



**HAL**  
open science

# Isolement, caractérisation et cibles de nouveaux Inhibiteurs de protéases pour la création de plantes transgéniques résistantes aux pucerons

Céline Deraison

► **To cite this version:**

Céline Deraison. Isolement, caractérisation et cibles de nouveaux Inhibiteurs de protéases pour la création de plantes transgéniques résistantes aux pucerons. Biochimie [q-bio.BM]. Université Paris Sud - Paris XI, 2002. Français. NNT: . tel-00006629

**HAL Id: tel-00006629**

**<https://theses.hal.science/tel-00006629>**

Submitted on 4 Aug 2004

**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**

**UNIVERSITE PARIS XI  
UFR SCIENTIFIQUE D'ORSAY**

**THESE**

**Présentée  
Pour obtenir**

**Le GRADE de DOCTEUR EN SCIENCES  
DE L'UNIVERSITE PARIS-XI ORSAY**

**PAR**

**Céline DERAISON MANUEL**

**Isolement, caractérisation et cibles de nouveaux inhibiteurs de  
protéases pour la création de plantes transgéniques résistantes  
aux pucerons**

**Soutenue le 27 juin 2002 devant la Commission d'examen**

**Michel DRON  
Michel BREHELIN  
Philippe BULET  
Philippe GIORDANENGO  
Lise JOUANIN  
Yvan RAHBE**

**Président  
Rapporteur  
Rapporteur  
Examineur  
Directeur de thèse  
Directeur de thèse**

Un grand Merci à

A Lise qui m'a confié ce sujet de thèse. Ses précieuses discussions, son enthousiasme inaltérable et son amitié m'ont beaucoup appris et aidé tout au long de ce travail. Merci surtout de m'avoir laissé cette grande autonomie dans la gestion de mon projet.

A Yvan qui a vu débarquer un jour une parisienne n'ayant jamais croisée un insecte. Je me souviendrai longtemps de tous ces moments d'échanges, de réflexion. Ta vision « modéliste » sur mes résultats, ta patience et ta frénésie positive m'ont continuellement soutenu pendant ces 3 années.

A Yves Chupeau, qui m'a accueillie au laboratoire de Biologie cellulaire de l'INRA de Versailles et à Gérard Febvay pour son accueil très chaleureux au laboratoire Biologie Fonctionnelle, Insectes et Interactions de l'INRA de Lyon.

Aux membres du jury qui ont accepté de juger ce travail : Philippe Bulet (Dr CNRS), Michel Bréhelin (Dr CNRS), Philippe Giordanengo (MC Université d'Amiens) et Michel Dron (Dr CNRS).

A mes amis Juliette, Thomas et Cédric avec qui j'ai partagé tant de moments de joie et parfois de tristesse. C'est l'occasion pour moi de leur témoigner ma profonde amitié.

A Frédéric sans qui la Biochimie serait toujours un mystère pour moi. Son savoir-faire, son esprit aiguisé et son regard extérieur sur mon travail ont été un inestimable appui. Tu sais combien c'est dur...

A Gaby pour toutes ces heures passées sous la loupe. Je crois que nous approchons des 5000 tubes digestifs...

A mes stagiaires grâce auxquels j'ai beaucoup appris aussi : Jérôme (aïe, aïe...) de licence, Sandrine (Formule 3000) de D.U., Gérard (star de l'immunoloc) de licence, Camille (c'est nouveau) du collège, Claire (les aléas de la papaine) 2<sup>ème</sup> année de l'INSA, Caroline (et ses anticorps) 3<sup>ème</sup> année de l'INSA et enfin Yannick (plus de secret pour les mini-prep) de BTS.

A toutes celles et ceux qui ont fait du laboratoire un lieu de vie où je me suis senti « comme un poisson dans l'eau » pendant toute la durée de ma thèse : à Versailles, Nathalie, Dominique, Christophe, Bruno, Line, Tatiana, Steph, Sylvère, Greg, Aymerick, J. Christophe, Rachid, Fanchon, Olivier, Téva, Daphnée ; à Lyon, Agnès, Fedé, Josette, Chaqué, Isabelle, Sandrine, Hubert, Bernard, Olivier. Merci pour ces moments d'allégresse et d'étonnement.

A la SNCF et ses TGV, sans qui cette thèse n'aurait pas ce parfum de voyage entre deux mondes.

A mes parents, à Candice qui m'ont toujours soutenu et respecté mes humeurs...

Et à toi Vitor, sans toi rien n'aurait été possible.

## Liste des abréviations et symboles utilisés

aa	acide aminé
A, mA	ampère, milliampère
ADN	acide désoxyribonucléique
ADNc	acide désoxyribonucléique complémentaire
ADN-T	acide désoxyribonucléique transféré
ARN	acide ribonucléique
ARNm	acide ribonucléique messenger
A.C.G.T	Adénosine, Cytosine, Guanidine, Thimidine
kb, bp	kilobase, paire de base
CaMV	<i>cauliflower mosaic virus</i> (virus de la mosaïque du chou-fleur)
CTAB	bromure d' hexadécyltriméthyl-ammonium
DEPC	diéthyl pyrocarbonate
dNTP	déoxynucléotides triphosphates
DMSO	diméthyl sulfoxyde
DTT	dithiothréitol
EDTA	éthylène diamine tétraacétique
Kg, g, mg, µg, ng	kilogramme, gramme, milligramme, microgramme, nanogramme
HPLC	<i>high performance liquid chromatography</i> (chromatographie liquide haute performance)
M, mM, µM	molaire, millimolaire, micromolaire
MS	<i>mass spectrometry</i> (spectrométrie de masse)
PCR	<i>polymerase chain reaction</i> (réaction de polymérisation en chaîne)
pH	potentiel hydrogène
pI	point isoélectrique
p/v	poids/volume
rpm	rotations par minute
R	résistant
RT	<i>reverse transcription</i> (rétrotranscription ou transcription inverse)
S	sensible
SDS	sodium dodécylsulfate
SSC	citrate salin standard
TAE	Tris Acetate EDTA
TE	Tris EDTA
U	unité enzymatique
V, mV	volts, millivolts
v/v	volume/volume
X	fois

<b>Introduction</b>	<b>5</b>
<b>1. Les inhibiteurs de protéases</b>	<b>6</b>
1.1. Classification et mode d'action des inhibiteurs de protéases	6
1.2. Le rôle des inhibiteurs de protéases chez les plantes	13
1.3. Gain de résistance par sur-expression d'inhibiteurs de protéases	16
1.4. L'effet de l'ingestion d'inhibiteurs de protéases sur les protéases des insectes	17
<b>2. Les pucerons</b>	<b>23</b>
2.1. Biologie et écologie des pucerons	23
2.2. Mode alimentaire et nutrition	25
2.3. La réponse moléculaire des plantes à la présence de pucerons.	30
2.4. Moyens de lutte contre les pucerons	33
<b>3. Objectif de la thèse</b>	<b>39</b>
<b>Chapitre I</b>	<b>40</b>
<b>Etude des protéases du tube digestif d'<i>A. gossypii</i></b>	<b>40</b>
<b>1. Contexte de l'étude</b>	<b>41</b>
Article 1	43
<b>Cloning and characterization of a gut specific Cathepsin L from the aphid <i>Aphis gossypii</i></b>	<b>43</b>
<b>2. Résultats complémentaires</b>	<b>66</b>
2.1 Analyse des séquences codant des protéases digestives putatives.	66
2.2 Dosages enzymatiques	67
2.3. Discussion	68
<b>Chapitre II</b>	<b>71</b>
<b>Utilisation &amp; Diversification des inhibiteurs de proteases à cystéine</b>	<b>71</b>
<b>1. Contexte de l'étude</b>	<b>72</b>
Article 2	73
<b>Effects of the cysteine protease inhibitor oryzacystatin (OC-I) on different aphids and reduced performance of <i>Myzus persicae</i> on OC-I expressing transgenic oilseed rape</b>	<b>73</b>
<b>II. Diversification des inhibiteurs de protéases à cystéine</b>	<b>92</b>
<b>Introduction</b>	<b>92</b>
1.1. Résistance dérivée du pathogène	92
1.2. Les inhibiteurs de protéases chez les insectes	93
<b>Matériel et Méthodes</b>	<b>97</b>
2.1. Matériel biologique	97

2.2. Préparation des extraits protéiques _____	97
2.3. Construction et analyse de la banque d'ADNc _____	98
2.4. Dosage des protéines _____	99
2.5. Méthodes d'analyse des protéines _____	99
2.6. Mise en évidence d'inhibiteurs de protéases _____	101
2.7. Méthodes de purification des protéines _____	102
2.8. Microséquençage _____	104
3.2. Inhibition de l'activité protéolytique du tube digestif par l'hémolymphe _____	106
3.3. Essai de clonage du gène codant un inhibiteur de protéase à cystéine. _____	107
3.4. Purification du peptide à activité IP-cystéine. _____	107
<b>Conclusions</b> _____	<b>116</b>
<b>Chapitre III</b> _____	<b>117</b>
<b>Etude du mode d'action &amp; modification de PsTI</b> _____	<b>117</b>
<b>Introduction</b> _____	<b>118</b>
1.1 Les inhibiteurs de protéases à sérine chez les plantes. _____	118
1.2. Toxicité des inhibiteurs de la famille des Bowman Birk envers les pucerons _____	118
<b>I. Caractérisation du mode de toxicité d'un inhibiteur de protéase à sérine, PsTI chez les pucerons <i>A. pisum</i> et <i>A. gossypii</i>.</b> _____	<b>120</b>
<b>Introduction</b> _____	<b>120</b>
<b>Matériel et Méthodes</b> _____	<b>121</b>
2.1. Origine et élevage des insectes _____	121
2.2. Origine de l'inhibiteur de protéase à sérine _____	121
2.3. Construction et criblage d'une banque d'ADNc _____	121
2.4. Caractérisation des activités protéolytiques du tube digestif du puceron du pois. _____	122
2.5. Chromatographie liquide en phase inverse _____	122
2.6. Composition et préparation des milieux artificiels _____	123
2.7. Principe des tests de toxicité in vitro _____	123
2.8. Immunolocalisation _____	124
<b>Résultats</b> _____	<b>125</b>
3.1. Mise en évidence des protéases du tube digestif du puceron du pois. _____	125
3.2. Interaction de PsTI avec les protéases du tube digestif _____	126
3.3. Effet de PsTI sur la survie et la croissance du puceron du melon. _____	127
3.4 Mécanismes potentiels de détoxification de PsTI chez le puceron du melon _____	128
3.5. Localisation de la cible de PsTI chez les pucerons du pois et du melon. _____	128
<b>Discussion</b> _____	<b>129</b>

<b>II. Mutagenèse dirigée et Expression hétérologue de PsTI et de PsCI</b>	<b>135</b>
<b>Introduction</b>	<b>135</b>
<b>Matériel et méthodes</b>	<b>137</b>
2.1 Synthèse du gène artificiel	137
2.2. Expression dans la levure	137
2.3. Expression de PsTI et PsCI <i>in planta</i>	139
<b>Résultats et Discussion</b>	<b>143</b>
3.1. Construction des séquences synthétiques	143
3.2. Expression de PsTI et PsCI dans la levure <i>Pichia pastoris</i>	144
3.3. Expression hétérologue de PsTI et PsCI dans la plante <i>Arabidopsis thaliana</i>	146
<b>Discussion Générale &amp; conclusions</b>	<b>149</b>
1. Les protéases digestives du puceron <i>A. gossypii</i>	151
2. L'utilisation des inhibiteurs de protéases pour lutter contre les pucerons.	153
3. Les inhibiteurs de protéases atteignent une nouvelle cible	156
<b>Bibliographie</b>	<b>158</b>
<b>Publication et Communications</b>	<b>173</b>
<b>Annexes</b>	<b>174</b>

Les insectes phytophages constituent un problème majeur pour l'agriculture. Actuellement, les moyens de lutte contre ces insectes sont essentiellement basés sur l'utilisation d'insecticides chimiques. Dans le cadre d'une agriculture plus respectueuse de l'environnement, il serait nécessaire de cultiver des plantes naturellement plus résistantes aux insectes phytophages. La résistance naturelle d'une plante à un insecte se définit par l'ensemble des facteurs, essentiellement d'origine génétique, conduisant à des infestations et/ou des symptômes moins importants que ceux observés sur une plante sensible. La résistance génétique aux insectes phytophages est une composante importante de la protection des cultures car elle répond aux préoccupations du consommateur en matière de qualité/sécurité nutritionnelle comme à celles des producteurs par la réduction des intrants. Les programmes d'amélioration génétique ont permis d'obtenir quelques résultats intéressants mais sont dépendants de l'identification de sources de résistance dans des espèces sauvages apparentées aux espèces végétales domestiques.

Aujourd'hui, les biotechnologies offrent de nouvelles capacités aux stratégies de contrôle des populations d'insectes nuisibles à l'agriculture. Parmi les métabolites synthétisés par la plante pour se protéger contre les ravageurs et les pathogènes, les peptides et protéines toxiques sont très étudiés pour leur potentialité à être utilisés en transgénèse. Les inhibiteurs de protéases font partie de ces outils de défense et sont souvent utilisés dans des approches de transgénèse afin d'accroître la résistance des plantes aux phytophages majeurs que sont les lépidoptères et coléoptères. D'une façon schématique, la perte d'activité protéolytique induit une déficience en acides aminés chez l'agresseur, entraînant une diminution de la croissance et de la reproduction, voire la mort de l'insecte dans certains cas.

Parmi les insectes phytophages, les pucerons sont particuliers car ils se nourrissent de sève élaborée, compartiment très spécifique de la plante, en général relativement pauvre en protéines. C'est pourquoi, dans ce contexte, les pucerons sont considérés comme ne possédant pas l'arsenal enzymatique permettant une utilisation efficace des protéines. Par conséquent, les stratégies utilisant les inhibiteurs de protéases comme polypeptide entomotoxique ne semblent pas adaptées *a priori* à ce groupe d'insecte. Nous verrons que cette description sommaire reste très approximative, et les objectifs de thèse, développés en fin d'introduction, concernent précisément les interactions entre les pucerons et les inhibiteurs de protéases, dont les cibles sont jusque là complètement ignorées.



**~~– INTRODUCTION~~**

# 1. Les inhibiteurs de protéases

## 1.1. Classification et mode d'action des inhibiteurs de protéases

Les inhibiteurs de protéases (IP) sont des protéines très répandues dans la nature, présentes chez les animaux, les plantes et les micro-organismes. Le nombre d'IP isolés et identifiés à ce jour est extrêmement large. Les fonctions des inhibiteurs de protéases sont essentielles pour tout organisme vivant :

(1) d'une part, ils préviennent une activité protéolytique incontrôlée. Les protéases font l'objet d'une régulation très étroite, limitant ainsi leur activité catalytique aux seuls processus et compartiments physiologiques où elles sont nécessaires ;

(2) d'autre part, ils protègent l'organisme d'une activité protéolytique exogène, qu'elle soit le fait de prédateurs ou d'organismes pathogènes.

On distingue classiquement quatre classes d'inhibiteurs de protéases qui inhibent respectivement les quatre classes majeures de protéases.

### 1.1.1. Les classes fonctionnelles de protéases

Les protéases<sup>1</sup> sont ordonnées en fonction des résidus d'acides aminés essentiels impliqués dans la catalyse :

- les **protéases à sérine** (EC. 3.4.21.-)<sup>2</sup> comportent généralement dans leur site actif trois résidus essentiels à la catalyse (une sérine, une histidine et un aspartate) formant la triade catalytique. Au cours de la catalyse, la sérine joue le rôle de nucléophile et l'histidine celui d'accepteur initial de protons. En fonction de leur séquence et de leur organisation structurale, les protéases à sérine sont organisées en quarante familles (Rawlings, 1998).

Chaque enzyme présente une spécificité particulière plus ou moins étroite vis à vis des résidus après lesquels ils coupent la chaîne peptidique. Par exemple, la trypsine coupe une

---

<sup>1</sup> Il existe deux groupes de protéines regroupés sous le terme "protéase" : les endoprotéases qui hydrolysent une liaison peptidique interne et les exopeptidases qui éliminent un ou deux acides aminés en N- ou C-terminal de la protéine. Dans cette étude, la distinction entre les endoprotéases (principal intérêt de l'étude) et les exopeptidases reposera sur l'utilisation des termes protéases et peptidases respectivement.

<sup>2</sup> Numérotation selon la nomenclature des enzymes (Commission on Biochemical Nomenclature, 1972), telle que décrite dans la base ENZYME (Bairoch A., 2000)

liaison peptidique du côté C-terminal d'un résidu basique tel que l'arginine ou la lysine ; la chymotrypsine choisit les liaisons peptidiques du côté carboxylique d'acides aminés à noyau aromatique comme la phénylalanine, la tyrosine ou le tryptophane, qui possèdent une charge latérale hydrophobe volumineuse (Schellenberger, 1991).

- les **protéases à cystéine** (E.C. 3.4.22.-) sont largement représentées tant dans le monde animal que végétal ainsi que chez les bactéries et les virus où elles sont impliquées dans de nombreux processus protéolytiques intra et extracellulaires. On distingue actuellement au minimum vingt six familles distinctes de protéases à cystéine (Barret, 1998), dont la plus grande et la plus connue est sans doute celle de la papaïne. La triade catalytique est dans ce cas formé de trois résidus (cystéine, histidine et asparagine), le résidu cystéine constituant le groupement nucléophile. La papaïne, protéase à cystéine caractéristique, possède un site actif composé de sept acides aminés (S1-S4-S1'-S3')<sup>3</sup> dont chacun est capable d'interagir avec un acide aminé du substrat (P1-P4 et P1'-P3') situé respectivement de part et d'autre des côtés N et C-terminaux du site de coupure. L'hydrolyse de la liaison peptidique intervient après un résidu arginine (P1) précédé de préférence d'un résidu non polaire (P2).

- les **protéases acides**<sup>4</sup> (EC.3.4.23.-) et les **métalloprotéases** (EC.3.4.24.-), agissent par l'intermédiaire d'une molécule d'eau activée, cette dernière constitue l'agent nucléophile permettant d'attaquer la liaison peptidique. Les résidus impliqués dans la catalyse sont deux résidus acides (aspartate ou glutamate) qui jouent le rôle de ligand de la molécule d'eau activée pour les protéases acides, alors que dans le cas des métalloprotéases, la molécule d'eau est liée à un ou deux ions métalliques, généralement du zinc.

---

<sup>3</sup> La nomenclature de Schechter et Berger (1967) H<sub>2</sub>N-P<sub>x</sub>...-P<sub>3</sub>-P<sub>2</sub>-P<sub>1</sub>-P'<sub>1</sub>-P'<sub>2</sub>-P'<sub>3</sub>-...-P'<sub>x</sub>-COOH définissant la topologie linéaire des interactions protéase-substrat peptidique est aussi généralement utilisée pour décrire l'environnement du site actif des IP. La liaison peptidique entre les résidus P<sub>1</sub> et P'<sub>1</sub> correspond alors au site actif de l'inhibiteur. Les acides aminés catalytiques de la protéase sont notés S<sub>3</sub>, S<sub>2</sub>, S<sub>1</sub> et S'<sub>1</sub>, S'<sub>2</sub>, S'<sub>3</sub>...

<sup>4</sup> Les protéases acides sont aussi nommées protéases à aspartyl.

### 1.1.2. Mécanisme d'action des inhibiteurs de protéases.

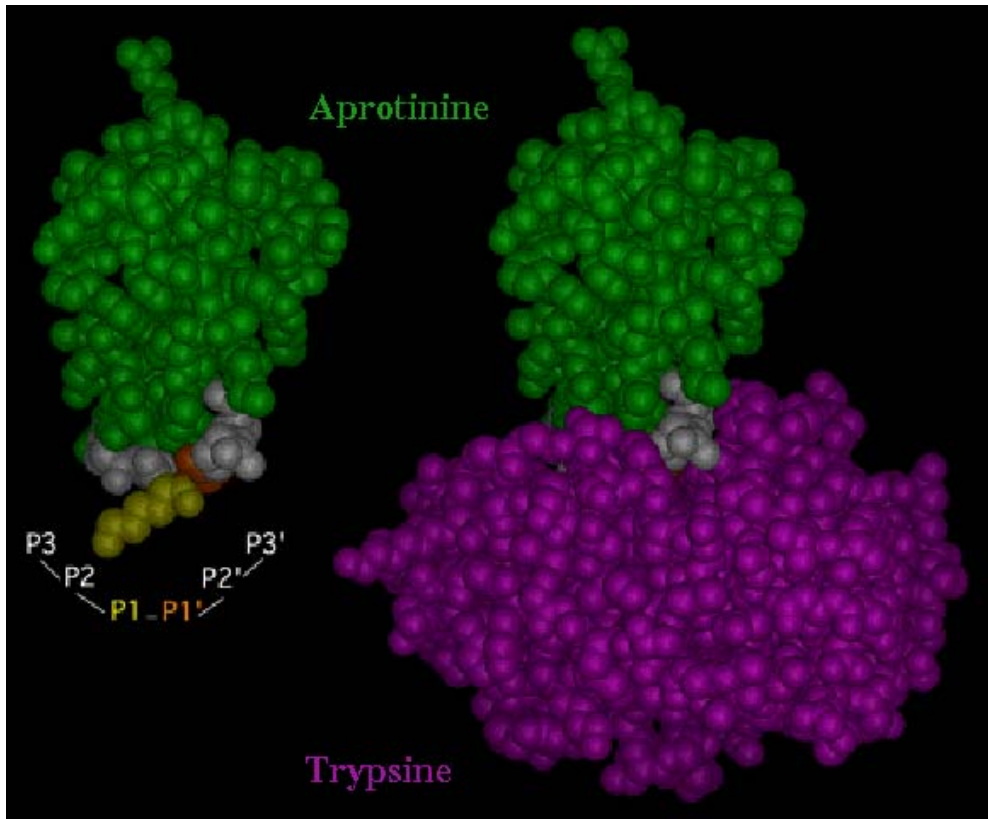
Les inhibiteurs de protéases sont le plus souvent spécifiques d'une classe de protéases mais peuvent aussi avoir des spécificités croisées ou ne présenter aucune spécificité particulière, c'est le cas des macroglobulines humaines, protéine majeure du plasma des vertébrés.

Les IP agissent selon différents mécanismes d'inhibition (Bode, 2000) :

- soit, ils bloquent le site actif en prenant la place du substrat selon un mécanisme de compétition ou en se fixant sur des acides aminés adjacents au site actif, avec cependant une association IP-protéase beaucoup plus forte ( $K_d$  de  $10^{-7}$  à  $10^{-13}$  M) que celle du complexe substrat-protéase. Le complexe ainsi formé est très stable du fait de nombreuses interactions intermoléculaires comme des liaisons hydrogènes ou électrostatiques ;

- soit, ils se fixent sur des motifs relativement éloignés du site actif et inhibent l'enzyme de façon allostérique, l'effecteur allostérique induisant dans la structure tertiaire de l'enzyme un changement de conformation qui se propage au site catalytique en modifiant son activité.

La nomenclature actuelle répartit les IP dans vingt et une familles. Ce sont des protéines en général de petites tailles (3 à 47 kDa) monomériques ou multimériques. Dans la plupart des cas, les membres d'une superfamille d'inhibiteurs sont dirigés contre des protéases cibles possédant le même mécanisme catalytique. Il existe très peu d'inhibiteurs bifonctionnels où les deux domaines, hétérologues, inhibent deux enzymes de classes différentes. A notre connaissance 2 cas sont décrits dans la littérature : l'inhibiteur du maïs 14K-CI qui inhibe une trypsine et une  $\alpha$ -amylase, et l'équistatine isolée d'une anémone de mer qui inhibe simultanément les enzymes de la famille de la papaïne et de la cathepsine D (protéase acide) (Lenarcic, 1997).



**Figure 1 : Complexe trypsine-aprotinine**

L'aprotinine (à gauche), interagit très fortement avec la trypsine par l'intermédiaire d'un site réactif inhibiteur hautement complémentaire du site actif de la trypsine. Les inhibiteurs de type canonique possèdent tous une boucle inhibitrice de conformation similaire par laquelle ils interagissent d'une manière commune avec les protéases à sérine (P3 à P'3). La liaison cible de la protéase est située entre les résidus P1 et P'1. Lorsque le complexe est formé (à droite), la lysine (en jaune) constituant le résidu P1 de la boucle inhibitrice forme une liaison électrostatique avec un résidu aspartate du fond de la poche primaire de la protéase (S3).

### 1.1.3. Les inhibiteurs de protéases à sérine

Les protéases à sérine constituent sans aucun doute les protéases les plus étudiées et le mieux caractérisées actuellement. Il n'est donc pas étonnant que la majorité des inhibiteurs naturels de protéases connus à ce jour soient des inhibiteurs de protéases à sérine. Une classification de ces inhibiteurs permet de les regrouper en treize familles distinctes basées sur les similarités de séquences, de topologie ainsi que sur les mécanismes de fixation des protéases. Cependant, en fonction du mécanisme d'inhibition, on peut distinguer trois principaux groupes (Laskowski, 1980) :

- les **serpines** (*Serine Proteinase Inhibitors*) appartiennent à une famille de glycoprotéines d'environ 400 résidus interagissant avec les protéases cibles par l'intermédiaire d'une boucle inhibitrice. Lors de l'interaction protéase-serpine, cette boucle inhibitrice subit un clivage protéolytique au niveau de la liaison P1-P1' qui a pour conséquence soit la formation d'un complexe irréversible, soit la formation d'un complexe transitoire qui se dissocie en libérant une forme clivée de la serpine mais dorénavant inactive en tant qu'inhibiteur (Wright, 1996). Elles fonctionnent donc comme des substrats suicides (Engh, 1995).

- les **inhibiteurs canoniques** des protéases à sérine sont les IP à sérine les plus nombreux. Bien que ces molécules aient des modes de repliement complètement différents, elles possèdent néanmoins toutes en commun une boucle inhibitrice exposée vers l'extérieur, de séquence variable mais dont la conformation P3-P3' (six résidus autour du site inhibiteur) est très similaire, et appelée de ce fait conformation canonique (figure 1). La structure 3D qui porte cette boucle, comporte un cœur hydrophobe très compact, stabilisé le plus souvent par plusieurs ponts disulfure, ce qui rend ces molécules remarquablement stables. On peut considérer que la formation du complexe protéase-inhibiteur résulte de l'association entre deux molécules rigides. L'interaction est limitée à la région de la boucle inhibitrice avec le site actif de l'enzyme, mais elle est très spécifique du fait de la complémentarité très élevée de forme entre les deux surfaces moléculaires. L'affinité de l'IP pour l'enzyme est très grande mais à la différence d'un substrat, la liaison peptidique sensible P1-P1' est hydrolysée extrêmement lentement.

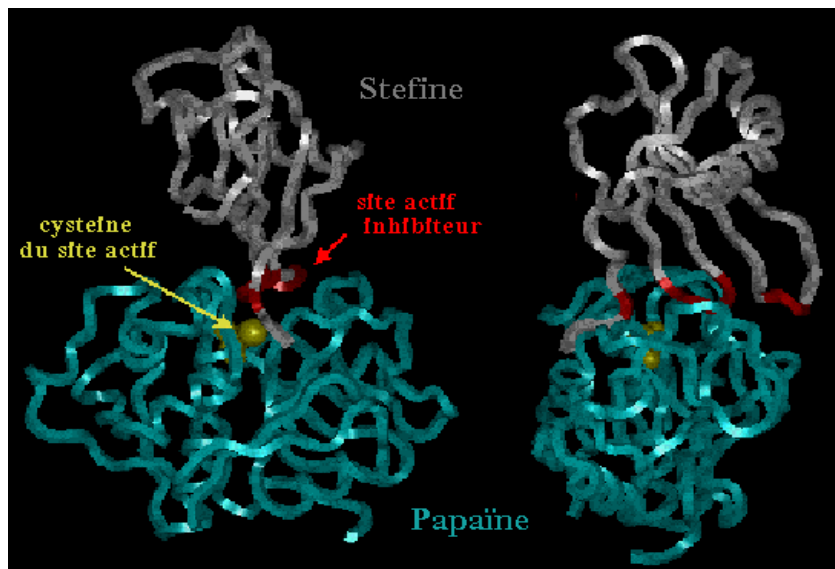
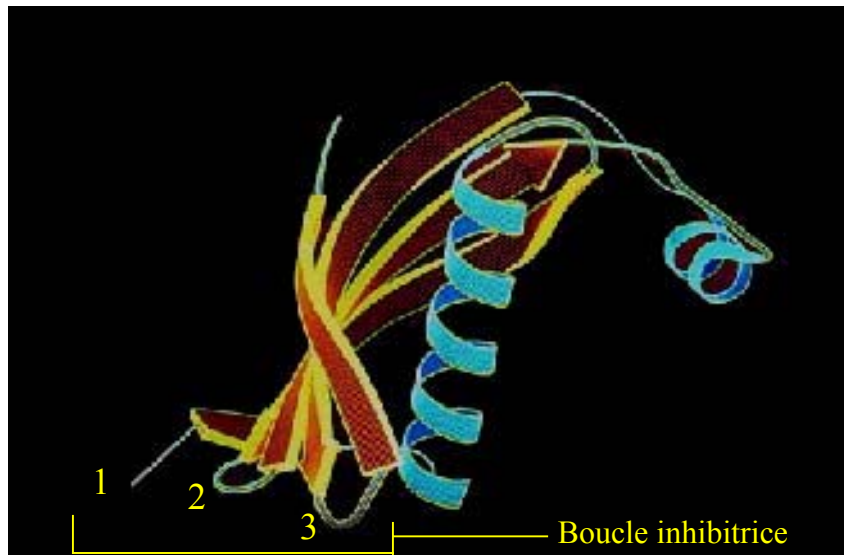
Les inhibiteurs de protéases à sérine des plantes appartiennent à cette super-famille des inhibiteurs canoniques. Richardson (1991) en se basant sur l'homologie structurale entre les séquences, sur la localisation des ponts disulfure et sur la position du site actif, a défini sept familles d'IP à sérine. Les deux familles qui renferment le plus grand nombre de séquences connues sont les inhibiteurs de la famille Bowman-Birk et de la famille Kunitz. Les inhibiteurs de type Bowman-Birk possèdent typiquement deux régions en tandem avec pour chacune un site réactif, et contiennent sept ponts disulfure (Birk, 1976). Les inhibiteurs de type Kunitz, par contre, ne possèdent qu'une seule chaîne avec un seul site réactif, et ne renferment que 2 ponts disulfure.

- l'hirudine est l'un des plus puissants inhibiteurs naturels de thrombine. Cette protéase est impliquée dans la dernière étape de la coagulation sanguine permettant de convertir le fibrinogène en un caillot de fibrine. Molécule de 65 acides aminés à trois ponts disulfures internes, l'hirudine complexe la thrombine de manière non covalente avec une affinité extrêmement élevée ( $K_d < 1 \text{ pM}$ , constituant l'une des plus fortes interactions entre deux protéines décrites à ce jour (Storc, 1996). L'hirudine forme des contacts intermoléculaires tout le long de la cavité du site actif de la thrombine. Comme une grande proportion de l'inhibiteur est impliquée dans l'interaction avec la thrombine, un grand nombre de contacts intermoléculaires sont formés avec la région du site actif et d'autres régions adjacentes, expliquant l'affinité très élevée entre les deux protéines (la surface de contact est deux à trois fois plus grande que pour les inhibiteurs canoniques). A la différence des inhibiteurs canoniques, la poche primaire de spécificité (S1) de l'enzyme n'est pas occupée par des résidus de l'inhibiteur mais par des molécules d'eau.

#### **1.1.4. Les inhibiteurs de protéases à cystéine**

Les principaux inhibiteurs endogènes des protéases de la famille de la papaïne sont les cystatines, ainsi nommées pour leur capacité à bloquer l'action des protéases à cystéine. Ces molécules sont des inhibiteurs protéiques à forte affinité et réversibles. La super famille des cystatines est divisée en trois familles selon le nombre de copies de segment cystatine et la présence de ponts disulfure (Barrett, 1987) :

- les **stéfines** (famille 1) représentent les molécules les plus primitives et les plus simples de la super-famille. Elles sont constituées d'une seule chaîne polypeptidique non



**Figure 2 : Le complexe cystatine-protéase à cystéine.**

**a. Structure 3D d'une molécule de la superfamille des cystatines, la cystatine de poulet.**

Le site inhibiteur de cette molécule est formé par la juxtaposition dans l'espace de 3 régions de la protéine, la région N-terminale (1), la région centrale comportant la séquence conservée de type QxVxG (2) et la région C-terminale (3).

**b. Complexe de la stéfine B humaine avec la papaïne.**

La stéfine se lie avec une forte affinité dans le site actif de la papaïne. La molécule de stéfine interagit avec les protéases à cystéine par l'intermédiaire de son site inhibiteur (régions 1, 2 et 3, colorées en rouge) hautement complémentaire du site actif des protéases (la cystéine catalytique est colorée en jaune).



glycosylée et d'environ 100 acides aminés (11 kDa). Elles ne possèdent pas de ponts disulfure et leur localisation est principalement intracellulaire.

- les **cystatines** (famille 2) sont des molécules un peu plus complexes constituées d'environ 115 acides aminés (13 kDa). Elles renferment en général deux ponts disulfure qui forment deux boucles à proximité de l'extrémité C-terminale de l'IP. Elles sont présentes dans beaucoup de fluides extracellulaires.

Les inhibiteurs de protéase à cystéine de plantes (phytocystatines) font partie de cette famille mais ils ont la particularité de ne pas posséder de ponts disulfure (Nagata, 2000). Les phytocystatines présentent une masse moléculaire de 12 à 16 kDa. Elles ont été purifiées à partir de nombreuses graines (Abe, 1987 ; Pernas, 1998 ; Yamada, 2000), mais la présence de ces IP a aussi été révélée dans d'autres tissus tels que les feuilles (Zhao, 1996), les fruits (Kimura, 1995) et le pollen (Rogers, 1993).

- les **kininogènes** (famille 3) apparaissent avec leur structure complexe comme les membres les plus évolués de la superfamille des cystatines. La chaîne lourde de ces molécules comporte trois domaines de type cystatine acquis vraisemblablement par duplication de gènes dont seulement deux sont fonctionnels. Les kininogènes sont des molécules complexes (335 acides aminés, 2 ponts disulfure, glycosylées) et multifonctionnelles qui peuvent fixer et inhiber de manière indépendante deux molécules de protéases à cystéine (Turk, 1996).

Rawlings et Baret (1990) suggèrent que ces familles dérivent d'un gène ancestral possédant deux ponts disulfure. Des phénomènes de duplication du matériel génétique et de mutation conduisant à la perte des cystatines ont eu lieu lors des 100 millions dernières années (Rawlings, 1990).

Les structures 3D obtenues par cristallographie de deux membres représentatifs de la superfamille, la cystatine de poulet (Bode, 1988) et la stéfine B humaine complexée avec la papaïne (Stubbs, 1990), ont permis de comprendre le mécanisme d'action de ces inhibiteurs (figure 2). La structure de ces molécules consiste notamment en un feuillet  $\beta$  antiparallèle composé de cinq brins. A l'une des extrémités du feuillet  $\beta$  se trouve une séquence consensus du type QxVxG, flanquée de part et d'autre de la région dipeptide PW (pour les stéfines ou PH pour les familles 2 et 3) et de la région N-terminale possédant une glycine invariante en position 9. Ces trois régions, formant le site inhibiteur, non contiguës dans la séquence, donnent à l'IP une forme en "coin", de nature hydrophobe, tout à fait complémentaire du site actif de la papaïne et d'autres protéases à cystéine. A la différence de substrats liés dans le

site actif de l'enzyme, la région N-terminale adopte une structure en épingle à cheveux qui lui permet de contourner la cystéine 25 du site catalytique, et ne peut donc pas être clivée (Bode, 2000).

Une nouvelle classe d'inhibiteurs de protéases à cystéine a été mise en évidence récemment formant la superfamille des **thyropines**, ainsi dénommés parce qu'ils représentent une homologie avec le motif de séquence appelé thyroglobuline de type 1. Plusieurs membres de cette famille ont été caractérisés, dont l'équistatine (Lenarcic, 1997) et le fragment peptidique P41 dérivé de la chaîne invariante Li associée aux molécules de classe II du complexe d'histocompatibilité (Guncar, 1999). La structure 3D de ce dernier en complexe avec la cathepsine L (protéase à cystéine lysosomiale) révèle un mode de repliement nouveau comprenant deux sous domaines stabilisés par des ponts disulfure. L'inhibition de la cathepsine L résulte de la fixation d'un motif en forme de " coin ". comme celui des cystatines mais des interactions additionnelles de nature électrostatiques font de ces inhibiteurs des molécules plus sélectives que les cystatines.

### **1.1.5. Les inhibiteurs de métalloprotéases et de protéases acides**

Ces dernières classes d'inhibiteurs ont été mises en évidence dans très peu d'organismes. De manière générale, très peu d'inhibiteurs de métalloprotéases ou de protéases acides sont connus à ce jour.

Chez les vertébrés, les inhibiteurs tissulaires de métalloprotéases (**TIMP**) sont des molécules d'environ 20 à 30 kDa multifonctionnelles car dotées aussi d'une activité mitogénique indépendante de leur activité inhibitrice (Gomez, 1997). Ce sont des inhibiteurs à forte affinité qui forment des complexes réversibles et stœchiométriques avec leur enzyme cible, bloquant physiquement l'accès des substrats au site actif. La molécule de TIMP est constituée de deux domaines, un domaine C-terminal exposant cinq boucles qui forment le site inhibiteur venant interagir avec le site actif de l'enzyme et un domaine N-terminal qui forme des liaisons de coordination avec l'atome de zinc catalytique.

Un inhibiteur de protéase acide a été caractérisé chez la courge (Christeller, 1998). Cette molécule de 10,5 kDa circule dans le phloème mais sa cible physiologique n'est pas connue. Cet IP est capable d'inhiber une protéase acide synthétisée par un champignon pathogène des Cucurbitacées suggérant un rôle possible dans la défense de la plante contre les pathogènes.

Neurath (1984) a suggéré que les organismes primitifs contrôlaient probablement les activités protéolytiques grâce aux inhibiteurs de protéases. Au cours de l'évolution, des systèmes biochimiques et physiologiques plus complexes ont été mis en place comme la régulation *via* l'activation de proenzyme et la compartimentation. Toutefois, bien que le mode d'action des inhibiteurs de protéases puisse paraître simple, les mécanismes moléculaires mis en jeu ont été perfectionnés et peuvent s'avérer très sophistiqués (*cf.* les serpinines).

## 1.2. Le rôle des inhibiteurs de protéases chez les plantes

Les IP dans les plantes, du moins pour ce qui concerne les serpinines et cystatines les plus étudiées, semblent remplir plusieurs fonctions.

### 1.2.1 Régulation des protéases endogènes

Certains IP sont impliqués dans la régulation des protéases endogènes de la plante :

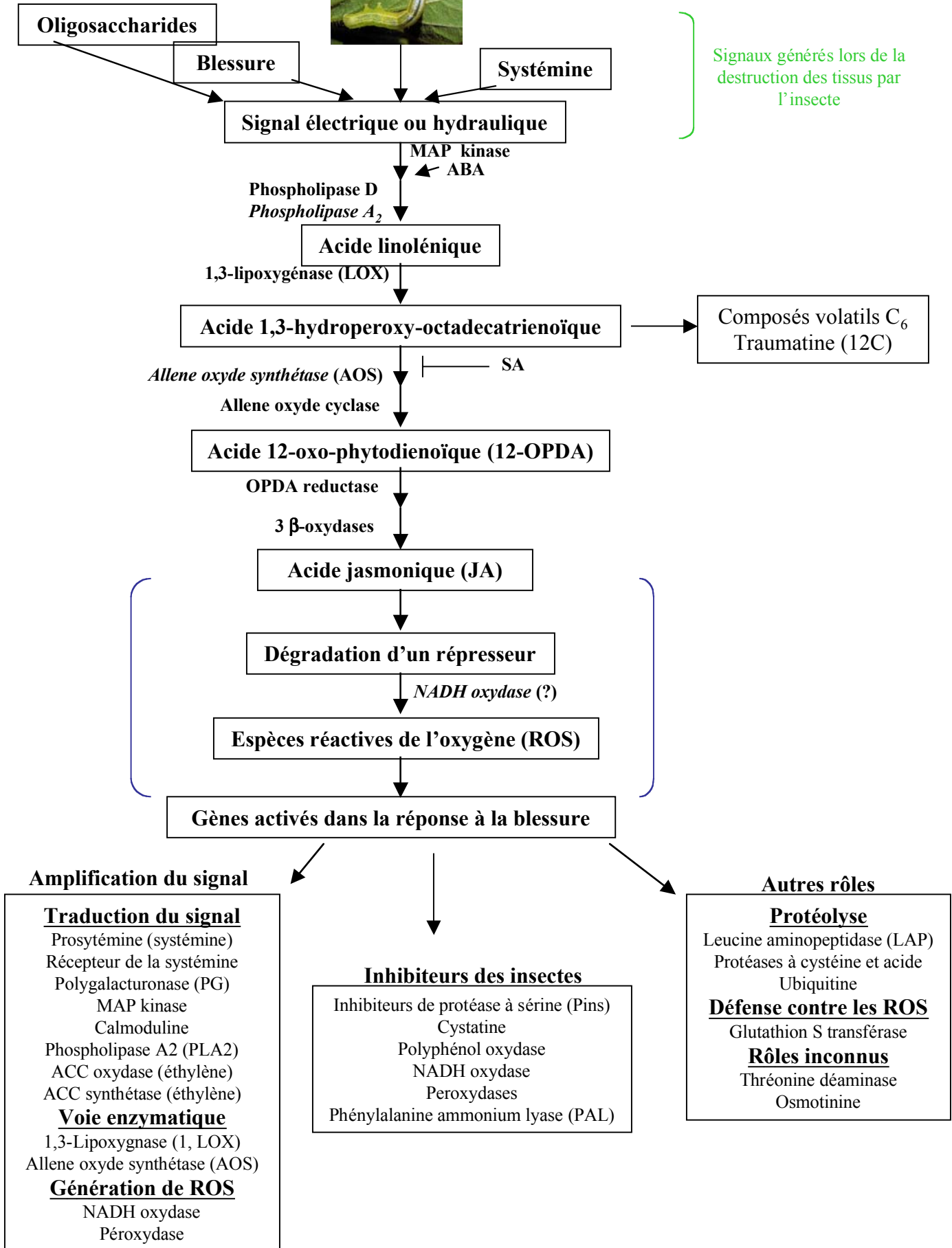
- **la régulation de la mort cellulaire programmée ou apoptose.** Des protéases à cystéine jouent un rôle essentiel dans l'apoptose induite par un stress oxydatif notamment dans le cas de la xylogénèse (Ye, 1996). Certains IP pourraient donc prévenir d'une mort cellulaire indésirable (Solomon, 1999).

- **la régulation de la mise en place des réserves dans la graine.** C'est notamment le cas de la cystatine OC-I mise en évidence chez le riz (Abe, 1987). L'oryzacystatine est synthétisée dans les graines immatures de riz dès que la floraison a débuté et atteint une quantité maximale (2-3 mg/kg de graines) quelques semaines après, période correspondant à la phase de remplissage du grain. L'accumulation des protéines de réserve dans la graine au cours de la maturation impose un contrôle de l'activité protéasique vacuolaire (Watanabe, 1991).

- **la régulation de la mobilisation des réserves** lors de la germination. Au cours de la germination, les protéines de réserve et les IP seraient hydrolysés par des protéases synthétisées *de novo* (Kondo, 1990) ou par d'autres protéases à spécificité différente (dégradation de serpinines par une protéase à cystéine lors de la germination du soja (Papastoitis, 1991) permettant l'assimilation des composés azotés dans les différentes voies



Signaux générés lors de la destruction des tissus par l'insecte



de biosynthèse. Ainsi, les IP joueraient également un rôle de réservoir d'acides aminés (surtout soufrés), mobilisables pendant la germination et le développement précoce de la plantule. Les IP caractérisés dans les graines sont surtout des IP à sérine qui possèdent des structures très compactes maintenues par des ponts disulfure.

### 1.2.2. Protection contre les phytopathogènes.

On trouve des IP en grande quantité dans les organes de réserve comme les graines ou les tubercules (10 à 15% des protéines totales), et dans les fruits immatures. En plus du rôle direct de stockage, la prévalence des IP dans les organes de réserve suggère **un rôle défensif contre les herbivores**. Dans les fruits verts de la tomate sauvage, *Lycopersicon peruvianum*, la très forte concentration en IP (en outre, deux IP à sérine : InhI et InhII) empêcherait la consommation des graines immatures par les herbivores. Après maturation de la graine, on observe une décroissance de la concentration en IP rendant ainsi les fruits comestibles, les herbivores pouvant ainsi faciliter la dispersion des graines de tomate (Pearce, 1988).

Beaucoup d'inhibiteurs de protéases sont le produit de familles multigéniques. Il est courant de trouver chez les plantes plusieurs isoformes possédant des spécificités différentes envers des protéases de mécanisme réactionnel identique.

De nombreuses études publiées mettent en évidence des voies signalétiques mises en place par les plantes après qu'elles aient subi des dommages tissulaires importants lors de la présence des larves d'insectes tel que les lépidoptères ou les coléoptères (figure 3).

Beaucoup de plantes répondent à l'attaque des herbivores en activant des gènes de défense dont les produits réduisent la qualité nutritionnelle des protéines ingérées (Polyphenol oxydase) et inhibent les protéases digestives de l'insecte (IP), ce qui peut être directement lié à une augmentation de résistance (Ryan, 1990).

---

### Figure 3 : Modèle des voies de transduction en réponse à la blessure (Walling, 2000).

Ce modèle est basé sur la voie des octadécanoïdes chez la tomate. Plusieurs effecteurs sont putatifs et indiqués entre crochets bleus, les enzymes clés sont indiquées en italique.

La blessure et/ou les sécrétions salivaires génèrent un signal électrique ou hydraulique qui est rapidement propagé à partir du site de dommage. Ces signaux pourraient stimuler la libération de composés (oligogalacturonides et systémine) qui amplifieraient en outre la cascade de signalisation à travers la plante (Walling, 2000). La figure 3 résume un modèle récent intégrant l'ensemble des données disponibles sur l'expression des gènes induits par la blessure. L'induction des gènes codant pour des IP est mise en place aussi bien dans les feuilles endommagées (au niveau local) (pour exemple : Saarikoski, 1996) que dans toute la plante (systémie) (pour exemple : De Leo, 2001).

Chez la tomate, les ARNm correspondants à des IP de différentes classes (sérine, aspartique et métallocarboxypeptidase) sont détectables dans la plante deux heures après la blessure et leurs niveaux continuent de croître lors des huit heures suivantes. En comparaison, les messagers codant pour la prosystémine sont détectables à partir de trente minutes. C'est pourquoi les IP sont considérés comme faisant partie de la réponse tardive des plantes (Ryan, 2000).

Les IP inductibles sont régulés au niveau de la transcription grâce à la présence de motifs spécifiques dans la région promotrice de leur gène : le motif ABRE (ABscissic acid Respons Element) ainsi que les boîtes reconnues par les protéines G sont indispensables à l'induction du gène par la blessure (Koiwa, 1997 ; Ceci, 1995 ; De Leo, 2001).

Le cas des IP à sérine comme élément essentiel de la défense des plantes contre les prédateurs est particulièrement clair. En effet, la classe majeure de protéases présentes dans les plantes, utilisées lors de processus comme la mobilisation des protéines de réserves, sont des protéases de type cystéine. Les protéases à sérine ne sont apparemment pas activées dans les processus de dégradation des protéines endogènes. C'est pourquoi la présence d'une quantité significative d'IP à sérine dans les plantes n'est certainement pas destinée à la régulation de l'activité des protéases endogènes. Cela suggère donc que les cibles de ces IP soient les protéases digestives des animaux phytophages (Brzin, 1996, Gatehouse, 2000).

La démonstration directe du rôle négatif des IP sur la croissance des larves d'insectes a été faite en 1987. Des lignées de tabac transgéniques résistantes à l'attaque des larves d'*Heliothis virescens* ont été créées par expression constitutive de l'inhibiteur trypsique du niébé (CpTI) (Hilder, 1987). Lorsque les larves de premier stade se nourrissent de feuilles dans lesquelles CpTI représente 1% des protéines solubles, on observe après 7 jours, une mortalité de 62,5% des larves. Les larves survivantes présentent un retard de croissance (-70%). Ces plantes transgéniques étaient aussi, par conséquent, beaucoup moins défoliées.

**bleau 1 : Plantes transgéniques résistantes aux insectes par expression d'inhibiteurs de protéases**

35S, promoteur de l'ARN 35S du virus de la mosaïque du chou fleur ; actI, promoteur du gène de l'actine I du riz ; pin2, promoteur du gène dePPI-II ; assu, promoteur de la petite sous unité de la Rubisco *d'A. thaliana*

L, lépidoptère ; C, coléoptère ; H, homoptère

Inhibiteur	Promoteur	Expression (% des protéines totales solubles)	Plante	Ravageur cible	Effet observé			Références
					Mortalité	Retard de croissance	Réduction des dégats	
CpTI	35S	1%	Tabac	<i>H. virescens</i> (L)	60%	60%	50%	(Hilder, 1987)
	35S	1%	Pomme de terre	<i>L. oleacera</i> (L)	15%	45%	non	(Gatehouse 1997)
	actI	1,3%	Riz	<i>C. suppressalis</i> (L)	-	-	20-80%	(Xu, 1996)
				<i>S. inferens</i> (L)	-	-	20-80%	(Xu, 1996)
PI-II	pin2	0,6-2,5%	Riz	<i>S. inferens</i> (L)	-	70%	80%	(Duan, 1996)
				<i>C. erisoma</i> (L)	non	40-50%	-	(McManus, 1994)
	35S	0,1-1%	Tabac	<i>S.litura</i> (L)	non	non	-	(McManus, 1994)
				<i>T.orichalcea</i> (L)	non	non	-	(McManus, 1994)
				<i>S. exigua</i>	non	non	-	(Jongsma, 1995)
TPI-II	35S	3,3%	Tabac	<i>M. sexta</i> (L)	non	65%	oui	(Johnson, 1989)
TPI-I	35S	1,1%	Tabac	<i>M. sexta</i> (L)	non	50%	oui	(Johnson, 1989)
TPI-I	35S	1,2%	Tabac	<i>M. sexta</i> (L)	non	15%	non	(Johnson, 1989)
SpTI-I	35S	0,2%	tabac	<i>S.litura</i> (L)	95%	80-97%	oui	(Yeh, 1997)
MTI-2	35S	1,6%	Tabac	<i>S. littoralis</i> (L)	20%	50%	-	(De Leo, 1998)
		0,5%		<i>S. littoralis</i> (L)	non	non	non	(De Leo, 1998)
SKTI	35S	0,05-2,5%	Riz	<i>N. lugens</i> (H)	60%	-	-	(Lee, 1999)
Na-PI	assu	0,05-0,1%	pois	<i>H.armigera</i> (L)	25%	oui	-	(Charity, 1999)
	assu	0.1-0,3%	tabac	<i>H.armigera</i> (L)	30%	oui	-	(Charity, 1999)
OC-I	35S	2%	Peuplier	<i>C. tremulae</i> (C)	30-40%	-	-	(Leplé, 1995)
	35S	0,05-0,2%	Colza	<i>P. coqueariae</i> (C)	non	non	non	(Girard, 1998 b)
				<i>C. assimilis</i> (C)	non	non	non	(Girard, 1998 c)
Atcys	35S	-	Peuplier	<i>C. populi</i> (C)	100%	oui	-	(Delledonne, 2001)
M. sexta IP	35S	0,1%	coton	<i>B. tabaci</i> (H)	33-80%	-	-	(Thomas, 1995 b)
	35S	0,05-0,1%	tabac	<i>B. tabaci</i> (H)	98%	-	-	(Thomas, 1995 a)

Toutefois, on observe rarement ces effets délétères sévères lorsque les insectes sont nourris sur des milieux nutritifs artificiels contenant des concentrations en IP similaires à celles produites par les plantes blessées (Jongsma, 1997). Cette observation suggère que les IP agissent sur l'insecte en interaction avec d'autres molécules produites par la plante, comme les composés phénoliques ou alcaloïdes. Par conséquent, les inhibiteurs de protéases seraient une des composantes de la réaction de défense des plantes, et non l'unique effecteur.

### 1.3. Gain de résistance par sur-expression d'inhibiteurs de protéases

Les premiers travaux de Hilder en 1987 sont une démonstration directe du rôle défensif des IP contre les insectes. Depuis, une quinzaine de plantes transgéniques exprimant des IP se sont révélées efficaces contre les insectes cibles (tableau 1). Ce sont surtout des inhibiteurs de protéases à sérine (les plus étudiés chez les plantes), efficaces contre les larves de lépidoptères, qui ont été introduits dans ces plantes. Les bio-essais montrent que la diminution du poids des larves se développant sur les plantes transgéniques est proportionnelle au niveau de sur-expression de l'IP. De nombreuses lignées de plantes transgéniques contiennent une concentration relative d'IP de  $200 \mu\text{g}\cdot\text{g}^{-1}$  de tissus (concentration à laquelle les IP sont induits dans les feuilles de tomate ou de pomme de terre (Graham, 1986)).

Pourtant les lignées de tabac exprimant PI-II (IP à sérine de la pomme de terre) à un niveau similaire ne présentent pas de gain de résistance contre les larves de la noctuelle *Chrysodeixis eriosoma* (Mc Manus, 1994), suggérant la spécificité de l'interaction entre l'IP et l'espèce d'insecte considérée. Les IP ont un spectre d'action relativement large, néanmoins l'efficacité d'un IP peut varier d'un insecte à un autre. Le pouvoir entomopathogène d'un IP n'est pas identique sur toute la gamme des insectes sur lequel il est susceptible d'être actif.

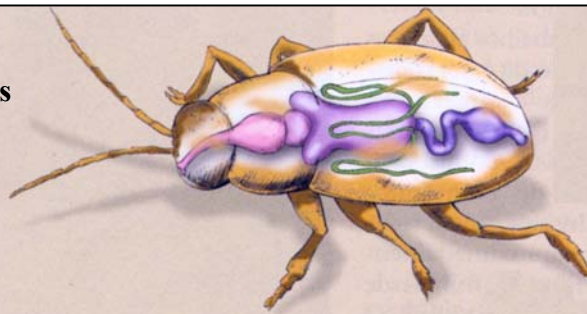
De plus, le niveau d'expression de l'IP introduit dans les plantes doit être souvent d'un niveau élevé (environ 1% des protéines solubles totales) (De Leo, 1998), bien que cela dépende de la sensibilité de l'insecte cible.

Bien que la gamme d'inhibiteurs de protéases, ainsi que celle des plantes d'intérêt agronomique transformées, soient en augmentation, la viabilité commerciale de cette stratégie n'est toujours pas démontrée.

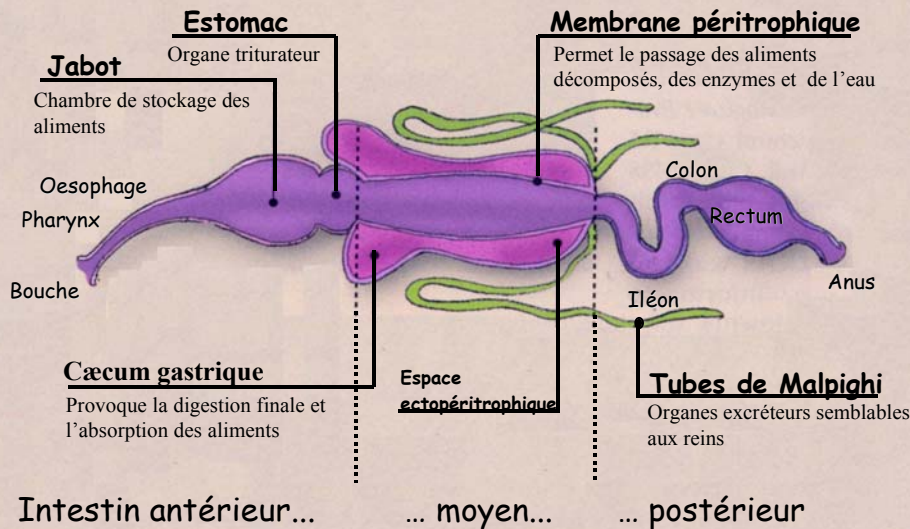
Le niveau de protection atteint grâce à la sur-expression d'IP dans les plantes transgéniques est dans certains cas supérieur à 50% en terme de réduction des dégâts, de mortalité des larves



## Le parcours et les transformations des aliments



a.

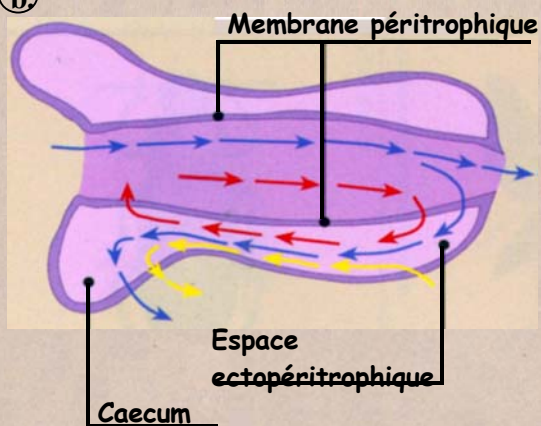


Après la mastication, l'aliment traverse l'**intestin antérieur** (recouvert de chitine et imperméable à la plupart des substances)

L'**intestin moyen** est le lieu où se déroule la production des enzymes, la plus grande partie de la digestion et de l'absorption des nutriments. L'aliment est attaqué par les enzymes et fractionné au niveau de la membrane péritrophique. Ensuite, les particules digérées traversent cette membrane et atteignent par un système de circulation les cellules qui absorbent les nutriments.

L'**intestin postérieur** intervient dans l'élimination des déchets, et dans l'équilibre de l'eau et des sels minéraux

b.



### Le recyclage des enzymes.

Les enzymes (flèche rouge) traversent la membrane péritrophique et arrivent dans l'espace ectopéritrophique accompagnant les aliments (flèche bleue) au fur et à mesure qu'ils sont fractionnés. Les enzymes retraversent ensuite la membrane, au niveau de la première partie de l'intestin moyen. L'eau (flèche jaune) passe par l'espace ectopéritrophique jusqu'au caecum dans lequel elle est absorbée.

Figure 4 : Représentation schématique de l'organisation du tube digestif des coléoptères et lépidoptères phytophages (d'après Terra, 1996).

a. Schéma de l'organisation du tube digestif.

b. Diagramme des flux aqueux et de la circulation des enzymes digestives.

ou de diminution de la biomasse. Mais ces niveaux ne sont pas considérés comme étant suffisants commercialement, ils n'atteignent pas ceux des variétés exprimant des toxines de *Bacillus thuringiensis* (quelques dizaines de microgrammes de toxine par gramme de tissus induisent plus de 95% de mortalité).

Des études au niveau biochimique et moléculaire ont donc été menées par de nombreux laboratoires afin de mieux comprendre les interactions entre les IP et la physiologie des insectes afin d'améliorer cette stratégie.

## **1.4. L'effet de l'ingestion d'inhibiteurs de protéases sur les protéases des insectes**

### **1.4.1. L'organisation de la digestion des protéines**

De nombreux travaux ont eu pour objet l'identification des enzymes digestives et la détermination de la structure fine du tube digestif chez des insectes d'ordres et de régimes alimentaires différents (pour revue : Terra, 1994). Le tube digestif des larves d'insectes phytophages holométaboles possède une structure relativement simple. Constitué d'un tube reliant la bouche à l'anus, il est bordé d'une couche mono-cellulaire de cellules formant un épithélium et se divise en trois parties (figure 4). Il semble que l'intestin moyen soit le site majeur de la digestion des protéines. Son organisation présente deux particularités structurales :

- (1) une membrane péritrophique délimite deux zones du tube digestif médian (l'espace endopéritrophique et l'espace ectopéritrophique). Cette membrane est constituée d'un entrelac de fibrilles de chitine pris dans une matrice de glycoprotéines. Elle est issue, dans la majorité des cas, de la condensation de sécrétions de l'épithélium du tube digestif médian. La membrane péritrophique assure une protection mécanique de la bordure en brosse (prévention de l'attachement non spécifique de matériel non digéré sur les hydrolases et/ou les transporteurs membranaires) qui constitue une barrière physique contre les micro-organismes et représente un tamis moléculaire grâce à sa propriété semi-perméable (Lehane, 1997). Ses fonctions sont similaires à celles du mucus gastro-intestinal des mammifères, et ses glycoprotéines constitutives possèdent des homologies avec les mucines du tube digestif des mammifères,

- (2) un épithélium constitué majoritairement de cellules dites colonnaires. Ces cellules sont pourvues à leur surface de microvillosités formant la “ bordure en brosse ”. Cet épithélium est présent sur toute la longueur du tube digestif médian.

La localisation des enzymes présents au sein de l'intestin moyen, compte tenu de son organisation structurale, permet de mettre en évidence les deux caractéristiques essentielles de la digestion chez les insectes phytophages holométaboles : la digestion est séquentielle et compartimentée (figure 4b). Les endoprotéases sont secrétées par les cellules de l'épithélium, puis libérées dans l'espace endopéritrophique où elles initient les premières étapes de la digestion des protéines. Les peptides, issus de cette première dégradation, passent dans l'espace ectopéritrophique où les exopeptidases (souvent liées à la surface de l'endothélium) achèvent la digestion en libérant les acides aminés.

L'existence d'un courant d'eau dans l'espace ectopéritrophique, progressant en sens inverse (postérieur⇒antérieur) du bol alimentaire situé dans l'espace endopéritrophique (antérieur⇒postérieur), permet une re-circulation des protéases (figure 4b). Les protéases de l'espace endopéritrophique passent dans l'espace ectopéritrophique au niveau de la partie postérieure et remontent vers la partie antérieure de l'intestin, puis entrent à nouveau dans l'espace endopéritrophique (figure 4b). Le flux d'eau ainsi créé permet de récupérer par simple effet de diffusion toutes les molécules présentes dans l'espace endopéritrophique et suffisamment petites pour passer à travers les pores de cette membrane ( $\leq 8\text{nm}$ ). Chaque enzyme peut donc effectuer plusieurs passages dans l'espace endopéritrophique.

Les protéases de classes différentes agiraient dans des zones différentes de l'espace endopéritrophique, en relation avec les valeurs locales du pH intestinal. De manière générale, des protéases à sérine (type trypsine, chymotrypsine, élastase), des leucines aminopeptidases et des carboxypeptidases A et B ont été mises en évidence chez les lépidoptères (Christeller, 1992 ; Broadway, 1986). La plupart des ARNm codant pour des protéases à sérine sont détectés au niveau de l'intestin moyen de *M. sexta* (Peterson, 1994). Ces protéases seraient secrétées par les cellules de l'intestin moyen puis traverseraient l'espace ectopéritrophique afin d'atteindre la lumière de l'intestin antérieur (Terra, 1996).

Les coléoptères possèdent des protéases à cystéine en association avec des protéases à sérine et/ou des protéases à aspartyl<sup>4</sup> (Thie, 1990 a ; Thie, 1990 b). Des leucines aminopeptidases et des carboxypeptidases A et B ont également été mises en évidence chez certains coléoptères. Les travaux de Thie et Houseman (1990) sur la larve de *Tenebrio molitor* ont démontré qu'un système complexe, combinant deux familles de protéases, pouvait exister au sein d'un même insecte. Ces activités endoprotéasiques localisées à plus de 90% dans le lumen, sont

spatialement séparées dans l'intestin moyen : une activité protéases à cystéine dans la première moitié de l'intestin moyen (pH 5-6) et une activité relative aux protéases à sérine dans la seconde moitié (pH 7-9). La localisation de ces activités est parfaitement corrélée avec les valeurs de pH déterminées pour chaque partie du tube digestif (Terra, 1985). Ce rapprochement montre donc qu'il existe un système de régulation fine du pH de la lumière intestinale. Ce système permet d'obtenir localement des valeurs de pH favorables à une activité optimale des endoprotéases digestives. Il est en outre suffisamment efficace pour que des valeurs de pH, éloignées entre elles de plus de deux unités, puissent être atteintes dans l'intestin moyen. Grâce à la circulation entre les espaces endopéritrophique et ectopéritrophique, il est probable que les protéases des deux classes fonctionnelles parcourent l'intestin moyen dans sa longueur. Les endoprotéases rencontrent donc tour à tour des conditions d'activité, notamment de pH, favorables et défavorables. Les insectes pourvus d'un tel système de digestion des protéines ont dû développer des endoprotéases présentant des spécificités fonctionnelles leur permettant de résister au passage dans des conditions de milieu *a priori* défavorables, voire inactivantes. Ainsi, un environnement réducteur nécessaire à une activité maximale des protéases à cystéine, est capable de dénaturer les trypsines de mammifères. L'activité endoprotéasique à sérine observée chez les coléoptères, semble être activée par les composés thiols (Thie, 1990 a). Cette diversité des mécanismes réactionnels de dégradation pourrait être un avantage adaptatif important. Les protéases à cystéine pourraient ainsi permettre de contourner la résistance liée à la production de serpins par la plante.

En général, on observe une excellente corrélation entre le pH intestinal de l'insecte et la classe des protéases digestives prédominantes. Les insectes utilisant des protéases à sérine (lépidoptères, orthoptères, carabes) ont un pH intestinal basique (8-11), alors que les coléoptères qui utilisent des protéases à cystéine présentent un pH intestinal légèrement acide (6-7).

## **1.4.2. Les effets complexes de l'ingestion des inhibiteurs de protéases sur la régulation des protéases.**

### *1.4.2.1. Déséquilibre du pool d'acides aminés*

De nombreuses études ont été menées afin de déterminer l'effet de l'ingestion d'inhibiteurs de protéases sur la survie et le développement de l'insecte. Ces essais ont été effectués *in vivo* grâce à l'incorporation de l'IP dans un milieu nutritif artificiel (Christeller, 1992 ; Christeller, 1994 ; Bown, 1998 ; Liang, 1991). En comparaison avec le milieu nutritif contrôle, la présence d'un IP efficace permet d'observer une augmentation de la mortalité des

larves, une perte de poids et un retard de développement (Gatehouse, 1999 ; pour revue Jongasma, 1997). La plupart de ces auteurs ont évalué l'inhibition des protéases présentes dans le tube digestif de l'insecte. Grâce à l'utilisation d'un substrat synthétique spécifique de chaque classe d'enzyme, l'inhibition spécifique de l'IP peut être mise en évidence. Ce type d'essais permet d'effectuer un criblage afin de sélectionner l'IP qui possède la meilleure affinité pour les protéases de l'insecte cible. Ces tests en alimentation artificielle ont révélé le potentiel des IP d'origine végétale pour l'inhibition des protéases digestives des insectes. Ces molécules inhibent les enzymes de la digestion primaire des protéines ingérées. Ce sont des substances antimétaboliques et antinutritives responsables d'une réduction du taux de protéolyse et donc d'une moindre disponibilité en acides aminés de l'alimentation conduisant à un retard de croissance larvaire (Burgess, 1994 ; Orr, 1994; De Leo, 1998). Un enrichissement du milieu nutritif en acides aminés soufrés (méthionine, cystéine) élimine les effets délétères des IP sur la croissance larvaire (Gatehouse, 1983 ; Oppert, 1993 ; Broadway, 1996). Ces auteurs montrent également que les retards de croissance observés sont corrélés à une hyper-production de protéases digestives en réponse à l'ingestion d'IP. Le taux de digestion des protéines *in vivo* peut ainsi être maintenu malgré la présence de l'IP. Cette hyper-production de protéases mobiliserait des acides aminés soufrés au détriment de la synthèse d'autres protéines essentielles. Le détournement métabolique ainsi induit, semble conduire à un déséquilibre du stock d'acides aminés disponibles qui aboutit finalement à l'apparition de carences en certains d'entre eux (certainement correspondant aux acides aminés faiblement représentés dans la nourriture ingérée). Les IP n'agissent donc pas forcément par simple blocage de l'assimilation des protéines mais aussi par mobilisation des réserves d'acides aminés de la larve en vue de la synthèse de protéines.

#### 1.4.2.2. *Stratégie adaptative des insectes face à l'ingestion d'IP*

Malgré les résultats encourageants que l'on observe grâce à l'expression d'inhibiteurs de protéases exogènes pour lutter contre les insectes phytophages, certaines lignées transgéniques ne présentent aucun gain de résistance, voire des effets inverses, les insectes se développent beaucoup mieux (effets probiotiques).

Les effets de l'ingestion d'inhibiteurs de protéases sur les insectes sont très complexes. Chaque espèce d'insectes, voire chaque population d'insecte, réagit différemment à la sur-expression d'inhibiteurs de protéases dans son alimentation (Girard, 1998 c). Grâce à l'étude de quelques cas, il a été mis en évidence trois stratégies adaptatives :

- (1) La détoxification de l'alimentation **par dégradation de l'IP**. Par exemple, les peupliers transgéniques exprimant constitutivement l'oryzacystatine (OC-I) présentent un gain de résistance aux larves du coléoptère *C. tremulae*. En effet, 40% des larves meurent après ingestion des feuilles de ces peupliers en fin de cycle. Pourtant, aucun effet délétère n'est observé lorsque les larves de la phédone du cresson (*P. cochleaciae*) se nourrissent sur des colzas transgéniques exprimant OC-I (Girard, 1998b). De plus, aucun changement au niveau de la nature des protéases ou de leur niveau d'activité n'est observé, aucune réponse de compensation n'est mise en place. La résistance des larves de phédone semble être médiée par la dégradation d'OC-I qui perd rapidement son activité. Ce système de détoxification met en jeu des protéases à sérine en association avec des leucine-aminopeptidases déjà présentes dans le tube digestif de la larve (Girard, 1998 a).

- (2) **L'adaptation favorable suite à un faible niveau d'expression de l'IP**. Lorsque l'inhibiteur tryptique de la moutarde MTI-2 est exprimé à un taux de 1,6% des protéines solubles dans les feuilles d'*A. thaliana* ou de tabac transgéniques, des effets délétères sur les larves du lépidoptère *S. littoralis* (20% de mortalité, 50% de réduction de croissance, 56% de réduction de la fertilité) sont observés ainsi qu'une réduction des dommages foliaires (De Leo, 1998 ; De Leo, 2002). Par contre, lorsque ces larves se nourrissent de feuilles de plantes transgéniques exprimant l'inhibiteur à un faible niveau (0,5% des protéines solubles), les larves se développent plus rapidement et plus efficacement, provoquant des dégâts beaucoup plus importants sur la plante (+ 26%). Ces observations sont corrélées à une sur-expression de protéases pré-existantes et sensibles à MTI-2 *in vitro*. L'accroissement de la consommation foliaire peut être la conséquence d'une diminution de la qualité nutritive de l'aliment du fait de la présence de l'IP, et/ou d'une activité protéolytique accrue chez ces larves.

- (3) **L'induction de protéases insensibles à l'IP**. De nombreuses équipes ont pu révéler la *néo*-synthèse de protéases insensibles à l'inhibiteur ingéré (Broadway, 1995 ; Jongsma, 1995 ; Bown, 1997). L'équipe de M. Jongsma a pu montrer, que contrairement aux résultats encourageants obtenus avec les tests en milieux artificiels, le développement des larves du lépidoptère *S. exigua* n'est pas affecté par l'expression constitutive de PI-II (inhibiteur de protéases à sérine de la pomme de terre) dans des tabacs transgéniques. Ces insectes mettent en place un mécanisme d'adaptation par induction de protéases insensibles à l'inhibiteur. Exposées continuellement à l'inhibiteur, les larves modifient la nature de leurs

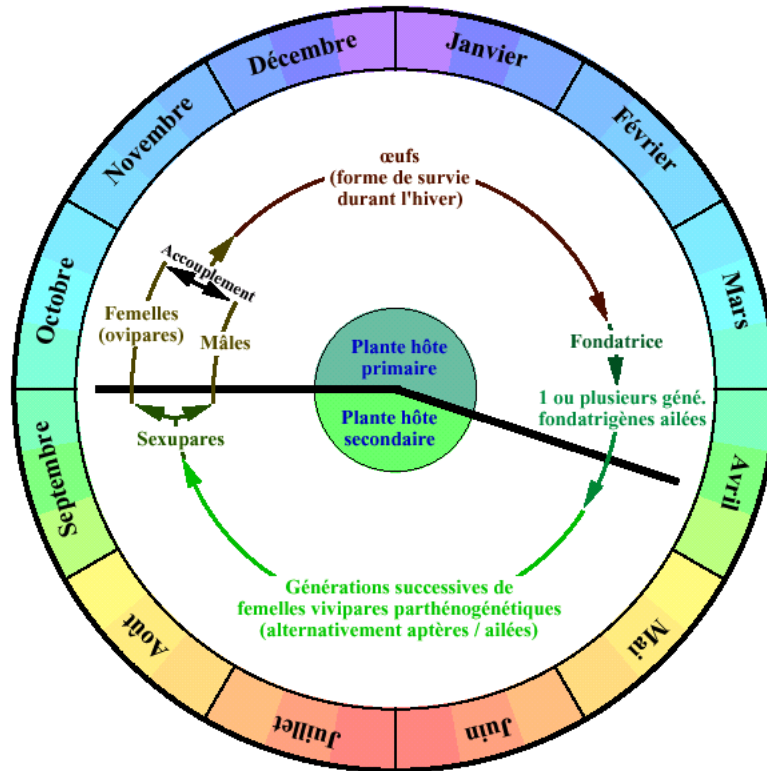
protéases, réduisant ainsi leur sensibilité et abolissant l'effet antimétabolique du produit du transgène.

Bown *et al.* (1997) ont pu éclaircir le mécanisme moléculaire impliqué dans cette réponse adaptative. Le génome du lépidoptère *H. armigera* renferme au moins vingt huit gènes codant des protéases à sérine. Bien que les résidus du site actif soient conservés, les résidus impliqués dans le contact protéase-inhibiteur sont variables entre les différentes isoformes (démontrant à nouveau la très fine spécificité d'interaction entre l'inhibiteur et sa cible). Ainsi, l'expression de ces différents gènes peut être différentiellement induite selon la nature de l'inhibiteur. En présence de 0,5% de SKTI (Soybean Kunitz Trypsin Inhibitor) dans le milieu de nutrition artificielle, la régulation de la synthèse des protéases est complexe : le niveau des messagers codant pour certaines chymotrypsines est augmenté alors que celui de certaines trypsines est réduit (d'autres présentent peu de variation).

Récemment, Mazumdar et Broadway (2001) ont mis en évidence la cinétique d'induction de ces différentes classes de protéases. La réponse des larves de lépidoptères se décompose en deux phases : tout d'abord, il y a sur-expression de protéases sensibles à partir d'un pool d'ARNm pré-existants, et ensuite induction d'une synthèse *de novo* de messagers codant pour des protéases insensibles. Ces protéases insensibles sont détectées deux heures après l'ingestion de l'inhibiteur (Broadway, 1997).

Enfin, l'étude de l'interaction entre *H. armigera* et sa plante hôte (chickpea) montre que les protéases de la larve sont capables de dégrader les inhibiteurs de protéases synthétisés par la plante pour se protéger (Giri, 1998). L'induction de protéases insensibles aux inhibiteurs de protéases endogènes, induits lors de la blessure, semble être un mécanisme fréquent grâce auquel les insectes peuvent contourner l'accumulation d'IP.

Ces résultats démontrent la rapidité de la réponse adaptative, la plasticité de la physiologie digestive des insectes, ainsi que la complexité de la régulation de tous ces gènes codant des protéases.



**Figure 5 : Représentation schématique du cycle de vie des pucerons en région tempérée.**

Dans une région donnée, l'adaptation des pucerons aux conditions particulières du milieu peut se traduire par de nombreuses variantes au niveau des phases du cycle (changement de plante hôte appartenant à une même espèce ou non ; perte de la possibilité de se reproduire par voie sexuée). La représentation ci-dessus est donc très schématique et ne traduit que partiellement la complexité des cycles chez les pucerons.



## 2. Les pucerons

Les pucerons ou aphides appartiennent à l'ordre des Homoptères<sup>5</sup> au même titre que les cicadelles, les psylles, les aleurodes et les cochenilles. Ces insectes phytophages constituent un groupe d'insectes extrêmement répandu dans le monde. Ils sont apparus il y a environ 280 millions d'années. Ce groupe s'est diversifié parallèlement à celui des angiospermes (plantes à fleurs) dont presque toutes les espèces sont hôtes. 4700 espèces de pucerons ont été répertoriées et bien qu'elles soient répandues dans le monde entier, c'est dans les zones tempérées que l'aphidifaune est la plus variée (Dixon, 1987). La plupart des genres de pucerons (90%) sont inféodés à une famille végétale particulière, c'est pourquoi ils sont dits monophages. Cependant, les pucerons s'attaquant aux cultures (seulement quelques centaines, (Leclant, 1994)) ont un régime alimentaire moins restreint. Ils sont en général oligophages, voire polyphages, et peuvent se nourrir sur des végétaux de familles très distinctes. Ainsi, le puceron vert du pêcher *Myzus persicae* est considéré comme l'un des pucerons les plus polyphages. Il peut se nourrir de plantes appartenant à plus de 40 familles différentes, dont la plupart sont d'intérêt économique (*i.e.* melon, tomate, pomme de terre, courgette, colza, épinard, laitue...).

### 2.1. Biologie et écologie des pucerons

Le cycle annuel complet « canonique » repose sur une alternance entre une génération sexuée, comportant des mâles et des femelles fécondables (génération amphisexuelle) et de nombreuses générations parthénogénétiques (figure 5). Les femelles fécondées sont toujours ovipares et donnent naissance à un œuf, alors que les femelles parthénogénétiques sont vivipares et donnent naissance à de jeunes larves ou nymphes. En fonction des espèces, ce cycle est effectué soit sur une même espèce de plante ou associé à une série de plantes de la même famille botanique, soit avec une alternance d'hôtes obligatoires. L'hôte primaire est généralement une espèce pérenne sur laquelle sont pondus les œufs. Au printemps, après quelques générations, des individus ailés assurent la dissémination sur les hôtes secondaires.

---

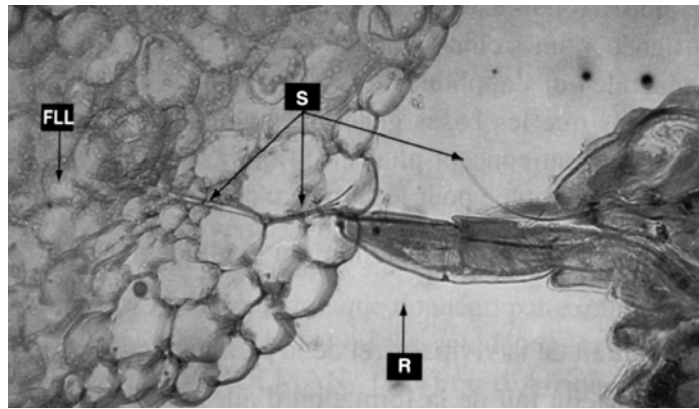
<sup>5</sup> Règne : - Animal, Embranchement : Arthropodes, Classe : Insectes, Sous-classe : Ptérygotes, super-ordre : Hémiptéroïdes, Ordre : Homoptères, Sous-ordre : Aphidomorpha ; Super-famille : Aphidoidea, Famille : Aphididae.

Les pucerons sont des insectes hémimétaboles. En effet, les différents stades larvaires (L1, L2, L3 et L4) ont la même morphologie (mis à part l'absence d'ailes chez les larves de futurs ailés) et le même mode de vie que les adultes. Le développement est rapide, chez *A. gossypii* et *M. persicae*, une semaine environ est nécessaire pour le développement larvaire aux températures optimales. L'apparition des ailes est effective après la mue imaginale au stade adulte et semble induite par certaines contraintes environnementales (densité de la population, qualité de la plante hôte, baisse de la photopériode et de la température) (Dixon, 1987).

Au cours du cycle, les différentes morphes (aptère/aillée) apparaissent et jouent un rôle particulier. La formation des ailes, qui correspond à une réponse rapide à la détérioration du milieu, permet d'établir en peu de temps de nouvelles colonies et de suivre la succession des végétaux en croissance au cours de l'année (vol de dissémination). Sur ces plantes hôtes, plus favorables, apparaissent ensuite des individus parthénogénétiques qui vont permettre une multiplication rapide de la population. Ce mode de reproduction représente un gain de temps considérable puisque la rencontre entre les deux sexes n'est pas nécessaire et surtout cela aboutit à la formation de femelles uniquement, multipliant par 2 le potentiel de reproduction d'une génération par rapport à un sexe ratio de 50/50. De plus, les embryons se développent bien avant que la mère n'ait atteint l'âge adulte. Une jeune larve, à peine née, porte déjà en elle les embryons de la génération suivante. Un calcul théorique<sup>6</sup> montre les potentialités démographiques exceptionnelles des pucerons : un puceron possède une fécondité moyenne d'une trentaine de larves dont la durée de développement (de la naissance jusqu'à la maturité sexuelle) est de 14 jours. A raison de neuf générations par an pendant la belle saison, un seul individu pourra donner naissance à  $600 \cdot 10^9$  individus (Robert, 1981). Ces insectes possèdent un pouvoir de multiplication très important quand les conditions trophiques et climatiques leur sont favorables, conduisant à des explosions démographiques.

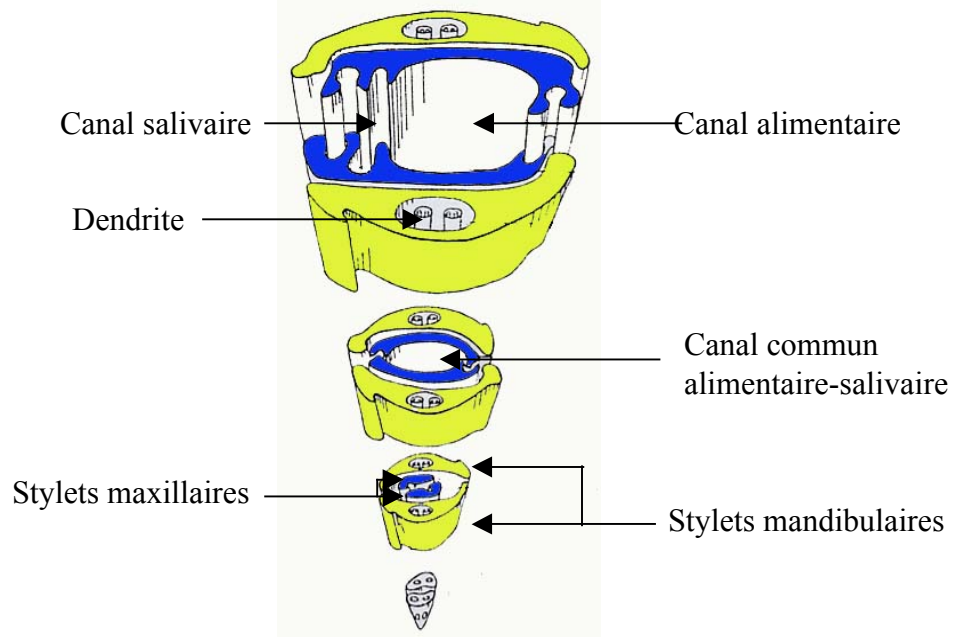
---

<sup>6</sup> Ce calcul est évidemment irréaliste puisqu'il ne tient compte que des facteurs favorables à la multiplication des pucerons.

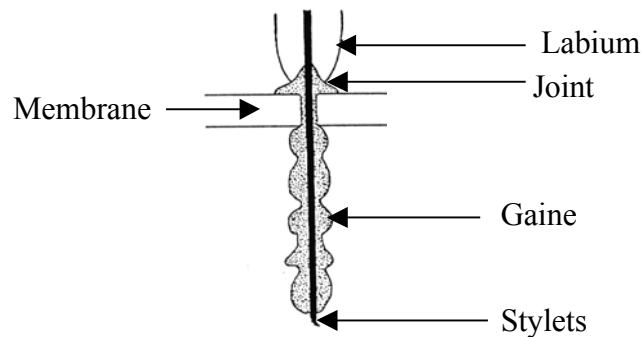


**Figure 6a: Puceron en phase d'alimentation (Leclant, 1981).**

FLL : Faisceau Libéro-Ligneux, S : Stylet, R : Rostre.



**Figure 6b: Représentation en perspective de l'extrémité des stylets d'un puceron (Leclant, 1981).**



**Figure 6c : Schématisation des relations des stylets avec la gaine de salive lors de l'alimentation du puceron à travers une membrane artificielle (Miles, 1999).**

## 2.2. Mode alimentaire et nutrition

Le développement de ce paragraphe devrait permettre de mettre en avant les particularités intrinsèques aux pucerons par rapport aux insectes phytophages tels que les lépidoptères ou les coléoptères déjà cités.

### 2.2.1. Un insecte phloémophage

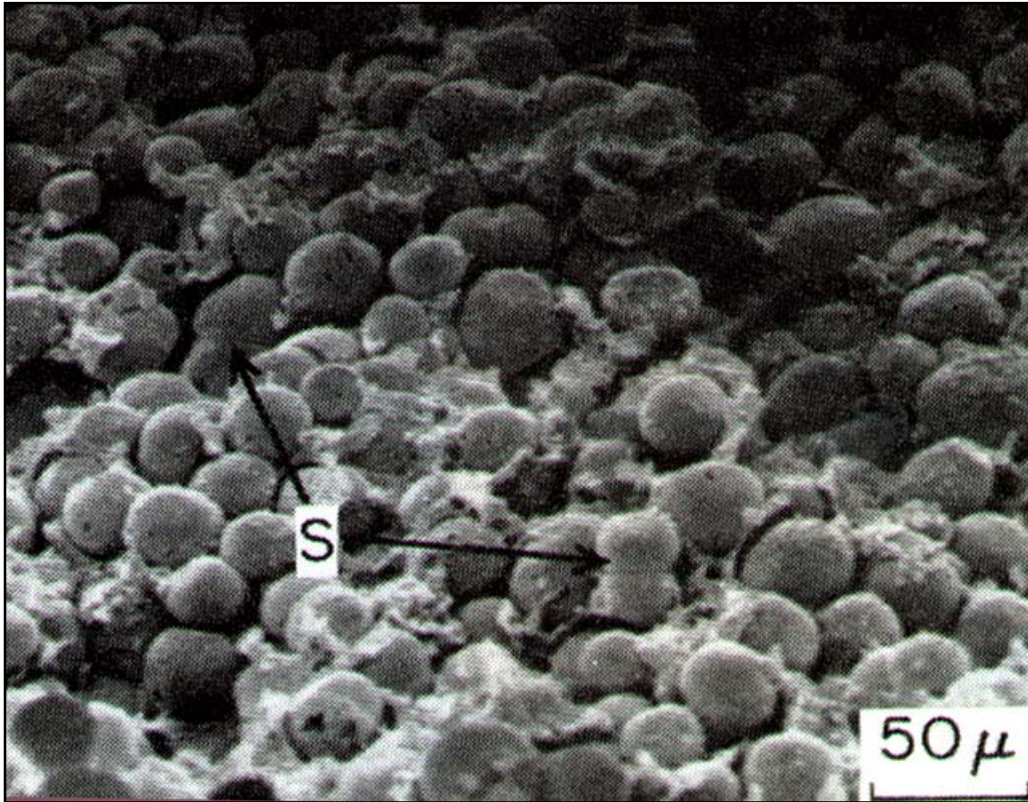
Les pucerons sont des insectes piqueurs-suceurs. Leurs pièces buccales sont transformées et adaptées au prélèvement d'une alimentation liquide non disponible en surface. Leur système buccal a été spécialisé pour percer la paroi du végétal et atteindre les faisceaux criblo-vasculaires du phloème où il prélèvera la sève élaborée (figure 6a). Les mandibules et les maxilles ont la forme de stylets (au nombre de quatre) qui coulissent les uns par rapport aux autres et ménagent entre eux deux canaux : un canal descendant provenant des glandes salivaires et un canal ascendant qui conduit la nourriture vers la tête puis vers le tube digestif (figure 6 b).

Lorsque le puceron est installé sur la surface de la feuille, il commence à collecter des informations provenant de stimuli physiques et chimiques (des composés volatiles, des cires épicuticulaires peuvent donner des informations sur l'état physiologique et sur l'espèce de plante) (cité dans Chen, 1996a). Il secrète alors une goutte de salive qui se gélifie rapidement au contact de l'air (figure 6c). Les enzymes hydrolytiques<sup>7</sup> de la salive permettent la dissolution de la cuticule et semblent n'être utilisées que pour l'exploration (Miles, 1999). Le puceron insère ensuite ses stylets dans la couche épithéliale de son hôte. Il utilise alors des chémorécepteurs situés dans le canal alimentaire, au niveau du pharynx, afin d'apprécier l'appétence de la plante (piqûre d'essai). Lorsqu'un site favorable a été choisi, le processus de pénétration des stylets dans les tissus de la plante commence. Les stylets doivent traverser la cuticule, l'épiderme et le mésophylle afin d'atteindre le site d'alimentation. La plupart des pucerons, grâce à la flexibilité de leurs stylets, traversent les tissus extracellulairement, entre les cellules du parenchyme.

La sève circule sous pression dans le phloème, et dès que l'insecte pique un faisceau, le flot s'engage dans le canal ascendant des stylets après une courte période de salivation (Tarn, 1982).

---

<sup>7</sup> Pectinases, cellulases, amylases, protéases, lipases, peroxydases, phosphatases alcaline et acide. La composition de la salive est complexe et dépend de chaque espèce de puceron (Miles, 1999)



**Figure 7 : Observation par microscopie électronique de *Buchnera aphidicola* à l'intérieur d'un bactériocyte d'*A.pisum* (McLean & Houk, 1973).**

Les flèches indiquent la taille variable et une division du symbiote (S).

Au fur et à mesure que l'insecte enfonce ses stylets dans la plante, il injecte de la salive qui durcit autour de ceux-ci, formant une gaine sétale de nature protéique : sorte de fourreau à l'intérieur duquel l'aphide peut manœuvrer ses stylets et les ressortir s'il est importuné ou effectuer une nouvelle piqûre. Cette gaine restera dans la plante après le retrait des stylets (Srivastava, 1987). Il semble que la gaine permette au puceron de limiter l'ingestion de toxines synthétisées par la plante en réponse à la blessure<sup>8</sup> (Miles, 1999).

Les aphides se nourrissent quasiment exclusivement de sève élaborée. Cet aliment étant relativement pauvre, les pucerons l'ingèrent en grande quantité. La sève phloémienne est riche en sucres (0,15-1 M), pratiquement dépourvue de lipides et très déséquilibrée en acides aminés libres (3-250mM, l'azote est véhiculé à plus de 70% sous forme d'amines, asparagine ou glutamine<sup>9</sup>). Un grand nombre d'insectes, se développant sur des milieux carencés du point de vue nutritionnel, vivent en symbiose intégrée avec des microorganismes intracellulaires (Buchner, 1965).

### 2.2.2. Une symbiose obligatoire

Les pucerons vivent tous en symbiose avec une bactérie intracellulaire, une Entérobactériacée du genre *Buchnera*, proche d'*E.coli* (Charles, 1999). Cette dernière possède une paroi cellulaire Gram-négative et présente un diamètre de 2,5 à 4 µm (figure 7). Cette bactérie est présente dans toutes les morphes d'aphides et chez toutes les espèces (hors familles des Lachnidae et Phyloxeridae, groupes basaux dans le sous-groupe des aphidimorpha).

Contrairement à d'autres associations microbiennes mises en évidence chez les pucerons, cette association à *Buchnera* est définie selon trois caractéristiques (Douglas, 1998) :

- les microorganismes sont intracellulaires et restreints au cytoplasme d'un type de cellule spécialisée, les bactériocytes. Les bactéries occupent 60% du cytoplasme du bactériocyte.

---

<sup>8</sup> Bien que les cellules parenchymateuses possèdent de grands espaces intercellulaires permettant une certaine mobilité lors de la pénétration du stylet, il arrive souvent que ce dernier traverse la paroi végétale et passe le long du plasmalemme ou le perce. Une toute petite quantité de salive est alors déposée afin de combler le trou avec un matériau biologiquement compatible, prévenant ainsi de la libération et/ou de production de composés de défense toxiques (Tjallingii et Esch, 1993 cité dans (Miles, 1999).

<sup>9</sup> Les 9 acides aminés essentiels (que l'animal ne sait pas synthétiser *de novo*) sont en concentration très faible. Ces acides aminés sont l'histidine, l'isoleucine, la leucine, la lysine, la méthionine, la phénylalanine, la thréonine, la valine et le tryptophane.

A leur naissance, les pucerons renferment environ une centaine de bactériocytes qui ne se diviseront plus ;

- les bactéries sont transmises par hérédité maternelle (transmission verticale) ;

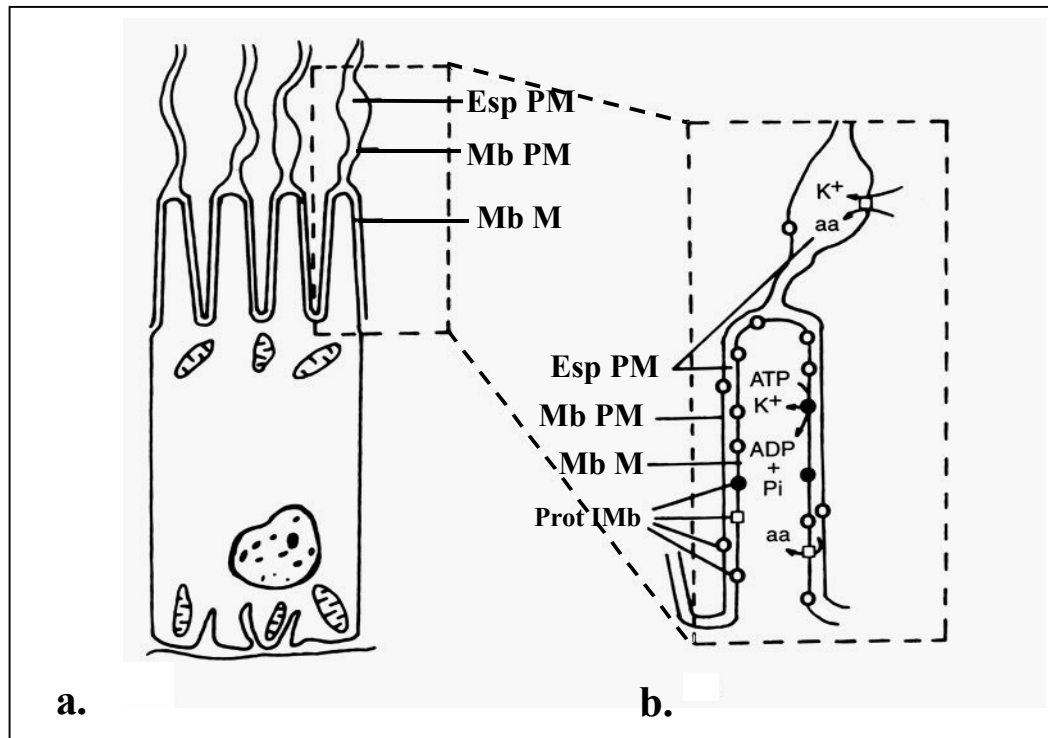
- c'est une association obligatoire aussi bien pour l'insecte que pour la bactérie. Cette dépendance a été mise en évidence grâce à l'obtention de pucerons rendus aposymbiotiques après injection ou ingestion d'un antibiotique (Rahbé, 1994). Ces pucerons sont viables mais de petite taille et stériles. C'est grâce à l'obtention de ces deux types de pucerons que le rôle de *Buchnera* a pu être élucidé. D'autre part, il est impossible de multiplier *Buchnera* en culture et cette bactérie n'a été rapportée dans aucune autre niche écologique que les pucerons.

La bactérie endosymbiotique tient un rôle primordial dans le métabolisme intermédiaire conduisant à une composition équilibrée et régulée du pool d'acides aminés libres utilisés par le puceron pour ses synthèses protéiques.

Le métabolisme azoté chez les pucerons est étudié à l'aide d'acides aminés marqués au  $^{14}\text{C}$  incorporé dans l'alimentation artificielle. Par exemple, le puceron du pois *A. pisum*, synthétise l'isoleucine, la lysine et la thréonine à partir du glutamate (Febvay, 1995) et tous les acides aminés essentiels à partir du saccharose (Febvay, 1999). Ces données physiologiques ont été confirmées grâce au séquençage du génome entier de *Buchnera* (Shigenobu, 2000). Le génome comprend un chromosome circulaire de 641 kb et deux plasmides circulaires, pLeu (7,8 kb, 7 ORF dont l'opéron *LeuABCD* qui code pour les enzymes spécifiques de la synthèse de la leucine) et pTrp (portant le gène *trpEG* codant pour la première enzyme de la voie de biosynthèse du tryptophane). Une des caractéristiques fonctionnelles de ce génome est la conservation de la plupart des voies de biosynthèse des acides aminés essentiels. Inversement, les voies de biosynthèses des acides aminés non-essentiels sont très incomplètes suggérant une auxotrophie de *Buchnera* pour ces acides aminés<sup>10</sup>. Le glutamate, composé azoté essentiellement ingéré par le puceron, pourrait être transaminé en aspartate dont la plupart des acides aminés essentiels sont dérivés. Ces 2 acides aminés pourraient être transportés jusqu'à la bactérie où la synthèse des acides aminés essentiels a lieu. Toutefois, lors de l'analyse du génome du symbiote, aucun gène codant pour un transporteur d'acide aminé n'a été annoté. Les voies de régulation de biosynthèse des acides aminés et leur régulation sont en cours d'analyse au laboratoire BF2I INRA/INSA de Lyon grâce au développement de puces à ADN.

---

<sup>10</sup> Ceci peut-être interprété comme l'établissement au cours de l'évolution d'interdépendance entre l'hôte et son symbiote (Zientz, 2001), la symbiose obligatoire étant très ancienne (250 millions d'années).



**Figure 8 : Modèle de la structure et du rôle physiologique de bordure en brosse des cellules du tube digestif des Hémiptères (Terra, 1996).**

(a) Représentation schématique d'une cellule du tube digestif ; (b) détail de l'apex. La membrane microvillaire (Mb M) est bordée par la membrane périmicrovillaire (Mb PM). Ces deux membranes délimitent un espace fermé, l'espace périmicrovillaire (Esp PM). La membrane microvillaire est riche en protéines intramembranaires (Prot IMb). La membrane transporte activement le potassium de l'espace périmicrovillaire dans les cellules du tube digestif, générant un gradient de concentration entre la lumière du tube digestif et l'espace périmicrovillaire. Ce gradient représente une force qui permet l'absorption active de composés organiques (ex : acides aminés) grâce à des transporteurs spécialisés à la surface de la membrane microvillaire.



### 2.2.3. La digestion

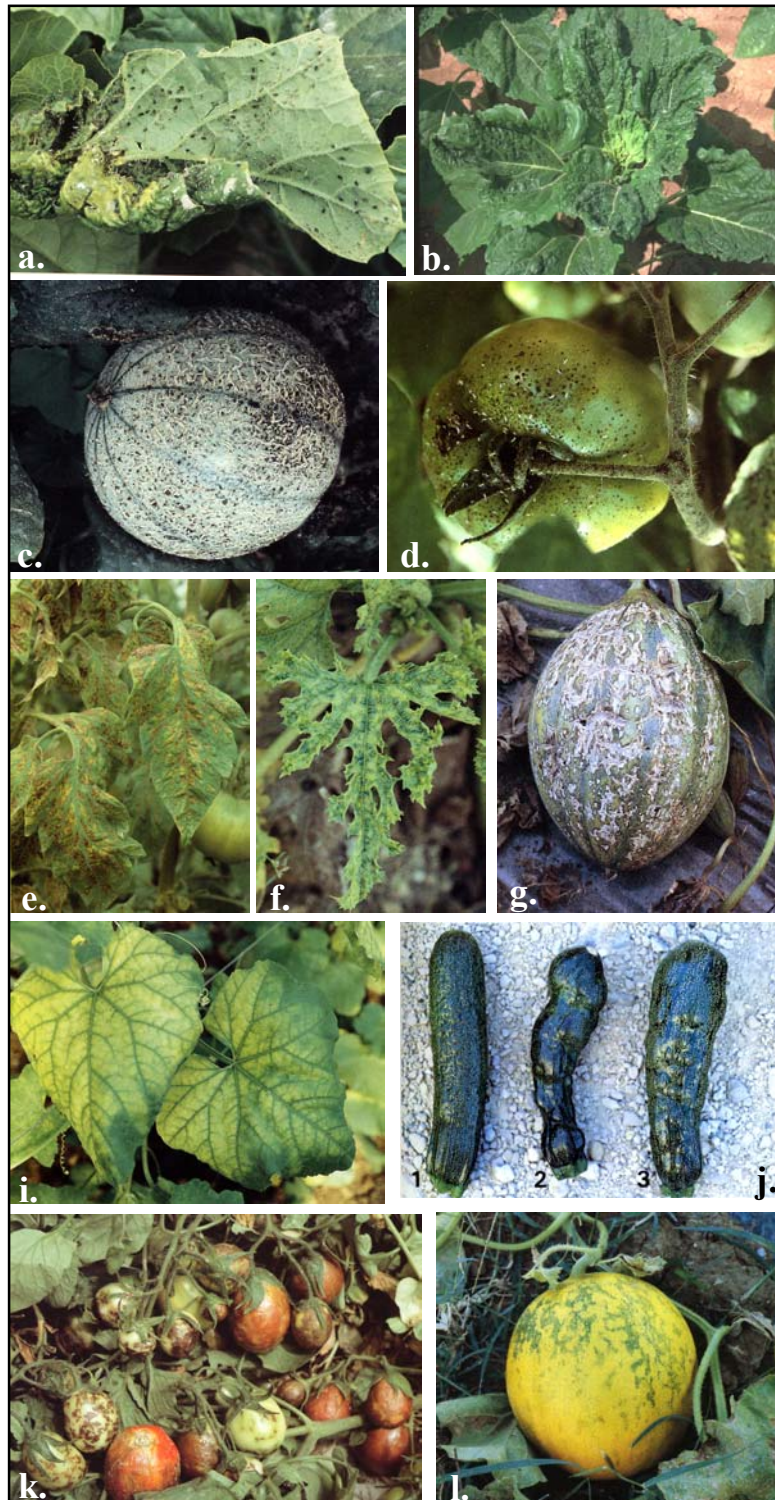
L'organisation générale du tube digestif des pucerons paraît relativement simple comparée à celle d'autres insectes (figure 4). La description structurale du tube digestif et des cellules qui le composent sont le fruit du remarquable travail de Ponsen (1972). Le tube digestif des pucerons est constitué uniquement d'un très fin œsophage, d'un estomac, d'un intestin et d'un « colon ». Ils ne possèdent ni de jabot, ni de membrane péritrophique, ni de système urinaire (tubes de Malpighi).

Les insectes piqueurs-suceurs se nourrissent d'une alimentation pauvre en protéines (moins de 1% des acides aminés circulant dans le phloème (Rahbé et Febvay, 1993) et de polymères carbonés. Aussi, excepté pour l'hydrolyse des sucres simples, aucune digestion des aliments ne semble nécessaire chez les Homoptères (Terra, 1994). Il semble qu'en raison de la spécificité du type de nutriments ingérés, les Homoptères aient subi une spécialisation de leur système digestif leur permettant d'absorber les acides aminés à partir d'une alimentation très diluée. Les microvillosités des cellules épithéliales de l'intestin moyen sont recouvertes (du côté luminal) par une membrane, la membrane périmicrovillaire. Cette dernière ne renferme pas de protéines intégrées. Par contre, la membrane microvillaire est riche en complexes protéiques. Des transporteurs actifs permettraient d'importer le potassium<sup>11</sup> de l'espace périmicrovillaire dans les cellules de l'intestin moyen, générant un gradient entre ces deux compartiments. De cette force ionique pourrait résulter une absorption active des composés tel que les acides aminés. Une fois dans l'espace périmicrovillaire, les acides aminés pourraient diffuser grâce à des transporteurs spécifiques intégrés dans les microvillosités (figure 8).

Quelques travaux sont en faveur de l'hypothèse émise par Terra sur l'absence de digestion protéique dans la lumière du tube digestif chez les pucerons : (Terra, 1994 ; Febvay, 1995 ; Mochizuki, 1998). Peu d'éléments sont disponibles sur le processus digestif des pucerons, c'est une question que l'on abordera plus en détail lors du développement du chapitre I.

---

<sup>11</sup> Le potassium est l'ion le plus représenté dans la sève.



**Figure 9 : Dégâts directs et indirects occasionnés par les pucerons sur plantes.**

**Dégâts directs.** a. et b. : déformations causées par la présence de pucerons chez le melon.

**Dégâts indirects causés par des champignons saprophytes et par les virus transmis par les pucerons selon les modes persistant ou non persistant.** c. et d. : fumagine sur melon et tomate. e. : symptômes causés par le PVY chez la tomate. f. : mosaïque déformante des feuilles de courgette due au CMV. g. et j. : déformations des fruits de melon et de courgette causées par le ZYMV. i. : chlorose, conséquence de l'infestation par le CABYV, le lutéovirus du melon. k. : nécroses et tâches noires sur fruits de tomate infectés par l'AMV. l. : viroses à CVMV et WMV2 chez le melon.

#### 2.2.4. L'importance des dégâts causés par les pucerons

Le mode de nutrition des pucerons et leur pouvoir de colonisation élevé font des pucerons un des problèmes majeurs des cultures en milieux tempérés. Il n'est pas d'organe de la plante qui échappent à la colonisation (figure 9). Leur présence sur la plante induit plusieurs types de dommages :

- **Une action spoliatrice par prélèvement de sève.** Lors des pullulations d'aphides, la ponction importante des produits de la photosynthèse entraîne un affaiblissement de l'hôte et peut provoquer un arrêt de la croissance de la plante (Miles, 1987).

- **Une action irritative et toxique.** La salive secrétée, lors des piqûres d'essais ou de l'alimentation, peut provoquer une réponse spectaculaire du végétal chez certaines espèces d'aphides (*Schizaphis graminum*, *Diuraphis noxia*, pucerons galligènes). La crispation des feuilles ou la formation de galles sont souvent observées lors de la colonisation (Forrest, 1987).

- **Une souillure.** Les produits de la digestion (très riches en sucre) forment le miellat rejeté sur les feuilles. Ce miellat non toxique en lui-même peut diminuer l'activité photosynthétique et respiratoire en obstruant les stomates. Lorsqu'il est en quantité importante, le miellat peut provoquer à la surface des feuilles un effet osmotique de nature à créer un appel d'eau à travers la couche épidermique semi-perméable. L'eau ainsi soutirée de la plante s'évapore très facilement. Le miellat tient le rôle d'agent dessiccateur très actif, rapidement mortel lorsque les conditions environnementales favorisent l'évaporation (Comeau, 1992).

Le miellat est également un milieu de culture très favorable au développement de champignons saprophytes, appelés fumagine. Les spores des champignons noircissent notamment les parties consommables de la plante et les rendent, par conséquent, impropres à la commercialisation.

- **Une dissémination de virus phytopathogènes.** Cette capacité à transmettre et à disséminer des virus représente une nuisibilité bien supérieure à celle des autres dégâts. Elle repose aussi bien sur le nombre de virus que les aphides sont susceptibles de transmettre que sur le nombre d'espèces d'aphides impliquées. Près de deux cents espèces de pucerons ont été reconnues vectrices. L'une d'entre elles, *Myzus persicae*, est capable d'inoculer plus de 120

**Tableau 2 : Principales propriétés des virus transmis par les aphides.**

Phases du cycle de transmission	Virus non circulants		Virus circulants
	Non persistants	Semi persistants	
<b>Acquisition</b>	Brève (piqûre d'essai)	Généralement longue	Longue (dans le phloème)
<b>Latence</b>	Non (le puceron est immédiatement infectieux)	Non	Oui (nécessité d'un circuit dans le corps de l'insecte)
<b>Rétention</b>	Brève (quelques heures maximum)	Assez longue (quelques jours)	Longue (plus de 6 heures à plusieurs jours, parfois pour toute la vie)
<b>Conservation après la mue</b>	Non	Non	Oui
<b>Spécificité de transmission</b>	Faible	Etroite	Etroite
<b>Multiplication</b>	Non	Non	Oui
<b>Autres appellations</b>	Virus du stylet, Virus externes		Virus persistants, Virus internes

Acquisition : le vecteur se charge en virus dans une plante malade.

Latence : laps de temps qui s'écoule entre l'acquisition et le moment précis où le vecteur sera apte à infecter une plante saine. Il deviendra alors infectieux. *La latence concerne le vecteur.*

Rétention : laps de temps durant lequel le vecteur demeure infectieux après un repas contaminant.

Inoculation : le vecteur infectieux introduit le virus dans la plante.

Transmission : ensemble des événements qui commence de l'acquisition du virus par le vecteur dans une plante malade et se termine à l'inoculation du pathogène dans une plante saine.

Incubation : laps de temps qui s'écoule entre l'inoculation et l'extériorisation des symptômes *sur la plante.*

virus (Blackman, 1985). Il faut très peu de pucerons pour qu'une culture soit rapidement contaminée et lorsque le virus se multiplie, les plantes jaunissent, s'incurvent, se cloquent. Les viroses sont des maladies incurables.

Il existe différents types de relations possibles entre le puceron et le virus, en rapport avec la transmission : les virus non persistants ou virus du stylet, les virus persistants ou circulants, les virus semi-persistants et les virus qui se multiplient dans l'insecte comme dans la plante sont appelés virus persistants propagatifs (tableau 2).

Les virus non persistants sont, de loin, les virus les plus nombreux transmis par les pucerons. Ils sont acquis et transmis au cours de piqûres brèves. Lorsque le puceron se nourrit, il effectue des piqûres d'exploration, enfonçant ses stylets au niveau du mésophylle pendant un temps très court (10 à 60 secondes) puis les retire et essaie à nouveau. Ces piqûres sont remarquablement efficaces pour l'acquisition et la transmission des particules virales qui se multiplient dans le parenchyme. Le puceron ne demeure pas longtemps infectieux après l'acquisition, le virus ne persiste pas dans le vecteur plus d'une heure.

Contrairement aux virus non persistants où la spécificité virus-vecteur est assez large, la spécificité pour les virus persistants est étroite. Ces virus sont acquis lors de la prise alimentaire dans le phloème. Le puceron n'est pas infectieux immédiatement, le virus suit l'itinéraire alimentaire passant par le canal ascendant des stylets, le système digestif, puis est transporté dans l'hémolymphe et envahit ensuite les glandes salivaires. Ce circuit dans le corps de l'insecte caractérise la période de latence. Au cours de cette période, le virus peut commencer à se multiplier (virus circulant propagatif). L'interaction spécifique pour les virus circulants consisterait en une interaction avec des protéines des cellules de glandes accessoires qui reconnaissent le virus, forment une vésicule de transport l'amenant vers le conduit salivaire (par pinocytose). Une fois dans le flot de salive, la particule est injectée dans les tissus végétaux (Cornuet, 1987).

### **2.3. La réponse moléculaire des plantes à la présence de pucerons.**

La majorité des études réalisées jusqu'à présent concerne l'étude des réponses induites par des agents pathogènes ou par des insectes défoliateurs qui causent d'importants dommages à la plante (Shenk, 2000 ; Stotz, 2000 ; Hermsmeier, 2001 ; Baldwin, 1999). La réponse induite par les insectes a souvent été comparée à celle engendrée par la blessure. Une étude comparative récente des gènes exprimés, chez *Arabidopsis thaliana*, entre la réponse à la

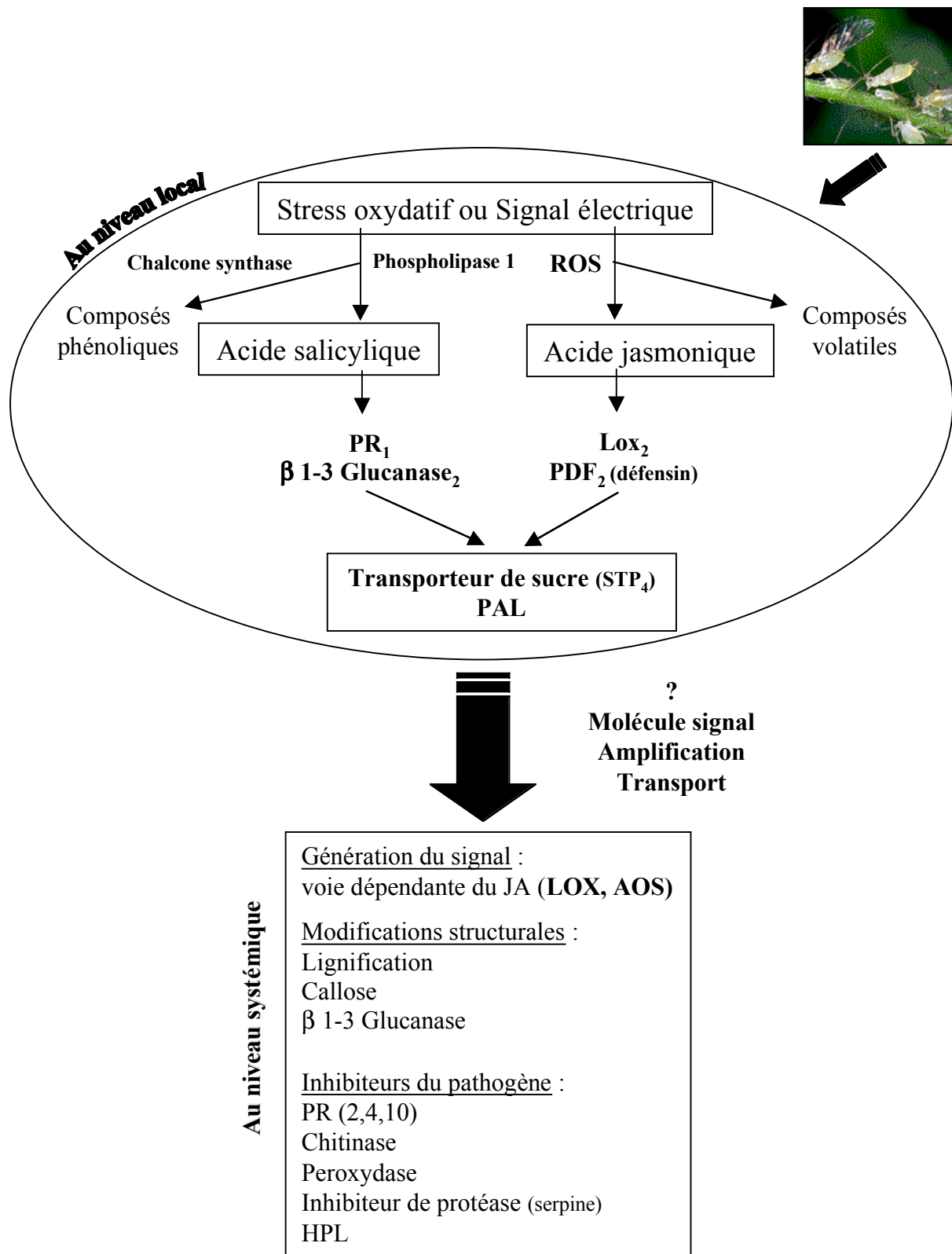


Figure 10 : Modèle de la réponse des plantes à la présence de pucerons (Walling, 2000).

blessure et la réponse à l'attaque par la larve du lépidoptère *Pieris rapae* a permis de mettre en évidence des différences (Reymond, 2000). Pour cette expérience de « microarray » 150 gènes ont été choisis. Parmi ceux ci, plusieurs familles de transcrits sont régulées après l'attaque par les insectes et la blessure. Un seul gène (HEL, gène de défense induit par des éliciteurs) est spécifique à la présence de l'insecte, confirmant l'hypothèse précédente.

Paradoxalement, au regard de l'importance des insectes en agronomie, la réponse induite par les pucerons a été beaucoup moins étudiée. Cependant quelques études permettent de mieux appréhender les mécanismes physiologiques et moléculaires de la réponse des plantes aux pucerons. En raison de leur mode d'alimentation, les aphides entraînent peu de dommages aux tissus de la plante ce qui les différencie des insectes défoliateurs. Une fois le site d'alimentation établi, il peut être utilisé pendant des heures voire des semaines. Il n'est donc pas surprenant que la réponse aux pucerons soit distincte de celle engendrée par les insectes mâcheurs. Les signaux générés par ces herbivores spécifiques sont complexes. Ils peuvent dériver du stress mécanique engendré par la pénétration des stylets dans les tissus. Les mouvements des stylets détruisent les connexions entre les cellules. Les salives émises (gélifiées et aqueuses) lors de l'alimentation peuvent aussi remplir la fonction d'éliciteur. La salive polymérise rapidement autour des stylets formant une gaine protectrice et limitant ainsi les contacts directs entre les stylets et l'apoplasme. Toutefois, la salive digestive contient de nombreuses molécules complexes ainsi que des enzymes. Une de ces molécules et/ou les produits de la réaction enzymatique peuvent induire une réponse de défense. Par exemple, les oligogalacturonides produits lors de la dégradation de la paroi végétale par les pectinases ont déjà été caractérisés comme éliciteurs des voies de défense des plantes (Walling, 2000). Quelques études nous permettent d'établir un schéma de la réponse des plantes aux pucerons (figure 10) :

- Chez la tomate, la présence de pucerons est responsable d'une augmentation des transcrits et de l'activité des plusieurs PR protéines, telles que les peroxydases et les chitinases, et d'une lipoxygénase (LOX) (Fidantsef, 1999 ; Stout, 1999). L'inhibiteur de protéase à sérine Pin2, déjà cité dans la réponse aux lépidoptères, ne semble pas induit par la présence de pucerons (Fidants, 1999).

- Chez le blé, les activités peroxydases solubles et liées à la paroi sont aussi accrues après infestation par les pucerons (Chaman, 2001). Ces activités diminuent très rapidement après que les pucerons se soient retirés suggérant un rôle spécifique de ces enzymes dans la

réaction de défense et non un rôle dans la réparation des dommages. Les composés phénoliques qui en résultent, peuvent agir comme éléments anti-appétants ou comme toxines.

- Chez l'orge, l'induction de  $\beta$ -1-3 glucanases est observée lors d'infestation par le puceron des céréales *R. padi*. Cette enzyme possède une activité sur les polymères de callose (Forslund, 2000). Elle pourrait participer au "turn-over" de la callose au cours de la mise en place de barrières callosiques autour des lésions et dans le phloème, limitant ainsi l'accès à la sève (Vanderwesthuizen, 1998).

D'autre part, une augmentation de l'activité d'inhibiteurs de protéases à sérine, et plus particulièrement d'un inhibiteur de chymotrypsine, est corrélée à une résistance non déterminée à ce puceron (Casaretto, 1998). Chez le concombre, la présence d'une serpine circulant de la phloème est aussi corrélée avec une diminution de la croissance et de la fécondité de *M. persicae*, bien que son action ne soit pas directe (Yoo, 2000).

- L'émission de composés volatils est également un élément de la réponse induite par les pucerons. L'hydroxyperoxydelyase (HPL) catalyse le clivage d'hydroperoxyde d'acides gras en aldéhydes et oxoacides. Cette enzyme utilise les produits de l'oxygénation d'acide gras (linoléate ou linolate) catalysée par les LOX. Ces dérivés d'acides gras sont impliqués dans la réponse des plantes aux insectes et aux micro-organismes (figure 10). Des pommes de terre transgéniques, où l'activité HPL est réduite par antisens, permettent un développement beaucoup plus favorable du puceron *M. persicae* (+ 30% de fécondité). L'activité HPL est constitutive chez le témoin, suggérant que l'émission d'aldéhydes volatiles est générée par une décompartmentalisation lors de la blessure (Vancanneyt, 2001).

- Les gènes activés lors de l'attaque d'herbivores sont très fortement corrélés avec leur mode de nutrition et avec le degré de dommage des tissus au niveau du site d'alimentation. Les pucerons, insectes piqueurs-suceurs, sont responsables de petites blessures. Il semble que la réponse des plantes aux pucerons s'apparente plus à la réponse induite par les agents pathogènes (Walling, 2000). Ces résultats ont été confirmés grâce à une étude moléculaire entreprise chez *Arabidopsis thaliana* (Moran, 2001). De plus, cette étude permet de mettre en évidence la différence de réponse aux niveaux local et systémique (figure 10).

Ces études ont permis de mettre en évidence que les pucerons stimulent les voies dépendantes du salicylate, de l'acide jasmonique et de l'éthylène. Les pucerons sont donc capables d'activer des réponses associées à la blessure et à la réponse aux agents pathogènes.



Toutefois, le schéma présenté est très incomplet. La molécule signal, qui permet une réponse systémique, n'est pas identifiée. Des études visant à rechercher les éliciteurs spécifiques chez les pucerons permettraient d'avoir une connaissance plus approfondie des mécanismes de résistance liés aux pucerons.

## **2.4. Moyens de lutte contre les pucerons**

Les méthodes de lutte actuellement utilisées reposent sur des approches conventionnelles basées principalement sur la lutte chimique, et à plus petite échelle, la lutte biologique et la lutte intégrée.

### **2.4.1. La lutte chimique**

Avant la seconde guerre mondiale, le contrôle des pucerons par lutte chimique se limitait à l'application de nicotine et de produits à base d'arsenic. Au début du 20<sup>ème</sup> siècle, les méthodes empiriques ont été remplacées par des procédés établis sur des bases scientifiques en utilisant des dérivés minéraux et des extraits de plantes. Avec les progrès réalisés en chimie, l'utilisation d'insecticides de synthèse s'est répandue à partir de la découverte des effets insecticides du DDT et de son effet neurotoxique. Les premières grandes familles développées après la seconde guerre mondiale (organochlorés et organophosphorés) possèdent des propriétés jugées à l'époque exceptionnelles : grande persistance et large spectre d'activité. L'emploi d'insecticides pour lutter contre les Homoptères représente aujourd'hui environ 27% de la quantité utilisée mondialement. Les insecticides les plus couramment employés contre les pucerons sont le Pirimicarb (carbamate à action spécifiquement aphicide), les organophosphates Demethon-S-methyl, le Dimethoate et l'Heptenophos en association avec les pyréthriinoïdes Deltamethrine (Thomas, 1996 cité dans Couty, 2000).

Après une vingtaine d'années d'utilisation, ces qualités se sont souvent muées en défauts rédhibitoires : la grande persistance s'est traduite par une pollution assez généralisée de l'environnement et par l'apparition rapide de résistances chez de nombreux ravageurs. Le large spectre d'activité a provoqué des effets écotoxicologiques liés au manque de spécificité, tels les effets néfastes sur la faune sauvage, l'élimination d'ennemis naturels des ravageurs... Les pucerons n'échappent pas à cette règle. Les premiers échecs de traitement sont apparus en 1955 aux Etats Unis et il existe aujourd'hui plus de 20 espèces de pucerons résistants aux

insecticides dans le monde (Devonshire, 1998). Il est extrêmement difficile de faire un état de la résistance aux insecticides chez les pucerons car la diversité des espèces, des cultures et des situations empêchent toute approche globale du problème. *Myzus persicae* et *Aphis gossypii* (2 pucerons très polyphages et très bons vecteurs de virus) font partie des rares espèces pour lesquelles il existe des données suffisamment précises.

Le principal mécanisme de résistance consiste en la capacité de production, par les souches résistantes, de grandes quantités de carboxylestérase (estérase E4). Cette enzyme a pour action de détoxiquer certains insecticides comme les organophosphorés mais aussi de piéger de très nombreux autres insecticides et ainsi de les empêcher d'atteindre leur cible. Cet accroissement d'activité est lié à un mécanisme d'amplification génétique. Ainsi chez les souches les plus résistantes, ce gène peut être amplifié plus de 64 fois (Field, 1997).

La pollution de la biosphère par les résidus d'organochlorés, aggravée par un phénomène de concentration au sein des chaînes alimentaires, a amené la FAO à interdire les insecticides les plus rémanents (Riba, 1989). Cependant, la lutte chimique est encore la méthode la plus utilisée par les agriculteurs, et la plus rentable pour les industries phytosanitaires dans le contrôle d'organismes nuisibles. L'action rapide des insecticides, alliée à la facilité d'utilisation et à la maîtrise des technologies de traitement d'application, font de la lutte chimique un outil apprécié par les producteurs et facile à recommander par les conseillers agricoles.

#### **2.4.2. La lutte biologique**

Le principe de la lutte biologique est simple : combattre les ravageurs des cultures en introduisant dans le milieu où ils vivent un de leurs ennemis. Ce concept fait également référence à toute modification de l'environnement dans le respect des règles écologiques de stabilité et d'équilibre qui conduisent au maintien des organismes nuisibles en dessous d'un seuil ayant un impact économique. Chaque espèce animale possède un grand nombre d'antagonistes associés de façon plus ou moins spécifique qui sont soit des animaux soit des microorganismes pathogènes.

A l'échelle mondiale, dans tous les cas où la lutte biologique est utilisée avec succès pour combattre les pucerons, les antagonistes sont des parasitoïdes membres de la famille de Aphidiidae (Hyménoptères) (Sauvion, 1995). L'utilisation de prédateurs tels que les Coccinelles, les Syrphes ou les Chrysopes, est fréquente pour le contrôle des pucerons en serre. Les coccinelles aphidiphages peuvent avoir un impact positif sur le contrôle des insectes nuisibles. La mise en place de techniques d'élevage efficaces et la connaissance de la

compatibilité avec d'autres agents de lutte permettent de croire qu'elles pourraient être utilisées rapidement comme une alternative aux insecticides chimiques.

Bien que suscitant un intérêt de plus en plus grand, la lutte biologique demeure négligeable comme moyen de lutte contre les insectes nuisibles aux cultures en plein champs.

### **2.4.3. La création variétale**

#### *2.4.3.1 Par sélection classique*

La résistance des plantes aux insectes peut être définie comme l'ensemble des caractéristiques d'une plante lui conférant la capacité de s'opposer à la multiplication de population d'insectes ravageurs ou aux dégâts provoqués par celles-ci. Cette caractéristique est toujours mesurée par rapport à une situation de référence (témoin dit sensible), et on restreint souvent l'usage du terme de « résistance » aux caractères héréditaires qui différencient certaines populations de plantes de la même espèce. La sélection et l'utilisation de ces caractères constituent la base de la stratégie de lutte variétale contre les pathogènes et les ravageurs des plantes cultivées.

La résistance des plantes a été classée en 3 catégories :

- la non-préférence (non acceptation, antixénose) : l'insecte n'accepte pas la plante pour s'alimenter et pour se reproduire, même en l'absence de choix. Ce n'est pas un phénomène de "tout ou rien" et la plupart des cas se traduit par des formes graduées d'acceptation de l'hôte. La forme extrême de la non-préférence est d'un grand intérêt quand la colonisation même brève entraîne des dégâts sévères (toxicoses, infection virale...).

- l'antibiose : englobe les effets négatifs que la plante est capable de produire sur la biologie et la physiologie de l'insecte l'utilisant comme hôte. Les effets typiques de l'antibiose sont des retards de croissance ou une altération du taux de reproduction de l'espèce nuisible cible.

- la tolérance : capacité à tolérer ou à récupérer des suites d'infestations même sévères. L'expression de cette tolérance est déterminée soit par la capacité génétique intrinsèque à ne pas subir les effets délétères d'une infestation (tolérance physiologique directe), soit par la capacité à compenser des dégâts réalisés ou des destructions tissulaires importantes par une croissance accrue après l'attaque. Contrairement aux deux catégories précédentes, la tolérance est plus une caractéristique physiologique de la plante qu'une composante à part entière de l'interaction plante-insecte (sauf cas particuliers comme des tolérances à des toxicoses spécifiques) (Chen, 1996 a).

**Tableau 3 : Exemple de gènes majeurs conférant une résistance aux pucerons**

<b>Plante</b>	<b>Pucerons</b>	<b>Gènes</b>	<b>Références</b>
Blé	<i>Schizaphis graminum</i>	<i>GB</i> (1 à 6)	(Porter, 2000)
	<i>Diuraphis noxia</i>	<i>DN</i> (1 à 6)	(Dutoit, 1989)
Orge	<i>Schizaphis graminum</i>	<i>RSG 1a</i>	(Edwards, 1985)
Cowpea	<i>Aphis craccivora</i>	<i>RSC</i> (1 et 2)	(Githiri, 1996)
Tomate	<i>Macrosiphum euphorbiae</i>	<i>MI</i>	(Rossi, 1998)
Melon	<i>Aphis gossypii</i>	<i>VATt</i>	(Pitrat, 1982)
Laitue	<i>Nasonovia ribisnigri</i>	<i>NR</i>	(Eenink, 1982)
Pêcher	<i>Mysus persicae</i>	<i>Résistance monogénique</i>	(Monet, 1994) ; (Sauge, 1998)
Pommier	<i>Dysaphis devectora</i>	<i>SD</i> (1 à 3)	(Alton, 1977)
	<i>D. plantaginea</i>	<i>SMH</i>	(Alton, 1970)

Les recherches sur les interactions plantes-pucerons ont permis de mettre en évidence de telles résistances dans diverses familles de plantes (Légumineuses, Graminées, Cucurbitacées, Composées...) (tableau 3). Deux gènes de résistance de type monogénique sont particulièrement bien caractérisés dans la littérature :

- **le gène *VAT*** (*virus aphid transmission*) a été identifié à l'INRA d'Avignon à la fin des années 70 (Lecoq, 1979 ; Pitrat, 1980). Ce gène (monogénique dominant) présente l'avantage majeur de conférer une résistance à la transmission par *A. gossypii* des principaux virus non-persistants des Cucurbitacées. La présence de cet allèle entraîne également une résistance à la colonisation des plantes suivant un mécanisme de non-préférence et d'antibiose. Différentes analyses montrent que cette résistance est constitutive et que les effets du gène *VAT* s'expriment dans les tissus préphloémiens et dans le phloème (Chen, 1996 b). La cartographie récente d'un allèle probable de ce gène (*Agr*) a permis de montrer qu'il se trouve dans un cluster de gènes de résistance aux pathogènes (notamment d'un gène codant pour une LOX) (Klingler, 2001).

- **le gène *MI***, chez la tomate, confère aussi une double résistance : une résistance au puceron *Macrosiphum euphorbiae* et une résistance aux nématodes (Rossi, 1998). Ce gène dominant a été cloné et séquencé. Il code pour une protéine possédant des motifs caractéristiques de gènes de résistance impliqués dans la reconnaissance spécifique d'agents pathogènes avirulents : nucléotide binding site (NBS) et un motif riche en leucine (LRR). De la même façon que pour la résistance conférée par *VAT*, il semble que les pucerons ne possèdent plus la capacité de s'alimenter durablement sur le phloème et n'arrivent plus à se nourrir (Chen, 1997 ; Kaloshian, 2000). Cette résistance est aussi spécifique à une espèce de puceron puisque *M. persicae* est capable de coloniser les tomates exprimant *MI* (Goggin, 2001).

Les sélectionneurs ont évidemment exploité ces mécanismes de défense et introgressé ces gènes de résistance dans de nombreuses variétés commerciales. Cependant, il n'existe pas de solutions génétiques pour toutes les espèces végétales, et même à l'intérieur d'une espèce, la résistance n'est pas efficace contre tous les pucerons. Dans les deux exemples développés ci-dessus, la présence du gène de résistance chez la plante hôte permet d'obtenir une résistance à une seule espèce de puceron suggérant une analogie avec la notion de race/cultivar décrite dans le modèle de Flor<sup>12</sup> (1971). Ces modes de défense, soutenus par des résistances

---

<sup>12</sup> Modèle gène à gène (Flor, 1971) : la reconnaissance spécifique de l'agent pathogène par la plante est souvent contrôlée par un gène de résistance chez la plante auquel correspond un gène d'avirulence chez l'agent pathogène.

monogéniques, sont souvent sujets à une adaptation rapide des insectes et à l'apparition de colonies sur les génotypes résistants<sup>13</sup>.

#### 2.4.3.2 Par transgénèse : une approche nouvelle pour lutter contre les pucerons

L'utilisation de la transformation génétique a également été envisagée dans la lutte contre les pucerons malgré la spécialisation des pucerons pour leur plante hôte (pour revue Jouanin, 2002).

Les premiers travaux publiés au sujet de plantes transgéniques résistantes aux insectes font part du transfert et de l'expression de gènes (*CRY*) dérivant de la bactérie *Bacillus thuringiensis* (Bt). Malgré la grande diversité des gènes codant pour une protéine CRY (130), peu de delta-endotoxines ont été testées contre les pucerons, et se sont malheureusement révélées inefficaces (Walters, 1995).

La majorité des plantes transgéniques testées contre les pucerons repose sur l'expression d'une lectine appartenant à la famille des lectines à mannose de monocotylédones, la GNA, extraite du perce neige. L'expression de la GNA a permis d'observer un retard de croissance et une baisse de la fécondité chez les pucerons (Hilder, 1994, Stoger, 1999), ce qui correspond au syndrome observé *in vitro* (Sauvion, 1996). La GNA a également été utilisée en combinaison avec une chitinase de haricot. Dans ce cas, la synergie entre les 2 protéines n'a pas été clairement démontrée (Gatehouse, 1996).

Le génie génétique, combiné à la transgénèse, a permis de franchir les barrières des espèces en introduisant des gènes d'origine bactérienne, virale, fongique et animale dans les plantes. Un gène bactérien codant pour une enzyme (IPT pour isopentenyl transferase), impliqué dans la biosynthèse des cytokinines, a été exprimé dans le tabac sous le contrôle du promoteur d'un inhibiteur de protéase (Smigocki, 1993). L'expression de ce transgène altère le développement de *M. persicae*. Une caractérisation partielle de la physiologie des plantes transgéniques suggère que la production, la sécrétion ou l'accumulation de métabolites secondaires dans les feuilles puissent être responsables de l'activité insecticide. Toutefois, ce travail n'a pas été exploité plus en avant.

Une dernière stratégie a été utilisée pour lutter contre les pucerons, l'extinction d'un gène. Grâce à la co-suppression d'un gène codant pour un cytochrome P450 spécifiquement exprimé dans les trichomes, les métabolites secondaires des exsudats de trichomes de tabac

---

<sup>13</sup> Des contournements du gène *Vat* ont déjà été observés en plein champs par un biotype d'*A. gossypii* guadeloupéen (Sauvion, comm. pers.).

ont été modifiés (Wang, 2001). Une plus grande concentration d'un diterpène (CBT-ol) permet d'observer une réaction de non-préférence de *Myzus nicotiana*. Cette réaction très forte de résistance suggère que la modification des composés secondaires émis par les trichomes soit aussi une possibilité à exploiter.

Les travaux utilisant la transgénèse dans la lutte contre les pucerons sont relativement récents et le facteur limitant reste le choix du gène candidat. Un programme de criblage systématique de protéines possédant une toxicité potentielle a été entrepris afin de mettre à jour de nouveaux gènes codant pour des protéines aphidicides (Rahbé, 1993 ; Rahbé, 1995). D'autre part, les programmes d'études des mécanismes de résistance des plantes aux pucerons devraient permettre de sélectionner de nouveaux candidats (gènes spécifiquement induits). Le clonage des gènes *VAT* et *MI* présente de nouvelles perspectives pour la protection des Cucurbitacées et des Solanacées.

#### 2.4.4. La lutte intégrée

La protection intégrée est un concept de lutte associant plusieurs armes. Prenant en compte les caractéristiques de l'écosystème considéré, la protection intégrée fait appel en priorité aux techniques alternatives à la lutte chimique. Elle utilise, autant que possible, les moyens de lutte biologique, et y associe l'emploi de variétés résistantes et la gestion des techniques culturales. Toutefois, la lutte intégrée n'exclut pas le recours à des pesticides chimiques. Elle en prévoit l'usage tout en garantissant le respect des insectes auxiliaires et pollinisateurs. Il n'est pas question d'éliminer les ravageurs jusqu'au dernier mais d'abaisser leurs effectifs de telle façon que les dégâts soient supportables et sans incidence économique majeure.

Une approche plus récente consiste à aménager l'environnement végétal de la culture de manière à enrichir la faune d'auxiliaires actifs sur les principaux ravageurs de la culture. L'intégration de plantes transgéniques dans les schémas de lutte intégrée apparaît également un point intéressant pour le futur.

Dans le cadre des " Services de recherches intégrées sur les productions Végétales et la protection des plantes ", ces stratégies sont testées grandeur nature par l'INRA. Des programmes de recherche sont actuellement engagés sur la protection intégrée des vergers de pommiers et de pêchers (contrôle des populations de *M. persicae*).

### 3. Objectif de la thèse

Actuellement, la lutte contre les pucerons présente un nombre de possibilités encore restreint en dehors de la lutte chimique. Les solutions utilisant la sélection classique sont très limitées en raison d'un faible nombre de gènes de résistances disponibles. Face à cette situation, un développement des approches utilisant la transgénèse peut apparaître comme une alternative de choix. C'est dans cette optique que la collaboration entre le laboratoire de Biologie Cellulaire et le laboratoire de Biologie Fonctionnelle, Insectes et Interactions a permis le développement de deux programmes utilisant la transgénèse pour l'obtention de plantes résistantes aux pucerons.

La première étude a été réalisée dans le cadre d'une convention CIFRE entre l'INRA et la Société Vilmorin, Clause & Cie et vise à évaluer l'expression tissu-spécifique de protéines candidates et l'efficacité aphicide de lectines de monocotylédones.

La seconde étude s'inscrit dans le cadre d'un programme européen qui a pour vocation de diminuer les applications d'insecticides en produisant des plantes d'intérêt agronomique transgéniques résistantes aux insectes grâce à l'expression de nouveaux gènes codant pour des inhibiteurs de protéases. Ce programme constitue le travail décrit dans cette thèse. Dans le cadre du programme européen, l'INRA est en collaboration avec un sélectionneur de semences de melon : la société Teizier S.A. du groupe Limagrain. C'est pourquoi notre étude repose sur le principal ravageur du melon le puceron *Aphis gossypii*.

Les pucerons sont parmi les rares insectes ne possédant pas *a priori* d'arsenal endoprotéasique important, ils consomment un aliment très riche en acides aminés libres. Pourtant, plusieurs inhibiteurs de protéases ont montré des toxicités paradoxales contre les pucerons (Rahbé, 1995 ; Rahbé, 2002). Cette étude a pour objectif d'accroître la disponibilité de gènes d'IP originaux et d'en comprendre le mode d'action afin d'en maîtriser l'introduction raisonnée dans des programmes de lutte variétale.

Nous avons articulé cette étude autour de 3 axes principaux constituant la base du travail de thèse et la structure du présent manuscrit :

- Une étude de l'arsenal endoprotéasique du puceron du melon
- L'utilisation et la diversification des inhibiteurs de protéases à cystéine pour lutter contre les pucerons
- Une étude du mode d'action et la modification d'un inhibiteur de protéase du pois.



## **CHAPITRE I**

### **- ETUDE DES PROTEASES DU TUBE DIGESTIF D'*A. GOSYPII* -**

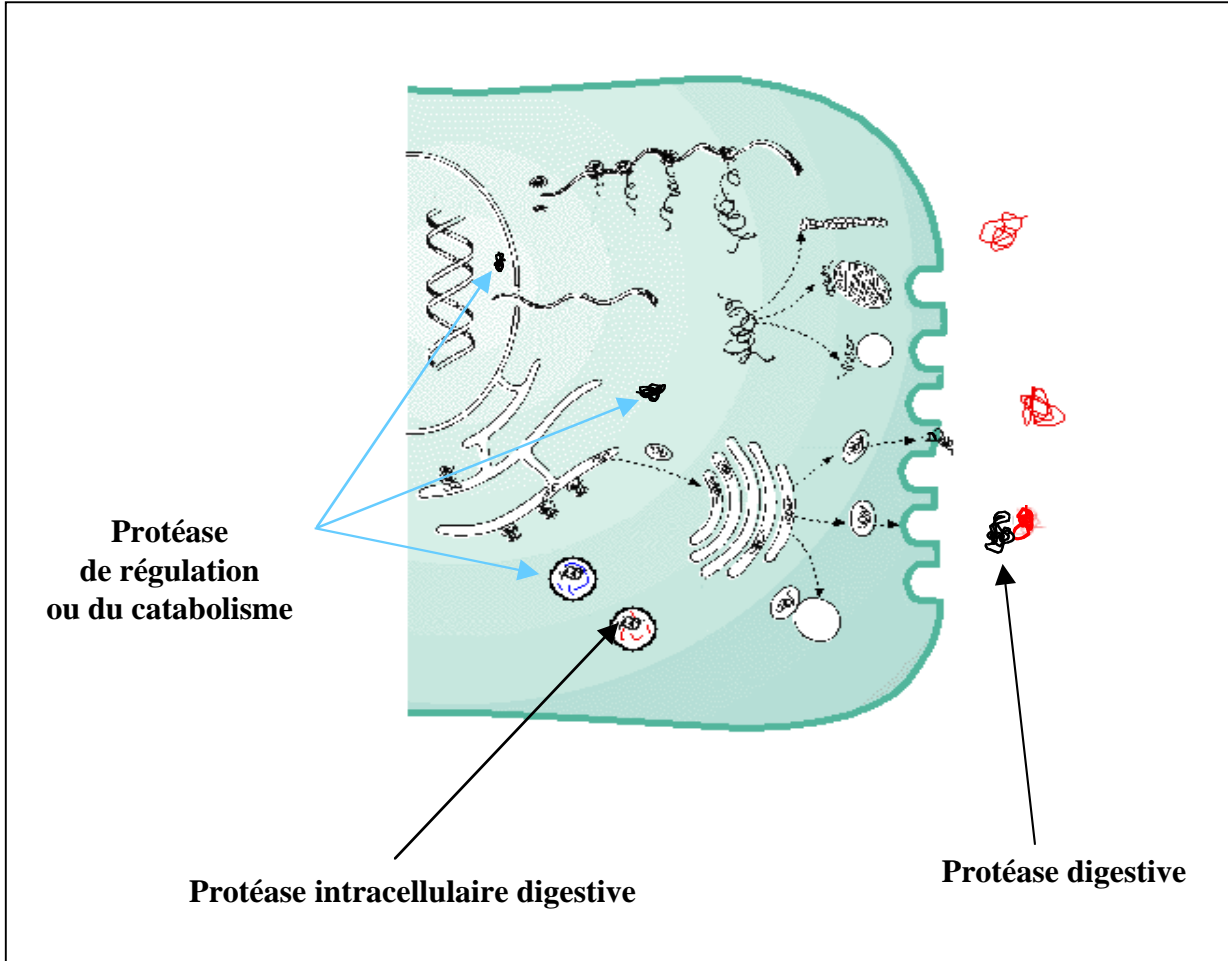
## 1. Contexte de l'étude

Dans le cadre d'une stratégie utilisant les inhibiteurs de protéases en tant qu'entomotoxine, il nous semble indispensable de rechercher à mieux caractériser les protéases de l'insecte afin de cibler correctement la famille d'inhibiteurs potentiellement actifs. Les pucerons se nourrissent exclusivement de sève élaborée, dont le rapport protéines sur acides aminés libres est très faible. Il semble qu'en raison de la spécificité du type de nutriments ingérés, les Homoptères aient subi une spécialisation de leur système digestif leur permettant d'absorber les acides aminés à partir d'une alimentation très diluée. C'est pourquoi, dans ce contexte, les pucerons sont considérés comme ne possédant pas de protéases digestives (Terra, 1996 ; Mochizuki, 1998).

Toutefois, dans le phloème de certaines espèces végétales, comme les cucurbitacées, circulent des protéines dont la concentration globale peut être très élevée, allant jusqu'à 60 mg.ml<sup>-1</sup> (Cronshaw, 1989). Séparés sur un gel d'électrophorèse 2D, plus de 200 polypeptides solubles sont dénombrés sur toute la gamme de points isoélectriques et de poids moléculaires (Chen, 1996 a ; Haebel, 2001). Les protéines du phloème des cucurbitacées sont essentiellement connues pour certaines classes qui semblent spécifiques à ce tissu, comme les protéine-P (P-proteins, ou protéines du phloème) (Sabnis, 1976), mais aussi des lectines (Dannenhoffer, 1997) ou des inhibiteurs de protéases qui sont aussi présents dans la sève élaborée (IP à sérine (Dannenhoffer, 2001 ; Yoo, 2000), cystatines (Haebel, 2001), inhibiteur de protéase acide (Christeller, 1998)). Ces protéines peuvent être ingérées et métabolisées, elles sont des sources potentielles de composés azotés. De plus le marquage associé à OC-I observé dans le tube digestif du puceron *Myzus persicae*, nourrie sur plantes de tabacs surexprimant cet IP, suggère une interaction avec des protéines du tube digestif et plus particulièrement avec des protéases du tube digestif (article 2, voir plus loin).

Nous avons mis en œuvre plusieurs méthodologies afin d'identifier le spectre complet des protéases digestives utilisées par le puceron. En effet, le tube digestif représente l'élément principal d'interaction entre l'insecte et son environnement.

Notre étude révèle que l'activité endoprotéasique présente *dans* le tube digestif du puceron *A. gossypii* est constituée à plus de 95% d'une activité de type protéases à cystéine. Les protéases à cystéine du tube digestif du puceron ont donc été caractérisées au niveau moléculaire et biochimique, démontrant que cette activité s'apparente au groupe des



**Figure 1.1 : Les différentes classes de protéases d'un entérocyte.**

Les cibles endogènes des protéases sont représentées en bleu et les cibles exogènes en rouge.

cathepsines. Les cathepsines sont des protéases lysosomiales très étudiées chez les vertébrés (pour revue : Turk, 2000).

Dans les entérocytes, il existe des protéases que l'on peut rattacher fonctionnellement, selon l'origine de leurs cibles, à trois grands processus physiologiques (figure 1.1) :

– (1) *cibles endogènes “ activables ”* : des protéases intracellulaires spécialisées dans la maturation des protéines (activation, hydrolyse de zymogènes...)

- (2) *cibles endogènes* : des protéases intracellulaires liées au catabolisme des protéines. La dégradation des protéines a lieu en permanence dans une cellule en équilibre métabolique. Il paraît exister deux mécanismes essentiels : la dégradation dans des organites spécialisés que sont les lysosomes (autophagie) et le catabolisme dans le cytoplasme (protéases à sérine, protéasome 20S). Les protéases lysosomiales sont variées : cathepsines B, D, H, L, amino et carboxypeptidases. L'acidité intérieure des lysosomes est un facteur de dénaturation des protéines qui y pénètrent. Les protéases lysosomiales sont efficaces à pH acide, par contre elles sont inactives au pH neutre du cytoplasme, qui est ainsi protégé de relargages occasionnels ou intempestifs.

(3) *cibles exogènes* : les protéases digestives secrétées, que l'on retrouve donc dans la lumière du tube digestif, ou celles des lysosomes à fonction digestive assez stricte (hétérophagie). Ces lysosomes sont le site de la dégradation de substances étrangères, captées par phagocytose.

A partir d'une banque d'ADNc du tube digestif de puceron, une séquence de protéases à cystéine a été isolée. L'étude phylogénétique réalisée place cette séquence dans le sous-groupe des cathepsines L. La répartition tissulaire, étudiée par immunomarquage, indique que cette protéase est exprimée spécifiquement dans les organes liés à l'assimilation des acides aminés (tube digestif et bactériocytes). Par contre, il semble que la protéine ne soit pas secrétée dans la lumière du tube digestif, et la protéase serait donc strictement intracellulaire. Le puceron du melon posséderait donc un système enzymatique lysosomal à expression digestive (Figure 1, classe 3), potentiellement capable de dégrader les protéines du phloème ingérées, bien que cela n'ait pas été démontré dans notre cas d'espèce.

**Article 1** : Deraison C. , Darboux I., Duportets L., Gorojankina T., Jouanin L., Rahbé Y.. soumis à *J. Biol. Chem.*

**Cloning and Characterization of a gut specific Cathepsin L from the aphid *Aphis gossypii*.**

## Article 1

### **Cloning and characterization of a gut specific Cathepsin L from the aphid *Aphis gossypii***

Deraison C<sup>1,2</sup>, Darboux I<sup>1</sup>, Duportets L.<sup>1</sup>, Gorojankina T.<sup>1</sup>, Jouanin L.<sup>1</sup>, Rahbé Y<sup>2</sup>..

1, Laboratoire de Biologie Cellulaire, INRA de Versailles, route de Saint Cyr, 78026 Versailles cedex

2, Laboratoire Biologie Fonctionnelle, Insectes et Interactions, Bat. Louis-Pasteur, 20 av. A. Einstein, 69621 Villeurbanne cedex

key words : Hemiptera, Aphididae, cysteine protease, digestion

## **Introduction**

Together with moths (Lepidoptera) and beetles (Coleoptera), aphids (Hemiptera, Aphididae) belong to one of the three major orders of insect agricultural pests. Damages resulting from direct aphid feeding and from the cost of controlling this pest family account for one of the largest sources of economic losses in agriculture. Currently, the main widely available method to control aphid population is the use of chemical insecticides. This approach leads to serious drawbacks, such as environmental and food chain contamination. Moreover, many aphid species are becoming more and more resistant to insecticides. The use of plant resistance genes is a common alternative to chemicals, but very few aphid resistance genes have been characterized to date (1-3), and only one cloned (4). Furthermore, most of these genes are specific to single aphid species, or even populations, suggesting an analogy to the race/host plant model initially described for plant pathogens (5).

Plant genetic engineering offers new possibilities to control aphid populations (6-9). Among the genes enhancing plant resistance to aphids, protease inhibitors (PI) have been reported to promote toxic effects when ingested by different aphid species (6,10,11). However, due to their peculiar digestive physiology, the mode of PI toxicity towards aphids is unknown, and the success of this control strategy relies on a better identification of PI targets in these insects.

Hemiptera is an insect order quite different and phylogenetically distant from most other pest or model insects, including Diptera (mosquito, *Drosophila*). In addition to lacking full metamorphosis, aphids share with scales and whiteflies the peculiarity of being strict phloem feeders. They penetrate plant tissues through a strict intercellular route between epidermal and mesophyll cells, and feed exclusively on photoassimilates translocated in phloem sieve elements. Therefore, their diet consists mainly of phloem sap, which is generally poor in proteins. Values reported for total proteins in sieve tube sap of various plants range from 0.2

to 2 mg.ml<sup>-1</sup> (12). This is the reason why aphids are not considered to possess proteolytic activity in their digestive track (13,14). However, some plants such as the cucurbitaceae, have a protein-rich phloem sap, in which protein content reaches up to 60 mg.ml<sup>-1</sup> (15). An important protein concentration could provide an adequate source of amino acids and nitrogen for phloem feeding insects if these proteins are ingested and metabolized.

Informations on protein digestion in aphids are extremely scarce in the literature. In order to examine the digestive physiology, a first step to understand the toxicity of some PIs to these insects, we analyzed the putative proteolytic activity in the gut of an aphid naturally feeding on a protein-rich diet. *Aphis gossypii* Glover is known as the cotton-melon aphid and is a very efficient vector of various plant viruses. It is one of the most widespread species of aphids, therefore being a major pest for many crop plants. In the present study, we demonstrate that the main digestive protease of this aphid is of the cysteine type, and a cDNA related to this family has been cloned. The encoded sequence, termed *AgCatL*, shows significant amino acid sequence similarity to cathepsin L-like cysteine proteinases (SCP) from the weevil *Sitophilus zeamais* (16), and to other cathepsin L cysteine proteinases including mammalian ones. Molecular and immunohistochemical studies allowed us to analyze the tissue distribution and expression of *AgCatL*, and suggests that it could play a role in digestion, although not being secreted in the digestive fluids as this occurs in some other insect species already shown to rely on cysteine protease digestion.

## **Experimental procedures**

### **Materials**

pH 4-10 universal indicator was obtained from Merck (Darmstadt, Germany). Restriction endonucleases, reverse transcriptase and TRIZol reagent were purchased from Gibco-BRL (Life science, N.Y., USA). SMART cDNA library construction kit and Prime a Gene™ labeling kit were obtained from Clontech (Palo Alto, USA) and from Promega (Madison, USA) respectively. Other chemicals were from Sigma (St Quentin Fallavier, France).

### **Insects**

The “melon strain” of *Aphis gossypii* (clone Ag-LM02) originated from INRA Montfavet was reared on *Cucumis melo* cv. Védraçais in the laboratory since 1991. Aphids were reared in ventilated Plexiglas cages (21°C, 70% r.h., L16:D8). A lab strain of the maize weevil *Sitophilus zeamais* reared on wheat was used as positive control for protease immunodetection.

### **Aphid dissection**

Guts from adult aphids (mean adult weight 100-150 µg) were dissected under a binocular microscope. Disrupted tracts were not used and entire midguts (“stomach”, median and posterior midguts, hindguts) were kept on ice until an adequate amount was collected, then used or stored at –80°C.

### **Determination of digestive tract pH**

The pH of the different gut parts was estimated using a universal indicator solution. Guts were either homogenized in four parts, or dipped entirely in the indicator and resulting tissue color was compared to pH values obtained using standards made from buffers, or to a color chart.

### **Determination of *A.gossypii* protease activities**

Dissected guts from adult were frozen and kept at –80°C until required. Digestive tracks were homogenized in 0.15M NaCl at 0°C. General protease activity was determined in the range of pH 4 and 9 using <sup>125</sup>I-globulin as a substrate protein between (17). Labeled substrate was obtained by coupling <sup>125</sup>Iodide to commercial purified γ-globulin through the iodobead



protocol (Pierce). Two gut equivalents were mixed with 5 µg ( $5 \cdot 10^5$  dpm) of labeled globulin, with 10 µl of water or inhibitors and buffer in a final volume of 300 µl. Inhibitors (E64, benzamidine, EDTA, pepstatine) or activator (DTT) were pre-incubated with gut extract at room temperature for 15 min prior to addition of substrate. The reaction mixture was then incubated at 37°C for 24 hours. 500 µl of chilled 20% TCA and 250 µl 2% caseine were added to stop the reaction, then centrifuged at 10 000g for 5 min to remove precipitated proteins. The radioactivity recovered in the supernatant was measured, and reported for each pH value.

### **Assay of protease activity on polyacrylamide gel electrophoresis**

Tissues were homogenized in cold extraction buffer (0.15 M NaCl and 1% SDS), incubated at room temperature for 2h and centrifuged for 5 min at 17000g and 4°C. SDS was eliminated by dialysis against 0.15M NaCl at 4°C overnight. Preliminary experiments on *A. gossypii* and another aphid species have showed that the use of SDS was necessary to solubilize endoproteolytic activity, and that other tested detergents did not result in substantial solubilization (18).

Samples containing two gut equivalents were diluted two-fold in Leammli buffer (62.5 mM Tris HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 1‰ (w/v) bromophenol blue). After migration on a 0.1% SDS-10% polyacrylamide gel (BioRad Mini Protean II™ system) conducted at 4°C, the proteins were electrotransferred on an agarose/gelatin replica gel where *in situ* proteolytic visualization is carried on, as described by Michaud (19).

### **RNA isolation, RT-PCR and Northern-blot hybridization**

Total RNA was isolated from both 200 dissected guts and the resulting carcasses (whole aphid without guts and embryos –containing gut tissues-). RNA was extracted using TRIzol™ reagent as described by the manufacturer from young adult aphids.

Five  $\mu\text{g}$  of total RNA were submitted to reverse transcription (RT) before PCR. The RT reaction was done at  $37^\circ\text{C}$  in the presence of oligo(dT) and M-MLV reverse transcriptase according to standard instructions. The PCR reaction was performed using  $1\ \mu\text{l}$  of the RT reaction, and  $10\ \text{pmol}$  of each degenerated primers (Flav1 :  $5'\text{CAR GGW SAM TGY GGV WSN TGY TGG DSM TT}^3'$ , Flav2 :  $5'\text{CCG WAD CCM ACN ACM WDM ACM CCR TGR T}^3'$ ).

For Northern blot analysis, total RNA from carcasses (body minus gut and embryos) or guts were prepared from fifty dissected or whole adult aphids. Three  $\mu\text{g}$  of total RNA were analyzed on 1.5% agarose gel containing 2.2M formaldehyde and transferred onto a nylon membrane Gene Screen Plus (NEN<sup>TM</sup>, Life science, Boston, USA). The RNA/DNA hybridization was performed overnight at  $65^\circ\text{C}$  according to Church and Gilbert (20) using a standard protocol with a probe corresponding to the entire cDNA *AgCatL* labeled with  $\alpha\text{-}^{32}\text{PdCTP}$  using the oligolabeling kit from Promega. The same filter was stripped and hybridized to a mitochondrial 18S ribosomal RNA probe to check for RNA loading. The autoradiograms were scanned using a densitometer and data points were corrected for input RNA variations.

### **cDNA cloning**

The gut cDNA library in pTriplex2<sup>TM</sup> vector (SMART cDNA library, Clontech) was constructed according to the manufacturer's instructions. About  $4 \times 10^6$  plaques were differentially screened by PCR using 2 specific *AgCatL* primers deduced from the sequence of the RT-PCR fragment isolated previously (GossF :  $5'\text{CTA CTG GAT CGC TGG AAG GAC}^3'$ , GossR :  $5'\text{GGG CAT CTT CAT CAC CTT CAG G}^3'$ ). The cDNA library was diluted and each quarter of Petri box was screened by PCR. The positive part was titered and subsequently diluted. This step was repeated until obtaining one phagemid clone. The cDNA containing phagemids were rescued from phage according to the kit's instructions. The cDNA was characterized by DNA sequencing on both strands.

### **Genomic Southern-blot hybridization**

Genomic DNA was extracted from whole mature aphid using a CTAB-based method (21). Ten µg of DNA were fully digested with *EcoRI* or *EcoRV* (restriction sites absent in *AgCatL* gene sequence) and run on a 0.8% agarose gel (Sambrook *et al.*, 1989). The DNA was transferred onto a nylon membrane Gene Screen Plus (NEN<sup>TM</sup>, Life science, Boston, USA) and hybridized overnight at 65°C according to standard protocol (20) with a cDNA *AgCatL* probe. The membrane was washed in 2xSSC, 0.1% SDS for 20 min at room temperature (2 washes) and two fold with the same buffer at 40°C, then exposed to X-ray film.

### **Phylogenetic analysis**

A phylogenetic analysis was performed on *AgCatL* to address the following questions: the correct functional assignment of this gene, its linkage to the main classes of vertebrate cysteine proteases (on which the functional classification is based); to compare it to all other known arthropod sequences; to position it relative to the model fully-sequenced insect cysteine proteases (full set of homologs from *Drosophila*).

In order to select vertebrate sequences, the HOVERGEN database was used (22). It contains all vertebrate sequences from GenBank (excluding ESTs), curated to address redundancy issues; homologous coding sequences have been classified in gene families and protein multiple alignments and phylogenetic trees have been computed for each family. One representative sequence was retained from each family of cysteine proteases (B, C, H, S and L ; NC-IUBMB enzyme classification). All these protein sequences were then blasted against GenbankP<sup>TM</sup> to extract homologous arthropod sequences, which were retained after verifying that they possess the three catalytic amino acids. Protein sequences were aligned using the CLUSTALW program (23) followed by manual refinement. Gaps and regions with ambiguous alignments were excluded from the analysis. The phylogenetic tree was generated using the Phylo\_win software (24) and the neighbor-joining algorithm with the PAM

distance. Confidence limits on grouping were determined by the Clustal W bootstrapping technique (500 repeats).

### **Molecular modeling**

As a basis for a spatial localization of critical residues, patterns and features of the aphid cathepsin-L molecule, a first modelling approach was performed through the ExPasy protein resource server (<http://www.expasy.ch>), using the Swiss-Model automated protein modelling facility (25). The target sequence used was the *AgCatL* prosequence (Fig. 4), and the selected templates were two PDB accessions for human Cathepsin-L crystal structures, 1CS8/A-chain (26) and 1CJL (27). Sequence similarity between target and template was high (52% identical and 69% positive residues between *AgCatL* and 1CJL template), ensuring robust alignment and global correctness of the model for the purposes needed herein.

### **Expression in *E. Coli***

The *AgCatL* coding region (1023 bp), excluding N-terminal 57 bp signal peptide sequence and stop codon, was inserted into the polylinker site downstream of the *lacZ* promoter of pET41b expression vector (Novagen, Darmstadt, Germany) in the same orientation as *lacZ* and in frame with several N- and C- terminal tags. Recombinant protein production was induced with isopropyl- $\beta$ -D-thiogalactopyranoside. Inclusion bodies were solubilized with denaturation buffer (6M guanidine-HCL, 300mM NaCl, 50 mM Na phosphate buffer pH 8 and 10mM DTT). Recovered solubilized proteins was purified by affinity chromatography on nickel column in denaturing conditions with imidazole elution (150mM, pH8), according to the manual's instruction (TALON, Clontech). The eluted fraction was then dialyzed against renaturation buffer (25mM Tris-HCL ; pH 8 ; 100mM NaCl ; 5 mM EDTA) and the pH was reduced to 5.5 by addition of HCl to induce autocatalytic processing.

### **Western blot analysis**

Total extracts of aphid guts and of *S. zeamais* at larval stage (positive sample) were prepared by homogenizing tissues in one volume of Laemmli buffer. Proteins were electrophoresed in 10% SDS polyacrylamide gel and electroblotted onto nitrocellulose membrane; the membrane was immunoreacted with antibody against SCP (1:500 dilution in PBS plus 5% milk) and the immunoreacted proteins were visualized using peroxidase conjugated anti-(rabbit IgG) antibody and an ECL kit (Amersham). Since the amino acid sequences of weevil (SCP) and aphid were highly similar (Fig. 4), the anti-SCP antibody (16) was expected to cross react with *AgCatL*.

### **Immunohistochemical study**

Adult aphids were fixed overnight in Bouin-Holland fixative (32% ethanol, 4% neutral formaldehyde, 2.6 mg picric acid, 4% ice cold acetic acid). After dehydration in graded ethanol, the cuticle was softened in 100% butanol for 48h. After rehydration in H<sub>2</sub>O for 15 min, material was included in 1.3% agarose, 0.04% neutral red in order to correctly position the aphid. Agarose blocks of 3 individuals were embedded in paraffin. Serial sections (6µm) were cut and mounted on poly-L-lysine coated slides, de-paraffinized, blocked with 1% bovine serum albumin in PBS and incubated at room temperature overnight with the anti-SCP antibody diluted 1:100 in PBS plus 0.5% BSA. The samples were rinsed in PBS at room temperature and incubated with the secondary antibody, fluorescein isothiocyanate (FITC)-conjugated swine anti-rabbit immunoglobulin, diluted 1:100 in PBS plus 0.5% BSA (one hour at room temperature and in the dark). After rinsing in PBS at room temperature, the sections were mounted in Citifluor mounting medium (Sigma, Steinheim, Germany) and observed under Leica TCSNT confocal laser scanning microscopy (Leica, Heidelberg, Germany) equipped with an argon/krypton laser (Omnichrome, Chino, Canada). To visualize the FITC fluorochrome we used a BP515 band pass filter for emission. All sections were first scanned at low magnification (x16 objective, electronic zoom x1 or x2.64).

## Results

The physiological pH along *A. gossypii* gut was examined by dipping the dissected organ in the pH indicator. This experiment revealed the presence of a pH gradient in the gut, with an acidic segment in the anterior “stomach” part (down to pH 5-6), a smooth gradient towards alkaline pH in the mid and posterior midgut (pH 8 around the mid/hindgut junction), and a last acidification (pH 6) occurring again at the level of hindgut (Fig. 1). These values evaluated visually were confirmed when the gut was dissected and homogenized according to its four cited anatomical parts. .

The pH-dependence of the proteolytic activity in midgut extracts of *A. gossypii* was examined using the <sup>125</sup>I-globulin substrate in a mixed buffer system. Midgut extracts from adult aphid showed a substantial activity between pH 4.5 and 6. The optimum for proteolysis was close to pH 4.5 (Fig. 2), with activity falling severely at lower pH values. Assays were performed in presence of inhibitors and activators specific to different protease types to discriminate this global activity into the four major mechanistic classes described for proteases (28). The proteolytic activity was substantially enhanced (+70%) in presence of DTT, a cysteine protease activator, and E-64 (a highly specific cysteine protease inhibitor) inhibited more than 95% of this activity. The other specific inhibitors did not affect the activity, revealing that cysteine proteases were responsible for most of the detectable protease activity.

Zymogram analysis of aphid gut extracts, performed on polyacrylamide gels transferred onto a 12% gel co-polymerized with gelatin, revealed multiple zones of proteolytic activity. Zymograms were incubated at pH 5 in order to assess aphid gut proteolytic activity at a physiologically relevant pH. Three major bands of activity were resolved, the estimated molecular weights of which being approximately 40 (doublet) and 55 kDa; all of these activities were E-64 sensitive, in agreement with the <sup>125</sup>I-globulin assays (Fig 3).

### **Cloning a full cDNA corresponding to a cysteine protease**

Gut RNA of *A.gossypii* was used first in RT-PCR with 2 sets of degenerated primers designed with insect cysteine proteases already published (29) (16) (30) (31). Several amplimers of 480pb were obtained and cloned. Analysis of all obtained sequences revealed that these RT-PCR fragments were belonging to a single sequence, and related to published thiol proteases. In order to isolate the entire cDNA encoding this cathepsin, a cDNA library was constructed using gut RNA and specific primers were designed according to the isolated PCR fragment and used to screen this library by PCR. The library contained more than  $6.10^6$  independent clones. A full length cDNA (1504 bp) including the previously identified PCR fragment was obtained. The sequence is deposited in EMBL/Genbank/DBJJ database under accession number EMBL : AJ489298 and is referred to as *AgCatL*.

### **Ag cathepsin L sequence analysis**

Comparing the deduced amino acid sequence (Fig. 4) with sequences of the Genbank™ database demonstrated that *AgCatL* belongs to the papain-family of cysteine proteases, and more precisely to the cathepsin L subfamily as unambiguously determined by the phylogenetic analysis (Fig. 5). The cDNA encodes a putative preproprotein, the pre-region being composed of a N-terminal sequence of 19 amino acids, characteristic of a signal peptide as predicted by several algorithms, such as SignalP or TargetP (32). Canonical cleavage of the leader peptide after the 19<sup>th</sup> residue (Ser<sup>-105</sup>) would leave a pro-enzyme of 322 amino acids. The amino acid sequence present an “E<sup>-78</sup>R<sup>-74</sup>[FY]<sup>-70</sup>N<sup>-67</sup>I<sup>-63</sup>N<sup>-59</sup>” motif and a “G<sup>-53</sup>N<sup>-44</sup>F<sup>-42</sup>D<sup>-40</sup>” motif (Fig. 4), both highly conserved in the propeptide segment of all cysteine proteases except the cathepsin B group (33). The “ERFNIN” motif is considered to be a functional unit probably involved in the inhibition of the enzyme activity by the propeptide (33). A single potential N-glycosylation site was observed in the propeptide region (N<sup>-17</sup>FTN<sup>-14</sup>). The potential cleavage site of the pro-region was estimated by homology to be between residues

Val<sup>-1</sup> and Ile<sup>+1</sup>, fitting the scheme of a conserved proline residue at position 2 in the mature enzyme, as for many members of the papain family. This proline may serve to prevent unwanted N-terminal proteolysis (46). The pro-enzyme may be autocatalytically cleaved to produce the mature enzyme consisting of 218 amino acid residues with a predicted molecular mass of 24 kDa, and a pI of 6.98. The catalytic triad Cys<sup>+25</sup>, His<sup>+164</sup> and Asn<sup>+185</sup> is fully conserved in *AgCatL*, and the six cysteine residues that form disulfide bonds in papain between positions 22-65, 56-98 and 157-207, conserved in other cysteine-proteases, are also present in the *A. gossypii* sequence. Overall the deduced amino-acid sequence of *A. gossypii* cathepsin-L is very similar to other insect cysteine-proteases (75% similarity to *Drosophila*) and to mammalian cathepsin-L (70% similarity with human, Fig. 4).

### **Structural inferences**

The tentative structure produced for *AgCatL* by homology modelling (not shown) differs significantly from its human template by three "junctional" regions easily identifiable by the gaps in the alignment: the  $\alpha 1/\alpha 2$  helix boundary of the prosegment (-80 region), the prosegment C-terminus (-15 region) and the putative light/heavy chain junction (+175 region). As they all occur at the periphery of the molecule, the catalytic consequences of such large differences might be reduced. A few questions were further addressed to better interpret the functional characteristics of the aphid enzyme. Catalytic activity and its regulation by the propeptide region were thoroughly investigated on the human enzyme (26,27,34). When exploring the important residues involved in the structural interactions between the enzyme and its N-terminal proregion, most of which are supported by aromatic and hydrophobic interactions (27), one ends up with the extreme conservation between the human, aphid and other arthropod sequences (14 conserved on 15 critical residues, highlighted grey in Fig. 5). The only missing hydrophobic interaction concerned residue K<sup>-66</sup> in the ERFNIN motif (M in human; variant in the 4 lower sequences of Fig. 4), and supports  $\alpha 1/\alpha 2$  stabilizing



interactions in the propeptide (27). When looking at the catalytic activity, through residues interacting with the E64 inhibitor (34), a similar pattern of conservation is revealed (24 conserved residues over 26 critical, boxed in Fig. 4). In this case, the two variant positions in the aphid sequence (Q<sup>145</sup> and T<sup>190</sup> resp. from L and E in human) share a common trend of variation among arthropod sequences, and both positions concern the so-called S' region of the catalytic cleft, which seems important for the selectivity of inhibitors between cathepsins of the B and L groups (34). One consequence could be that insect cathepsins display non-standard behaviour towards specific inhibitors of mammalian cathepsin families. Finally, two acidic residues crucial for the pH-dependency of the activation of cathepsin L are also conserved in the insect sequences, highlighting probable common activation features for all these enzymes (except the crustacean, see D<sup>-40</sup> and E<sup>-35</sup> in the  $\alpha$ 3 helix of prosegment, Fig. 4). The last functional inference from *AgCatL* structure lies in the identification of potential lysosomal targetting signals, which associate in the human sequence a N-glycosylation site (N<sup>108</sup>, blue in Fig. 4) to a lysine tandem responsible for the recognition by the phosphotransferase specific to the mannose-phosphate targetting pathway (K in green in Fig. 4). In *AgCatL*, one potential glycosylation site exists in the proregion, and is conserved in four of the six presented arthropod sequences (Fig. 4). This site is located at the same pole of the cathepsin L molecule when compared to the human N-glycosylation site (only 20 Å separate them in the model structures), despite being quite apart on the primary sequence. Also, lysine tandems may be identified in the aphid prosequence, which could fit the structural constraints set-up for the human recognition signal: K-72 and K-26 in *AgCatL* are 33 Å apart, close to the 34 Å rule designed for the human Cat D/Cat L signal (35). In contrast to this scheme, no potential lysosomal targetting signal may be identified in the *Phaedon* or *Homarus* sequences (Fig. 4).

### **Phylogenetic affiliation**

To further investigate the relationships between the cloned cathepsin sequence and those from other insect and animal organisms, a protein phylogenetic tree was constructed. The alignment displayed high similarities only in the immediate vicinity of the active site residues Cys<sup>+25</sup>, His<sup>+164</sup> and Asn<sup>+185</sup>, as described earlier (36). The phylogenetic analysis was extended to include 40 representative vertebrate cathepsin sequences as well as all the arthropod cysteine protease sequences known at that time (including 24 insect sequences and all the “relevant” cysteine protease genomic sequences from *Drosophila*; a couple of related sequences from this species, at the cathepsin-B boundary lacking active site residues from the canonical catalytic triad, were excluded from the analysis to ensure correct alignment and sufficient informative sites). The selected alignment runs over the 311 residues delineating the mature enzyme, resulting in a data set of 86 informative sites. Four major clusters were identified (Fig. 5). The first cluster contained an assortment of vertebrate and insect cathepsin L-related enzymes (*s.l.*). Being located very close to the *Sitophilus zeamais* (weevil) sequences, *AgCatL* belonged to this group, and more precisely to a well-defined subgroup containing several insect and arthropod sequences (from the shrimp *Artemia* to the fruitfly *Drosophila*). This subgroup is likely to represent the ortholog group to the human cathepsin L gene (*s.s.*). In two of the other following clusters (cathepsin H and C), no insect representative was included. The last cluster, corresponding to cathepsin B-like enzymes, included five other insect sequences, two of which belonged to the *Drosophila* genome. The cathepsin B family having diverged from the cathepsin L family even before the divergence of cysteine proteases from plant and from lower organisms (36), the cathepsin B cluster was chosen as the outgroup for our phylogenetic analysis.

### **Analysis of cysteine proteinase gene family**

The results of Southern hybridization experiments suggest the existence of a gene family encoding cysteine proteases in *A. gossypii* (Fig. 6a). Although only one high hybridizing band

was detected at high stringency, prolonged exposure (17 days at  $-80^{\circ}\text{C}$ ) revealed two other weekly hybridizing signals for each digestion.

Oligonucleotide primers corresponding to 5' and 3' segments of *AgCatL* cDNA sequence were used for PCR amplification of *A. gossypii* genomic DNA. The fragment amplified from genomic DNA was found to be similar in size to the cDNA product, indicating that the genomic locus does not harbor introns in the whole region encoding the preproprotein.

### **Pattern of tissue expression**

Northern-blot experiments were performed to determine the expression profile of the *AgCatL*. Total RNA isolated from whole insect, gut, and carcass (whole body minus gut and embryos) were hybridized with a radiolabeled *AgCatL* probe. Two transcripts of 1.6 and 2.0 kb were expressed at a high level in the gut, and at least ten fold higher than the levels detected in carcasses (Fig. 6b).

### **Expression in *E.coli***

The fusion protein of 70 kDa (corresponding to the expected molecular weight) was observed on SDS PAGE and western blots from *E. coli* cells after induction (Fig. 6c). The polypeptide was accumulated in inclusion bodies. Recovered solubilized fusion protein was purified by affinity chromatography in denaturing conditions. The pH of eluted fraction was lowered to 5.5 to promote potential autocatalytic processing (37). Maturation of cathepsin proteases needs propeptide processing, and is generally achieved by autocatalytic activation at acidic pH, or by proteolytic activation. Anti-SCP antibodies detected 2 bands in this acidic extract, one corresponding to the fusion protein and the other to the putative mature enzyme (about 30 kDa), since it migrated at the same level as the lower band observed in aphid and *S. zeamais* guts (Fig. 6c). However, no enzymatic activity could be detected in these extracts.

In total extracts from *S. zeamais* larvae, as in total extract from *A. gossypii* midgut, the antibody recognized two bands in immunoblot analyses (Fig. 6c). The upper band (about 49 kDa) is likely to represent the proenzyme form of CatL and the other (approximately 30 kDa) should correspond to the mature *AgCatL* enzyme, by analogy with SCP (16). Both forms are therefore detectable in aphid midguts. However, in contrast to *S. zeamais* gut extract, *A. gossypii* guts did not contain equal amounts of the 49 and 30 kDa forms, but contained significantly more of the large *AgCatL* 49 kDa (pro-)form.

### **Tissue distribution of *AgCatL***

In order to localize *AgCatL* in the aphid, adult *A. gossypii* reared on their host plant (melon) were fixed with paraformaldehyde, included in paraffin and sectioned. Fixed sections were immunoreacted with antiSCP antibodies. An intense and specific signal was detected intracellularly in aphid digestive tissues, with a clear gradient decreasing along the digestive tract (Fig. 7) ; the anterior “stomach” part was always strongly labeled. Homogeneity of cell staining indicated an intracytoplasmic localization of *AgCatL*. This protease was not clearly detected in the reddish-coloured apical microvillar structures, and therefore does not seem to be secreted.

A strong signal was also observed at the level of the bacteriocytes, which contained *A. gossypii* endosymbionts (*Buchnera aphidicola*, an intracellular obligate enterobacterial symbiont of aphids (38)). However, since antibodies were raised against SCP expressed in *E. coli* cells, the staining could result from artifactual cross-detection of contaminants. A control with a batch of affinity purified antibody was therefore performed, and although the signal was lower in these slides, it remained strong enough to confirm the presence of CatL at the bacteriocyte level. CatL activity was also directly detected in this tissue by enzymatic assay on dissected bacteriocytes with Z-Phe-Arg pNA substrate, specific to cysteine proteinase (39). By contrast, the presence of *AgCatL* was not clearly identified in fat body or other tissues

(Fig. 7), which may differ from the situation described for *Sarcophaga* (40), *Drosophila* (41) or *Sitophilus* (16).

## **Discussion**

In spite of the relative abundance of proteins in the phloem of some plant species, which include preferred hosts of the cotton-melon aphid *A. gossypii*, polypeptides are not usually considered to be a major source of dietary nitrogen for most phloem-feeding insects (13, 14, 42). Recently however, a whitefly species, although being also a strict phloem-feeding Homoptera, was shown by radiotracer experiments to ingest plant proteins, to degrade part of them to free amino acids and to use them for *de novo* protein synthesis (43). Therefore, the previous inability to detect any protease activity in digestive tracts of Homoptera does not seem to reflect the reality. The main reasons for this misinterpretation, at least for aphids, was probably the insolubility and the dependence on a single catalytic class of their digestive tract endoproteases.

In this study, we have shown that midgut extracts of *A. gossypii* adults do possess clear endoproteolytic activity for both proteins and synthetic substrates. Acidic pH optimum, enhancement by reducing agents and sensitivity to a cysteine protease inhibitors clearly suggests a strong dominance of cysteine proteases in this aphid. While in-gel assays may restrict the detection to those proteases able to renature after SDS/electrophoretic treatments, all our enzymatic and molecular results are consistent with the occurrence of this single major class in aphid midguts. These results are confirmed by a more extensive study of the hydrolytic enzymes of the model and larger pea aphid, for which a cysteine protease has been purified and is now available for deeper enzymological characterization (18).

One cDNA encoding a putative cysteine protease has been cloned from *A. gossypii* midguts. Thanks to its gut-specific expression pattern, to its identification as a regular cathepsin L

enzyme, compatible with the substrate specificity shown by whole midgut extracts (preferential activity on Z-Phe-Arg substrate), and to the presence of a single major homologous gene encoding CatL in *A.gossypii*, this gene is probably responsible for the main activity observed in aphid midguts. The presence of two transcripts may reflect the use of two consensus polyadenylation sites. A preliminary systematic sequencing of a subtracted cDNA library (guts vs carcasses) from *A. gossypii* confirms the strict reliance of aphid guts on cysteine proteinase, in that only three endopeptidase-related transcripts were identified in this library, all of the cysteine type (*AgCatL* plus two additional cathepsin B-related cDNAs; Deraison, unpublished results). Within *AgCatL* cDNA sequence, a single long open reading frame was found which encoded a typical CatL preproprotein of 341 amino acids. Propeptide structure seemed canonical, processing was observed both *in vivo* and *in vitro*, although no resulting activity was obtained from the *E. coli* expressed/processed *AgCatL* gene product. This may be due to an incorrect folding process in *E. coli*, which involves the formation of three potential disulfide bridges and relies on chaperone properties of the propeptide (44). The overall pattern of structure and expression of *AgCatL* is reminiscent of what occurs in *Drosophila*, where a single orthologous gene exists (CATL\_DROME, or CP1) and has been analysed (31, 41, 45), showing also partial gut-specific expression. One difference may lie in the absence of introns in the aphid gene (3 in the *Drosophila* gene), but this has to be confirmed by a more complete analysis of the genomic sequence of *AgCatL*.

Regarding the catalytic properties of *AgCatL*, as derived from comparative sequence and structure information, the striking point is the extreme conservation of all residues that were shown to be critical to the activity of a canonical mammalian cathepsin L (except the S' subsite residues already discussed, and featuring specificities common to all insect orthologous CatL). Among various subfamilies of lysosomal cysteine proteases (B, H or L) cathepsin L favors aromatic residues at P<sub>2</sub>, which distinguishes it for example from the

closely related enzymes cathepsin S and K. Human Ala<sup>135</sup> (as opposed to Gly found in cathepsin S), is strongly responsible for this discrimination (46), and this diagnostic conservation is present as Ala<sup>136</sup> in *AgCatL* (Fig. 4). This extreme conservation is not the rule for all arthropod CatL, but restricted to our so-called orthologous group, featuring *Drosophila*, *Aphis* or *Sitophilus* sequences: the *Phaedon* sequence does not conserve the critical Ala<sup>136</sup> residue, and the *Homarus* sequence does not conserve residue E<sup>135</sup>, previously quoted as crucial for the acid autoactivation of procathepsin L (Fig. 4). When comparing the phylogenetic grouping of all arthropod cathepsin L sequences (Fig. 5), it is striking to note that three groups emerge, and share within their members both phylogenetic and functional features that may indicate specific “physiological” recruitments for these sequences during the evolution of the family:

- i) the crustacean sequences may represent a distinct “digestive” recruitment from a basal cathepsin L lineage, having experimentally proved to be secreted proteins from the digestive tract of their source organism (47);
- ii) in contrast to this, none of the members of the *AgCatL* group have been shown biochemically to be present in the digestive secretions of its source organism, and although different members are preferentially expressed in the digestive tract, including *AgCatL*, they may be regarded as a whole as “standard”, and functionally conserved, insect lysosomal enzymes (therefore called the insect orthologous CatL group);
- iii) finally, a group of quite divergent enzymes from Coleoptera, probably secreted, may have appeared from the same basal CatL group, but alternatively also from more distantly related CatL-like groups, such as the CatF lineage or other genes present at this point in the *Drosophila* genome (Fig. 6).

As for the subcellular expression/localisation of *AgCatL*, the situation might be complex, as pointed out already by the presence of two related transcripts in gut mRNAs. On one hand, *AgCatL* features share many standard traits of lysosomal cathepsins:

- i) structural/functional conservation indicative of strong selective pressures by structural constraints on (endogenous, conserved ?) substrates,
- ii) propeptide organisation, including acid activation characteristic of the lysosomal compartment in order to protect cells from the potentially disastrous consequences of uncontrolled degradative activity, essentially all known cellular proteolytic enzymes are synthesized as inactive precursors. Activation of the enzyme is accomplished by cis- and trans- cleavage of the propeptide (48) in very acidic conditions (37).
- iii) strict intracellular expression, as observed by immunolocalisation and enzymatic assays
- iv) presence of *potential* lysosomal targeting signals through a mannose-phosphate pathway.

The latter trait may not be consensual, and have been challenged in a insect heterologous cell expression system (49), but was shown here to be strikingly correlated with the three functional/phylogenetic groups of arthropod CatL discussed above: the presence of a *conserved* N-glycosylation pattern in the prosequence region –20 (a necessary though not sufficient condition for mannose-phosphate targeting) is present in all insect sequences of the AgCatL clade, while it is absent from all sequences in both the crustacean clade and the divergent coleopteran group (*Hypera*, *Phaedon*, *Diabrotica*), all encoding putative secreted cathepsins (Fig. 5).

On the other hand, some other traits might not fit a strict lysosomal expression, such as:

- i) the striking insolubility of the CatL/Z-Phe Arg activity in absence of SDS,
- ii) the consistent gut-specific expression of many insect CatL genes, associating this molecule with (intracellular ?) digestive functions,
- iii) the detection of a major pro-enzyme pool in aphid and other insect cells by immunoblotting (16,50),
- iv) the localization of cathepsin L by EM/immunogold labelling in vesicular *and* apical membrane structures of an aphid enterocyte (18).



Indeed, this complex expression pattern is also a typical feature of mammalian CatL, with multiple expression phenotypes (lysosomal vs secretory) depending on cell context, associated with potential alternative splicing and numerous post-transcriptional signals (N-/C-terminal propeptide and mature enzyme structures necessary for folding and secretion, signalling for mannose-phosphate dependent and independent lysosomal targeting...). Insect cathepsins from the orthologous CatL group might share this expression complexity, which renders them genes and molecules of interest to study in model (*Drosophila*, *Bombyx*) or non-model organisms (*Sitophilus*, *aphids*), to better understand their involvement in physiological functions as diverse as digestion, development and may be in other processes (defense, detoxication ?).

## Acknowledgements

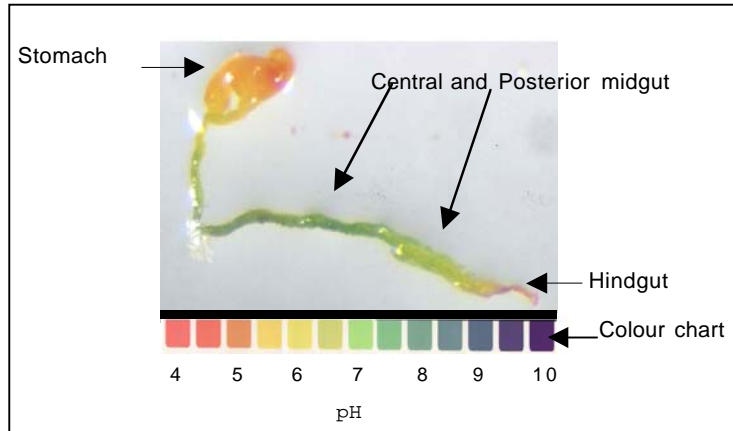
We thank Marie-Gabrielle Duport for her help in aphid rearing and dissection, and gratefully acknowledge the European Union for a doctoral fellowship to CD, through a FAIR collaborative program (FAIR6-CT98-4239).

## References

1. Pitrat, M., and Lecoq, H. (1982) *Agronomie* **2**, 503-508
2. Porter, D. R., Burd, J. D., Shufran, K. A., and Webster, J. A. (2000) *Environ. Entomol.* **29**(4), 1315-1318
3. Khush, G. S., and Brar, D. S. (1991) *Adv. Agron.* **45**, 223-274
4. Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., and Williamson, V. M. (1998) *Proc. Natl. Acad. Sci. USA* **95**(17), 9750-9754
5. Flor, H. H. (1987) *Annu. Rev. Phytopathol.* **25**, 59-66
6. Rahbé, Y., and Febvay, G. (1993) *Entomol. Exp. Appl.* **67**(2), 149-160

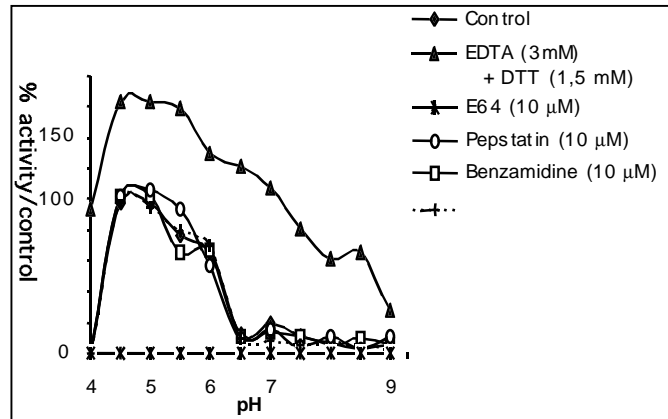
7. Gatehouse, A. M. R., Shi, Y., Powell, K. S., Brough, C., Hilder, V. A., Hamilton, W. D. O., Newell, C. A., Merryweather, A., Boulter, D., and Gatehouse, J. A. (1993) *Phil. Trans. R. Soc. Lond. B* **342**(1301), 279-286
8. Thomas, J. C., Adams, D. G., Keppenne, V. D., Wasmann, C. C., Brown, J. K., Kanost, M. R., and Bohnert, H. J. (1995) *Plant Cell Rep.* **14**(12), 758-762
9. Jouanin, L., Bonadé-Bottino, M., Girard, C., Lerin, J., and Pham-Delegue, M. H. (2000) in *Recombinant protease inhibitors in plants* (Michaud, D., ed), pp. 179-190, Landes Bioscience, Austin, TX (USA)
10. Lee, S. I., Lee, S. H., Koo, J. C., Chun, H. J., Lim, C. O., Mun, J. H., Song, Y. H., and Cho, M. J. (1999) *Mol. Breed.* **5**(1), 1-9
11. Tran, P., Cheesbrough, T. M., and Keickhefer, R. W. (1997) *J. Econ. Entomol.* **90**(6), 1672-1677
12. Ziegler, H. (1975) in *Encyclopedia of plant physiology* (PIRSON A., Z. M. H. E., ed) Vol. 3, pp. 59-100, Springer Verlag, Berlin (GER)
13. Terra, W. R., and Ferreira, C. (1994) *Comp. Biochem. Physiol.* **109B**(1), 1-62
14. Mochizuki, A. (1998) *Appl. Entomol. Zool.* **33**(3), 401-407
15. Cronshaw, J., and Sabnis, D. D. (1989) in *Sieve elements - comparative structure, induction and development* (Behnke, H. D., and Sjolund, R. D., eds), pp. 257-283, Springer, Berlin
16. Matsumoto, I., Emori, Y., Abe, K., and Arai, S. (1997) *J. Biochem.* **121**(3), 464-476
17. Sarath, G., Delamotte, R. S., and Wagner, F. W. (1989) in *Proteolytic enzymes. A practical Approach* (Beynon, R. J., Bond, J.S., ed), pp. 25-56, IRL Press, Oxford, G.B.
18. Cristofolletti, P. T., Ribeiro, A. F., Rahbé, Y., Deraison C. and Terra, W. R. (2002) *J. Insect Physiol.* **in prep**
19. Michaud, D., Faye, L., and Yelle, S. (1993) *Electrophoresis* **14**, 94-98
20. Church, G. M., and Gilbert, W. (1984) *Proc Natl Acad Sci U S A* **81**(7), 1991-5.
21. Doyle, J. J., Doyle, J. L., Brown, A. H., and Grace, J. P. (1990) *Proc. Natl. Acad. Sci. U S A* **87**(2), 714-7.
22. Duret, L., Mouchiroud, D., and Gouy, M. (1995) *Nucl. Acids. Res.* **22**, 2360-2365
23. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucl. Acids Res* **22**(22), 4673-80.
24. Galtier, N., Gouy, M., and Gautier, C. (1996) *Comput. Applic. Biosci.* **12**, 543-548
25. Guex, N., Diemand, A., and Peitsch, M. C. (1999) *TIBS* **24**(9), 364-7.
26. Groves, M. R., Coulombe, R., Jenkins, J., and Cygler, M. (1998) *Proteins* **32**(4), 504-14.
27. Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J. S., and Cygler, M. (1996) *Embo J.* **15**(20), 5492-503.
28. Dunn, B. M. (1989) in *Proteolytic enzymes. A practical Approach* (Beynon, R. J., Bond, J.S., ed), pp. 57-104, IRL Press, Oxford, G.B.
29. Girard, C., and Jouanin, L. (1999) *Insect Biochem. Molec. Biol.* **29**(12), 1129-1142

30. Yamamoto, Y., Zhao, X., Suzuki, A. C., and Takahashi, S. Y. (1994) *J. Insect Physiol.* **40**(5), 447-454
31. Tryselius, Y., and Hultmark, D. (1997) *Insect Mol. Biol.* **6**(2), 173-81
32. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng.* **10**(1), 1-6
33. Karrer, K. M., Peiffer, S. L., and DiTomas, M. E. (1993) *Proc. Natl. Acad. Sci. U S A* **90**(7), 3063-7.
34. Fujishima, A., Imai, Y., Nomura, T., Fujisawa, Y., Yamamoto, Y., and Sugawara, T. (1997) *FEBS Lett.* **407**(1), 47-50.
35. Cuozzo, J. W., Tao, K., Cygler, M., Mort, J. S., and Sahagian, G. G. (1998) *J. Biol. Chem.* **273**(33), 21067-76.
36. Berti, P. J., and Storer, A. C. (1995) *J. Mol. Biol.* **246**(2), 273-83.
37. Jerala, R., Zerovnik, E., Kidric, J., and Turk, V. (1998) *J Biol Chem.* **273**(19), 11498-504.
38. Charles, H., Heddi, A., and Rahbé, Y. (2001) *C. R. Acad. Sci.* **324**, 489-494
39. Rahbé, Y., Deraison, C., Bonadé-Bottino, M., Girard, C., Nardon, C., and Jouanin, L. (2001) *Plant Sci.* **in prep.**
40. Fujii-Taira, I., Tanaka, Y., Homma, K. J., and Natori, S. (2000) *J. Biochem.* **128**(3), 539-42
41. Matsumoto, I., Watanabe, H., Abe, K., Arai, S., and Emori, Y. (1995) *Eur. J. Biochem.* **227**(1-2), 582-7
42. Rahbé, Y., Sauvion, N., Febvay, G., Peumans, W. J., and Gatehouse, A. M. R. (1995) *Entomol. Exp. Appl.* **76**, 143-155
43. Salvucci, M. E., Rosell, R. C., and Brown, J. K. (1998) *Arch. Insect Biochem. Physiol.* **39**(4), 155-165
44. Matsumoto, I., Abe, K., Arai, S., and Emori, Y. (1998) *J. Biochem.* **123**(4), 693-700
45. Gray, Y. H., Sved, J. A., Preston, C. R., and Engels, W. R. (1998) *Insect Mol. Biol.* **7**(3), 291-3
46. Barrett, A. J., Rawlings, N. D., and Woessner, J. F. (1998) *Handbook of proteolytic enzymes*, CD-ROM edition Ed., Academic Press, London (UK)
47. Laycock, M. V., Hiramata, T., Hasnain, S., Watson, D., and Storer, A. C. (1989) *Biochem J.* **263**(2), 439-44.
48. Menard, R., Carmona, E., Takebe, S., Dufour, E., Plouffe, C., Mason, P., and Mort, J. S. (1998) *J. Biol. Chem.* **273**(8), 4478-84.
49. Aeed, P. A., and Elhammer, A. P. (1994) *Biochemistry* **33**(29), 8793-8797
50. Homma, K., Kurata, S., and Natori, S. (1994) *J Biol Chem* **269**(21), 15258-64.



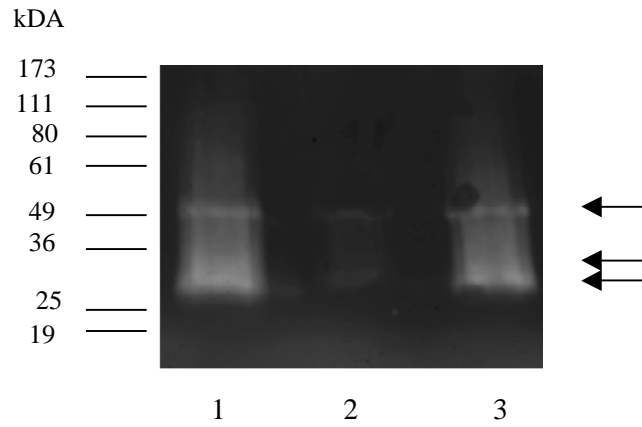
**Fig 1, Physiological pH in *A. gossypii* digestive tract.**

After dissection, the gut was dipped in universal indicator solution. Colour change was observed under binocular and compared with standards made from buffers of known pH.



**Fig 2, pH-dependence of proteolytic activity of *A. gossypii* guts.**

Whole extracts from the digestive tract were incubated with  $^{125}\text{I}$ -globulin solution (pH4-9) and different diagnostic inhibitors or activator at 37°C for 24h. Hydrolysis was measured by the resulting radioactivity in supernatants from TCA precipitation.



**Fig 3, Protease activity gel.**

Two midgut equivalents were loaded in each lane; (1, standard conditions ; 2, 3, E64 10  $\mu$ M or PMSF 10  $\mu$ M respectively were added to the sample 15 min before loading)

-120 -110 -100 -90 -80 -70

Homo s. MN-PTLLILAAFLCLGIASA-TLTFDHSLEAQWTKWKAMHNRLY-GMNEEGWRRAVWEKINMKM  
 Homarus a. MK----VVALFLFGLALA-----AANPSWEEFKGKFGRKYVDLEEERYRLNVFLDNLQY  
 Phaeton c. MK-LIIALAALI-VVINAA-----SDQELWADFKKTHARTYKSLREELKRFNIFQDILLRQ  
 Drosophila m. MRTAVL-LPLLA-LLAVAQAVSFADVVMEEWHTFKLEHRKNYQDETEERFRLLKIFNENKHK  
 Sitophilus z. MK-LF--LILAA-VVISCQAVSFYDLVQEQWSSFKMQHSKNYDSETEERFRMKIFMENAHK  
 Bombyx m. MK-CLVLLLC--AVAAVSAVQFF-DLVKEEWSAFKLOHRLNYKSEVEDNFRMKIYAEHKHI  
 Aphis g. MKVVIV-LGLVAFAISTVSSINLNEVIEEWSLFIQFKKLYEDIKETEFRKKVYLDNKLK  
 Patterns x1 E---R---F--N---

-60 -50 -40 -30 -20 -10

Homo s. IELHNQEYREKHSFTMAMN--AFGDMTSEEFQVMNGFQNRKPR-----K GK  
 Homarus a. IEEFNKKYERGEVTYNLAIN---QFSDMTNEKFNAMVMKGYKKGPR-----PAA  
 Phaeton c. IAEHNVKYENGESTYLLAIN---KFSDITDEEFRDMLMKNEASRPN-----LEGL  
 Drosophila m. IAKHNQRFAGEKVSFKLAVN---KYADLLHHEFRQLMNGFNNTL--HKQLRAADESFKGV  
 Sitophilus z. VAKHNKLFVQGFVKFKLGLN---KYADMLHHEFVSTLNGFNKTK--NNILKGSGLN-DAV  
 Bombyx m. IAKHNQKYEMGLVSYKLGMSNSWHEGDMHHEFVKTMNGFNKTAKHNKNLYMKGGSVRGA  
 Aphis g. IAGHNKLYESGEETAYALEMN---HFGLDMQHEYTKMMNGFKPSLAGGDRNFNTND---EAV  
 Patterns I---N G-----N---F-D NXX

1 10 20 30 40 50

Homo s. VFQEPLFYEAPRSVDWREKGYVTPVKNQGGCGSCWAFSATGALEGGMFRKGTGRLISLSEQ  
 Homarus a. VFTSTDAAPESTEVDWRTKGAVTPVKDQGGCGSCWAFSTTGGLEGQHFLKTGRLVSLSEQ  
 Phaeton c. EVADLTVGAAPEIDWRSKGVVLPVRNQGEGCGSCWALSTAAAIESQSAIKSGSKVPLSPQ  
 Drosophila m. TFISPANVTLPKSVDWRTKGAVTAVKDQGHCGSCWAFSSTGALEGGHFRKSGVLSLSEQ  
 Sitophilus z. RFISPANVTLPTVDWRTKGAVTAVKDQGHCGSCWAFSSTGALEGGHFRKSGVLSLSEQ  
 Bombyx m. KFISPANVTLPEQVDWRTKHGAVTDIKDQGGCGSCWAFSSTGALEGGHFRQSGVLSLSEQ  
 Aphis g. TFLKSENVVTPKSVWDRKKGAVTAVKQGGCGSCWAFSSTGALEGGHFRKGTGRLVSLSEQ  
 Patterns y1 QXXCGXCWXXSA

60 70 80 90 100 110

Homo s. NLVDCSG-PQGNENGGLMDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYSVANDT  
 Homarus a. QLVDCAGGSYINQGGNGGWERAIMYVRDNGGVDTESSYPYEARDNTRCFNSNTIGATCT  
 Phaeton c. QLVDCST-SYGNHGCNGGFAVNGFEYVKDN-GLSDADYPYSGKEDKCKANDKSRSVVEL  
 Drosophila m. NLVDCST-KYGNNGCNGGLMDNAFRYIKDNGGIDTEKSYPYEAIDDSCHFNGKTVGATDR  
 Sitophilus z. NLVDCSG-RYGNNGCNGGLMDNAFRYIKDNGGIDTEKSYPYLAEDEKCHYKAQNSGATDK  
 Bombyx m. NLIDCSE-QYGNNGCNGGLMDNAFRYIKDNGGIDTEQAYPYEGVDDKCRYNPKNTPGAEDV  
 Aphis g. NLIDCSR-KYGNNGGEGGLMDLAFKYIKSNKGLDTEKSYPYEAEDDKCRYNPENSGATDK  
 Patterns NXT

120 130 140 150 160 170

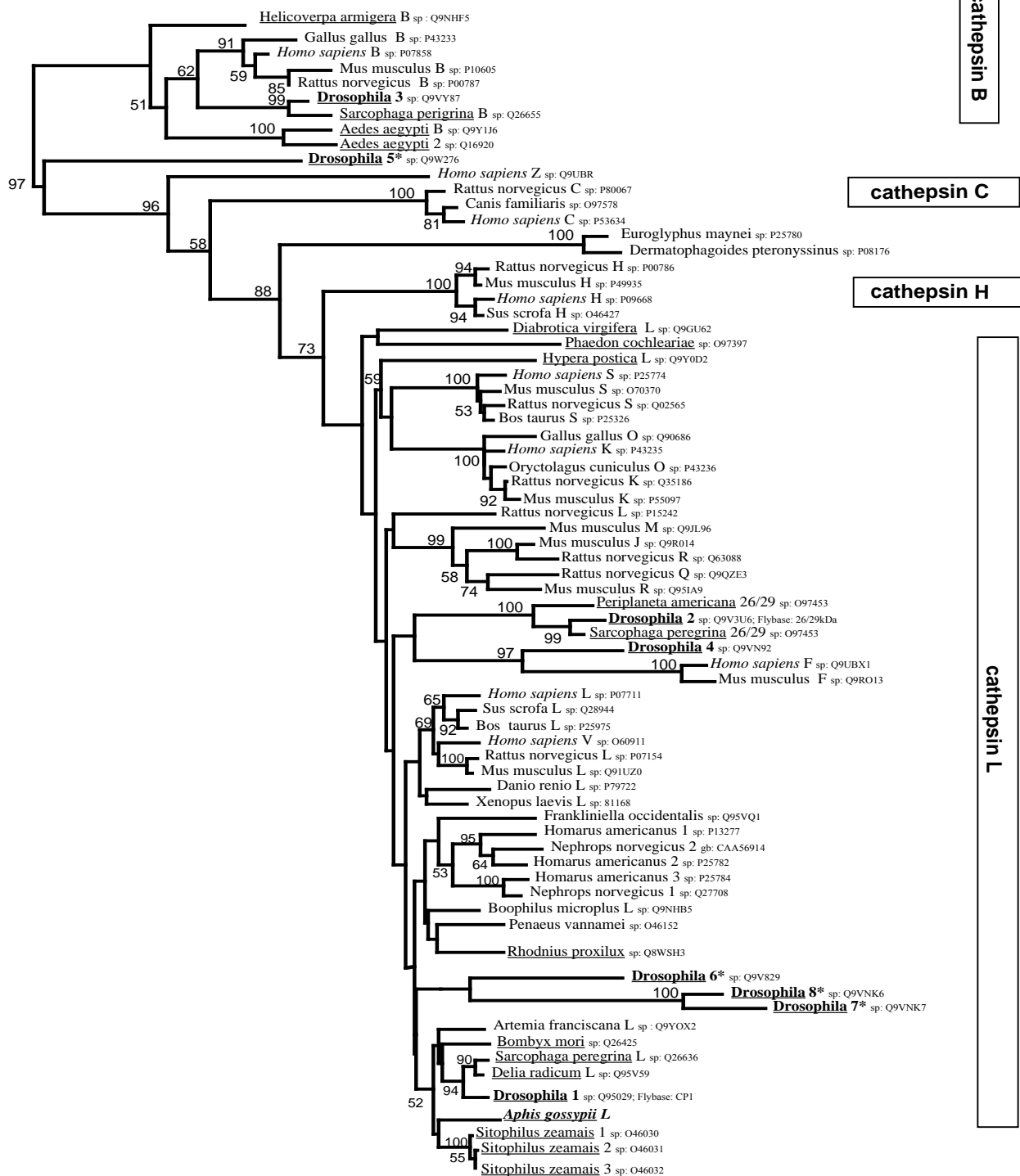
Homo s. GFVDIPK-QEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDHGVLVVG  
 Homarus a. GYVGIAQGESALKTATRDIGPISVAIDASHRSFQSYTGVYEPSCSSQLDHAVALVG  
 Phaeton c. TGYKKVTASETSLKEAVGTIGPISAVVFG--KFMKSYGGIFDSSCLGDNLHHGVNVVG  
 Drosophila m. GFTDIPQGDQKMAEAVATVGPVSVDAIDASHESFQFYSEGVYNEPQCDANLDHGVLVVG  
 Sitophilus z. GFVDIEEANEDDLKAAVATVGPVSVDAIDASHETFQLYSDGVYSDPECSSQLDHDHGVLVVG  
 Bombyx m. GFVDIPEGDEQKLMEAVATVGPVSVDAIDASHTHFQLYSSGVYNEEECSSTDLHDHGVLVVG  
 Aphis g. GFVDIPEGDEDALMHALATVGPVSVDAIDASSEKFFQFYKKGVFYNPRCSSTELDHDHGVLVVG  
 Patterns X LXHGVLAVG

180 190 200 210 220

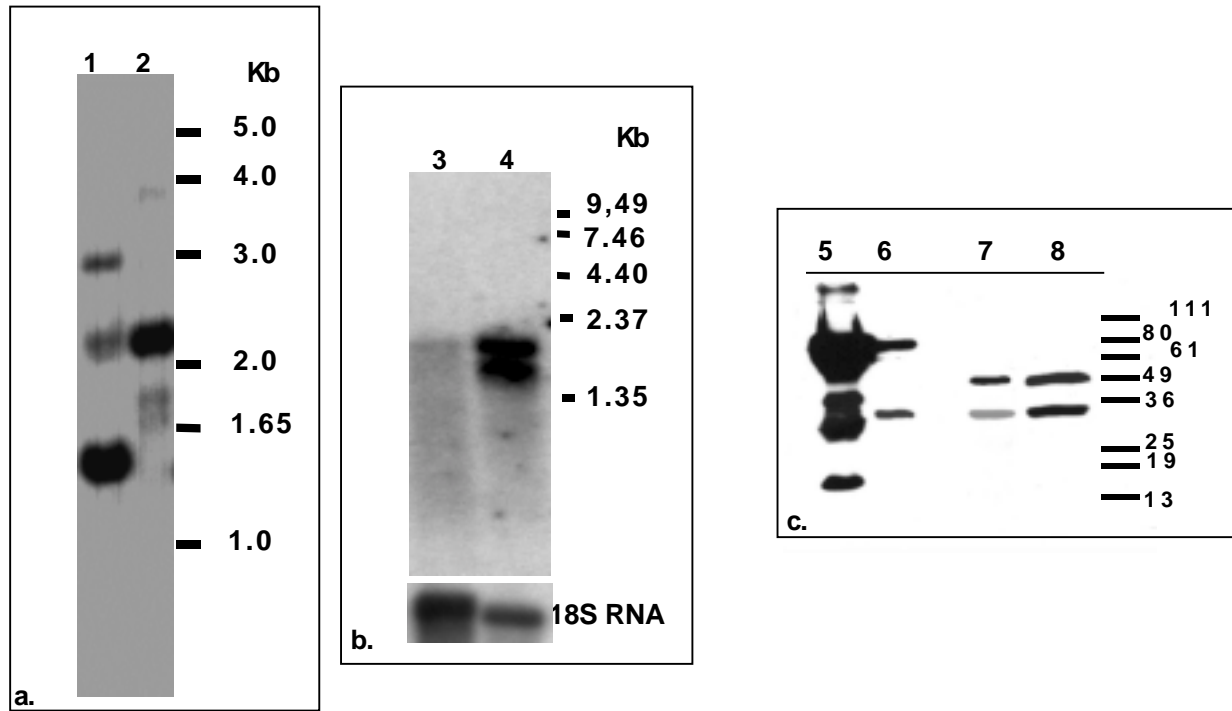
Homo s. YGFESTESDNKNYWLVKNSWGEWGMGGYVKMAKDRRNHCGIASAASYPTV-  
 Homarus a. YG----SEGGQDFWLVKNSWATSWGESGYIKMARNRNNNCGIATDACYPTV-  
 Phaeton c. YG----IENGQKYWIINKTWGADWGESGYIRLIRDTHSCGVEKMASYPLA  
 Drosophila m. FG---TDESGEDYWLKNSWGTWGDGKFIKMLRNKENQCGIASASSYPLV-  
 Sitophilus z. YG---TSDDGQDYWLKNSWGPSWGLNGYIKMARNQDNMCGVASQASYPLV-  
 Bombyx m. YG---TDEQGVYWLKNSWGRSWGELGYIKMIRNKNRNCGIASSYPLV-  
 Aphis g. FG---SDKKGGDYWLVKNSWGTWGDGYIMMARNKKNRNCGVASSYPLV-  
 Patterns XG z1 YWIXKNSWXXXWGXXGYIXM

**Fig 4, Deduced amino acid sequence of AgCatL (predicted mature sequence numbering) and comparison with selected cysteine protease sequences from the Cathepsin L subgroup** (swissprot accessions *Homo sapiens* P07711 ; *Homarus americanus* P13277 ; *Pheadon cochleariae* O97397; *Drosophila melanogaster* Q95029 ; *Sitophilus zeamais* O46030 ; *Bombyx mori* Q26425 ). Gaps were introduced by clustalW to maximise the homology between the prosequences, and indicated by dots. Active site residues C<sup>25</sup>, H<sup>164</sup> and N<sup>185</sup> (AgCatL numbering) are highlighted in yellow, and Prosite consensus patterns are reported below. Yellow also highlights conserved carboxylic side chains important for acid autocatalytic processing of L cathepsins. In grey are residues potentially implicated in tight intramolecular interactions, either of the covalent - disulfide bridges-, aromatic, or hydrophobic types. When present, the potential lysosomal-addressing/N-glycosylation site is highlighted in blue and its pattern reported; likewise, lysine residues involved in the mammalian lysosomal mannose-phosphate processing signal are highlighted in green. The site of predicted signal peptide cleavage is shown as pattern x1, that of the activation peptide as y1 (by homology to human sequence), and the putative site of cleavage between eventual light and heavy chains as z1 (1 marking the potential downward leaving chain N-terminus). Residues involved in the active site cleft are boxed in the human sequence, as well as their variants in the aphid sequence





**Fig. 5, Phylogenetic tree based on cysteine protease sequences, constructed from a PAM matrix alignment and the neighbor-joining algorithm. The tree was rooted from the CatB outgroup. Numbers near the nodes indicate % of bootstraps supporting a given node (500 replicates). The scale bar correspond to 0.1 amino acid substitution per site. Insect sequences are underlined, *Drosophila* sequences are boldfaced and numbered for the different loci in the genomic sequence (\* for sequences found in ESTs but lacking one of the catalytic domain signatures), human sequences are italicized. Cysteine protease classes are indicated at the right. Sequences used are reported with their originating species followed by cathepsin class and accession numbers (gb : GenBank, sp : Swissprot).**

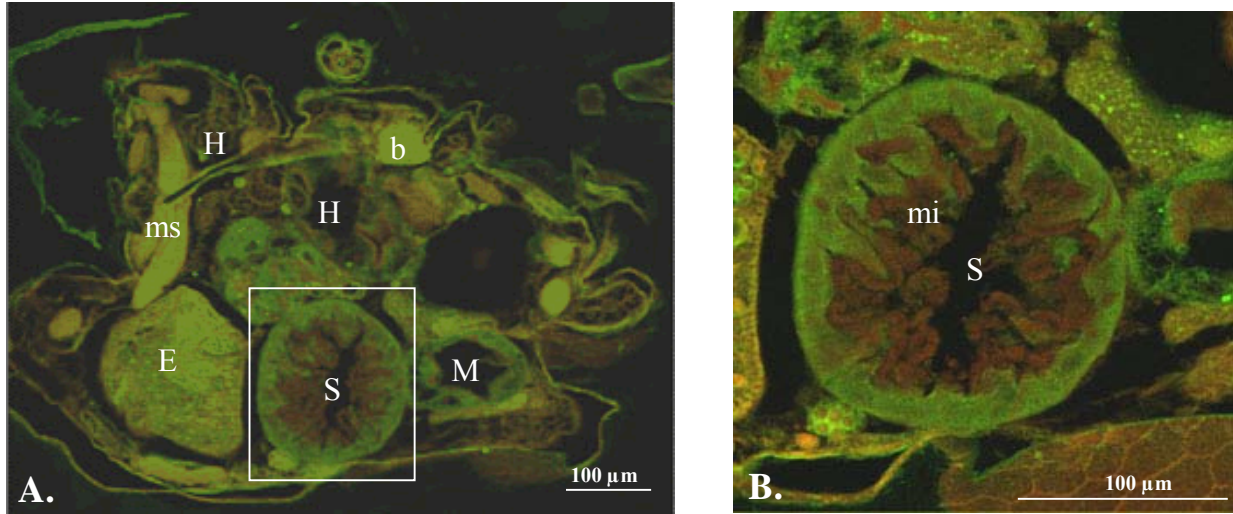


### Fig 6, Molecular analysis of AgCatL

a. Southern-blot hybridization with radiolabeled AgCatL cDNA probe (1 and 2, ten micrograms of genomic DNA digested with EcoRI and EcoRV).

b. Northern blot hybridization with radiolabeled AgCatL cDNA probe (3 and 4, total RNA isolated from carcass and guts respectively).

c. Western-blot hybridization with affinity purified antibodies against SCP (5, inclusion bodies solubilized in denaturation buffer ; 6, eluted fraction from Nickel chromatography and dialysed under acidic condition ; 7, total protein extract from one *A. gossypii* gut equivalent; 8, total protein extract from one larva of *S. zeamais*).



**Fig 7, Immunohistological localization of AgCatL in *Aphis gossypii*.**

Detection of section-bound primary reagents was accomplished using secondary immunohistochemical reagents labeled with fluorochrome (FITC-conjugated (green) anti rabbit IgG), as described in the material & methods section. A. Transverse sections of adult aphid. B. Magnification of the stomach. *stomach (S)*, *midgut (M)*, *hindgut (H)*, *bacteriocytes (b)*, *embryos (E)*, *haemolymph (H)*, *muscle (ms)*, *microvillosities (mi)*.

## **2. Résultats complémentaires**

Afin de cloner d'autres protéases digestives du puceron, une banque soustractive spécifique des messagers exprimés dans le tube digestif a été construite (Clontech).

Un certain nombre de clones (n=107) présentant un insert de taille supérieure à 500 pb ont été séquencés.

L'analyse des 107 séquences récupérées, grâce à l'algorithme blastX exécuté par le logiciel GCG (1982), révèle que 53 d'entre elles (49%) ne présentent aucune homologie (au seuil de E value =  $10^{-2}$ ) avec les séquences de la base de données GenBankP. Une grande quantité de ces séquences présente une fréquence élevée de codons stop, dans tous les différents cadres de lecture, et de répétition de motifs de polyadénylation. Elles semblent correspondre à des régions transcrites, non traduites.

Parmi les 54 séquences présentant une homologie significative, la fonction putative de 49 d'entre elles a pu être assignée (Tableau 1.1). Uniquement trois séquences indépendantes codant pour des protéases putatives ont été isolées : au moins deux gènes de cathepsine B et un gène d'aminopeptidase s'expriment dans les cellules du tube digestif. La présence dans la banque soustractive du messager codant la cathepsine L a été confirmée par PCR (résultat non présenté).

### **2.1 Analyse des séquences codant des protéases digestives putatives.**

#### **2.1.1 Cathepsine B**

L'analyse des séquences déduites des ADNc codant des endoprotéases à cystéine révèle qu'elles sont très homologues à celles des cathepsines B déjà clonées chez d'autres insectes, avec 40% d'identité en moyenne (figure 1.2). En comparaison avec les cathepsines B clonées chez d'autres organismes, la séquence du clone 2 serait représentative du propeptide complet, alors que le clone 1 serait représentatif de la séquence mature de la protéase à cystéine.

Les deux séquences qui ont été isolées présentent dans leur région commune que 42% de similitude, il semble donc que chez le puceron *A. gossypii* deux gènes de cathepsines B s'expriment dans le tube digestif.

**Tableau 1.1 : Analyse des différentes séquences obtenues par séquençage aléatoire d'une banque soustractive spécifique du tube digestif d'*A. gossypii***

Fonction Putative	Score	E value	Organisme le plus proche	Taille de la séquence cible (a.a.)	%age d'identité (a.a.)
<b>Cathepsin B (CatB 1)</b>	103	4,00E-36	<i>Sarcophaga peregrina</i>	705	80/196
<b>Cathepsin B (CatB 2)</b>	78	5,00E-18	<i>Bombyx mori</i>	337	35/92
Cytochrome p450	173	3,00E-43	<i>Myzus persicae</i>	219	85/110
Cytochrome B	42	0,002	<i>Homo sapiens</i>	263	16/36
s.u S19 ribosomiale	153	7,00E-37	<i>Spodoptera frugiperda</i>	153	72/102
s.u. 60S ribosomiale	146	1,00E-34	<i>Drosophila melanogaster</i>	204	69/109
s.u. S2 ribosomiale	199	2,00E-50	<i>Drosophila melanogaster</i>	411	91/121
S-adenosylmethionine decarboxylase	108	2,00E-23	<i>Acyrtosiphon pisum</i>	335	50/58
5'-nucleotidase	105	8,00E-23	<i>Drosophila melanogaster</i>	1095	58/136
ABC transporteur	81	9,00E-15	<i>Danio rerio</i>	1273	36/51
<b>Aminopeptidase N</b>	99	2,00E-20	<i>Plutella xylostella</i>	946	44/76
Aspartate transaminase cytosolique	163	4,00E-40	<i>Rattus norvegicus</i>	413	68/114
ATP synthase, H+ transporting, mito.	66	1,00E-10	<i>Drosophila melanogaster</i>	157	32/64
ATPase vacuolaire	144	6,00E-38	<i>Manduca sexta</i>	226	78/126
B-cell receptor-associated protein	55	0,0000004	<i>Drosophila melanogaster</i>	228	28/55
Esterase	50	0,000004	<i>Drosophila melanogaster</i>	593	20/35
Glutathione S transferase 2	236	1,00E-61	<i>Blattella germanica</i>	200	115/186
Heat shock-like protein	68	2,00E-11	<i>Ceratitis capitata</i>	653	34/34
inconnue	61	9,00E-17	<i>Ralstonia solanacea</i>	434	35/91
inconnue	67	3,00E-11	<i>Homo sapiens</i>	201	30/54
inconnue	40	0,005	<i>Drosophila melanogaster</i>	487	20/54
inconnue	41	0,00008	<i>Mus musculus</i>	212	19/43
inconnue	53	0,000001	<i>Drosophila melanogaster</i>	109	28/60
inconnue	64	6,00E-10	<i>Drosophila melanogaster</i>	391	33/66
Maltase	67	8,00E-11	<i>Drosophila melanogaster</i>	564	35/106
Malate deshydrogenase, cytoplasmique	55	0,0000005	<i>Medicago sativa</i>	332	24/37
NADH-ubiquinone oxidoreductase acyl-carrier	144	2,00E-34	<i>Drosophila melanogaster</i>	152	65/99
Nuclear Transport Factor 2	119	2,00E-26	<i>Drosophila melanogaster</i>	130	54/80
Phosphatase acide	69	3,00E-11	<i>Homo sapiens</i>	386	36/93

Fonction Putative	Score	E value	Organisme le plus proche	Taille de la séquence cible (a.a.)	%age d'identité (a.a.)
Phosphoenolpyruvate carboxykinase	87	3,00E-17	<i>Drosophila</i>	647	38/66
Phosphoserine aminotransferase	46	0,00000000	<i>Listeria innocua</i>	363	24/46
		2			
Protein p53	216	2,00E-55	<i>Drosophila melanogaster</i>	205	114/181
Ring canal protein, kel-P1	127	4,00E-29	<i>Drosophila melanogaster</i>	689	65/143
Ring canal protein, kel-P2	107	6,00E-28	<i>Drosophila melanogaster</i>	1477	50/100
<b>Serpine</b>	149	2,00E-35	<i>Drosophila melanogaster</i>	374	84/202
Thioredoxine	60	0,00000000	<i>Cyanidioschyzon merolae</i>	102	27/68
		3			
Transporteur d'UDP-galactose	49	0,000008	<i>Homo sapiens</i>	322	25/50
Transporteur de sucre	65	3,00E-10	<i>Drosophila melanogaster</i>	460	38/90
Transporteur de sulfate	80	9,00E-15	<i>Drosophila melanogaster</i>	742	43/83
Transporteur de vitamine C sodium dépendant	111	1,00E-24	<i>Drosophila melanogaster</i>	573	52/70
UDP-glucuronosyltransferase	59	0,00000000	<i>Drosophila melanogaster</i>	480	39/133
		4			
UDP-glycosyltransferase	214	5,00E-55	<i>Drosophila melanogaster</i>	559	109/215

Fonction putative de la séquence déduite de l'annotation du meilleur hit donné par BlastX. Score et E.value relatif au meilleur hit donné par blatsX

10 20 30 40 50 60 70 80 90

*Ceanorhabditis e.*-----MKYLILAAALVAVTAGLVIPLVPKTQEAITERYVNSKQSLWKAIEIPKDITIEQVKKRLMRTEFVAPHTPDVEVVKHDINEDT-----  
*Homo sap.* MWQLWASLCLLVLAN---ARSRPSFHPVSDLVNIVNKRNTTWQAGHNF-YNVDMSYLKRLCGTFLGGPK---PPQVRMFTEDL-----  
*Sarcophaga p.* MRQHFVIIICIAFLAFGQVLANLDAENDLLSDEFLEIVRSKAKTWTVPGRNYDKSVPRSHFRRLMGVHPDAHKFTLHEKSLVLGEEVGLADS  
*Drosophila m.* -----MNLVVATAASVAALTSGEPSLLSDEFIEVRSKAKTWTVGRNFDASVTEGHIRRLMGVHPDAHKFALPDKREVLGDLYVNSVD  
**Aphis g. 1** -----  
**Aphis g.** -----SVVAAEVHLCCLKSIVLYTIAFRNTHG---RLEISIQVFLMFHVSWEFYQGINVKRTLYLHTIHPAGS---E

X1

100 110 120 130 140 150 160 170 180

*Ceanorhabditis e.* -IPATFDARTQWPNCMSINNIRDQSDCGSCWAFAAEAASDRFCIASNGAVNTLLSAEDVLSCCSN-CGYGCEGGYPINAWKYLKVS GFC  
*Homo sap.* KLPASFDAREQWPQCPTIKEIRDQSGCGSCWAFGAVEAISDRICIHNTAHVSVEVSAEDLLTCCGSMCGDGCNGGYPAEAWNFWTRKGLV  
*Sarcophaga p.* DVPEEFDARKAWPNCPTIGEIRDQSGCGSCWAFGAVEAMSDRLCIHSNATIHFHFSADDLVSCCHT-CGFGCNGGFPGAAWAYWTRKGI  
*Drosophila m.* ELPEEFDARKAWPNCPTIGEIRDQSGCGSCWAFGAVEAMSDRVCIHSGGKVNHFHFSADDLVSCCHT-CGFGCNGGFPGAAWSYWTRKGI  
**Aphis g. 1** -LPINF DARKRWPNCPSIGHIYNQGNCRSCYAVISASAVTDRICIHNETKNPIMSAQQIISCCYL-CGYGCDGGSQFESWDFYRRHGFV  
**Aphis g. 2** PLPESYDVTQTWSECKSVVSIRDQSNCGSCWALSTASAFSGRLCIASNMDFNIVLSGEYINS CCNGKCGDGCNGGHPEKAWKYIKKNGLC  
+1

190 200 210 220 230 240 250 260 270

*Ceanorhabditis e.* TGGSYEAQFGCKPYS LAPCGETVGNVTWPSPD-DGYDTPACVNKCTNKNYNAVAYTADKHF GSTAYAVGKKV S QIQAEIIAHGPVEAAFT  
*Homo sap.* SGGLYESHVGRPYSIP EHHV-NGSRPPCTG-E-GDTPKCSKIC-EPGYSPTYKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFTS  
*Sarcophaga p.* SGGPYGSSQGCRPYEIA EHHV-NGTRPPCDG-EHGKTPSCRHEC-QKSYDVDYKTDKHF GSKSY SVKRNVKDIQKEIMQNGPVEGAFT  
*Drosophila m.* SGGPYGSNQGCRPYEIS EHHV-NGTRPPC-A-HGGRTPKCSHVC-QSGYTVDYAKDKHF GSKSY SVRRNVREIQEEIMTNGPVEGAFT  
**Aphis g. 1** SGGDYNSNQGCQPYMIPPCKLINENSPRHSC TTYNREETPACEIKCNNPNYSSFKTDIYKG-KYYQVYPFMA--MKEIFDNGPITTFY  
**Aphis g. 2** TGGEYNSNEGCPYSIFPCPRNSNSCSK-----ENEDTPQCYKNQ-----

250 260 270 280 290 300

*Ceanorhabditis e.* VYEDFYQYKTGVYVHTTGQELGGHAIRILGWGT--DNGTPYWLVA NSWNVNWGENGYFRIIRGTNECGIEHAVVGGVPKV-----  
*Homo sap.* VYSDFLLYKSGVYQHVTGEMMGGHHAIRILGWGV--ENGTPYWLVA NSWNTDWDGNGFFKILRGQDHCIESEVVAGIPRTDQYWEKI  
*Sarcophaga p.* VYEDLILYKDG VYQHVGRELGGHAIRILGWGV--ENKTPYWLVA NSWNTDWDGNGFFKMLRGEDHCIE SAIAAGLPKV-----  
*Drosophila m.* VYEDLILYKDG VYQHEHGKELGGHAIRILGWGVWEEKIPYWLIGNSWNTDWDGHDGFFRILRGQDHCIESSISAGLPKL-----  
**Aphis g. 1** MYRDLIDYKSGVYQYRRG-----HILRMIS SLYKREDYRLGRKMNGWLV SQFIWYRLGRIRGPX-----  
**Aphis g. 2** -----

**Figure 1.2 : Comparaison des séquences protéiques déduites des cathepsines B d'*A. gossypii* avec quelques-unes des cathepsines B connues chez d'autres organismes**

(numéro d'accèsion Swiss Prot (sp) ou GenBank (gb), *Caenorhabditis e.* gb : NP504682, *Homo sap.* : sp P07858, *Sarcophaga p.* gb : BAA041103, *Drosophila m.* sp : Q9VY87)

Des gaps ont été introduits par le programme Clustal W afin de maximiser l'alignement. Les résidues du site active Cys, His et Asn sont surlignés en jaune. Le site de clivage du peptide signal est noté x1, le site de clivage du propeptide est surligné en gris, la boucle occlusive d'un trait (-), le site potentiel de glycosylation est surligné en rose, les acides aminés responsables de l'activité exopeptidasiques (His-His) de la boucle occlusive sont surlignés en rouge.



```

           550       560       570       580       590       600       610       620       630       640
Manduca s.  WVISSNVNFSDTSPQAWILPTFPATAVDVPGLSNADWYIFNKKQTGYRVNYDVENWVALARVLNNSHEIIVHVLNRAQIVDDAFNLRNGRLHYKNAFEI
Bombyx m.   WVLSSNIDFSDTKPGQWIPPSFPATSIDIPGLANAWEWFIENKQQTGYRVNYDPENWAALAKILQTNHAVIHLLNRAQILDDSFNMARNGRLLNYNLPFEI
Plutella x. WVLSTDVNFNDTRPMAWLPPQLAAEAVQVPGLQNAWEWFIENKQQTGYRVNYDPENWRALAKVLDNTHHEIIVHLLNRAQLIDDSFNLRNGRLDYSLAFDL
Aphis g.   -----VVISVAPRYIFNIQSTGYRVNYTEENWLSLIKQLNDSHKDVIHVLNRAQLIDDSFSLAKVGLLNYTIPMSM

           650       660       670       680       690       700       710       720       730       740
Manduca s.  SRYLEMEKDYIPWAAAANPAFNLYLDIVLSGANSYNLYRYLLNLTA PMFEDLGFDFVKS GEEFVTPYHRNIILDINCRFGNQRCISRQEIILQAFKNNPNQR
Bombyx m.   SRYLIN EKDYIPWAAINPAFNLYLDIVLTGSSVYNLFREYLLTLTAPLYDEIGWEATANEEHV MAYHRNIILDINCR LGNQRCVTRAQELLEQFRNPTQR
Plutella x. SRYL VQERDYIPWAAAANA AFNLYLNSVLSGSSVHPLFQEYLLFLTAPLYQRLGFNAATGEEHVTPFHRNII LNINCLHGNE DCVSTAETLLQNF RDNPTQT
Aphis g.   AEYLLKKE DDIIPWYRRRCYTRKLWNERSSSIAVGRKHC IWWSR KSDYFWRKLRSEYWL SHVIANERPISQSHTPKWYLPGGR-----

           750       760       770       780
Manduca s.  PNPDIQTLVYCS SLRAGNVENFNFLWNMYLGTSDSSEQSI
Bombyx m.   LNPDLQNTVYCSGLRGGDRDNFNFLWEQYLATSDSSGQNI
Plutella x. LNPDIQTTFVCSGLRGGDVDFNFNLWARYTATQDSSEQSI
Aphis g.   -----

```

**Figure 1.3 : Comparaisons de la séquence déduite de l'Aminopeptidase-N d'*A. gossypii* avec quelques-unes des aminopeptidases-N d'insectes.**

Numéro d'accèsion GenBank, *Manduca sexta* : X97877, *Bombyx mori* BAA 32140, *Plutella xylostella* : X97878.  
 Les résidus impliqués dans le site actif sont surlignés en jaune.

L'alignement de CatB 2 avec les autres protéases à cystéine suggère la présence d'une prorégion de 66 résidus aminés (la séquence serait incomplète au niveau du peptide signal). L'enzyme mature serait une protéine de 196 acides aminés. Les résidus de la triade catalytique Cys<sup>+25</sup>, His<sup>+159</sup> et Asn<sup>+175</sup> sont conservés dans la séquence du puceron.

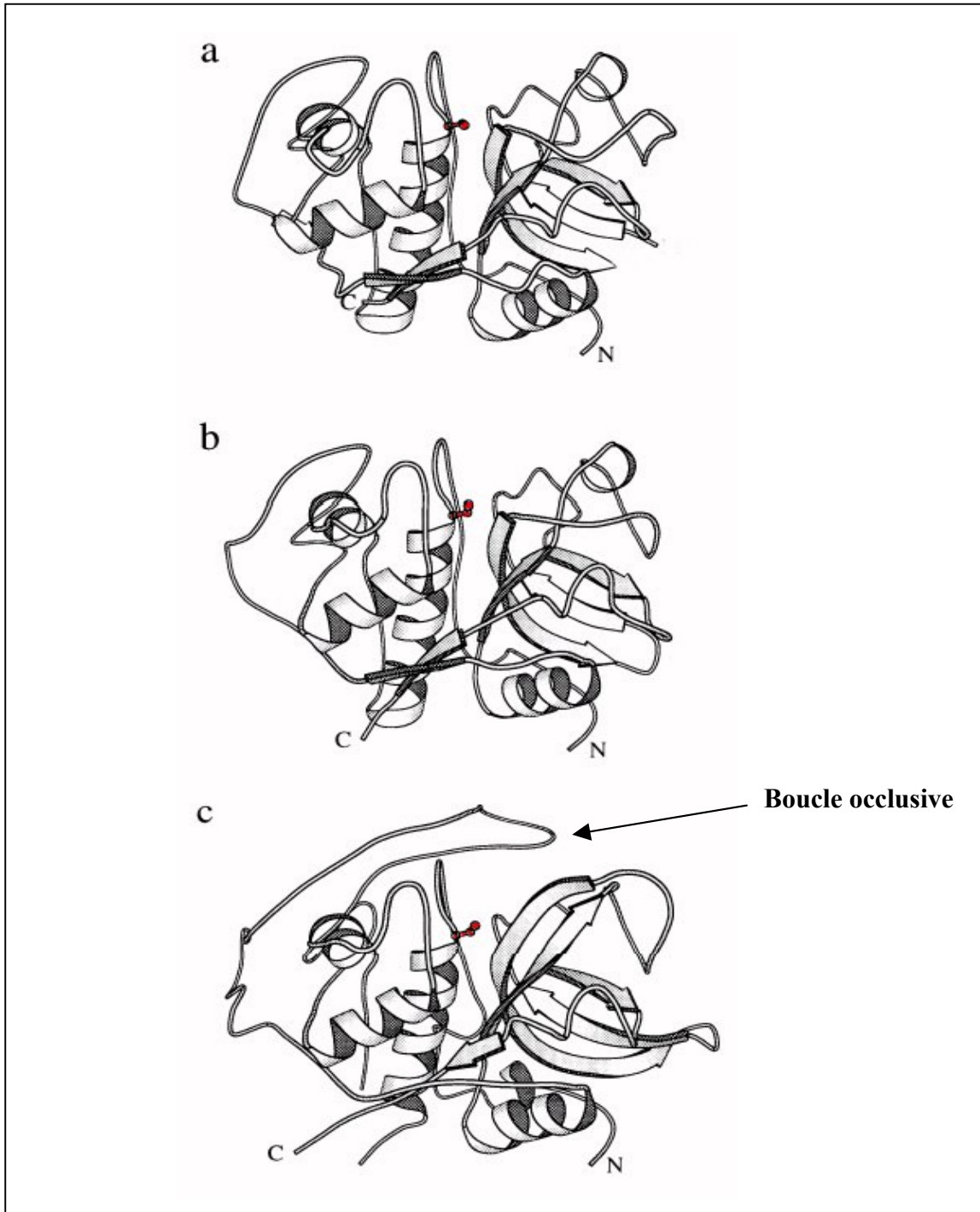
Les protéines déduites possèdent un motif de N-glycosylation dans la région mature de l'enzyme. Ce site est présent chez les cathepsines B de mammifères, et les cathepsines B de la drosophile et de *Sarcophaga*. Chez les mammifères, l'adressage des protéases aux lysosomes repose sur la reconnaissance d'un signal constitué d'un oligosaccharide N lié (avec un résidu mannose terminal) et d'un motif structural impliquant des lysines, présent dans le propeptide (Cuozo, 1995). Grâce à ce motif, la GlcNAc phosphotransférase reconnaît les protéases lysosomiales dans l'appareil de Golgi et transfère les résidus GlcNAc-P en position 6 de plusieurs résidus mannose sur les oligosaccharides N-liés à la protéine. Ensuite la phosphoglycosidase élimine le résidu GlcNAc par clivage, créant le marqueur du mannose 6-phosphate. Ces motifs sont reconnus par des transporteurs à mannose 6-phosphate qui permettent l'acheminement de la protéase dans les lysosomes. Les cathepsines B du puceron possèderaient les signaux permettant une localisation lysosomiale.

### **2.1.2 Aminopeptidase N**

La séquence protéique déduite correspond à une aminopeptidase N. Bien que la séquence soit partielle, elle présente une homologie forte (58%) avec les séquences d'aminopeptidases clonées chez certains lépidoptères (figure 1.3). Le site actif, basé sur le motif d'acides aminés HExxH, caractéristique des peptidases, est présent dans la séquence partielle clonée.

## **2.2 Dosages enzymatiques**

Bien que la structure globale des cathepsines B soit identique à la papaïne ou aux cathepsines L, elle possède une longue insertion au voisinage du résidu 95 (numérotation selon la papaïne) ; cette région est appelée boucle occlusive car elle bloque l'accès du substrat dans site actif (figure 1.4). Les cathepsines B possèdent une activité endo et exopeptidasique. La délétion de la boucle occlusive par mutagenèse dirigée entraîne la perte de l'activité carboxypeptidasique. La présence de deux résidus Histidine adjacents dans ce domaine est responsable de cette activité (Rawlings, 1998).



**Figure 1.4 : Structure globale de cathepsines (d'après Fujishima, *et al.* 1997)**

Cathepsine L humaine (a), papaïne (b), cathepsine B humaine (c). N et C représentent les extrémités amino- et carboxy-terminales. Les cystéines catalytiques sont représentées selon le modèle "ball-and-stick".

La spécificité des cathepsines B repose sur le sous site S2 du site catalytique (Glu 245/Ser chez les autres protéases à cystéine), qui lui permet d'accueillir des acides aminés très larges tels que l'arginine.

Grâce à l'utilisation de substrats et d'inhibiteurs irréversibles spécifiques, l'activité relative des différentes cathepsines présentes dans un tissu peut être quantifiée. Toutefois, au vu des différentes affinités que présentent les cathepsines pour un substrat donné, nous avons fait le choix de mettre en évidence ces différentes activités avec un seul substrat et différents inhibiteurs, jugés plus spécifiques (Inubushi, 1994). En effet, si le substrat chromogène Z-Arg-Arg-pNA est donné comme assez spécifique des cathepsines B, le substrat Z-Phe-Arg-pNA est hydrolysé par les cathepsines L, B, J et S (ainsi que par certaines trypsines) ; et le substrat Arg-PNA est quant à lui plus spécifique des aminopeptidases (dont l'activité exopeptidases des cathepsines H) (Rawlings, 1998).

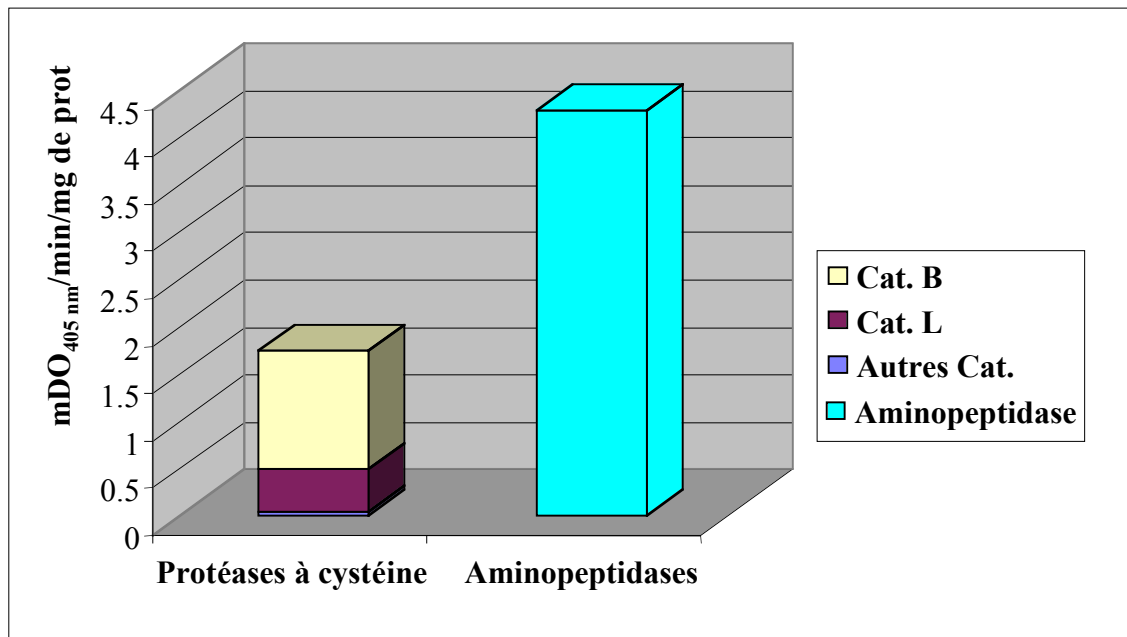
L'activité aminopeptidase a également été quantifiée par la même méthode. Cette activité est presque 3 fois plus importante que l'activité protéase à cystéine (figure 1.5).

Avec l'addition de l'inhibiteur CA074 spécifique des cathepsines B (Murata, 1991) dans le mélange réactionnel, l'activité relative aux différentes cathepsines a pu être quantifiée. L'activité cathepsine B représente 70% de l'activité protéases à cystéine du tube digestif. L'activité cathepsine L constitue 27% de l'activité totale et les 3% restant sont représentatifs d'autres activités cathepsines, relatives à des activités non inhibées par l'E64 (figure 1.5). L'activité cathepsine B semble donc l'activité protéase à cystéine majoritaire dans le tube digestif du puceron.

### **2. 3. Discussion**

Le séquençage aléatoire de la banque soustractive, spécifique des ARNm présents dans le tube digestif d'*A. gossypii*, nous permet d'avoir une vision plus globale de la physiologie digestive du puceron, même si elle reste très partielle par l'ampleur modeste du séquençage systématique.

Le puceron se nourrit dans le phloème et prélève une alimentation essentiellement composée de sucres. Un grand nombre de séquences isolées présentent des homologies avec des protéines impliquées dans le transport et l'utilisation des sucres simples (31% des séquences identifiées). Aucune enzyme putative de la dégradation des polysaccharides n'a été isolée (pectines, amidon, cellulose), confirmant la particularité du puceron dans le groupe des



**Figure 1.5 : Quantification des différentes activités protéasiques du tube digestif d'*A. gossypii***

Afin de mesurer l'activité des différentes protéases à cystéines (cathepsines), les protéines d'un tube digestif sont incubées avec 10  $\mu\text{M}$  d'inhibiteur (CA074 ou E64) pendant 15 min à température ambiante, puis la réaction est débutée par l'addition de 200 $\mu\text{M}$  du substrat Z-Phe-Arg-pNA dans un tampon phosphate 0,1M à pH 5 à 37°C.

L'activité enzymatique des aminopeptidases-N est mesurée par l'hydrolyse du substrat Arg-pNA par les protéines d'un tube digestif dans les mêmes conditions.

insectes phytophages. La situation se différencie clairement à cet égard d'insectes s'alimentant sur feuilles entières, et étudiées de façon analogue, comme la phédonne du cresson (Girard, 1999).

Le nombre de séquences possédant une homologie avec des protéases ne représente que 4,3% des séquences obtenues. Cette proportion est beaucoup plus faible que les 20% de protéases clonées à partir du séquençage aléatoire d'une banque d'ADNc de tube digestif d'*Helicoverpa armigera* (Bown, 1997) ou même des 6% chez la phédonne du cresson (Girard, 1999). Cette différence est renforcée par le fait que, d'une part, cette analyse est menée sur une banque soustractive (processus qui enrichit la banque en ADNc exprimés spécifiquement dans le tube digestif)<sup>14</sup>. D'autre part, uniquement deux classes de protéases ont été caractérisées.

Un ADNc partiel présente une homologie forte dans la région du site actif avec les aminopeptidases clonées chez les lépidoptères, suggérant une faible divergence avec l'ancêtre commun. Chez ces insectes, l'aminopeptidase présente un motif d'ancrage GPI (glycosylphosphatidylinositol) à la membrane, situé dans la région C-terminale (Denolf, 1997 ; Hua, 1998 ; Emmerling, 2001). Il a pu être montré que cette enzyme était attachée à la membrane plasmique et possédait un signal de sécrétion, présentant ainsi son site actif vers la lumière du tube digestif (Denolf, 1997). Cette enzyme est impliquée dans l'étape ultime de digestion des protéines, les exopeptidases produisant des résidus aminés simples. Les pucerons utiliseraient cette activité afin de dégrader soit les peptides ingérés<sup>15</sup>, soit les peptides résultant de l'hydrolyse terminale de protéines alimentaires, si cette dernière a bien lieu dans des proportions significatives.

Par ailleurs, deux ADNc partiels codant pour deux cathepsines B différentes, ont été isolés. Lors de l'analyse phylogénétique des endoprotéases à cystéine, toutes les séquences de protéases à cystéines d'insecte ont été prises en compte afin d'établir des relations d'orthologie avec les différents sous-groupes des cathepsines de vertébrés. Un groupe de séquences d'insectes appartenant à différents ordres se place dans le sous-groupe des cathepsines B, suggérant que le puceron puisse posséder également un gène de cathepsine B orthologue de ce groupe de gènes. Le séquençage aléatoire de la banque soustractive a permis d'isoler deux séquences d'endoprotéases à cystéine putatives, l'analyse de leurs séquences

---

<sup>14</sup> Initialement, un séquençage aléatoire direct de la banque d'ADNc de tube digestif d'*A. gossypii* ne nous avait permis d'isoler que des séquences présentant des homologies avec des gènes de ménage.

<sup>15</sup> Ce compartiment alimentaire reste à l'heure actuelle totalement inconnu chez les pucerons, et seule l'approche peptidomique initiée récemment sur du phloème de Cucurbitaceae (Haebel, 2000) pourra en révéler correctement l'importance quantitative et qualitative.

révèle qu'elles appartiennent au groupe des cathepsine B confirmant ainsi l'hypothèse formulée d'après l'analyse phylogénétique. La cathepsine L isolée précédemment (et identifiée par PCR dans la banque soustractive) constitue une troisième protéase à cystéine identifiée chez les pucerons. Ces séquences présentent un motif de N-glycosylation qui représentent un signal potentiel d'adressage au lysosome.

Aucune séquence homologue à des protéases à sérine, ni à des protéases acides n'a été isolée. Ce résultat confirme la particularité du puceron, l'activité endoprotéolytique du tube digestif étant constituée à plus de 95% de protéases à cystéine appartenant au sous groupe des cathepsines (article 1).

Ces résultats soutiennent l'hypothèse que nous avons déjà formulée lors de l'étude de la cathepsine L, et les pucerons pourraient utiliser des protéases lysosomiales afin de dégrader les protéines phloémiennes ingérées.

**CHAPITRE II**

**–UTILISATION & DIVERSIFICATION**

**DES**

**INHIBITEURS DE PROTEASES A CYSTEINE–**



## 1. Contexte de l'étude

Depuis le développement des techniques de transfert de gènes, l'INRA a développé des programmes de défense des plantes contre les insectes phytophages. Des travaux ont été réalisés sur différents modèles végétaux : le tabac (*Nicotiana tabacum* L.), le chou (*Brassica oleracea* L.), le cotonnier (*Gossypium hirsutum*), le peuplier et le colza. Les insectes cibles étaient des lépidoptères comme *S. littoralis* pour le cotonnier et le tabac et *M. brassicae* pour le chou, et des coléoptères tel que *C. tremulae* pour le peuplier et les charançons des siliques pour le colza. Les gènes codant différentes protéines entomotoxiques, d'origine bactérienne (Bt) ou végétale (inhibiteurs de protéases), ont été utilisés. Plus récemment, un projet a également été initié pour tester la faisabilité de la création de plantes transgéniques résistantes aux pucerons. Les travaux visant à utiliser la transgénèse dans la lutte contre les pucerons sont relativement récents et le facteur limitant reste le choix des gènes candidats. Un programme de criblage de protéines végétales possédant des propriétés aphiditoxiques ont permis d'identifier de nouveaux candidats tel que des lectines ou des inhibiteurs de protéases (Rahbé 1993 ; Rahbé, 1995).

Cette étude présente la mise en évidence des effets entomotoxiques de l'Oryzacystatine-I (OC-I) sur les pucerons. L'oryzacystatine-I est un inhibiteur de protéase à cystéine qui a été mis en évidence dans les graines immatures de riz (Abe., 1987). La disponibilité, au laboratoire, de plantes de colza transgéniques exprimant constitutivement cet IP à cystéine (Jouanin, 2000) a permis d'envisager d'évaluer la possibilité d'obtenir un gain de résistance contre le puceron *Myzus persicae*. Une diminution du poids des adultes et de la fécondité est observée chez les pucerons se nourrissant sur plantes transgéniques.

Ce travail est à l'origine de la collaboration entre le laboratoire de Biologie Cellulaire de l'INRA de Versailles et le laboratoire de Biologie Fonctionnelle, Insectes et Interactions de l'INRA de Lyon. Ce travail avait été initié avant mon arrivée au laboratoire. Dans le cadre de cette étude, j'ai participé à la caractérisation des cibles de l'IP et la compréhension de son mécanisme de toxicité. Cette étude a constitué la base de mon projet de thèse car il avait été démontré la faisabilité d'exprimer des inhibiteurs de protéases dans le phloème, et d'accroître ainsi la résistance des plantes aux pucerons.

Article 2 : Yvan Rahbé, Céline Deraison, Michel Bonadé-Bottino, Cécile Girard, Christiane Nardon, Lise Jouanin, soumis à *Plant Science*.

**Effects of the cysteine protease inhibitor oryzacystatin (OC-I) on different aphids and reduced performance of *Myzus persicae* on OC-I expressing transgenic oilseed rape.**

## Article 2

# Effects of the cysteine protease inhibitor oryzacystatin (OC-I) on different aphids and reduced performance of *Myzus persicae* on OC-I expressing transgenic oilseed rape

**YVAN RAHBE (1), CELINE DERAISON (1,2), MICHEL BONADE-BOTTINO (2), CECILE GIRARD (2), CHRISTIANE NARDON (1) AND LISE JOUANIN (2)**

(1) INRA- INSA, UMR BF2I Biologie Fonctionnelle Insectes et Interactions, 20 av. A. Einstein, 69621 Villeurbanne cedex, France.

(2) INRA-CNRS, Laboratoire de Biologie Cellulaire, route de St Cyr, 78026 Versailles Cedex, France.

**CORRESPONDING AUTHOR :**

**YVAN RAHBE**

INRA-INSA, UMR BF2I Biologie Fonctionnelle Insectes et Interactions 406, Bat Louis-Pasteur, 20 av. A. Einstein, 69621 Villeurbanne, France.

tel : +33-4 72 43 84 76

fax : +33-4 72 43 85 34

e-mail : rahbe@jouy.inra.fr

*Running title:* Effect on aphids of transgenic oilseed rape expressing oryzacystatin

## Summary

*When administered in artificial diets, the cysteine protease inhibitor oryzacystatin I (OC-I) induced only moderate though significant growth inhibition on the pea aphid (*Acyrtosiphon pisum* Harris), the cotton/melon aphid (*Aphis Gossypii* Glover) and the peach potato aphid (*Myzus persicae* Sulzer). In transgenic oilseed rape plants (*Brassica napus* L. cv. Drakkar) expressing the OC-I sequence under control of a modified cauliflower mosaic virus 35S RNA promoter, OC-I was detected in both leaves and phloem sap. Transgenic plants from three independent homozygous lines were used to test the effect of in planta-expression of OC-I on *M. persicae*. Mean adult weight and fecundity were significantly reduced, and aphid biomass produced in two weeks was decreased by 25 to 40 % for aphids fed transgenic plants, when compared to those fed control plants. The effects of OC-I on *M. persicae* were correlated with the decrease of a major cathepsin L/H-type cysteine protease activity, detected in whole insects extracts. Immuno-histological analysis showed OC-I labeling along the gut epithelium, but also its association with aphid oenocytes and bacteriocytes. These results suggest that OC-I affect *M. persicae* by reaching the haemolymph, thereby inhibiting extra-digestive proteolytic activities and interacting with functions related to aphid reproduction. Therefore, it appears that protease inhibitors can display deleterious effects against phloem-feeding insects, although through different mechanisms, in addition to their activity on leaf-feeding insects.*

**Key-Words :** Homoptera, Aphididae, host plant resistance to insects, phytocystatin, insecticidal protease inhibitors, peptidases

## 1. Introduction

Availability of genetic resources for plant resistance to pests and pathogens is an important component of integrated control strategies. This component is even more important with pests for which few pesticide classes are effective due to biological peculiarities, such as specialised feeding habits (stem/pod-borers or phloem-feeding insects). Oilseed rape, as other *Brassica* crops, hosts such pest species: the cabbage seedpod weevil *Ceutorhynchus assimilis* (Paykull), or the aphids *Myzus persicae* (Sulzer), *Lipaphis erysimi* (Kaltenbach) or *Brevicoryne brassicae* (Linnaeus), which are respectively the peach-potato aphid, the turnip aphid and the cabbage aphid [1, 2]. Due to the fast evolution of cropping systems, to the genetic plasticity of economically important *Brassica* species, and to the cost of extensive germplasm testing for insect resistance, very few natural resistance sources are available for aphid control in this botanical group [3, 4, 2]. In such a situation, genetic transformation of crops with genes conferring resistance is an attractive alternative.

To date, strategies aimed at improving plant resistance against insect pests through gene transfer have relied on a limited set of genes [5], most of which either derived from the toxin of the soil bacterium *Bacillus thuringiensis* (*Bt*) or were protease inhibitors (PIs). In addition to the lack of *Bt* toxins active against oilseed rape pests, diversifying the panel of available genes is highly desirable. PI-based strategies targeted pests, mainly leaf-feeding insects, using digestive proteases to degrade ingested proteins. On the contrary, phloem-feeding insects do not rely extensively on protein digestion to fulfill their nutritional needs, due to the abundance of free amino acids in their natural food, and to the presence of symbionts supplying them with essential amino-acids [6]. They are insensitive or moderately sensitive to many peptidic protease inhibitors [7, 8]. Therefore, so far, strategies aimed against sap-sucking insects have been exclusively based on lectins, mainly from the mannose or glucose-mannose specificity group [9]. Surprisingly, the pea aphid was found to be quite sensitive to one subclass of

inhibitors of the serine Bowman-Birk family [7], especially those targeting chymotrypsin (Quillien and Rahbé, unpublished). This suggested that a defense strategy against aphids may benefit from the use of protease inhibitor-encoding genes like those used so far against pests using digestive proteases.

The cysteine protease inhibitor oryzacystatin I (OC-I), characterized and cloned from rice [10, 11], has been used successfully in plant protection against pests possessing cysteine proteases, either coleopteran insects [12, 13] or nematodes [14]. As such, it was chosen to test its potential effects against several aphids. A first series of tests examined the influence of purified OC-I on aphid viability and growth when incorporated to artificial diets. Then, the use of OC-I for oilseed rape protection against *M. persicae* was examined. The expression of OC-I in phloem was monitored with the corresponding coding sequence placed under control of a promoter derived from the constitutive 35S CaMV promoter. The question of expression levels of OC-I in the phloem appeared to be critical since it has been recently reported that very different effects can be observed on target insects depending on levels of expression of PIs in plants [15]. Also, we searched for possible adaptation mechanisms in the aphid since it was recently shown that different adaptive mechanisms were used by crucifer-eating insects displaying cysteine digestive proteinases in order to counteract the effects of ingesting OC-I [16, 17]. Finally, we conducted a complimentary immunohistological study on *M. persicae* ingesting OC-I in order to get information on its localization in the aphid, as a first step in elucidating its physiological targets and mode of action.

## 2. Methods

### 2.1. Insects

The *M. persicae* clone used in this study originated from a greenhouse sample collected in Lyon on tobacco in 1994, has been reared on oilseed rape since then, and is recorded as Mp-L04-Bo. Neonate aphids (aged 0-18h) were used for all the artificial diet and plant assay experiments. For protease and OC-I detection assays, aphids fed on control or transgenic plants were from another *M. persicae* clone, collected on oilseed rape and kindly provided by Dr G. Biache (Laboratoire de lutte biologique, INRA Guyancourt, France).

The pea aphid *Acyrtosiphon pisum* Harris (clone Ap-LL01-Vf) and two clones of the cotton/melon aphid *Aphis gossypii* Glover (clones Ag-LM02-Cm and Ag-LB04-Gh) were from laboratory cultures and reared on broad bean, melon and cotton plants respectively.

### 2.2. Plants

OC-I-expressing oilseed rape plants from the spring cv. Drakkar (Serasem) were obtained through *Agrobacterium tumefaciens* gene transfer using a pKYΩ'OC-I/GUS plasmid vector [18]. This vector is a modification of pKYΩ'OC-I, described in Leplé *et al.* [12], with an intron-containing-β-glucuronidase sequence, under control of the cauliflower mosaic virus 35S RNA promoter [19], cloned between the *nptII* (kanamycin resistance) and the *OC-I* genes [18]. Homozygous plants, obtained through three generations of self-pollination, and containing one insertion site of the corresponding T-DNA, were compared on the basis of their foliar OC-I expression level. Lines 4B, 16B and 29A were used for the aphid bioassays. Line 4B, the highest OC-I leaf expressing line was retained for the phloem-expression and enzyme control experiments. Non-transformed Drakkar plants were used as control.

### 2.3. Proteins and chemicals

OC-I protein was purified from the supernatant of an *E. coli* expression system [12]. All chemicals used for artificial diets and enzymatic analysis were purchased from either Fluka or Sigma (St-Quentin Fallavier, France). Chemical abbreviations used are as follows: BApNA = N $\alpha$ -benzoyl-L-arginine p-nitroanilide, BTpNA = N $\alpha$ -benzoyl-L-tyrosine p-nitroanilide, L-Leu-pNA = L-leucine p-nitroanilide, L-Arg pNA = L-arginine p-nitroanilide, PMSF = Phenyl methyl sulfonyl fluoride and E64 = L-3-trans-carboxiran-2-carbonyl L-leucyl-gmatin.

### 2.4. Bio-assays

Toxicity assays followed the standard procedure described previously [7]: neonate aphids (0-18h) were given access for 7 days to artificial diets, mortality was checked daily, and adult aphids were weighed at the end of experiment. Diets (with 0, 10, 50 or 250  $\mu$ g/ml purified OC-I) were changed twice and environmental conditions were the following (as for the plant bioassays): temperature 21°C, r.h. 70%, photoperiod 16L/8D.

Plant evaluation assays were conducted in three independent experiments involving 3 (*exp. 1 and 2*) or 2 transgenic lines (*exp. 3*), plus control cv. Drakkar. Transgenic lines used were lines 4B, 16B and 29A, obtained as described above (in *exp. 3*, only lines 4B and 16B were used). For all experiments, plants were grown under controlled conditions (21 $\pm$ 1°C, r.h. 60-70%, photoperiod 16L/8D, under 3400-4400 Lux of hort-gro fluorescent light), and used at stage 5-7 fully expanded leaves. Plants were watered daily and fertilised twice a week with a NPK 20/20/20 solution. For all experiments (except end of *exp. 3*), plants were caged in plexiglas cylinders (diam. 25 cm, with muslin top) and aphids were allowed to move freely within plants, although most of them remained near their initial feeding site. For *exp. 1*, 15 neonate aphids were deposited on a top leaf (n=5 per genotype), and aphid mortality was recorded twice a week. Adult as well as larval counts and weighing (global weights per plant, adults and larvae separately) were performed on each plant fourteen days after infestation.

The same protocol was used for *exp. 2*, with the following modifications: 10 plants were used per genotype, aphids were deposited per batches of 5 on paper discs inserted on the plant stem, below leaf 4-5, and plant position was randomized on shelves; experiment was stopped at day 12 and adult aphids were weighed individually, while progeny was counted and weighed globally per plant. For *exp. 3*, 10 plants were used per selected genotype, and aphids were laid directly on leaves to avoid handling (from 3 mothers per plant; 15/22 larvae per plant). On day 7, cages were removed, adult aphids were weighed individually, except 3 large individuals that were left on separate leaves (labeled). Daily fecundity was scored on these aphids from day 8 to 15.

For phloem sap sampling, aphid breeding for honeydew collection and protease analysis, transgenic 4B and control plants were sown and grown simultaneously in a greenhouse under natural light, with an approximate photoperiod of 15h/9h day/night; temperature and relative humidity varied between 30°C/50% and 20°C/70% respectively (day/night). Plants were used when they had developed 9 true leaves, just before flowering.

## **2.5. Plant sampling**

Phloem sap exudation was performed on mature leaves located at mid-stem (fifth leaf) taken from 4 individual plants per line tested. Leaves were severed from plants at the same time in the morning, and treated individually. Leaf petioles were cut, immediately plunged in extraction buffer (5mM phosphate - pH 7.0, 5mM EDTA), and cut again in the buffer to prevent air contact with the phloem. Petioles were placed in tubes containing 1.5 ml extraction buffer, subsequently sealed with parafilm. All leaves were then placed under high moisture conditions, at 20°C under illumination for 9 hours. Sap-containing extraction buffer was aliquoted and kept at -80°C until needed.

Protein extraction was performed on the leaves used for phloem collection. Leaves were grinded under liquid nitrogen, then 2ml extraction buffer and 250mg Polyclar (Serva) were



added per g of tissue. The mix was allowed to thaw slowly at 4°C and sustained a 12,000 g centrifuge at 4°C for 10 min. The supernatant was heated for 10 min at 65°C then centrifuged again. Protein quantification on sap and leaf extracts (with and without heating) were performed according to a Bradford assay using bovine serum albumin (BSA) as control.

Honeydew was collected by placing an aluminum sheet under aphid colonies for 24 h. Sheets were then washed in extraction buffer and proteins assayed before use.

## **2.6. Electrophoresis and blots**

Western blot hybridization were performed on leaf, phloem and honeydew extracts, using *E. coli*-purified OC-I [12] as control. Heat-denatured proteins were run on 20% acrylamide SDS-PAGE gels, then immobilized on a nitrocellulose membrane by electro-transfer and hybridised with anti-OC-I rabbit anti-bodies [12]. Presence of anti-OC-I antibodies was revealed through ECL following hybridization with a commercial anti rabbit Ig-G antibody from goat coupled with peroxidase (Sigma).

## **2.7. Protease assays**

Adult aphids (from control or transgenic line 4B) were collected and ground in cold 0.15 M NaCl, then centrifuged at 16,500 g (4°C, 5 min). Supernatants were kept either 2/3 days at 4°C or stored at -20°C for later use. For each test, 40-aphid batches were used (n=3) and total protein was measured. For proteolytic pH-profiles, total caseinolytic activity was measured at pH ranging from 5 to 12 as described in Leplé et al. [12], without or with  $\beta$ -mercaptoethanol ( $\beta$ ME, 5 mM final). Comparison of total proteolytic activity between aphids fed on control and 4B plants was conducted at pH 6 in 100mM sodium phosphate buffer containing 5 mM  $\beta$ ME. Specific amidolytic activities were measured with p-nitroanilide substrates (1 mM final) in 100 mM sodium phosphate buffer, pH 8.0 with 50 $\mu$ g protein equiv. of extract. Micro-titer plates were used to follow chromophore release at 405 nm, at 1 min

intervals for 15 min after addition of the substrate. BApNA, BTpNA, L-Leu-pNA and L-Arg-pNA were used to detect respectively trypsin, chymotrypsin, leucine aminopeptidase (LAP) and cysteine proteinase-like activities. For the latter, pH 6 and 5 mM  $\beta$ ME were chosen following results for major (cysteine) protease activities on gels, and substrate use was described in Bolter and Jongsma [20].

One or two-step gelatin/PAGE [21], modified as described in De Leo et al. [15], was used for the separation of major proteases present in whole extracts of *M. persicae*. Gels were incubated at 37°C for three hours with constant shaking in the buffers used for caseinolytic activity tests [12] with addition of Triton X100 (2.5% final). When used,  $\beta$ -ME (5mM final) and OC-I (0.1mg/ml final) were added after migration in washing and incubation buffers. E64 (0.1mM final) or PMSF (10 mM in isopropanol) were incubated with extracts for 10min at 37°C before migration and was subsequently added to washing and incubation buffers.

## **2.8. Histology**

Encaged groups of 3 plants of each tested line were infested manually on the fifth leave with approximately 50 young adult insects. Insects were allowed to move freely on the plants and the population was allowed to grow for 3 weeks before young adults were collected and fixed in ethanolic Bouin solution and embedded in paraffin. Serial sections (5  $\mu$ m) were mounted on poly-L-lysine-coated microscope slides, de-paraffined and hydrated through xylene-ethanol-water series. Sections were probed with rabbit anti-OC-I antiserum using Vectastain Elite ABC kit (Vector). Sections were first treated with 2% H<sub>2</sub>O<sub>2</sub> in methanol for 30 min and washed with Tris-buffered saline (TBS: 20 mM Tris-HCl pH 7.4, 150 mM NaCl). The sections were then incubated with 2% normal goat serum in TBS (NGS-TBS) for 30 min to block non-specific binding of antibodies. Afterwards, they were placed into anti-OC-I antiserum diluted in NGS-TBS (1  $\mu$ g.ml<sup>-1</sup>) for 1 h, and subsequently washed in TBS. The

sections were then incubated with the secondary antibody (biotinylated goat anti-rabbit IgG) in NGS-TBS for 1 h and subsequently washed thoroughly in TBS. Finally, incubation with the avidin-biotin horseradish peroxidase complex in TBS (ABC reagent) resulted in conjugation of tissue-bound OC-I to peroxidase. Extensive washing in TBS was followed by peroxidase staining by the DAB substrate (3,3'-diaminobenzidine), and sections were counterstained with 0.1% toluidine blue. In control sections, preimmune serum replaced anti-OC-I antiserum.

### 3. Results

When administered *in vitro*, purified oryzacystatin-I did not induce significant nymphal mortality to the four aphid species tested. However, all aphids displayed a slight but significant growth inhibition at a dose of 250 µg/ml (21 µM), or even at lower doses (Figure 1). Although no significant species-specific trend was observed (two-way Anova with *species* x *dose* as factors,  $p=0.08$  for *species* factor,  $p<0.0001$  for *dose* factor), the growth of the *A. gossypii* melon strain was inhibited by over 40% at 250 µg/ml, whereas the growth of other species were only decreased by about 20%. *M. persicae* exhibited a significant decrease in performance at the lowest dose used (10 µg/ml, 0.8 µM), but no further dose-response effect.

From western blot experiments conducted with phloem sap (Figure 2), OC-I was shown to be expressed in phloem at levels comparable to those obtained from whole-leaf extracts. OC-I was also clearly detected in western blots from honeydew extracts of aphids reared on transgenic 4B line (Figure 2). Western blot quantification showed that OC-I accounted for 0.2% of total soluble proteins in phloem samples, whereas it represented 0.3% of total soluble proteins in leaves. The amount of OC-I found in leaves was in accordance with previous estimations made both by western blot and by biochemical papain inhibition assays [16].

The biological parameters obtained with the aphid *M. persicae* in our three independent experiments *in planta* were comparable, and were consistent with the *in vitro* results. They are summarized by the general data of Figure 3 (*exp 2*) and the fecundity data of Figure 4 (*exp 3*). In all experiments, survival data did not reveal any difference between control and transgenic lines, as shown in Figure 3 (non parametric Kruskal & Wallis test,  $p > 0.05$ ). In contrast, mean individual weights of adults reared on lines 4B, 16B and 29A were all significantly different from the control (ANOVA,  $p = 0.0002$  and Fisher's PLSD tests  $p < 0.008$  for all three post-hoc tests for comparisons to control; Figure 3). The resulting biomass after 12 days was also lower on OC-I expressing lines than on control plants (ANOVA  $p = 0.08$  and PLSD test  $p = 0.01$  for line 16B *vs* control). This effect was related to the 25% reduction of individual fecundity observed on insects fed on transgenic lines when compared to insects fed on control plants (*Exp 3*, Figure 4). Therefore, in spite of the low deleterious effects on first generation aphids (as was observed *in vitro*), significant depression of aphid reproduction occurred on OC-I lines.

In order to identify the potential target proteases of OC-I in *M. persicae*, the protease activity in whole aphids extracts was monitored (Figures 5 and 6). When assayed against the non-specific substrate azocasein, the protease profile showed that protease activity was optimum near pH 7 (Figure 5A). At this pH, 5mM  $\beta$ -ME dramatically increased the proteolytic activity (Figure 5A), whereas further adding the cysteine protease inhibitor E64 (0.1mM) completely inhibited the proteolytic activity (result not shown). These results demonstrate that protease activity in whole *M. persicae* extracts is essentially of the cysteine type. Proteases responsible for the mentioned activity were resolved on acrylamide gels through two-step gelatin/PAGE (Figure 5B). At least five proteolytic bands could be observed on gel. Two of them (about 52kDa and 30kDa) were highly activated by the addition of  $\beta$ -ME in the incubation buffer (5mM), the one with the highest molecular weight becoming the most

active protease of all under these conditions. These two proteases were similarly inhibited by E64 (0.1mM) and OC-I (0.1mg/ml) showing the high sensitivity of cysteine proteases toward OC-I in whole *M. persicae* extracts. No qualitative difference could be observed in one or two-step gelatin/PAGE between extracts from aphids fed control plants and extracts from aphids fed plants from the 4B line, and all detected bands were similarly inhibited to E-64 and insensitive to PMSF (result not shown).

An immuno-histological analysis was conducted on aphids grown during their whole pre-reproductive life on plants expressing OC-I (Figure 7B). When compared to aphids reared on control plants (Figure 7A), the main differences in staining were located along the whole midgut epithelium (i), the bacteriocytes (b) and the oenocytes (o). In control aphids, the latter cells exhibit some background staining, as was also observed on the first -“stomach”- part of the digestive tract (st). However, bacteriocytes never showed such non-specific staining (Figure 7A).

#### **4. Discussion**

Many plant disease vectors, such as aphids, whiteflies or mealybugs feed almost exclusively on phloem sap. Producing transgenic plants with deleterious effects on such insects require efficient phloem sap expression. Some studies have provided indirect evidence that constitutive promoters may lead to the expression of entomotoxins in the phloem sap. Indeed, honeydew of phloem-feeding insects reared on transgenic tobacco or potato was shown to contain the snowdrop lectin GNA [9]. In transgenic poplar tree, results of aphid performance on OC-I expressing plants also led to suspect efficient phloem expression (Delplanque, personal communication). The data presented here are the first to clearly demonstrate the expression of a heterologous protease inhibitor in the phloem sap of a crop plant. Many plants have been reported to harbour natural phloem-based protease inhibitors,

but all of these are classified as serine or aspartic protease inhibitors [22]. Cystatins were only recently identified by a peptidomic MALDI-TOF screening of a cucurbit phloem sample [23]. One other class of proteins specifically related to phloem physiology, the filament-forming P-proteins of cucurbits, was shown to be structurally related to a multicystatin inhibitor, although no activity was demonstrated for these high molecular weight polypeptides [24]. It is therefore the first time that a functional cysteine protease inhibitor is reported in the phloem of a non cucurbit plant.

Since hemipteran insects were shown to rely mainly on this protease class for digestion [25], and although no published data was available for strict phloem-feeding insects such as aphids, we decided to test the effects of transgenic oilseed rape on an aphid's growth and reproductive functions. *In vitro* toxicity assays demonstrated that oryzacystatin-I was able to impair the growth (not the viability) of *M. persicae*, an aphid pest of oilseed rape. The effect was also observed with two other aphid species, but did not induce strong harmful effects comparable to other entomotoxic proteins such as the mannose-binding lectins GNA or concanavalin A, for which lethal effects are observed within the dose range used in the present study [26]. In spite of this weak toxic effect, transgenic plants expressing OC-I repeatedly exhibited a deleterious effect on *M. persicae*, which was mainly detectable on reproductive performance. With GNA-expressing potatoes, similar experiments have shown that the lectin reduced total aphid fecundity and delayed by 1.5 or 2 days the peak of reproduction. With OC-I, no detectable influence was observed on peak reproductive time (Figure 4), or on the development of progeny (mean weight of larvae at end of experiment, Figure 3).

It also came out from this study that the prominent proteolytic activity in *M. persicae* was of the cysteine type, although the enzyme was not demonstrated here to be located in the digestive tract. Furthermore, an L-Arg-pNA activity was detected in the aphid and was

completely inhibited by OC-I, suggesting that it may belong to the cathepsin H or L families [27]. As it was also significantly depleted in aphids reared on OC-I expressing plants (Fig. 6), this activity is a good candidate for being a target of this inhibitor in the insect. Since a complete *in vitro* inhibition of this activity was achieved by purified OC-I, stronger *in vivo* deleterious effects may be expected with higher levels of expression of OC-I in the sap of transgenic plants. For instance, expression of GNA in transgenic rice plants at levels of up to 2% of total protein was achieved using the maize ubiquitin promoter [28]. Expression strength and promoter use are critical for retrieving expected phenotypes, and to correctly correlate results obtained *in vitro* and *in planta*, as stressed in a recent work on non-target aphicidal effects of nematode-resistant plants expressing other cystatins [29].

The immunolocalisation of OC-I within the aphids feeding on diets containing this inhibitor (either on transgenic plants or on artificial diets -not shown-) resulted in a picture quite different from that obtained with the mannose-binding lectin ConA fed to the pea aphid [30]. Although some OC-I staining could be detected lining the digestive epithelia, it was observed along the whole tract, in contrast to the ConA lectin which binds almost exclusively to the stomach epithelium. Moreover, whereas no trace of ConA was detected neither within the aphid body, nor by Elisa assays on haemolymph of intoxicated aphids [30], we observed in the present study a clear stain associated with two aphid cell types: the oenocytes and the bacteriocytes. It must be noted however that caution must be taken when comparing those results with other insect/lectin pairs, since for example in the case of the rice brown planthopper *Nilaparvata lugens*, an important binding of lectin GNA to the luminal surface of the midgut epithelial cells have been observed, but it was also observed in the fat body and the ovarioles [31]. Oenocytes constitute an ectodermal cell line found in all true insects, which are very diverse in their secretory and biosynthetic functions, although they have not been extensively studied in aphids or any other homopteran insect [32, 33]. Bacteriocytes are

the specialised cell structures which harbor the obligate intracellular symbionts of aphids, the enterobacteria-like *Buchnera aphidicola* [6]. The structure of OC-I is similar to that of animal cystatins [34], and this molecule shows strong inhibitory activity against cathepsin L [27]. Bacteriocyte labeling, as demonstrated in this work, co-localises with a clear signal for cathepsin-L in these cells (Deraison, unpublished); we confirmed this link by measuring cysteine protease activity on dissected bacteriocytes from *in vitro* intoxicated aphids (*A. pisum* reared to adulthood on diets incorporating 250 µg/ml oryzacystatin), which confirmed both the presence of cysteine proteinase in bacteriocytes, and a 25% decrease of this activity in oryzacystatin-fed aphids. While the reasons governing the localisation of OC-I within such cell types are still obscure, three points may be emphasised: i) the proteinase inhibitor does not accumulate in a specific portion of the digestive tract; ii) it is able to cross the intestinal barrier in an immunologically active form; iii) it displays a tropism for two tissues involved in quantitative transport/ secretion of metabolites (amino-acids for the bacteriocytes, probably lipids and other cuticular-related precursors for the oenocytes). It may also be pointed out that bacteriocytes are crucial for the reproductive function of aphids, as aposymbiotic aphids are reproductively sterile. Therefore, the effect of oryzacystatin on aphid reproduction, and its potential interaction with *Buchnera* symbiosis, may be best studied in long term surveys on transgenic plants, as artificial diets do not allow multiple generation rearing of most clones of aphids.

Finally, although one cannot exclude indirect effects of OC-I expression on aphid performance through minute modifications of phloem physiology due to pleiotropic effects of transgene expression [35], the presented results confirm the validity of the phloem-based strategy to build-up plant resistance towards sap-sucking insects, as first shown with a lectin-expressing potato [9]. Phloem sap is difficult to analyse outside very few model-plant species (cucurbits, lupine, ricinus...), and there is no indication to date that the expression of toxic



polypeptides in this compartment has been used as a natural means of protection against such insects. However, since potentially toxic lectins were described in the phloem of leek for example [36], this hypothesis is highly plausible. Exploring this field will involve the availability of independent genes (allowing enhanced protection or durability by combining them), and will need a better understanding of the genetical and physiological factors governing phloem-specific expression (under control of different promoter types). Also, non-target effects of transgene expression are to be regarded critically, as reported in a recent survey of oryzacystatin-expressing plants on the biology of insect parasitoids [37].

## Acknowledgements

We would like to thank Gabrielle Duport for her technical help in aphid artificial diet experiments, and Gérard Biache for providing us with aphids.

## References

1. G.D. Buntin, P.L. Raymer, Pest status of aphids and other insects in winter canola in Georgia, *J. econ. Entomol.* 87 (1994) 1097-1104.
2. M.A.N. Sedyama, A.T.A. Junior, D.J.H. Dasilva, M.C. Picanco, C.D. Cruz, A study of biochemical markers of kale (*Brassica oleracea* var *acephala*) clones resistant to aphids (*Brevicoryne brassicae*), *Rev. bras. Genet.* 18 (1995) 435-438.
3. G.S. Khush, D.S. Brar, Genetics of resistance to insects in crop plants, *Adv. Agron.* 45 (1991) 223-274.
4. R. Singh, P.R. Ellis, D.A.C. Pink, K. Phelps, An investigation of the resistance to cabbage aphid in *Brassica* species, *Ann. appl. Biol.* 125 (1994) 457-465.
5. L. Jouanin, M. Bonadé-Bottino, C. Girard, G. Morrot, M. Giband, Transgenic plants for insect resistance, *Plant Sci.* 131 (1998) 1-11.
6. A.E. Douglas, Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera*, *Annu. Rev. Entomol.* 43 (1998) 17-37.
7. Y. Rahbé, G. Febvay, Protein toxicity to aphids - an in vitro test on *Acyrtosiphon pisum.*, *Entomol. exp. appl.* 67 (1993) 149-160.

8. P. Tran, T.M. Cheesbrough, R.W. Keickhefer, Plant proteinase inhibitors are potential anticereal aphid compounds, *J. econ. Entomol.* 90 (1997) 1672-1677.
9. A.M.R. Gatehouse, R.E. Down, K.S. Powell, N. Sauvion, Y. Rahbé, C.A. Newell, et al., Transgenic potato plants with enhanced resistance to the peach-potato aphid *Myzus persicae*, *Entomol. exp. appl.* 79 (1996) 295-307.
10. K. Abe, Y. Emori, H. Kondo, K. Suzuki, S. Arai, Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin), *J. biol. Chem.* 262 (1987) 16793-16797.
11. J.S. Womack, J. Randall, J.D. Kemp, Identification of a signal peptide for oryzacystatin-I, *Planta* 210 (2000) 844-7.
12. J.C. Leplé, M. Bonadé-Bottino, S. Augustin, G. Pilate, V.D. Letan, A. Delpinque, et al., Toxicity to *Chrysomela tremulae* (Coleoptera: chrysomelidae) of transgenic poplars expressing a cysteine proteinase inhibitor, *Mol. Breed.* 1 (1995) 319-328.
13. A. Lecardonnell, L. Chauvin, L. Jouanin, A. Beaujean, G. Prevost, X. Sangwan-Norreel, Effects of rice cystatin I expression in transgenic potato on Colorado potato beetle larvae, *Plant Sci.* 140 (1999) 71-79.
14. P.E. Urwin, H.J. Atkinson, D.A. Waller, M.J. McPherson, Engineered oryzacystatin-I expressed in transgenic hairy roots confers resistance to *Globodera pallida*, *Plant J.* 8 (1995) 121-131.
15. F. De Leo, M. Bonadé-Bottino, L.R. Ceci, R. Gallerani, L. Jouanin, Opposite effects on *Spodoptera littoralis* larvae of high expression level of a trypsin proteinase inhibitor in transgenic plants, *Plant Physiol.* 118 (1998) 997-1004.
16. C. Girard, M. Le Metayer, M. Bonadé-Bottino, M.H. Phamdelegue, L. Jouanin, High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae, *Insect Biochem. Molec. Biol.* 28 (1998) 229-237.
17. M. Bonadé-Bottino, J. Lerin, B. Zaccomer, L. Jouanin, Physiological adaptation explains the insensitivity of *Baris coerulescens* to transgenic oilseed rape expressing oryzacystatin I, *Insect Biochem. Molec. Biol.* 29 (1999) 131-138.
18. M. Bonadé-Bottino, Défense du Colza contre les insectes phytophages déprédateurs: étude d'une stratégie basée sur l'expression d'inhibiteurs de protéases dans la plante. 1993, PhD Thesis, pp. 1-91, Paris-Orsay.
19. G. van Canneyt, R. Schmidt, A. O'Conner-Sanchez, L. Willmitzer, M. Rocha-Soza, Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in *Agrobacterium*-mediated plant transformation, *Mol. gen. Genet.* 220 (1990) 245-250.
20. C.J. Bolter, M.A. Jongsma, Colorado potato beetles (*Leptinotarsa decemlineata*) adapt to proteinase inhibitors induced in potato leaves by methyl jasmonate, *J. Insect Physiol.* 41 (1995) 1071-1078.

21. D. Michaud, L. Faye, S. Yelle, Electrophoretic analysis of plant cysteine and serine proteinases using gelatin-containing polyacrylamide gels and class-specific proteinase inhibitors, *Electrophoresis* 14 (1993) 94-98.
22. J.T. Christeller, P.C. Farley, R.J. Ramsay, P.A. Sullivan, W.A. Laing, Purification, characterization and cloning of an aspartic proteinase inhibitor from squash phloem exudate, *Eur. J. Biochem.* 254 (1998) 160-167.
23. S. Haebel, J. Kehr, Matrix-assisted laser desorption/ionization time of flight mass spectrometry peptide mass fingerprints and post source decay: a tool for the identification and analysis of phloem proteins from *Cucurbita maxima* Duch. separated by two-dimensional polyacrylamide gel electrophoresis, *Planta* 213 (2001) 586-593.
24. A.M. Clark, K.R. Jacobsen, D.E. Bostwick, J.M. Dannenhoffer, M.I. Skaggs, G.A. Thompson, Molecular characterization of a phloem-specific gene encoding the filament protein, phloem protein 1 (PP1), from *Cucurbita maxima*, *Plant J.* 12 (1997) 49-61.
25. W.R. Terra, Evolution of digestive systems of insects, *Annu. Rev. Entomol.* 35 (1990) 181-200.
26. Y. Rahbé, N. Sauvion, G. Febvay, W.J. Peumans, A.M.R. Gatehouse, Toxicity of lectins and processing of ingested proteins in the pea aphid *Acyrtosiphon pisum*, *Entomol. exp. appl.* 76 (1995) 143-155.
27. M. Abe, K. Abe, K. Iwabuchi, C. Domoto, S. Arai, Corn cystatin I expressed in *Escherichia coli*: investigation of its inhibitory profile and occurrence in corn kernels, *J. Biochem. (Tokyo)* 116 (1994) 488-92.
28. K.V. Rao, K.S. Rathova, T.K. Hodges, X. Fu, E. Stoger, D. Sudhakar, et al., Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper, *Plant J.* 15 (1998) 469-477.
29. S.E. Cowgill, C. Wright, H.J. Atkinson, Transgenic potatoes with enhanced levels of nematode resistance do not have altered susceptibility to nontarget aphids, *Mol. Ecol.* 11 (2002) 821-7.
30. N. Sauvion, Effects and mechanisms of toxicity of two lectins of the glucose-mannose group towards the pea aphid, *Acyrtosiphon pisum* (Harris). Potential use of plant lectins for creating transgenic plant resistant to aphids. 1995, PhD Thesis, pp. 1-353. INSA-Lyon, Villeurbanne (FRA).
31. K.S. Powell, J.M. Spence, M. Bharathi, J.A. Gatehouse, A.M.R. Gatehouse, Immunohistochemical and developmental studies to elucidate the mechanism of action of the snowdrop lectin on the rice brown planthopper, *Nilaparvata lugens* (Stal.), *J. Insect Physiol.* 44 (1998) 529-539.
32. V.B. Wigglesworth, *The principles of insect physiology*, 1972, Chapman and Hall, London (UK).

33. G.A. Kerkut, L.I. Gilbert, *Comprehensive insect physiology, biochemistry and pharmacology*, 1985, Elsevier, Oxford (UK).
34. K. Nagata, N. Kudo, K. Abe, S. Arai, M. Tanokura, Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. japonica, *Biochemistry* 39 (2000) 14753-60.
35. R. Gutierrez Campos, J. Torres Acosta, J.D. Perez Martinez, M.A. Gomez Lim, Pleiotropic effects in transgenic tobacco plants expressing the oryzacystatin I gene, 36 (2001) 118-119.
36. W.J. Peumans, K. Smeets, K. Vannerum, F. Vanleuven, E.J.M. Vandamme, Lectin and alliinase are the predominant proteins in nectar from leek (*Allium porrum* L.) flowers, *Planta* 201 (1997) 298-302.
37. T.H. Schuler, I. Denholm, L. Jouanin, S.J. Clark, A.J. Clark, G.M. Poppy, Population-scale laboratory studies of the effect of transgenic plants on nontarget insects, *Mol. Ecol.* 10 (2001) 1845-1853.

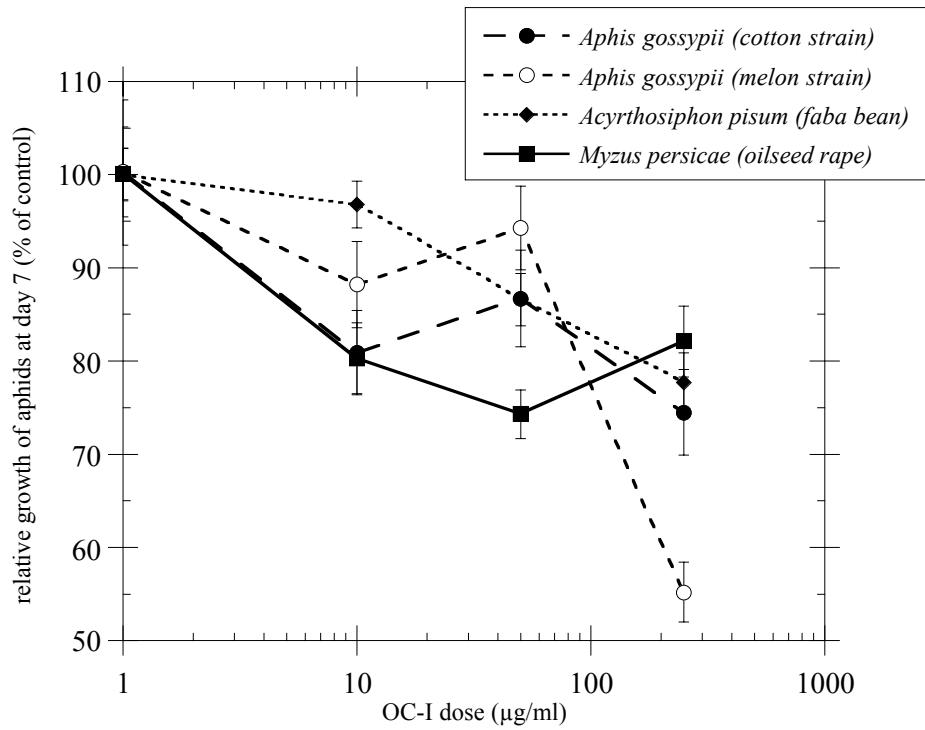


Figure 1: Toxicity (growth inhibition) of oryzacystatin to four aphid clones living on different host plants. Aphid weights expressed as % of weight of species on control diet (means  $\pm$  standard errors).

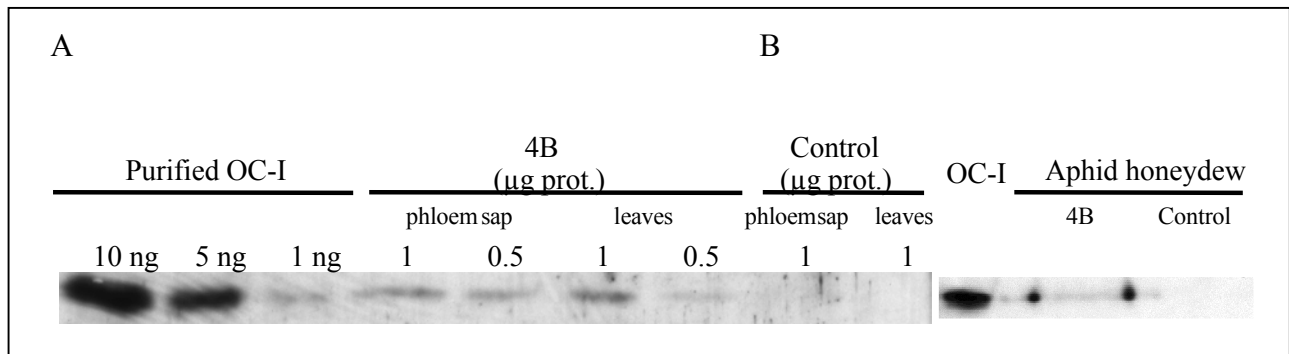


Figure 2: Western blot analysis of phloem sap samples (lanes « phloem sap ») and total leaf proteins from transgenic 4B or control oilseed rape lines (A), and from honeydew of aphids reared on these lines (B). Lanes were loaded with the indicated amounts of either purified OC-I (ng) or total proteins ( $\mu\text{g}$ ).

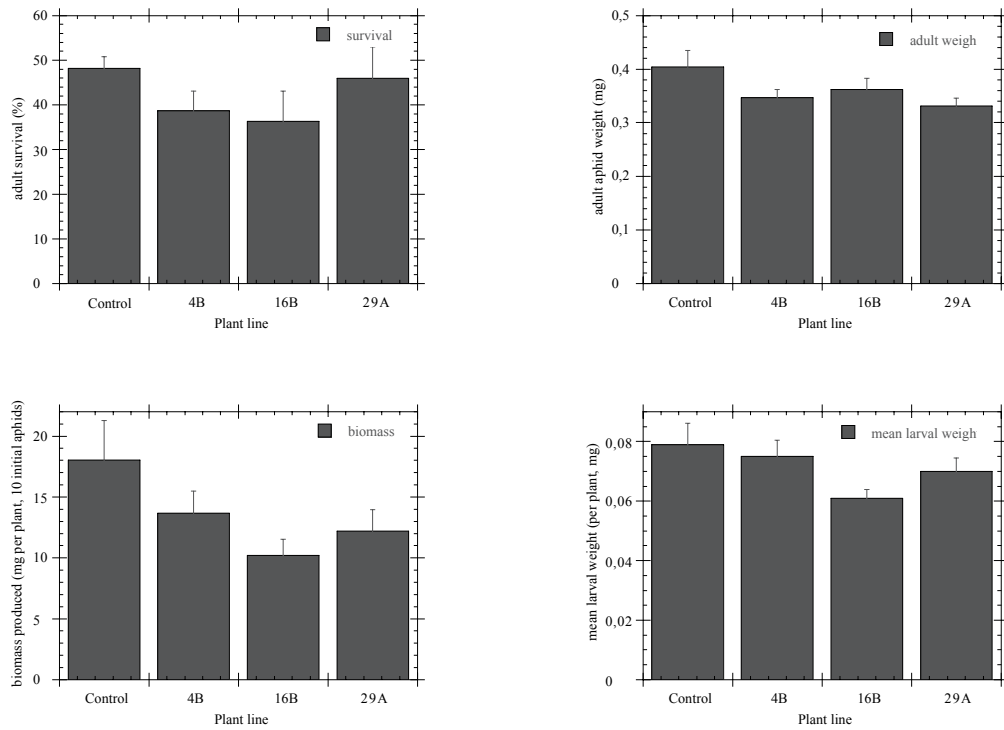


Figure 3: Aphid biological parameters obtained in experiment 2 from three transgenic oilseed rape lines and the control Drakkar cv (left bars). Aphid survival through the 12 days of experiment (top left); mean adult aphid weight (top right); mean aphid biomass produced per plant (bottom left); mean weight of larvae produced on a given plant (bottom right). Error bars are standard errors (per plant data, n=10)

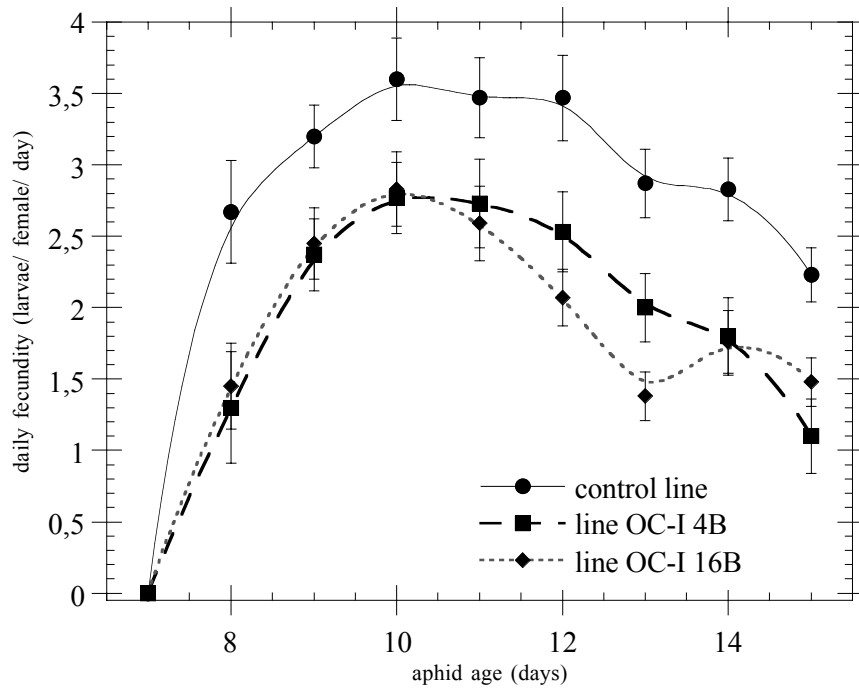


Figure 4: Daily fecundity of the first generation of *Myzus persicae* reared on control (solid line) or transgenic (dashed lines) oilseed rape plants, lines 4B and 16B. Error bars are standard errors; 30 initial females were used per modality; fecundity is reported per surviving female.



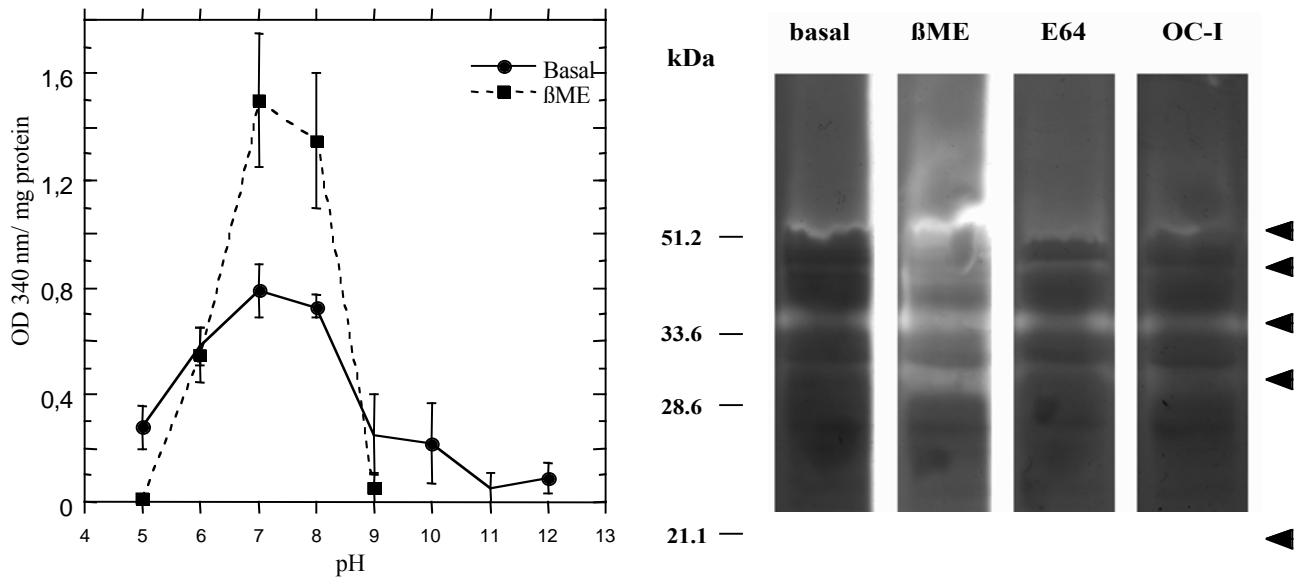


Figure 5: Protease activity in whole *M. percisae* extracts. A: pH sensitivity of caseinolytic activity, measured on 100µg proteins for 3 hours at 37°C either without (basal) or with 5mM β- mercaptoethanol (β-ME - pH 5 to 9). B: Zymograms (35µg proteins per well). Incubation was performed at pH 7, without any addition (basal), in presence of 5mM β-ME, or in presence of 5mM β-ME and 0.1mM E64 (E64) or 0.1mg/ml OC-I (OC-I). Arrows show major protease activities detected.

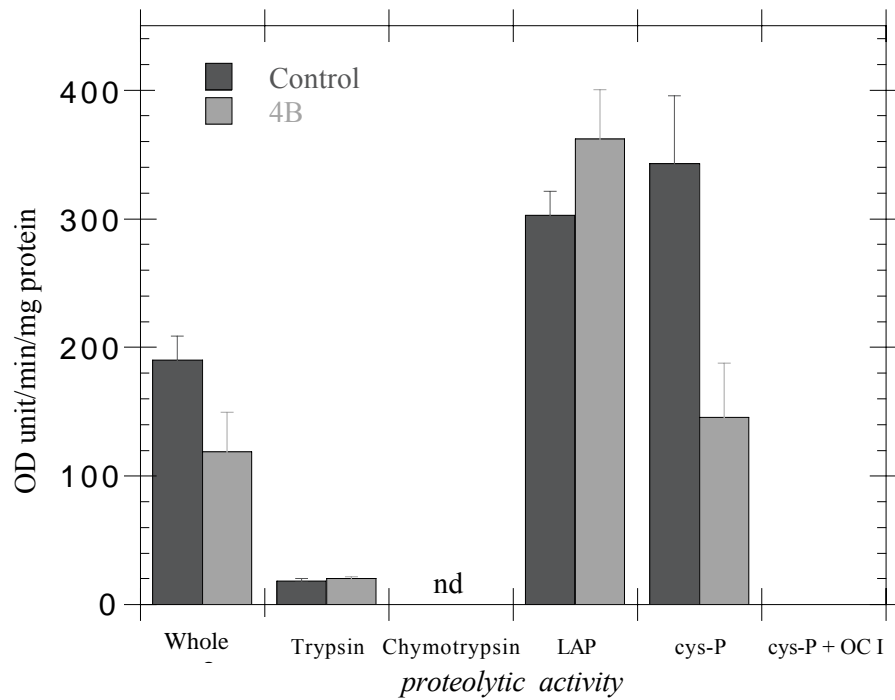


Figure 6: Total and specific proteolytic activities of whole protein extracts from *Myzus persicae* reared on control or OC-I 4B transgenic oilseed rape lines. Whole proteolytic activity (whole) was measured using azocasein as substrate (O.D. 340 nm,  $\times 10^1$ ). Specific activities were measured using p-nitroanilide substrates (O.D. 405 nm,  $\times 10^3$ ): trypsin (T), chymotrypsin (C, *nd* not detected), leucine aminopeptidase (LAP), cysteine protease either without (cys-P) or with 0.1 mg/ml OC-I (cys-P + OC-I).

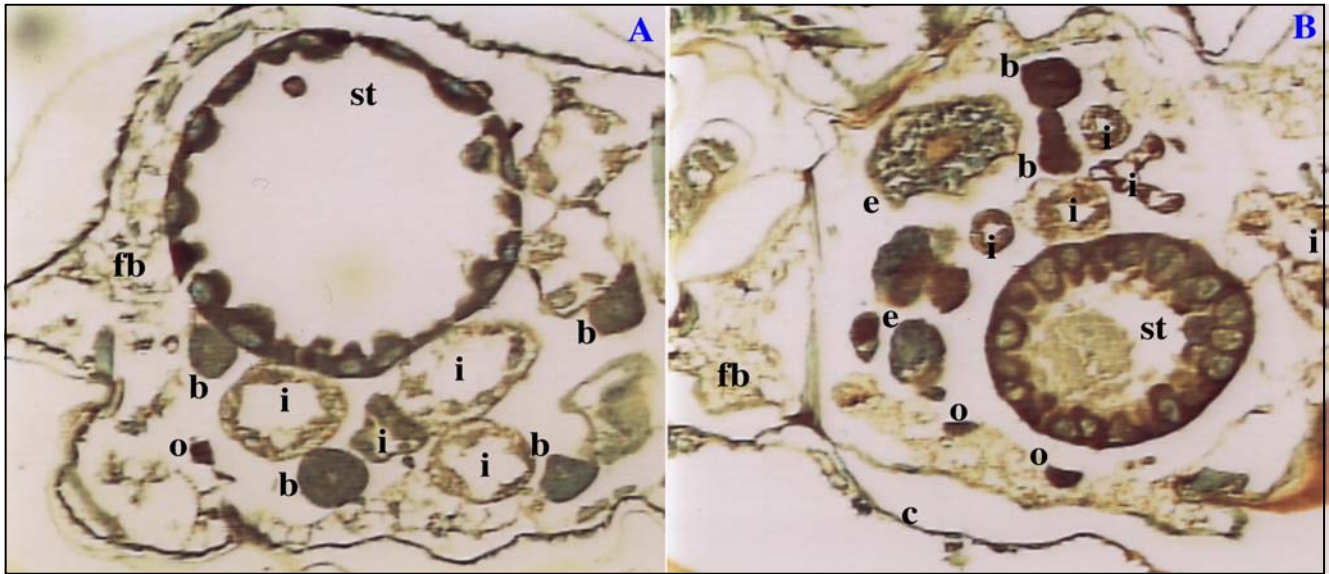


Figure 7: Immunohistological detection of oryzacystatin in adult *Myzus persicae* aphids reared on oilseed rape plants expressing (B) or not (A) the OC-I gene. Transverse sections at the level of *stomach* (st) show the brownish peroxidase stain coating the *digestive tract* (B: st, i: *intestine*), and labelling aphid *oenocytes* (o) and *bacteriocytes* (b, compare B to A). c: *cuticle*, e: *embryos*, fb: *fat-body*.

## II. Diversification des inhibiteurs de protéases à cystéine

### Introduction

A ce jour, le nombre de gènes disponibles pour lutter contre les pucerons *via* la transgénèse est très restreint. L'expression de l'oryzacystatine permet d'observer une diminution de la fécondité du puceron *M. persicae* mais ne confère pas un niveau de protection commercialement intéressant dans l'état actuel de la demande professionnelle (article 2). Bien que les inhibiteurs de protéases représentent une nouvelle voie, il est nécessaire d'accroître la diversité de ces molécules entomotoxiques en sélectionnant des protéines susceptibles de conférer aux plantes une résistance accrue face aux attaques de pucerons.

Il est possible d'obtenir de nouveaux inhibiteurs de protéases grâce aux techniques de mutagenèse dirigée ou de "phage display" à partir des séquences d'IP d'origine végétale (Ceci, 2002). Ces nouvelles molécules ainsi créées peuvent être sélectionnées selon leur affinité avec les protéases de l'insecte cible. Un test en nutrition artificielle permet ensuite de mettre en évidence leurs effets délétères sur la physiologie de cet insecte. L'approche par mutagenèse dirigée a été développée avec un inhibiteur de protéase à sérine du pois, PsTI (chapitre III).

Des séquences d'inhibiteurs de protéases peuvent être aussi isolées d'organismes éloignés des plantes. Une équipe américaine a obtenu des plantes transgéniques montrant une résistance élevée à *Bemisia tabaci* par expression constitutive d'inhibiteurs de protéases originaires de l'hémolymph de larves de *Manduca sexta* (Thomas, 1995a ; Thomas, 1995b).

### 1.1. Résistance dérivée du pathogène

Historiquement, le concept de résistance dérivée du pathogène (PDR pour Pathogen-Derived Resistance) a été émis par Sanford et Johnston en 1985. Ces auteurs ont proposé d'induire une résistance à un pathogène en exprimant dans un organisme sensible des séquences dérivées de ce pathogène. Cette stratégie est très utilisée pour l'obtention de plantes transgéniques résistantes aux virus (pour revue : Berthomé, 2000). De nombreux transgènes viraux ont été introduits dans les plantes et ont permis de protéger différentes espèces végétales contre des virus appartenant à de nombreuses familles, grâce à deux types de

mécanismes de résistance : (1) *résistance conférée par l'ARN viral*, obtenue par l'expression de transcrits interférant avec la réplication de l'ARN viral ou induisant la dégradation séquence – spécifique de l'ARN viral,

(2) *résistance conférée par la protéine*, obtenue par l'expression d'une protéine virale fonctionnelle ou tronquée, codée par le transgène (protéine capsidiale, protéine de mouvement ou ARN polymérase).

Ce concept peut néanmoins être dérivé afin de développer la diversité des inhibiteurs de protéases pour lutter contre les pucerons. Nous souhaitons utiliser les insectes comme source de séquence d'inhibiteurs de protéases entomotoxiques. L'isolement d'inhibiteurs de protéases du puceron pourrait permettre théoriquement d'utiliser des molécules très affines pour les protéases endogènes de cet insecte.

## 1.2. Les inhibiteurs de protéases chez les insectes

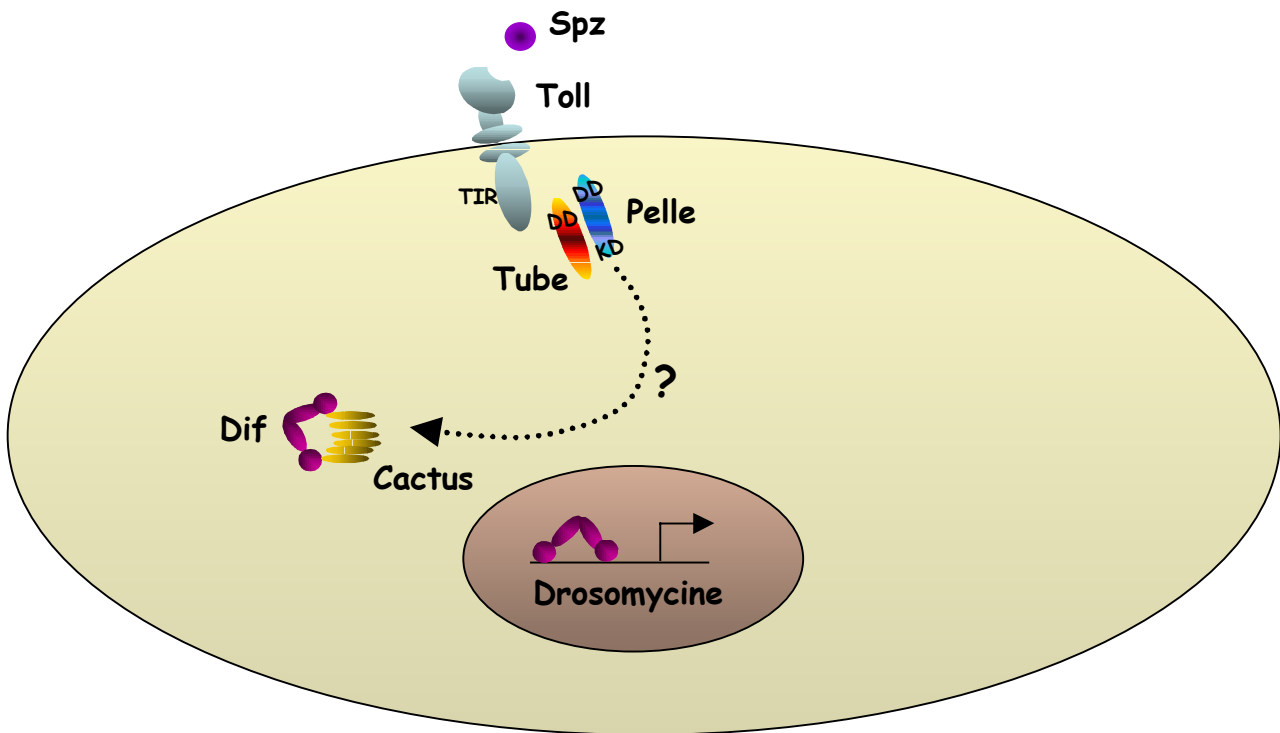
### 1.2.1. Les inhibiteurs de protéases à sérine

Les informations concernant le rôle physiologique des inhibiteurs de protéases chez les insectes reposent surtout sur l'étude des inhibiteurs de protéases à sérine. Il existe quatre familles d'inhibiteurs de protéases à sérine chez les insectes :

- la famille des serpins dont les membres possèdent un poids moléculaire d'environ 40 kDa ;
- la famille des « Kunitz inhibitors » comprend des protéines plus petites (10 kDa environ) (pour revue : Polanowski, 1996) ;
- la famille des « *Locusta* inhibitors » ; ces protéines isolées chez le criquet sont des peptides de 35 acides aminés possédant une activité antichymotrypsine et/ou prophénoloxydase (Boigegrain, 2000 ; Kanost, 1999) ;
- enfin, la famille des « *Ascaris* inhibitors » renferme des inhibiteurs de la chymotrypsine qui sont des peptides d'environ 6 kDa possédant cinq ponts disulfure (Bania, 1999).

Une des fonctions des IP de l'hémolymphe est probablement la régulation des protéases impliquées dans **les réactions immunitaires** telles que la synthèse de peptides antimicrobiens et la synthèse de mélanine :

- la réaction immunitaire en réponse aux micro-organismes est particulièrement bien étudiée chez la drosophile grâce aux outils génétiques que constituent les lignées mutantes.



**Figure 2.1 : Contrôle d'une voie de la synthèse inductible de peptides antimicrobiens chez la drosophile adulte (Imler et Hoffmann, 2001).**

Le récepteur TOLL est activé par le facteur de croissance SPAETZLE (SPZ). Le récepteur TOLL induit la phosphorylation de l'inhibiteur CACTUS par l'intermédiaire des protéines TUBE et PELLE. La dégradation de CACTUS permet le passage dans le noyau du facteur de transcription DIF (Drosal-related immune factor) qui induit la synthèse du peptide antifongique DROSOMYCINE

DD, death domain ; KD, kinase domain ; TIR, toll/IL-1 receptor homology domain ; IL, interleukin

Pour se défendre contre les pathogènes, la drosophile met en jeu une réponse humorale et cellulaire. La réponse humorale est basée sur l'induction et la sécrétion dans l'hémolymphe de nombreux peptides. Ces derniers possèdent une activité antifongique (DROSOMYCINE, METTCHNIKOWIN) ou antibactérienne (DIPTERICINE, DROSOCINE, CECROPINE, ATTACINE, DEFENSINE). La voie de signalisation permettant l'activation du gène codant la DROSOMYCINE (figure 2.1) met en jeu une cascade de régulation de plusieurs effecteurs dont des protéases à sérine (Imler et Hoffman, 2001). Bien que l'on ne connaisse pas encore comment la cascade est déclenchée en amont, des antigènes présents à la surface du champignon<sup>16</sup> pourraient être reconnus par un récepteur activant en cascade le clivage du propeptide de plusieurs protéases à sérine (codées par les gènes *GASTULATION DEFENSIVE*, *SNAKE* et *EASTER*). La protéase *EASTER* clive le polypeptide *SPAETZLE* qui active alors le récepteur transmembranaire *TOLL*. La réaction entre ces deux protéines n'est pas encore complètement élucidée mais elle conduit à la réponse cellulaire : l'interaction entre le ligand extracellulaire *SPAETZLE* et le récepteur *TOLL* induit la phosphorylation de l'inhibiteur *CACTUS* via la protéine *TUBE* et la kinase *PELLE*<sup>17</sup>. La phosphorylation de *CACTUS* représente un signal pour sa dégradation et permet ainsi le passage du facteur de transcription *DROSAL* du cytoplasme au noyau induisant la transcription du gène de la *DROSOMYCINE*.

Des inhibiteurs de protéases régulent cette voie de signalisation. Chez le mutant de drosophile déficient pour le gène *spn43ac* codant pour un inhibiteur de protéase à sérine, on observe une expression constitutive de la *DROSOMYCINE* ainsi que le clivage permanent de *SPAETZLE* (Levashina, 1999). Des inhibiteurs de protéases sont donc des modulateurs de la réponse immunitaire des insectes.

- la mélanisation est un autre mode de neutralisation du pathogène. Lorsque l'organisme étranger est trop gros pour être phagocyté par les hémocytes, il est encapsulé grâce à de multiples couches de mélanine. Ces capsules de mélanine isolent, immobilisent et tuent le pathogène. L'encapsulation est généralement une réponse à la présence de protozoaires ou de métazoaires comme des parasites et des œufs de parasitoïdes (Gillespie, 1997).

---

<sup>16</sup> Chez le ver à soie, les antigènes reconnus à la surface du champignon sont des  $\beta$ 1-3 glucanes. Des protéines qui reconnaissent spécifiquement ces sucres ont été isolées de l'hémolymphe de *Bombyx mori* (cité dans Gillespie et Kanost, 1997) et semblent actuellement appartenir à la famille des PGRP.

<sup>17</sup> La kinase *PELLE* ne phosphoryle pas directement *CACTUS*.

Une des enzymes clé de la mélanisation est la prophénoloxydase (PO). Cette enzyme catalyse l'oxydation de phénols en quinones qui polymérisent spontanément en mélanine. Bien que les étapes directes de reconnaissance d'antigènes tels que des lipopolysaccharides, des  $\beta$ 1-3 glucanes ou des peptidoglucanes, ne sont pas encore connues, l'activation de la PO est médiée *via* une cascade protéasique. La PO est présente dans l'hémolymphe ou dans les hémocytes sous forme inactive (proPO), ce zymogène est clivé par une cascade de protéases à sérine en amont.

Une seule protéase a été isolée à ce jour, chez le ver à soie, c'est la protéase responsable de la dernière étape. L'enzyme activatrice de la prophénoloxydase (PPAE) est une protéase à sérine, homologue à la protéase EASTER (Sato, 1999).

L'hémolymphe du moustique contient au moins cinq protéases à sérine dont quatre sont induites par la présence de bactéries. L'une d'entre elles présente plusieurs domaines de reconnaissance potentielle du pathogène (Gorman, 2001).

Certains inhibiteurs de protéases à sérine de l'hémolymphe pourraient réguler la mélanisation. Par exemple, Boigegrain *et al.* (1992) ont observé que les peptides isolés de l'hémolymphe du criquet inhibaient non seulement la chymotrypsine et l'élastase, mais aussi la cascade d'activation de la PO chez cet insecte.

Les IP à sérine ont également un **rôle défensif direct**. Les microorganismes produisent des enzymes protéolytiques leur facilitant la pénétration à travers la cuticule. Eguchi *et al.* (1993) ont observé que les IP de l'hémolymphe ou de la cuticule du ver à soie inhibaient fortement l'activité des protéases synthétisées par les champignons entomopathogènes. Un inhibiteur de protéase à sérine (Fungal Protease Inhibitor-F) hémolympatique inhibe aussi la sporulation du champignon pathogène *Beauveria bassiana* (Yoshida, 1990 cité dans Eguchi, 1993). D'autre part, lorsque la larve du lépidoptère *Hyphantria cunea* est envahie par des bactéries, une serpine est sur-exprimée une heure après l'infection par *E. coli* (Shin, 1998). Ces inhibiteurs de protéases pourraient protéger l'insecte des activités protéasiques exogènes lors de l'infection par des pathogènes.

Enfin, les IP de l'hémolymphe des larves d'une espèce inhibent fortement les protéases digestives de la même espèce (Eguchi, 1993 ; Girard, 1998a). Ces IP possèderaient une fonction de **protection tissulaire**. Ils inhiberaient des protéases libérées dans le fluide de la mue à la suite de la destruction des tissus lors de la métamorphose ou après lésion du tube digestif par les pathogènes.



### 1.2.2. Les inhibiteurs de protéases à cystéine

Seuls trois inhibiteurs de protéase à cystéine ont été isolés chez les insectes. L'hémolymphe de *Sarcophaga peregrina* renferme deux inhibiteurs de protéases à cystéine A $\alpha$  (10 kDa) et A $\beta$  (9,5 kDa) (Suzuki, 1985). Ces deux protéines sont le résultat d'une modification post-traductionnelle du produit d'un seul gène (Saito, 1989). On retrouve ces deux IP soit sous la forme libre, soit complexée (2A $\alpha$  –A $\beta$ ) dans l'hémolymphe. Il semble que le complexe présente une plus faible capacité d'inhibition des protéases présentes dans les hémocytes (majoritairement des protéases à cystéine) que les formes libres. Les conditions d'association/dissociation du complexe pourraient donc réguler l'activité de ces IP. L'expression du gène codant ces IP est activée transitoirement au stade pupa. Une fonction possible de cet inhibiteur serait de protéger les tissus de l'adulte en développement, tel que le disque imaginal, des protéases hémocytaires.

Chez la drosophile, un gène codant pour un inhibiteur de protéase à cystéine a été cloné (Delbridge, 1990), et trois autres membres de cette famille sont présents dans le génome de cette espèce. Les séquences des IP de drosophile et de *Sarcophaga* montrent qu'ils appartiennent à la famille des cystatines (Brown, 1997), caractérisée par la présence de motifs spécifiques de l'inhibition des protéases à cystéine (QxVxG et PW) et pour certaines d'un pont disulfure.

### 1.2.3. Les autres inhibiteurs de protéases

Un seul inhibiteur de métalloprotéase a été mis en évidence, chez la teigne de la ruche, *Galleria mellonella*. Cet IP est inductible et présent dans l'hémolymphe de l'insecte uniquement après immunisation contre les champignons. C'est une protéine de 8 kDa qui renferme cinq ponts disulfure. Aucune métalloprotéase n'ayant été détectée dans l'hémolymphe de *Galleria*, il semble donc que cet IP intervienne dans l'inhibition de protéases synthétisées par les pathogènes (Wedde, 1998).

Notre étude a pour but d'évaluer la diversité des IP de l'hémolymphe du puceron, leur capacité à inhiber les protéases digestives et d'élargir le spectre d'IP disponibles pour la transformation des plantes grâce à l'isolement d'un IP du puceron.

## Matériel et Méthodes

### 2.1. Matériel biologique

Cette étude repose sur la lutte contre le puceron du melon *Aphis gossypii*. Mais pour des raisons de faisabilité (récoltes d'hémolymphe, quantité de pucerons) nous avons utilisé, pour la caractérisation biochimique d'un inhibiteur de protéase à cystéine, le puceron du pois, *Acyrtosiphon pisum*. Ce puceron est le puceron modèle du laboratoire BF2I de Lyon mais il est surtout notablement plus gros qu'*A. gossypii* (4-5 mg pour des adultes de *A. pisum*, 300-400 µg pour *A. gossypii*).

Le puceron du pois *A. pisum* a été initialement introduit dans le laboratoire BF2I de l'INRA-INSA de Lyon par un échantillon collecté sur luzerne en 1986 (Rahbé, 1988). Les pucerons sont élevés sur des jeunes fèves (*Vicia faba*, cv aqualduce) à 21°C, en lumière artificielle (16h de lumière) et à 60-70% d'hygrométrie. Les fèves sont utilisées à un âge de trois à cinq semaines.

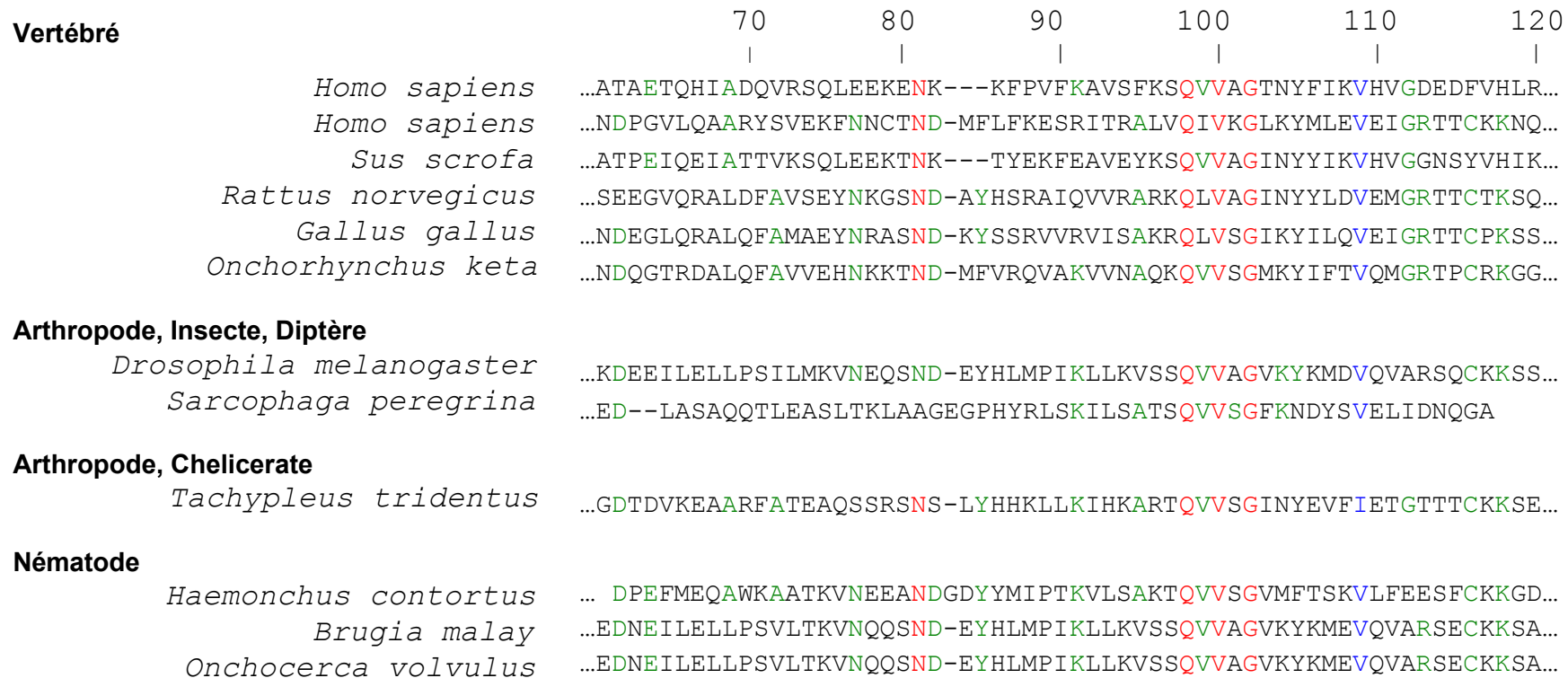
### 2.2. Préparation des extraits protéiques

#### 2.2.1. Récolte d'hémolymphe

Pour prélever l'hémolymphe, les pucerons sont fixés vivants sur une lame grâce à un adhésif double face, une antenne est coupée avec des micro-ciseaux au niveau du deuxième article. La goutte exsudant de la coupure est prélevée avec un micro-capillaire. Trois à quatre capillaires sont plongés dans un tube Eppendorf contenant 20 µl d'une solution de NaCl 0,15 M saturée en PTU<sup>18</sup>, et une centrifugation à basse vitesse permet ensuite de faire descendre le contenu des capillaires et de dissoudre ainsi l'hémolymphe dans la solution saline. Chaque puceron permet ainsi la récolte d'environ 100 nl d'hémolymphe.

---

<sup>18</sup> Phényl-thio urée, inhibiteur de la mélanisation



**Figure 2.2 : Comparaison de séquences protéiques d'inhibiteurs de protéase à cystéine de quelques organismes vivants.**  
 La barre indique la localisation de l'oligonucléotide dégénéré QV.

### 2.2.2. Extraction des protéines

Des jeunes adultes de *A. pisum* congelés sont broyés à l'UltraTurrax grâce à 3 séquences de 10 secondes, à 4°C, dans une solution de méthanol 50% (10 ml par g de puceron). L'extraction protéique est ensuite poursuivie en agitant le broyat 2 h à 4°C. L'extrait ainsi obtenu est centrifugé 10 min à 11000 g, et le surnageant résultant est conservé à -20°C jusqu'à utilisation. Le méthanol est ensuite éliminé par évaporation sous vide d'air, et l'échantillon est repris dans 1 ml d'eau. Cet extrait est ensuite déposé sur une minicolonne (Sep-Pak, Waters) contenant 1 ml de phase stationnaire en C18 préalablement équilibrée dans l'eau. La colonne est lavée avec 5 ml d'eau, puis l'extrait est élué par 3 ml de méthanol à 50% (l'activité IP à cystéine se retrouve dans cette fraction). Après élimination du méthanol par évaporation sous vide (Speed-Vac), l'extrait ainsi obtenu est conservé dans l'eau à 4°C.

### 2.3. Construction et analyse de la banque d'ADNc

De jeunes adultes aptères d'*A. gossypii* sont prélevés délicatement de leur plante hôte et immédiatement congelés dans l'azote liquide, puis conservés à -80°C jusqu'à utilisation.

Une banque d'ADNc représentative des ARNm du puceron entier *A. gossypii* a été construite selon la procédure décrite pour la construction de la banque d'ADNc de tube digestif (article 2). Elle contient  $1,4 \cdot 10^6$  clones indépendants dont 89% sont recombinants.

Plusieurs couples d'oligonucléotides dégénérés correspondant à des séquences consensus de cystatine (figure 2.2) ont été utilisés afin de cribler la banque, ou utilisés en RT-PCR.

L'oligonucléotide cystatine f correspond au domaine QVVAG conservé des cystatines de nématodes jusqu'à celles de l'homme. Les oligonucléotides QVr, QVf et PWr correspondent au domaine QxVxG et PW du site actif des cystatines de l'hémolymphe de *Sarcophaga peregrina* et *Drosophila melanogaster*.

L'ADN génomique de *Drosophila melanogaster* constitue un témoin positif de la réaction de PCR entre les oligonucléotides QV f et PW r.

Dans le tableau 2.1, les différentes combinaisons d'oligonucléotides utilisées en vue de clonage d'une cystatine sont présentées. Les oligonucléotides dégénérés sont utilisés dans un couple avec une autre amorce dégénérée ou bien avec une amorce spécifique du vecteur de clonage de la banque.

**Tableau 2.1 : Oligonucléotides utilisés en vue du clonage de l'IP à cystéine.**

	5' seq	3' seq	polyT	PWr
Cystatine f GTA ITT (CG)AI ICC I(TG)I IAC IAI CTG		X		
QVr CCI (CG)(ATC)I ACI AC(TC) TGI (CG)(ATC)I GT	X		X	
QVf ACI (ATG)(GC)I CA(AG) GTI GTI (ATG)(GC)I GG		X		X
PWr A(GA)C CAI GGI TGI (CG)(AT)C CA(ATG) AT	X		X	

Les oligonucléotides 5' seq et 3' seq sont spécifiques du vecteur dans lequel la banque d'ADNc a été clonée de façon orientée (*cf* Article 2), et l'oligonucléotide polyT a été utilisé pour générer des ADNc en RT-PCR.

Différentes températures d'hybridation (45° à 55°C) pendant 30 secondes ainsi que différents programmes (Touch-up ou Touch-down, incrémentation d'un degré à chaque cycle) ont été utilisés.

## 2.4. Dosage des protéines

La concentration en protéines des différents extraits a été déterminée selon la méthode décrite par Bradford (1976), en utilisant la Bovine Serum Albumin (BSA) comme référence.

## 2.5. Méthodes d'analyse des protéines

### 2.5.1. Electrophorèse Tris-Tricine.

Nous avons utilisé une électrophorèse en Tris-Tricine adaptée de Schagger (1987) afin d'obtenir une meilleure séparation des peptides inférieurs à 20 kDa.

Le tampon de charge est un tampon Tris-HCl 50 mM, pH 6,8, contenant 12% de glycérol et 4% de sodium dodécyl sulfate (SDS), du bleu de Coomassie G-250 et 2% de  $\beta$ -mercaptoéthanol. Le tampon pour les gels de concentration et de séparation est un tampon Tris-HCl 3 M pH 8,45 ; SDS 0,3%. Le tampon de migration de la cathode est constitué de Tris-HCl 0,1 M ; tricine 0,1 M ; SDS 0,1%. Le tampon de migration pour l'anode est un tampon Tris-HCl 0,2 M pH 8,9.

Les échantillons sont séparés sur PAGE en conditions dénaturantes à l'aide d'un appareillage vertical Mini-protean II (Biorad). Les gels de polyacrylamide 10 x 10 cm sont de 0,5 mm d'épaisseur.

Les échantillons biologiques sont mélangés volume à volume avec le tampon de charge, chauffés 5 min à 100°C et déposés sur le gel d'électrophorèse. Le gel de séparation est à 10,6% d'acrylamide et le gel de concentration à 4% d'acrylamide. La migration des échantillons est réalisée à voltage constant (60 V pour le gel de concentration et 120 V pour le gel de séparation).

A l'issue de l'électrophorèse les gels sont soit colorés au bleu de Coomassie soit au nitrate d'argent selon la quantité de protéines déposées.

### **2.5.2. Electrophorèse bidimensionnelle**

L'électrophorèse bidimensionnelle consiste à séparer dans un premier gel les protéines en fonction de leur point isoélectrique par isofocalisation. Le gel obtenu après cette première séparation est repris et déposé sur un second gel SDS-PAGE où les protéines sont séparées selon leurs poids moléculaires.

Les électrophorèses bidimensionnelles sont réalisées en système Mini-Protean II 2D (BioRad) avec le protocole standard décrit par le fournisseur (gamme de pH 3-10 pour la première dimension d'isofocalisation, 15% d'acrylamide pour la SDS-PAGE).

### **2.5.3. Transfert des protéines sur membranes ou sur gel de polyacrylamide.**

Pour les transferts sur membranes, le gel et la membrane (PVDF) sont équilibrés 15 min dans le tampon Tris 25 mM ; glycine 100 mM pH 8,3 ; méthanol 20%. Les protéines sont électrotransférées sur la membrane (30 min à 15 V) grâce à un appareil de transfert semi-sec (Biorad).

Dans le cas de transfert gel à gel, le gel est incubé 1h à 4°C dans une solution de triton X100 2,5%. Les protéines sont ensuite transférées par électrotransfert (60V, 20 min) à l'aide d'un appareil de transfert (Biorad) en milieu liquide (Tris 25 mM ; glycine 100 mM pH 8,3) sur un gel d'activité (12% acrylamide, 0,1% gélatine) afin de révéler la présence d'un IP.

## **2.6. Mise en évidence d'inhibiteurs de protéases**

### **2.6.1. Dosage de l'activité IP à cystéine**

Afin de préserver la quantité de matériel, nous avons utilisé un test en microplaque de 96 puits qui permet d'utiliser très peu de matériel et de tester de nombreuses conditions en une seule fois.

Le principe consiste à mettre en contact des quantités croissantes d'un extrait de l'échantillon à doser avec de la papaïne en excès. Les extraits sont incubés pendant 15 min à température ambiante, en présence de 50 ng de papaïne<sup>19</sup> dans un volume final de 40 µl de tampon phosphate de sodium 0,1 M pH6,5 ; EDTA 10 mM ; DTT 2 mM.

La réaction commence par l'addition du substrat Z-phe-Arg-pNA à 200 µM final et l'homogénéisation des 200 µl de la réaction enzymatique. L'apparition du produit d'hydrolyse est mesurée à 410 nm toutes les 5 minutes pendant 90 minutes sur un lecteur de microplaques (ELX 50, Biotek Instruments).

Les résultats sont exprimés en unités de papaïne inhibée (PUI) par µg de protéine, une unité correspondant à une augmentation de DO de 0,01 unité d'absorbance par minute à 410 nm. La courbe d'inhibition est tracée en portant les variations d'absorbance par minute et par quantité de protéine dans l'échantillon, la pente de la droite de régression de cette courbe correspondant à la valeur de PUI par µg de protéines. Seules les valeurs comprises entre 10 et 80% d'inhibition de la papaïne ont été prises en considération. En effet, en dehors de ces limites la relation entre l'absorbance et la quantité d'IP n'est pas linéaire et ne permet donc pas d'effectuer un dosage.

---

<sup>19</sup> La papaïne est conservée à -20°C sous forme d'aliquote concentrée 100 fois (250 µg.ml<sup>-1</sup>) avec 0,25% de Triton X100. Deux heures avant le test enzymatique, l'enzyme est décongelé lentement dans le tampon d'activité.

## **2.6.2. Mise en évidence de l'activité des IP sur gel**

Une électrophorèse SDS-PAGE est réalisée à 4°C sur un gel d'acrylamide à 15% additionné de gélatine 0,1%. Le tampon de charge est dépourvu de réducteur (Tris-HCl 100 mM, pH 6,8 ; glycérol 12% ; SDS 4% ; bleu de bromophénol 0,02%) et les échantillons ne sont pas chauffés avant dépôt.

Après migration, le gel est placé dans une solution de Triton X100 2,5% pendant 1 h à 4°C afin d'éliminer le SDS. Il est ensuite incubé pendant 4 h à 37°C dans 25 ml soit dans un tampon Na<sub>2</sub>HPO<sub>4</sub> 100 mM pH 6,8 contenant 12,5 µg de papaïne, soit dans un tampon Tris-HCl 100 mM pH 8 contenant 12,5 µg de trypsine. Le gel est ensuite coloré au bleu de Coomassie. La protéase digère la gélatine ainsi que les protéines présentes dans le gel sauf au niveau de l'IP qui apparaît sous forme d'une large bande foncée sur fond clair.

## **2.7. Méthodes de purification des protéines**

### **2.7.1. Chromatographie d'échange d'ions**

Les protéines sont fractionnées sur des colonnes échangeuses d'ions constituées de 1 ml de phase stationnaire (Hi-Trap Sépharose, Pharmacia). Les colonnes d'échange d'anions faibles (HiTrap DEAE) et fortes (HiTrap Q) sont équilibrées dans le tampon AN (HEPES 50 mM pH 8), puis l'extrait préalablement repris dans 1 ml de tampon AN est déposé. La colonne est lavée avec 3 ml de tampon AN. Les protéines sont ensuite éluées par palier de concentrations croissantes en NaCl (0,1 M ; 0,3 M ; 0,8 M puis 1,5 M), à raison de 3 ml de solution par concentration en sel. Le débit est de 1 ml.min<sup>-1</sup> et des fractions de 3 ml correspondant à chaque palier sont récoltées. Pour les colonnes d'échanges de cations faibles (HiTrap CM) ou fortes (HiTrap SP), le mode opératoire est le même excepté que le tampon AN est remplacé par un tampon CA (Tris 20 mM pH 8).

### **2.7.2. Chromatographie d'exclusion de taille**

Le fractionnement par chromatographie d'exclusion de taille est effectué sur une colonne de 50 x 3 cm remplie de gel (Bio-Gel P4, Biorad ; taille d'exclusion : 0,8 à 4 kDa, volume mort : 32 ml) à 4°C. La phase mobile est constituée d'un tampon phosphate de potassium 10 mM pH 6,5. L'extrait, préalablement amené à sec, est repris par 100 µl du même tampon et



déposé sur la colonne. Le débit d'éluion utilisé est de  $0,1 \text{ ml.min}^{-1}$  et les fractions collectées ont un volume de 8 ml. L'éluion est suivie par spectrophotométrie à 280 nm.

Les fractions sont concentrées sous vide (Speed-Vac), reprises dans l'eau et conservées à  $4^{\circ}\text{C}$ .

### **2.7.3. Chromatographie d'affinité**

La papaïne est greffée de manière covalente sur un support de Sépharose 4B. Cette matrice de papaïne carboxyméthylée nous a été fournie par le laboratoire de M. Jongma (Wageningen, Pays Bas). Le gel est coulé dans une colonne de  $10 \times 1 \text{ cm}$  dans le tampon AF (Tris 20 mM pH 7,5 ; NaCl 0,5 M). La phase stationnaire est équilibrée à un débit de  $1 \text{ ml.min}^{-1}$  par le tampon d'activation (MES 50 mM pH 6,5 ; DTT 2,5 mM) qui permet d'activer la papaïne.

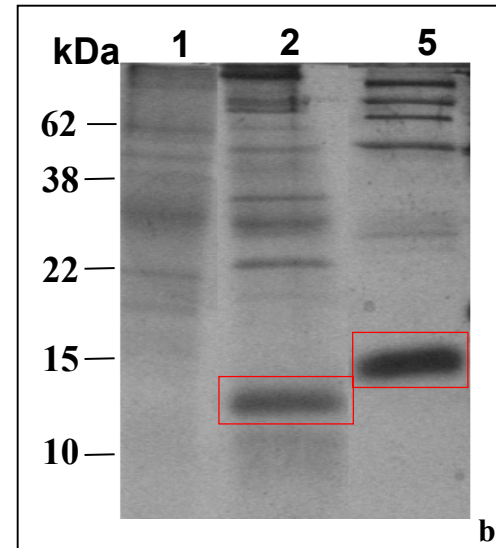
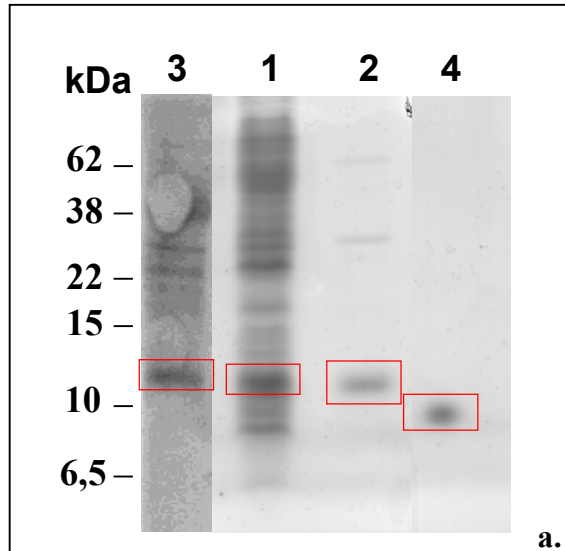
L'échantillon à activité IP-cystéine amené à sec est repris dans le tampon d'activation, puis déposé sur la colonne et recyclé pendant 2 h à température ambiante. La colonne est ensuite lavée avec le tampon AF à un débit de  $1 \text{ ml.min}^{-1}$  jusqu'à retrouver une densité optique à 280 nm équivalente à la ligne de base. Le déplacement des sites est effectué pendant 1h à température ambiante dans un tampon constitué de E64  $10\mu\text{M}$  ; Tris 20 mM pH 10,3 ; DMSO 20%. Le débit d'éluion est de  $1 \text{ ml.min}^{-1}$  et des fractions de 1 ml sont collectées.

### **2.7.4. Chromatographie liquide en phase inverse**

Les protéines ont été fractionnées par chromatographie liquide en phase inverse sur une colonne analytique C18 (Hypersil,  $250 \times 4,6 \text{ mm}$ ,  $5 \mu\text{m}$ , ThermoQuest, France). L'éluion est réalisée par un gradient linéaire de 0 à 70% de mélange B ( $\text{CH}_3\text{CN}$  / acide trifluoroacétique (TFA) 0,1%) dans le mélange A ( $\text{H}_2\text{O}$  / TFA 0,1%) en 35 min. Le débit est de  $1 \text{ ml.min}^{-1}$  et des fractions de 1 ml sont collectées. Les produits sont détectés en sortie de colonne par leur absorbance dans l'UV à 210 et 280 nm à l'aide d'un détecteur à barrettes de diodes (DAD 440, Kontron, France). Chaque fraction est ensuite amenée à sec par évaporation sous vide (Speed-Vac) et lavée deux fois avec 1 ml de méthanol 50% afin d'éliminer le TFA. Les résidus secs finaux sont repris dans une solution de CHAPS 0,2%.

## **2.8. Microséquençage**

Après électrophorèse SDS-PAGE et transfert sur membrane de PVDF, la bande d'intérêt est découpée et le microséquençage N-terminal du peptide est réalisé par la société Genosphere Biotech (France), selon la méthode de dégradation d'Edman. Un essai de coupure au bromure de cyanogène a été ensuite réalisé (coupure potentielle aux méthionines internes), suivi d'un autre cycle de microséquençage.



**Figure 2.3 : Les inhibiteurs de protéases du puceron.**

50 µg de protéines d'extrait total d'*A. pisum* (1) ou d'*A. gossypii* (3), ou 2 µl d'hémolymphe d'*A. pisum* (2) sont séparées sur un gel SDS-PAGE contenant de la gélatine. Le gel est ensuite incubé avec de la papaïne (a) ou avec de la trypsine (b) pendant 4h à 37°C. Après révélation au bleu de Coomassie, les IP apparaissent sous forme de bandes foncées épaisses sur fond clair et sont encadrés de rouge (les bandes très fines correspondent aux protéines de l'extrait non digérées par la protéase). Les témoins positifs correspondent à 20 ng d'oryzacystatine (4) et 2 µl d'hémolymphe de *Spodoptera litoralis* (5).

## Résultats et Discussion

### 3.1. Mise en évidence d'une activité inhibitrice de protéases à cystéine et à sérine chez le puceron

La faisabilité du projet repose d'abord sur la mise en évidence préalable d'une activité inhibitrice de protéase chez le puceron. L'hémolymphe contient, comme le sérum des vertébrés, différents polypeptides capables d'inhiber l'activité protéolytique des enzymes (Eguchi, 1993). L'activité IP a donc été recherchée sur gel d'activité, permettant de révéler les différentes protéines selon leur spécificité pour une protéase et leur masse moléculaire.

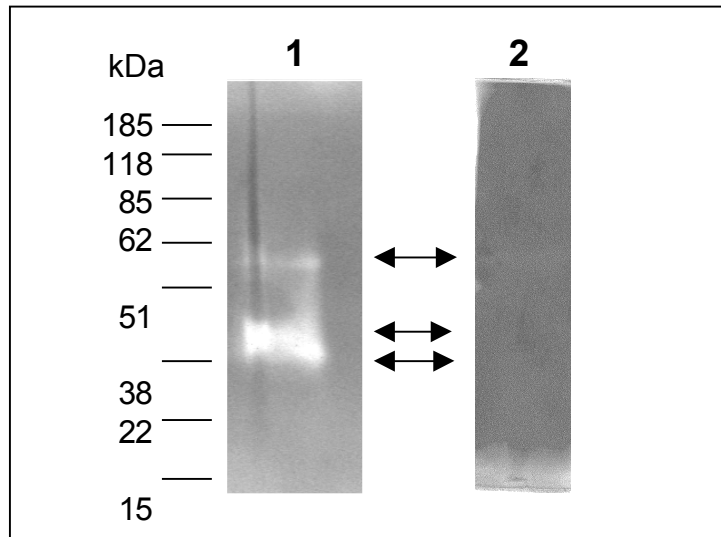
Un extrait de protéines solubles totales dans du NaCl 0,15 M et un extrait d'hémolymphe de puceron ont été utilisés afin de mettre en évidence, sur gel, la présence d'inhibiteurs de protéases à cystéine ou à sérine. Le zymogramme, présenté en figure 2.3a (activité anti-papaïne), révèle une activité anti-protéase à cystéine dans l'hémolymphe de puceron. Cette molécule présente une masse moléculaire apparente de 11 kDa.

Un inhibiteur de protéases à sérine, de masse moléculaire différente du peptide présentant l'activité antipapaïne a aussi été détectée dans l'hémolymphe de puceron (figure 2.3b). L'IP à sérine présente une masse moléculaire apparente de 13 kDa. Lorsque le gel a été incubé avec de la chymotrypsine, aucune bande spécifique d'une activité IP n'a été révélée dans l'hémolymphe de puceron (résultats non présentés). L'activité IP à sérine détectée semble donc spécifique de la trypsine.

Chez le puceron *A. pisum*, il circule donc dans l'hémolymphe au moins un IP à cystéine et un IP à sérine. La présence d'inhibiteur de protéase à sérine a été révélée chez de nombreux insectes et la plupart d'entre eux appartiennent au type serpine ou Kunitz (pour revue : Polanowski, 1996). En se basant sur sa masse moléculaire, l'IP à sérine du puceron de 13 kDa pourrait s'apparenter à ceux de la famille des Kunitz.

Par contre, la présence d'inhibiteurs de protéases à cystéine est beaucoup moins étayée. Chez le puceron, cette protéine présentant aussi une faible masse moléculaire (11 kDa) pourrait appartenir à la famille des cystatines (Brown, 1997).

D'autre part, nous n'avons pas détecté ces activités IP dans le tube digestif (résultats non présentés), ces molécules semblent donc présenter une spécificité tissulaire.



**Figure 2.4 : Inhibition des protéases du tube digestif par l'hémolymphe.**

Les protéines de deux tubes digestifs sont séparées en SDS-PAGE puis électrotransférées sur un gel contenant de la gélatine. Le gel est incubé toute la nuit dans un tampon phosphate de sodium (0,1 M pH 4,5 ; DTT 2 mM) (1) ou est incubé au préalable 1h avec 5  $\mu$ l d'hémolymphe (2). Le gel est ensuite coloré au bleu de Coomassie, l'activité protéasique apparaît sous forme de bandes blanches signalées par une flèche.

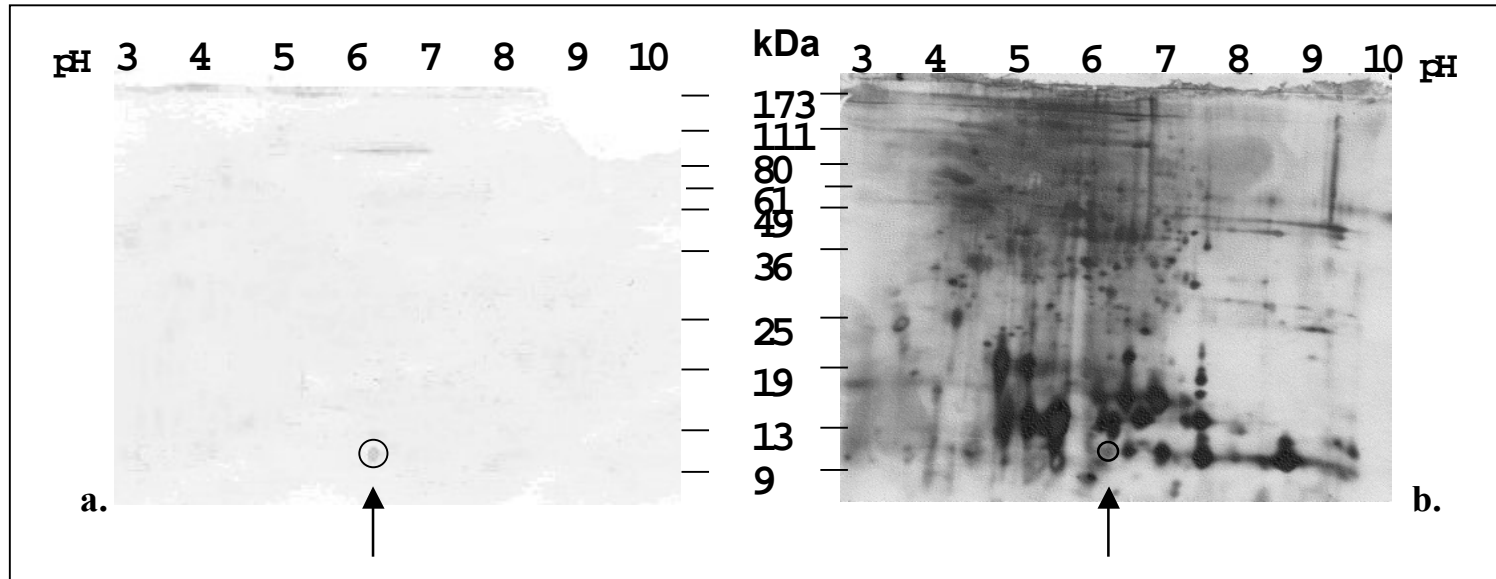
### 3.2. Inhibition de l'activité protéolytique du tube digestif par l'hémolymphe

Dans l'optique de l'utilisation potentielle des séquences d'inhibiteurs de protéases isolées du puceron pour lutter contre le puceron *A. gossypii* en transgénèse, il était indispensable de vérifier que les IP hémolympatiques étaient actifs contre les protéases digestives du puceron cible *A. gossypii*.

L'activité inhibitrice de l'hémolymphe sur les protéases digestives du puceron *A. gossypii* a donc été vérifiée. Après séparation des protéines représentatives de deux tubes digestifs en SDS-PAGE et électrotransfert sur gel d'activité, les gels contenant les protéases du tube digestif ont été incubés 1 h à 4°C en présence de 5 µl d'hémolymphe.

Les zymogrammes obtenus indiquent clairement que l'activité protéasique présente dans le TD de puceron est totalement inhibée par la présence d'hémolymphe (figure 2.4). Une courbe dose-réponse réalisée par l'incubation de l'équivalent d'un TD avec 50 à 1000 nl d'hémolymphe a permis de définir une valeur d'IC50 approximativement de 100 nl (données non présentées).

Ces résultats montrent donc que l'activité protéolytique présente dans le TD de puceron peut être totalement inhibée par l'addition d'hémolymphe. Les protéases du TD sont représentées à plus de 95% par des protéases de type cystéine (article 1). Il est probable que cette inhibition soit due à la présence du peptide de 11 kDa, principal inhibiteur de protéase à cystéine détectable sur gel d'activité. Cet inhibiteur hémolympatique de protéase à cystéine de 11 kDa semble donc être un bon candidat dans la recherche d'IP actifs contre le puceron. Le clonage du gène codant cet IP a donc été entrepris, en développant tout d'abord une approche moléculaire.



**Figure 2.5 : Correspondance entre l'activité IP à cystéine et le peptide hémolymphatique sur électrophorèse en 2D**

- a. Les protéines de 5  $\mu$ l d'hémolymphe ont été séparées sur électrophorèse 2D. Le gel a été transféré sur un gel d'activité contenant de la gélatine. Le gel a ensuite été incubé 4 h à 37°C avec la papaïne puis coloré au bleu de Coomassie. L'inhibiteur apparaît sous forme de tache ronde (entouré de noir).
- b. La même quantité d'hémolymphe a été séparée sur électrophorèse 2D et les protéines ont été colorées au nitrate d'argent. La relation entre le spot d'activité (gel a) et le peptide correspondant est faite grâce aux coordonnées graphiques, le spot d'intérêt est entouré d'un cercle et indiqué par une flèche.

### 3.3. Essai de clonage du gène codant un inhibiteur de protéase à cystéine.

Après avoir vérifié la présence d'un IP à cystéine chez le puceron *A. gossypii* possédant un comportement électrophorétique identique à celui mis en évidence chez *A. pisum* (figure 2.3a), une banque d'ADNc d'*A. gossypii* total a été construite afin de cloner un gène de cystatine.

Afin d'obtenir la séquence nucléotidique de l'inhibiteur de protéase à cystéine, des amorces dégénérées ont été définies en tirant profit de l'homologie des séquences de cystatines au niveau des boucles inhibitrices QxVxG et PW. Ces amorces ont été utilisées pour cribler une banque d'ADNc d'*A. gossypii* total. Dix bandes discrètes ont été clonées et séquencées. Bien que le témoin positif (ADN génomique de drosophile (Delbridge, 1990)) ait donné une bande discrète de taille attendue (104pb), aucune séquence correspondant à un inhibiteur de protéase à cystéine n'a été isolée de la banque d'ADNc. Des expériences de RT-PCR ont aussi été réalisées à partir des ARN totaux d'*A. gossypii* mais, à nouveau, le séquençage de quatre bandes discrètes n'a pas permis le clonage d'une séquence homologue à un IP à cystéine.

Il est probable que les messagers correspondant à cette protéine soit très faiblement représentés dans le pool d'ARN messagers. D'autre part, la séquence de la cystatine du puceron peut avoir beaucoup divergé au niveau nucléotidique par rapport aux séquences de drosophile et de *S. peregrina* (figure 2.2).

L'approche moléculaire s'étant révélée infructueuse, la purification du peptide par voie biochimique a donc été entreprise.

### 3.4. Purification du peptide à activité IP-cystéine.

#### 3.4.1. A partir de l'hémolymphe

Afin de pouvoir micro-séquencer cette protéine, une électrophorèse bidimensionnelle a été réalisée avec 10 µl d'hémolymphe. Afin de repérer le spot à prendre en compte, un gel miroir a été électrotransféré sur un gel d'activité (figure 2.5). Après digestion enzymatique, un seul spot est révélé (figure 2.5a). Il correspond à une protéine de masse moléculaire 11 kDa et présentant un pI de 6,5. La correspondance de ce spot d'activité sur le gel coloré au nitrate d'argent permet de localiser un spot cible (figure 2.5b).



Malgré la réduction et l'alkylation irréversible des cystéines subies par les protéines préalablement à l'électrophorèse en SDS-PAGE, l'IP est toujours actif. La séquence de la cystatine qui a été étudiée chez la drosophile, ainsi que celle de *S. peregrina* possèdent deux cystéines dans la région C-terminale de la protéine. Ces résidus peuvent former un pont disulfure. Le rôle des ponts disulfures dans le maintien de l'activité inhibitrice des cystatines a été étudié sur le modèle que représente la cystatine de poulet (Björk, 1992). Cette protéine renferme deux ponts disulfure dans la région C-terminale. La réduction ménagée par le DTT entraîne l'ouverture d'un pont disulfure mais aucune perte d'activité. Par contre, la réduction irréversible après alkylation des deux ponts disulfure induit un changement de conformation de la protéine plus important et résulte en une perte de son activité inhibitrice de 40%. Il est ainsi possible que malgré la réduction du pont disulfure de l'inhibiteur de protéase à cystéine du puceron (putativement identifié comme une cystatine), ce dernier soit suffisamment actif pour permettre l'inhibition de la papaïne et ainsi sa localisation sur le gel d'électrophorèse 2D.

L'intensité de coloration du spot correspondant à l'IP à cystéine ne permet pas d'envisager un microséquençage de ce peptide (figure 2.5b). Une quantité plus importante d'hémolymphe (12,5  $\mu$ l) a donc été déposée, mais la résistance de l'échantillon étant trop élevée, aucune migration n'a pu être effectuée. Il est possible que la trop grande concentration d'un composant fait qu'il cristallise dans ces conditions.

Au vu de l'impossibilité d'utiliser plus de matériel pour isoler cette protéine en électrophorèse 2D et de la difficulté à collecter de l'hémolymphe de puceron, nous avons fait le choix de purifier cet IP à partir du puceron entier.

### **3.4.2. A partir du puceron entier**

#### *3.4.2.1. Obtention d'un extrait enrichi*

La première étape de l'étude a consisté à rechercher un tampon qui permette d'extraire le maximum d'activité antipapaïne du puceron en préservant la structure de l'inhibiteur. Les protéases du tube digestif du puceron possèdent une affinité pour l'IP. Les cystatines sont des inhibiteurs réversibles et compétitifs mais leur affinité pour leur enzyme cible peut être suffisamment élevée pour que ces propriétés soient difficiles à démontrer (Barett, 1987). Nous avons ainsi cherché à limiter l'interaction entre l'IP et les protéases du puceron qui peut avoir lieu lors de la destruction des compartimentations tissulaires et cellulaires.

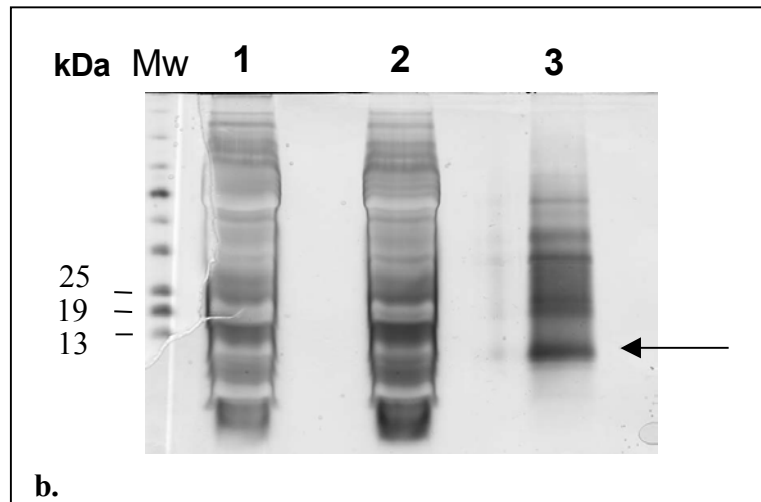
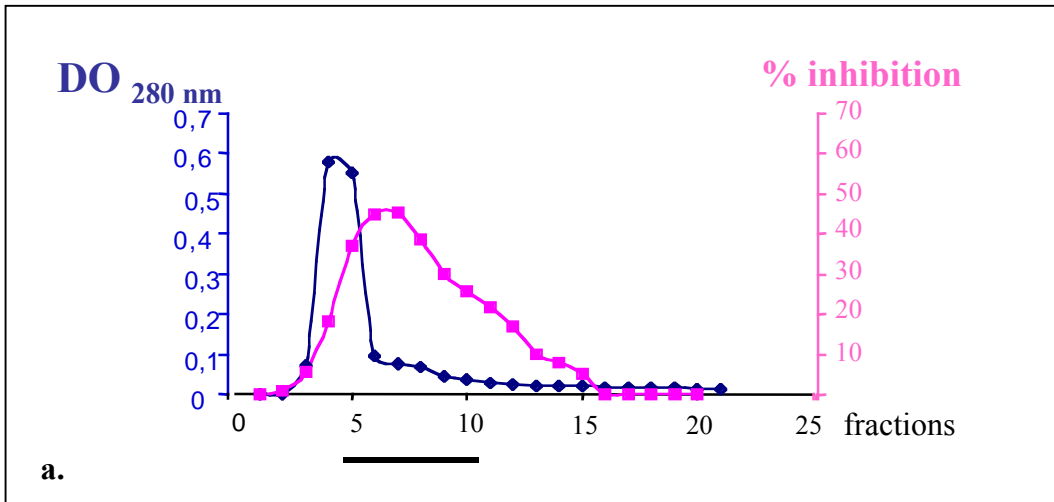
Les pucerons sont broyés dans un mélange d'inhibiteur de protéases dont l'E64 (inhibiteur spécifique des protéases à cystéine qui se fixe comme les cystatines dans le site actif). L'extrait protéique est passé sur une membrane d'ultrafiltration avec une limite de 30 kDa afin de séparer l'IP des protéases, puis l'extrait est passé sur une membrane d'ultrafiltration présentant un seuil de coupure de 5 kDa afin d'éliminer le E64 libre restant.

Cependant, un recouvrement de l'activité de 24% uniquement est observé dans le rétentat de la seconde ultrafiltration, l'ensemble des autres fractions obtenues ne présentant qu'une activité résiduelle de 2,8%. De plus, la perte en protéines totales est de l'ordre de 55%. Il semble qu'un grand nombre de protéines soient restées fixées sur les membranes. L'utilisation de la dialyse (avec une membrane présentant un seuil de coupure de 5 kDa) permet de récupérer presque complètement les protéines totales, mais on observe une récupération de l'ordre de 200% de l'activité IP-cystéine. Il semble donc que la dialyse, même extensive, de l'extrait ne permette pas d'éliminer l'E64 (en large excès par rapport au peptide à activité IP). Cette approche par extraction en présence de E64, même si elle semble permettre la récupération du peptide d'intérêt et la suppression d'interactions avec les protéases endogènes, a donc cependant dû être écartée.

Nous avons donc cherché ensuite à séparer IP et protéases par précipitation différentielle, sachant que les cystatines, au contraire des protéases, sont de petites molécules stables. Les pucerons sont donc broyés selon différentes conditions :

- dans un tampon phosphate. Après centrifugation, le surnageant est amené à une concentration finale de 50% de méthanol (E1),
- dans 100% de méthanol. Après centrifugation, le culot est solubilisé dans 50% de méthanol et le surnageant obtenu est conservé (E2),
- dans 50% de méthanol. Après centrifugation, le surnageant obtenu est conservé (E3).

La fraction E3 présente l'activité IP la plus élevée (tableau 2.2). L'inhibiteur de protéase est donc soluble dans le méthanol à 50%. Dans les fractions E1 et E2, le rendement est de 60 et 19,7% respectivement, il est possible que des molécules d'IP soient entraînées dans le culot avec leurs protéases affines insolubles dans le méthanol. D'autre part, la diminution de la concentration en méthanol (E2) ne permet pas la récupération de toute l'activité antipapaïne. La fraction E3 présente une activité spécifique comparable à celle obtenue avec le puceron broyé dans le NaCl 0,15 M, et permet la récupération effective de l'activité IP-cystéine. Le



**Figure 2.6 : Purification de l'inhibiteur de protéase à cystéine par chromatographie d'affinité**

a. Profil d'éluion des protéines affines pour la papaïne détectées à 280nm (en bleu). Un test d'activité IP à cystéine est ensuite réalisée sur chaque fraction récoltée et est reportée en rose (% inhibition de la papaïne)

Les fractions 4 à 10 ont été regroupées (barre noire).

b. Electrophorèse en SDS-PAGE des fractions séparées par affinité, colorées à l'argent. Mw : marqueur de masse moléculaire ; (1) 1/250<sup>ème</sup> de l'extrait soluble dans le méthanol 50% ; (2) 1/250<sup>ème</sup> de la fraction « lavage » ; (3) 1/10<sup>ème</sup> de la réunion des fractions 4 à 10 affines pour la papaïne. Le peptide de 11 kDa est indiqué par une flèche.

rendement de 130%, supérieur donc au témoin, peut s'expliquer par une meilleure récupération de l'IP, par la suppression des interactions avec les protéases dans le méthanol. La solubilisation de l'IP directement dans le méthanol 50% a donc été choisie comme méthode d'extraction.

**Tableau 2.2 : Effet des conditions d'extraction sur le recouvrement de l'activité antipapaïne.**

Conditions d'extraction des protéines de puceron	Activité antipapaïne (PUI)	Rendement (%)	Activité spécifique (PUI.μg <sup>-1</sup> de protéines)
NaCl 0,15 M (témoin)	55	100	0,17
E1	33,2	60	0,19
E2	10,8	19,7	0,18
E3	71,5	130	0,16

100 mg de pucerons ont été broyés dans 1ml de tampon phosphate pH 6,5 0,1 M. L'activité spécifique obtenue dans ces conditions est utilisée comme référentiel. E1 représente la fraction soluble obtenue après précipitation au méthanol 50%. E2 représente la fraction soluble obtenue après solubilisation dans le méthanol 50% du culot résultant du broyage des pucerons dans le méthanol 100%. E3 représente la fraction soluble obtenue après broyage des pucerons dans le méthanol 50%.

#### 3.4.2.2. Première étape de purification.

La chromatographie d'affinité utilisant les protéases cibles couplées à un support solide a souvent été employée pour la purification d'inhibiteurs de protéases. Avec cette méthode, des extraits bruts ou partiellement purifiés sont utilisés.

La fraction soluble dans le méthanol 50% est fractionnée selon son affinité pour la papaïne. La fraction éluée présentant une activité anti-papaïne maximale, révèle un enrichissement pour un polypeptide de 11 kDa (figure 2.6). Cet extrait a été analysé par spectrométrie de masse (ES/MS) au laboratoire de technologie de l'INRA de Rennes. Toutefois, son analyse en MS s'est révélée infructueuse à cause de la présence de pigments.

Une ou des étapes supplémentaires de purification sont donc nécessaires avant d'envisager l'obtention de l'inhibiteur de protéase à cystéine. Afin d'éliminer la contamination de l'extrait par des composants non-protéiques empêchant l'analyse du peptide, différentes techniques ont été mises en oeuvre dans le but d'obtenir une fraction enrichie en protéines et non d'obtenir une purification en terme de protéines.

♦ *Essai de clarification par des méthodes physico-chimiques*

Afin d'éliminer les composés non-protéiques (pigments, lipides, sucres complexes...) de la fraction soluble dans le méthanol 50%, différents solvants ont été utilisés dans le but de faire précipiter les protéines et d'éliminer dans le surnageant des molécules beaucoup plus apolaires. L'adsorption des pigments a aussi été tentée grâce à une matrice active (Bio-beads SM adsorbents, BioRad). Les billes de polymères non polaires, macroporeuses, permettent l'adsorption de composés organiques d'un poids moléculaire inférieur à 2 kDa. Une séparation des composés non-protéiques de l'activité IP a aussi été testée en s'appuyant sur leur différence d'hydrophobicité sur une mini-colonne C18.

La précipitation au chloroforme, à l'acétone 80% ou au benzène n'a pas permis d'éliminer les pigments de la fraction IP. L'utilisation des Bio-beads a permis d'éliminer les pigments de la phase aqueuse, mais aucune activité IP n'est retrouvée dans cette fraction.

La fraction soluble dans le méthanol 50%, est déposée sur une colonne C18 Sep-Pak. Le lavage de la phase stationnaire par l'eau permet d'éliminer des pigments en quantité importante. Après élution de la colonne avec du méthanol 50%, toute l'activité IP est recouvrée dans cette fraction. Bien que la solution révèle encore la présence de pigments, une part importante en a toutefois été éliminée.

Ainsi, avec l'extrait obtenu, il est désormais possible d'envisager des techniques de chromatographie.

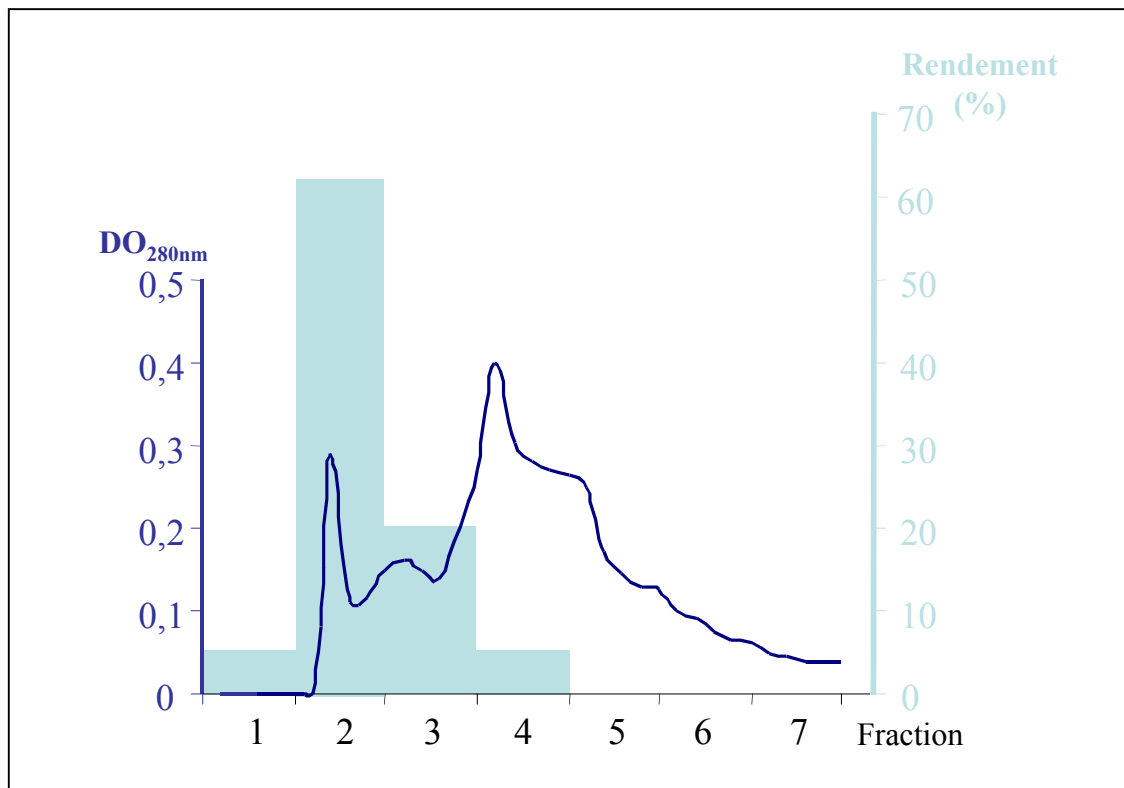
♦ *Essai de purification par chromatographie d'échange d'ions.*

La charge électrostatique intrinsèque aux molécules est un caractère utilisable pour les séparer. La fraction obtenue après séparation sur Sep-Pak est chargée sur différentes colonnes échangeuses d'ions à pH 8 (tableau 3).

La séparation des protéines sur colonne d'échange de cations permet le recouvrement de 62% de l'activité dans la fraction des protéines non retenues. Cependant, la totalité des pigments co-purifient avec l'activité IP.

Après séparation de molécules sur colonne d'échange d'anions, l'activité IP est retrouvée dans la fraction éluée avec 0,3 M NaCl mais elle ne représente que 25% de l'activité initiale, un très faible taux de récupération du peptide ne permet pas une purification effective. De plus, des pigments sont encore présents dans cette fraction.

L'accrochage du peptide sur colonne d'anions mais pas sur colonne de cations à pH 8 est en accord avec le pI de l'inhibiteur précédemment déterminé expérimentalement par



**Figure 2.7 : Séparation des molécules de l'extrait par filtration sur gel (Bio gel P4)**

Profil d'éluion des protéines en sortie d'une colonne de chromatographie d'exclusion de taille (seuil de coupure de 0.8 à 4 kDa), suivies par leur absorbance à 280 nm (courbe bleue).

Un test d'activité IP à cystéine est ensuite réalisée sur chaque fraction récoltée, et le recouvrement de l'activité IP à cystéine dans les différentes fractions est indiquée en vert (en % de l'activité injectée).

électrophorèse (6,5). Néanmoins, aucune purification réelle du peptide n'ayant été obtenue, la chromatographie d'échange d'ions n'a donc pas été retenue.

**Tableau 2.3 : Séparation des molécules sur colonnes d'échange d'ions à pH 8**

Méthode de fractionnement	Fractions	Rendement en activité IP (%)	Présence des pigments
Echange d'anions	Non retenue	0	+
	0,1 M NaCl	6	+
	0,3 M NaCl	25,6	++
	0,8 M NaCl	2,7	+
	1,5M NaCl	0	+
Echange de cations	Non retenue	62,6%	+++
	0,1 M NaCl	6	+
	0,3 M NaCl	0	-
	0,8 M NaCl	0	-
	1,5M NaCl	0	-

L'appréciation de la présence de pigments est donnée par l'observation visuelle de la couleur présente dans chaque fraction. L'activité enzymatique obtenue à partir de l'extrait Sep-Pak est utilisée comme référentiel. (-) aucun pigment visible, (+) légère, (+++) intense coloration de la solution.

♦ *Essai de purification par filtration sur gel.*

Nous avons cherché à éliminer les petites molécules telles que les pigments. L'extrait protéique contenant l'activité IP a ainsi été fractionné par chromatographie d'exclusion de taille, sur une matrice permettant de séparer des molécules de 0,8 à 4 kDa. Les fractions obtenues sont dessalées sur Sep-Pak avant d'effectuer le test enzymatique, des essais préliminaires ayant montré que l'activité papaïne était sensible aux sels résultant de la filtration sur gel. Après élution et dessalage, 80% de l'activité sont retrouvés dans les fractions 3 et 4 correspondant au volume mort de la colonne (figure 2.7), un comportement attendu puisque le peptide présente un poids moléculaire de 11 kDa, donc supérieur au seuil de coupure du gel utilisé. La perte de rendement est due à l'élimination de fractions en fin de pic présentant une faible activité. Si cette étape de sélection des molécules en fonction de leur poids moléculaire ne permet pas de purifier le peptide d'un point de vue protéique (la presque totalité des protéines de l'extrait se retrouvant dans le volume mort avec l'activité IP),

l'élimination totale des derniers pigments et des petites molécules constitue cependant une avancée assez importante pour que cette étape de filtration moléculaire soit conservée.

♦ *Essai de purification par dénaturation à la chaleur*

Une des caractéristiques qui différencie les inhibiteurs de protéases d'insectes de ceux des mammifères est leur résistance à la chaleur (Eguchi, 1993). Cette propriété est également une caractéristique générale des cystatines végétales. Après avoir éliminé les pigments, l'extrait est chauffé à 100°C pendant 10 minutes. L'élimination des molécules thermolabiles n'est effectuée qu'après avoir éliminé tous les pigments afin de limiter les possibilités de liaisons entre les pigments et l'IP à cystéine. Après centrifugation, toute l'activité anti-papaïne est présente dans le surnageant. Ce traitement n'entraîne aucune perte d'activité. Ainsi, avec un extrait composé désormais majoritairement de protéines, l'étape finale de purification peut être envisagée.

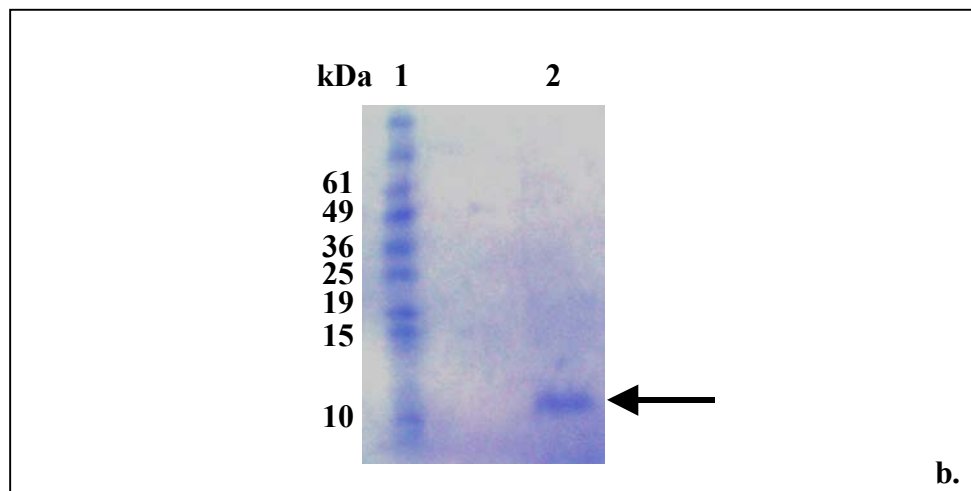
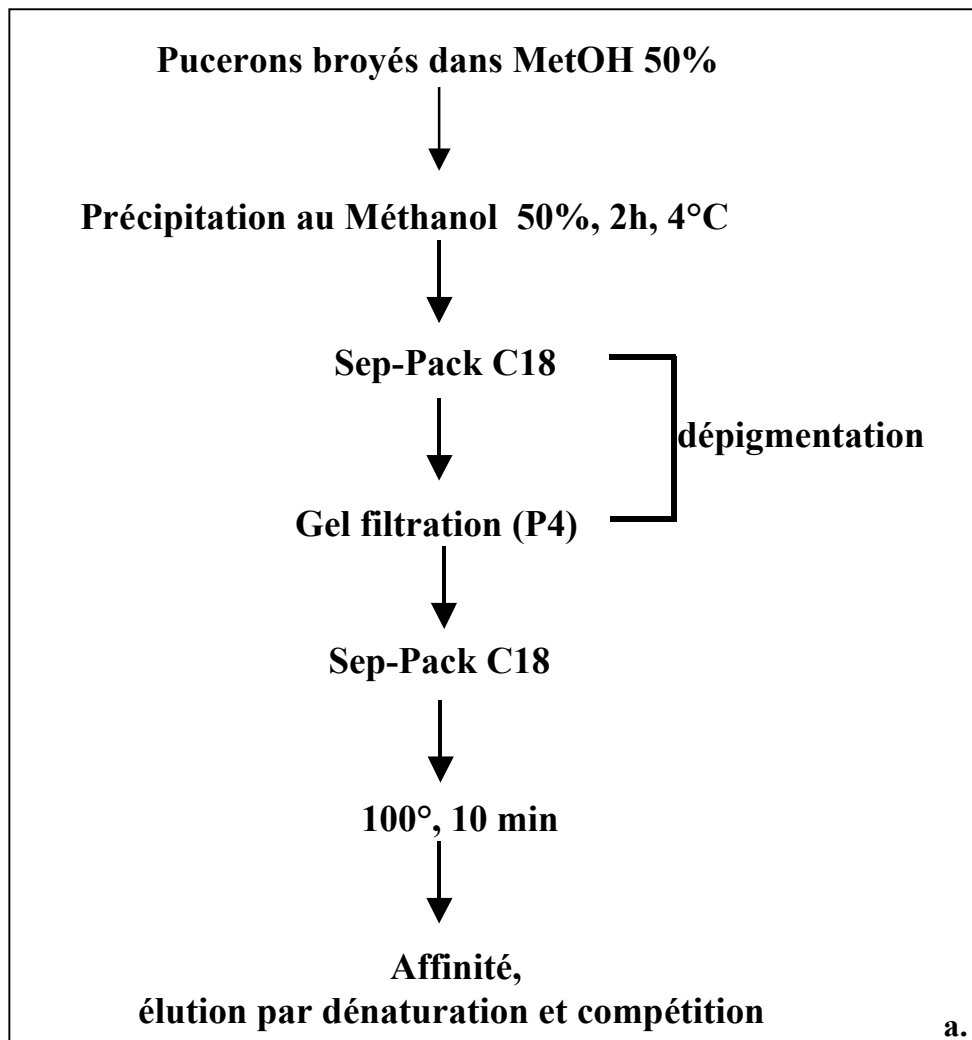
*3.4.2.3. Dernière étape de purification*

♦ *Par CLHP*

L'IP à cystéine a révélé une certaine apolarité puisqu'il est retenu sur la mini-colonne Sep-Pak C18. Afin d'obtenir l'inhibiteur de protéase à homogénéité, les molécules ont été séparées en fonction de leur polarité sur une colonne greffée en C18. Les différents tests préliminaires ont pu montrer que l'activité papaïne était sensible à la concentration de TFA (0,1%) utilisée au cours de la CLHP. Ainsi, après élution, le TFA est éliminé de chaque fraction grâce à deux lavages au méthanol 50%. Toutefois, aucune activité IP substantielle n'est recouvrée après cette séparation.

Il est couramment observé, au cours de la purification de protéines, de fortes pertes lors de l'étape finale du processus. Ces pertes souvent dues à l'adsorption de la protéine hydrophobe d'intérêt, une fois pure sur la verrerie utilisée. Devant la nature potentiellement apolaire de l'IP (solubilité dans le méthanol 50%), les fractions issues de la CLHP ont dans un second temps été reprises dans une solution contenant un détergent. Après des tests d'activité témoins effectués en présence de différents détergents, les fractions sont reprises dans du CHAPS 0,2%. Mais à nouveau aucune activité n'a pu être mise en évidence, suggérant une perte de fonctionnalité du peptide au cours de la chromatographie plutôt qu'une perte du peptide lui-même.





**Figure 2.8 : Purification à homogénéité d'un peptide de 11 kDa**

a. Schématisation du mode opératoire à partir de pucerons entiers

b. Electrophorèse SDS-PAGE de la fraction affine pour la papaine. (1) marqueur de poids moléculaire, (2) fraction affine pour la papaine

Le peptide de 11 kDa (indiqué par la flèche) a été ensuite microséquencé.

♦ *Par affinité*

La chromatographie d'affinité a souvent permis l'isolement d'inhibiteurs de protéases avec un taux de purification élevé. L'extrait est déposé sur une colonne d'affinité composée d'une matrice de papaine carboxyméthylée. Après lavage, le déplacement des sites est effectué pendant 1h à température ambiante dans un tampon constitué de E64 10 $\mu$ M ; Tris 20 mM pH 10,3 ; DMSO 20% puis les protéines affines sont éluées de la colonne.

Après élution, les fractions obtenues sont déposées sur électrophorèse SDS-PAGE afin de contrôler la présence de l'IP. Après coloration au bleu de Coomassie (figure 2.8), une seule bande est observée. Cette protéine présente une masse moléculaire apparente de 11 kDa, ce qui est en accord avec les résultats obtenus précédemment.

La séparation des molécules, issues du broyage des pucerons dans le méthanol 50%, sur minicolonne en phase inverse puis sur gel filtration permet d'enrichir en protéine et d'éliminer toutes les molécules interférentes avec l'analyse de l'activité IP à cystéine. La purification de l'activité IP à cystéine a ensuite été envisagée selon plusieurs techniques mais seule l'affinité a permis d'obtenir un peptide de 11 kDa à homogénéité.

### **3.4.3. Microséquençage**

Après électrotransfert des protéines du gel SDS-PAGE obtenu après chromatographie d'affinité sur membrane de PVDF, la bande d'intérêt est découpée et le séquençage effectué par la société Génosphère Biotech en vue d'obtenir la séquence N-terminale du peptide. Les quantités molaires d'acides aminés successives obtenues au cours du séquençage ont été très faibles, malgré l'envoi d'un spot significativement coloré au bleu de Coomassie. Une séquence de douze acides aminés a tout de même été obtenue : (A/E/S) A E A S L x K G Q x L.

Les alignements réalisés avec cette séquence contre le génome de la drosophile révèlent quelques hits faiblement significatifs avec des ORF de petit poids moléculaire, dont :

- 1 : une protéine de 139 acides aminés (Q9VG63), homologue d'un inhibiteur de phosphatase (I-T protein) étudié chez la drosophile (Q9VGJ2). Le hit ne se trouve pas en région N-terminale de cette protéine (segment 89-100) et présente 75% d'homologie (50% d'identité).

```
query :      AAEASLxKGQxL
Q9VG63 :    089AAEFTITRGQKL100
Q9VGJ2 :    125EAEFTICKGHKL135
```

- 2 : une protéine de 119 acides aminés présentant une faible homologie si les x représentent des cystéines. C'est un peptide de la famille des CHH (Crustacean Hyperglycemic Hormone).

```
query :      AAEASLCKGQCL
Q95SV5 :    66TSIHRLCKKDCF77
```

Toutefois le spectre obtenu après hydrolyse ne reflète qu'une quantité infime de protéines, en regard de l'intensité de la coloration au bleu de Coomassie (figure 2.8). La protéine majoritaire pourrait donc présenter un N-terminal bloqué. Des coupures de la protéine au CNBr ont donc été effectuées afin d'obtenir des peptides internes mais sans succès. Si dans le mélange co-existent deux espèces protéiques (une minoritaire, séquençée, et une majoritaire, bloquée en N-terminale), ces deux protéines ne semblent donc pas présenter de méthionines accessibles.

Ces résultats ne permettent malheureusement pas de lever l'incertitude sur la nature de l'activité IP à cystéine : est-ce que ce peptide de 11 kDa présente des homologies avec un peptide présent dans l'hémolymphe des arthropodes dont la fonction a dérivé à partir des fonctions des deux protéines décrites ci-dessus (la fonction hormonale ou la fonction inhibitrice d'une autre enzyme), ou est-ce une cystatine ou une protéine inhibitrice d'une autre famille, bloquée en N-terminal ?

## **Conclusions**

Cette étude montre que l'hémolymphe de puceron contient un inhibiteur de protéase à sérine (trypsine) et un inhibiteur de protéase à cystéine. Ces deux molécules présentent des masses moléculaires apparentes de 13 et 11 kDa respectivement. Elles semblent majoritairement présentes dans l'hémolymphe.

L'inhibiteur de protéase à sérine pourrait remplir le rôle de régulateur des voies d'immunoréactivité humorale comme cela a été montré chez d'autres insectes.

Peu d'inhibiteurs de protéases à cystéine ont été mis en évidence dans l'hémolymphe d'insecte et leur rôle n'est pas très développé dans la littérature. Nos résultats montrent que les IP de l'hémolymphe sont actifs sur les protéases digestives du puceron. Or l'activité protéolytique du tube digestif est due à une activité protéase à cystéine. C'est pourquoi l'IP à cystéine nous a particulièrement intéressé. Bien que la séquence ne soit pas encore disponible, les différentes stratégies que nous avons développées pour purifier cet IP, ont révélé quelques caractères physico-chimiques particuliers de cette protéine. C'est un peptide qui présente une masse moléculaire de 11 kDa et qui s'avère être très stable. Ce peptide est résistant à la chaleur, c'est une protéine soluble et relativement apolaire puisqu'elle ne précipite pas dans le méthanol 50%.

Récemment, un inhibiteur de protéase à cystéine a été isolé de l'hémolymphe du ver à soie (Yamamoto, 1999 a). C'est également une protéine de petit poids moléculaire (10,5 kDa) et résistante à la chaleur. Sa particularité est qu'elle présente une faible homologie avec les cystatines, mais la séquence de ce peptide est très homologue aux proséquences de plusieurs précurseurs de protéases à cystéine (Yamamoto, 1999 b).

On ne peut pas écarter la possibilité de l'existence d'une telle protéine dans l'hémolymphe du puceron. Elle expliquerait en outre l'échec des PCR recherchant des homologues de cystatines dans la banque de puceron. C'est pourquoi, une nouvelle préparation de l'inhibiteur de protéase à cystéine est nécessaire afin d'effectuer un microséquençage interne (après digestion trypsique par exemple, si elle est effective). De nouveaux oligonucléotides pourront être définis afin de cloner l'ADNc. Le peptide pourra être exprimé en système hétérologue et son pouvoir entomotoxique contre les pucerons pourra être évalué en nutrition artificielle, comme nous l'avions prévu initialement.

## **CHAPITRE III**

### **– ETUDE DU MODE D’ACTION & MODIFICATION DE PsTI –**

**Tableau 3.1 : Les familles et quelques exemples d'inhibiteurs de protéases à sérines chez les plantes**

T, trypsine ; C, chymotrypsine ; E, élastase

Famille	Inhibiteurs types	Enzyme inhibée (nombre de sites)	Taille (kDa)
<b>Kunitz</b>	Soybean Trypsin Inhibitor (SBTI)	T (1)	21
	Windged Bean Trypsin Inhibitor (WTI)	T (1)	19
	Mustard Trypsin Inhibitor	T (1)	6,5
<b>Bowman-Birk</b>	Soybean Bowman-Birk Inhibitor (BBI)	T (1)/C (1)	8
	Pisum sativum Trypsin Inhibitor	T (1)/C (1)	7,5
	Wheat Germ Trypsin Inhibitor I (WGI-I)	T (2)	14,5
	Limabean Trypsin Inhibitor (LBTI)	T (1)/ C (1)	9
	Cowpea Trypsin Inhibitor (CpTi)	T (1)	8,5
	C-II	T (1)/C (1)/E (1)	8
<b>Barley trypsin inhibitor</b>			
<b>Potato Inhibitor-I</b>	Potato Inhibitor-I (PPI-I)	C (2)	8
	Englin c	E (1)	8,1
<b>Potato Inhibitor-II</b>	Potato Inhibitor-II (PPI-II)	T (1)/C (1)	10,5
<b>Squash inhibitor</b>			
	<i>C. maxima</i> Trypsin Inhibitor III (CMTI-III)	T (1)	3
	Pumkin Fruit Proteinase Inhibitor (PFPI)	T (1 ?)	< 6

## **Introduction**

Parmi les métabolites synthétisés par la plante pour se défendre contre les insectes phytophages, les protéines sont beaucoup étudiées pour leur toxicité et leur potentiel biotechnologique, car elles sont plus facilement et directement manipulables que des voies de biosynthèse d'autres catégories de métabolites toxiques.

La recherche de protéines entomotoxiques est depuis quelques années un des points clés pour l'élaboration de stratégies de lutte contre les insectes *via* la transgénèse. De nombreuses études reposent sur l'utilisation des inhibiteurs de protéases à sérine.

### **1.1 Les inhibiteurs de protéases à sérine chez les plantes.**

Les inhibiteurs de protéases présents dans les organes de réserves jouent un rôle important dans la détermination de leur valeur nutritionnelle. De nombreuses recherches en nutrition humaine et animale, en physiologie végétale, en biochimie et ingénierie des protéines ont été réalisées afin de déterminer les effets potentiels résultant de l'inhibition de diverses protéases de l'homme et des espèces domestiques. Les inhibiteurs de protéases de plantes les plus étudiés, et donc les mieux caractérisés, sont ceux qui sont responsables de l'inhibition des protéases à sérine (chymotrypsine et surtout trypsine), qui sont les protéases digestives majeures des mammifères (avec la pepsine).

Toutes les familles d'inhibiteurs de protéases à sérine isolés chez les plantes (Tableau 4.1) sont des inhibiteurs compétitifs et inhibent leurs protéases cibles selon le même mécanisme. Ils appartiennent à la super-famille des " inhibiteurs canoniques ".

### **1.2. Toxicité des inhibiteurs de la famille des Bowman Birk envers les pucerons**

Malgré l'importance agronomique que représente la présence de pucerons dans les cultures, les recherches de protéines présentant une toxicité contre les Homoptères sont peu nombreuses.

Un crible de protéines entomotoxiques réalisé avec le puceron du pois *Acyrtosiphon pisum* a permis de mettre en évidence que la plupart des inhibiteurs de protéases possèdent une très faible toxicité pour les pucerons (Rahbé, 1993 ; Rahbé, 1995). Toutefois, des molécules de la famille des Bowman Birk présentaient un effet délétère très intéressant sur cet insecte (Rahbé, 1993).

PsTI (*Pisum Sativum* Trypsin Inhibitor) est un inhibiteur de la famille des Bowman-Birk (Ferrasson, 1995). Dans la graine de pois, les inhibiteurs de trypsine sont représentés par 6 isoformes de PsTI (PsTI-1, 2, 3, 4a, 4b, et 5). Ils sont le résultat de l'expression d'au moins deux gènes/loci et d'une modification post-traductionnelle (clivage en C-terminal) qui engendre les différentes isoformes (Domoney, 1993 ; Quillien, 1997). Les IP de type PsTI sont caractérisés par 7 ponts disulfures et deux sites réactifs. La structure tridimensionnelle de ces inhibiteurs montre que les 2 sites réactifs sont situés aux extrémités de la protéine, les boucles inhibitrices peuvent réagir simultanément et indépendamment avec deux molécules de protéases (une trypsine et/ou une chymotrypsine) (Werner, 1991).

Les 6 isoformes de PsTI sont très toxiques pour le puceron du pois (au 7<sup>ème</sup> jour, LC50 = 48 µM (Rahbé, 2002)). Pour compléter ces données et comprendre le mécanisme de la toxicité inattendue de ces molécules, des peptides synthétiques correspondant aux boucles inhibitrices ont été synthétisés lors de l'étude citée. L'ingestion de la boucle antichymotrypsine à une concentration de 1 mg/ml engendre une très forte mortalité : après six jours d'alimentation, 100% des pucerons sont morts. La même quantité de peptide antitrypsine n'entraîne la mort que de 40% des pucerons. Il a donc été clairement démontré que la toxicité de cet inhibiteur reposait sur l'activité de la boucle antichymotrypsine.

Notre étude repose sur la toxicité de PsTI<sup>20</sup>. Nous avons dans un premier temps essayé de comprendre le mode d'action de cet inhibiteur de protéases chez le puceron. Nous avons ensuite décidé d'accroître éventuellement le pouvoir entomotoxique de cette protéine par mutagenèse dirigée et d'évaluer le potentiel de ces protéines pour leur utilisation en transgénèse.

---

<sup>20</sup> Les 6 isoformes de PsTI sont aussi actives envers le puceron, nous avons travaillé avec l'isoforme PsTI-2, purifiée par L. Quillien (INRA Nantes), que nous appellerons plus simplement PsTI dans cette étude.



## I. Caractérisation du mode de toxicité d'un inhibiteur de protéase à sérine, PsTI chez les pucerons *A. pisum* et *A. gossypii*.

### Introduction

Chez les insectes tel que les lépidoptères et les coléoptères, les inhibiteurs de protéases ingérés durant l'alimentation réagissent avec les protéases du tube digestif entraînant une perte de poids, un retard de développement et, dans les cas où l'IP est très actif, la mort des insectes.

Les pucerons, insectes phloémophages, se nourrissent sur un milieu pauvre en protéines et c'est pourquoi ils sont considérés comme ne possédant pas de protéases digestives. Les études préliminaires sur le puceron du pois, *A. pisum*, soutenaient cette hypothèse (Rahbé, 1995). Toutefois, l'étude que nous avons entreprise sur le puceron du melon, *A. gossypii*, démontre que les pucerons possèdent des activités endoprotéolytiques (chapitre I) ; aucune activité protéase à sérine n'a cependant été mise en évidence ni chez *A.gossypii* (article 1), ni chez *A.pisum* (v. ci-dessous).

Nous avons engagé une étude sur la toxicité paradoxale de PsTI sur le puceron du pois et le puceron du melon. Dans le cadre du programme européen, nous sommes en partenariat avec un sélectionneur de melon, il était donc indispensable de vérifier la toxicité de cet IP avant d'engager les expériences de transformation du melon. Le puceron du melon est un puceron polyphage qui s'est avéré être moins sensible à PsTI que l'espèce modèle *A. pisum*. Afin d'élucider le mode de toxicité de PsTI, nous avons recherché la présence de protéases à sérine, les cibles potentielles de PsTI, et étudié la localisation de l'inhibiteur chez les deux pucerons après ingestion, afin de mettre en évidence une différence éventuelle du mécanisme d'action de PsTI.

## Matériel et Méthodes

### 2.1. Origine et élevage des insectes

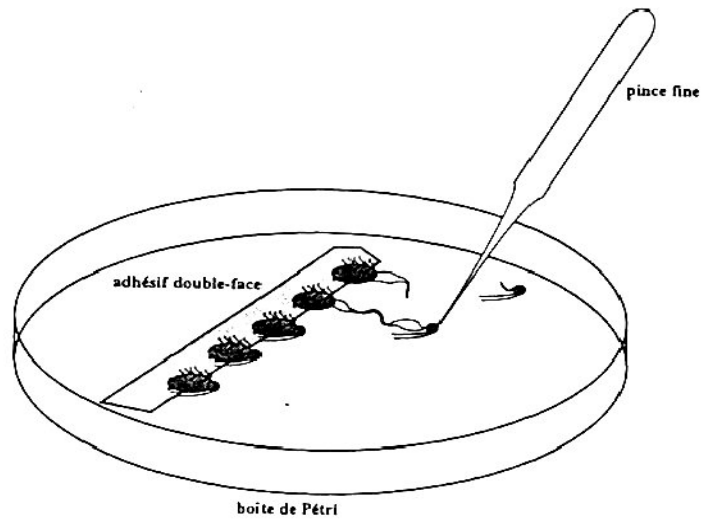
Avant chaque test, les populations d'aphides sont homogénéisées par synchronisation de leur production. Des adultes ailés matures sont transférés sur des jeunes plants. Après 48h, ils sont retirés et les larves issues de cette ponte sont laissées sur les plantes le temps de leur développement complet (4 stades larvaires, 7 à 9 jours suivant les espèces de pucerons), à faible densité de population. Les jeunes adultes aptères obtenus sont transférés sur de nouveaux plants et laissés à pondre 24 h. Les larves issues de cette ponte sont alors transférées sur les milieux artificiels à tester. Nous prélevons les larves les plus grosses que nous supposons âgées de 12 à 24h.

### 2.2. Origine de l'inhibiteur de protéase à sérine

La protéine PsTI nous a été fournie par le laboratoire de Biochimie et Technologie des protéines de l'INRA de Nantes. Elle provient de la graine de pois d'hiver (c.v. Frilene), et a été purifiée à partir de la farine par précipitation au sulfate d'ammonium, gel filtration, puis chromatographies échangeuses d'ions selon le protocole de purification décrit par E. Ferrasson (Ferrasson, 1997). La protéine est obtenue sous forme lyophilisée, sans sel et de pureté protéique supérieure à 90%. Le laboratoire de Biochimie et Technologie des protéines de l'INRA de Nantes nous a également fourni les anticorps polyclonaux anti-PsTI.

### 2.3. Construction et criblage d'une banque d'ADNc

De jeunes adultes aptères d'*Acyrtosiphom pisum* sont prélevés délicatement de leur plante hôte. Le tube digestif est disséqué rapidement dans une solution de NaCl 0,15M stérile et immédiatement congelé dans l'azote liquide, puis conservé à  $-80^{\circ}\text{C}$  jusqu'à leur utilisation. Une banque d'ADNc représentative des ARNm du tube digestif d'*A. pisum* a été construite. La procédure est identique à celle suivie pour la construction de la banque d'ADNc de tube digestif d'*A.gossypii* (article 1). Elle contient  $6 \times 10^5$  clones indépendants dont 81% sont



**Figure 3.1 : Schématisation de la dissection des tubes digestifs de pucerons.**

- (1) Les pucerons vivants sont fixés dans une boîte de Pétri sur un adhésif double face ;
- (2) ils sont recouverts d'une solution de NaCl 0,15M.
- (3) Extraction du tube digestif : la tête est arrachée, tout le tube digestif est étiré, l'œsophage est coupé juste au-dessus de l'estomac. A l'aide d'une pince fine, les tubes digestifs sont transférés avec précautions dans un tube contenant du NaCl 0,15M.

recombinants. Cette banque a été criblée avec des oligonucléotides dégénérés selon l'homologie des séquences de protéases à sérine des insectes : Trypf (TGG GTI GTI ACI GCI GCI CA(C/T) TG) correspond au domaine conservé autour de l'histidine 57<sup>21</sup> impliqué dans le site catalytique et l'oligonucléotide Trypr (A(A/G)I GGI CCI CCI (G/C)(A/T)(A/C) TCI CC) correspond au domaine conservé autour de la sérine 195 catalytique des protéases à sérine. Ces oligonucléotides nous ont été fournis par l'équipe de J. Gatehouse (Durham, Grande Bretagne).

#### **2.4. Caractérisation des activités protéolytiques du tube digestif du puceron du pois.**

Les tubes digestifs sont prélevés par dissection de jeunes adultes prélevés délicatement de leur plante hôte (figure 3.1). Ils sont aussitôt plongés dans 20 µl de NaCl 0,15M par lot de 20. Ils sont conservés à -20°C jusqu'à utilisation. Au moment de l'expérience, les tubes digestifs sont décongelés lentement dans la glace, puis homogénéisés au potter dans la glace.

Le profil protéolytique en fonction du pH ainsi que les zymogrammes des protéases du tube digestif ont été obtenus selon le protocole décrit dans l'article 1.

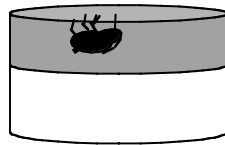
#### **2.5. Chromatographie liquide en phase inverse**

Vingt tubes digestifs sont broyés dans 50 µl de tampon phosphate de sodium (pH 7, 200mM). L'extrait obtenu est homogénéisé aux ultrasons pendant 5 minutes. Soixante quinze microgrammes de PsTI sont ajoutés à l'extrait dans un volume final de 100 µl. L'incubation est effectuée à 37°C pendant 24h. La réaction est arrêtée par l'ajout d'acide acétique à une concentration finale de 10%, ces conditions acides favorisant la dissociation du complexe.

Les protéines ont été fractionnées par chromatographie en phase inverse (CLPH) sur une colonne semi-préparative C4 (ThermoQuest, France). L'élution est réalisée par un gradient linéaire de 15 à 40% de mélange B (CH<sub>3</sub>CN/acide trifluoroacétique (TFA) 0,1%) dans le mélange A (H<sub>2</sub>O/TFA 0,1%) en 25 min. Le débit est de 1 ml.min<sup>-1</sup>. Les produits sont détectés

---

<sup>21</sup> Numérotation selon la séquence de la trypsine bovine



**Figure 3.2 Représentation schématique du dispositif permettant une alimentation des pucerons sur milieu artificiel.**

en sortie de colonne par leur absorbance dans l'UV à 210 et 280 nm (DAD 440, Kontron, France).

## 2.6. Composition et préparation des milieux artificiels

Le milieu artificiel Ap2 (annexe 1) est le milieu de base dans lequel les protéines à tester sont solubilisées. Ce milieu permet un développement optimal des pucerons sur une génération complète<sup>22</sup>. Les solutions contenant les protéines à tester sont préparées par dilution d'une solution mère à 1mg.ml<sup>-1</sup> stérilisée par filtration à 0,22 µm. Une série de 5 concentrations comprises entre 10 µg.ml<sup>-1</sup> et 800 µg.ml<sup>-1</sup> est préparée. Un volume de 70 µl de solution est ensuite déposé entre deux feuilles de Parafilm<sup>®</sup> tendues sur un support cylindrique en PVC (35 mm de diamètre, 20 mm de hauteur), en conditions stériles.

## 2.7. Principe des tests de toxicité *in vitro*

Au jour J<sub>0</sub>, des larves néonates sont déposées sur des milieux artificiels contenant la protéine purifiée aux concentrations 10 à 800 µg.ml<sup>-1</sup> (n=3 répétitions de 18 pucerons par concentration testée) (figure 3.2). Une heure après le dépôt, le pourcentage de larves fixées sur les milieux est relevé. C'est un indice d'acceptabilité du milieu par les pucerons (Rahbé, 1988). Les larves non fixées sont remplacées. Les pucerons sont laissés sur les milieux jusqu'à l'âge adulte (7 jours). Les milieux sont changés au jour J<sub>4</sub> (ou si une contamination est observée).

La mortalité des larves est observée chaque jour. Au jour J<sub>7</sub>, les pucerons aptères survivants sont pesés individuellement sur une balance Seteram (précision ± 1µg, Lyon France). La toxicité de la protéine pour chaque espèce de puceron est exprimée *in fine* par deux indices, la CL50 (concentration en protéine nécessaire pour provoquer la mort de la moitié des individus de la population à J<sub>7</sub>) et l'IC50 (concentration en protéine induisant une inhibition de croissance de 50% à J<sub>7</sub> par rapport aux individus sur les milieux témoins) (Rahbé, 1993).

---

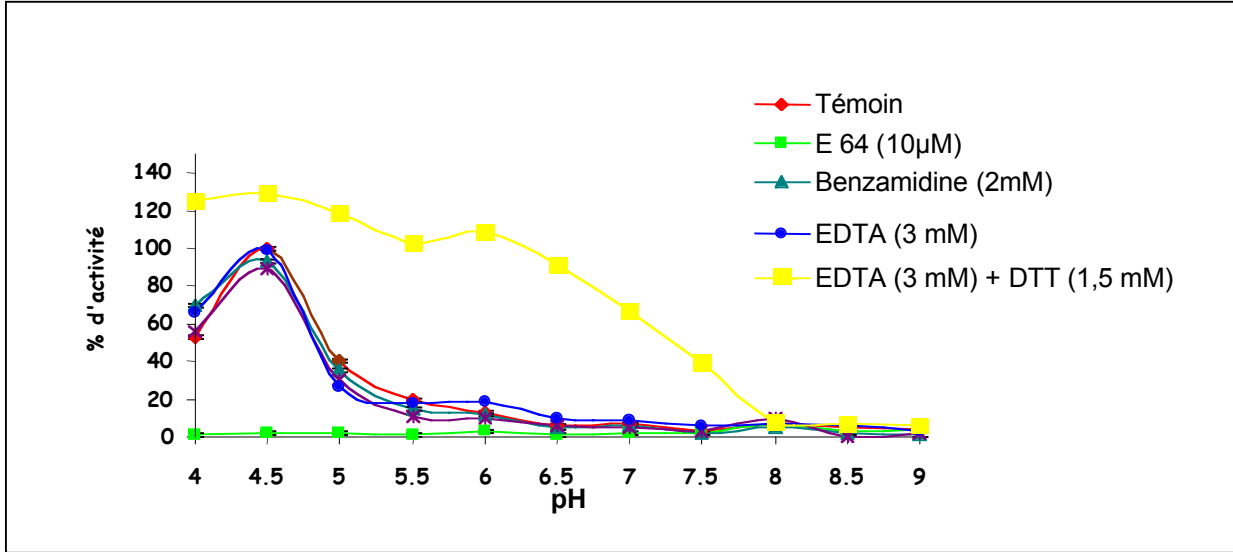
<sup>22</sup> Ils sont cependant généralement plus petits que ceux élevés sur plante hôte favorable.

## **2.8. Immunolocalisation**

Les pucerons sont intoxiqués pendant 7 jours avec PsTI à la dose de  $160 \mu\text{g}\cdot\text{ml}^{-1}$  ( $22 \mu\text{M}$ ). Les pucerons sont prélevés à  $J_{7+2}$  (après intoxication pendant 7 jours, les pucerons sont placés 2 jours sur milieu témoin pour épuiser l'inhibiteur localisé dans le bol alimentaire). Ils sont fixés par immersion dans une solution de Bouin-Holland (32% ethanol, 4% formol neutre, 2,6 mg acide picrique, 4% acide acétique glacial ; Nardon, communication personnelle), toute la nuit. Les inclusions et les coupes sont effectuées selon le protocole décrit dans l'article 1.

Les coupes sont ensuite déparaffinées grâce à 2 bains dans le cyclohexane puis réhydratées par des bains d'alcool de concentration décroissante (95%, 70%, 0%). Les lames sont incubées 30 minutes dans une solution PBS contenant 1% de BSA (PBS-BSA) afin de masquer les éventuels sites non spécifiques. Elles sont mises en présence de l'anticorps anti-PsTI dilué au  $1/100^{\text{ème}}$  dans la solution PBS-BSA pendant 1 heure. Après 3 rinçages de 5 minutes avec une solution de PBS, l'anticorps secondaire couplé au fluorochrome FITC (anticorps polyclonal anti-lapin, Sigma) est dilué au  $1/100^{\text{ème}}$  dans la solution de PBS additionnée de 0,5% de BSA et déposé sur les lames. L'incubation se déroule pendant 1 heure dans l'obscurité complète.

Après rinçage avec une solution de PBS, les lames sont montées en milieu aqueux contenant un anti-fading (Citifluor, Sigma). Les coupes sont ensuite observées au microscope confocal (Leica). L'utilisation du microscope confocal nous permet de nous affranchir d'une difficulté fréquente en immunolocalisation avec les insectes qui est la fluorescence endogène. Ainsi, nous réalisons les réglages avec les coupes témoins (coupes de pucerons élevés sur milieu artificiel témoin, hybridées avec l'anticorps anti-PsTI puis l'anticorps secondaire), tous les seuils (Excitation à 488 nm, émission de la fluorescence FITC à  $530 \pm 30$  nm) sont ajustés au plus bas (il ne reste qu'un signal faible au niveau des structures cuticulaires). Les coupes de pucerons intoxiqués sont alors analysées ; seule la fluorescence spécifique de la fluoresceine associée à la présence de PsTI dans le puceron est observée.



**Figure 3.3 : Profil de l'activité protéolytique du tube digestif d'*A. pisum* en fonction du pH.**

Les protéines de 2 tubes digestifs sont mises en contactes avec 5 µg de <sup>125</sup>I-globuline pendant 24h. La réaction est stoppée par précipitation des protéines au TCA (6%), la radioactivité présente dans le surnageant est reportée pour chaque valeur de pH.



## Résultats

Les inhibiteurs de protéases ont été utilisés pour protéger plus ou moins complètement de nombreuses espèces végétales contre leurs insectes ravageurs. Leur mécanisme de toxicité implique une interaction stœchiométrique avec les enzymes protéolytiques du tube digestif de l'insecte cible, limitant ainsi ses capacités d'assimilation protéique. Afin de comprendre le mécanisme de toxicité de PsTI, les protéases du tube digestif d'*A. pisum* ont d'abord été caractérisées et une activité protéase à sérine a été recherchée plus particulièrement.

### 3.1. Mise en évidence des protéases du tube digestif du puceron du pois.

L'activité protéolytique du tube digestif en fonction du pH a été déterminée avec un substrat non spécifique radiomarké (globuline-<sup>125</sup>I). L'utilisation d'inhibiteurs spécifiques permet d'effectuer un diagnostic des modes de protéolyse et de mettre en évidence l'activité de chaque classe de protéase. L'activité protéolytique est maximale à pH relativement acide, dont le pic d'activité à pH 4,5 (figure 3.3). Cette activité à pH acide est complètement inhibée par l'ajout d'E64 dans la solution. Par contre, l'inhibiteur spécifique des protéases à sérine, la benzamidine, est sans effet sur l'activité protéolytique ; cela confirme des essais préliminaires, de moindre sensibilité, d'hydrolyse de l'azocaséine qui s'est révélée insensible au PMSF (autre inhibiteur spécifique des protéases à sérine). L'activation de l'activité protéolytique par l'apport d'agent réducteur tel que le DTT confirme la présence d'une activité protéasique majoritairement représentée par des protéases à cystéine.

L'activité protéase à sérine a aussi été testée grâce à un test avec des substrats spécifiques. L'étude préliminaire de protéases du tube digestif du puceron du pois fait au cours de la caractérisation de la toxicité de PsTI, a révélé une très faible activité hydrolytique du substrat BapNA (plus spécifique des trypsines) et aucune activité BtpNA (plus spécifique des chymotrypsines) n'a pu être détectée (Rahbé, 2002).

L'activité enzymatique a été mesurée en fonction de la dégradation d'un substrat Suc-AAPFpNA, substrat synthétique chromogène à 405 nm. Ce dernier a été montré comme étant plus adapté à la caractérisation de protéases à sérine d'insectes, en l'occurrence de

lépidoptères (Johnston, 1995). Des extraits de tube digestif (de 1 à 10 équivalent TD par essai) ont été incubés avec 200 $\mu$ M de substrat. Toutefois aucune activité enzymatique n'a pu également être mise en évidence avec ce substrat (résultat non présenté).

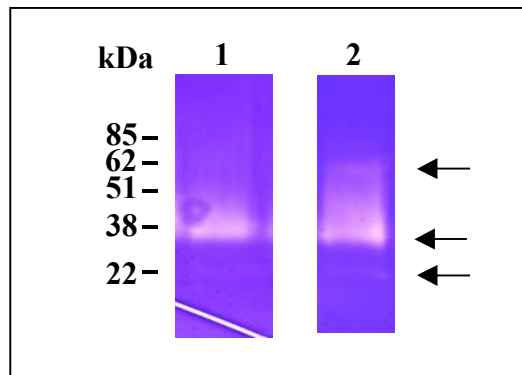
Une banque d'ADNc construite avec les messagers du tube digestif d'*A. pisum* a été criblée avec des oligonucléotides dégénérés complémentaires des séquences codantes des protéases à sérine d'insectes. Aucune amplification n'a été obtenue.

Ces résultats tendent à confirmer que les protéases du tube digestif du puceron sont en très grande majorité des protéases du type protéases à cystéine. Le puceron du pois présente donc dans son tube une activité endoprotéolytique similaire à celles mises en évidence dans le cas du puceron du melon. PsTI pourrait interagir avec une protéase à sérine très faiblement exprimée et/ou dont la spécificité de substrat serait très éloignée des protéases à sérine d'insecte déjà caractérisées.

### **3.2. Interaction de PsTI avec les protéases du tube digestif**

Les inhibiteurs de protéase à sérine de la famille des inhibiteurs canoniques se comportent comme des substrats très spécifiques des enzymes qu'ils inhibent. L'enzyme interagit avec l'inhibiteur pour former rapidement un complexe enzyme-inhibiteur très stable. Ce complexe peut se dissocier et libérer l'inhibiteur natif ou modifié. L'inhibiteur modifié est spécifiquement clivé au niveau de sa liaison peptidique entre les résidus P1 et P'1. Il n'est pas possible de séparer ces deux entités sur gel d'électrophorèse en conditions dénaturantes car ces inhibiteurs possèdent des structures très compactes maintenues par des ponts disulfures et des liaisons hydrogènes enfouies au cœur de la molécule (Jensen, 1996). Par contre, il est possible d'observer la formation de l'inhibiteur modifié en HPLC, deux pics sont alors observés sur le chromatogramme (Sreerama, 1997).

Après 24h de contact avec les protéines du tube digestif, un seul pic est à nouveau observé en sortie de chromatographie, correspondant à l'inhibiteur sous sa forme native (résultat non présenté). PsTI n'est donc pas clivé par un extrait total de tube digestif d'*A. pisum*. Ainsi, d'après cette étude préliminaire, la cible de PsTI ne semble pas être une protéase sécrétée dans la lumière intestinale, contrairement à tout ce qui a été montré de nombreuses fois chez les lépidoptères.



**Figure 3.4 : Activités protéasiques sur gel.**

Les protéines d'un tube digestif de puceron nourrie sur milieu artificiel témoin (1) ou sur milieu artificiel contenant  $160 \mu\text{g.ml}^{-1}$  de PsTI (2) sont séparées sur gel, transférés sur un gel 0,1% gélatine/10% polyacrylamide, puis l'hydrolyse de la gélatine est réalisée à pH 5 à  $37^\circ\text{C}$  pendant 16 heures.

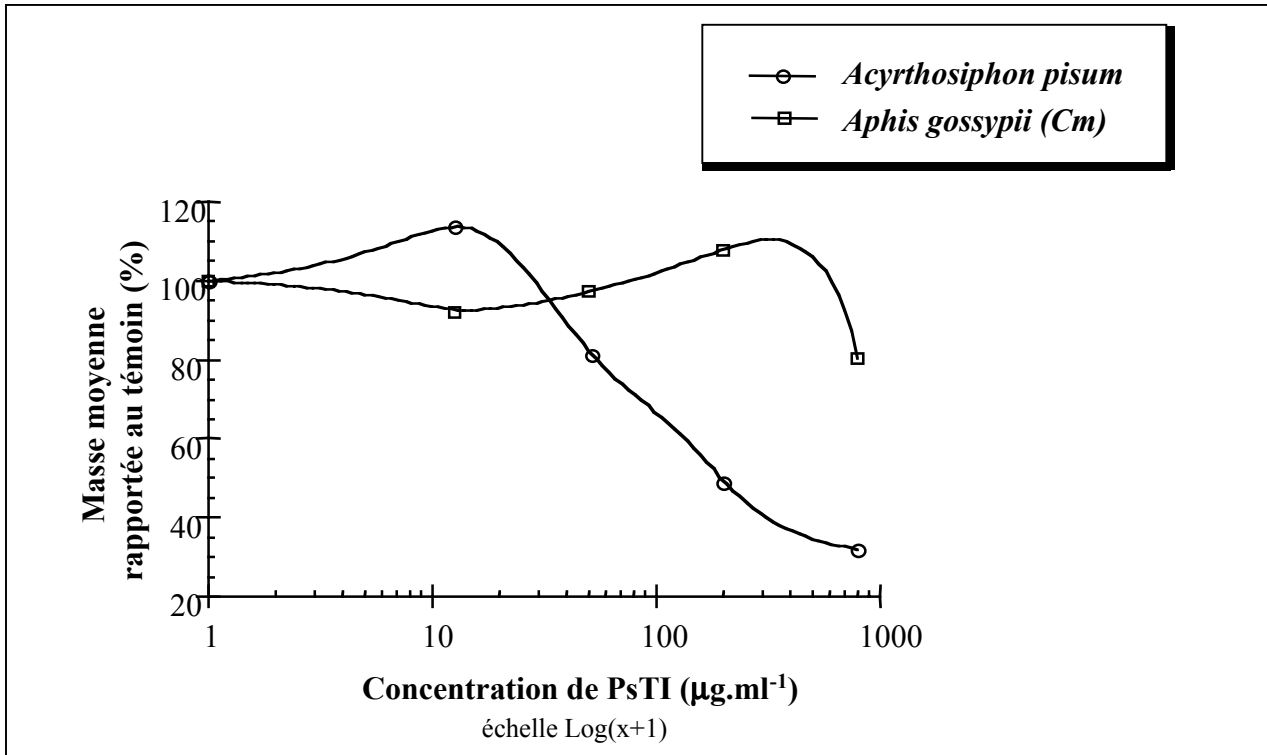
Les flèches indiquent les bandes caractéristiques d'une activité protéasique.

Les inhibiteurs de protéases ne sont parfois pas toxiques de façon directe mais leurs effets délétères peuvent reposer sur des phénomènes de compensation. Ces effets indirects sont consécutifs à la surproduction d'enzyme nécessaire à compenser le pool protéique « immobilisé » par l'inhibiteur. Un zymogramme a donc été réalisé à partir de tubes digestifs de pucerons intoxiqués par PsTI pendant 7 jours. L'activité des protéases a été révélée à pH 5 correspondant au pH optimal des activités protéasiques (figure 3.3). Trois bandes majeures ont été séparées sur gel, elles présentent un poids moléculaire apparent d'environ 23, 30 et 55 kDa (figure 3.4). Ces bandes sont spécifiques d'une activité protéases à cystéine, aucune activité supplémentaire n'a été révélée à pH 8 (résultat non présenté). D'autre part, malgré la présence de l'inhibiteur dans l'alimentation, on n'observe pas d'inhibition d'une bande en particulier et aucune bande supplémentaire n'est révélée. Il semble donc que PsTI n'agit pas non plus de façon indirecte, aucune adaptation physiologique ne semblant être induite au niveau du profil protéolytique du tube digestif.

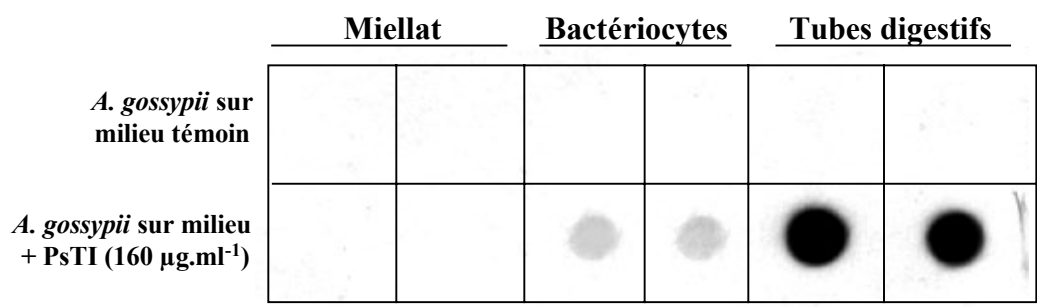
Afin d'approfondir le mécanisme de toxicité de PsTI et de tester la possibilité que des pucerons élevés sur une plante hôte distincte du pois présentent une sensibilité différente, la toxicité de PsTI pour le puceron du melon a été évaluée.

### **3.3. Effet de PsTI sur la survie et la croissance du puceron du melon.**

Des larves d'*A. gossypii* ont été intoxiquées par PsTI à des concentrations comprises entre 10 et 800  $\mu\text{g.ml}^{-1}$ , pendant 7 jours. A ces concentrations, aucune mortalité des larves n'a été observée. PsTI s'avère donc présenter un tout autre profil de toxicité que celui observé chez le puceron du pois (figure 3.5). Une diminution de croissance par rapport aux pucerons contrôles n'est observée qu'avec la dose la plus élevée. La présence de PsTI à 800  $\mu\text{g.ml}^{-1}$  dans le milieu d'alimentation est cependant responsable d'une réduction significative de la masse des pucerons, de 20% environ. Bien que ce puceron présente un profil protéolytique identique à celui d'*A. pisum* (figure 3 et article 1), le puceron du melon est beaucoup moins sensible à la présence de l'inhibiteur de protéase à sérine dans son alimentation.



**Figure 3.5 : Effet de PsTI sur la croissance du puceron du pois (*A. pisum*) et du puceron du melon (*A. gossypii*).**



**Figure 3.6 : Analyse en Dot-blot de la présence de PsTI.**

Les extraits déposés correspondent au matériel obtenus à partir de 25 pucerons.

### **3.4 Mécanismes potentiels de détoxification de PsTI chez le puceron du melon**

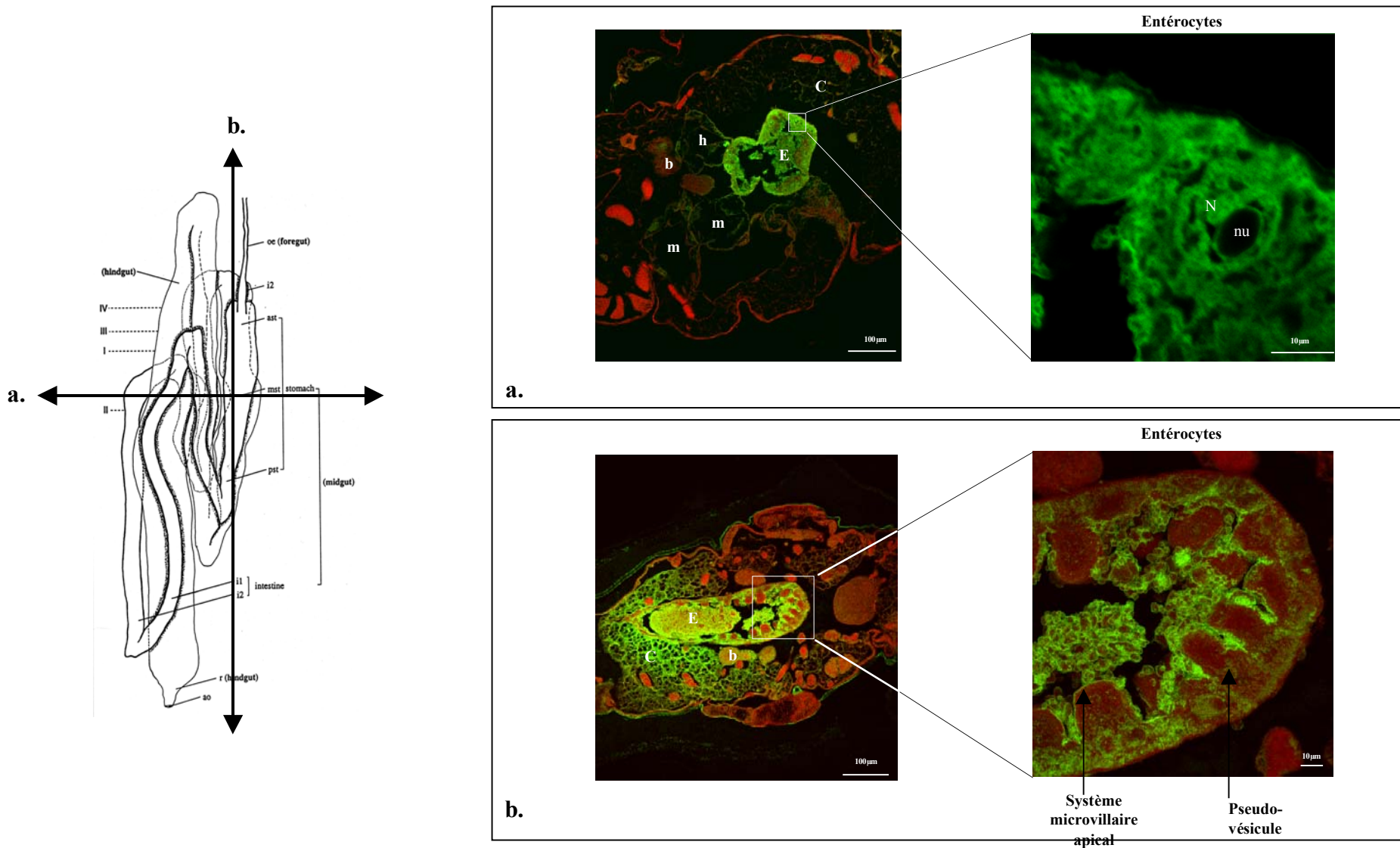
Le puceron du melon pourrait avoir développé un système plus efficace pour détoxifier son alimentation, notamment au niveau des polypeptides présents à des concentrations plus élevées que la moyenne des plantes-hôtes de pucerons, et ceci selon plusieurs mécanismes possibles : la possibilité de dégrader ou d'excréter PsTI.

Afin de tester si le puceron du melon possède la capacité de dégrader l'IP, une expérience identique à celle effectuée pour tester les interactions de PsTI et les protéases du tube digestif du puceron du pois a été mise en place pour *A. gossypii*. Après 24h de contact avec les protéines du tube digestif, on n'observe toujours qu'un seul pic en sortie de chromatographie, correspondant à l'inhibiteur sous sa forme native (résultat non présenté). PsTI n'est ainsi ni clivé par une protéase cible, ni dégradé par des protéases insensibles, et aucune différence entre les deux espèces de pucerons n'est détectée par cette approche, évaluant la stabilité de PsTI au contact avec des extraits de tube digestif.

Afin de mettre en évidence un éventuel processus différentiel d'excrétion/assimilation chez le puceron du melon, le miellat correspondant à 7 jours d'élevage de 25 pucerons sur milieu contenant PsTI a été analysé en dot blot. Aucun signal n'est détecté dans le miellat (figure 3.6). Ce résultat suggère que PsTI est complètement absorbé/utilisé par le puceron. La présence de l'inhibiteur au niveau des bactériocytes suggère que, comme l'oryzacystatine (article 2) et contrairement aux lectines à mannose (Sauvion, 1995 ; Ripoll, 2001), PsTI traverse la barrière épithéliale. Cet IP pourrait agir selon un mécanisme original ; il traverserait la barrière intestinale et inhiberait une protéase à sérine non digestive et indispensable à la survie de l'insecte sensible.

### **3.5. Localisation de la cible de PsTI chez les pucerons du pois et du melon.**

Des pucerons élevés pendant 7 jours en présence de PsTI ( $160 \mu\text{g.ml}^{-1}$ ) puis 2 jours sur milieu contrôle ont été utilisés pour la recherche des sites de fixation de l'inhibiteur dans l'insecte, y compris hors du tube digestif. Chez les 2 pucerons, un gradient de marquage est



**Figure 3.7 : Détection par immunolocalisation de PsTI chez les larves de puceron intoxiquées par PsTI ( $160 \mu\text{g.ml}^{-1}$ ).**

a. Coupe transversale d'*Acyrthosiphon pisum*.

b. Coupe longitudinale d'*Aphis gossypii*.

Le niveau des coupes dans le tube digestif est indiquée sur le schéma à droite. La fluorescence spécifique associée à PsTI est représentée par la couleur verte, l'autofluorescence des structures est visible en rouge. Corps gras (C), bactériocytes (b), estomac (E), intestin moyen (midgut, M), intestin postérieur (hindgut, H), noyau (N), nucléole (nu). Une représentation schématique des coupes (Ponsen, 1972) est donnée en annexe 2.

observé le long du tube digestif : l'estomac est très fortement marqué à l'opposé du rectum qui ne présente qu'une intensité faible de fluorescence (figure 3.7a). Le marquage est très intense à l'intérieur des cellules, PsTI est donc internalisé par les cellules épithéliales du tube digestif. D'autre part, on retrouve PsTI dans l'hémolymphe et au niveau des bactériocytes. L'inhibiteur de protéase est donc capable de traverser la barrière épithéliale.

C'est au niveau cellulaire de l'épithélium intestinal que la différence entre les deux espèces de pucerons est la plus importante. Chez le puceron du pois, un marquage associé à PsTI est observé uniformément dans les cellules épithéliales du tube digestif. L'inhibiteur est présent dans le cytoplasme, mais aussi dans le noyau et il semble que seul le nucléole ne soit pas marqué.

Par contre, chez le puceron du melon un marquage uniquement à la périphérie des entérocytes est observé. PsTI ne semble pas s'accumuler de la même façon dans les cellules du tube digestif de cette espèce. Au niveau intracellulaire, l'inhibiteur de protéase semble dans ce cas être pris en charge par de petites vésicules (figure 3.7b).

## **Discussion**

PsTI est un inhibiteur de protéase à sérine de la famille des Bowman-Birk isolé chez le pois. Cette molécule de 7,5 kDa s'est révélée être toxique pour le puceron du pois (Rahbé, 2002). Ses propriétés aphidotoxiques reposent essentiellement sur l'activité anti-chymotrypsine de l'inhibiteur bifonctionnel.

Au contraire des insectes broyeur de l'ordre des lépidoptères et des coléoptères, les pucerons, insectes piqueurs/suceurs de l'ordre des Homoptères ne semblent posséder que des protéases à cystéine dans leur tube digestif. Cela fait de PsTI une molécule entomotoxique s'appuyant sur un mécanisme de toxicité différent de celui invoqué jusqu'ici pour les inhibiteurs de protéases utilisés en protection des plantes contre les insectes.

Le tube digestif représente le premier élément cellulaire d'interaction pour les molécules ingérées, c'est pourquoi l'arsenal enzymatique du tube digestif a été étudié.

Le tube digestif du puceron du pois renferme bien des activités protéolytiques. Cette activité est optimale à pH acide, et de type protéases à cystéine (figure 3.3), comme c'est le cas pour le puceron du melon (article 1).



Il est important de noter, comme dans le cas du puceron du melon, que l'activité protéolytique est difficilement solubilisable (traitement du broyat de tube digestif pendant 2h à température ambiante avec 1% de SDS). C'est pourquoi le déficit d'activités endoprotéolytiques, souligné pour les pucerons dans les différentes études comparatives antérieures (Terra, 1996 ; Mochizuki, 1998), devrait être nuancé par la mention plus exacte que les pucerons ne possèdent pas d'endoprotéases *secrétées* dans la lumière du tube digestif.

Aucune activité protéase à sérine n'a été mise en évidence, que ce soit avec des inhibiteurs ou un substrat spécifiques. Il est possible que les protéases cibles de l'inhibiteur ne soient pas stables en solution et qu'elles soient inactivées après la destruction des tissus. De plus, lors du broyage, la compartimentation cellulaire est supprimée. Une potentielle protéase à sérine cible peut être alors inhibée par un inhibiteur endogène et ne plus être détectable. Lors du séquençage aléatoire de la banque soustractive spécifique des messagers exprimés dans le tube digestif, une séquence homologue aux serpinines a été isolée (chapitre I). Si elle est bien exprimée dans les cellules du tube digestif et non sécrétée, cette protéine pourrait bien inhiber une activité protéase à sérine présente à faible niveau.

Le criblage de la banque de messagers représentatif du transcriptome du tube digestif d'*A. pisum* avec des oligonucléotides complémentaires des séquences de protéases à sérine d'insectes ne nous a pas permis d'isoler de séquence d'ADNc de cette classe d'endoprotéases. La protéase cible de PsTI pourrait cependant être représentée très faiblement dans les cellules du tube digestif. Le manque de succès de cette expérience pourrait aussi reposer sur une plus grande divergence des protéases à sérine du puceron. Des différences entre les protéases à sérine des insectes et celles des mammifères (trypsine bovine) ont été soulignées. Ces protéases sont insensibles aux ions calcium, alors que ceux-ci sont indispensables à la trypsine de mammifère (Purcell, 1992). Récemment une protéase à sérine a été clonée chez un Homoptère, *Nilaparvata lugens* (Foissac, 2002). Toutefois, cet insecte n'est peut-être pas aussi strictement phloémophage que les pucerons, bien que ce tissu soit un site d'alimentation privilégié également chez cet insecte. La physiologie digestive de *Nilaparvata* se différencie aussi de celle du puceron car l'activité protéolytique intestinale est principalement représentée par des protéases à sérine. La séquence codante de cette protéase à sérine présente plus d'homologies avec les séquences de mammifères qu'avec celle d'insectes déjà clonées et biochimiquement caractérisées, à l'exception d'une séquence probablement orthologue présente dans le génome de drosophile (Foissac, 2002). Ce point renforce notre hypothèse sur la spécificité des oligonucléotides utilisés lors de notre criblage par PCR et sur la variabilité potentielle de ces protéases à sérine au sein des différentes familles d'insectes. Des insectes

souvent présentés comme proches par leurs caractéristiques trophiques et alimentaires, peuvent s'avérer reposer sur des physiologies digestives divergentes, renforçant d'ailleurs l'intérêt d'approches comparatives assez larges.

Les figures d'immunolocalisation montrent que PsTI s'accumule dans les cellules épithéliales de l'intestin du puceron. Ce type d'observations a déjà été fait chez des animaux très éloignés des pucerons. En effet, BBI qui, comme PsTI, est aussi un inhibiteur bifonctionnel de protéases à sérine de la famille des Bowman-Birk, possède des propriétés anti-cancérogènes mises en évidence aussi bien *in vitro* qu'*in vivo*. Il a été montré que BBI était capable d'inhiber la prolifération de cellules irradiées de fibroblaste, et que le domaine anti-chymotrypsine de l'inhibiteur était responsable des effets observés (Yavelow, 1985). Des études sur les propriétés de BBI sur le cancer du colon ont également été menées (Billings, 1991). Des cultures de cellules épithéliales d'intestin de rat sont capables d'internaliser l'inhibiteur après 15 min d'incubation. On retrouve un marquage intense dans toute la cellule. Un fractionnement cellulaire montre que BBI est majoritairement présent dans le cytoplasme des cellules en culture sous forme active, démontrant que l'inhibiteur est sous forme libre dans le cytosol (Billings, 1991).

Chez les pucerons, aucune interaction avec les protéases du tube digestif n'a pu être mise en évidence : PsTI ne semble pas être dégradé, il n'est pas clivé par une protéase cible (figure 3.4), il n'inhibe aucune protéase active sur gel (figure 3.5) suggérant que PsTI resterait comme chez les mammifères sous forme majoritairement active dans le cytosol.

Le domaine anti-chymotrypsine, responsable des propriétés entomotoxiques, pourrait interagir avec un autre effecteur cellulaire ; ses capacités d'inhibiteur de protéases ne seraient alors pas mises en jeu, et sa toxicité reposerait alors sur sa structure tridimensionnelle particulière. Pourtant, des formes variantes obtenues par " phage display " d'un inhibiteur de protéase à sérine (Mustard Trypsin Inhibitor) possèdent également des capacités aphidotoxiques (Ceci, 2002). Le niveau de toxicité des molécules est très fortement corrélé à l'affinité de ces inhibiteurs pour la chymotrypsine bovine, alors que l'inhibiteur parent sauvage, anti-trypsique, possède peu d'activité aphicide (Ceci, 2002). Ces molécules sont de petits polypeptides de 36 résidus aminés mais qui ne présentent aucune homologie structurale avec d'autres inhibiteurs de protéases à sérine de plante (Menegatti, 1992). Ces résultats étayaient donc l'hypothèse que la toxicité de PsTI reposerait bien sur sa capacité à inhiber une chymotrypsine, et non sur un trait caractéristique de sa structure en deux boucles rigides.

Bien que le mécanisme précis responsable de la toxicité de ces inhibiteurs ne soit pas connus, la séquestration de PsTI par les cellules du tube digestif de puceron est indiscutable. Le signal, spécifique de la présence de l'IP dans les cellules, décroît le long du tube digestif. Nous pouvons nous demander s'il existe une relation entre les fonctions des cellules intestinales et leur sensibilité à l'inhibiteur. Ces cellules, à la fois sécrétrices et absorbantes offrent des différences d'une région à l'autre, qui reflètent des rôles physiologiques diversifiés : sécrétion (enzymes digestives<sup>23</sup>, substances diverses comme des oligosaccharides), absorption, excrétion<sup>24</sup>, régulation osmotique.

Ponsen (1972, 1991) considère que l'« estomac » est le siège de nombreuses sécrétions et possède des cellules digestives très actives. Dans la région médiane du mésentéron, les cellules de sécrétions sont absentes. La fonction essentielle des cellules de cette région serait l'absorption (eau, acides aminés...). La fonction d'excrétion serait remplie par les cellules du stomodeum (rectum). Chez les pucerons, cette partie du tube digestif est très développée, il semble que cette région joue un rôle dans l'osmorégulation, pouvant être rapproché de celui de la « chambre filtrante » observée chez d'autres homoptères. La régulation de la pression osmotique est un processus essentiel chez ces insectes qui ingèrent de grandes quantités de nourriture liquide, dont la concentration en solutés peut varier considérablement.

Un marquage très intense est observé au niveau de l'estomac, PsTI n'est donc pas pris en charge de la même façon que les acides aminés absorbés au niveau du mésentéron. Dans les cellules en culture de colon de mammifères, aucun transporteur spécifique en surface des cellules n'a été mis en évidence (Billings, 1991). Il est peu probable qu'une molécule de 8 kDa, sans propriétés hydrophobes particulières, puisse pénétrer dans la cellule par simple diffusion passive. D'autre part, lorsque la vie cellulaire est ralentie (cellules à 4°C), le transport de BBI est inhibé (Billings, 1991), et ces résultats suggèrent que le passage de ces molécules à l'intérieur des cellules mammaliennes est basé sur un mécanisme de transport actif.

Toutefois, un marquage est observé aussi dans les autres parties du tube digestif. A cette concentration de PsTI, l'estomac peut se trouver saturé en sa capacité d'incorporation de la molécule. La protéine transiterait alors le long du tube digestif et serait internalisée par les cellules dans des parties postérieures. Pour tester cette hypothèse, des tests avec une

---

<sup>23</sup> amino et oligopeptidases, estérases, glucosidases

<sup>24</sup> ammonium ou autres composés riches en azote comme l'arginine.

concentration plus élevée d'IP pourraient être effectués afin d'observer une augmentation de l'intensité du marquage dans les cellules de l'intestin moyen et postérieur.

Toutes les cellules du tube digestif du puceron possèdent le mécanisme de transport de ce polypeptide. Il ne semble pas en effet que ce soit un processus restreint aux cellules digestives de l'estomac, ni même éventuellement un processus spécifique des inhibiteurs de protéases à sérine. On trouve également un marquage spécifique de l'oryzacystatine à l'intérieur des cellules du tube digestif (article 2).

L'IP se trouve aussi être capable de traverser la barrière intestinale. En effet, un signal est observé dans l'hémolymphe de même qu'au niveau des bactériocytes (figure 3.7). Le tube digestif pourrait n'être qu'un lieu de transit et non la cible tissulaire principale, et PsTI pourrait inhiber une protéase à sérine d'importance critique et localisée hors du tube digestif.

Le puceron du melon s'est révélé être beaucoup moins susceptible à la présence de PsTI dans l'alimentation, mais des différences de sensibilités entre le puceron du pois et le puceron du melon ont cependant déjà été soulignées dans le cas de l'intoxication par la lectine GNA (Sauvion, 1995 ; Ripoll, 2001). Il pourrait s'agir d'un phénotype général lié à la biologie de cette espèce, et à sa polyphagie bien plus importante que celle du puceron du pois. Rappelons qu'*Aphis gossypii*<sup>25</sup> est la seule espèce aphidienne capable de coloniser les cucurbitacées, c'est à dire des plantes ayant une singularité certaine au niveau des teneurs protéiques phloémiennes.

Les études sur la nutrition des pucerons ignorent largement la composition en protéines dans les sources d'alimentation, et sont souvent focalisées sur les composés de petits poids moléculaires tel que les acides aminés, les sucres, les minéraux ou les vitamines. D'un point de vue nutritionnel cela est facilement justifiable car pour beaucoup d'espèces de plantes, les protéines ne comptent que pour 1 à 2% des acides aminés totaux dans le phloème comme c'est le cas pour le pois ou le lupin. Mais les plantes de la famille des cucurbitacées renferment un phloème riche en protéines dont la concentration peut atteindre 60 mg. ml<sup>-1</sup>. En particulier, plusieurs familles d'inhibiteurs de protéases y ont été décrites (Yoo, 2000 ; Haebel, 2001). Ainsi *A. gossypii*, inféodé au melon, pourrait présenter un mécanisme de prise en charge des macromolécules différent des autres pucerons, lui permettant tout à la fois une utilisation accrue des protéines alimentaires et une meilleure détoxification de peptides de

---

<sup>25</sup> Et plus précisément certaines souches au sein de cette espèce.

défense présents dans son alimentation. Ces capacités de neutralisation des facteurs antinutritionnels tels que les IP ne semblent cependant pas reposer sur des mécanismes de dégradation ou d'excrétion plus efficaces puisque PsTI ne semble pas dégradé, et qu'aucune trace de l'IP n'est retrouvé dans le miellat d'*A. gossypii*.

Toutefois, *A. gossypii* possède la cible de PsTI puisqu'une dose importante d'IP entraîne une diminution de croissance, d'environ 20% par rapport aux conditions témoins (figure 3.5). L'étude des activités protéolytiques de son tube digestif au niveau qualitatif ne révèle pas de différence avec celle du puceron du pois (article 2 et figure 3.3). Les figures d'immunomarquage confirment la présence de l'inhibiteur au niveau du tube digestif, des bactériocytes et dans l'hémolymphe. Par contre, chez le puceron du melon, *A. gossypii*, PsTI ne semble pas s'accumuler de la même façon dans les cellules du tube digestif, et la plus faible sensibilité d'*A. gossypii* pourrait plutôt reposer sur une séquestration et/ou utilisation particulière des protéines.

Bien que nos résultats ne fournissent pas de réponse claire et définitive à la question de la cible moléculaire de PsTI, celle-ci pourrait être :

- soit une activité endogène à la cellule intestinale (e.g. protéasome ?), non révélable avec les substrats/sensibilités utilisés par nous, et mieux protégée chez *A.gossypii* par un processus d'« exclusion cytoplasmique » des polypeptides d'origine exogène.

- soit une cible endoprotéasique extradigestive (chymotrypsine ?), également protégée (ou modifiée ?) chez *A. gossypii* par un bilan de transit intestinal des polypeptides alimentaires natifs assez différent de celui qui prévaut chez un puceron s'alimentant sur des phloèmes moins riches en protéines (*A. pisum*).

## **II. Mutagenèse dirigée et Expression hétérologue de PsTI et de PsCI**

### **Introduction**

Au cours de ce travail, les bases moléculaires de la toxicité de PsTI n'ont pas pu être élucidées, toutefois cet inhibiteur de protéase s'est révélé très toxique pour le puceron du pois. Seuls les inhibiteurs bifonctionnels de type Bowman-Birk possédant un site actif anti-chymotrypsine sont toxiques pour ce puceron. La toxicité a pu être corrélée à l'affinité de l'inhibiteur pour la chymotrypsine bovine (Rahbé, 2002).

Afin d'élargir le nombre de gènes utilisables en transgénèse pour lutter contre les pucerons, nous avons décidé de créer une nouvelle molécule à activité biologique aphicide. Nous avons souhaité créer un nouvel inhibiteur de chymotrypsine à partir de la molécule PsTI, afin d'accroître sa toxicité et ainsi son spectre d'action. Dans cet objectif, le site antitrypsine est changé en un potentiel site antichymotrypsine, par mutagenèse dirigée.

Ce choix est dû au fait que d'une part, à notre connaissance aucun inhibiteur portant une double spécificité pour la chymotrypsine n'existe dans la nature. D'autre part, nous souhaitons conserver la structure particulière de PsTI : 2 domaines symétriques organisés grâce à un réseau de sept ponts disulfures. En effet, lors de l'étude des relations structure/toxicité, des peptides synthétiques mimant les boucles de site actif se sont révélés moins toxiques que l'inhibiteur natif. La toxicité de cette famille d'inhibiteur repose probablement sur d'autres traits structuraux portés par la protéine native (Rahbé, 2002). C'est pourquoi nous avons choisi de créer un variant de PsTI plutôt que de réaliser un peptide simple antichymotrypsine. De plus, le fait d'avoir éliminé le site antitrypsine permet d'écartier un élément d'interaction avec une trypsine qui mobiliserait l'IP de façon inefficace en regard de la toxicité.

Chez le pois, les inhibiteurs de protéases de la famille des Bowman-Birk sont codés par une famille multigénique (Welham, 1998). Ils possèdent entre eux 78 à 85% d'homologie (Domoney, 2002). Pour être sûr de disposer du gène codant pour le peptide PsTI-2, il nous a semblé plus facile de créer un gène synthétique, plutôt que de cloner le gène codant pour PsTI-2.

Après une étude structurale, nous avons modifié le domaine antitrypsine afin d'obtenir un gène codant pour une nouvelle isoforme de PsTI, appelée PsCI (mutant double anti-chymotrypsine).

Nous avons construit des vecteurs permettant l'expression de ces deux isoformes dans la levure afin de produire en grande quantité ces protéines, dans le but d'effectuer une première évaluation du gain de toxicité de PsCI en test d'alimentation artificielle. L'expression d'un inhibiteur tryptique de la famille des Bowman-Birk ne peut pas être réalisée dans la bactérie *E. coli* (Le Tan-Dumanois, 1994). Le système d'expression eucaryote, contrairement à la bactérie, permet la maturation des protéines, impliquant des processus tel que la glycosylation, des protéolyses éventuelles ou la formation de ponts disulfure qui sont dans notre cas critique pour la conformation active de la molécule.

Parallèlement, des constructions permettant l'expression de PsTI ou de PsCI *in planta* ont été réalisées. Dans le cadre de cette thèse *Arabidopsis thaliana* a été utilisé comme espèce modèle. En effet, cette plante présente l'avantage d'être facile à transformer. Le niveau de stabilité de la protéine et sa maturation correcte dans une autre espèce végétale que le pois seront ainsi évalués. D'autre part, *A. thaliana* est une plante hôte du puceron du pêcher, *Myzus persicae* (Moran, 2001). Après l'obtention et la caractérisation de lignées transgéniques, les gains de résistance pour le puceron pourront être évalués *in planta*.

Nous avons choisi d'exprimer les 2 gènes synthétiques sous le contrôle du promoteur 35S<sup>2</sup> du virus de la mosaïque du chou-fleur car il permet l'expression d'une protéine exogène de manière constitutive, et notamment dans le phloème où elle devient accessible au puceron, insecte piqueur phloémophage (article 2 ; Ripoll, 2001 ; Lee, 1999).

## PsTI-2

GDDVKSACCDTC LCTKSNPPTCRVVDVGETCHSACLSCICAYSNPPKCQFDTHFCYKACHN

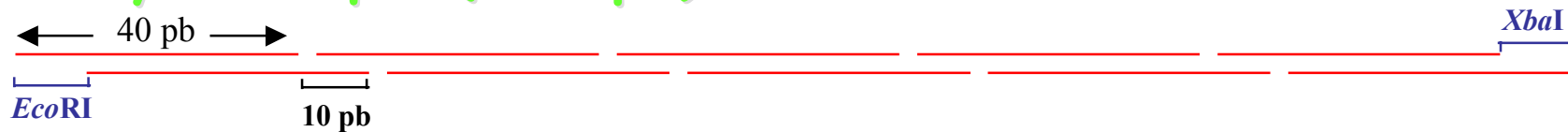
anti-trypsine anti-chymotrypsine

## PSTI-1 :

GDDVKSACCDTCLCTKSNPPTCRVVDVRETSCHSACDSCICAYSNPPKCQFDTHFCYKACHN

aga → ggt    gac → ctt

## Gene synthétique (180 pb)



**Figure 3.8 : Représentation schématique de la démarche utilisée pour définir les oligonucléotides permettant la synthèse du gène artificiel codant PsTI.**

PsTI-2 et PSTI-1 sont deux isoformes dont la variabilité repose sur deux acides aminés (résidus 28 et 36, notés en rose). La séquence nucléotidique de *PsTI-1* étant disponible dans la base de donnée Genbank, les oligonucléotides ont été définis en fonction de cette séquence *PsTI-1*. Seule la séquence des codons correspondant aux résidus 28 et 36 a été changée afin de correspondre à la séquence de PSTI-2. Des séquences correspondant au site de restriction de l'enzyme *EcoRI* et *XbaI* ont été additionnées aux extrémités de la séquence du gène synthétique pour permettre son clonage. Les 10 oligonucléotides de 40 pb sont chevauchant sur 30 pb avec un oligonucléotide complémentaire et sur 10 pb avec l'oligonucléotide complémentaire suivant. Les boucles correspondant au site antitrypsine ou antichymotrypsine sont encadrées de rouge, les acides aminés catalytiques sont notés en rouge.



## Matériel et méthodes

Les mêmes méthodologies ont été appliquées pour l'expression de PsTI et du mutant PsCI.

### 2.1 Synthèse du gène artificiel

Connaissant la séquence protéique de PsTI ainsi que la séquence nucléotidique d'une isoforme de PsTI, nous avons défini 5 oligonucléotides de 50 bases correspondant au brin codant et 5 oligonucléotides identiques au brin non-codant. Chaque amorce sens est complémentaire d'une amorce antisens sur 40 paires de bases et complémentaire de la suite sur 10 nucléotides (figure 3.8).

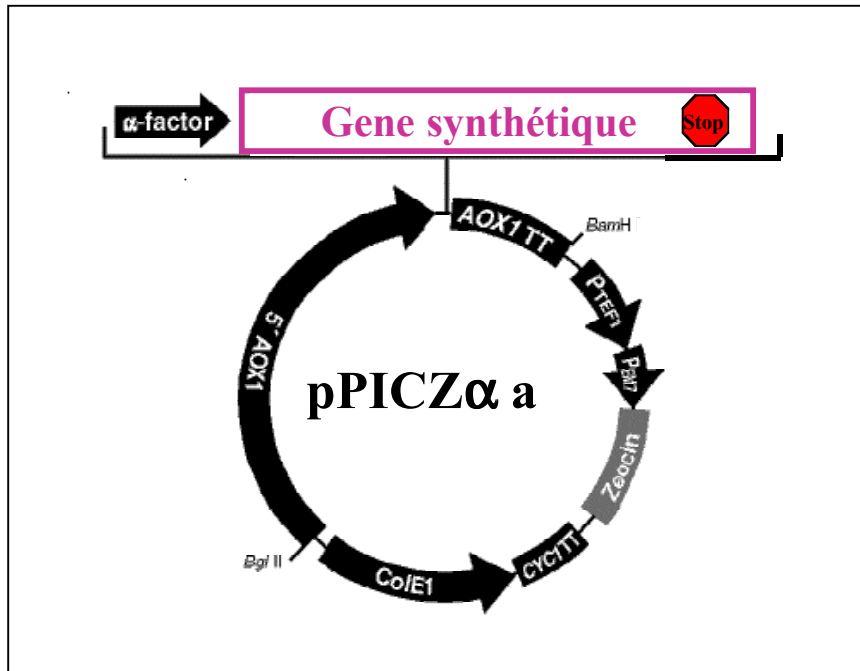
Les 10 oligonucléotides phosphorylés (20 pmol) sont mélangés dans un volume de 100  $\mu$ l final et chauffé pendant 15 minutes à 90°C. Puis la température est diminuée de 10° toutes les 15 minutes. Lorsque la température atteint 60°C, 10 unités de ligase thermostable (Epicentre) sont ajoutées au mélange. La ligation des amorces se poursuit pendant 1 heure à 60°C puis 1 heure à 50°C. La ligase catalyse la formation d'une liaison hydrogène entre l'extrémité 3'OH d'un oligonucléotide et l'extrémité 5'Phosphate de l'oligonucléotide suivant.

Le produit de la réaction est cloné dans le vecteur pBS KS<sup>-</sup> digéré par les endonucléases de restriction *EcoRI* et *XbaI*. Le respect de la séquence codante a été vérifié par séquençage du plasmide pBS-IP avec les amorces universelles (Forward et Reverse).

### 2.2.Expression dans la levure

#### 2.2.1. Construction du vecteur d'expression

Le plasmide pBS-IP est digéré par les enzymes de restriction *EcoRI* et *XbaI* afin d'isoler le gène synthétique. Le gène est inséré dans le vecteur pPICZ $\alpha$ A (Invitrogen) digéré au préalable par les mêmes enzymes. Ce vecteur est un vecteur qui permet l'expression et la sécrétion de protéines recombinantes dans la levure *Pichia pastoris* (figure 3.9). La protéine est exprimée sous le contrôle d'un promoteur inductible par le méthanol et en fusion avec un signal de sécrétion (le facteur  $\alpha$  de *Saccharomyces cerevisiae*). La sélection des bactéries et



**Figure 3.9 : Représentation schématique du vecteur d'expression dans la levure *Pichia pastoris*.**

Le gène codant la protéine d'intérêt est sous le contrôle du promoteur *AOX1* inducible par le méthanol. La protéine hétérologue est fusionnée avec le facteur  $\alpha$ , permettant sa sécrétion dans le milieu de culture.

Les levures recombinantes sont sélectionnées grâce à leur résistance à l'antibiotique Zéocine porté par le vecteur pPICZ.

des levures recombinantes s'effectue grâce à l'expression du gène de résistance à la Zéocine™<sup>26</sup> (Invitrogen), porté par le vecteur pPICZαA.

Le vecteur d'expression pPICZ-IP ainsi obtenu, est amplifié dans *E. coli* puis linéarisé par digestion de l'enzyme *SacI* et introduit dans la souche GS115 de la levure *Pichia pastoris* (Invitrogen) selon le protocole recommandé par le fournisseur.

Les transformants sont sélectionnés sur milieu YPD-sorbitol (annexe 3) contenant 100 µg.ml<sup>-1</sup> de Zéocine™ à 30°C.

### 2.2.2. Expression des protéines recombinantes

Dix clones de levures recombinantes sont pris en compte pour chaque construction (PsTI et PsCI). Une colonie est utilisée pour inoculer 20 ml de milieu BMGY (annexe 3) additionné de 100 µg.ml<sup>-1</sup> de Zéocine™ dans une fiole de 250 ml. Les levures sont incubées à 30°C, sous agitation de 250 rpm, jusqu'à atteindre une densité optique à 600nm de 2 à 6 (soit environ 24 heures). Les cellules sont culottées puis reprises dans un milieu de culture BMM (annexe 3) afin d'atteindre une densité optique à 600 nm égale à 1 dans un volume final de 20 ml. L'expression de la protéine recombinante est induite grâce à l'addition de méthanol à une concentration finale de 0,5% (v/v) dans le milieu de culture.

Tous les jours, 1 ml de culture est prélevé et remplacé par le même volume de milieu BMM-15% méthanol, pendant 4 jours. Les échantillons sont centrifugés à 1500 g pendant 5 minutes. Le surnageant est dialysé contre un tampon Tris -HCl 100 mM pH 8 grâce à une membrane présentant une limite d'exclusion de 5 kDa (Millipore). Le dialysat est lyophilisé et conservé à 4°C. Au moment de l'utilisation, la poudre est pesée et reprise dans un volume d'eau désirée.

L'expression de la protéine recombinante est estimée en Dot-blot selon le protocole décrit au paragraphe 2.3.3.2. pour le niveau d'expression de l'IP dans les plantes transgéniques.

### 2.2.3. Vérification de l'activité biologique des IP.

L'activité inhibitrice des protéines recombinantes a été vérifiée par un test d'inhibition des protéases à sérine. Une quantité croissante de surnageant de culture de levure dialysé (0 à 30µl) est incubée avec les enzymes trypsine (100 ng) ou chymotrypsine (500 ng). Les

---

<sup>26</sup> La zéocine est un antibiotique isolé des *Streptomyces* et appartient à la famille de la bléomycine.

activités enzymatiques résiduelles sont mesurées par hydrolyse d'un substrat chromogénique spécifique à 500  $\mu$ M, le  $\alpha$ N benzoyl-L arginine p-nitroanilide (BapNA) pour la trypsine et le  $\alpha$ N benzoyl-L tyrosine p-nitroanilide (BTpNA) pour la chymotrypsine à 37°C dans un tampon Tris-HCl 100 mM pH 8.

## **2.3. Expression de PsTI et PsCI *in planta***

### **2.3.1. Construction de la cassette d'expression**

La phase ouverte de lecture a été obtenue grâce à une réaction de PCR sur le plasmide pBS-IP au moyen de 2 amorces : l'oligonucléotide ATG (AAG CTT ATG GGT GAT GAT GTC AAA TC) et l'oligonucléotide STOP (GAA TTC TCA GTT GTG ACA TGC TTT ATA G), complémentaires respectivement de l'extrémité 5' et 3' du gène synthétique. L'utilisation de l'amorce ATG a permis l'obtention d'un site de restriction *Hind*III en amont de l'ATG. Le produit de PCR est cloné dans le vecteur pGEM-T(Promega) et le respect de la phase de lecture a été vérifié par séquençage du plasmide pGEM-IP avec les amorces universelles (Forward et Reverse).

Le plasmide pGEM-IP est digéré par les enzymes de restriction *Hind*III et *Pst*I afin d'isoler la séquence codante de l'inhibiteur. Le gène est inséré dans le plasmide pLBR19 ( plasmide dérivé de pUC18 contenant le promoteur 35S avec une séquence d'amplification doublée du CaMV et le terminateur du CaMV) préalablement digéré par les mêmes enzymes de restriction.

La cassette d'expression (p35S<sup>2</sup>::IP::ter) est ensuite obtenue par digestion du plasmide pLBR19-IP, avec les enzymes de restriction *Kpn*I et *Bgl*II. Cette cassette est clonée dans le vecteur binaire pCambia 2300 digéré par les enzymes *Kpn*I et *Bam*HI afin d'obtenir le plasmide pCambia-IP permettant l'expression constitutive *in planta* de l'inhibiteur (figure 3.10).

Les vecteurs binaires ainsi obtenus sont ensuite introduits dans la souche d'*Agrobacterium tumefaciens* C58/pMP90 par électroporation (Koncz, 1986).

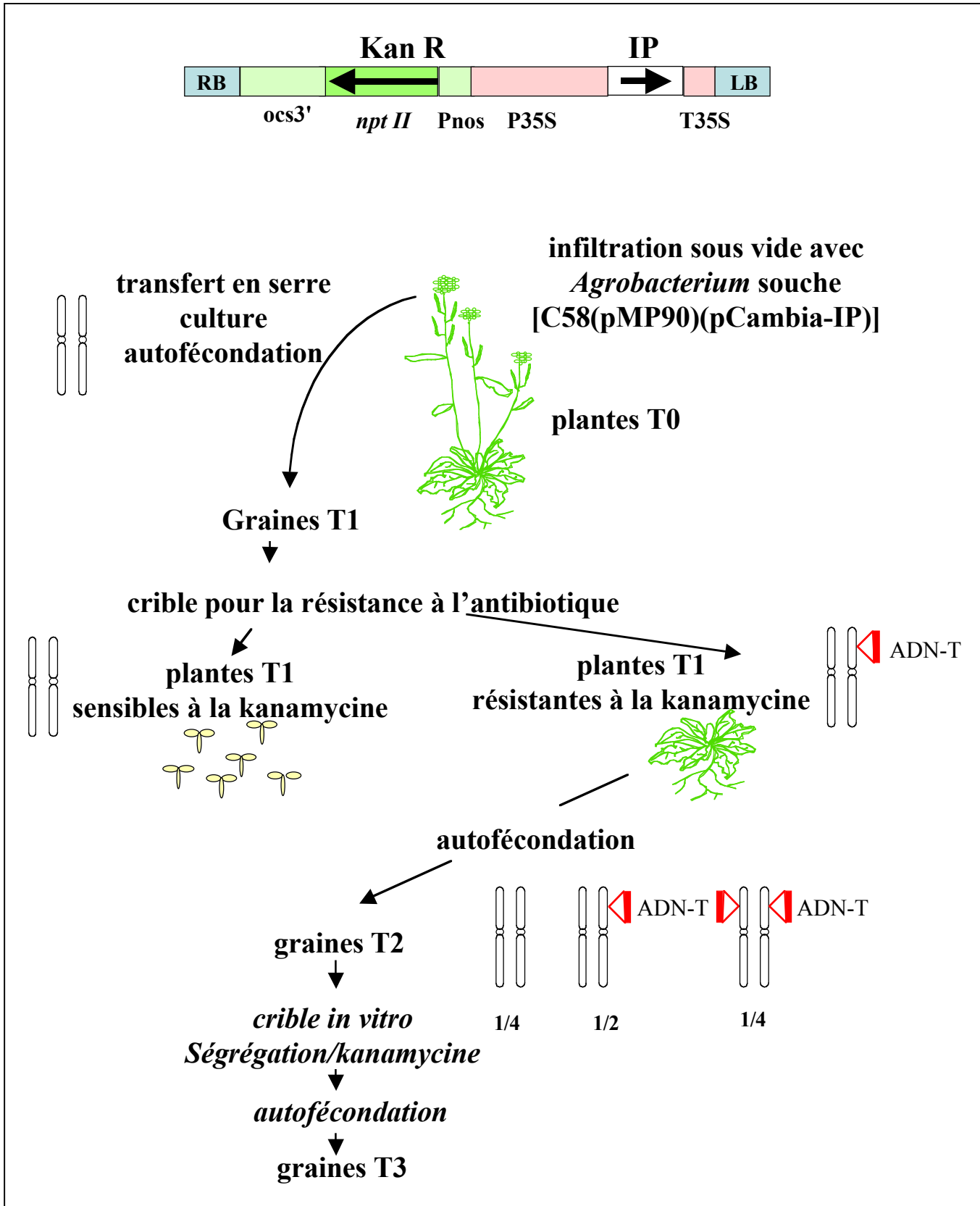


Figure 3.10 : Schéma de l'ADN-T dans le vecteur pCambia-IP et Principe de la génération de transformants d'*A. thaliana*.

### **2.3.2. Transformation d'*Arabidopsis thaliana* et obtention de lignées homozygotes**

L'écotype Wassilevskija (WS) d'*Arabidopsis thaliana* a été transformé selon le protocole décrit par Clough et Bent (Clough, 1998). Les fleurs d'*A. thaliana* âgées de 2 à 3 semaines sont trempées pendant 2 min dans une solution de saccharose à 5% et de Silwet L-77 à 0,03% contenant la suspension de bactéries issues d'une culture de densité optique de 0,8 à 600 nm en milieu LB additionné de gentamycine (20 mg.l<sup>-1</sup>), de rifampicine (50 mg.L<sup>-1</sup>) et de kanamycine (50 mg.l<sup>-1</sup>).

Les graines des plantes ayant subi la transformation sont récoltées. Elles sont stérilisées pendant 5 à 10 min dans une solution contenant 90% d'éthanol et 10% d'une solution de Barychlor® (un comprimé pour 40 ml d'eau) sous agitation. Les graines sont ensuite rincées dans de l'éthanol absolu, puis mises à sécher toute la nuit en conditions stériles. Elles sont ensuite semées sur un milieu de culture (annexe 4) contenant 100 µg.ml<sup>-1</sup> de kanamycine afin de sélectionner les transformants primaires. Après 2 à 3 semaines de cultures *in vitro*, les plantes résistantes à l'antibiotique sont repiquées dans du terreau puis transférées en serre.

Les graines T2 issues de l'autofécondation des transformants primaires sont semées sur un milieu sélectif contenant 100 µg.ml<sup>-1</sup> de kanamycine afin de sélectionner les lignées présentant une ségrégation 3:1 (Kana<sup>R</sup>: Kana<sup>S</sup>) caractéristique de l'insertion du ou des ADN-T à un seul locus dans le génome. Afin d'acquérir des lignées T3 homozygotes, huit plantes Kana<sup>R</sup> (pour chaque lignée T2 sélectionnée) sont transférées en serre afin d'obtenir les graines issues de l'autofécondation. Ces graines sont à nouveau semées sur milieu sélectif et les lignées homozygotes pour la résistance à l'antibiotique sont sélectionnées (figure 3.10).

### **2.3.3. Détermination du nombre d'insertion de l'ADN-T par Southern-blot.**

Ces expériences n'ont été menées que pour les lignées exprimant PsTI.

#### *2.3.3.1.Extraction d'ADN génomique*

L'ADN génomique est extrait selon le protocole de Doyle et Doyle (1990), 0,5 à 2 g de matériel végétal est broyé au mortier en poudre fine dans l'azote liquide. La poudre est transférée dans un tube dans lequel est ajouté 7,5 ml de tampon d'extraction CTAB (CTAB 2% (w/v) ; NaCl 1,4 M ; EDTA 20 mM ; Tris-HCl 100 mM pH 8 ; β-mercaptoéthanol 0,2%

(v/v)) préalablement chauffé à 60°C. Après 30 minutes d'incubation à 60°C, 7,5 ml d'un mélange chloroforme:alcool isoamylique (24:1) est ajouté. Le mélange est ensuite centrifugé (10 min à 4100 g). La phase aqueuse contenant les acides nucléiques est reprise et additionnée de 6 ml d'isopropanol avant de subir une nouvelle centrifugation (20 min à 8000 g). Le culot d'ADN est lavé avec un mélange d'éthanol 76% et d'acétate d'ammonium 10mM, puis centrifugé (5 min à 8000 g). L'ADN est repris dans 500 µl de tampon TE (Tris-HCl 10 mM pH 8 ; EDTA 1 mM pH 8). L'ADN est précipité une nouvelle fois avec 50 µl d'acétate de sodium 3 M et 500 µl d'isopropanol puis centrifugé (5 min 8000 g). Après un dernier rinçage à l'éthanol 70%, le culot est séché et repris dans 100 µl de tampon TE par gramme de tissus végétaux de départ.

### *2.3.3.2. Analyse en Southern Blot*

Un microgramme d'ADN génomique est digéré par 5 unités d'endonucléase de restriction *XbaI* dans le tampon du fabricant, en présence de spermidine 5 mM. Les produits de digestion sont séparés sur gel d'agarose à 1,2%. Après électrophorèse toute la nuit à 30 mV, Les molécules d'ADN sont dépurinées par un traitement de 15 minutes dans un bain de HCl 0,15M. Le transfert alcalin sur membrane de Nylon Hybond N+™ (Amersham) est réalisé avec une solution de NaOH 0,4 N pendant 4 à 6 heures. La membrane est rincée dans une solution de 2X SSC (NaCl 300 mM ; Tris-sodium citrate 30 mM) pendant 5 minutes, la membrane est conservée à 4°C jusqu'à l'hybridation.

L'hybridation est réalisée selon le protocole décrit par Church et Gilbert (1984). La membrane est pré-hybridée pendant 3 à 4 heures à 65°C avec la solution d'hybridation (SDS 7% ; Na<sub>2</sub>HPO<sub>4</sub> 0,25 M pH 7,4 ; EDTA 2mM ; Héparine 200 µg.ml<sup>-1</sup> ; ADN de sperme de saumon soniqué et dénaturé 100 ng.ml<sup>-1</sup>). La sonde PsTI obtenue par PCR et marquée au α-<sup>32</sup>P avec le kit de marquage aléatoire (Promega) selon les conditions recommandées par le fournisseur est ajoutée après dénaturation alcaline. L'hybridation s'effectue pendant 16 heures, à 65°C, sous agitation.

La sonde radioactive en excès est éliminée par une succession de 2 rinçages à 65°C avec une solution de 1X SSC/0,1% SDS puis avec un solution de 0,1X SSC/0,1% SDS. La membrane est ensuite placée dans une cassette d'autoradiographie en contact avec le film Kodak X-Ray et placée entre 2 écrans intensificateurs à -80°C. Le film est développé après 4 jours d'exposition.

### **2.3.4. Etude de l'expression du transgène**

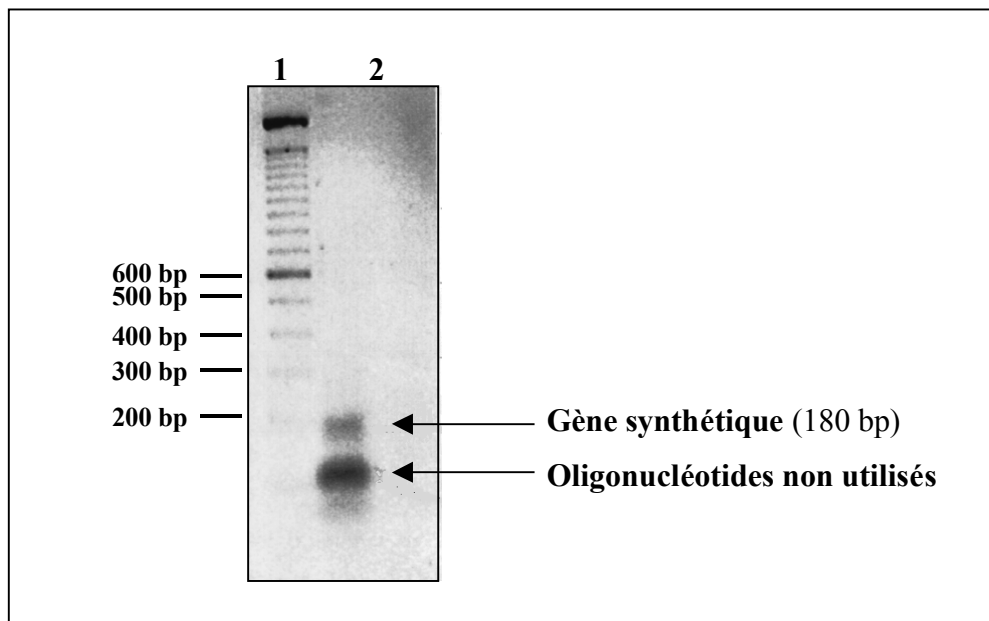
Afin de quantifier l'expression de PsTI dans les lignées recombinantes par analyse en Dot-blot, les protéines solubles sont extraites de 300 µg de parties aériennes des plantes avec 500 µl de tampon d'extraction (Tris-HCl 100 mM pH 7,6 ; EDTA 50 mM ; NaDETC 0,5% (w/v) ; PVPP 10%) après broyage dans l'azote liquide. Après centrifugation (15 min, 8000g, 4°C), les protéines sont quantifiées selon la méthode de Bradford (1976).

Une quantité croissante d'extrait végétal (1 à 10 µg) ainsi que la protéine PsTI pure (50 à 1000 ng) sont diluées dans 100 µl de tampon d'extraction. Le transfert des protéines sur une membrane de nitrocellulose Hybond-C (Amersham) est réalisé au moyen du Bio-Dot® (Bio-Rad) en condition semi-sèche d'après le protocole préconisé par le fournisseur. La membrane est incubée 1 heure dans une solution PBS contenant 0,5% de Tween 20 et 5% de BSA (PBS-T-BSA) afin de masquer les éventuels sites non spécifiques. La membrane est ensuite mise en présence de l'anticorps anti-PsTI dilué au 1/1000<sup>ème</sup> dans la solution PBS-BSA pendant 1 heure à température ambiante puis toute la nuit à 4°C. Après 3 rinçages de 5 minutes à température ambiante avec une solution de PBS-T, l'anticorps secondaire couplé à la peroxydase (anticorps polyclonal anti-lapin, Sigma) est dilué au 1/10000<sup>ème</sup> dans la solution de PBS-T. L'incubation se déroule pendant 1 heure à température ambiante. Cette étape est suivie de 3 lavages de 5 minutes avec la solution de PBS-T. La révélation est ensuite réalisée au moyen des réactifs Covalight A et B (Valbiotech) selon les consignes du fabricant. Les signaux sont visualisés après exposition d'un film Kodak X-Ray.

### **2.3.5. Quantification du niveau d'expression de PsTI.**

Les niveaux d'expression des lignées ont été estimés grâce au logiciel Molecular Analyst v.2.1.1. software (Bio-rad), par une analyse densitométrique du film autoradiographique obtenu lors de l'analyse en Dot-Blot. L'intensité des signaux obtenus pour les différentes lignées transgéniques est comparée à l'intensité des signaux correspondants à des quantités de PsTI connues.





**Figure 3.11 : Synthèse du gène synthétique**

1. Marqueur de poids moléculaire
2. Produits de la réaction de ligation des 10 oligonucléotides. Ces oligonucléotides ont été mélangés en quantité équimolaire. Après dénaturation à 80°C, les oligonucléotides s'hybrident entre eux au cours de la décroissance de la température. Les extrémités de deux oligonucléotides consécutifs sont liées entre elles grâce à une ligase thermostable.

## Résultats et Discussion

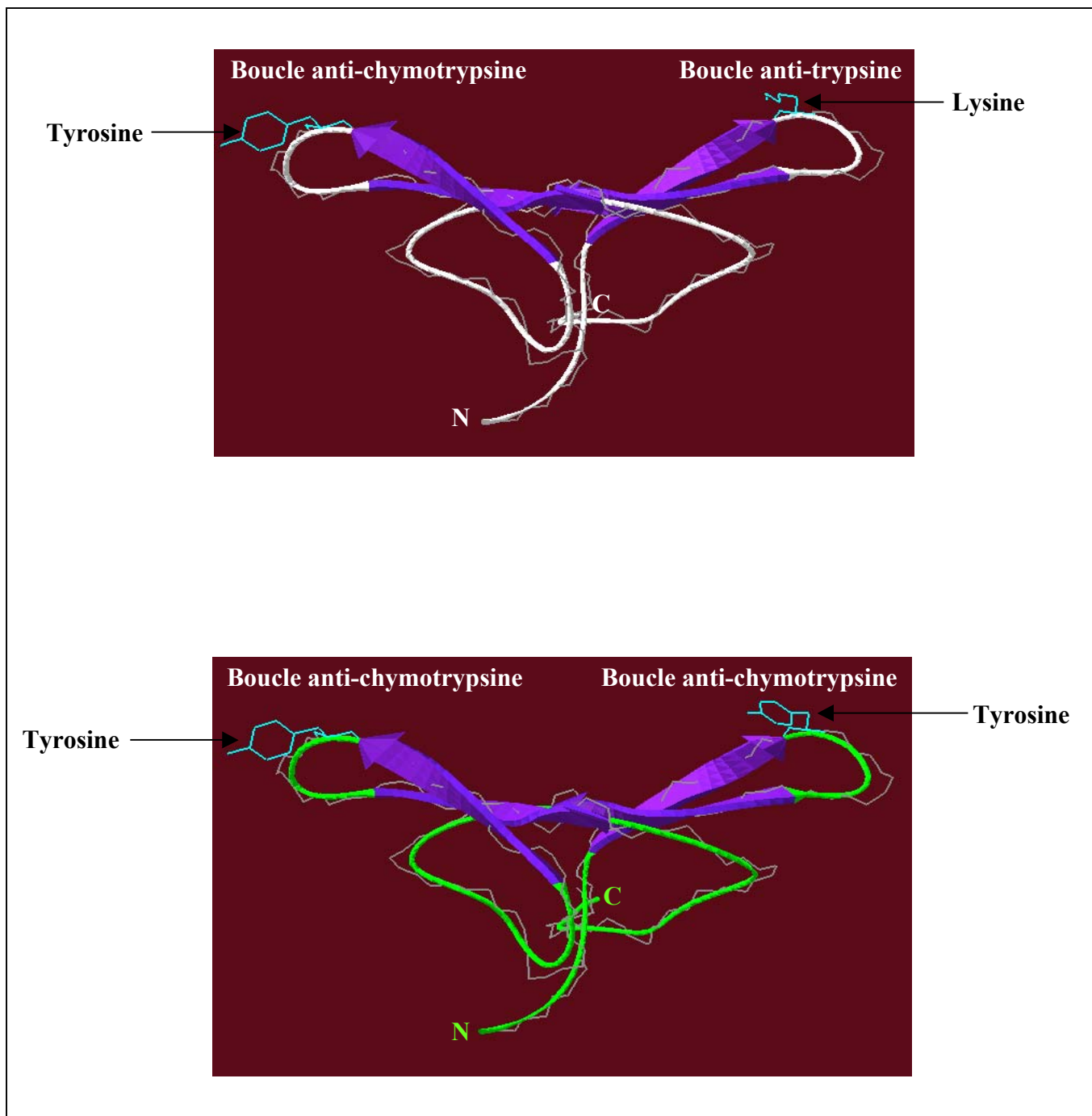
### 3.1. Construction des séquences synthétiques

#### 3.1.1. Le gène codant *PsTI*

Les progrès importants réalisés ces dix dernières années concernant les techniques de synthèse d'oligonucléotides ont rendu possible la réalisation de longs gènes totalement synthétiques. L'ADN à synthétiser est découpé sous forme d'oligonucléotides simple brin, synthétisés chimiquement, qui sont assemblés. La méthode d'assemblage classique développée par Khorana en 1978 n'a pas subi de changements substantiels et est toujours la technique la plus utilisée dans le domaine. Celle-ci nécessite la synthèse d'un ensemble d'oligonucléotides phosphorylés chevauchants recouvrant l'ensemble des deux brins. Grâce à leur complémentarité, la dénaturation suivie d'un refroidissement progressif de la solution d'oligonucléotides résulte de l'appariement correct et la formation d'une molécule hybride. Occasionnellement des "gaps" peuvent être introduits du fait de la faible température à laquelle s'effectue la réaction d'hybridation mais optimale pour la ligase (15 à 22°C). L'utilisation d'une ligase thermostable (active à 65°C) permet d'accroître la spécificité d'hybridation.

Dix oligonucléotides de 30 à 40 résidus nucléotidiques ont été définis selon la séquence protéique de PsTI-2 en s'appuyant sur la séquence nucléotidique connue d'une isoforme de PsTI (PsTI-1, X83211, Domoney, 1995). Ce choix permet d'éliminer la redondance de la 3<sup>ème</sup> base du codon, et d'optimiser l'usage des codons chez les eucaryotes (figure 8).

Une quantité équimolaire de chaque oligonucléotide a été mélangée et leur liaison a été réalisée à haute température (60°C) grâce la ligase thermostable. Après vérification sur électrophorèse, une seule bande de 190 pb est observée (figure 3.11). Cette molécule d'ADN a été clonée dans le vecteur pBS KS<sup>-</sup> et la séquence a été vérifiée. Aucune mutation n'a été observée dans le gène synthétique.



**Figure 3.12 : Structure comparée de PsTI et de PsCI.**

PsTI est représentée en blanc (1PBI) et PsCI en vert (modélisation). Les acides aminés catalytiques (site P1) sont représentés avec la structure complète de leur chaîne latérale.

### 3.1.2 Mutagenèse dirigée afin d'obtenir le gène PsCI.

Les inhibiteurs de protéases de la famille de Bowman-Birk possèdent des ponts disulfure qui entraînent la formation d'une structure rigide et symétrique basée sur 2 domaines (Werner, 1992). Chaque boucle inhibitrice est constituée de neuf acides aminés fermés par un pont disulfure. L'identité du résidu aminé en position P1 de chaque domaine inhibiteur détermine la spécificité de la protéase à sérine inhibée, il possède une lysine ou une arginine pour la trypsine, et une leucine ou une tyrosine pour la chymotrypsine (Ikenaka, 1986).

Afin de créer une nouvelle boucle antichymotrypsine, nous souhaitons modifier des acides aminés dans la boucle antitrypsine afin d'accroître son affinité pour la chymotrypsine. Nous avons donc défini deux nouveaux oligonucléotides complémentaires afin de changer la lysine de la boucle antitrypsine en tyrosine (résidu P<sub>1</sub> de la boucle) mimant la boucle antichymotrypsine de PsTI. La séquence de la boucle putative antichymotrypsine est donc : CTYSNPPTC.

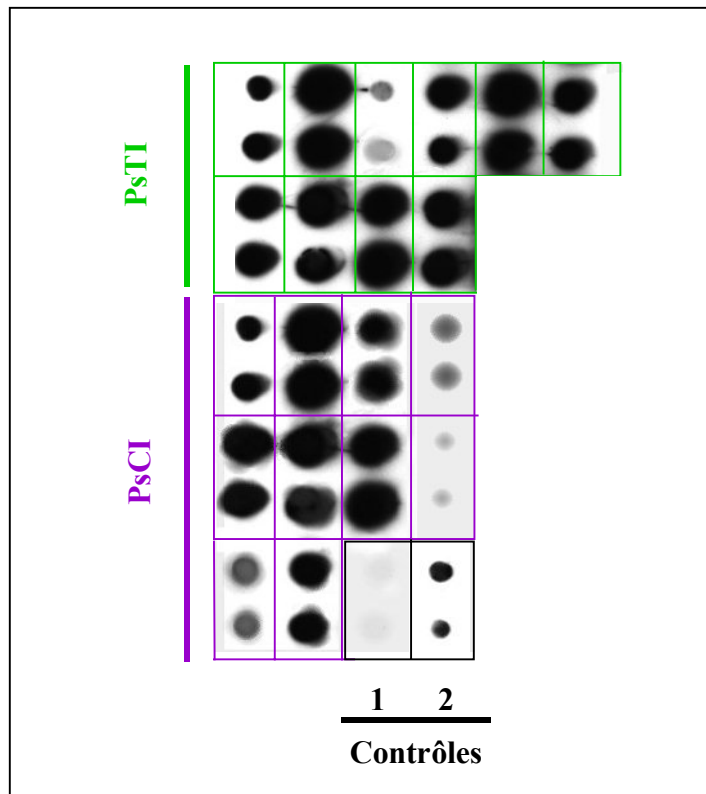
La reconnaissance des inhibiteurs de type Bowman-Birk avec la chymotrypsine a été étudiée grâce à 8000 variants générés à partir de la séquence des 9 acides aminés, correspondant à la boucle antitrypsine de BBI (McBride, 1996). Les séquences qui présentent l'affinité la plus forte pour la chymotrypsine sont les séquences qui renferment, en complément de la tyrosine en position P<sub>1</sub>, une thréonine en P<sub>2</sub>. La thréonine est un petit acide aminé dont la structure est parfaitement complémentaire du site actif de l'enzyme et peut ainsi établir de nombreuses liaisons hydrogènes avec les acides aminés de la poche afin de maintenir la conformation de la boucle inhibitrice (rigidité) (McBride, 1998). Ainsi la boucle que nous avons créée présente normalement une séquence idéale pour inhiber la chymotrypsine (figure 3.12).

Le gène synthétique PsCI a été construit de la même façon que celui de PsTI.

## 3.2. Expression de PsTI et PsCI dans la levure *Pichia pastoris*

### 3.2.1 Construction des vecteurs binaires

Le système d'expression dans la levure *Pichia pastoris* a été utilisé afin de produire en masse les deux isoformes de l'inhibiteur. Les gènes synthétiques *PsTI* et *PsCI* ont été clonés entre les sites *EcoRI* et *SalI* dans un vecteur binaire pPICZ $\alpha$ A permettant l'expression de la protéine exogène.



**Figure 3.13 : Mise en évidence des protéines PsTI ou PsCI dans le milieu de culture de levure en dot blot.**

Chaque extrait de milieu de culture (100µl) est déposé en deux répétitions sur le Dot-blot

- (1) Contrôle négatif : milieu de culture de levure non transformée (100 µl).
- (2) PsTI pure (50 ng)

Le gène synthétique a été fusionné avec un peptide de sécrétion responsable de l'adressage de la protéine recombinante dans le milieu de culture. La présence de l'inhibiteur de protéase dans le milieu extérieur permet d'éliminer les interactions, et ainsi la mobilisation de l'IP par les protéases à sérine cellulaires lors de la destruction des cellules au cours de l'extraction.

L'utilisation du site *EcoRI* situé après le site de coupure du peptide d'excrétion entraîne l'addition d'une séquence de 12 nucléotides codant 4 acides aminés (YVEF). Lors de la maturation de la protéine, le peptide signal est coupé par l'enzyme KeX2C et l'extrémité N-terminale de l'IP est donc additionnée de ces 4 résidus. La conservation du codon stop dans la séquence du gène synthétique permet de ne pas modifier la protéine en C-terminal avec une queue histidine ou un épitope myc ; c'est pourquoi nous avons conservé le codon stop du gène codant l'IP. Nous possédons des anticorps, ainsi que toutes les caractéristiques biochimiques permettant de purifier l'IP.

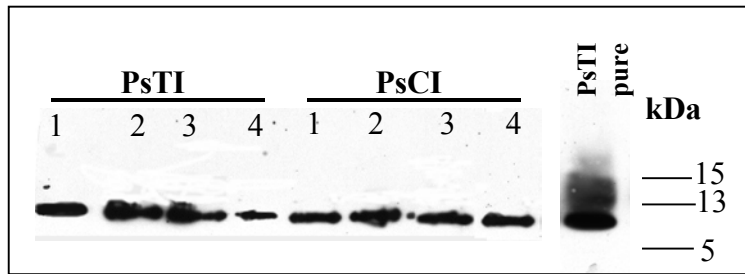
Les acides aminés ajoutés en N-terminal ne devraient pas être responsables d'une inactivité puisque de nombreux inhibiteurs de la famille de Bowman-Birk possèdent une extension en N-terminal. Par contre, l'extrémité C-terminale vient se replier vers le site antitrypsine. L'addition de motif spécifique à la purification pourrait être responsable d'une diminution d'activité par encombrement stérique.

### **3.2.2. Expression et activité biologique des protéines recombinantes**

Le plasmide pPICZ $\alpha$ A –IP linéarisé au niveau du site de restriction *SalI* a été introduit dans la souche GS115 de *Pichia*. Les clones recombinants, résistants à la Zéocine, exprimant l'IP à un niveau élevé ont été sélectionnés (figure 3.13). Dix clones pour chaque construction ont été analysés en Dot-blot afin d'estimer le niveau d'expression de la protéine recombinante dans le milieu de culture. La quantité d'inhibiteur a été estimée en comparaison avec une gamme de PsTI pur. Après induction de la transcription de l'IP par addition de méthanol dans le milieu de culture pendant 4 jours, la quantité d'IP est comprise entre 0,4 à 1-2 mg.l<sup>-1</sup> de milieu de culture selon les clones (résultat non présenté).

Les protéines du surnageant d'une culture des 4 clones exprimant au niveau le plus important l'IP ont été analysées en Western-blot. Une bande unique est observée à 7 kDa (figure 3.14).

Un test d'inhibition de la trypsine ou de la chymotrypsine a été réalisé afin de vérifier l'activité biologique des protéines recombinantes. Aucun de ces clones ne présente d'activité inhibitrice de la trypsine ou de la chymotrypsine (résultats non présentés).



**Figure 3.14 : Mise en évidence des protéines PsTI et PsCI dans le milieu de culture de levure**

50  $\mu$ l de milieu de culture des levures recombinantes exprimant les protéines PsTI et PsCI sont déposés sur le gel.

La bande unique de 7 kDa que l'on observe en Western-blot à partir d'un extrait de surnageant de culture de levure révèle qu'aussi bien PsTI que PsCI sont sécrétés dans le milieu extérieur. D'autre part, la taille observée correspond à la masse moléculaire de la protéine native. Le peptide signal de 88 acides aminés est donc correctement clivé. Par contre aucune activité IP à sérine n'a pu être mise en évidence. La structure des domaines inhibiteurs est engendrée par la présence de deux prolines qui permettent de former une boucle ainsi que par la présence d'un pont disulfure qui maintient et referme la boucle. La protéine possède une structure compacte et rigide soutenue par sept ponts disulfure. L'absence d'activité pourrait résulter d'un agencement incorrect de la structure tridimensionnelle de la protéine, bien que l'obtention de protéines actives renfermant plusieurs ponts disulfures soit possible par expression hétérologue dans la levure *Pischia pastoris* (Volpicella, 2000). Afin de remédier à ce problème, plusieurs solutions peuvent être considérées. Des techniques de dénaturation puis renaturation peuvent être envisagées. Une étape de purification par affinité peut être attractive. En effet, au contact de la protéase, la molécule inactive change de conformation et acquiert une activité inhibitrice (Flecker, 1989), une colonne de d'anhydrotrypsine et/ou d'anhydrochymotrypsine pourrait être employée afin de renaturer ces protéines recombinantes.

### **3.3. Expression hétérologue de PsTI et PsCI dans la plante *Arabidopsis thaliana***

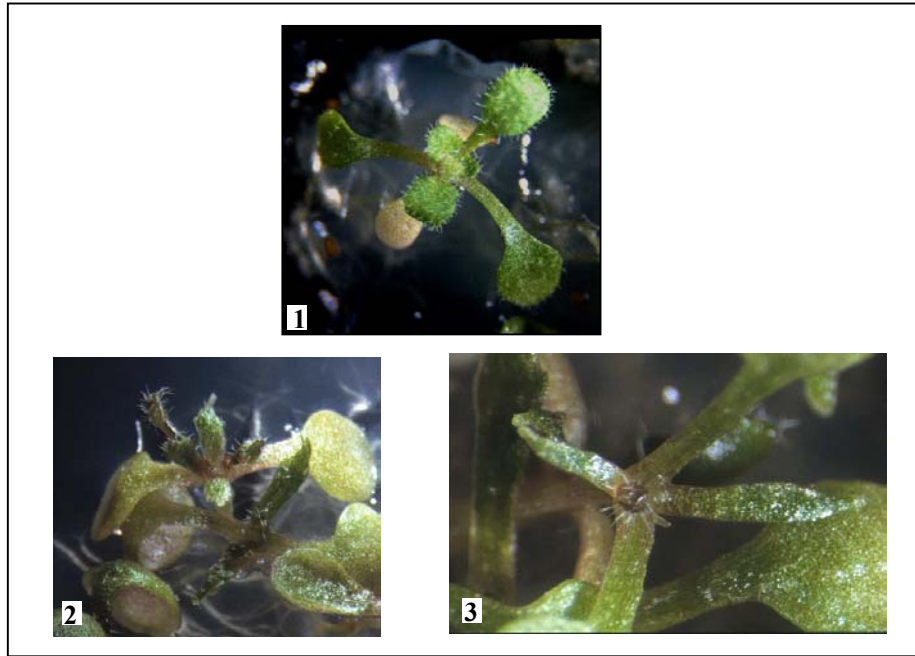
#### **3.3.1 Transformation d'*Arabidopsis thaliana* et sélection des lignées homozygotes**

Deux vecteurs binaires, contenant *PstI* ou *PsCI*, ont été réalisés afin de permettre leur expression dans des plantes d'*A. thaliana*. Les gènes synthétiques ont été placés sous le contrôle du promoteur 35S fusionné avec une séquence amplificatrice pour permettre une expression constitutive des IP.

*Arabidopsis thaliana* été transformé par l'intermédiaire d'*Agrobacterium tumefaciens*, le gène de résistance à la kanamycine et le gène codant l'IP ont été ainsi introduits dans le génome de façon aléatoire.

Dans le cas de *PsTI*, vingt deux transformants primaires ont été sélectionnés (tableau 1). Parmi les plantes T2 obtenues, quatre lignées sont stériles et deux lignées présentent des





**Figure 3.15 : Phénotype anormal des transformantes 35S<sup>2</sup>::PsTI**

1. Plantule témoin.
- 2, 3. Plantules 35S<sup>2</sup>::PsTI. Les cotylédons sont normaux alors que les premières feuilles sont lancéolées.

anomalies structurales, les feuilles étaient lancéolées (figure 3.15). De plus, ces plantes ne se développent pas en serre.

Dans le cas de *PsCI*, malgré la répétition de l'expérience de transformation, seulement trois lignées recombinantes ont pu être sélectionnées. Ces plantes ne présentent aucun phénotype.

**Tableau 3.1 : Récapitulatif des résultats de transformation d'*A. thaliana***

lignées	Nombre de transformants primaires	Lignées à ségrégation 3:1 (Kana <sup>R</sup> :Kana <sup>S</sup> )	Nombre de lignées T3 homozygotes sélectionnées
35S <sup>2</sup> ::PsTI	22	6	3
35S <sup>2</sup> ::PsCI	3	2	2

### 3.3.2. Caractérisation des lignées T3 homozygote contenant PsTI.

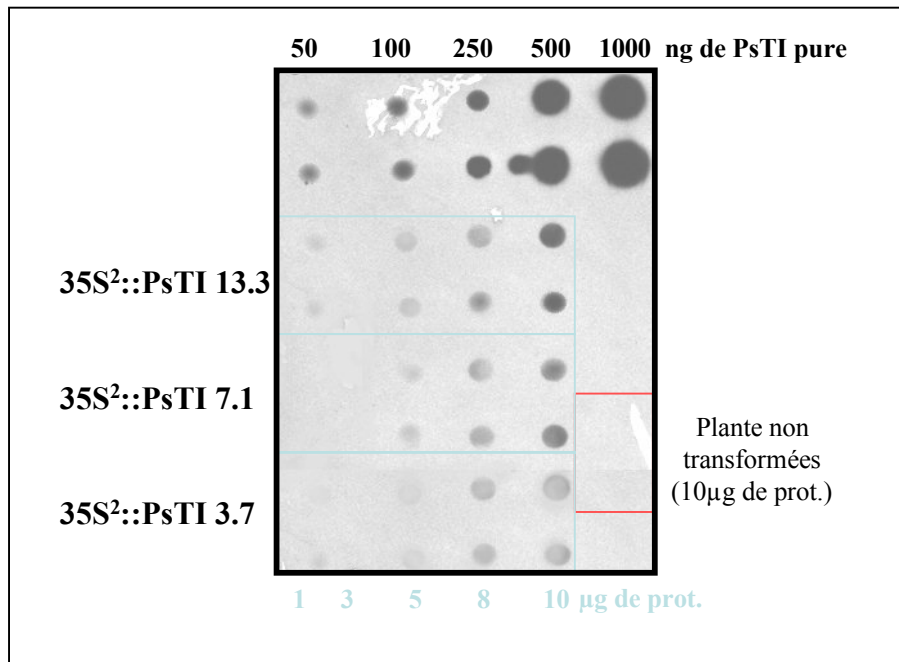
L'étude de l'expression de PsTI dans les plantes transformées a permis la sélection de 3 lignes T3 homozygotes. Elles permettent d'établir une gamme d'expression de l'IP afin de pouvoir corrélérer le gain de résistance éventuel de la plante avec la quantité d'IP exogène, lors de tests insectes effectués ultérieurement.

L'analyse en Dot-blot permet d'estimer les niveaux d'expression compris entre 0,8 et 2% des protéines totales solubles (figure 3.16). Le nombre d'ADN-T insérés dans le génome de la plante varie de 1 à 3 dans ces différentes lignées (tableau 3.2).

L'analyse des trois lignées contenant le gène *PsCI* est en cours.

**Tableau 3.2 : Caractérisation moléculaire des lignées homozygotes exprimant PsTI**

Lignées T3 homozygotes	Nombre de copies d'ADN-T	Niveaux d'expression (% protéines totales solubles)
35S <sup>2</sup> ::PsTI 3.7	1	0,8
35S <sup>2</sup> ::PsTI 7.1	1	1,3
35S <sup>2</sup> ::PsTI 13.3	3	2



**Figure 3.16 : Quantification en Dot-blot de l'expression de PsTI dans les plantes transformées d'*A. thaliana*.**

Les extraits de protéines végétales solubles totales sont déposées en répétition (n=2) sur la membrane. La quantification du signal est faite grâce à une gamme de PsTI pure déposée en quantité croissante.

Les résultats obtenus montrent qu'il est possible d'exprimer des inhibiteurs de la famille des Bowman-Birk chez *A. thaliana*. Les niveaux d'expression sont comparables avec ceux obtenus chez d'autres modèles végétaux exprimant CpTI (inhibiteur antitrypsine) sous le contrôle du même promoteur (Hilder, 1987 ; Gatehouse, 1997).

Cependant la caractérisation de ces lignées exprimant PsTI ou PsCI a permis de mettre en évidence certaines anomalies liées à l'expression de ces IP.

A notre connaissance, aucun inhibiteur de chymotrypsine n'a été sur-exprimé chez *A. thaliana* ou dans une autre plante. Les lignées d'*A. thaliana* exprimant l'inhibiteur anti-trypsine MTI-2 ne présentent aucun phénotype (De Leo, 1998). Les phénomènes de stérilité, de phénotypes de désorganisation de la croissance foliaire chez les lignées PsTI et le faible nombre de transformants primaires obtenus lors de la transformation avec le gène *PsCI*, révèlent peut-être une forme de toxicité du domaine antichymotrypsine. Comme tout organisme vivant, les plantes possèdent des protéases impliquées dans le recyclage des protéines mais aussi dans des processus beaucoup plus spécialisés tel que la sénescence (Solomon, 1999), la xylogénèse (Groove, 1997), la réponse aux stress environnementaux, et la germination (dégradation des protéines de réserve). Le domaine antichymotrypsine pourrait interagir avec une protéase impliquée dans le contrôle de la croissance foliaire et/ou la formation des graines.

Des comparaisons de séquences entre le gène codant l'IP et le génome complet d'*A. thaliana* ont été réalisées, aucune homologie de séquences n'a été observée. Ces altérations du développement ne paraissent donc pas devoir être liées à des phénomènes d'extinction de gènes.

Après avoir vérifié l'activité biologique des inhibiteurs à partir d'extraits foliaires, des tests insectes seront mis en place afin d'estimer le gain de résistance des plantes transgéniques face à la présence des pucerons.

– **DISCUSSION GENERALE & CONCLUSIONS** –

**Tableau 4.1 : Récapitulatif des principaux résultats obtenus**

	<i>A. pisum</i> oligophage	<i>A. gossypii</i> polyphage	<i>M. persicae</i> très polyphage
Activité portéolytique principale du tube digestif	Cystéine endoprotéase	Cystéine endoprotéase	Cystéine endoprotéase
Protéases clonées	Cat L	CatL, 2 CatB	CatL
Effet d'OC-I <i>in planta</i>	non testé	à tester sur melon	> 25% de réduction de la fécondité
Effet d'EIM <i>in vitro</i>	LC50=193 µg.ml <sup>-1</sup>	LC50=204 µg.ml <sup>-1</sup>	LC50>1000 µg.ml <sup>-1</sup>
Effet de PsTI <i>in vitro</i>	IC50=320 µg.ml <sup>-1</sup>	IC50>1000 µg.ml <sup>-1</sup>	IC50>1000 µg.ml <sup>-1</sup>

Les pucerons sont des ravageurs majeurs des cultures en climat tempéré. Leur pullulation occasionne périodiquement des dégâts directs importants sur la plupart des espèces végétales cultivées. Malgré leur importance économique, les pucerons ont été peu étudiés dans le contexte des plantes transgéniques résistantes aux insectes du fait de leur adaptation à une niche alimentaire extrêmement spécialisée, la sève phloémienne des plantes. Ce trait pose à la fois des problèmes d'expression ciblée, mais aussi d'identification de toxines adéquates par des bioessais plus délicats que ceux développés pour des insectes phytophages généralistes. Cet aliment est très riche en sucres, dépourvu de lipides et très déséquilibré en acides aminés<sup>27</sup>. Comme les pucerons étaient considérés comme ne possédant pas de protéases digestives, les stratégies de transgénèse basées sur les inhibiteurs de protéases, largement étudiés par ailleurs comme facteurs de défense contre divers insectes phytophages, ne paraissaient pas adaptées à ce groupe d'insectes. Or, au cours de l'évaluation d'un certain nombre de protéines pour leur intérêt dans la défense contre les pucerons, certains IP ont révélé un potentiel intéressant. Ces molécules représentent donc des gènes originaux pour lutter contre les pucerons *via* la transgénèse. Il est toutefois indispensable d'en comprendre le mode d'action afin d'en maîtriser l'introduction raisonnée dans des programmes de lutte variétale.

Cette thèse a donc un double objectif : combler l'absence totale d'informations sur la protéolyse digestive des Homoptères et accroître le nombre de gènes d'inhibiteurs de protéases disponibles pour lutter contre le puceron du melon, *A. gossypii*. Ce travail a parfois été élargi à deux autres pucerons : le puceron du pois, *A. pisum*<sup>28</sup> et, dans le cadre du programme européen, le puceron du pêcher *M. persicae*<sup>29</sup>. Une présentation des principaux résultats est présentée sous forme de tableau récapitulatif (tableau 4.1). Ces trois pucerons sont inféodés à des plantes d'espèces différentes et sont, selon l'espèce, oligophages ou polyphages. L'obtention de résultats souvent similaires chez ces trois pucerons nous permet parfois de généraliser certaines conclusions (IP à cystéines), mais montre également leurs particularités (IP à sérine).

---

<sup>27</sup> 70% de l'azote circule sous forme de quelques acides aminés (généralement les amides Asn ou Gln), et les protéines représentent moins de 1% des acides aminés circulant.

<sup>28</sup> espèce modèle en physiologie et génétique des Homoptères.

<sup>29</sup> qui est l'espèce d'aphide la plus nuisible au niveau mondial.

## 1. Les protéases digestives du puceron *A. gossypii*

Notre étude a permis d'apporter un regard nouveau sur la physiologie digestive du puceron. Nous avons mis en œuvre plusieurs méthodologies afin d'identifier le spectre complet des protéases digestives utilisées par le puceron. Les pucerons ont manifestement adopté un processus de digestion spécifique ; l'activité endoprotéasique majeure n'y est pas sécrétée, et existe préférentiellement dans les cellules du tube digestif, et plus particulièrement du mésentéron antérieur. Aucune activité protéase à sérine n'a été mise en évidence dans ce tissu, et la protéolyse y semble assurée exclusivement par des protéases à cystéine. L'activité protéase à cystéine chez le puceron du melon est représentée majoritairement par les cathepsines B (70%, deux ADNc clonés) et les cathepsines L (20%, un ADNc isolé).

Chez les vertébrés, les cathepsines sont exprimées dans un grand nombre de types cellulaires, et leur fonction principale repose sur une dégradation non sélective des protéines à l'intérieur du lysosome. Toutefois, certaines cathepsines possèdent des spécificités tissulaires révélant leur implication dans des processus cellulaires beaucoup plus spécialisés que le catabolisme des protéines (Turk, 2000). L'étude de la proséquence de la cathepsine L du puceron du melon suggère un adressage de la protéine aux lysosomes<sup>30</sup>, les figures d'immunolocalisation de la cathepsine L montrent que la protéase est intracellulaire. Sa localisation sub-cellulaire pourra être vérifiée par immunolocalisation au niveau cellulaire (microscopie électronique). Ce travail a été initié grâce à une collaboration avec un laboratoire brésilien, et montre qu'en plus d'une localisation intracellulaire vésiculaire, la cathepsine L du puceron semble associée à des structures membranaires apicales (Cristofolletti, en préparation).

La répartition tissulaire de la cathepsine L, observée par immunomarquage en microscopie confocale, révèle que cette protéase est exprimée au niveau de « l'estomac », ainsi que dans les bactériocytes, cellules spécialisées qui abritent le symbiote obligatoire des pucerons, *Buchnera aphidicola*. Ces deux types cellulaires sont impliqués dans la physiologie de l'assimilation des acides aminés chez le puceron, pour l'instant fortement centrée sur le transport actif des acides aminés libres<sup>31</sup>. La cathepsine L, et les cathepsines B clonées à partir

---

<sup>30</sup> bien que ce trait ne soit que potentiel, et basé sur un mode d'adressage lysosomal qui n'a pas été formellement identifié chez les insectes (Aeed *et al.*, 1994).

<sup>31</sup> transport bidirectionnel pour ces deux tissus : assimilation des acides aminés exogènes ou produits par le symbiote (flux vers l'hôte), ou excrétion azotée et alimentation du symbiote en métabolites aminés pour lequel il est auxotrophe (flux venant de l'hôte).



d'une banque soustractive spécifique des messagers du tube digestif, sont probablement impliquées dans un processus cellulaire spécialisé que représente la digestion. Les pucerons utiliseraient les protéases lysosomiales afin de dégrader les protéines phloémiennes ingérées, comme cela a été montré pour un aleurode (voir plus loin).

Nos expériences d'immunomarquage associées à OC-I ou à PsTI dans le puceron révèlent que ces protéines sont ingérées et pénètrent dans les cellules du tube digestif. D'autre part, en condition d'alimentation artificielle, PsTI n'est pas retrouvé dans le miellat, suggérant qu'il est utilisé par le puceron, contrairement à de nombreux polypeptides testés dans des conditions analogues (Rahbé, 1995). La possibilité de métabolisation des protéines ingérées a été étudiée chez un autre homoptère, *Bemisia argentifolii*. Grâce à des protéines de feuilles de coton radiomarquées et solubilisées dans l'alimentation artificielle, les auteurs ont pu montrer que les protéines sont ingérées et présentes dans le tube digestif de l'aleurode. Elles sont digérées, dégradées en acides aminés libres et utilisées pour la synthèse *de novo* de protéines (Salvucci, 1998). Les homoptères semblent donc bien posséder un système digestif fonctionnel de protéolyse.

A la lumière de ces résultats, il est légitime de s'interroger sur les mécanismes de prise en charge des polypeptides pour pénétrer dans les entérocytes et sur les mécanismes de transport jusqu'au lysosome. Chez les vertébrés, certaines protéines extracellulaires sont captées dans les cellules par un processus d'endocytose. Les vésicules d'endocytose sont fusionnées avec les lysosomes et leur contenu est dégradé. Ce mécanisme très efficace pourrait aussi exister chez le puceron.

Toutefois, la digestion des protéines ne doit pas représenter une source suffisante d'acides aminés essentiels, car tous les pucerons vivent en symbiose obligatoire avec une bactérie intracellulaire<sup>32</sup>. Ces bactéries symbiotiques des Aphididae ont un rôle manifeste dans le métabolisme intermédiaire, conduisant à une composition équilibrée et régulée du pool d'acides aminés libres utilisés par le puceron pour ses synthèses protéiques. Certes, les

---

<sup>32</sup> Cela est même plus largement le cas de tous les homoptères phloémophages, obligatoirement symbiotiques (dont les aleurodes). Le cas de taxons d'aphides écologiquement et phylogénétiquement divergents, comme le phyloxera (Aphidomorpha : Phylloxeroidea) pourraient bien éclairer l'histoire évolutive de la digestion chez ce groupe d'insectes, car ils ne sont ni phloémophages ni symbiotiques, et s'alimentent sur les contenus cellulaires riches en protéines. L'étude de leur physiologie protéolytique serait d'une aide précieuse pour comprendre la situation prévalant chez les pucerons modernes (Aphidomorpha : Aphidoidea : Aphididae).

pucerons possèdent une digestion protéolytique, mais est-ce un reliquat d'un processus physiologique présent chez leur ancêtre et chez tout animal, ou bien est-ce plutôt un mécanisme de défense permettant d'éliminer les polypeptides entomotoxiques véhiculés dans le système vasculaire de la plante pour se défendre ? L'implication des cathepsines dans ce mécanisme de détoxification pourrait être testé *via* la création de pucerons " transgéniques ". Un virus d'invertébré circulant pourrait être utilisé afin d'éteindre le gène de la cathepsine par un phénomène de RNAi, et la survie *in vitro* ou *in planta* des pucerons infectés pourrait être analysée.

## 2. L'utilisation des inhibiteurs de protéases pour lutter contre les pucerons.

Les inhibiteurs de protéases sont des composés essentiels dans les mécanismes de défense des plantes contre les insectes phytophages. Ainsi, la sur-expression d'inhibiteurs de protéases grâce à la transgénèse ne fait qu'accentuer une stratégie déjà existante chez les plantes. Les voies signalétiques, induites chez les plantes lors d'une attaque de pucerons, ne sont pas encore totalement élucidées. Il a pu être mis en évidence que, bien que n'ayant pas d'effet direct sur la physiologie du puceron, une serpine phloémienne est produite lors d'une attaque de la courge par les pucerons (Yoo, 2000). Les inhibiteurs de protéases font donc partie de l'arsenal de défense mis en place par la plante pour lutter contre les pucerons. Terra (1996) a émis l'hypothèse que les Héteroptères ne possédaient plus de protéases à sérine afin de s'adapter à une alimentation végétale (sèves) riche en inhibiteurs de protéases à sérine.

Dans une première étude, il a pu être montré qu'il était possible d'exprimer des inhibiteurs de protéases à cystéine dans le phloème du colza. Les pucerons, qui alimentent sur les plantes transgéniques exprimant l'oryzacystatine I présentent une fécondité réduite. Ce paramètre est très important en génétique des populations, une baisse de la fécondité peut permettre de limiter l'expansion de la population et de préserver ainsi une partie des cultures. Les inhibiteurs de type phytocystatines peuvent être considérés comme des gènes de défense anti-pucerons. Actuellement, les stratégies de lutte contre les insectes par transgénèse sont basées sur la stratégie "forte dose" permettant d'éliminer rapidement tous les insectes cibles. Toutefois ce n'est pas la seule stratégie envisageable pour la création et/ou la sélection

variétale. Le cumul de plusieurs gènes, à faible potentiel individuel, peut permettre d'observer un gain de résistance significatif et d'abaisser la pression de sélection conduisant à l'apparition d'insectes résistants.

Bien qu'OC-I ne soit pas responsable d'effets drastiques sur la population du puceron *M. persicae*, les inhibiteurs de protéases à cystéine sont tout de même des molécules prometteuses car il semble que dans le tube digestif, première barrière obligée pour ce type de molécules, ne soient présentes que des protéases à cystéine. Les cibles de ces inhibiteurs de protéases à cystéine sont désormais bien caractérisées (cathepsines B et L). Nous avons essayé d'élargir le nombre de gènes d'inhibiteurs de protéases à cystéine disponibles pour lutter contre les pucerons. Si les transgènes basés sur des gènes d'origine végétale sont considérés comme présentant "une acceptabilité" plus élevée par le public, nous avons tout de même cherché à isoler un gène de puceron. Cette démarche n'est pas terminée, car bien que purifié, l'inhibiteur mis en évidence n'a pas encore pu être identifié et cloné.

Toutefois, l'élargissement du criblage à des protéines extérieures au monde végétal, dans le cadre du programme européen, a permis de révéler la capacité aphicide de l'équistatine.

Cette protéine a été isolée de l'anémone de mer, *Actinia equinia* (Lenarcic, 1997), et appartient à une nouvelle classe d'inhibiteurs de protéases à cystéine, les thyropines (Lenarcic, 1998). C'est une protéine de 22 kDa qui possède un domaine d'inhibition très efficace des protéases à cystéine<sup>33</sup> et un domaine d'inhibition des protéases acides (cathepsine D), qui sont d'autres protéases lysosomiales de mammifères.

L'équistatine possède des capacités entomotoxiques. Lorsque l'équistatine est badigeonnée sur des feuilles de pommes de terre, 50% des larves de scarabées meurent après 4 jours d'alimentation sur ces plantes (Gruden, 1998). Les tests préliminaires que nous avons réalisés montrent que cet IP inhibe 95% de l'activité endoprotéolytique présente dans le tube digestif du puceron, soit très significativement plus que l'oryzacystatine (50%). Les tests de toxicité en alimentation artificielle montrent également que l'équistatine est létale à faible concentration pour le puceron de melon, la LC50 est d'environ 200  $\mu\text{g}\cdot\text{ml}^{-1}$  (14  $\mu\text{M}$ ). Cette molécule est aussi active que les lectines à mannose qui représentaient jusqu'à ce jour les seules protéines utilisables en transgénèse pour lutter contre les pucerons.

Nous disposons au laboratoire de plantes d'*A. thaliana* exprimant l'équistatine, les tests de gain de résistance seront réalisés très prochainement. La société Teizier, sélectionneur de

---

<sup>33</sup>  $K_i = 0,57 \text{ mM}$  pour la papaïne,  $K_i = 0,051 \text{ nM}$  pour la cathepsine L,  $K_i = 1,4 \text{ nM}$  pour la cathepsine B.

semences de melon, avec laquelle nous collaborons, a réalisé la transformation génétique de melon avec le gène codant pour l'équistatine mais aucun transformant primaire n'a pu être obtenu pour l'instant. D'autres plantes d'intérêt agronomique (pomme de terre, tomate, maïs) ont été transformées avec un vecteur contenant ce gène. Il sera possible de les utiliser afin de confirmer la protection de ces plantes contre d'autres pucerons qui leur sont inféodés.

Compte-tenu des activités de l'équistatine, les pucerons doivent ingérer des milieux contenant  $200 \mu\text{g}.\text{ml}^{-1}$  de toxine pour être significativement intoxiqués, en admettant que nous pouvons extrapoler ces résultats aux conditions *in planta*. Les plantes sont-elles capables d'exprimer des taux élevés de protéines dans leur sève phloémienne ? Les résultats que nous avons obtenus montrent que l'utilisation du promoteur constitutif 35S<sup>2</sup> permet de solubiliser l'oryzacyistine à un taux de 0,2% des protéines phloémiennes dans les lignées de colza étudiées. La sève phloémienne d'un végétal est généralement pauvre en protéines au regard de ses autres composants tels que les acides aminés ou les sucres. Leur concentration mesurée dans des exsudats de types divers, varie de 0,2 à 2 mg. ml<sup>-1</sup> pour les espèces " standards " qui ont été étudiées, qui ne comprennent cependant pas le colza (cités dans Sauvion, 1995). Néanmoins, chez les cucurbitacées ce taux peut être cinquante fois plus élevé (Cronshaw, 1989) bien que toutes ces protéines, dites P-protéines, ne participent pas de la même façon au compartiment protéique soluble et mobile, le seul vraiment accessible par le puceron. Le niveau d'expression que nous avons obtenu avec le promoteur 35S<sup>2</sup> n'est apparemment pas suffisant pour exprimer les concentrations nécessaires à une véritable toxicité. L'utilisation de promoteurs forts spécifiques du phloème, tel que le promoteur WDV (Ripoll, 2001) devrait permettre de nous rapprocher des concentrations toxiques et d'observer une mortalité des pucerons se nourrissant sur les plantes transgéniques afin de répondre aux demandes des professionnels.

Des niveaux d'expression faible de la protéine dans la sève, tels ceux actuellement accessibles, imposent de poursuivre les efforts de recherche pour identifier des inhibiteurs de protéases toxiques à des concentrations plus faibles que ne le sont les IP que nous avons déjà testé.

La mutagenèse dirigée d'IP présente une solution possible afin d'accroître leur toxicité. L'expression hétérologue du mutant point PsCI n'a pas permis l'obtention d'un inhibiteur actif, et les tests en alimentation artificielle n'ont pu être réalisés. Si l'activité de ce mutant se manifeste dans les plants d'*Arabidopsis* transformés, les tests de résistance *in planta* devront nous permettre d'estimer le gain de résistance lié à l'expression de PsCI. Dans le cadre du

programme européen, la mutagenèse d'autres inhibiteurs de protéases a permis de confirmer la faisabilité et le succès de cette stratégie (Ceci, 2002).

Une seconde approche serait d'utiliser une autre classe d'inhibiteurs, n'ayant pas encore été très développée, qui est la pro-région des protéases. La pro-région possède deux rôles, premièrement de chaperon moléculaire qui guide le repliement de l'enzyme dans une conformation active, et deuxièmement, le pro-peptide se replie dans le sillon du site actif inhibant ainsi la protéase lors de son transport vers le lysosome. *In vitro*, l'inhibition des cathepsines par leur pro-région est très efficace, et les constantes d'inhibition observées sont inférieures à 0,1 nM (Coulombe, 1996). Bien que les mécanismes de reconnaissance spécifique de la pro-région avec sa protéase, après séparation, ne soient pas connus, la spécificité entre les deux peptides est très grande (Turk, 2000). Ces peptides pourraient donc représenter de nouveaux outils très spécifiques pour inactiver les cathepsines lysosomiales de divers organismes, et donc potentiellement pour lutter contre les insectes. La disponibilité des ADNc de la cathepsine L et des cathepsines B nous permettrait d'effectuer rapidement les constructions pour l'expression hétérologue des pro-régions afin d'obtenir les peptides solubles et de tester leur activité biologique aphicide en régime artificiel. Si leur toxicité s'avérait positive, leur expression *in planta* pourrait être envisagée sous le contrôle d'un promoteur phloème spécifique, bien que la question de la stabilité de ces inhibiteurs puisse poser problème *a priori*.

### 3. Les inhibiteurs de protéases atteignent une nouvelle cible

Les capacités entomotoxiques des inhibiteurs de protéases que nous avons testées, ont été révélées en test d'alimentation artificielle. Cette alimentation minimise le rôle des protéases à fonction digestive stricte (si elles existent), puisqu'elle ne renferme que des acides aminés libres dans des proportions permettant un développement optimal des pucerons. Dans ces conditions, les inhibiteurs de protéases sont capables de traverser la barrière épithéliale, et ils peuvent donc cibler d'autres protéases impliquées dans divers processus cellulaires. Le puceron est donc un bon modèle permettant de mettre en évidence un mode original de toxicité des IP. Les études d'immunolocalisation que nous avons réalisées tant avec OC-I que PsTI montrent que l'inhibiteur est bien présent dans les cellules du tube digestif, mais également par delà la barrière intestinale. Un signal spécifique est observé dans l'hémolymphe (PsTI), comme au niveau des bactériocytes (OC-I, PsTI). C'est la première

fois que la présence d'un inhibiteur de protéases exogène est révélée dans des tissus autres que digestifs. Chez les autres insectes, l'inhibiteur est mobilisé par les protéases cibles de l'espace ectopéritrophique (Sutherland, 2002), et éventuellement dégradé par un système efficace de protéases non cibles.

Dans le cas d'OC-I, l'identification de la cible non digestive a pu être établie. OC-I inhibe une protéase à cystéine présente dans les bactériocytes (qui est probablement la même que dans le tube digestif, CatL). Le bactériome intégré est un organe clé de l'adaptation des pucerons à leur niche écologique. La bactérie apporte au symbiocosme une grande partie de sa robustesse écologique, et n'est actuellement la cible d'aucune stratégie de lutte. L'utilisation d'inhibiteur de protéases représente aujourd'hui un nouvel outil afin d'atteindre ce métabolisme stratégique.

– **BIBLIOGRAPHIE** –

## Bibliographie

- Abe K., Emori Y., Kondo H., Suzuki K. and Arai S.** (1987). Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). *J. Biol. Chem.*, 262 : 16793-16797.
- Aced PA and Elhammer AP.** (1992) Glycosylation of recombinant prorenin in insect cells: the insect cell line Sf9 does not express the mannose 6-phosphate recognition signal. *Biochemistry*.33 : 8793-8797.
- Alston F. H. and Briggs J. B.** (1970) Inheritance of hypersensitivity to rosy apple aphid *Dysaphis plantaginea* in apple. *Can. J. Genet. Cytol.* 12 : 257-258
- Alston F. H. and Briggs J. B.** (1977) Resistance genes in apple and biotypes of *Dysaphis devectora*. *Ann. Appl. Biol.* 87 : 75-81
- Bania J., Stachowiak D. and Polanowski A.** (1999) Primary structure and properties of cathepsin G/chymotrypsin inhibitor from the larval hemolymph of *Apis mellifera*. *Eur. J. Biochem.* 262 : 680-687
- Bairoch A.** (2000) The ENZYME database in 2000. *Nucleic Acids Research.* 28 (1) : 304-305
- Baldwin I. T. and Preston C. A.** (1999). The eco-physiological complexity of plant responses to insect herbivores. *Planta*, 208 : 137-145.
- Barrett A. J.** (1987) The cystatins : a new class of peptidase inhibitors. *TIBS* 12 : 193-196
- Berthomé R., Teycheney P.Y. and Tepfer M.** (2000) Mécanismes de résistance aux virus dans les plantes transgéniques. *Virologie* 4 : 49-60
- Billings P. C., Brandon D. L. and Habres J. M.** (1991). Internalisation of the Bowman-Birk Protease Inhibitor by Intestinal Epithelial Cells. *Eur. J. Cancer*, 27 : 903-908.
- Birk, Y.** (1976). Proteinase inhibitors from plant sources. *Methods Enzymol*, 45 : 695-7.
- Björk I and Ylinenjärvi** (1992) Different roles of the two disulfide bonds of the cysteine proteinase inhibitor, chicken cystatin, for the conformation of the active protein. *Biochemistry*, 31: 8597-8602
- Blackman R. L. and Eastop V.F.** (1985) Aphids on the world's crop : an identification guide. Chichester, Royaume Uni. 466 p.
- Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A., Brzin J., Kos J. and Turk V.** (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.*, 7, 2593-2599
- Bode W. and Huber R.** (2000) Structural basis of the endoproteinase-protein inhibitor interaction. *Bioch. Biophys. Acta.* 1477 : 241-252
- Boigegrain R. A., Pugnère M., Paroutaud P., Castro B. and Brehelin M.** (2000). Low molecular weight serine protease inhibitors from insects are proteins with highly conserved sequences. *Insect Biochem. Molec. Biol.*, 30 : 145-152.
- Boigegrain R.A., Matras H, Brehelin M, Paroutaud P, Coletti-Previero MA. (1992)** Insect immunity: two proteinase inhibitors from hemolymph of *Locusta migratoria*. *Biochem Biophys Res Commun.* 189 : 790-3.
- Bown D. P., Wilkinson H. S. and Gatehouse J. A.** (1997). Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem. Molec. Biol.*, 27 : 625-638.



## Bibliographie

- Bown D. P., Wilkinson H. S. and Gatehouse J. A.** (1998). Midgut carboxypeptidase from *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae: enzyme characterisation, cDNA cloning and expression. *Insect Biochem. Molec. Biol.*, 28 : 739-749.
- Bradford M.M.** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72 : 248-254
- Broadway R. M., Duffey S. S., Pearce G. and Ryan C. A.** (1986). Plant proteinase inhibitors: A defense against herbivorous insects? *Entomol. exp. appl.*, 41 : 33-38.
- Broadway R. M.** (1995). Are insects resistant to plant proteinase inhibitors? *J. Insect Physiol.*, 41 : 107-116.
- Broadway R. M.** (1996). Dietary proteinase inhibitors alter complement of midgut proteases. *Arch. Insect Biochem. Physiol.*, 32 : 39-53.
- Broadway R. M.** (1997). Dietary regulation of serine proteinases that are resistant to serine proteinase inhibitors. *J. Insect Physiol.*, 43 : 855-874.
- Brown W. M. and Dziegielewska K. M.** (1997). Friends and relations of the cystatin superfamily - new members and their evolution. *Prot. Sci.*, 6 : 5-12.
- Brzin J. and Kidri, M.** (1996). Proteinases and their inhibitors in plants: role in normal growth and in response to various stress conditions. In *Biotechnology and Genetic Engineering Reviews*, ed. M. P. TOMBS, Vol. 13, Intercept Ltd. PO Box 716, Andover, Hants, England SP10, pp. 421-467.
- Buchner P.** (1965). Aphids. In *Endosymbiosis of animals with plant microorganisms*, ed. P. BUCHNER, Interscience. New York (USA), pp. 297-332.
- Burgess E. P. J., Main C. A., Stevens P. S., Christeller J. T., Gatehouse A. M. R. and Laing W. A.** (1994). Effects of protease inhibitor concentration and combinations on the survival, growth and gut enzyme activities of the black field cricket, *Teleogryllus commodus*. *J. Insect Physiol.*, 40 : 803-811.
- Casaretto J. A. and Corcuera L. J.** (1998). Proteinase inhibitor accumulation in aphid-infested barley leaves. *Phytochemistry*, 49 : 2279-2286.
- Ceci, L. R., Spoto, N., de Virgilio, M. & Gallerani, R.** (1995). The gene coding for the mustard trypsin inhibitor-2 is discontinuous and wound-inducible. *FEBS Letters*, 364 : 179-81.
- Ceci L., Volpicella M., Rahbé Y. Gallerani R., Jongsma M. and Beekwilder.** (2002) Selection by phage display of a variant mustard trypsin inhibitor toxic against aphids. *The Plant J. in prep*
- Chaman M. E., Corcuera L. J., Zuniga G. E., Cardemil L. and Argandona V. H.** (2001). Induction of soluble and cell wall peroxidases by aphid infestation in barley. , 49 : 2249-2253.
- Charity J. A., Anderson M. A., bitismch D. J. , Whitecross M. and Higgins T. J. V.** (1999) Transgenic tobacco and peas expressing a protease inhibitor from *Nicotiana glauca* increased insect resistance. *Mol. Breed.* 5 : 357-365
- Charles H. and Ishikawa H.** (1999). Physical and genetical map of the genome of *Buchnera*, the primary endosymbiont of the pea aphid *Acyrtosiphon pisum*. *J. mol. Evol.*, 48 : 142-150.
- Chen J. Q.** (1996 a). Caractérisation biologique et chimique de la résistance monogénique du melon (gène vat, pour virus aphid transmission) au puceron *Aphis gossypii*. INSA Lyon, pp. 1-155.
- Chen J. Q., Delobel B., Rahbé Y. and Sauvion N.** (1996 b). Biological and chemical characterisation of a genetic resistance of melon to the melon aphid. *Entomol. exp. appl.*, 80 : 250-253.

## Bibliographie

- Chen J. Q., Rahbé Y., Delobel B., Sauvion N., Guillaud J. and Febvay G.** (1997). Melon resistance to the aphid *Aphis gossypii*: behavioural analysis and chemical correlations with nitrogenous compounds. *Entomol. exp. appl.*, 85 : 33-44.
- Christeller J. T., Laing W. A., Markwick N. P. and Burges, E. P. J.** (1992). Midgut protease activities in 12 phytophagous lepidopteran larvae - dietary and protease inhibitor interactions. *Insect Biochem. Molec. Biol.*, 22 : 735-746.
- Christeller J. T., Markwick N. P. and Burgess E. P. J.** (1994). Midgut proteinase activities of three keratinolytic larvae, *hofmannophila pseudospretella*, *tineola bisselliella*, and *anthrenocerus australis* and the effect of proteinase inhibitors on proteolysis. *Arch. Insect Biochem. Physiol.*, 25 : 159-173.
- Christeller J. T., Farley P. C., Ramsay R. J., Sullivan P. A. and Laing W. A.** (1998). Purification, characterization and cloning of an aspartic proteinase inhibitor from squash phloem exudate. *Eur. J. Biochem.*, 254 : 160-167.
- Church G. M. and Gilbert W.** (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81 : 1991-1995
- Clough S. J. and Bent A. F.** (1998) Floral dip : a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16 : 735-743
- Comeau A.** (1992) La résistance aux pucerons : aspects théoriques et pratiques. In : la lutte biologique. Morin, G. Eds *Boucherville*, Canada ; 433-449
- Cornuet P.** (1987) Eléments de virologie, Eds *INRA*.
- Couty A.** (2000). Impact environnemental de pommes de terres transgéniques, exprimant une lectine contre les pucerons, sur des parasitoïdes de pucerons. In *IRBI*, Université François-Rabelais de Tours. Tours (France), pp. 91.
- Cuozzo J. W., Tao K., Wu Q., Young w., and Sahagian G. G.** (1998) Lysine-based structure in the proregion of procathepsin L is the recognition site for mannose phosphorylation. *J. Biol. Chem.* 270 : 15611-15619
- Cristofolli P. T., Ribeiro A.F., Deraison C., Rahbé Y., Terra w. R.** Midgut adaptation and enzyme distribution in a phloem sucking insect, the pea aphid *Acyrtosiphon pisum*. soumis à *J. Insect Phys*
- Cronshaw J. and Sabnis D. D.** (1989). Phloem proteins. In *Sieve elements - comparative structure, induction and development*, ed. H. D. BEHNKE, SJOLUND, R.D., Springer. Berlin, pp. 257-283.
- Dannenhoffer J. M., Schulz A., Skaggs M. I., Bostwick D. E. and Thompson G. A.** (1997). Expression of the phloem lectin is developmentally linked to vascular differentiation in cucurbits. *Planta*, 201 : 405-414.
- Dannenhoffer J. M., Suhr R. C. and Thompson G. A.** (2001). Phloem-specific expression of the Pumpkin Fruit Trypsin Inhibitor. , 212 : 155-162.
- De Leo F., Bonadé-Bottino M., Ceci L. R., Gallerani R. and Jouanin L.** (1998). Opposite effects on *Spodoptera littoralis* larvae of high expression level of a trypsin proteinase inhibitor in transgenic plants. *Plant Physiol.*, 118 : 997-1004.
- De Leo F., Ceci L. R., Jouanin L. and Gallerani R.** (2001). Analysis of mustard trypsin inhibitor-2 gene expression in response to developmental or environmental induction. *Planta*, 212 : 710-717.
- De Leo F. and Gallerani R.** (2002) The mustard trypsin inhibitor 2 affects the fertility of *Spodoptera littoralis* larvae fed on transgenic plants. *Insect Biochem. Mol. Biol.* 32: 489-496

## Bibliographie

- Delbridge M. L. and Kelly L. E.** (1990). Sequence analysis, and chromosomal localization of a gene encoding a cystatin-like protein from *Drosophila melanogaster*. *FEBS Lett.*, 274 : 141-145.
- Delledonne M., Allegro G., Belenghi B., Balestrazzi A., Picco F., Levine A., Zelasco S., Calligari P. and Conoulorieri M.** (2001) Transformed of white poplar (*Populus alba* L.) with a novel *Arabidopsis th.* cysteine proteases inhibitor and analysis of insect pest resistance. *Mol. Breed.* 7 :35-42
- Denolf P., Hendrickx K., Vandamme J., Jansens S., Peferoen M., Degheele D. and Vanrie J.** (1997). Cloning and characterization of *Manduca sexta* and *Plutella xylostella* midgut aminopeptidase N enzymes related to *Bacillus thuringiensis* toxin-binding proteins. *Eur. J. Biochem.*, 248 : 748-761.
- Devonshire A. L., Field L. M., Foster S. P., Moores G. D., Williamson M. S. and Blackman R. L.** (1998). The evolution of insecticide resistance in the peach-potato aphid, *Myzus persicae*. *Phil. Trans. R. Soc. Lond. B*, 353 : 1677-1684.
- Dixon A. F. G., Klingauf F. A. and Kawada K.** (1987). Biology, chapter 4. In *World crop pests, Aphids their biology, natural enemies and control, volume A*, Amsterdam: Elsevier, pp. 197-297.
- Domoney C., Welham T. and Sidebottom C.** (1993). Purification and characterization of pisum seed trypsin inhibitors. *J. exp. Bot.*, 44 : 701-709.
- Domoney C., Welham T., Sidebottom C. and Firmin J. L.** (1995). Multiple isoforms of pisum trypsin inhibitors result from modification of two primary gene products. *FEBS Letters*, 360 : 15-20.
- Domoney C., Welham T., Ellis N., Mozzanega P. and Turner L.** (2002) Three classes of proteinases inhibitor gene have distinct but overlapping patterns of expression in *Pisum sativum* plants. *Plant Mol. Biol.* 48 : 319-329
- Douglas A. E.** (1998). Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera*. *Annu. Rev. Entomol.*, 43 : 17-37.
- Doyle J. J. and Doyle J. L.** (1990) DNA extraction from *Arabidopsis*. *Focus* 12 : 13-15
- Duan X. L., Li X. G., Xue q. Z., Aboelsaad M., Xu D. P. and Wu R.** (1996) Transgenic rice plants harboring a introduced potato proteinase inhibitor II gene are insect resistant. *Nature Biotech* 14 : 494-498
- Dutoit F.** (1989) Inheritance of resistance in two *Triticum aestivum* lines to Russian wheat aphid (Homoptera : Aphididae). *J. Econ. Entomol.* 82 : 1251-1253
- Edwards L. H., Smith E. L., Pass H. and Morgan G.** (1985) Registration of post barley. *Crop Sci.* 25 : 363
- Eenink A. H. and Dieleman F. L.** (1982). Resistance of lactuca accessions to leaf aphids: components of resistance and exploitation of wild lactuca species as sources of resistance, pp. 349-355.
- Eguchi M.** (1993). Protein protease inhibitors in insects and comparison with mammalian inhibitors. *Comp. Biochem. Physiol.*, 105B : 449-456.
- Emmerling M., Chandler D. and Sandeman M.** (2001) Molecular cloning of three cDNAs encoding aminopeptidases from the midgut of *Helicoverpa punctigera*, the Australian native budworm. *Insect Biochem. Mol. Biol.* 31 : 899-907
- Engl R. A., Huber R., Bode W. and Schulze A. J.** (1995). Divining the serpin inhibition mechanism: a suicide substrate "springe" ? *TIBTECH*, 13 : 503-510.
- Estelle m. A. and Summerville C.**(1987) Auxin-resistant mutants of *Arabidopsis thaliana* with an altered morphology *Mol. Gen. Genet.* 206 : 200-206

## Bibliographie

- Febvay G., Liadouze I., Guillaud, J. and Bonnot G.** (1995). Analysis of energetic amino acid metabolism in *Acyrtosiphon pisum*: a multidimensional approach to amino acid metabolism in aphids. *Arch. Insect Biochem. Physiol.*, 29 : 45-69.
- Febvay G., Rahbé Y., Rynkiewicz M., Guillaud J. and Bonnot G.** (1999). Fate of dietary sucrose and neosynthesis of amino acids in the pea aphid, *Acyrtosiphon pisum*, reared on different diets. *J. exp. Biol.*, 202 : 2639-2652.
- Ferrasson E., Quillien L. and Gueguen J.** (1995). Amino acid sequence of a Bowman-Birk proteinase inhibitor from pea seeds. *J. Prot. Chem.*, 14 : 467-475.
- Ferrasson E., Quillien L. and Gueguen J.** (1997). Proteinase inhibitors from pea seeds: Purification and characterization. *J. agric. Food Chem.*, 45 : 127-131.
- Fidantsef A. L., Stout M. J., Thaler J. S., Duffey S. S. and Bostock R. M.** (1999). Signal interactions in pathogen and insect attack: expression of lipoxygenase, proteinase inhibitor II, and pathogenesis-related protein P4 in the tomato, *Lycopersicon esculentum*. *Physiol. molec. Plant Pathol.*, 54 : 97-114.
- Field L. M. and Devonshire A. L.** (1997). Structure and organization of amplicons containing the E4 esterase genes responsible for insecticide resistance in the aphid *Myzus persicae* (Sulzer). *Biochem. J.*, 322 : 867-871.
- Flecker P.** (1989) a new and general procedure for refolding mutant Bowman-Birk-type proteinase inhibitors on trypsin-sepharose as a matrix with complementary structure. *FEBS letter* 252 : 153- 157
- Flor H. H.** (1971) Current status of the gene for gene concept. *Annu. Rev. Phytopathol.* 9 : 275-296
- Foissac X., Edwards M., Du J.P., Gatehouse A.M.R. and Gatehouse J.A.** (2002) putative protein digestion in a sap sucking homopteran plant pest (rice brown plant hopper ; *Nivaparvata lugens* : Delphacidae) – identification of trypsin-like and cathepsin B proteases. *Insect Biochem. Mol. Biol, sous press*
- Forrest J.M.S.** (1987) Gallling aphids. In Aphids, their biology, natural enemies and control. Amsterdam : A.K. Minks & P. Harrewjin, Ed. *Elsevier* Vol ; 2A, p. 341-354
- Forslund K., Pettersson J., Bryngelsson T. and Jonsson L.** (2000). Aphid infestation induces PR-proteins differently in barley susceptible or resistant to the birdcherry-oat aphid (*Rhopalosiphum padi*). *Physiol. Plant.*, 110 : 496-502.
- Gatehouse A. M. R. and Boulter D.** (1983). Assessment of antimetabolic effects of trypsin inhibitors from Cowpea (*Vigna unguiculata*) and other legumes on development of the bruchid beetle *Callosobruchus maculatus*. *J. Sci. Food Agric.*, 34 : 345-350.
- Gatehouse A. M. R., Down R. E., Powell K. S., Sauvion N., Rahbé Y., Newell C. A., Merryweather A., Hamilton W. D. O and Gatehouse J. A.** (1996). Transgenic potato plants with enhanced resistance to the peach-potato aphid *Myzus persicae*. *Entomol. exp. appl.*, 79 : 295-307.
- Gatehouse A. M. R., Davidson G. M., Newell C. A., Merryweather A., Hamilton, W. D. O., Burgess E. P. J., Gilbert R. J. C and Gatehouse J. A.** (1997). Transgenic potato plants with enhanced resistance to the tomato moth, *Lacanobia oleracea*: growth room trials. *Mol. Breed.*, 3 : 49-63.
- Gatehouse, A. M. R., Norton, E., Davison, G. M., Babbe, S. M., Newell, C. A. & Gatehouse, J. A.** (1999). Digestive proteolytic activity in larvae of tomato moth, *Lacanobia oleracea*; effects of plant protease inhibitors in vitro and in vivo. *J. Insect Physiol.*, 45 : 545-558.

## Bibliographie

- Gatehouse J.A., Gatehouse A. M. R. and Bown D.** (2000) Control of phytophagous insect pests using serine proteinases inhibitors in recombinant protease inhibitors in plants. Michaud D. Eds. *Academic press*. p 9-26
- Genetics Computer Group** (1982) sequence analysis software package. University research park, Madison, Wisconsin
- Gillespie J. P. and Kanost M. R** (1997) Biological mediators of insect immunity. *Ann. Rev. Entomol.* 42: 611-643
- Girard C.** (1998 a.) Evaluation de l'impact de colza transgéniques exprimant des inhibiteurs de protéases sur des insectes ravageurs et sur l'abeille. Thèse de l'INA PG
- Girard C., Le Metayer M., Bonadé-Bottino M., Phamdelegue M. H. and Jouanin, L.** (1998 b.). High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochem. Molec. Biol.*, 28 : 229-237.
- Girard C., Bonadé-Bottino M., Phamdelegue M. H. and Jouanin L.** (1998 c). Two strains of cabbage seed weevil (Coleoptera: Curculionidae) exhibit differential susceptibility to a transgenic oilseed rape expressing oryzacystatin I. *J. Insect Physiol.*, 44 : 569-577.
- Girard C. and Jouanin L.** (1999). Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. *Insect Biochem. Molec. Biol.*, 29 : 1129-1142.
- Giri A. P., Harsulkar A. M., Deshpande V. V., Sainani M. N., Gupta V. S. and Ranjekar P. K.** (1998). Chickpea defensive proteinase inhibitors can be inactivated by podborer gut proteinases. *Plant Physiol.*, 116 : 393-401.
- Goggin, F. L., Williamson, V. M. and Ullman, D. E.** (2001). Variability in the response of *Macrosiphum euphorbiae* and *Myzus persicae* (Hemiptera : Aphididae) to the tomato resistance gene Mi. *Environ. Entomol.*, 30 : 101-106.
- Gomez D. E., Alonso D. F., Yoshiji H. and Thorgeirsson U. P.** (1997) Tissue inhibitors of metalloproteinases : structure, regulation and biological functions. *Eur. J. Cell. Biol.* 74 : 111-122
- Gorman M.J. and Paskewitz S.M.** (2001) Serine proteases as mediators of mosquito immune responses. *Insect Biochem. Mol. Biol.* 31: 257-262
- Graham J. S., Hall G., Pearce G. and Ryan C. A.** (1986). Regulation of synthesis of proteinase inhibitors I and II mRNAs in leaves of wounded tomato plants. *Planta*, 169 : 399-405.
- Groover A., Dewitt N., Heidel A. and Jones A.** (1997) Programmed cell death of plant tracheary elements differentiating in vitro. *Protoplasma* 196 : 197-211
- Gruden K., Strukelj B., Popovic T., Lenarcic B., Bevec T., Brzin J., Kregar I., Herzog-Velikonja J., Stiekema W. J., Bosch D. and Jongasma M. A.** (1998). The cysteine protease activity of Colorado potato beetle (*Leptinotarsa decemlineata* Say) guts, which is insensitive to potato protease inhibitors, is inhibited by thyroglobulin type-1 domain inhibitors. *Insect Biochem. Molec. Biol.*, 28 : 549-560.
- Guncar G., Pungercic G., Klemencic I., Turk V. and Turk D.** (1999) Crystal structure of MHC class II-associated p41 li fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. *EMBO J.* 18, 793-803
- Haebel S. and Kehr J.** (2001). Matrix-assisted laser desorption/ionization time of flight mass spectrometry peptide mass fingerprints and post source decay: a tool for the identification and analysis of phloem

## Bibliographie

- proteins from *Cucurbita maxima* Duch. separated by two-dimensional polyacrylamide gel electrophoresis. *Planta*, 213 : 586-593.
- Hermesmeier D., Schittko U. and Baldwin I. T.** (2001). Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. I. Large-scale changes in the accumulation of growth- and defense-related plant mRNAs. , 125 : 683-700.
- Hilder V. A., Gatehouse A. M. R., Sheerman S., Barker R. F. and Boulter D.** (1987). A novel mechanism of insect resistance engineered into tobacco. *Nature*, 330 : 160-163.
- Hilder V. A., Powell K. S., Gatehouse A. M. R., Gatehouse J. A., Gatehouse L. N., Shi Y., Hamilton W. D. O., Merryweather A., Newell C. A., Timans J. C., Peumans W. J., Vandamme E. J. M. and Boulter D.** (1994). Expression of snowdrop lectin in transgenic tobacco plants results in added protection against aphids. *Transgenic Res.*, 4 : 18-25.
- Hua G., Tsukamoto K. and Ikezawa H.** (1998) Cloning and sequence analysis of the aminopeptidase N isozyme (APN2) from *Bombyx mori* midgut. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 11 : 213-222
- Ikenaka T. and Norioka S.** (1986) Bowman-Birk family serine proteinase inhibitors. In Proteinase Inhibitors. Amsterdam : Barrett A. J. and Salvesen G.S. Eds Elsevier Science Publishers : p. 23-51
- Imler J.L. and Hoffman J.A.** (2001) Toll receptors in innate immunity. *Trends in Cell Biol.* 11-7 : 304-311
- Inubushi, T., Kakegawa, H., Kishino, Y. and Katunuma, N.** (1994). Specific assay method for the activities of cathepsin L-type cysteine proteinases. *J. Biochem.*, 116 : 282-284.
- Jensen B., Unger K. K., Uebe J., Gey M., Kim Y. M. and Flecker P.** (1996). Proteolytic cleavage of soybean Bowman-Birk inhibitor monitored by means of high-performance capillary electrophoresis. Implications for the mechanism of proteinase inhibitors. *J. biochem. biophys. Methods*, 33 : 171-85.
- Johnston, K. A., Lee, M. J., Brough, G., Hilder, V. A., Gatehouse, A. M. R. and Gatehouse, J. A.** (1995). Protease activities in the larval midgut of *Heliothis virescens*: evidence for trypsin and chymotrypsin-like enzymes. *Insect Biochem. Molec. Biol.*, 25 : 375-383.
- Jouanin L., Bonadé bottino M., Girard C., Lerin J. and Pham Delègue M.** (2000). Expression of proteinase inhibitors in rapeseed in Recombinant protease inhibitors in plants. Michaud D. Eds. *Academic press*. p 9-26
- Jouanin L., Ripoll C. and Deraison C.** (2002) Genetic engineering for insect resistance in plants ; dans Plant Genetic Engineering Vol 4 R.P. Singh Editeur Sci. Tech. Pub., Houston, USA, sous press
- Jongsma M. A., Bakker P. L., Peters J., Bosch D. and Stiekema W. J.** (1995). Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. natl. Acad. Sci. USA*, 92 : 8041-8045.
- Jongsma M. A. and Bolter C. J.** (1997). The adaptation of insects to plant protease inhibitors. *J. Insect Physiol.*, 43 : 885-895.
- Kaloshian, I., Kinsey, M. G., Williamson, V. M. & Ullman, D. E.** (2000). Mi-mediated resistance against the potato aphid *Macrosiphum euphorbiae* (Hemiptera : Aphididae) limits sieve element ingestion. *Environ. Entomol.*, 29 : 690-695.
- Kanost M. R.** (1999). Serine proteinase inhibitors in arthropod immunity. *Develop. comp. Immunol.*, 23 : 291-301.

## Bibliographie

- Kimura M., Ikeda T., Fukumoto D., Yamasaki N., Yonekura M.** (1995) Primary structure of a cysteine proteinase inhibitor from the fruit of avocado (*Persea americana* Mill). *Biosci. Biotechnol. Biochem.* 59 : 2328-2329
- Klingler J., Kovalski I., Silberstein L., Thompson G. A. and Perl Treves R.** (2001). Mapping of cotton-melon aphid resistance in melon. *J. am. Soc. hort. Sci.*, 126 : 56-63.
- Koiwa, H., Bressan, R. A. et Hasegawa, P. M.** (1997). Regulation of protease inhibitors and plant defense. *Trends Plant Sci.*, 2 : 379-384.
- Kondo H., Abe K, Nishimura I., Watanabe H., Emori Y and Arai S** (1990) two distinct cystatin species in rice seed with different specificities against cystein proteinases. *J.Biol.Chem.* 265 : 15832-15837
- Koncz C. and Schell J.** (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204 : 383-396
- Laskowski M. and Kato I.** (1980) Protein inhibitors of proteinases. *Ann. Rev. Biochem.* 49 : 593-626
- Leclant F. and Deguine J. P.** (1994). Aphids (Hemiptera: Aphididae). In *Insect pests of Cotton*, ed. G. Matthews, Tunstall, J., CAB International. Oxon (GBR), pp. 285-323.
- Lecoq H., Cohen S., Pitrat M. and Labonne, G.** (1979). Resistance to Cucumber mosaic virus transmission by aphids in Cucumis melo. *Phytopathology*, 69 : 1223-1225.
- Lee S. I., Lee S. H., Koo, J. C., Chun H. J., Lim C. O., Mun J. H., Song Y. H. and Cho M. J.** (1999). Soybean Kunitz trypsin inhibitor (SKTI) confers resistance to the brown planthopper (*Nilaparvata lugens* Stal) in transgenic rice. *Mol. Breed.*, 5 : 1-9.
- Lehane M. J.** (1997) Peritrophic matrix structure and function, *Annu. Rev. Entomol.*, 42 :525-550
- Lenarcic B., Ritonja A., Strukelj B., Turk B. and Turk V.** (1997). Equistatin, a new inhibitor of cysteine proteinases from *Actinia equina*, is structurally related to thyroglobulin type-1 domain. *J. biol. Chem.*, 272 : 13899-903.
- Lenarcic B. and Bevec T.** (1998) Thyropins-News structurally related proteinase inhibitors. *Biol. Chem.* 379 : 105-111
- Léplé J. C., Bonadé-bottino M., Augustin S., Pilate G., Letan V. D., Delplanq a., Cornu D. and Jouanin L.** (1995) Toxicity of *chrysomela tremulae* (Coleoptera : chrysomelidae) of transgenic poplars expressing a cysteine protease inhibitor. *Mol Breed.* 1 : 319-328
- LeTan-Dumanois V.** (1994) Défense du cotonnier contre les insectes ravageurs : étude d'une stratégie basée sur l'expression conjointe d'inhibiteurs de protéases et de toxines de *Bacillus thuringiensis* dans la plante. Thèse de l'université Paris XI
- Liang C., Brookhart G., Feng G. H., Reeck G. R. and Kramer K. J.** (1991). Inhibition of Digestive Proteinases of Stored Grain Coleoptera by Oryzacystatin, a Cysteine Proteinase Inhibitor from Rice Seed. *FEBS Letters*, 278 : 139-142.
- Mazumdar-Leighton S. and Broadway R.M.** (2001) Transcriptional induction of diverse midgut trypsins in larval *Agrotis ipsilon* and *Helicoverpa zea* feeding on the soybean trypsin inhibitor. *Insect Biochem. Mol. Biol.* 31 :645-657

## Bibliographie

- McBride J. D., Brauer A. B., Nievo M. and Leatherbarrow R. J.** (1998). The role of threonine in the P2 position of Bowman-Birk proteinase inhibitors: studies on P2 variation in cyclic peptides encompassing the reactive site loop. *J. Mol. Biol.*, 282 : 447-58.
- McBride, J. D., Freeman N., Domingo G. J. and Leatherbarrow R. J.** (1996). Selection of chymotrypsin inhibitors from a conformationally-constrained combinatorial peptide library. *J. Mol. Biol.*, 259 : 819-827.
- McManus, M. T., White, D. W. R. & McGregor, P. G.** (1994). Accumulation of a chymotrypsin inhibitor in transgenic tobacco can affect the growth of insect pests. *Transgenic Res.*, 3 : 50-58.
- Menegatti, E., Tedeschi, G., Ronchi, S., Bortolotti, F., Ascenzi, P., Thomas, R. M., Bolognesi, M. and Palmieri, S.** (1992). Purification, inhibitory properties and amino acid sequence of a new serine proteinase inhibitor from white mustard (*Sinapis alba* L.) seed. *FEBS Letters*, 301 : 10-4.
- Miles, P. W.** (1999). Aphid saliva. *Biol. Rev.*, 74 : 41-85.
- Miles, P. W.** (1987). The responses of plants to the feeding of aphidoidea: Principles. In *Aphids, their biology, natural enemies and control*, eds. A. K. Minks & P. Harrewijn, Vol. C, Elsevier. Amsterdam (NDL), pp. 1-21.
- Mochizuki, A.** (1998). Characteristics of digestive proteases in the gut of some insect orders. *Appl. Entomol. Zool.*, 33 : 401-407.
- Moran, P. J. and Thompson, G. A.** (2001). Molecular responses to aphid feeding in *Arabidopsis* in relation to plant defense pathways. *Plant Physiol.*, 125 : 1074-1085.
- Murata, M., Miyashita, S., Yokoo, C., Tamai, M., Hanada, K., Hatayama, K., Towatari, T., Nikawa, T. & Katunuma, N.** (1991). Novel epoxysuccinyl peptides. Selective inhibitors of cathepsin B, in vitro. *FEBS Lett.*, 280 : 307-10.
- Nagata, K., Kudo, N., Abe, K., Arai, S. & Tanokura, M.** (2000). Three-dimensional solution structure of oryzacystatin-I, a cysteine proteinase inhibitor of the rice, *Oryza sativa* L. japonica. *Biochemistry*, 39 : 14753-60.
- Neurath H.** (1984) Evolution of proteolytic enzymes. *Science* 224 : 350-363
- Oppert, B., Morgan, T. D., Culbertson, C. & Kramer, K. J.** (1993). Dietary mixtures of cysteine and serine proteinase inhibitors exhibit synergistic toxicity toward the red flour beetle, *Tribolium castaneum*. *Comp. Biochem. Physiol.*, 105C : 379-385.
- Orr G. L., Strickland J. A. and Walsh T. A.** (1994). Inhibition of Diabrotica larval growth by a multicystatin from potato tubers. *J. Insect Physiol.*, 40 : 893-900.
- Papastoitis G. and Wilson K.A.** (1990) Initiation of the degradation of the soybean Kunitz and Bowman-Birk inhibitors by cysteine protease. *Plant Physiol.* 96 : 11086-1092
- Pearce, G., Ryan, C. A. and Liljegren, D.** (1988). Proteinase inhibitors I and II in fruit of wild tomato species: transient components of a mechanism for defense and seed dispersal. *Planta*, 175 : 527-531.
- Pernas, M., Sanchezmonge, R., Gomez, L. & Salcedo, G.** (1998). A chestnut seed cystatin differentially effective against cysteine proteinases from closely related pests. *Plant molec. Biol.*, 38 : 1235-1242.
- Peterson A. M., Barillas-Mury C. V. and Wells M. A.** (1994) Sequence of three cDNAs encoding an alkaline midgut trypsin from *Manduca sexta*. *Insect. Biochem. Mol. Biol.* 24, 463-71



## Bibliographie

- Pitrat, M. & Lecoq, H.** (1980). Inheritance of resistance to cucumber mosaic virus transmission by *Aphis gossypii* in *Cucumis melo*. *Phytopathology*, 70 : 958-961.
- Pitrat M. and Lecoq H.** (1982). Relations génétiques entre les résistances par non-acceptation et par antibiose du melon à *Aphis gossypii*. Recherche de liaisons avec d'autres gènes. *Agronomie*, 2 : 503-508.
- Polanowski A. and Wilusz T.** (1996) Serine proteinase inhibitors from insect hemolymph. *Acta Biochem. Polon.* 43 : 445-454
- Ponsen, M. B.** (1972). The site of potato leaf virus multiplication in its vector, *Myzus persicae*. An anatomy study. thèse d'état : Université de Wageningen (NDL) 147 p.
- Ponsen, M. B.** (1991). Structure of the digestive system of aphids. *WAU papers*, 91.5 .
- Porter D. R., Burd J. D., Shufran K. A. & Webster J. A.** (2000). Efficacy of Pyramiding Greenbug (Homoptera: Aphididae) Resistance Genes in Wheat. *Environ. Entomol.*, 29 : 1315-1318.
- Purcell J. P., Greenplate J. T. and Sammons R. D.** (1992). Examination of Midgut Luminal Proteinase Activities in 6 Economically Important Insects. *Insect Biochem. Molec. Biol.*, 22 : 41-47.
- Quillien L., Ferrasson E., Molle D. and Gueguen J.** (1997). Trypsin inhibitor polymorphism: multigene family expression and post-translational modification. *J. Prot. Chem.*, 16 : 195-203.
- Rahbé Y., Febvay G., Delobel B. and Bournoville R.** (1988). *Acyrtosiphon pisum* performance in response to the sugar and amino acid composition of artificial diets, and its relation to lucerne varietal resistance. *Entomol. exp. appl.*, 48 : 283-292.
- Rahbé Y. and Febvay G.** (1993). Protein toxicity to aphids - an in vitro test on *Acyrtosiphon pisum*. *Entomol. exp. appl.*, 67 : 149-160.
- Rahbé Y., Delobel B., Febvay G. and Chantegrel B.** (1994). Aphid specific triglycerides in symbiotic and aposymbiotic *Acyrtosiphon pisum*. *Insect Biochem. Molec. Biol.*, 24 : 95-101.
- Rahbé Y., Sauvion N., Febvay G., Peumans W. J. and Gatehouse A. M. R.** (1995). Toxicity of lectins and processing of ingested proteins in the pea aphid *Acyrtosiphon pisum*. *Entomol. exp. appl.*, 76 : 143-155.
- Rahbé Y., Ferrasson E., Rabesona H. and Quillien L.** (2002) Groth inhibition of the pea aphid *Acyrtosiphon pisum* by isoforms and fragments of the Bowman-Birk protease inhibitor from pea seeds. soumis à *J. Insect Physiol.*
- Rawlings N. D. and Barrett A. J.** (1990). Evolution of proteins of the cystatin superfamily. *J. Mol. Evol.*, 30 : 60-71.
- Rawlings N.D., Barrett A.J. and Woessner J.F.** (1998) Handbook of Proteolytic Enzymes. Eds. *Academic press*, 1665 p.
- Reymond P., Weber H., Damond M. and Farmer E. E.** (2000). Differential Gene Expression in Response to Mechanical Wounding and Insect Feeding in *Arabidopsis*. *Plant Cell*, 12 : 707-720.
- Riba G. et Silvy C.** (1989) Combattre les ravageurs des cultures - Enjeux et perspectives. Paris :Ed°INRA
- Richardson M.** (1991). Seed storage proteins : the enzyme inhibitors. *Methods in plant Biochemistry* 5 : 259-305
- Ripoll C.** (2001) Evaluation de stratégie transgéniques pour l'obtention de plantes résistantes aux pucerons. Thèse de l'université Paris VI
- Robert Y.**, ACTA, (1981) Les pucerons des cultures, Paris, Eds. INRA

## Bibliographie

- Rogers B. L., Pollock J., Klapper G. and Griffith I. J.** (1993) Sequence of the proteinase-inhibitor cystatin homologue from the pollen of *Ambrosia artemisiifolia* (short ragweed). *Gene* 133 : 219-221
- Rossi M., Goggin F. L., Milligan S. B., Kaloshian I., Ullman D. E. and Williamson V. M.** (1998). The nematode resistance gene Mi of tomato confers resistance against the potato aphid. *Proc. natl. Acad. Sci. USA*, 95 : 9750-4.
- Ryan C. A.** (1990). Protease inhibitors in plants: genes for improving defenses against insect and pathogens. *Annu. Rev. Phytopathol.*, 28 : 425-449.
- Ryan C. A.** (2000) The systemin signaling pathway : differential activation of plant defensive genes. *Bioch. Biophys. Acta* 1477 : 112-121
- Saarikoski P., Clapham D., von Arnold S.** (1996) A wound-inducible gene from salix viminalis coding for a trypsin inhibitor. *Plant Mol Biol.* 31 : 465-478
- Saito H., Suzuki T., Ueno K., Kubo T. and Natori S.** (1989) Molecular cloning of cdNA for Sarcocystatin A and analysis of the expression of the sarcocystatin a gene during development of *Sarcophaga peregrina*. *Biochem.* 28 : 1749-1755
- Sabnis D. D. and Hart J.** (1976). A comparative analysis of phloem exudate proteins from Cucumis melo, Cucumis sativa and Cucurbita maxima by polyacrylamide gel electrophoresis and isoelectric focusing. *Planta*, 130 : 211-218.
- Salvucci M. E., Rosell R. C. and Brown J. K.** (1998). Uptake and metabolism of leaf proteins by the silverleaf whitefly. *Arch. Insect Biochem. Physiol.*, 39 : 155-165.
- Sanford J.C. and Johnston S.A.** (1985) The concept of parasite-derived resistance-deriving resistance genes from the parasite's own genome. *J. Theor. Biol.* 11 : 395-405
- Satoh D. Horii A., Ochiai M. and Ashida M.** (1999). Prophenoloxidase-activating enzyme of the silkworm, *Bombyx mori*. Purification, characterization, and cDNA cloning. *J. Biol. Chem.*, 274 : 7441-53.
- Sauvion N.** (1995) Effets et modes d'actions de deux lectines à mannose sur le puceron du pois, *Acyrtosiphon pisum* (Harris), Thèse de l'INSA
- Sauvion N., Rahbé Y., Peumans W. J., Vandamme E. J. M., Gatehouse J. A. and Gatehouse A. M. R.** (1996). Effects of GNA and other mannose binding lectins on development and fecundity of the peach-potato aphid *Myzus persicae*. *Entomol. exp. appl.*, 79 : 285-293.
- Schagger H. and Jagow G. V.** (1987). Tricine-Sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, 166 : 368-379.
- Schellenberger V., Braune K., Hofmann H. J. and Jakubke H. D.** (1991) The specificity of chymotrypsin. *Eur. J. Biochem.* 199: 623-636
- Shenk P.M., Kazan K., Wilson I., Anderson J.P., Richmond T., Sommerville S. C. and Manners J.M.** (2000) Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. *Proc. Natl. Acad. Sci.* 97 (21) : 11655-11660
- Shigenobu, S., Watanabe, H., Hattori, M., Sasaki, Y. & Ishikawa, H.** (2000). Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature*, 407 : 81-86.
- Shin S., Park S., Park D., Kim M., Kim S., Brey P. T. and Park H.** (1998). Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR-based differential display and subtractive cloning. *Insect Biochem. Molec. Biol.*, 28 : 827-837.

## Bibliographie

- Smigocki A., Neal J. W., Mccanna I. and Douglass L.** (1993). Cytokinin-Mediated insect resistance in nicotiana plants transformed with the Ipt gene. *Plant molec. Biol.*, 23 : 325-335.
- Solomon, M., Belenghi, B., Delledonne, M., Menachem, E. and Levine, A.** (1999). The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *Plant Cell*, 11 : 431-443.
- Sreerama Y. N. and Gowda L. R.** (1997). Antigenic determinants and reactive sites of a trypsin/chymotrypsin double-headed inhibitor from horse gram (*Dolichos biflorus*). *Biochim Biophys Acta*, 1343 : 235-42.
- Srivastava P. N.** (1987). Nutritional physiology. In *Aphids - Their biology, natural enemies and control*, eds. A. K. Minks & P. Harrewijn, Vol. 2A, Elsevier. Amsterdam, pp. 99-130.
- Stone S. R. and Hofsteenge J.** (1986). Kinetics of the inhibition of thrombin by hirudin. *Biochem.* 25 : 4622-4628
- Stoger E., Williams S., Christou P., Down R. E. and Gatehouse J. A.** (1999). Expression of the insecticidal lectin from snowdrop (*Galanthus nivalis agglutinin*; GNA) in transgenic wheat plants: effects on predation by the grain aphid *Sitobion avenae*. *Mol. Breed.*, 5 : 65-73.
- Stotz H. U., Pittendrigh B. R., Kroymann J., Weniger K., Fritsche J., Bauke A. and Mitchell-Olds T.** (2000). Induced Plant Defense Responses against Chewing Insects. Ethylene Signaling Reduces Resistance of Arabidopsis against Egyptian Cotton Worm But Not Diamondback Moth. *Plant Physiol.*, 124 : 1027-1038.
- Stout M. J., Fidantsef A. L., Duffey S. S. and Bostock R. M.** (1999). Signal interactions in pathogen and insect attack: systemic plant-mediated interactions between pathogens and herbivores of the tomato, *Lycopersicon esculentum*. *Physiol. molec. Plant Pathol.*, 54 : 115-130.
- Stubbs M. T., Laber B., Bode W., Huber R., Jerala R., Lenarcic B., Turk V.** (1990) The refined 2.4 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* 9 : 1939-1947
- Sutherland P. W., Burgess E. P. J., Philip B. A., McManus M. T., Watson L. and Christeller J. T.** (2002) Ultrastructural changes to the midgut of the black field cricket (*Teleogryllus commodus*) following ingestion of potato protease inhibitor II. *J. Insect Phys.* 48 : 27-336
- Suzuki T. and Natori S.** (1985) Purification and characterisation of an inhibitor of the cysteine protease from the hemolymph of *Sarcophaga peregrina* larvae. *J. Biol. Chem.* 8 : 5115-5120
- Tarn T. R., Adams J. B.** (1982) Aphids probing and feeding, electronic monitoring, and plant breeding in Pathogens, vectors and plant diseases, *Harris K. F. and Maramorosch* (Eds.)
- Terra W. R. and Ferreira C.** (1994). Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol.*, 109B : 1-62.
- Terra W. R., Ferreira C. and Bastos F.** (1985). Disaccharidases and the spatial organization of digestion in the *Tenebrio molitor* larvae. *Insect Biochem.*, 15 : 443-449.
- Terra W. R., Ferraira C. and Baker J. E.** (1996) Compartmentalization of digestion. In *Biology of the insect midgut*. Eds. *Chapman & Hall* 485p
- Thie N. M. R. and Houseman J. G.** (1990a). Cysteine and Serine Proteolytic Activities in Larval Midgut of Yellow Mealworm, *Tenebrio-Molitor* L (Coleoptera, Tenebrionidae). *Insect Biochem.*, 20 : 741-744.

## Bibliographie

- Thie N. M. R. and Houseman J. G.** (1990b). Identification of Cathepsin-B, Cathepsin-D and Cathepsin-H in the Larval Midgut of Colorado Potato Beetle, *Leptinotarsa-Decemlineata* Say (Coleoptera, Chrysomelidae). *Insect Biochem.*, 20 : 313-318.
- Thomas J. C., Adams D. G., Keppenne V. D., Wasmann C. C., Brown J. K., Kanost M. R. and Bohnert H. J.** (1995 a). *Manduca sexta* encoded protease inhibitors expressed in *Nicotiana tabacum* provide protection against insects. *Plant Physiol. Biochem.*, 33 : 611-614.
- Thomas J. C., Adams D. G., Keppenne V. D., Wasmann C. C., Brown J. K., Kanost M. R. and Bohnert H. J.** (1995 b). Protease inhibitors of *Manduca sexta* expressed in transgenic cotton. *Plant Cell Rep.*, 14 : 758-762.
- Tjallingii W. F. and Esch T. H.** (1993). Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiol. Entomol.*, 18 : 317-328.
- Turk B., Stoka V., Turk V., Johansson G., Cazzulo J. J. and Bjork I.** (1996) High-molecular-weight kininogen binds two molecules of cysteine proteinases with different rate constants, *Febs lett.* 391, 109-112
- Turk B., Turk D. and Turk V.** (2000) Lysosomal cysteine proteases : more than scavengers. *Bioch. Biophys. Acta* 1477 : 98-111
- Vancanneyt G., Sanz C., Farmki T., Paneue M., Ortego F., Castanera P. and Sanchez-Serrano J.J.** (2001) Hydroxyperoxide lyase depletion in transgenic potato plants leads to an increase in aphid performance. *Proc. Natl. Acad. Sci.* 98 (14) : 8139-8144.
- Vanderwesthuizen A. J., Qian X. M. and Botha A. M.** (1998). beta-1,3-glucanases in wheat and resistance to the Russian wheat aphid. *Physiol. Plant.*, 103 : 125-131.
- Volpicella M., Schipper A., Jongsma M. A., Spoto N., Gallerani R. and Ceci L. R.** (2000). Characterization of recombinant mustard trypsin inhibitor 2 (MTI2) expressed in *Pichia pastoris*. *FEBS Letters*, 468 : 137-41.
- Walling L.** (2000) The myriad plant responses to herbivores. *Journal of Plant Growth Regulation* 19 : 195-216
- Walters F. S. and English L. H.** (1995). Toxicity of *Bacillus thuringiensis* d-entotoxins toward the potato aphid in an artificial diet bioassay. *Entomol. exp. appl.*, 77 : 211-216.
- Wang E., Wang R., DeParasis J., Loughrin J., Dan S., and Wagner G.** (2001) Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance. *Nature Biotech.* 19 : 371-374
- Watanabe H., Abe K., Emori Y., Hosoyama H. and Arai S.**(1991) Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (Oryzains). *J.B.C.* 266 : 16897-16902
- Wedde M., Weise C., Kopacek P., Franke P. and Vilcinskas A.** (1998). Purification and characterization of an inducible metalloprotease inhibitor from the hemolymph of greater wax moth larvae, *Galleria mellonella*. *Eur. J. Biochem.*, 255 : 535-43.
- Welham T., O'Neill M., Johnson S., Wang T. L. and Domoney C.** (1998). Expression patterns of genes encoding seed trypsin inhibitors in *Pisum sativum*. *Plant Sci.*, 131 : 13-24.
- Werner, M. H. and Wemmer, D. E.** (1991). 1H assignments and secondary structure determination of the soybean trypsin/chymotrypsin Bowman-Birk inhibitor. *Biochemistry*, 30 : 3356-3364.

## *Bibliographie*

- Werner M. H. and Wemmer D. E.** (1992) Three-dimensional structure of soybean trypsin/chymotrypsin inhibitor Bowman-Birk inhibitor in solution. *Biochemistry* 31 : 999-1010
- Wright T. H.** (1996) The structural puzzle of how serpin serine proteinase inhibitors work. *BioEssays* (18) 6 : 453-464.
- Xu D. P., Xue Q. Z., Mc Elroy D., Mawal Y., Hilder v. A. and Wu R.** (1996) Constitutive expression of a cowpea trypsin inhibitor gene, CpTI in transgenic rice plants confers resistance to two major rice insect pests. *Current Contents* 2 : 167-173
- Yavelow J., Collins M., Birk Y., Troll W. and Kennedy A.R.** (1985) Nanomolar concentrations of Bowman-Birk soybean protease inhibitor suppress x-ray induced transformation in vitro. *Proc. Natl. Acad. Sci. USA* 82 : 5395-5399
- Yamamoto Y., Watabe S., Kageyama T. and Takahashi S.** (1999 a.) A novel inhibitor protein for *Bombyx* cysteine proteinase is homologous to propeptide regions of cysteine proteinases. *FEBS lett.* 448: 257-260
- Yamamoto Y., Watabe S., Kageyama T. and Takahashi S. Y.** (1999 b.). Purification and Characterization of Bombyx Cysteine Proteinase Specific Inhibitors From the Hemolymph of *Bombyx mori*. *Arch. Insect Biochem. Physiol.*, 42 : 119-129.
- Ye Z. H. and Varner J.E.** (1996) Induction of cysteine and serine proteases during xylogenesis in *Zinnia elegans*.. *Plant Mol. Bol.* 30 : 1233-1246
- Yoo B. C., Aoki K., Xiang Y., Campbell L. R., Hull R. J., Xoconostle-Cazares B., Monzer J., Lee J. Y., Ullman D. E. and Lucas W. J.** (2000). Characterization of *Cucurbita maxima* Phloem Serpin-1 (CmPS-1). A Developmentally regulated elastase inhibitor. *J. Biol. Chem.*, 275 : 35122-35128.
- Zhao Y., Botella M. A., Subramanian L., Niu X. M., Nielsen S. S., Bressan R. A. and Hasegawa P. M.** (1996). Two wound-inducible soybean cysteine proteinase inhibitors have greater insect digestive proteinase inhibitory activities than a constitutive homolog. *Current Contents*, 111 : 1299-1306.
- Ziegler H.** (1975). Nature of transported substances. In *Encyclopedia of plant physiology*, ed. Z. M. H. E. PIRSON A., Vol. 3, Springer Verlag. Berlin (GER), pp. 59-100.
- Zientz E., Silva F. J., Gross R.** (2001) Genome interdependence in insect-bacterium symbioses. *Genome Biology*, 2 (12) : 1032.1-102.6

## **Publication et Communications**

**Publication** : Jouanin L., Ripoll C., Deraison C. Genetic engineering for insect resistance in plants ; dans Plant Genetic Engineering Vol4 R.P. Singh Editeur Sci. Tech. Pub., Houston, USA, sous press

### **Colloques et Posters** :

Présentation orale : The use of protease inhibitors for aphid control. 6<sup>th</sup> International congress of Plant Molecular Biology, Quebec-Canada, juin 2000

### Posters :

Deraison C., Duportets L., Rahbé Y., Jouanin L.. Mise en évidence de protéases et d'inhibiteurs de protéases chez le puceron *Aphis gossypii*. XIII colloque de physiologie de l'insecte. Versailles, mai 2000

Deraison C., Duportets L., Maba M., Quillien L., Rahbé Y., Jouanin L.. Use of protease inhibitors for aphid control 6<sup>th</sup> International congress of Plant Molecular Biology, Quebec, juin 2000

Deraison C., Duportets L., Duport G., Rahbé Y., Jouanin L.. Examination of protease and protease inhibitor activities in the aphid, *Aphis gossypii*. XXI International congress of Entomology, Iguassu-Brasil, août 2000

**— ANNEXES —**

## Composition du milieu standard Ap2

Ordre		Pds.Mol.	Milieu mM	Qté pour 100 ml (mg)
<b>1</b>	<b>Saccharose</b>	<b>342,3</b>	<b>584</b>	<b>20000</b>
<b>2</b>	<b>L-Amino acides</b>			
	Alanine	89,09	20,06	178,71
	β-alanine	89,10	0,70	6,22
	Arginine	174,20	14,06	244,90
	Asparagine H <sub>2</sub> O	150,14	19,88	298,55
	Acide aspartique	133,11	6,63	88,25
	Cystéine	121,16	2,44	29,59
	Acide glutamique	147,13	10,15	149,36
	Glutamine	146,15	30,49	445,61
	Glycine	75,07	22,19	166,56
	Histidine HCl H <sub>2</sub> O	209,63	6,49	136,02
	Isoleucine (allo free)	131,18	12,56	164,75
	Leucine	131,18	17,65	231,56
	Lysine HCl	182,65	19,22	351,09
	Méthionine	149,21	4,85	72,35
	Ornithine HCl	168,62	0,56	9,41
	Phénylalanine	165,19	14,04	231,93
	Proline	115,13	11,23	129,33
	Sérine	105,09	11,83	124,28
	Thréonine (allo free)	119,12	10,67	127,16
	Tryptophane	204,23	2,09	42,75
	Tyrosine	181,19	2,13	38,63
	Valine	117,15	16,29	190,85
	B Ala Tyr	288,00	3,79	109,15
	<b>Total</b>		<b>260</b>	<b>3567</b>
	<b>Rapport Sacch./A.A.</b>		<b>2.25</b>	
<b>3</b>	<b>Divers</b>			
	Citrate de calcium			10,00
	Benzoate de Cholestérol			2,50
	MgSO <sub>4</sub> 7H <sub>2</sub> O			242,00
<b>4</b>	<b>Vitamines</b>			
	Acide p-aminobenzoïque			10,00
	Acide L-Ascorbique			100,00
	Biotine			0,10
	D-Pantothénate de Calcium			5,00
	Chlorure de Choline			50,00
	Acide Folique			1,00
	i-Inositol anhydre			42,00
	Amide nicotinique			10,00
	Pyridoxine HCl			2,50
	Riboflavine			0,50
	Thiamine HCl			2,50
<b>5</b>	<b>Métaux traces</b>			
	CuSO <sub>4</sub> 5H <sub>2</sub> O			0,47
	FeCl <sub>3</sub> 6H <sub>2</sub> O			4,45
	MnCl <sub>2</sub> 4H <sub>2</sub> O			0,65
	NaCl			2,54
	ZnCl <sub>2</sub>			0,83
<b>6</b>	<b>KH<sub>2</sub>PO<sub>4</sub></b>			<b>250</b>

Le pH est ajusté à 7,5 avec KOH (5M). Il faut alors vérifier que la pression osmotique soit d'environ 1200 mOsM ([Febvay, 1988 #2407]). La solution est ensuite filtrée à travers un filtre de 0,45 µm (Disposable Filterwave, Nalgène) et stockée par aliquotes de 5 ou 10 ml à -20°C jusqu'à l'utilisation (cinq à six semaines maximum).



## Composition des différents milieux de croissance de levures

- YPD (Yeast Extract Peptone Dextrose Medium)

<b>Composition</b>	<b>Concentration</b>
Extrait de levure	1 %
Peptone	2 %
Glucose	2 %
Agar	2%

Le glucose n'est additionné qu'après avoir stérilisé le milieu à 120°C pendant 20 min.

- BMGY (Buffered Glycerol-complex Medium)

<b>Composition</b>	<b>Concentration</b>
Extrait de levure	1 %
Peptone	2 %
Phosphate de potassium, pH6	100 mM
Yeast Nitrogen base	1,34%
Biotine	$4.10^{-5}$ %
Glycerol	1 %

Le glucose et la biotine ne sont additionnés qu'après avoir stérilisé le milieu à 120°C pendant 20 min.

- BMM (Buffered Minimal Methanol)

<b>Composition</b>	<b>Concentration</b>
Phosphate de potassium, pH6	100 mM
Yeast Nitrogen base	1,34%
Biotine	$4.10^{-5}$ %
Methanol	0,5 %

Le méthanol n'est additionné qu'après avoir stérilisé le milieu à 120°C pendant 20 min.

## Composition du milieu de croissance in vitro d'arabidopsis

(d'après Estelle et Summerville, 1987)

Composition	Concentration
<b>KNO<sub>3</sub></b>	5 mM
<b>KH<sub>2</sub>PO<sub>4</sub></b>	2,5 mM
<b>MgSO<sub>4</sub> (7H<sub>2</sub>O)</b>	2mM
<b>Ca(NO<sub>3</sub>)<sub>2</sub></b>	2 mM
<b>Microéléments</b>	
H <sub>3</sub> BO <sub>3</sub>	70 µM
MnCl <sub>2</sub> (4H <sub>2</sub> O)	14 µM
CuSO <sub>4</sub> (5H <sub>2</sub> O)	0,5 µM
Na <sub>2</sub> MoO <sub>4</sub> (2H <sub>2</sub> O)	0,2 µM
NaCl	10 µM
ZnSO <sub>4</sub> (7H <sub>2</sub> O)	1 µM
CoCl <sub>2</sub> (6H <sub>2</sub> O)	0,01 µM
<b>Vitamines de Morel et Wetmore</b>	
Myo-Inositol	5000 µg
Panthotenate de calcium	50 µg
Niacine	50 µg
Pyridoxine	50 µg
Thiamine HCl	50 µg
Biotine	0,5 µg
<b>BCP</b>	0,8 □
<b>MES</b>	0,07 %
<b>Saccharose</b>	1% (p/v)
<b>Agar</b>	0,7 % (p/v)

Le pH est ajusté à 5,7, et le milieu est stérilisé pendant 20 min à 120°C sous pression. Le fer citrate ammoniacal 1% est ajouté avant de couler le milieu