.
- [HIGGINS '86] T. J. Higgins *et al.*, Gene structure, protein structure, and regulation of the synthesis of a sulfur-rich protein in pea seeds. *The Journal of biological chemistry*, 1986, **261**, 24, pp. 11124.
- [HOU '03] X. Hou, P. G. Fields, Granary trial of protein-enriched pea flour for the control of three stored-product insects in barley. *J Econ Entomol*, 2003, **96**, 3, pp. 1005.
- [HOU '06] X. Hou, W. Taylor, P. Fields, Effect of pea flour and pea flour extracts on *Sitophilus oryzae*. *The Canadian Entomologist*, 2006, **138**, 01, pp. 95.

- [IRELAND '06] D. C. Ireland, M. L. Colgrave, D. J. Craik, A novel suite of cyclotides from *Viola odorata*: sequence variation and the implications for structure, function and stability. *The Biochemical journal*, 2006, **400**, 1, pp. 1.
- [IRELAND '08] D. C. Ireland, C. K. Wang, J. A. Wilson, K. R. Gustafson, D. J. Craik, Cyclotides as natural anti-HIV agents. *Biopolymers*, 2008, **90**, 1, pp. 51.
- [IYER '11] S. Iyer, K. R. Acharya, Tying the knot: The cystine signature and molecular-recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. *FEBS Journal*, 2011, **278**, 22, pp. 4304.
- [JEAN-MARIE TREMBLAY '68] Jean-Marie Tremblay, *Friedrich Engels 1883, Dialectique de la nature*. Paris, ed. Éditions sociales, 1968, pp. 367.
- [JENNINGS '01] C. Jennings, J. West, C. Waine, D. Craik, M. Anderson, Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proceedings of the National Academy of Sciences of the United States of America*, 2001, **98**, 19, pp. 10614.
- [JI '05] W. Ji *et al.*, Expression and purification of Huwentoxin-I in baculovirus system. *Protein Expr Purif*, 2005, **41**, 2, pp. 454.
- [JMP '89-'] v. JMP, SAS Institute Inc., Cary, NC. 1989-2007.
- [JOUVENSALE '03] L. Jouvensal *et al.*, PA1b, an insecticidal protein extracted from pea seeds (*Pisum sativum*): 1H-2-D NMR study and molecular modeling. *Biochemistry*, 2003, **42**, 41, pp. 11915.
- [KAPUST '99] R. B. Kapust, D. S. Waugh, Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein science : a publication of the Protein Society*, 1999, **8**, 8, pp. 1668.
- [KING '92] King, Possee, *The baculovirus expression system : a laboratory guide*. London ; New York, N.Y. : Chapman & Hall, 1992., England, ed. 1st ed., 1992.
- [KOLMAR '08] H. Kolmar, Alternative binding proteins: biological activity and therapeutic potential of cystine-knot miniproteins. *The FEBS journal*, 2008, **275**, 11, pp. 2684.
- [KRATZNER '05] R. Kratzner *et al.*, Structure of Ecballium elaterium trypsin inhibitor II (EETI-II): a rigid molecular scaffold. *Acta crystallographica. Section D, Biological crystallography*, 2005, **61**, Pt 9, pp. 1255.
- [KRISTENSEN '99] A. K. Kristensen *et al.*, Processing, disulfide pattern, and biological activity of a sugar beet defensin, AX2, expressed in *Pichia pastoris*. *Protein Expr Purif*, 1999, **16**, 3, pp. 377.
- [LAVALLIE '93] E. R. LaVallie *et al.*, A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/technology (Nature Publishing Company)*, 1993, **11**, 2, pp. 187.
- [LE-NGUYEN '90] D. Le-Nguyen *et al.*, Molecular recognition between serine proteases and new bioactive microproteins with a knotted structure. *Biochimie*, 1990, **72**, 6-7, pp. 431.
- [LE-NGUYEN '89] D. Le-Nguyen, D. Nalis, B. Castro, Solid phase synthesis of a trypsin inhibitor isolated from the Cucurbitaceae *Ecballium elaterium*. *International Journal of Peptide and Protein Research*, 1989, **34**, 6, pp. 492.
- [LECOQ '96] Lecoq, Les pucerons: de redoutables vecteurs de virus des plantes. *PHM Revue horticole*, 1996, **369**, pp. 25.
- [LOUIS '04a] S. Louis, Diversité structurale et d'activité biologique des Albumines entomotoxiques de type 1b des graines de Légumineuses. L'Institut National des Sciences Appliquées de Lyon, 2004a.

- [LOUIS '07] S. Louis *et al.*, Broad screening of the legume family for variability in seed insecticidal activities and for the occurrence of the A1b-like knottin peptide entomotoxins. *Phytochemistry*, 2007, **68**, 4, pp. 521.
- [LOUIS '04b] S. Louis *et al.*, Molecular and biological screening for insect-toxic seed albumins from four legume species. *Plant Science*, 2004b, **167**, 4, pp. 705.
- [LYNCH '00a] M. Lynch, J. S. Conery, The Evolutionary Fate and Consequences of Duplicate Genes. *Science*, 2000a, **290**, 5494, pp. 1151.
- [LYNCH '00b] M. Lynch, A. Force, The probability of duplicate gene preservation by subfunctionalization. *Genetics*, 2000b, **154**, 1, pp. 459.
- [LYNCH '01] M. Lynch, M. O'Hely, B. Walsh, A. Force, The probability of preservation of a newly arisen gene duplicate. *Genetics*, 2001, **159**, 4, pp. 1789.
- [MASKOS '96] K. Maskos, M. Huber-Wunderlich, R. Glockshuber, RBI, a one-domain alpha-amylase/trypsin inhibitor with completely independent binding sites. *FEBS Lett*, 1996, **397**, 1, pp. 11.
- [MAURIZI '92] M. R. Maurizi, Proteases and protein degradation in Escherichia coli. *Experientia*, 1992, **48**, 2, pp. 178.
- [MERRIFIELD '63] R. B. Merrifield, Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society*, 1963, **85**, 14, pp. 2149.
- [MICHAEL HASS '81] G. Michael Hass, C. A. Ryan. Carboxypeptidase inhibitor from potatoes. in *Methods in Enzymology*, L. Laszlo, Ed. Academic Press, 1981, vol. Volume 80, pp. 778-791.
- [MONTEIRO '12] F. Monteiro, N. Carinhas, M. J. T. Carrondo, V. Bernal, P. M. Alves, Towards system-level understanding of baculovirus host cell interactions: from molecular fundamental studies to large-scale proteomics approaches. *Frontiers in Microbiology*, 2012, **3**.
- [MOORE '05] R. C. Moore, M. D. Purugganan, The evolutionary dynamics of plant duplicate genes. *Current opinion in plant biology*, 2005, **8**, 2, pp. 122.
- [MULVENNA '06] J. P. Mulvenna *et al.*, Discovery of cyclotide-like protein sequences in graminaceous crop plants: ancestral precursors of circular proteins? *The Plant cell*, 2006, **18**, 9, pp. 2134.
- [MURRAY-RUST '93] J. Murray-Rust *et al.*, Topological similarities in TGF-beta 2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure (London, England : 1993)*, 1993, **1**, 2, pp. 153.
- [NGAMO '07] L. S. T. Ngamo, T. Hance, Diversité des ravageurs des denrées et méthodes alternatives de lutte en milieu tropical. *TROPICULTURA*, 2007, **25**, 4, pp. 215.
- [NISHIHARA '98] K. Nishihara, M. Kanemori, M. Kitagawa, H. Yanagi, T. Yura, Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli. *Applied and environmental microbiology*, 1998, **64**, 5, pp. 1694.
- [NOLAN '89] K. E. Nolan, R. J. Rose, J. R. Gorst, Regeneration of Medicago truncatula from tissue culture: increased somatic embryogenesis using explants from regenerated plants. *Plant Cell Reports*, 1989, **8**, 5, pp. 278.
- [NORTON '98] R. S. Norton, P. K. Pallaghy, The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon : official journal of the International Society on Toxinology*, 1998, **36**, 11, pp. 1573.
- [NOWAK '97] M. A. Nowak, M. C. Boerlijst, J. Cooke, J. M. Smith, Evolution of genetic redundancy. *Nature*, 1997, **388**, 6638, pp. 167.

- [**NYGREN '94**] P. A. Nygren, S. Stahl, M. Uhlen, Engineering proteins to facilitate bioprocessing. *Trends in biotechnology*, 1994, **12**, 5, pp. 184.
- [**OHNO '70**] S. Ohno, Evolution by gene duplication. *George Allen and Unwin, London*, 1970.
- [**OSBORNI '88**] T. C. Osborni, D. C. Alexander, S. S. Sun, C. Cardona, F. A. Bliss, Insecticidal activity and lectin homology of arcelin seed protein. *Science*, 1988, **240**, 4849, pp. 207.
- [**PALLAGHY '94**] P. K. Pallaghy, K. J. Nielsen, D. J. Craik, R. S. Norton, A common structural motif incorporating a cystine knot and a triple-stranded beta-sheet in toxic and inhibitory polypeptides. *Protein science : a publication of the Protein Society*, 1994, **3**, 10, pp. 1833.
- [**PETERSEN '11**] T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods*, 2011, **8**, 10, pp. 785.
- [**PETI '07**] W. Peti, R. Page, Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. *Protein Expr Purif*, 2007, **51**, 1, pp. 1.
- [**PETIT '06**] J. Petit, Etude structure/fonction d'une albumine entomotoxique de type A1b du pois chez le riz : application a la protection contre le ravageur des stocks Sitophilus oryzae., Université Montpellier II Sciences et techniques du Languedoc, 2006.
- [**PETIT '05**] J. Petit, G. Duport, Y. Rahbé, E. Guiderdoni, J. C. Breitler, paper presented at the 5th International Rice Genetics Symposium & 3rd International Rice Functional Genomics Symposium, 2005-00-00
2005 2005.
- [**PINES '99**] O. Pines, M. Inouye, Expression and secretion of proteins in E. coli. *Molecular biotechnology*, 1999, **12**, 1, pp. 25.
- [**PLAN '08**] M. R. Plan, I. Saska, A. G. Cagauan, D. J. Craik, Backbone cyclised peptides from plants show molluscicidal activity against the rice pest Pomacea canaliculata (golden apple snail). *Journal of agricultural and food chemistry*, 2008, **56**, 13, pp. 5237.
- [**POSSEE '93**] Possee, Hirst, Jones, Bishop. Opportunities for Molecular Biology in Crop Production. in *BCPC Monogram*, D. J. Beadle, Bishop, D.H.L., Copping., D. L.G., G.K. and Hollomon, D.W., Eds. The British Crop Protection Council, Farnham, Surrey., 1993, vol. 55.
- [**PRETHEEP-KUMAR '04**] P. Pretheep-Kumar, S. Mohan, K. Ramaraju, Protein-enriched pea flour extract protects stored milled rice against the rice weevil, Sitophilus oryzae. *Journal of Insect Science*, 2004, **4**, pp. 26.
- [**PRINZ '97**] W. A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm. *The Journal of biological chemistry*, 1997, **272**, 25, pp. 15661.
- [**RANKIN '88**] C. Rankin, B. G. Ooi, L. K. Miller, Eight base pairs encompassing the transcriptional start point are the major determinant for baculovirus polyhedrin gene expression. *Gene*, 1988, **70**, 1, pp. 39.
- [**REES '04**] D. Rees, *Insects of Stored Products*. CSIRO, 2004, pp. 183.
- [**REES '80**] D. C. Rees, W. N. Lipscomb, Structure of the potato inhibitor complex of carboxypeptidase A at 2.5-A resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 1980, **77**, 8, pp. 4633.

- [REES '82] D. C. Rees, W. N. Lipscomb, Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution. *Journal of molecular biology*, 1982, **160**, 3, pp. 475.
- [RIJPKEMA '07] A. S. Rijpkema, T. Gerats, M. Vandenbussche, Evolutionary complexity of MADS complexes. *Current opinion in plant biology*, 2007, **10**, 1, pp. 32.
- [ROHRMANN '86] G. F. Rohrmann, Polyhedrin structure. *The Journal of general virology*, 1986, **67** (Pt 8), pp. 1499.
- [ROHRMANN '92] G. F. Rohrmann, Baculovirus structural proteins. *The Journal of general virology*, 1992, **73** (Pt 4), pp. 749.
- [ROSE '99] R. J. Rose, K. E. Nolan, L. Bicego, The Development of the Highly Regenerable Seed Line Jemalong 2HA for Transformation of *Medicago truncatula* — Implications for Regenerability via Somatic Embryogenesis. *Journal of Plant Physiology*, 1999, **155**, 6, pp. 788.
- [SAETHER '95] O. Saether *et al.*, Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. *Biochemistry*, 1995, **34**, 13, pp. 4147.
- [SAVOLDELLI '13] Savoldelli, paper presented at the 9th Conference on Integrated Protection of Stored Products IPSP 2013, Agora - University Bordeaux 1 TALENCE, FRANCE, 1 - 4 July, 2013 2013.
- [SCHAFFNER '01] J. Schaffner, J. Winter, R. Rudolph, E. Schwarz, Cosecretion of chaperones and low-molecular-size medium additives increases the yield of recombinant disulfide-bridged proteins. *Applied and environmental microbiology*, 2001, **67**, 9, pp. 3994.
- [SHAO '99] F. Shao *et al.*, A new antifungal peptide from the seeds of *Phytolacca americana*: characterization, amino acid sequence and cDNA cloning. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1999, **1430**, 2, pp. 262.
- [SHENKAREV '08] Z. O. Shenkarev *et al.*, Divalent cation coordination and mode of membrane interaction in cyclotides: NMR spatial structure of ternary complex Kalata B7/Mn²⁺/DPC micelle. *Journal of inorganic biochemistry*, 2008, **102**, 5-6, pp. 1246.
- [SMITH '83] G. E. Smith, M. D. Summers, M. J. Fraser, Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Molecular and Cellular Biology*, 1983, **3**, 12, pp. 2156.
- [TAM '99] J. P. Tam, Y. A. Lu, J. L. Yang, K. W. Chiu, An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proceedings of the National Academy of Sciences of the United States of America*, 1999, **96**, 16, pp. 8913.
- [TAYLOR '04] W. G. Taylor, D. H. Sutherland, D. J. Olson, A. R. Ross, P. G. Fields, Insecticidal components from field pea extracts: sequences of some variants of pea albumin 1b. *Journal of agricultural and food chemistry*, 2004, **52**, 25, pp. 7499.
- [TERLAU '04] H. Terlau, B. M. Olivera, Conus venoms: a rich source of novel ion channel-targeted peptides. *Physiological reviews*, 2004, **84**, 1, pp. 41.
- [TOGOLA '13] Togola, paper presented at the 9th Conference on Integrated Protection of Stored Products IPSP 2013, Agora - University Bordeaux 1 TALENCE, FRANCE, 1 - 4 July, 2013 2013.
- [TOHMÉ '07] G. Tohmé, H. Tohmé, Illustrated Flora of Lebanon. *NATIONAL COUNCIL FOR SCIENTIFIC RESEARCH*, 2007.

- [UDVARDI '01] M. K. Udvardi, Legume models strut their stuff. *Molecular plant-microbe interactions : MPMI*, 2001, **14**, 1, pp. 6.
- [VAIL '71] P. V. Vail, G. Sutter, D. L. Jay, D. Gough, Reciprocal infectivity of nuclear polyhedrosis viruses of the cabbage looper and alfalfa looper. *Journal of invertebrate pathology*, 1971, **17**, 3, pp. 383.
- [VAUGHN '77] J. L. Vaughn, R. H. Goodwin, G. J. Tompkins, P. McCawley, The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In vitro*, 1977, **13**, 4, pp. 213.
- [VETSCH '00] M. Vetsch, I. Janzik, A. Schaller, Characterization of prosystemin expressed in the baculovirus/insect cell system reveals biological activity of the systemin precursor. *Planta*, 2000, **211**, 1, pp. 91.
- [VLAK '82] J. M. Vlak, G. E. Smith, Orientation of the Genome of *Autographa californica* Nuclear Polyhedrosis Virus: a Proposal. *Journal of virology*, 1982, **41**, 3, pp. 1118.
- [VOGEL '04] C. W. Vogel, D. C. Fritzinger, B. E. Hew, M. Thorne, H. Bammert, Recombinant cobra venom factor. *Molecular immunology*, 2004, **41**, 2-3, pp. 191.
- [WALTER '03] G. H. Walter, *Insect Pest Management and Ecological Research*. Cambridge University Press, 2003.
- [WATANABE '94] Y. Watanabe *et al.*, A peptide that stimulates phosphorylation of the plant insulin-binding protein. Isolation, primary structure and cDNA cloning. *European journal of biochemistry / FEBS*, 1994, **224**, 1, pp. 167.
- [WENDEL '00] J. F. Wendel, Genome evolution in polyploids. *Plant molecular biology*, 2000, **42**, 1, pp. 225.
- [WERLE '07] M. Werle, K. Kafedjiiski, H. Kolmar, A. Bernkop-Schnurch, Evaluation and improvement of the properties of the novel cystine-knot microprotein McoEeTI for oral administration. *International journal of pharmaceuticals*, 2007, **332**, 1-2, pp. 72.
- [WHITE '94] C. E. White, N. M. Kempf, E. A. Komives, Expression of highly disulfide-bonded proteins in *Pichia pastoris*. *Structure (London, England : 1993)*, 1994, **2**, 11, pp. 1003.
- [WOJCIECHOWSKI '04] M. F. Wojciechowski, M. Lavin, M. J. Sanderson, A phylogeny of legumes (Leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. *American journal of botany*, 2004, **91**, 11, pp. 1846.
- [XUDONG MAI '06] C. W. Xudong Mai, Wei Wang, Xiaobo Li, Peptides from Plants: A New Source for Antitumor Drug Research. *Asian Journal of A New Source for Traditional Medicines*, 2006, pp. 6.
- [YAMAZAKI '03] T. Yamazaki *et al.*, A possible physiological function and the tertiary structure of a 4-kDa peptide in legumes. *European journal of biochemistry / FEBS*, 2003, **270**, 6, pp. 1269.
- [YOUNG '11] N. D. Young *et al.*, The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature*, 2011, **480**, 7378, pp. 520.
- [YOUNG '03] N. D. Young, J. Mudge, T. H. Ellis, Legume genomes: more than peas in a pod. *Current opinion in plant biology*, 2003, **6**, 2, pp. 199.
- [ZHANG '01] Zhang. 2001.
- [ZHANG '13] Y. Zhang *et al.*, Evolution of a horizontally acquired legume gene, albumin 1, in the parasitic plant *Phelipanche aegyptiaca* and related species. *BMC evolutionary biology*, 2013, **13**, pp. 48.

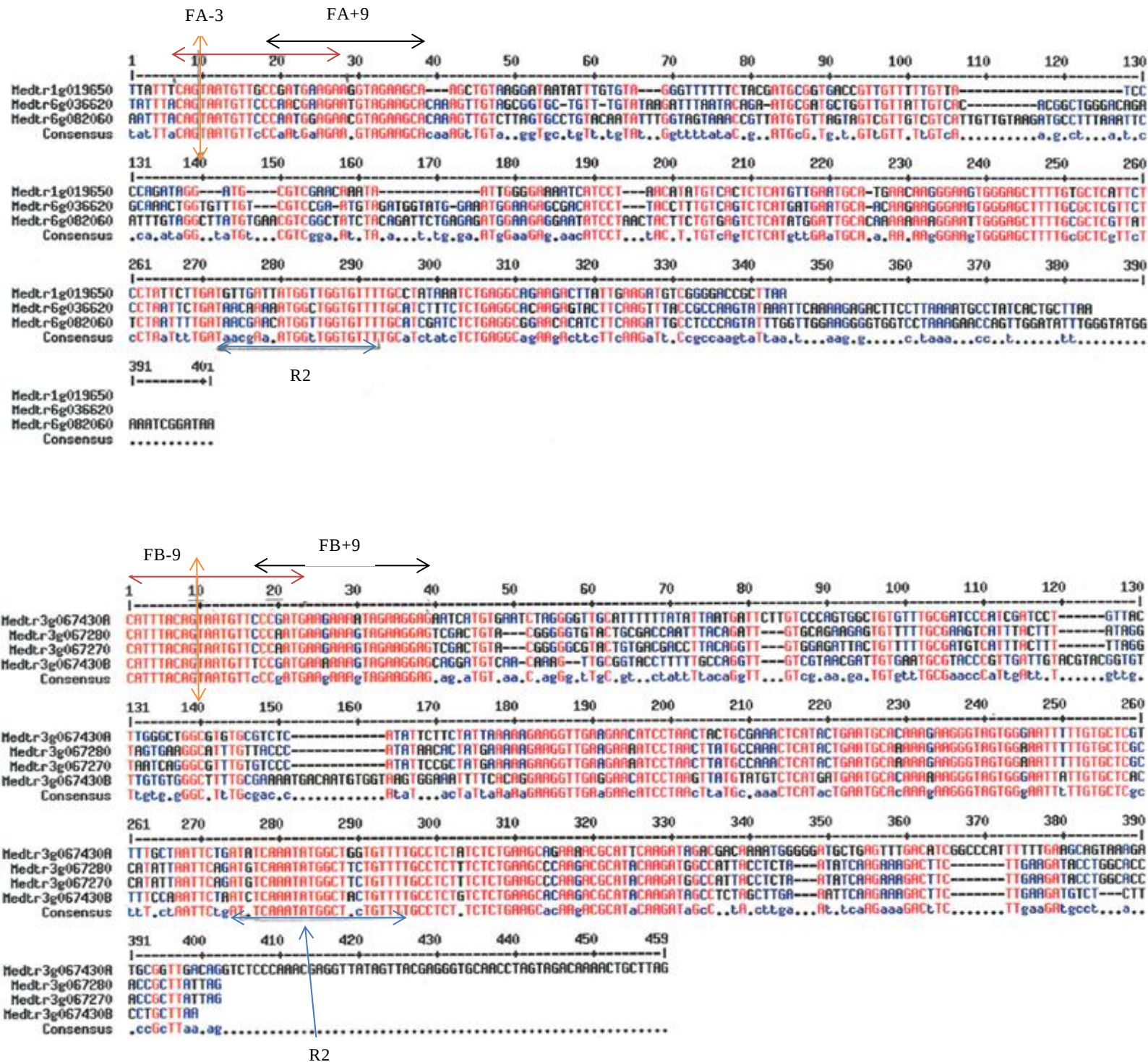
Annexes

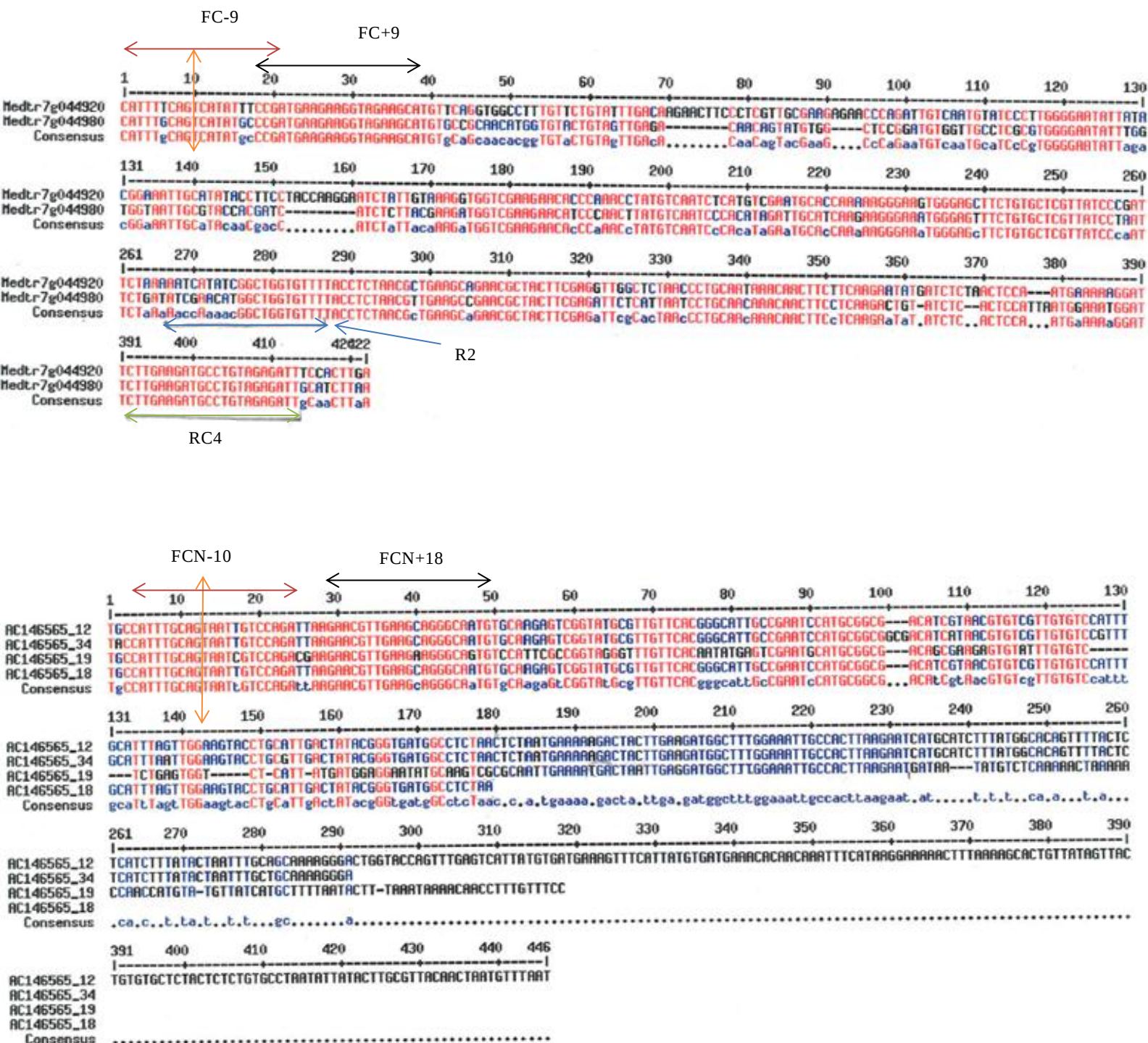
Annexe 1 : Liste totale des espèces endémiques sélectionnées du patrimoine végétale Libanais

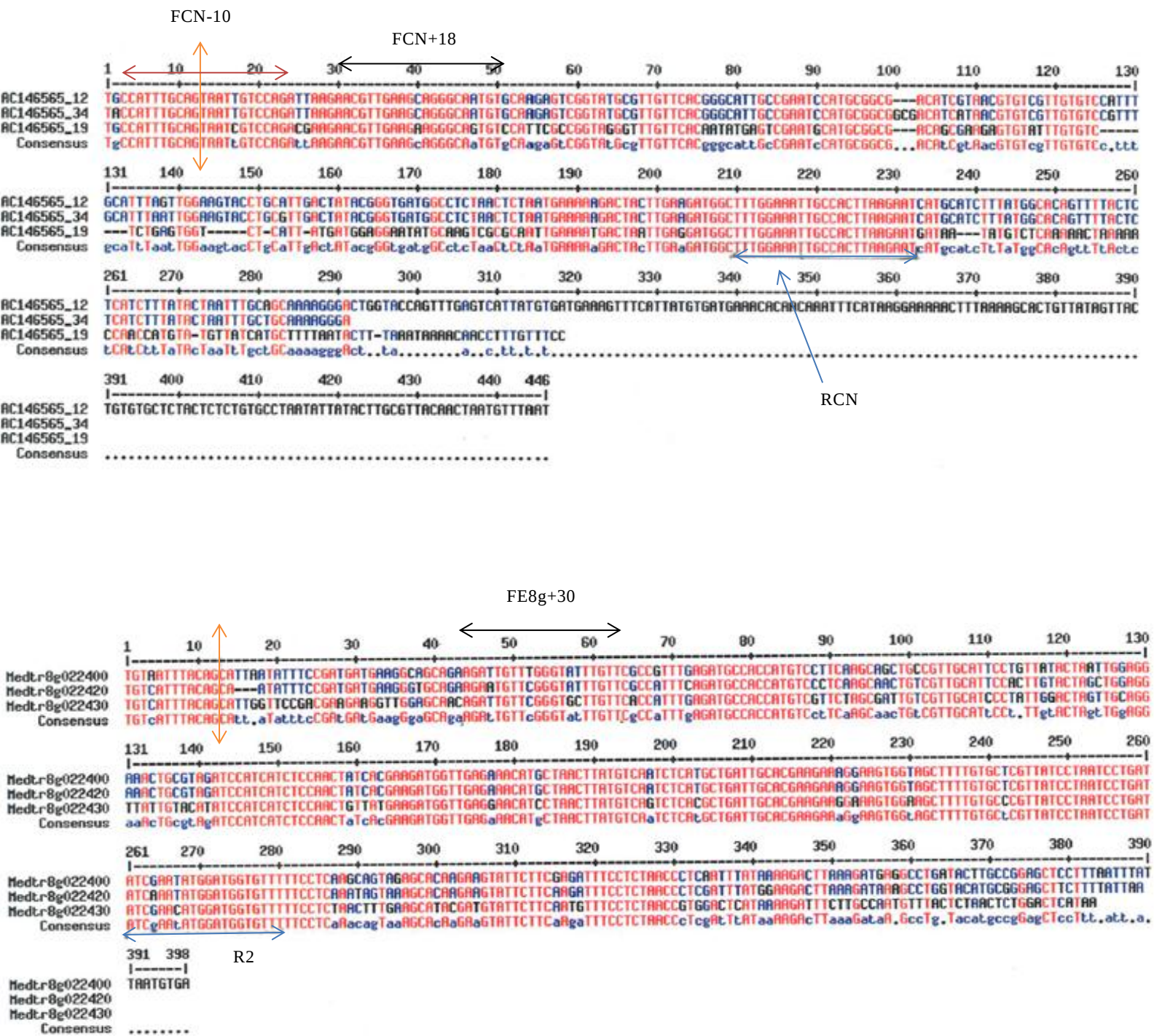
Famille	Sous-famille	Tribu	Genre	Espèces	Période de Floraison (mois)
Fabaceae	Papilionoideae	Galegeae	Astragalus	<i>angulosus</i>	5,6,7
				<i>ehdenensis</i>	5,6
				<i>hirsutissimus</i>	6,7
				<i>trichopterus</i>	6,7
				<i>cephalotes</i>	Donnée indisponible
				<i>zachlensis</i>	5,6,7,8
				<i>berytheus</i>	2,3
		Genisteae	Colutea	<i>cilicica</i>	4,5,6,7
			Spartium	<i>junceum</i>	Donnée indisponible
			Cytisus	<i>syriacus</i>	5,6,7,8
			Genista	<i>libanotica</i>	5,6,7
			Adenocarpus	<i>complicatus</i>	5,6,7
			Calycotome	<i>villosa</i>	1,2,3,4
		Trifolieae	Medicago	<i>globosa</i>	Donnée indisponible
				<i>rotata</i>	Donnée indisponible
				<i>radiata</i>	Donnée indisponible
			Trifolium	<i>angustifolium</i>	Donnée indisponible
				<i>billardieri</i>	Donnée indisponible
				<i>lagrangei</i>	Donnée indisponible
				<i>pilulare</i>	Donnée indisponible
				<i>scabrum</i>	Donnée indisponible
				<i>xerocephalum</i>	Donnée indisponible
			Ononis	<i>adenotricha</i>	6,7,8
	Vicieae	Lathyrus	Vicia	<i>libani</i>	5,6
				<i>canescens</i>	Donnée indisponible
		Cicereae	Cicer	<i>incisum</i>	6,7,8
	Loteae	Coronilla	Lotus	<i>varia libanotica</i>	5,6,7,8,9
				<i>angutissimus</i>	Donnée indisponible
	Caesalpinioideae	Phaseoleae	Lablab	<i>purpureus</i>	Donnée indisponible
		Cercideae	Cercis	<i>siliquastrum</i>	2,3,4

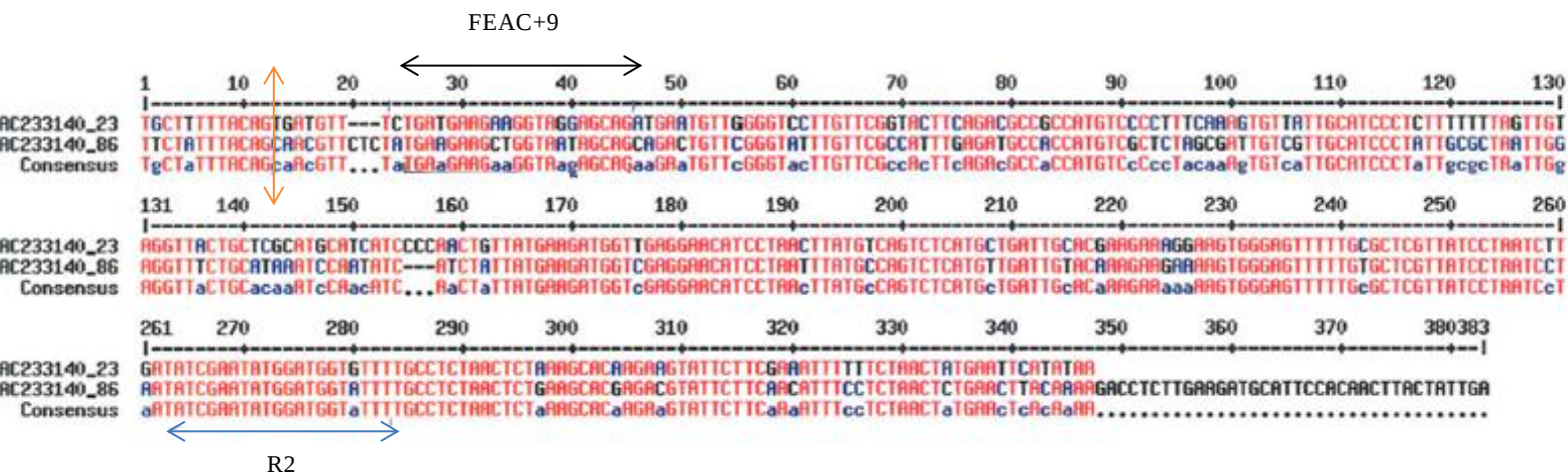
Annexe 2 Choix des amorces selon les différents clades phylogénétiques des A1bs de *Medicago truncatula* (génom version 3.5).

A noter que l'axe vertical est l'axe qui sépare la fin de l'intron et le début de l'exon 2.









Annexe 3: Résultats du séquençage des homologues de PA1b

A noter, les séquences trouvées sont bordées de part et d'autre par des sites de restriction de l'enzyme EcoRI (en rouge). Ce site se trouve dans le vecteur plasmidique pCR™ 2.1-TOPO (Invitrogen), vecteur dans lequel la séquence du gène a été insérée lors du clonage par homologie.

1_LABPU :

GAATTCGCCCTTCAGTAATGTTGCCGATGAAGAAGATA-
 GAAGCTGTAGACTGTTTCAGGTGCATGTTTCACCGTTT-
 GAGGTGCCACCATGCGGGTCACGTGATTGTCGCTGCATACC-
 TATTGCACTATTTGTTGGTTTTTGCATTTATCCAACCTGGACTTTTCAT
 CATCTGTGGGAAAGATGATAGATGAACATCCCAACTTATGT-
 GAATCTCATGAAGAATGCATGAAGAAAG-
 GAAGTGGAACCTTCTGTGCTCGTTATCCCAATCACTA-
 TATCGACCACGGATGGTGCTTAAGGGC**GAATTC**

2_ASTEH :

GAATTCGCCCTTAAGCACCAACCATGATCGATAT-
 TAGGATTAGGATAACGAGCA-
 CAAAAGCTTCCACTTCCTTTTTTCGTGCAGTCAGCAT-
 GAGACTGACATAAGTTAGGATGTTCTCAACCATCTTCATAGTAC
 GTCCAACCTGGATATATGCAAAAACCCACTACTAGTCCAATGGG-
 TACACAACGACAGTCTCTAGAGCGACATGGTGG-
 TATCTCAAATGGGGAACAAGCTCCCGAACAATCCACTGCTTCTAC
 CTTCTCCATCGGCAACATTACTGAAGGGC**GAATTC**

3_MEDRO :

GAATTCGCCCTTAAGCACCAR-
 CCGTGKTCGATATCAGGATTAGGATAACGAGCA-
 CAAAAGCTCCCACTTCCCTTCTTCGTGCAATCAGCAT-
 GAGACTGACATAAGTTAGGATGTTCTCAACCATCTTCATAACAG
 CTGGAGATGATGGACATCTACAATAACCTCCAAATATTTTCAT-
 TAGGGATACAACGACAGTCGCTAGAATGACATGG-
 TAGCATCTCAAATGGCGAACAAGCGCCCGAACAATATGTT-
 GCTCCTACCTTTTTTCATCGGAAACATTACTGTAAATGAAGGGC**GA**
ATTC

4_MEDRO :

GAATTCGCCCTTCCATTTGCAGTAATCGTCCA-
 GACGAAGAACGTTGAAGCAGGGCAATGTCCATCCGCCGGTAT-
 GATTTGTTCAACATTAAATCCAATGATGTGTGGCGACTACTTA-
 GAGTGTCAAGTGTGTCCCTTCGTCTGATTTTGGATCCGGCGGTATTT
 GCGGTGACATGAATTTGCACCTAACTCTAATGAAAAAGAG-

TACTTGAAGATGGCTTTGGAAATTGCCACTTAAGAA-
TAAGGGCGAATTC

5_MEDRO :

GAATTCGCCCTTAGAATGTTTGGGTGCTTGTTGCCATTT-
GAGATGCTACCATGTCATTCTAGCGACTGTCGTTGTATCCCTAAT-
GAAATATTTGGAGGTTATTGTAGATGTCCATCATCTCCAGCTGT-
TATGAAGATGGTTGAGGAACATCCTAACTTATGTCAGTCTCATGCT
GATTGCACGAAGAAGGGAAGTGGGAGCTTTTGTGCTCGT-
TATCCTAATCCTGATATCGACCATGGTTGG-
TGCTTAAGGGCGAATTC

6_ASTZA :

GAATTCGCCCTTAAGCAC-
CAGCCATGTTTCGATATCAGCATTAGGGTAACGAGCGCA-
GAAGCTTCCACTTTCCTTCTAGACACAGTCGTCATGAGACTTGCA-
TAAGTTAGGATGTTTCTCGACCATCTTTGTGACATGTTTCAGATCCA
GTTGGATGTATGCATTGACCCGCAAA-
TAAACCCCATGGGATGCAACGACAGTCAGCAGAACCACATGG-
TGGCATCTCGAATGGCGAACAAGCACCCAAACATTCTAAGGGCGA
ATTC

7_LABPU :

GAATTCGCCCTTCCGATGGA-
GAAGGTAGAAGCAGTACTTTGTTTCAGGTGTTTGTTCAACTTTT-
GAGTTTCCACCATGCGGGTCAACTGATTGTCGCTGTATCCCTATT-
GGAATATTTGTTGGTTTCTGCACTAATCCAAGTGGACTTTYATCTG
CGGCAAAGATRATAGATGAACATCCCAACTTATGTCAATCTCAT-
GATGAATGCATGAAWAAAGGAAGTGGAAACTTTTGTGCTCGT-
TATCCCAATCACTATATAGACCACGGTTGG-
TGCTTAAGGGCGAATTC

8_ASTAN :

GAATTCGCCCTTCCGACGGA-
GAATGTAGAAGCAACGGATTGTTTCGGGTGCTT-
GTTTCGCCATTCCAAATGCCACCATGTGGTTCTACCGATTGTCGTT-
GTATCCCATGGGGTTTACTTGTGGGTGAATGCAAATATCCAACTG
GATCTGAACGGGTTACAAAAATGGTCGA-
GAAACATCCCAACTTATGTCAGTCTCATGACGATTGTGTCCA-
GAAGGAAAGTGGAAAGCTTCTGTGCTCGTTACCCTAATGCT-
GATATCGASTACGGATGGTGCTTAAGGGCGAATTC

9_ASTZA :

GAATTCGCCCTTAATGGGTTACGCTGGGTTCGCATCTTT-
GGCTCTCTTCTGTTTGGCCACATTCTGTATGTCTATATTTA-
CAGCCTTCTATTCTTCTTTATTTCTCATTCATTCATGCAT-

TTCTTTTTCATGGGAAATCAGAATTAGTATA-
 TACTCTCAGTTAATCTATTGCATGCTTGACATTATGTATA-
 TATGTATTTATTCAAATTAATTAACCTTTGGCACTACCTTTGGATTT-
 GGCAGTGATGTTTCCGGCGAAGAATGTTGAAGGAACGGATTGTTC
 GGGTGCTTGTTCCGCCATTCGAGATGCCACCATGTGGTTCTGCT-
 GACTGTCGTTGCATCCCATGGGGTTTATTTGCGGGTCAATGCATA-
 CATCCAACCTGGATCTGAACATGTCACAAAGATGGTTCGA-
 GAAACATCCTAACTTATGCAAGTCTCATGACGACTGTGTCTAGAA
 GGAAAGTGGAAGCTTCTGCGCTCGTTACCCTAATGCT-
 GATATCGAATACGGTTGGTGCTTAAGGGC**GAATTC**

10_ASTTR :

GAATTCGCCCTTAAGCAC-
 CAGCCGTGTTTCGATATCAGCATTAGGG-
 TAACGAGCGCAAAAACCTCCACTTTCCTTCTGCACA-
 CAATCGGCATGAGACTGACATAAGTTAGGATGTTTTTCGACCATCT
 TGGTAACACGTTTCAGATCCAGTTGGATATTTGCATCCACCCA-
 CAAATAAACCCCATGGGATACAACGGCAATCAGTAGAGCCA-
 CATGG-
 TGGCATCTCAAATGGCGAACAACACCCGAACAATCCGCTGCTTC
 AACATTGTTTCGAGGGAACATCACTGCAAATCCAAAGGTGG-
 TGCCAAAGTTACTTTGAATAATTATACATAAATACATAAGCATA-
 GAAGAATAGATTAACCTGAGAGACAGAAAGAAGAA-
 TAGAAGGCTACAAACAATATAGACATACAGAATGTGGCAACCAA
 GAAAAGAGCCAAAGGGC**GAATTC**

11_ASTAN :

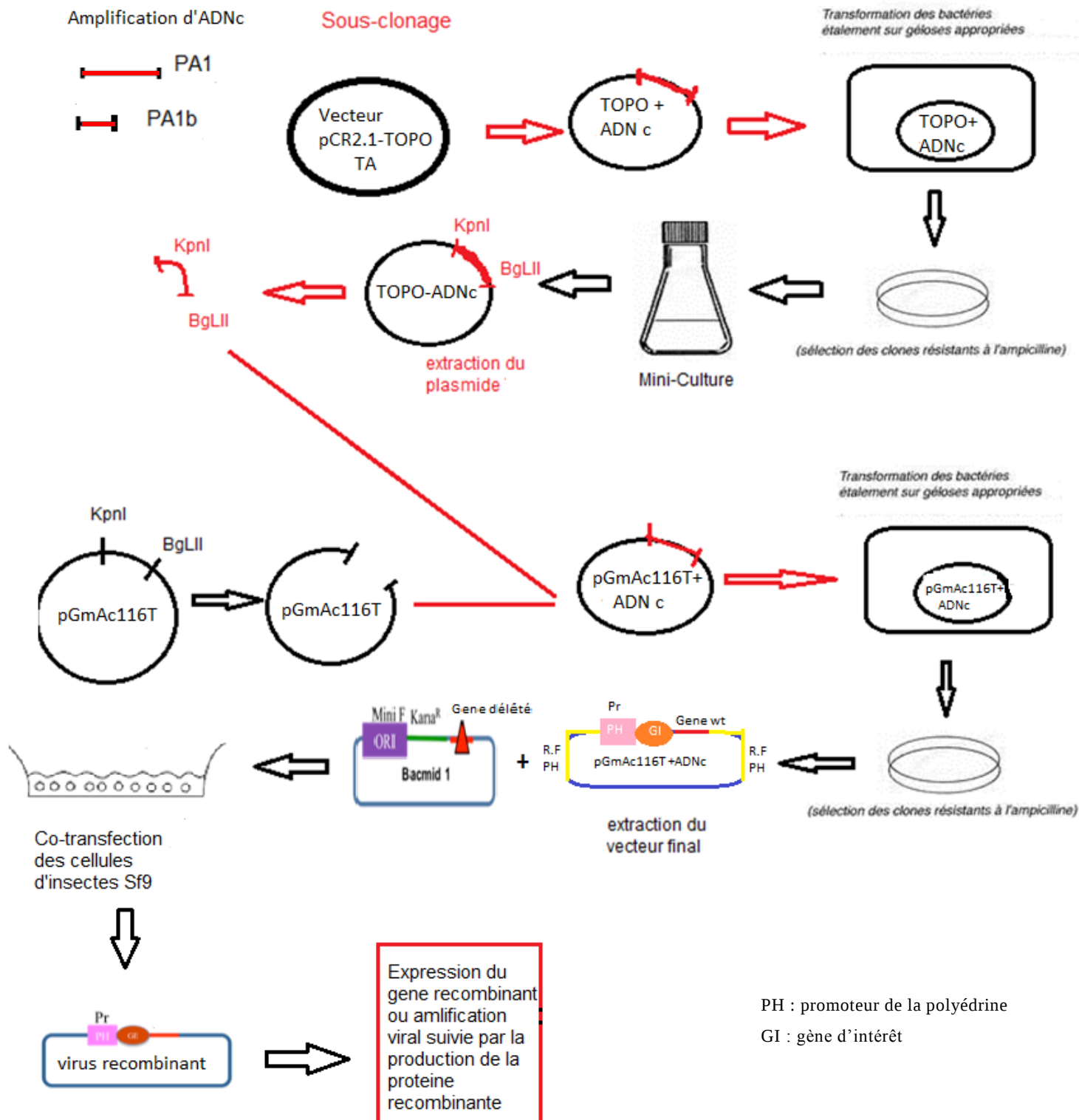
GAATTCGCCCTTAAGCACCAGCCATGG-
 TCGATATCAGCATTAGGGTAACGAGCACA-
 GAAGCTTCCACTTTCCTTCTGGACACAATCGTCATGAGACTGACA-
 TAAGTTGGGATGTTTCTCGACCATTTTTGTAAACCCGTTTCAGATCCA
 GTTGGATATTTGCATTACCCACAAGTAAACCCCATGGGATA-
 CAACGACAATCGGTAGAACCACATGGTGGCATT-
 GGAATGGCGAACAAGCACCCGAACAATCCGTT-
 GCTTCAACATTCTTCGCGGGGAACATCACTGCAAATCCAAAGGTG
 GTTGGTGCCAAAGTTAATTTAAATAAATACMTATATACA-
 CAATGTCAWGCATGCWATAGATCAATTGAGAG-
 TATMTCTRCKMGTYCYGYKTCCKTKSCTG

12_ASTZA :

GAATTCGCCCTTTGGCTCTATTCCTGCTTGCCACATTCTG-
 TATGTCTATATTTGCAGCCTTCTATTCTTCTT-
 TATTTCTCATCATTCATGCATTTCTATTTTCATGGGAAATCAGAAT-
 TAGTATATACTCTCAATTAATCTGTTGCATGCTTGACATATTATGT

ATATATGTATTTATTCAAATTAACCTTTGGTACCACCTTTGGATTT-
GCAGTGATGTTCCCCGCGAAGAATGTTGAAGCATGTTTCGGG-
TGCTTGTTCGCCATTTGAGATGCCAC-
CATGTGGCTCTACCGACTGTCGTTGTGTCCCATGGGGTTTATTTGC
GGGTCAATGCATAAATCCAACCTGGATCTGAACGTGTTA-
CAAAGATGGTTGAGAAACATCCTAACTTATGTCAGTCTCAT-
GACGATTGTGTTCAGAAGGAAAGTGGAAGTTTCTGCGCTCGT-
TACCCTAATGCTGATATCGAATACGGGTGGTGCTTAAGGGC**GAAT**
TC

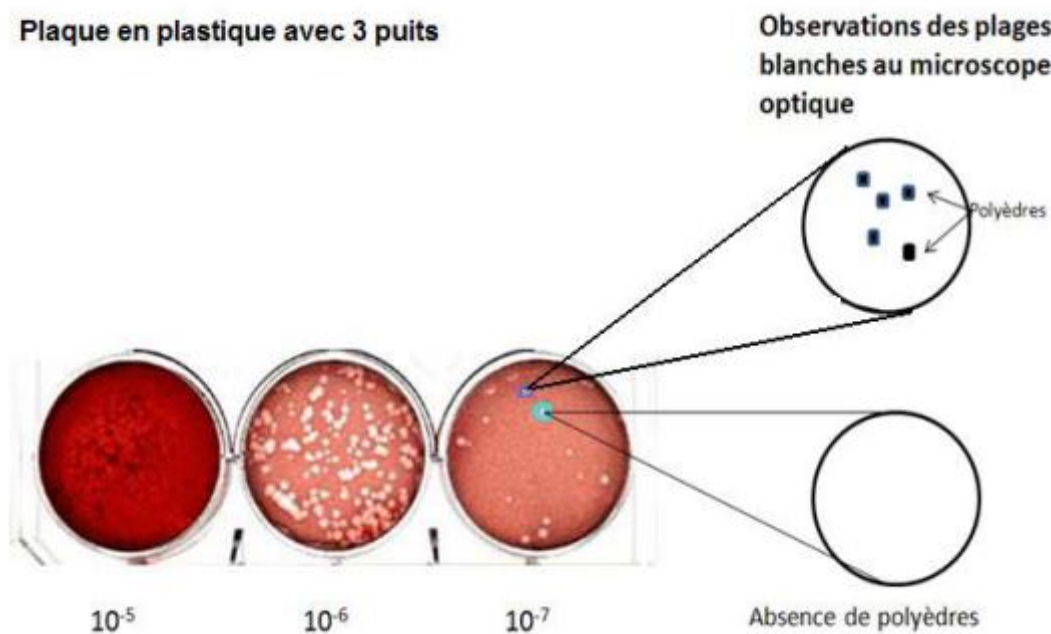
Annexe 4 : Etape de production de la protéine recombinante par le système baculovirus/cellules d'insectes



Annexe 5 Sélection des baculovirus recombinants par la technique des plages de lyse

Etapes :

- Utilisation d'une plaque en plastique formée de puits tapissés d'une culture de cellules d'insecte (formation d'un tapis cellulaire)
 - Ajout d'une suspension de virus obtenus par phénomène de recombinaison homologe. Différentes dilutions de la suspension virale sont testées
 - Ajout d'une couche d'agarose à 2,5 %
 - Après 5 jours à 28°C en atmosphère humide, ajout d'un colorant vital, le rouge neutre, qui ne se concentre que dans les cellules vivantes.
- Observation des plages blanches pour rechercher les virus recombinant



Annexe 6 Le Réactif CellTiter-Blue®

Le réactif CellTiter-Blue® est une solution tamponnée contenant du résazurine hautement purifié. Les ingrédients ont été optimisés pour une utilisation pour des tests de viabilité cellulaire.

Les propriétés du réactif CellTiter-Blue® changent lors d'une réduction de la résazurine en résorufine (Figure 35). La résazurine est de couleur bleu foncé et peu fluorescent jusqu'à ce qu'il soit réduit en résorufine, qui est rose et très fluorescent ($579_{Ex}/584_{Em}$). L'absorption maximale de la résazurine est de 605nm et de la résorufine est de 573nm. La fluorescence ou l'absorbance peuvent être utilisées pour enregistrer les résultats, mais la méthode de fluorescence est la préférée car elle est plus sensible et implique moins de calculs de données.

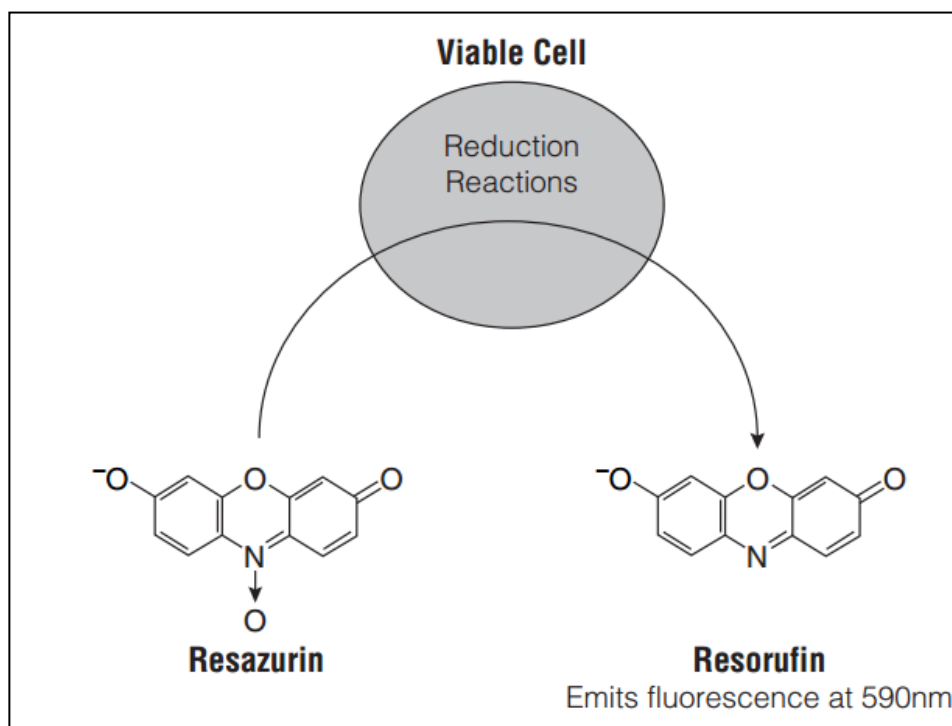
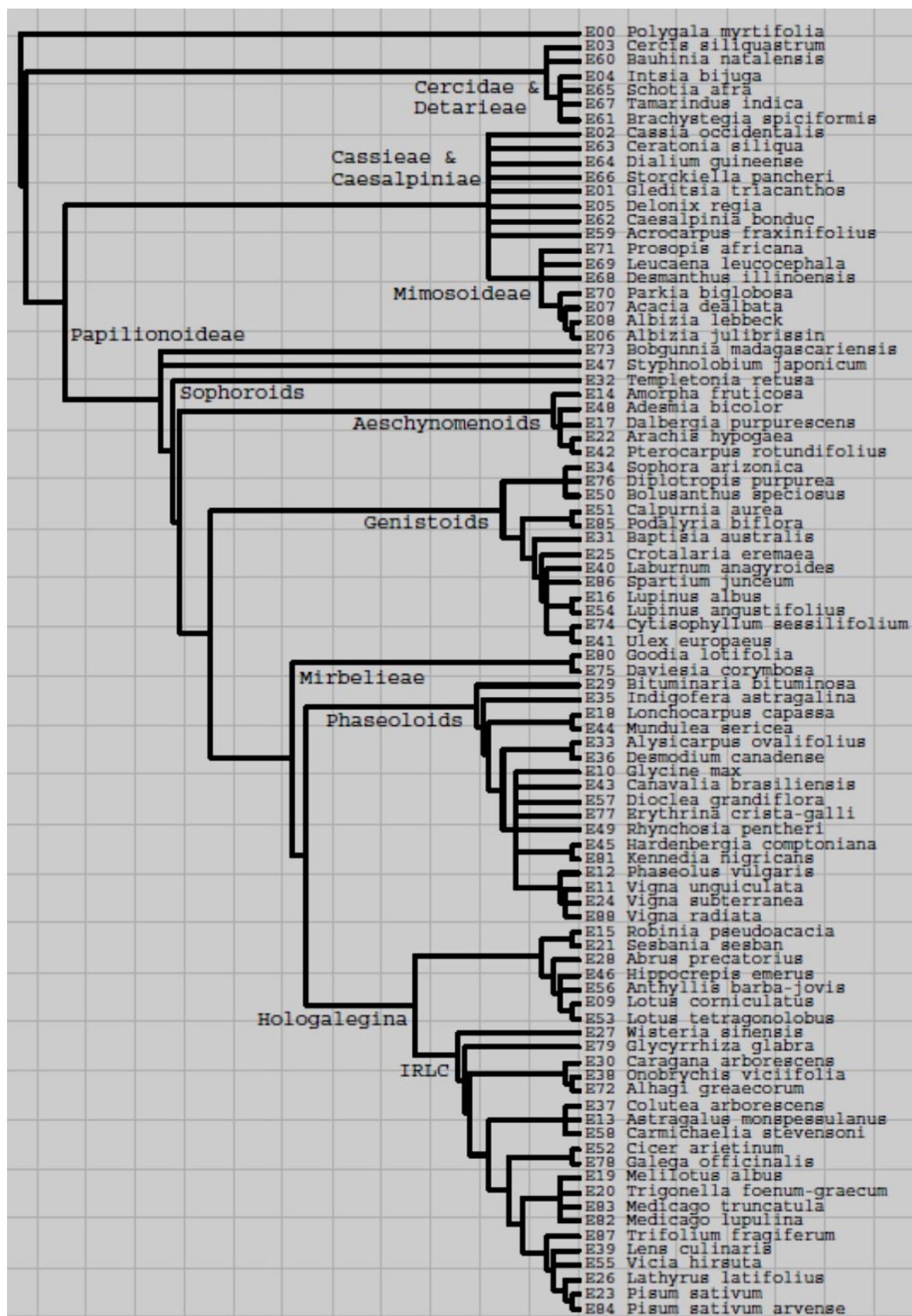


Figure 35 : Conversion de la résazurine en résorufine par des cellules métaboliquement actives entraînant la génération d'un produit fluorescent. La fluorescence produite est proportionnelle au nombre de cellules viables (CellTiter-Blue®Cell Viability Assay-INSTRUCTIONS FOR USE OF PRODUCTS G8080, G8081 AND G8082-Promega).

Annexe 7 Arbre phylogénétique des légumineuses [S LOUIS '07]



FOLIO ADMINISTRATIF
THESE SOUTENUE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES
DE LYON

NOM : Karaki	DATE de SOUTENANCE : 12 Décembre 2013
Prénoms : Lamis	
TITRE : Diversité génétique et fonctionnelle des molécules homologues de PA1b chez <i>Medicago truncatula</i> Gaertn. ainsi qu'au sein de légumineuses originaires du Liban.	
NATURE : Doctorat	Numéro d'ordre : AAAISALXXXX
Ecole doctorale : Evolution, Ecosystèmes, Microbiologie, Modélisation	
Spécialité : Biologie Evolutive, Biologie des Populations, Ecophysiologie	
Code B.I.U. – Lyon : T 50/210/19 / et bis CLASSE :	
RESUME : Les peptides Albumines 1 b sont des membres de la famille structurale des knottines et présentent un potentiel intéressant en tant que composés insecticides. À ce jour, leur diversité parmi les Fabacées a été essentiellement étudiée en utilisant des approches biochimiques et moléculaires. Les ressources bioinformatiques (le séquençage complet du génome, les bases de données transcriptomiques (EST...), les atlas d'expression...) de l'espèce modèle des légumineuses, <i>Medicago truncatula</i> Gaertn. (<i>Mtr</i>), nous a permis de développer une approche génomique de cette biodiversité. Deux objectifs principaux nous ont guidé à : 1) déchiffrer l'histoire évolutive de la famille A1 dans cette espèce et 2) explorer la biodiversité naturelle afin de découvrir de nouvelles molécules bioactives. L'exploration du génome de <i>Mtr</i> a révélé une remarquable expansion, à travers des duplications en tandem, des loci A1 qui retiennent presque toute la même structure génique canonique (2 exons et 1 intron). L'analyse phylogénétique nous a permis de comprendre l'évolution des gènes A1 intraspécifique et l'analyse de leur expression (EST, puces à ADN), et a révélé la distribution de la famille des gènes A1 dans les organes de la plante (tissus) : Cette dernière s'est révélée bien plus diverse que celle connue chez les autres espèces examinées de légumineuses, où la famille était jusque-là principalement graine-spécifique. Selon plusieurs critères, certains peptides ont été sélectionnés puis chimiquement synthétisés et repliés <i>in vitro</i> et testés pour leur activité biologique. Parmi eux, un peptide, nommé AG41, isoforme MtrA1013 et issu de l'EST orpheline TA24778_3880, a révélé un pouvoir insecticide élevé et inattendu. L'analyse à grande échelle en présence d'homologues d'A1 des légumineuses a montré l'ancestralité de la fonction insecticide et l'âge de la famille est estimé à plus de 58 million d'années. Notre étude s'est aussi orientée vers l'analyse de cette famille peptidique chez des légumineuses originaires du Liban. Cette approche par biologie moléculaire nous a permis de caractériser 9 nouveaux gènes chez 6 espèces de Papilionoideae. L'étude plus approfondie de ces gènes au niveau structural et fonctionnel est envisagée. Afin de relier les variations de structure et d'activité, un système d'expression hétérologue (baculovirus/cellules d'insectes <i>Sf9</i>) a été mis au point. Le peptide recombinant de référence PA1b (Pea Albumin 1 sub-unit b), même exprimé en faible quantité, présente une activité biologique et une masse bien conforme ainsi qu'une structure bien repliée. Ce système a permis, de même, de produire la proprotéine PA1, forme intermédiaire entre la préproprotéine et le peptide PA1b mature. Cette proprotéine, identifiée pour la première fois, ne présente aucune toxicité envers les cellules <i>Sf9</i>.	
MOTS CLES : Albumine 1, A1b, Fabacées, peptide de défense, résistance des plantes, expression hétérologue, knottine, <i>Medicago truncatula</i>	

Laboratoire(s) de recherches : UMR Biologie fonctionnelle, Insectes et Interactions (BF2I) INRA-INSA de Lyon. Laboratoire de Phytoécologie, bioclimatologie et écophysiologie végétale (UL) Fanar-Liban.
Directeur de thèse : Yvan Rahbé et Samir Safi
Président du jury : Composition du jury : Judith Burstin, Magda Bou Dagher-Kharrat, Raghida Abou Merhi, Miguel Lopez-Ferber, Yvan Rahbé, Corinne Royer, Francine Rizk et Samir Safi