THESE

Présentée à l'Université Louis Pasteur, Strasbourg I Faculté des Sciences de la Vie pour obtenir le titre de :

Docteur de l'Université Louis Pasteur de Strasbourg

Domaine : Biologie Cellulaire et Moléculaire

par

Francine LEVY

Analyse de la réponse immunitaire de la drosophile par une approche protéomique

Soutenue le 08 juillet 2005 devant la commission d'examen:

Mme Catherine FLORENTZ (Rapporteur interne, Strasbourg)

M. Eric VIVIER (Rapporteur externe, Marseille)

M. Hélian BOUCHERIE (Rapporteur externe, Bordeaux)

M. Alan ADEREM (Membre invité, Seattle)

M. Philippe BULET (Membre invité, Genève)

M. Jules HOFFMANN (Examinateur, Strasbourg)

Mme Laurence SABATIER (Directeur de thèse, Strasbourg)

A ma maman,

REMERCIEMENTS

Monsieur Hoffmann,

Je vous remercie de m'avoir accueillie dans votre laboratoire et permis de réaliser cette thèse dans un environnement scientifique pluridisciplinaire exceptionnel. Merci également pour l'opportunité que vous m'avez donnée d'aller à Seattle dans le cadre du "projet Tacoma". Merci pour l'intérêt que vous avez porté à mes travaux et pour votre soutien pendant toutes ces années.

Madame Florentz, Monsieur Boucherie, Monsieur Vivier,

Je vous suis reconnaissante d'avoir accepté de lire et de juger cette thèse malgré une charge de travail déjà importante. Veuillez trouver ici l'expression de ma profonde reconnaissance.

Mister Aderem,

Thank you very much for participating to this jury. I also thank you for having welcomed me in your laboratory and allow me to realize the "Tacoma project". Please transmit my thanks to Adrian Ozinsky and Sam Donohoe for their efficient help on the ICAT strategy.

Philippe,

Je te remercie d'être présent dans ce jury. Merci de m'avoir accueillie dans ton équipe pour mon DEA, d'avoir continué de suivre mon travail après ton départ et de m'avoir encouragée.

Laurence,

Merci de m'avoir confié ce projet et de l'avoir supervisé d'une manière exceptionnelle. Merci d'avoir cru en moi, de m'avoir constamment encouragée et d'avoir été toujours disponible. J'ai beaucoup appris durant ces cinq années à tes côtés, de ta rigueur scientifique, ton organisation, ton enthousiasme et ton honnêteté. Tu as rempli ta mission avec excellence et humanité. Je te remercie de m'avoir comprise dans les moments difficiles, de m'avoir écoutée et soutenue...

Merci à Estelle, Marie-Eve, Marie-Céline, Rachel, Sebahat et toutes les autres personnes qui m'ont apporté une aide technique tout au long de ces années. Merci à Christelle et Philippe pour votre efficacité lors des analyses MALDI et pour votre sympathie.

Martine,

Un merci tout particulier pour ton aide précieuse et ton efficacité pour les bactéries, les séances de prélèvements et le tri des mouches. Mais merci également de ta gentillesse, d'avoir gardé le contact, pour les repas et les gâteaux au citron...un vrai réconfort tout au long de cette thèse! Enfin merci de ton affection, spécialement envers Hana.

Marie (L), Léna, Marie (M), Christine, Jean-Marc, Charles, Daniel, Dominique, Jean-Luc, Julien, merci pour votre aide, vos remarques, conseils et collaborations de tous ordres.

Vincent (L),

Merci pour ta gentillesse et pour avoir pris le temps de relire ce manuscrit... tes conseils ont été très précieux.

Phil,

Merci pour les nombreux coups de main que tu m'as apportée surtout pour la correction des textes en anglais. Merci aussi pour les repas, les gâteaux, les sorties, les cinés, le sport...toujours partante!

Laurent,

Je te remercie pour ton assistance en informatique, mais aussi pour m'avoir accompagnée à Seattle, tu as été d'une grande aide durant ce séjour.

Paule,

Merci pour tout ce que tu fais pour le laboratoire, ton professionnalisme et ton aide inégalable dans les méandres administratifs. Merci de ta gentillesse et de ta manière si maternelle de gérer le laboratoire.

Merci à tous les membres du laboratoire, ceux d'aujourd'hui et ceux qui sont partis vers de nouvelles aventures : merci à Danielle, Nathalie, Maurice, Jacopo, Claude (C), Valérie, Annie, Miquette, Reine, Bernard, Nadège, Akira, Sylvia, Cosimo, Cécile (V), Vincent (B), Anne-Isa, Manu, Daniel (D), Hana, Servane, Catherine, Yann, Shin-Hong, Rajeev, Cécile (F), Jean-Michel, Nicolas, Maria, Zacharia, aux sportives Vanessa, Tatiana, Marie (G), et à la relève : Jessica, Richard, Safia, Annabelle et Tunde. Je vous souhaite à tous bon courage pour la suite. Enfin, j'ai une pensée particulière pour Ahmet...

Laure, Nadine et Raghida,

Un merci spécial à vous trois pour votre gentillesse, vos sourires, les repas libanais...mais tout particulièrement pour votre tendresse envers Hana, David et moi. Il y a des rencontres qu'on n'oublie pas.

Ma famille,

Un énorme merci à mon papa, mes sœurs et mon frérot, sans votre aide et vos encouragements, je ne serais sûrement pas arrivée jusque là. Je pense très fort à ma maman qui m'accompagne à chaque instant et qui j'espère est fière de moi...

Mes amours,

David et Hana, merci à vous deux d'éblouir ma vie. Votre soutien est inestimable...

TABLE DES MATIERES

REMERCIEMENTS	
TABLE DES MATIERES	5
INDEX DES FIGURES	
ABREVIATIONS	
INTRODUCTION	
A. La réponse immunitaire des insectes	15
A.1. Les mécanismes cellulaires	15
A.1.1. La phagocytose	15
A.1.2. L'encapsulement	16
A.2. Les mécanismes humoraux	16
A.2.1. La coagulation	17
A.2.2. La mélanisation	17
A.2.3. La synthèse des peptides antimicrobiens	
B. La régulation de l'expression des peptides antimicrobiens de la drosophile	20
B.1. La voie IMD	20
B.2. La voie Toll	23
B.2.1. Signalisation intracellulaire	23
B.2.2. Signalisation extracellulaire	24
B.2.2.1. Activation de la voie Toll par les bactéries à Gram-positif	25
B.2.2.2. Activation de la voie Toll par les champignons	26
C. Etudes post-génomiques de la réponse immunitaire de la drosophile	26
C.1. Approche différentielle transcriptomique	27
C.2. Approche différentielle peptidomique	
C.2.1. Mise en évidence des "Drosophila Immune-induced Molecules"	
C.2.1.1. DIMs présentant des similitudes avec des molécules déjà connues	29
C.2.1.2. DIMs sans similitudes avec des molécules connues	
C.2.2. Mise en évidence des molécules induites par une infection virale	
D. Objectif de ce travail	
MATERIEL ET METHODES	
A. Matériel biologique	
A.1. Souches de drosophiles	37
A.2. Infections expérimentales	37
A.2.1. Infections par piqûre septique	

A.2.2. Infections naturelles	
A.2.3. Infections virales	
A.3. Collecte de l'hémolymphe	
B. Approche par gels d'électrophorèse bidimensionnelle	
B.1. Préparation des échantillons	
B.1.1. Solubilisation des protéines	
B.1.2. Dosage des protéines	
B.2. Gels d'électrophorèse 2D	40
B.2.1. Focalisation isoélectrique	40
B.2.1.1. Réhydratation	41
B.2.1.2. Migration	41
B.2.2. Electrophorèse sur gel de polyacrylamide en conditions dénaturantes	42
B.2.2.1. Réduction/alkylation des molécules	42
B.2.2.2. Migration	42
B.2.3. Coloration des protéines	43
B.2.3.1. Coloration au nitrate d'argent	43
B.2.3.2. Coloration au bleu de Coomassie colloïdal	43
B.3. Analyse qualitative et quantitative informatisée des gels	44
B.3.1. Détection automatique des spots et nettoyage des images	45
B.3.2. Appariement des spots (ou "matching")	45
B.3.3. Normalisation des gels et analyse des données	45
B.4. Identification des protéines par spectrométrie de masse	46
B.4.1. Digestion des protéines et extraction des peptides générés	46
B.4.2. Analyse des échantillons par spectrométrie de masse MALDI-TOF	48
B.4.3. Identification des protéines par recherche dans les banques de données	48
C. Approche par marquage isotopique des protéines	49
C.1. Marquage et digestion des protéines	50
C.2. Prépurification des peptides sur cartouche échangeuse d'ions	51
C.3. Chromatographie d'affinité et coupure de l'étiquette biotine	51
C.4. Fractionnement des échantillons par chromatographie échangeuse d'ions	
C.5. Chromatographie liquide couplée à la spectrométrie de masse en tandem	
C.6. Analyse des résultats	53
D. La RT-PCR quantitative en temps réel	
D. 1. Extraction des ARN totaux.	
D.2. Transcription inverse	54
D.3. PCR quantitative en temps réel	
E. Inactivation de gène par injection d'ARN double brin	55
Di Intervation de gene par injection d'Arriv double of filmanianianianianianianianianianianianiania	
RESULTATS	58

Chapitre 1 : Recherche de nouvelles molécules impliquées dans la réponse immunitaire de la
drosophile : approche par gels d'électrophorèse bidimensionnelle
A. Mise en place de la méthodologie58
A.1. Electrophorèse bidimensionnelle
A.1.1. Optimisation de la solution de solubilisation des protéines et de la gamme de pH58
A.1.2. Choix du sexe des individus60
A.1.3. Cartes de grande dimension
A.2. Mise en place de l'analyse différentielle
A.2.1. Etude cinétique : choix du temps d'infection
A.2.1.1. Infections bactériennes
A.2.1.2. Infections fongiques
A.3. Etude de la reproductibilité des cartes 2D
B. Analyses différentielles des protéines de l'hémolymphe de drosophiles
B.1. Coloration au nitrate d'argent63
B.2. Gammes étroites de pH64
B.3. Coloration au bleu de Coomassie colloïdal
C. Etude complémentaire : Analyse de la transferrine
C.1. Régulation de la transferrine chez des drosophiles mutantes
C.2. Inactivation du gène de la transferrine
C.3. Spécificité des fragments de transferrine
Chapitre 2 : Recherche de nouvelles molécules impliquées dans la réponse immunitaire de la
drosophile : approche par marquage isotopique différentiel ICAT
A. Description de l'approche ICAT71
A.1. Principe méthodologique71
A.2. Principe des logiciels d'analyse des données71
B. Etude différentielle des protéines de l'hémolymphe après infection expérimentale
B.1. Mise en place de la méthodologie
B.1.1. Choix du seuil de sensibilité
B.1.2. Calcul du taux d'abondance relatif des protéines74
B.1.3. Analyse des protéines de l'hémolymphe de drosophiles75
B.2. Infections spécifiques
B.2.1. Choix des agents microbiens et des temps d'infection75
B.2.2. Identification des protéines76
DISCUSSION
A. Comparaison des deux approches
A.1. Caractéristiques des protéines détectées

A.2. Nature des variations
B. Molécules dont le rôle dans la réponse immunitaire est connu ou fortement supposé
B.1. Peptides antimicrobiens et DIMs
B.2. Protéases
B.2.1. Protéases à sérine à domaine clip
B.2.2. Molécules potentiellement impliquées dans l'activation de la prophénoloxydase
B.2.3. Métallopeptidases à zinc
B.3. Inhibiteurs de protéases
B.4. Molécules de reconnaissance du pathogène91
B.4.1. Gram-negative binding proteins91
B.4.2. Peptidoglycan Recognition Proteins92
B.4.3. Facteurs de croissance des disques imaginaux93
B.5. Molécules similaires aux protéines du complément94
C. Nouveaux candidats de la réponse immunitaire95
C.1. Molécules sensorielles
C.2. Peptidylglycine α-Hydroxylating Monooxygenase
C.3. Molécules de liaison à la phosphatidyléthanolamine97
C.4. Molécules de la famille des immunophilines97
C.5. Molécules de liaison au calcium98
C.5.1. Sarcoplasmic Calcium-binding Protein et SCP-related protein99
C.5.2. Alpha-amylase
C.6. Molécules antioxydantes et métabolisme des lipides101
C.7. Molécules impliquées dans le métabolisme du fer
CONCLUSION ET PERSPECTIVES 107
REFERENCES BIBLIOGRAPHIQUES
Annexe 1 - Liste des amorces utilisées 132
Annexe 2 : Gel synthétique de l'ensemble des spots identifiés par approche 2D . 133
Annexe 3 : Liste des protéines détectées par ICAT et régulées après infection par
un mélange bactérien 134
Annexe 4 : Liste des protéines détectées par ICAT et régulées après infection par
différents pathogènes
Annexe 5 : Publications 136

INDEX DES FIGURES

- Figure 1 : Représentation schématique de la cascade de coagulation chez la limule. (Page 17)
- Figure 2 : Représentation schématique de la cascade de mélanisation des Arthropodes. (Page 18)
- Figure 3: Les sept familles de peptides antimicrobiens de la drosophile. (Page 19)
- Figure 4 : Représentation schématique de la voie IMD de la drosophile. (Page 21)
- Figure 5 : Représentation schématique de la voie Toll de la drosophile. (Page 24)
- Figure 6 : Diagramme de Venn représentant le nombre de gènes induits au cours de la réponse immunitaire de la drosophile. (Page 27)
- Figure 7 : Spectre de masse MALDI-TOF de l'hémolymphe d'une drosophile contrôle ou infectée expérimentalement par un mélange bactérien. (Page 28)
- Figure 8 : Stratégie d'identification des protéines contenues dans les spots des gels 2D. (Page 49)
- Figure 9 : Structure schématique du réactif clivable ICAT. (Page 50)
- Figure 10: Comparaison des gels 2D de gammes de pH 3-10 et 5-8. (Page 59)
- Figure 11 : Cartes 2D de la fraction soluble des protéines de l'hémolymphe de drosophile. (Page 60)
- Figure 12 : Cartes 2D de l'hémolymphe de drosophile mâles et femelles. (Page 60)
- Figure 13 : Carte 2D de petite dimension de l'hémolymphe de drosophiles contrôles. (Page 61)
- Figure 14 : Carte 2D de grande dimension de l'hémolymphe de drosophiles contrôles. (Page 61)
- Figure 15 : Cartes 2D colorées au nitrate d'argent des protéines de l'hémolymphe de drosophiles contrôles ou infectées par *B. bassiana*. (Page 63)
- Figure 16 : Cartes 2D de gamme étroite de pH. (Page 64)
- Figure 17 : Cartes 2D colorées au bleu de Coomassie colloïdal des protéines de l'hémolymphe de drosophiles contrôles ou infectées. (Page 65)

Figure 18 :	Diagramme	de	Venn	représentant	la	répartition	du	nombre	de	spots
	régulés en fo	oncti	on du	type d'infectio	on.	(Page 66)				

- Figure 19: La transferrine et ses fragments de clivage. (Page 67)
- Figure 20 : Régulation de la transferrine chez des drosophiles mutantes. (Page 68)
- Figure 21: Inactivation de la transferrine par injection d'ARN double brin. (Page 69)
- Figure 22 : Représentation schématique de la stratégie ICAT. (Page 71)
- Figure 23 : Choix de la probabilité minimale d'identification des protéines par ProteinProphet. (Page 73)
- Figure 24 : Nombre de peptides ayant servi à l'identification des protéines. (Page 77)
- Figure 25 : Nombre de protéines identifiées en commun dans une, deux, trois, quatre ou cinq infection(s) en fonction du nombre moyen de peptides utilisés pour l'identification. (Page 77)
- Figure 26 : Diagramme d'Euler-Venn représentant la distribution des molécules identifiées par la stratégie ICAT en fonction des infections. (Page 78)
- Figure 27 : Structure schématique et distribution phylogénétique des GNBPs. (Page 91)
- Figure 28 : Structure schématique de la région thioester des TEPs. (Page 94)
- Figure 29 : Structure schématique du peptidoglycane. (Page 100)

INDEX DES TABLEAUX

- Tableau I :Récapitulation des caractéristiques structurales obtenues sur les DIMs.
(Page 29)
- **Tableau II :** Récapitulatif du nombre de spots régulés d'au moins un facteur cinq aprèsinfection. (Page 63)
- **Tableau III :** Identité des protéines dont l'abondance varie après infection d'aprèsl'analyse des gels 2D. (Page 66)
- Tableau IV : Identification des protéines quantifiées pour toutes les infections issues de l'analyse ICAT. (Page 79)

ABREVIATIONS

2D :	bidimensionnel (le)
2D-DIGE :	2D-difference gel electrophoresis
ACE :	Angiotensin Converting Enzyme (Mammifères)
ACER :	Angiotensin Converting Enzyme Related
ACN :	Acétonitrile
ADN :	Acide désoxyribonucléique
ANCE :	Angiotensin Converting Enzyme (drosophiles)
ARN :	Acide ribonucléique
ASAPRatio :	Automated Statistical Analysis of Protein abundance Ratio
ASB-14 :	Amidosulfobétaine-14
βGRP :	βGlucan Recognition Proteins
CBB :	Coomassie Brilliant Blue
CG :	Celera Gene
CHAPS :	3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulfonate
CLHP:	Chromatographie Liquide Haute Performance
CyP:	Cyclophiline
DAP :	Diaminopimélique (Acide)
DCV :	Drosophila C Virus
DD :	Death Domain
DIF :	Dorsal-related Immune Factor
DIM :	Drosophila Immune-induced Molecule
DREDD :	Death-Related ced-3/Nedd2-like protein
DTT :	Dithiothréitol
EDTA :	Ethylènediaminetétraacétique
EM:	Expectation-Maximization
ESI :	Electrospray ionization
FADD :	Fas Associated Death Domain protein
FKBP :	FK-506 Binding Proteins
GlcNAc :	N-acétylglucosamine
GNBP :	Gram-Negative Binding Proteins
GRAVY :	Grand average of hydropathy
ICAT :	Isotope Coded Affinity Tag
IDGF :	Imaginal Disc Growth Factor
IEF :	Focalisation isoélectrique
IMD :	Immune deficiency
IPG :	Immobilized pH gradient
IRC :	Immune-Regulated Catalase
IRD5 :	Immune Response Deficient 5
ISB :	Institute for Systems Biology
I-κB :	Inhibitor of NF-κB
JNK :	Jun Kinase
LC-MS/MS :	Chromatographie liquide couplée à la spectrométrie de masse en tandem
LPS :	Lipopolysaccharide
MALDI-TOF	: Matrix Assisted Laser Desorption Ionisation-Time Of Flight

MPAC :	Maturated Prodomain of Attacin C
MS/MS :	Spectrométrie de masse en tandem
MurNAc :	Acide N-acétylmuramique
NEC :	Necrotic
NEP :	Néprilysine
$NF-\kappa B$:	Nuclear Factor- κB
OBP :	Odorant Binding Protein
OTAP :	Ovotransferrin Antimicrobial Peptide
PBS :	Phosphate Buffered Saline
PCR :	Polymerase Chain Reaction
PDA :	Pipérazine diacrylamide
PEBP :	Phosphatidyléthanolamine Binding Protein
PGN :	Peptidoglycane
PGRP :	Peptidoglycan Recognition Protein
PHK :	Phérokine
PHM :	Peptidylglycine α-Hydroxylating Monooxygenase
pI :	Point Isoélectrique
PO:	Phénoloxydase
PPO :	Prophénoloxydase
PPAE :	Prophenoloxydase Activating Enzyme
PPIase :	Peptidyl-Prolyl cis/trans Isomérase
PSH :	Perséphone
RC/DC :	Reducing agent Compatible / Detergent Compatible
RCL :	Reactive Center Loop
RFABP :	Retinoid and Fatty Acid Binding Protein
RIP :	TNF-Receptor-Interacting Protein
RKIP :	Raf-1 Kinase Inhibitor Protein
RP49 :	Ribosomal Protein 49
RT-PCR :	Reverse Transcriptase-Polymerase Chain Reaction
SCP :	Sarcoplasmic Calcium-binding Protein
SCPr :	SCP-related protein
SDS :	Dodécylsulfate de sodium
SDS-PAGE :	SDS-Polyacrylamide Gel Electrophoresis
SPZ:	Spaetzle
TAB2 :	TAK1-associated kinase
TAK-1 :	Transforming growth factor-β Activated Kinase 1
TBP :	Tributylphosphine
TCEP :	Tris (2-carboxyethyle) phosphine
TEMED :	N,N,N',N'-tétraméthyléthylènediamine
TEP:	Thiolester Containing Proteins
TIR :	Toll/Interleukine-1 Receptor
TNF :	Tumor Necrosis Factor
TSE ·	Transferrine

TSF: Transferrine

INTRODUCTION

INTRODUCTION

Les organismes pluricellulaires ont développé des systèmes de défense très efficaces contre les infections microbiennes auxquelles ils sont confrontés. La réponse immunitaire innée, commune à tous les animaux, représente la première ligne de défense contre ces agressions. Elle repose sur le déclenchement rapide de mécanismes efficaces tels que la phagocytose ou la synthèse de molécules à activité antimicrobienne. Chez les Vertébrés, elle stimule et oriente la réponse adaptative. Cette deuxième ligne de défense implique l'amplification de lignées lymphocytaires exprimant des récepteurs réarrangés et résulte en la production d'un large répertoire de protéines, comme les anticorps, reconnaissant spécifiquement des antigènes microbiens. Contrairement à la réponse innée, les mécanismes adaptatifs mettent plusieurs jours à se mettre en place même si les réarrangements génétiques, conservés dans des cellules mémoires, permettent de répondre plus rapidement à une réexposition à un agent pathogène. Depuis plusieurs décennies, la plupart des études sur l'immunité est tournée vers l'immunité adaptative. Ce n'est qu'avec le début du XXI^{ème} siècle que la contribution de l'immunité innée dans le système de défense général a vraiment été appréciée. Ceci a été initié par les avancées récentes faites sur la caractérisation moléculaire de la réponse immunitaire innée de nombreux organismes pluricellulaires, des insectes jusqu'aux Mammifères, révélant des similitudes frappantes appuyant leur évolution commune. Une étape cruciale a été le séquençage du génome de la drosophile (Adams et al., 2000), Drosophila melanogaster, dépourvue de réponse adaptative et dont la survie repose uniquement sur l'efficacité de sa réponse immunitaire innée. La drosophile s'est révélée être un excellent modèle d'étude des mécanismes de l'immunité innée. La découverte du récepteur transmembranaire Toll chez cet insecte, jouant un rôle crucial dans la réponse antifongique notamment, a motivé la recherche d'homologues chez l'homme. La découverte des "Toll-like receptors" a ainsi ouvert de nouvelles perspectives dans le domaine de la recherche sur l'immunité des Mammifères (Hoffmann et al., 1999).

Au cours de cette introduction, les différents aspects de la réponse immunitaire des insectes seront discutés en détaillant plus précisément les données sur la drosophile.

A. La réponse immunitaire des insectes

Suite à une infection, les insectes mettent en place différents types de défense qui reposent sur des mécanismes cellulaires et humoraux pouvant être détaillés en trois composantes majeures : (i) des réactions cellulaires médiées par les hémocytes (cellules sanguines) conduisant à la phagocytose ou l'encapsulement des particules étrangères ; (ii) l'activation d'une série de cascades protéolytiques conduisant à la coagulation, la mélanisation et la libération localisée de radicaux libres et (iii) l'activation de la synthèse par le corps gras (équivalent fonctionnel du foie des Mammifères) d'une batterie de peptides antimicrobiens sécrétés dans l'hémolymphe (équivalent du sang des Mammifères) (Hoffmann, 2003). A ce jour, seul ce troisième mécanisme a été démontré comme étant indispensable à la résistance aux infections par des microorganismes.

A.1. Les mécanismes cellulaires

La drosophile possède un système circulatoire ouvert qui contient l'hémolymphe pompée par le vaisseau dorsal. La production des hémocytes est assurée par les glandes de la lymphe, organe hématopoïétique qui se développe au stade larvaire et disparaît au stade pupal. Les cellules sanguines de drosophiles adultes dérivent donc des hémocytes larvaires. Ces derniers se répartissent en trois catégories :

- les plasmatocytes qui sont des cellules phagocytaires,

- les lamellocytes qui sont impliqués dans l'encapsulement des particules étrangères trop volumineuses pour être phagocytées,

- les cellules à cristaux qui contiennent les enzymes nécessaires à la mélanisation.

Seuls les plasmatocytes sont retrouvés au stade adulte (Lanot et al., 2001; Meister, 2004).

A.1.1. La phagocytose

Les plasmatocytes présentent des similarités fonctionnelles avec les monocytes et les macrophages des Mammifères et ont la capacité de phagocyter les cellules apoptotiques et les bactéries. Il a été montré qu'un récepteur de type "scavenger", dSR- CI (Drosophila Scavenger Receptor Class 1), est impliqué dans la phagocytose des bactéries à Gram-positif et Gram-négatif (Ramet et al., 2001). Cependant, les mécanismes précis de la phagocytose restent encore mal connus. Ils pourraient impliquer des protéines à site thioester (TEPs pour ThiolEster Containing Proteins), molécules présentant des similarités avec la superfamille des α_{2} macroglobulines/facteur C3 du complément (Lagueux et al., 2000). En effet, six gènes codant des TEPs ont été caractérisés chez la drosophile. Ils pourraient participer à la réponse immunitaire soit en tant qu'inhibiteur de protéase, à la manière des α_2 macroglobulines, soit en tant qu'opsonine en favorisant la phagocytose, à la manière du facteur C3 du complément. Cette dernière hypothèse est soutenue par l'identification d'un homologue de cette famille chez le moustique Anopheles gambiae et par son implication dans l'opsonisation et la phagocytose de bactéries à Gram-négatif par les macrophages (Blandin et al., 2004; Levashina et al., 2001).

A.1.2. L'encapsulement

L'encapsulement est observé lorsque le pathogène est de taille trop importante pour être phagocyté par les plasmatocytes. C'est le cas lorsque la guêpe *Leptopilina boulardi* pond ses œufs dans la larve de drosophile (Carton and Nappi, 2001). Les lamellocytes, qui ne sont pas observés dans les larves saines, se différencient dans ce cas et viennent entourer les parasites, constituant ainsi la capsule. Cette capsule est alors mélanisée. Le parasite meurt ensuite, probablement par asphyxie, mais également sous l'effet cytotoxique de certains produits de la réaction de mélanisation, comme des résidus oxygénés (Nappi et al., 1995; Nappi et al., 2000).

A.2. Les mécanismes humoraux

Au niveau du site d'infection, la première réaction de défense humorale est l'activation de cascades protéolytiques qui conduisent à la coagulation et la mélanisation limitant ainsi la perte d'hémolymphe et la propagation de l'infection. Parallèlement à ces phénomènes, un second processus inductible se met en place aboutissant à la synthèse et à la sécrétion dans l'hémolymphe de peptides à activité antimicrobienne.

A.2.1. La coagulation

La coagulation a été plus largement étudiée chez les Arthropodes aquatiques, notamment chez plusieurs espèces de limule, où elle est activée par les lipopolysaccharides (LPS) ou les β -1,3-glucanes des agents microbiens (Iwanaga and Kawabata, 1998). Une cascade protéolytique se met alors en place dans l'hémolymphe convertissant une protéine soluble (coagulogène), normalement contenue dans les hémocytes mais rapidement libérée dans l'hémolymphe après l'infection, en une protéine insoluble (coaguline) (Figure 1). La cascade est basée sur quatre zymogènes de protéase à sérine impliquées dans deux voies d'activation (une par le LPS, l'autre par le β -1,3glucane) qui convergent vers une protéase à sérine, la proenzyme de coagulation. L'activation de la cascade permet la coupure de la proenzyme en enzyme de coagulation, permettant ainsi la conversion du coagulogène en coaguline.

Chez la drosophile, aucune cascade de protéases impliquée dans la coagulation n'a été mise en évidence. La transglutaminase, orthologue du facteur de coagulation XIIIa de Vertébrés, est le seul facteur de coagulation détecté dans le génome de cet insecte. Une récente étude protéomique sur l'hémolymphe de larves de drosophile avant et après coagulation a permis de mettre en évidence des candidats pouvant jouer un rôle de substrat de la transglutaminase ou encore un rôle de facteur de coagulation. On peut citer, entre autres, une protéase, deux phénoloxydases, une lipophorine et une gelsoline (Karlsson et al., 2004).

A.2.2. La mélanisation

En plus de la coagulation, la rupture de la cuticule par une blessure déclenche la mélanisation qui se caractérise par une tache noire localisée à l'endroit de la blessure. Ce phénomène résulte également d'une cascade protéolytique qui aboutit à la coupure de la prophénoloxydase (PPO), circulante dans l'hémolymphe sous forme de zymogène, en phénoloxydase (PO). L'enzyme responsable de ce clivage est une sérine protéase appelée PPAE pour "prophenoloxydase activating enzyme" (Ashida and Brey, 1995; Chosa et al., 1997; Jiang et al., 1998). La PO ainsi activée va catalyser la réaction



Figure 1 : Représentation schématique de la cascade de coagulation chez la limule.

Après contact avec du LPS et du β -1,3-glucane, une cascade protéolytique est activée très rapidement pour convertir le coagulogène, soluble, en coaguline, insoluble. La perte d'hémolymphe est ainsi stoppée et les microorganismes entrant dans l'hémocoele sont stoppés. *D'après Iwanaga et Kawabata 1998*.

d'oxydation des phénols en quinones qui, en se polymérisant, vont former la mélanine (Soderhall and Cerenius, 1998) (Figure 2). La PPAE de drosophile n'est pas encore caractérisée.

Chez les arthropodes, selon des études biochimiques, la présence de nombreuses molécules inhibitrices de sérines protéases (serpines) régule la cascade protéolytique et évite une hyperactivation du système (Jiang and Kanost, 2000; Park et al., 2000). Récemment, le rôle d'une de ces molécules, la serpine 27A (Spn27A), a été mis en évidence dans la régulation de la cascade de mélanisation chez la drosophile, par le laboratoire de Bruno Lemaître (De Gregorio et al., 2002a) et le notre (Ligoxygakis et al., 2002b). L'inactivation génétique de cette serpine conduit à l'activation constitutive de la PO et entraîne une mélanisation spontanée. D'autre part, la Spn27A agit sur la cascade de mélanisation, probablement en inhibant directement la PPAE de drosophile (De Gregorio et al., 2002a). La dégradation de la Spn27A est elle-même régulée par la voie Toll, une voie de signalisation qui dirige aussi l'expression de peptides antimicrobiens et dont nous parlerons par la suite.

A.2.3. La synthèse des peptides antimicrobiens

Nous savons, depuis les travaux de Metalnikow et de Paillot au début du XXème siècle, qu'une piqûre septique induit chez les insectes supérieurs (holométaboles) l'apparition transitoire d'une forte activité antimicrobienne dans l'hémolymphe. Les travaux des dernières années ont montré que cette activité traduit la sécrétion par le corps gras (équivalent fonctionnel du foie) d'une batterie de peptides antimicrobiens, souvent très puissants et pouvant agir sur une large gamme de microorganismes (Bulet et al., 1999). Leur concentration varie et peut atteindre 100 µM dans l'hémolymphe. Leur cinétique d'induction et leur spectre d'action sont variables. Ces peptides sont pour la plupart amphipathiques et cationiques, ce qui leur permettrait de s'accumuler sur la paroi bactérienne, chargée négativement, et de déstabiliser la membrane plasmique en s'y insérant, entraînant la lyse bactérienne (Shai, 1999). Des études réalisées au laboratoire ont mis en évidence chez la drosophile l'induction de gènes codant sept familles de peptides à activité antibactérienne et/ou antifongique (Figure 3) :



Figure 2 : Représentation schématique de la cascade de mélanisation des Arthropodes. La reconnaissance de motifs microbiens par des molécules de reconnaissance permet de déclencher une cascade protéolytique qui aboutit à l'activation de la phénoloxydase, enzyme catalysant la formation de mélanine, un polymère de quinones. *D'après Soderhall et Cerenius 1998*.

- Deux types de peptides, la Défensine et la Drosomycine, ont une structure complexe comportant une hélice amphipathique reliée par des ponts disulfure à un feuillet β formé respectivement de deux ou trois brins.

- Les cinq autres types de peptides ne comportent pas de cystéines ; ils sont soit organisés en deux hélices α , comme les Cécropines, soit non structurés et riches en résidus proline, comme la Metchnikowine et la Drosocine ou riches en résidus glycine, comme la Diptéricine et l'Attacine.

L'analyse du génome de la drosophile a révélé qu'il existe sept gènes de Drosomycine, quatre de Cécropine, quatre d'Attacine, deux de Diptéricine, un de Défensine, un de Metchnikowine et un de Drosocine. L'absence de détection de tous les variants peptidiques, au cours des différentes analyses biochimiques réalisées, suggère que tous ces gènes ne sont pas induits au cours de la réponse immunitaire de l'insecte ou que les quantités produites dans l'hémolymphe sont trop faibles pour être détectées.

Les peptides antimicrobiens sont produits par le corps gras lors d'une infection systémique, mais peuvent aussi être exprimés par les épithélia dans le cadre d'une réponse immunitaire locale. Chez la drosophile, il a été montré en utilisant des lignées transgéniques portant la "Green Fluorescent Protein" sous le contrôle des promoteurs des différents peptides antimicrobiens, que ceux-ci peuvent être activés suite à une infection dans certains tissus : le tractus respiratoire, la région orale et le tractus digestif, les tubules de Malpighi et les tractus reproductifs mâle et femelle (Ferrandon et al., 1998; Tzou et al., 2000). Ces tissus sont le lieu d'importants processus physiologiques impliquant une interaction avec le milieu extérieur et dans lesquels les cellules sont susceptibles d'être en contact avec des microorganismes. Cette première ligne de défense pourrait permettre de combattre localement les infections et ainsi d'éviter de déclencher une réponse systémique coûteuse en énergie.

Comme nous le verrons plus loin, différentes données expérimentales nous amènent à penser que les peptides antimicrobiens ne sont cependant qu'une des facettes de la résistance aux infections.



Figure 3 : Les sept familles de peptides antimicrobiens de la drosophile. Les activités et concentrations dans l'hémolymphe après infection bactérienne sont indiquées.

B. La régulation de l'expression des peptides antimicrobiens de la drosophile

L'étude de l'induction des gènes codant les peptides antimicrobiens et notamment l'analyse des régions promotrices des gènes codant la Cécropine (Engstrom et al., 1993) et la Diptéricine (Kappler et al., 1993; Meister et al., 1994) a permis de montrer que leur régulation est due, entre autres, à la présence de séquences homologues aux sites de fixation du facteur de transcription Nuclear Factor- κ B (NF- κ B) des Mammifères. Les facteurs de transcription de la famille NF-KB jouent, chez les Mammifères, un rôle central dans la réponse inflammatoire et dans la réponse à divers types de stress. Ce facteur de transcription est un dimère de protéines à domaine Rel. En l'absence de signal, le facteur NF-κB est retenu dans le cytoplasme par l'association à un inhibiteur I- κ B (Inhibitor of NF- κ B). Chez les Mammifères, la fixation d'un ligand à un récepteur transmembranaire conduit à l'activation d'une cascade de transduction du signal par des interactions protéine-protéine assurées par les "Death Domains" (DD) de protéines adaptatrices. Celles-ci activent ensuite une cascade de phosphorylation aboutissant à la phosphorylation par le complexe IKK (I-κB kinase) de l'inhibiteur I-κB (Inhibitor of NF- κ B) qui est alors dégradé par le protéasome. En conséquence, le facteur NF-kB est libéré de son ancre cytoplasmique et migre dans le noyau (Silverman and Maniatis, 2001). Des études génétiques effectuées dans notre laboratoire ont mis en évidence qu'au moins deux voies de régulation, indépendantes et adaptées à la nature des microorganismes infectieux, existent chez la drosophile (Lemaitre et al., 1995) : les voies Toll et IMD. Ces voies sont toutes deux de type NF- κ B.

B.1. La voie IMD

Les peptides antimicrobiens de drosophiles contrôlés par cette voie (Diptéricine, Drosocine, Cécropine et Attacine) ont un spectre d'action principalement dirigé contre les bactéries à Gram-négatif. Des mutants perte de fonction de la voie IMD présentent un défaut de résistance aux infections par ces bactéries tandis qu'elles se comportent normalement vis-à-vis des infections par des champignons ou des bactéries à Grampositif. De plus, la synthèse des peptides antimicrobiens contrôlés par cette voie est fortement diminuée chez ces mutants.

Le facteur de transcription cible de la voie IMD est Relish, une protéine qui contient un domaine Rel N-terminal et des domaines ankyrines, caractéristiques des inhibiteurs I- κ B, dans la partie C-terminale (Dushay et al., 1996). L'activation de la voie de signalisation entraîne la coupure endoprotéolytique de Relish, séparant les domaines Rel et ankyrines. Ceci permet la translocation nucléaire du domaine Rel qui peut ainsi activer l'expression de ses gènes cibles (Stoven et al., 2000). On peut noter que contrairement aux Mammifères, la dégradation de la région "I- κ B-like" de Relish est indépendante du protéasome.

Très peu d'éléments sont connus à l'heure actuelle quant aux molécules impliquées dans le déclenchement de la cascade de signalisation. Une protéine transmembranaire de la famille des PGRPs, (PGRP-LC pour Peptidoglycan Recognition Protein-LC), agit comme récepteur de la voie IMD (Gottar et al., 2002; Ramet et al., 2002) (Figure 4). Les molécules intervenant en aval de PGRP-LC jusqu'à l'activation de Relish ne sont pas toutes connues mais des études génétiques ont permis d'identifier une série de protéines impliquées dans la voie de signalisation intracellulaire et qui ont toutes un homologue chez les Mammifères. On trouve tout d'abord trois protéines à DD : (1) IMD, dont la mutation a permis d'identifier cette voie, possède de fortes similitudes avec la protéine RIP (TNF-Receptor-Interacting Protein) des Mammifères, (2) l'homologue de FADD (Fas Associated Death Domain Protein) DmFADD et (3) la caspase DREDD (death-related ced-3/Nedd2-like protein), homologue de la caspase-8 de Mammifères. Ces protéines semblent former un complexe adaptateur en aval de PGRP-LC. En effet, DmFADD peut interagir à la fois avec IMD et DREDD par l'intermédiaire de leur DD. On retrouve ensuite, en aval d'IMD, un homologue de TAK1 (Transforming growth factor-β Activated Kinase 1) des Mammifères, probablement impliqué dans l'activation du signalosome IKK qui va phosphoryler Relish et permettre son clivage. Ce complexe IKK regroupe les homologues de la kinase IKK- β et IKK- γ (I- κ B kinase β et γ) des Mammifères : IRD5 (Immune Response Deficient 5) et Kenny, respectivement (Lu et al., 2001; Rutschmann et al., 2000b; Silverman et al., 2000). L'activation complète de Relish dépend également de la caspase DREDD.



Figure 4 : Représentation schématique de la voie IMD de la drosophile. La voie IMD induit la transcription du gène de la Diptéricine en réponse à une infection par les bactéries à Gram-négatif. (Voir texte pour les détails).

Une étude au laboratoire a permis de mettre en évidence, au cours d'une mutagenèse du chromosome II, un gène dont la mutation a été nommée *galère*. Comme les mutants *imd*, les mouches *galère* étaient sensibles aux infections par les bactéries à Gram-négatif mais pas aux autres types d'infections. Le gène *galère* code une protéine présentant des similitudes de séquence avec la protéine TAB2 (TAK1-associated kinase) de Mammifères et pourrait former un complexe avec TAK1 (Gobert et al., soumis).

L'interaction directe de PGRP-LC avec IMD a été récemment démontrée, de même que l'importance du domaine cytoplasmique de PGRP-LC dans la transduction du signal (Choe et al., 2005).

Le phénotype des mutants perte de fonction PGRP-LC est moins sévère que celui des mutants *kenny*, suggérant que PGRP-LC n'est pas le seul élément activant la voie IMD (Gottar et al., 2002; Ramet et al., 2002). La surexpression d'un autre membre de la famille des PGRPs, PGRP-LE, conduit dans les larves à l'expression constitutive de la Diptéricine, le marqueur de l'activation de la voie IMD (Takehana et al., 2002). L'étude du phénotype d'individus perte de fonction pour cette PGRP circulante, a permis de montrer qu'elle agit en parallèle et en synergie avec PGRP-LC lors de la résistance aux infections par *E. coli*. De plus, la surexpression de PGRP-LE active la cascade de mélanisation (Takehana et al., 2004).

La voie IMD, en plus de contrôler la synthèse de peptides antimicrobiens au travers de Relish, pourrait être impliquée dans d'autres processus. Plusieurs études transcriptomiques globales par puces à ADN ont révélé que cette voie contrôle l'expression d'environ 250 gènes (De Gregorio et al., 2001; De Gregorio et al., 2002b; Irving et al., 2001). D'autre part, l'activation de la voie IMD peut induire des mécanismes d'apoptose, peut-être par l'activation de la voie Jun Kinase (JNK) en aval de TAK-1 (Boutros et al., 2002; Park et al., 2004).

B.2. La voie Toll

Cette deuxième voie de signalisation est impliquée dans la résistance aux infections par les champignons et les bactéries à Gram-positif. Les peptides antimicrobiens contrôlés par cette voie sont principalement la Drosomycine (antifongique) et la Défensine (anti Gram-positif). La synthèse de ces peptides est considérablement réduite chez des mutants perte de fonction pour cette voie. Ces mutants sont incapables de résister à des infections par les champignons et les bactéries à Gram-positif tandis qu'ils ne présentent pas de susceptibilité vis-à-vis des infections par les bactéries à Gram-négatif.

Le gène *Toll* a été mis en évidence au début des années 1980 pour son rôle dans la formation de la polarité dorso-ventrale au cours de l'embryogenèse de la drosophile (Anderson and Nusslein-Volhard, 1984; Belvin and Anderson, 1996). Lors de ce processus développemental, la coupure de la protéine Spaetzle (SPZ) résulte d'une cascade protéolytique impliquant Nudel, Gastrulation defective, Snake et Easter (Dissing et al., 2001; LeMosy et al., 2001). Finalement Easter coupe SPZ (DeLotto and DeLotto, 1998). La protéine SPZ clivée devient alors le ligand du récepteur Toll qui, une fois activé, induit une cascade de signalisation intracellulaire qui implique les protéines Tube, Pelle et le complexe Cactus-Dorsal.

En 1996, Bruno Lemaitre au laboratoire a observé que la cassette de gènes *spaetzle*, *Toll*, *tube*, *pelle* et *cactus* sont également impliqués dans la réponse immunitaire de la drosophile adulte (Lemaitre et al., 1996). Par la suite, une autre molécule impliquée dans les deux processus a été identifiée, l'homologue de MyD88, DmMyD88 (Horng and Medzhitov, 2001; Tauszig-Delamasure et al., 2002).

B.2.1. Signalisation intracellulaire

Toll est un récepteur transmembranaire dont le domaine extracellulaire contient des répétitions riches en leucine alors que le domaine intracytoplasmique présente des similitudes avec celui du récepteur de l'interleukine 1 (domaine TIR pour Toll/Interleukine-1 Receptor). Une fois activé par son ligand SPZ, Toll interagit avec un complexe adaptateur composé de DmMyD88, Tube et Pelle, molécules possédant toutes trois un DD (Sun et al., 2002) (Figure 5). En plus de son DD, DmMyD88 possède un domaine TIR qui interagit avec celui de Toll. L'activation de ce complexe aboutit indirectement à la phosphorylation de l'inhibiteur Cactus, homologue de I-κB, et à sa dégradation par le protéasome. Bien que Pelle possède une activité kinase, elle ne phosphoryle pas directement Cactus (Grosshans et al., 1994) ; la kinase impliquée dans ce phénomène n'a pas encore été identifiée. La dégradation de Cactus libère un facteur de transcription de la famille des protéines Rel, permettant sa translocation dans le noyau et l'activation de ses gènes cibles. Deux protéines Rel sont activées par la voie Toll : Dorsal et DIF (Dorsal-related Immune Factor). Dorsal est impliqué dans la mise en place de l'axe dorso-ventral chez l'embryon. Chez l'adulte, seul DIF est impliqué dans la réponse immunitaire en contrôlant l'expression des peptides antimicrobiens (Ip et al., 1993; Manfruelli et al., 1999; Rutschmann et al., 2000a).

B.2.2. Signalisation extracellulaire

Chez la drosophile adulte, il est établi qu'une cascade protéolytique déclenchée par l'infection, conduit à la coupure de SPZ qui devient alors un ligand capable d'interagir avec le récepteur Toll. Au début de ma thèse, aucun zymogène de la cascade n'était connu ; seul l'inhibiteur de sérine protéase (serpine) Necrotic (NEC) était identifié (Levashina et al., 1999) (Figure 5). L'absence de NEC dans l'hémolymphe est corrélée avec la coupure constitutive de SPZ, se traduisant par l'expression de Drosomycine en l'absence de stimulus immun. Le phénotype des mutants *nec* se traduit également par l'observation de taches de nécrose sur le corps et à la base des pattes, un gonflement de l'abdomen par de l'hémolymphe et une létalité quelques jours après émergence.

Depuis, une étude effectuée au laboratoire par le groupe de Jean-Marc Reichhart a permis de caractériser une protéase à sérine impliquée dans l'activation de la voie Toll : la protéase Perséphone (PSH, (Ligoxygakis et al., 2002a)). Cette protéase agit en aval de NEC et pourrait être la protéase cible de cette serpine. De manière surprenante, les drosophiles mutantes pour PSH présentent toutes les caractéristiques d'un mutant de la voie Toll suite à une infection fongique, mais se comportent comme des mouches sauvages vis-à-vis des infections par des bactéries à Gram-positif. Ces résultats ont



Figure 5 : Représentation schématique de la voie Toll de la drosophile. La voie Toll induit l'expression du gène de la Drosomycine en réponse à des infections par des bactéries à Gram-positif ou par des champignons. (Voir texte pour les détails).

permis de montrer qu'au moins deux cascades peuvent contrôler l'activation de Toll au cours de la réponse immunitaire : l'une est activée par les champignons et implique la protéase PSH et la serpine NEC, l'autre est activée par les bactéries à Gram-positif.

B.2.2.1. Activation de la voie Toll par les bactéries à Gram-positif

Les travaux réalisés au laboratoire pour rechercher les molécules impliquées dans la reconnaissance des pathogènes ont mené les équipes de Julien Royet et de Dominique Ferrandon à étudier les protéines de la famille des PGRPs (Peptidoglycan Recognition Proteins) et des GNBP/βGRPs (Gram-Negative Binding Proteins/βGlucan Recognition Proteins). Ainsi, une mutation identifiée sur PGRP-SA, une protéine de la famille des PGRPs courtes, augmente la susceptibilité des drosophiles vis-à-vis des bactéries à Gram-positif, mais n'affecte pas la réponse immunitaire vis-à-vis des champignons (Michel et al., 2001).

Parallèlement, une étude sur GNBP1 a révélé que les mutants pour cette protéine présentaient le même phénotype que ceux pour PGRP-SA. Il a été montré depuis que ces deux molécules pourraient agir en formant un complexe afin de déclencher la réponse immunitaire spécifique des bactéries à Gram-positif (Gobert et al., 2003; Pili-Floury et al., 2004).

Par ailleurs, il a été observé que des infections par certaines bactéries à Grampositif peuvent activer la voie Toll indépendamment du complexe PGRP-SA/GNBP1, suggérant un ou des récepteur(s) alternatif(s). Ainsi, un autre membre des PGRPs courtes, PGRP-SD, a été étudié au laboratoire par Vincent Bischoff et collaborateurs. Il a été montré que PGRP-SD peut agir comme récépteur pour certaines bactéries à Grampositif, telles que *Staphylococcus aureus*, *Staphylococcus saprophyticus* ou *Streptococcus pyogenes*, et qu'il existe une coopération entre PGRP-SD et le complexe PGRP-SA/GNBP1 (Bischoff et al., 2004). Ainsi, trois protéines circulantes peuvent reconnaître les bactéries à Gram-positif et déclencher une réponse immune efficace : PGRP-SA, PGRP-SD et GNBP1 (cf. Figure 5). Certaines bactéries à Gram-positif peuvent être reconnues par un seul type de récepteur tandis que d'autres nécessitent une combinaison de plusieurs protéines.

B.2.2.2. Activation de la voie Toll par les champignons

Une autre molécule de la famille des GNBP, GNBP3, a été étudiée au laboratoire. Son implication dans la reconnaissance des champignons a été démontrée (Gottar et al., 2002). Lors de l'études des mutants pour GNBP3, il a été constaté que GNBP3 n'intervient pas en amont de PSH pour activer la voie Toll mais définit une autre voie de signalisation extracellulaire (Gottar et al., soumis). L'infection par des champignons tués à la chaleur a permis de montrer que GNBP3 reconnaît des motifs de la membrane des champignons, tels que les β -1,3-glucanes, tandis que PSH semble être activée par des protéases secrétées par les champignons.

Les études récentes menées au laboratoire ont ainsi permis de montrer qu'il existe au moins quatre systèmes de reconnaissance des pathogènes, deux pour les champignons et deux pour les bactéries à Gram-positif, aboutissant à la coupure protéolytique de SPZ dans l'hémolymphe (cf. Figure 5).

Comme dans le cas de la voie IMD, la fonction de la voie Toll ne se limite pas à l'induction des peptides antimicrobiens. D'après des analyses transcriptomiques par puces à ADN, près de 400 gènes semblent régulés suite à une infection bactérienne ou fongique (De Gregorio et al., 2001; De Gregorio et al., 2002b; Irving et al., 2001). Ces gènes pourraient participer aux mécanismes de défense de la drosophile.

C. Etudes post-génomiques de la réponse immunitaire de la drosophile

Grâce au séquençage du génome de la drosophile en mars 2000 (Adams et al., 2000), les analyses post-génomiques sur cet insecte ont connu un essor exceptionnel. L'étude de l'expression des gènes et de leurs produits peut être abordée selon deux approches : l'analyse transcriptomique, qui consiste à mesurer le niveau quantitatif des ARN messagers, et l'analyse peptidomique/protéomique, qui s'attache directement à l'étude des peptides/protéines présents à un instant donné. Ces deux approches sont complémentaires, cependant seule la seconde permet d'étudier les modifications post-

traductionnelles, jouant un rôle essentiel dans la fonctionnalité des protéines, ou encore les interactions au sein de complexes multiprotéiques fonctionnels.

L'étude des mécanismes de défense de la drosophile par des analyses transcriptomiques, à l'aide de puces à ADN (microarrays), et peptidomiques, basées sur la spectrométrie de masse, a été mise en place au laboratoire. Je résume dans cette partie de l'introduction les résultats obtenus grâce à ces deux approches.

C.1. Approche différentielle transcriptomique

Afin d'étudier le profil transcriptionnel global de drosophiles adultes après infection expérimentale, le groupe de Charles Hétru a entrepris en 2000 une analyse par puces à ADN. Cette technique repose sur la synthèse *in situ* d'oligonucléotides correspondant à chacun des 13600 gènes prédits dans le génome de la drosophile. L'ARN de drosophiles adultes saines ou infectées extrait est alors hybridé sur les puces. Cette approche permet de comparer les valeurs correspondant au niveau d'expression de chaque gène de la drosophile avant et après une infection expérimentale.

La fiabilité de cette méthodologie a été validée en vérifiant la corrélation entre l'induction des peptides antimicrobiens sur les puces et les nombreuses informations obtenues lors des travaux antérieurs. Ces expériences ont permis de mettre en évidence l'induction, par un facteur d'au moins deux, d'environ 400 gènes autres que ceux codant les peptides antimicrobiens. Seule la moitié de ces gènes sont connus ou présentent des similitudes avec des gènes dont les fonctions sont établies.

En faisant une analyse comparée des profils transcriptionnels induits par une infection fongique (*Beauveria bassiana*) ou bactérienne (*Micrococcus luteus*, *Escherichia coli*), les points suivants ont pu être observés (Figure 6) : (i) 13% des gènes induits le sont par toutes ces infections; (ii) 50% des gènes sont induits par les infections bactériennes, mais non par les champignons ; (iii) un grand nombre de gènes ne sont induit que par les infections par les bactéries à Gram-négatif (10%) ; (iv) 23% des gènes ne sont induits que par les infections fongiques ; (v) enfin, 11% des gènes sont induits de façon concomitante par



Figure 6 : Diagramme de Venn représentant le nombre de gènes induits au cours de la réponse immunitaire de la drosophile suite aux infections, soit par une bactérie à Gram-négatif (*Escherichia coli*), soit par une bactérie à Gram-positif (*Micrococcus luteus*), soit par un champignon (*Beauveria bassiana*).

les champignons et les bactéries à Gram-positif et 2% par les champignons et les bactéries à Gram-négatif.

Cette étude a également permis de mettre en avant un certain nombre de gènes qui n'avaient pas été détectés précédemment comme étant induits par la réponse immunitaire chez la drosophile mais dont l'homologie avec des systèmes de Mammifères en font des candidats très intéressants. On peut noter par exemple l'induction de gènes codant des protéines de type kallikréine, corine ou hepsine dont le rôle dans la réponse inflammatoire, la circulation sanguine ou la coagulation, respectivement, est connu chez les Mammifères.

C.2. Approche différentielle peptidomique

Dans le but de rechercher de nouvelles molécules de la réponse immunitaire de la drosophile, le groupe de Philippe Bulet a entrepris dès 1997 au laboratoire une approche différentielle peptidomique afin de réaliser la cartographie de l'ensemble des peptides induits par une infection expérimentale. Cette approche se basait sur la détection la plus exhaustive possible des petites molécules présentes dans l'hémolymphe.

C.2.1. Mise en évidence des "Drosophila Immune-induced Molecules"

Une première étude différentielle a été réalisée par spectrométrie de masse MALDI-TOF (Matrix Assisted Laser Desorption Ionisation - Time Of Flight) entre l'hémolymphe d'une drosophile contrôle et d'une drosophile infectée expérimentalement (Uttenweiler-Joseph et al., 1998). Vingt quatre "*Drosophila* Immune-induced Molecules" (DIMs), de masse moléculaire comprise entre 1,5 et 15 kDa, ont ainsi été détectées dans le sang des drosophiles en réponse à l'infection expérimentale (Figure 7). Quatre de ces molécules correspondaient à certains des peptides antimicrobiens déjà identifiés (DIMs 9 et 11, Drosocines mono et disacharidiques ; DIM 17, Metchnikowine et DIM 19, Drosomycine), toutes les autres représentaient de nouvelles substances, dont la structure et le rôle exact dans la réponse immunitaire restaient à définir. Une étude cinétique a permis de montrer que ces DIMs sont présentes dans



Figure 7 : Spectre de masse MALDI-TOF de l'hémolymphe d'une drosophile contrôle ou infectée expérimentalement par un mélange bactérien. L'hémolymphe d'une mouche adulte (contrôle ou 24 h après infection par un mélange de *Escherichia coli* et *Micrococcus luteus*) est prélevée et analysée par spectrométrie de masse MALDI-TOF. 24 DIMs (*Drosophila* Immune-induced Molecules) sont trouvées dans l'hémolymphe des drosophiles infectées. Quatre d'entre elles correspondent à des peptides antimicrobiens déjà identifiés : la Drosocine monosaccharidique (1S) ou disaccharidique (2S), la Metchnikowine et la Drosomycine.

l'hémolymphe de drosophile dès trois heures après l'infection. De plus, deux phases d'induction ont été observées : l'une où les peptides ont une concentration maximale six heures après l'infection (DIMs 1 à 8, 14, 15, 22 et 23), l'autre avec une concentration maximale à 24 h (DIMs 9 à 13, 16 à 19 et 24). De plus, l'analyse de mutants pour les voies Toll ou IMD a permis de montrer que la plupart de ces molécules sont régulées par la voie Toll. Seuls les DIMs 9, 11, 15 et 16 semblent régulées par la voie IMD. DIM 17, qui correspond à la Metchnikowine, peut être régulée par l'une ou l'autre des deux voies.

Dans le but de caractériser les DIMs, une seconde étude différentielle a été réalisée, combinant l'analyse par chromatographie liquide haute performance (CLHP) de l'hémolymphe de drosophiles contrôles ou infectées, suivie d'une analyse par spectrométrie de masse MALDI-TOF de toutes les fractions obtenues. Cette seconde étude a permis de détecter plus de 75 DIMs, de masse moléculaire comprise entre 1,5 et 15 kDa, induites dans l'hémolymphe par une infection expérimentale.

Les DIMs ont été purifiées en combinant des techniques de CLHP (phase inverse, échange d'ions, exclusion stérique). La structure primaire a été étudiée grâce au séquençage N- et C-terminal, à des traitements protéolytiques, à la spectrométrie de masse en tandem, au clonage cDNA et à la recherche dans les banques de données. Trente-trois DIMs ont pu être caractérisées ainsi (Tableau I). En plus des peptides antimicrobiens déjà retrouvés par spectrométrie de masse (DIMs 9, 11, 17 et 19), les DIMs 25 et 26 correspondent aussi à deux peptides antimicrobiens connus, la Cécropine et la Défensine, respectivement. Concernant les autres DIMs, certaines présentent des similitudes avec des molécules connues et d'autres sont totalement originales. Les principales caractéristiques des DIMs sont récapitulées ci-dessous.

C.2.1.1. DIMs présentant des similitudes avec des molécules déjà connues

DIM 15 correspond à une forme trisaccharidique de la Drosocine. Le troisième sucre est porté au niveau d'un second site de O-glycosylation (sérine en position 7) (Rabel et al., 2004).

DIM 16 a été récemment caractérisée comme étant le prodomaine maturé d'une isoforme d'Attacine, l'Attacine C (MPAC pour Maturated Prodomain of Attacin C).
DIMs	Identité du gène	Masses (Da)	Identité / homologie	Séquences
DIMs présentant des similitudes avec des molécules connues				
17	CG8175	3046	Metchnikowine	HRRQGPIFDTRPSPFNPNQPRPGPIY
19	CG10810	4889	Drosomycine	DCLSGRYKGPCAVWDNETCRRVCKEEGRSSGHC SPSLKCWCEGC
25	CG1365	4156	Cécropine	GWLKKIGKKIERVGQHTRDATIQGLGIAQQAAN VAATAR <i>amidé</i>
26	CG1385	4354	Défensine	ATCDLLSKWNWNHTACAGHCIAKGFKGGYCNDK AVCVCRN
9	CG10816	2401	Drosocine monosaccharidique	GKPRPYSPRP <u>T</u> SHPRPIRV
11	CG10816	2564	Drosocine disaccharidique	GKPRPYSPRP <u>T</u> SHPRPIRV
15	CG10816	2767	Drosocine trisaccharidique	GKPRPY <u>S</u> PRP <u>T</u> SHPRPIRV
16	CG4740	2971	Prodomaine maturé de l'Attacine C	ZRPYTQPLIYYPPPP <u>T</u> PPRIYRA
21	CG10810	5934	Drosomycine N-glycosylée	DCLSGRYKGPCAVWD <u>N</u> ETCRRVCKEEGRSSGHC SPSLKCWCEGC
27	CG12763	9196	Diptéricine O-glycosylée	DDMTMKPTPPPQYPLNLQGGGGGQSGDGFGFAV QGHQKVWTSDNGRHEIGLGGGQHLGGPYGNSEP SWKVGSTYTYRFPNF
29	CG9046	2603	fragment de VM26Ab	ETIQLQPTQGILIPAPLAENIRVS
30	CG11129	3117	fragment de Yp3	EVNAKSPFGQRSPAHKQAAYHGMHHAQN
33	CG16712	6794	Inhibiteur de protéase de type Kunitz	LKDPICGLPAGIDGNGLIKCAAFIPSFSYHPET NSCEKFIYGGCGGNENRFGTQELCEQKCKE
DIMs sans similitude avec des molécules connues				
1	CG18108	1666		GNVIINGDCRVCNVHG amidé
2	CG18106	1688		GNVVINGDCKYCNVHG amidé
3	CG16844	1700		GNVIINGDCRVCNVRA
23	CG15066	9482		GNVIIGGVCQDCSPPVAENVVVGGQSYRTGRPG QGTVYINSPGAYPGALDGPIRRTGAGGGGGGGGT RYPDGYSGRLPGGTYLHLKDCVGCSISGGGD
10	CG18279	2520		ZLHVARPDRTVTIGNGGVYIORS
12	CG18279	2572		ZFHVERPGRTVDVGNGGFYIQRG
13	CG18279	2650		ZFHVERPDRTVDFGNGGFSAQRF
24	CG18279	10031		ZHTYDGRNGPHVFGSPGNQVYIRGQNEGTYSVP GVGGQFQNAPQRGEHVYTDEAGNTFVNRKNAGG PASHTISGPNFSAKNLGPNGAKSVGIPORA
4	CG15231	1722		GTVLIQTDNTQYIRT amidé
14		2694		GTQVIHAGGHTLIQTDRSQYIRKN
5	CG18279	1914		ZFHVERPDRTVDFGNGG
6	CG18279	1955		ZFHVERPGRTVDVGNGGF
8	CG18279	2348		ZFHVERPDRTVDFGNGGFSAQ
18	CG10332	4625		SRHTGPGNGSGSGAGSGNPFRSPSSQQRPLYYD APIGKPSKTMYA
31	CG31509	11935	TotA	YSDEDREADSLRIAEIIKNAQDDDSKINSTQEL LDIYRRLYPSLTPEERESIDKFVNEHTDAIIID GVPIQGGRKARIVGKIVSPGVKGLATGFFEELG SKLAOLFT

Tableau I : Récapitulation des caractéristiques structurales obtenues sur les DIMs. Les DIMs ont été numérotées de 1 à 24 dans l'ordre de leur apparition sur les spectres de masse et de 25 à 33 dans l'ordre chronologique de leur caractérisation. Les DIMs 7, 20, 22, 28 et 32 n'apparaissent pas dans ce tableau par manque d'information de séquence, excepté pour DIM 20 numéroté par erreur alors qu'il correspond à l'ion dichargé du DIM 24. Les résidus soulignés indiquent les positions des modifications post-traductionnelles identifiées.

Alors que sa séquence possède de fortes similitudes avec la Drosocine, MPAC possède *in vitro* une activité antimicrobienne restreinte contre quelques souches de bactéries à Gram-négatif, même a des concentrations élevées. En revanche, l'addition de faibles doses de Cécropine (en dessous du seuil d'activité de cet autre peptide antimicrobien de drosophile) confère au mélange une activité antimicrobienne supérieure contre certains germes à Gram-négatif (Rabel et al., 2004).

DIMs 21 et 27 sont également des isoformes de peptides antimicrobiens. DIM 21 présente la même séquence N-terminale que la Drosomycine, avec un écart de masse compatible avec une modification de l'asparagine située dans un site consensus de N-glycosylation. DIM 27 semble être une forme O-glycosylée de Diptéricine.

DIMs 29 et 30 sont des produits de clivage de protéines connues. DIM 29 correspond à un fragment N-terminal d'une protéine de la membrane vitelline (VM26Ab) tandis que DIM 30 est un fragment du produit du gène codant la vitellogenine III (Yp3). Des exemples de protéines ayant des similitudes structurales avec des vitellogénines ont déjà été décrites comme étant impliquées dans des mécanismes de la réponse immunitaire (Hall et al., 1999; Lee et al., 2000). De plus, un fragment du produit du gène *CG11064*, présentant des similitudes avec les vitellogénines, est surexprimé dans l'hémolymphe de larves de drosophiles après stimulation par du LPS (Vierstraete et al., 2004). L'apparition de DIM 29 et 30 uniquement chez des drosophiles ayant subi une infection expérimentale, peut résulter d'activités protéolytiques d'enzymes activées par l'infection.

DIM 33 est un inhibiteur de protéase à sérine de type Kunitz. Des études globales d'expression des ARN messagers ont montré que plusieurs gènes codant ce type d'inhibiteurs sont surexprimés suite à une infection microbienne (De Gregorio et al., 2001; Irving et al., 2001; Vierstraete et al., 2004).

C.2.1.2. DIMs sans similitudes avec des molécules connues

Les DIMs 1, 2, 3 et 23 appartiennent à une même famille. Les DIMs 1, 2 et 3 sont de petits peptides présentant de fortes similitudes de séquence (Uttenweiler-Joseph et al., 1998). Les transcrits de ces DIMs ont été trouvés fortement surexprimés lors d'analyses par puces à ADN (Boutros et al., 2002; De Gregorio et al., 2001). DIMs 1 et

2 sont amidés à leur extrémité C-terminale. DIM 23 est une molécule de taille plus importante, présentant des similitudes avec les DIMs 1, 2 et 3 dans sa partie N-terminale. Cette famille de DIMs ne partage aucune similarité avec des molécules connues.

Les DIMs 10, 12, 13 et 24 appartiennent aussi à une même famille. Ces DIMs sont caractérisées par la présence d'un acide pyroglutamique à leur extrémité Nterminale. L'analyse du génome de la drosophile a révélé que ces molécules résultaient de la maturation d'un même gène multiprécurseur. Trois autres DIMs isolées appartiennent à cette famille. DIMs 5 et 8 sont des produits de la dégradation Cterminale de DIM 13 et DIM 6 correspond à la partie N-terminale de DIM 12. Des travaux de surexpression et d'invalidation du gène codant ces peptides ont été réalisés au laboratoire. Ni l'expression constitutive ni la répression de cette famille de DIMs n'a eu d'effet sur la survie des drosophiles après diverses infections (D. Rabel, données non publiées). En revanche, une étude transcriptomique sur des mouches où l'expression de ce groupe de DIM est bloquée montre qu'un gène codant une protéine de la famille des spondines (la fat-spondine) est fortement exprimée suite à une infection par la bactérie à Gram-positif M. luteus. A l'inverse, chez des mouches normales pour l'expression de ce muliprécurseur, l'infection diminue l'expression du gène codant la fat-spondine (D. Rabel et C. Hétru, communication personnelle). Des souris invalidées pour un gène de la famille des spondines présentent une susceptibilité accrue à certaines infections microbiennes (He et al., 2004). L'implication de la fat-spondine dans la réponse immunitaire de la drosophile reste à déterminer.

DIMs 4 et 14 sont deux peptides présentant des similitudes de séquence entre elles mais pas avec d'autres molécules déjà décrites. DIM 4 possède une amidation C-terminale.

DIMs 7, 18 et 31 ne présentent ni particularité structurale ni similitude avec des molécules connues. Le gène *totA*, qui code pour DIM 31, est exprimé en réponse à une infection microbienne et également suite à d'autres stress (UV, chaleur, oxydants) (Ekengren and Hultmark, 2001).

Il est intéressant de noter que DIMs 2 et 4 sont également exprimées dans le système nerveux central de larves errantes de drosophiles (Baggerman et al., 2005;

Baggerman et al., 2002). Ces molécules pourraient être produites et stockées dans des cellules du système nerveux central et secrétées dans l'hémolymphe après infection.

C.2.2. Mise en évidence des molécules induites par une infection virale

Outre les modèles d'infection bactérienne ou fongique, un modèle d'infection virale a été mis en place au laboratoire. Laurence Sabatier, en collaboration avec l'équipe de Jean-Luc Imler, a réalisé une étude comparative en spectrométrie de masse MALDI-TOF de l'hémolymphe de drosophiles ayant subi ou non une injection de virus C de la drosophile (DCV pour *Drosophila* C Virus). Il a été montré que, contrairement aux infections bactériennes et fongiques, l'infection virale déclenche une réponse humorale très discrète. En effet, seule la Pherokine-2 (PHK-2), peptide de 12,8 kDa, a été détectée dans l'hémolymphe après l'infection virale. Elle est sécrétée spécifiquement dans l'hémolymphe des drosophiles 48 h après la stimulation (Sabatier et al., 2003). Les peptides antimicrobiens sécrétés en réponse aux infections bactériennes ou fongiques ne sont pas détectés dans l'hémolymphe des mouches infectées par le DCV. Des drosophiles mutantes pour les voies Toll ou IMD présentent la même sensibilité à l'infection par le virus que des mouches de type sauvage, suggérant que la défense antivirale éventuelle de la drosophile est indépendante de ces voies de signalisation.

PHK-2, encodée par le gène *pebIII*, est exprimée dans le bulbe éjaculateur des drosophiles mâles (Dyanov and Dzitoeva, 1995). Avec le produit des gènes *CG6642* (PHK-1) et *CG9358* (PHK-3), elle fait partie d'une famille pour laquelle un rôle dans l'olfaction ou la chémoperception a été proposé.

Une analyse du profil transcriptionnel global de la réponse immunitaire de la drosophile face à une infection virale a été récemment réalisée grâce à des puces à ADN. Une centaine de gènes dont l'expression augmente au moins d'un facteur deux ont été identifiés suite à l'infection par le DCV. Les deux tiers de ces gènes ne sont pas régulés après des infections bactériennes ou fongiques, suggérant une réponse virale spécifique (Dostert et al., soumis). D'ailleurs, les gènes codant les peptides antimicrobiens régulés par les voies Toll ou IMD ne sont pas induits suite à l'infection virale, confirmant les résultats obtenus précédemment par spectrométrie de masse. Parmi les gènes fortements induits par le DCV mais pas par piqûre septique, *vir-1* (pour

Virus induced RNA -1) a été plus amplement étudié. L'analyse globale de deux autres modèles d'infection virale (le "Flock House Virus" et le virus de Sindbis) montre que les transcrits de *vir-1* sont également fortement induits après infection. Les études génétiques réalisées au laboratoire indiquent que la réponse antivirale coopère avec la voie JAK/STAT, impliquée dans le développement, mais que cette voie n'est pas suffisante pour contrôler l'expression de *vir-1*.

D. Objectif de ce travail

L'importance prise par l'analyse génomique engendre une somme croissante d'informations sur les séquences, à laquelle s'ajoutent des données quantitatives portant sur l'expression des ARN messagers étudiées notamment grâce à la technique des puces à ADN. Cette technique permet de quantifier, de manière globale, l'expression des gènes dans différentes conditions expérimentales. Cependant, nous savons qu'il n'existe pas obligatoirement de corrélation entre la quantité d'ARN messagers et l'abondance des protéines correspondantes. De plus, il a été montré que des mécanismes de modifications post-traductionnelles peuvent contrôler le taux de synthèse, la demi-vie des protéines ainsi que leur fonctionnalité. Il semble par conséquent indispensable de compléter l'étude de l'expression des ARN messagers par l'analyse directe des peptides/protéines et ceci n'est accessible que par l'analyse peptidomique/protéomique. En effet, par rapport à la technique des microarrays, la seconde approche permet d'étudier des changements dynamiques au niveau des peptides/protéines en réponse à des perturbations externes ou internes (Corthals et al., 2000). Le terme de protéomique a été défini pour la première fois par Wilkins en 1995 (Wilkins et al., 1995). Cette discipline, également connue sous le nom de "projet protéome", consiste à caractériser toutes les PROTEines exprimées par le génOME d'un organisme. A l'origine, il n'y avait pas eu de distinction faite entre l'analyse des petites protéines (peptides) et des plus grandes. Face à l'ampleur des données regroupées par la protéomique, la peptidomique s'est alors imposée comme une discipline à part entière et le terme à été adopté en 2001 (Schulz-Knappe et al., 2001).

Une analyse peptidomique par spectrométrie de masse MALDI-TOF avait initié cette stratégie au laboratoire en étudiant les molécules de faible masse moléculaire

(inférieure à 15 kDa) présentes dans l'hémolymphe suite à une infection microbienne. Le but de mon travail de thèse était de mettre en place une stratégie complémentaire permettant d'étudier les molécules de masse moléculaire supérieure à 15 kDa impliquées dans la réponse immunitaire de la drosophile. Dans cet objectif, il était nécessaire de trouver une méthode séparative adaptée à cette gamme de molécules, et qui permette l'analyse d'un mélange protéique complexe (hémolymphe de drosophile).

La première approche que j'ai mise en place au cours de ma thèse est basée sur la combinaison d'électrophorèse bidimensionnelle (2D) et d'identification des protéines par digestion, analyse des fragments par spectrométrie de masse et recherche dans les banques de données. L'électrophorèse 2D combine la séparation de molécules suivant leur point isoélectrique et leur masse moléculaire (Hochstrasser et al., 1988). Cette méthode est très performante et très sensible, puisque si l'on utilise une coloration appropriée (comme celle au nitrate d'argent), il est possible de détecter plusieurs centaines de protéines à des abondances de l'ordre du nanogramme. Ceci est très important compte tenu de la faible quantité d'hémolymphe présente dans une drosophile (0,1 µL) et des faibles concentrations auxquelles certaines protéines régulatrices peuvent être présentes. De plus, elle permet de séparer sur le gel une même protéine selon qu'elle porte ou non des modifications post-traductionnelles. Au début de ma thèse, je me suis attachée à la mise au point de la méthodologie de l'électrophorèse 2D, notamment à l'obtention de cartes protéiques reproductibles. J'ai ensuite validé le protocole d'identification des protéines issues de ces cartes. J'ai enfin axé mes recherches sur l'étude des molécules induites ou réprimées lors de la réponse immunitaire de la drosophile. Ces travaux seront exposés dans le premier chapitre de la partie résultats.

Par la suite, j'ai eu l'opportunité d'étudier les molécules de l'hémolymphe de drosophiles par une seconde stratégie protéomique puissante, complémentaire à celle par électrophorèse 2D, utilisant le marquage de protéines à l'aide de réactifs isotopiques (ICAT pour Isotope Coded Affinity Tag). Cette méthode est basée sur l'introduction d'un isotope lourd sur les protéines d'un premier échantillon, tandis que le second est marqué par l'isotope léger. Les deux échantillons sont alors mélangés et traités simultanément jusqu'à une analyse chromatographique couplée à la spectrométrie de masse en tandem (LC-MS/MS). Le rapport d'intensité des ions porteurs des marqueurs

lourds/légers permet de calculer les concentrations relatives d'une protéine donnée dans les deux échantillons. J'ai eu la chance de pouvoir appliquer cette stratégie performante dans le cadre d'une collaboration avec les Drs Alan Aderem et Adrian Ozinsky à l'Institut for Systems Biology (Seattle). Le deuxième chapitre de ce manuscrit rassemblera les données obtenues par cette seconde approche protéomique.

L'ensemble des résultats obtenus à partir de ces deux techniques sera discuté à la fin de ce manuscrit.

MATERIEL ET METHODES

MATERIEL ET METHODES

A. Matériel biologique

A.1. Souches de drosophiles

L'élevage des drosophiles est réalisé au laboratoire dans des flacons maintenus à 25°C, sur un milieu gélosé semi-solide composé d'extrait de malt, de farine de soja et de maïs, de mélasse, d'agar et de levure de bière sèche. J'ai travaillé, dans la plupart des cas, avec la souche cinnabar brown (cnbw/cnbw) qui est sauvage pour les chromosomes I, III et IV et qui porte, sur le chromosome II les marqueurs cinnabar (cn) et brown (bw) à l'origine, chez les individus homozygotes, d'une absence de pigmentation de l'œil (œil blanc).

A.2. Infections expérimentales

A.2.1. Infections par piqûre septique

Au cours de mon travail, j'ai utilisé les souches bactériennes suivantes : Bactéries à Gram-négatif : *Escherichia coli* 1106 et *Agrobacterium tumefaciens*. Bactéries à Gram-positif : *Micrococcus luteus*, *Staphylococcus saprophyticus* et *Enterococcus faecalis*.

La préparation du concentré bactérien (milieu d'infection), se fait à partir de l'ensemencement d'une colonie du microorganisme considéré dans 50 mL de milieu Luria-Bertani. Après agitation une nuit à 37°C, ces cultures sont centrifugées (3500 g, 10 min) et le surnageant est éliminé. Les culots bactériens obtenus servent à l'infection des drosophiles par inoculation. Les drosophiles adultes mâles âgées de 3 à 4 jours sont anesthésiées au CO_2 et infectées par piqûre septique au niveau du thorax, sous l'insertion alaire, à l'aide d'une minutie (fine aiguille) en tungstène préalablement trempée dans le concentré bactérien. Des drosophiles piquées à l'aide d'une aiguille propre (piqûre non septique) sont analysées en tant que contrôles. Les drosophiles sont conservées à 25°C après piqûre jusqu'à la collecte de l'hémolymphe.

A.2.2. Infections naturelles

Les infections naturelles sont réalisées avec le champignon entomopathogène *Beauveria bassiana*. L'infection se fait en introduisant et en agitant les drosophiles, préalablement anesthésiées au CO₂, dans une boîte de Pétri contenant une culture sporulante du champignon filamenteux. Les drosophiles sont alors transférées dans un tube de milieu frais et conservées à 29°C, température optimale à laquelle les spores déposées sur la cuticule de l'insecte peuvent germer. Des drosophiles non infectées, conservées à la même température, servent de contrôle lors des infections naturelles.

A.2.3. Infections virales

Les stocks de suspension de virus C de la drosophile (DCV) sont préparés dans une solution de Tris/HCl 10 mM (pH 7,5). Les infections sont réalisées par injection à l'aide d'un nanoinjecteur (Nanoject, Drummond Scientific, Broomall, PA) de 4,6 nL de suspension virale (2.10¹¹ LD50/mL) dans le thorax des drosophiles adultes. Des mouches injectées avec le même volume de Tris/HCl 10 mM (pH 7,5) servent de contrôle. Les drosophiles sont incubées à 25°C.

A.3. Collecte de l'hémolymphe

L'hémolymphe des drosophiles adultes est prélevée à l'aide d'un nanoinjecteur utilisé comme porte capillaire. Un fin capillaire de verre, introduit au niveau du thorax sous l'insertion alaire, permet de collecter l'hémolymphe, environ 0,1 μ L par individu. L'hémolymphe peut être prélevée à différents temps après l'infection (de 30 secondes à 72 heures).

B. Approche par gels d'électrophorèse bidimensionnelle

Les protocoles nécessaires à la mise en place d'une approche protéomique diffèrent selon le type d'échantillon étudié et nécessitent de nombreuses mises au point.

Bien que de nombreux protocoles soient disponibles dans la littérature, ils doivent toujours être adaptés selon le type d'échantillon pour une optimisation des résultats. Les différentes conditions opératoires testées au cours de ma thèse seront récapitulées dans la première partie des résultats. Seuls sont décrits ici les protocoles finaux.

Les produits chimiques et appareils décrits dans cette partie proviennent pour la plupart de BioRad (Hercules, CA), sinon le fournisseur est indiqué entre parenthèses.

B.1. Préparation des échantillons

B.1.1. Solubilisation des protéines

L'hémolymphe collectée est immédiatement diluée à 4°C dans une solution de solubilisation des protéines adaptée à l'électrophorèse 2D, composée de 8 M urée (Q-Biogene, Carlsbad, CA; USA), 4% (m/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulfonate (CHAPS, Sigma, St-Louis, MO; USA), 0,5% (v/v) Triton X-100 (Sigma, déionisé sur résine AG[®]501-X8) et 0,5% ampholytes porteurs (Biolytes[®], pH 3-10 ou 5,5-6,7). Du dithiothréitol (DTT) est ensuite ajouté à la solution de solubilisation juste avant utilisation de manière à obtenir une concentration finale de 20 mM.

L'urée est un agent chaotrope qui permet de casser les liaisons hydrogène intra- et intermoléculaires. Le CHAPS, détergent zwitterionique, et le Triton X-100, détergent non ionique, participent à la solubilisation et à la prévention de la précipitation des protéines. Les ampholytes augmentent la solubilité et assurent une conductibilité uniforme au cours de la focalisation. Enfin, le DTT permet de couper les ponts disulfure intra- et inter-moléculaires permettant le dépliement total de la protéine.

Chaque lot d'hémolymphe est conservé à –20°C.

B.1.2. Dosage des protéines

La quantité de protéines nécessaire pour chaque expérience dépend de la taille du gel d'électrophorèse, de la gamme de pH étudiée et de la coloration utilisée.

La concentration protéique des échantillons est dosée grâce au kit RC/DCTM. Ce dosage colorimétrique est adapté du protocole de Lowry en incluant des modifications

le rendant compatible avec les réactifs utilisés pour les gels 2D (RC : Reducing agent Compatible / DC : Detergent Compatible). Le dosage est réalisé suivant les recommandations du fournisseur en ajoutant une étape de sonication de 4 min lors de la suspension du culot. Le dosage de chaque échantillon est obtenu par comparaison avec une gamme-étalon effectuée à partir d'albumine sérique bovine. Cette gamme est composée de cinq concentrations réparties régulièrement entre 0,125 et 1 g/L.

B.2. Gels d'électrophorèse 2D

L'électrophorèse 2D permet de séparer les protéines de l'hémolymphe de drosophiles selon deux dimensions. Les protéines sont séparées, tout d'abord, selon leur point isoélectrique (première dimension ou focalisation isoélectrique), puis selon leur masse moléculaire (deuxième dimension ou électrophorèse sur gel de polyacrylamide en conditions dénaturantes). Les protéines sont ensuite visualisées par coloration au nitrate d'argent ou au bleu de Coomassie colloïdal. Cette technique est appliquée à l'hémolymphe de drosophiles saines ou soumises à une infection expérimentale par différents agents microbiens.

B.2.1. Focalisation isoélectrique

Nous avons choisi de travailler sur des bandelettes de gel précoulé à gradient de pH immobilisé (IPG pour Immobilized pH Gradient), développées par Bjellqvist et collaborateurs (Bjellqvist et al., 1982), très reproductibles et permettant d'éviter les phénomènes de dérive cathodique que l'on pouvait rencontrer avec les capillaires.

La focalisation isoélectrique (IEF) permet de séparer les protéines selon leur point isoélectrique (pI), pH auquel la charge électrique globale de la protéine est nulle. L'application d'un champ électrique permet aux protéines de migrer jusqu'à atteindre leur pI et de s'y arrêter. Cette méthode permet donc de séparer de nombreuses protéines différentes mais également de discriminer entre des molécules qui auraient subi une mutation ou une modification post-traductionnelle changeant ainsi leur charge électrique. La première dimension est effectuée sur l'appareil Protean IEF Cell. Nous avons utilisé des bandelettes IPG de 7 ou de 17 cm de longueur, couvrant une gamme de pH allant de 3 à 10, de 5 à 8 ou de 5,5 à 6,7. Plus la bandelette est longue et la gamme de pH étroite, plus la quantité d'échantillon déposée peut être grande.

B.2.1.1. Réhydratation

De 15 μ g de protéines, pour les bandelettes IPG 7 cm, à 300 μ g, pour les bandelettes 17 cm, sont déposés au centre d'une rigole du plateau de focalisation et une bandelette IPG est délicatement déposée dessus. L'ensemble est protégé de la déshydratation et de l'oxydation par environ 2 mL d'huile minérale. La réhydratation s'effectue en appliquant un courant de 50 V, pendant un minimum de 10 h à 20°C.

B.2.1.2. Migration

Une mèche de papier imbibée d'eau MilliQ est placée entre la bandelette réhydratée et chacune des électrodes du plateau de focalisation. Ces mèches ont pour but de faciliter le passage du courant entre l'électrode et le gel et piéger les sels issus de l'échantillon. Le programme de focalisation nécessite trois étapes à 20°C dont le voltage diffère selon l'utilisation d'une bandelette 7 ou 17 cm :

- une phase de conditionnement (15 min à 250 V), permettant d'éliminer l'excès de sels,
- une montée linéaire en voltage (de 250 à 4000 V en 2 h pour une bandelette 7 cm, de 250 à 10000 V en 5 h pour une bandelette 17 cm)

- une phase de focalisation (5 h à 4000 V pour une bandelette 7 cm ou 8 h à 10000 V pour une bandelette 17 cm).

Enfin une étape à 500 V est appliquée jusqu'à l'arrêt du programme afin d'éviter la diffusion des protéines au sein du gel.

Les bandelettes peuvent alors être directement utilisées pour la deuxième dimension ou stockées à -80°C.

B.2.2. Electrophorèse sur gel de polyacrylamide en conditions dénaturantes

Les bandelettes IPG sont mises en présence de dodécylsulfate de sodium (SDS) qui se lie aux protéines en leur conférant une charge négative importante, noyant complètement les charges propres de la protéine. Cette propriété permet, sous l'action d'un courant, la migration des protéines de l'échantillon vers l'anode. Le gel d'acrylamide permet un tamisage moléculaire (les grosses molécules migrent plus lentement que les petites) et les protéines se répartissent le long du trajet de migration en fonction de leur masse moléculaire.

B.2.2.1. Réduction/alkylation des molécules

Les bandelettes sont équilibrées 15 min dans 2 mL d'un tampon contenant 6 M urée, 2% (m/v) SDS, 375 mM Tris-HCl (Sigma) pH 8,8, 20% (v/v) glycérol et 2% (m/v) DTT.

Cette étape permet de saturer le gel en SDS et de réduire les protéines. Les bandelettes sont ensuite placées 20 min dans le même tampon en substituant le DTT par 2,5% (m/v) d'iodoacétamide (Sigma), permettant l'alkylation des protéines et du DTT résiduel. 75 μ L de bleu de bromophénol 0,05% sont ajoutés à la solution d'alkylation afin de colorer la bandelette et suivre la migration des protéines.

B.2.2.2. Migration

La bandelette équilibrée est placée sur un gel composé de 11% acrylamide et 0,29% pipérazine diacrylamide (PDA), sans gel de concentration.

Le PDA est plus résistant que le bisacrylamide et conduit à une meilleure résolution pour la coloration à l'argent. La polymérisation de l'acrylamide et du PDA est catalysée par l'addition de persulfate d'ammonium et de N,N,N',N'-tétraméthylethylènediamine (TEMED). La bandelette est fixée en haut du gel d'acrylamide avec une solution contenant 1% d'agarose à bas point de fusion diluée dans du tampon de migration. Les bandelettes 7 cm sont déposées sur un gel de petite dimension (8 x 8 cm) et les bandelettes 17 cm sur un gel de grande dimension (20 x 20

cm). 10 μ L de marqueurs de masse moléculaire de faible gamme (21,5 à 97,4 kDa) sont déposés sur un côté du gel.

La migration des protéines est réalisée dans un tampon composé de 25 mM Trisbase, 192 mM glycine et 0,1% SDS.

Pour les petits gels, une tension de 20 V est appliquée pendant 30 min puis 150V jusqu'à la fin de la migration.

Pour les grands gels, un ampérage de 32 mA est appliqué pendant 1 h puis 48 mA jusqu'à la fin de la migration.

B.2.3. Coloration des protéines

B.2.3.1. Coloration au nitrate d'argent

La coloration utilisée est celle décrite par Shevchenko (Shevchenko et al., 1996). C'est une coloration dite "de surface", qui permet une meilleure compatibilité avec l'analyse des spots issus des gels 2D par spectrométrie de masse.

A la fin de la seconde dimension, les gels sont rincés à l'eau 2 x 3 min puis fixés dans une solution de méthanol (50%) et d'acide acétique (5%) pendant 20 min. Les gels sont rincés avec une solution de méthanol (50%) pendant 10 min puis à l'eau pendant une nuit. Les gels sont ensuite placés 1 min dans une solution de sensibilisation composée de thiosulfate de sodium (0,02% m/v, Sigma), lavés à l'eau 1 min et incubés 20 min dans une solution de nitrate d'argent (0,01% m/v, Sigma) à 4°C et sous agitation. Après un lavage à l'eau 2 x 1 min, les gels sont placés dans la solution de révélation (0,04% v/v formaldéhyde (Sigma) dans 2% m/v carbonate de sodium (Sigma)), sous agitation, jusqu'à apparition des spots. La réaction est arrêtée en plongeant les gels dans de l'acide acétique 5%. Les gels sont conservés dans une solution d'acide acétique 1%.

B.2.3.2. Coloration au bleu de Coomassie colloïdal

La coloration utilisée est adaptée de celle de Neuhoff et collaborateurs (Neuhoff et al., 1988) avec des modifications reportées par Matsui et al. (Matsui et al., 1999). Les gels sont plongés 20 min dans une solution de fixation composée de 45% méthanol et

1% acide acétique, puis dans une solution de coloration (méthanol 34%, sulfate d'ammonium 17% (m/v), acide acétique 0,5%, et bleu de Coomassie (CBB) G-250 (Serva, Heidelberg, Allemagne) 0,1% (m/v)) pendant 18 à 24 h. Les gels sont alors décolorés à l'eau chaude (55°C), en changeant l'eau toutes les 20 min, jusqu'à ce que le rapport entre l'intensité des spots et le bruit de fond soit satisfaisant et que l'équilibre avec l'eau de décoloration soit obtenu. Cette coloration est très compatible avec l'analyse par spectrométrie de masse.

B.3. Analyse qualitative et quantitative informatisée des gels

Les cartes bidimensionnelles des protéines de l'hémolymphe de drosophiles saines ou infectées expérimentalement sont comparées entre elles et quantifiées afin de rechercher d'éventuelles variations dans le contenu protéique en fonction du contexte d'infection. Cette étape est réalisée à l'aide du logiciel PDQuest v7.1. Ce logiciel convertit un signal provenant de données biologiques en données numériques et décompose chaque objet détecté selon le nombre de pixels qui le représente. PDQuest attribue à chaque pixel des coordonnées x, y et z. Les coordonnées x et y représentent la position du pixel sur un axe horizontal et vertical et la valeur z représente l'intensité du signal. Cette dernière valeur permet de quantifier un objet (nos spots) en groupant les pixels qui le composent. L'intensité de signal moyenne est calculée en additionnant les valeurs d'intensité de chaque pixel et en divisant le total par le nombre de pixels qui compose l'objet. Pour qu'un objet soit visible et quantifiable, il faut que l'intensité totale de cet objet soit supérieure à l'intensité des pixels qui représentent le bruit de fond de l'image.

Pour effectuer la comparaison qualitative et quantitative des gels scannés, l'analyse par le logiciel PDQuest peut être détaillée en plusieurs étapes dont certaines sont automatiques et d'autres doivent être validées et complétées par l'utilisateur.

B.3.1. Détection automatique des spots et nettoyage des images

Après avoir scanné les gels (GS-800 Calibrated Imaging Densitometer), il faut définir les paramètres qui vont permettre au programme de fixer un seuil de détection (spot de plus faible intensité, de plus petite et de plus grande taille ainsi qu'une zone représentative du bruit de fond). Après la détection automatique des spots de nos gels, une étape de "nettoyage" manuelle est alors nécessaire. En effet, des traînées, des particules d'argent, des bulles et des poussières sont également détectées comme des spots et il est nécessaire de les écarter de l'analyse. Les spots résultants de cette étape sont appelés "spots valides".

B.3.2. Appariement des spots (ou "matching")

Cette étape consiste à associer, sur les différents gels analysés, les spots identiques entre eux. Ceci est effectué automatiquement mais nécessite que l'utilisateur définisse des points d'ancrage (ou "landmarks") qui permettent au programme de fixer des points de repère et de compenser d'éventuelles déformations des gels. Pour un appariement correct, il faut indiquer manuellement au moins 10% de spots valides comme "landmarks", répartis sur l'ensemble du gel. Cependant, dans les zones contenant beaucoup de spots, il est parfois nécessaire d'en assigner une plus grande quantité ou de corriger manuellement certaines erreurs d'appariement du programme. Une fois que le logiciel a apparié la totalité des spots, l'analyse quantitative peut être réalisée. Pour cela, un des gels est désigné comme gel de référence afin de créer un gel "master" synthétique regroupant les données de tous les spots des gels de l'analyse.

B.3.3. Normalisation des gels et analyse des données

Une certaine variabilité dans la taille des spots ou leur intensité peut être due à des facteurs expérimentaux (perte d'échantillon au cours des préparations, mauvais transfert des protéines de la première vers la deuxième dimension, variation de temps de coloration des gels, température ambiante, etc.). Afin d'effectuer une comparaison fiable des taux d'abondance des protéines, il faut pouvoir compenser ces variations

indépendantes du contexte biologique. Pour cela, une normalisation est réalisée pour chaque expérience. Il existe plusieurs manières de normaliser les gels : par rapport à un groupe de spots dont on sait que l'intensité ne varie pas, par rapport à la densité totale du gel ou par rapport à l'ensemble des spots valides. Ne connaissant pas les spots invariants, nous avons choisi de normaliser par rapport aux spots valides.

La totalité des informations de chaque spot détecté (surface, densité, quantité) est regroupée sur le gel "master". Le logiciel peut alors mettre en évidence les spots dont l'intensité augmente ou diminue d'un gel à l'autre selon un facteur déterminé. L'ensemble de ces données peut être exporté sous forme de fichiers Excel pour être analysé en détail.

B.4. Identification des protéines par spectrométrie de masse

Les spots d'intérêt sont excisés des gels en utilisant un embout de pipette coupé comme emporte pièce. Un morceau de gel de taille similaire, mais ne contenant aucune protéine, est également excisé. Il sera traité dans les mêmes conditions que les échantillons et servira de blanc lors des analyses par spectrométrie de masse. Les spots excisés sont conservés à -20°C dans des microtubes.

L'identification d'un spot protéique issu d'une analyse protéomique sur la drosophile est facilitée par le fait que le génome de ce modèle est entièrement séquencé. Ainsi, l'empreinte trypsique de la protéine comparée à l'empreinte virtuelle contenue dans les banques de donnée peut suffire à l'identification.

B.4.1. Digestion des protéines et extraction des peptides générés

La préparation de l'échantillon est très importante pour l'identification des protéines par spectrométrie de masse. Il a été montré que la coloration à l'argent n'est pas la plus compatible pour l'analyse des fragments trypsiques par spectrométrie de masse. Le fait de décolorer les spots avant de les digérer élimine les problèmes dus aux restes de colorant et améliore la détection des signaux peptidiques (Gharahdaghi et al., 1999). J'ai donc utilisé cette méthode lors de toutes les identifications de protéines issues de spots colorés au nitrate d'argent. Les spots dont l'intensité est modulée sont

excisés du gel puis décolorés dans 50 μ L d'une solution composée du mélange 1 : 1 de ferricyanure de potassium (Sigma) 30 mM et de thiosulfate de sodium 100 mM. Les morceaux de gel sont agités jusqu'à la disparition totale de la coloration (1 à 2 min) puis rincés un minimum de cinq fois à l'eau MilliQ.

Les morceaux de gels décolorés ou ceux colorés au CBB sont immergés dans 50 μ L d'acétonitrile (ACN, Merck, Darmstadt, Allemagne) 50% et mis 15 min sous agitation. Cette opération est répétée une seconde fois et les surnageants sont éliminés.

Les morceaux de gels sont alors déshydratés dans 50 μ L d'ACN pur (5 min, sous agitation). Après élimination du surnageant, les échantillons sont réhydratés 5 min dans 50 μ L de tampon NH₄HCO₃ (Fluka, Buchs, Suisse) 25mM, pH 7,8 puis 50 μ L d'ACN pur sont ajoutés. Après 5 autres min sous agitation, les surnageants sont éliminés et les morceaux de gel séchés sous vide quelques minutes (Speed-vac, Savant).

Les échantillons sont alors réduits puis alkylés. La réduction se fait par incubation 45 min à 56°C en présence de 50 μ L de tampon NH₄HCO₃ 25 mM additionné de DTT 10 mM. Les surnageants sont éliminés puis les protéines sont alkylées : les morceaux de gel recouverts immédiatement avec 50 μ L de tampon NH₄HCO₃ 25 mM additionné d'iodoacétamide 55 mM. Les échantillons sont agités pendant 30 min à l'obscurité et les surnageants éliminés.

Les gels sont lavés deux fois par 50 μ L d'une solution comprenant 50% d'ACN et 50% de tampon NH₄HCO₃ 25 mM (sous agitation, 5 min). Un troisième lavage est réalisé avec de l'ACN pur. Après élimination des surnageants, ces échantillons déshydratés sont séchés sous vide.

La digestion "in gel" consiste à réhydrater les morceaux de gel μ L par μ L (6 μ L au total) avec une solution de trypsine (Promega, Madison, WI) à 15 ng/ μ L dans du tampon NH₄HCO₃ 25mM. Après absorption complète de cette solution, 2 μ L de tampon NH₄HCO₃ 25 mM sont ajoutés. Les échantillons sont incubés une nuit à 37°C.

Le lendemain, les peptides sont extraits par 4 μ L d'une solution composée de 70% d'ACN et 3% d'acide formique et les échantillons sont soniqués 10 min. Les peptides sont dessalés et concentrés sur microcône ZipTip C₁₈ (Millipore) puis élués dans 4 μ L d'ACN 50%.

B.4.2. Analyse des échantillons par spectrométrie de masse MALDI-TOF

L'analyse des fragments trypsiques se fait par spectrométrie de masse MALDI-TOF (Matrix Assisted Laser Desorption Ionisation - Time Of Flight), sur un BiflexIII (Bruker, Bremen, Allemagne). Cette technique, particulièrement sensible, permet de mesurer la masse de peptides présents dans un mélange à une concentration subpicomolaire. De plus, grâce à une calibration interne et dans certaines conditions d'analyse (mode réflectron), une précision de masse très élevée (< 50 ppm) peut être obtenue. Un tel niveau de précision est indispensable à l'utilisation de la spectrométrie de masse comme technique précise d'identification de protéines (Corthals et al., 2000).

Cette technique est basée sur l'analyse de l'échantillon co-cristallisé avec une matrice (acide α-cyano hydroxycinnamique) selon la préparation "sandwich" décrite par Kussman et collaborateurs (Kussmann et al., 1997). L'ensemble est bombardé par un rayon laser. La matrice, en excès par rapport à l'échantillon, le protège des radiations laser et isole les molécules à analyser les unes des autres, prévenant la formation de conglomérats. L'énergie du laser est absorbée par les co-cristaux et provoque le passage en phase gazeuse de la matrice et de l'analyte (Yates, 1998). Il y a un mécanisme de transfert de protons de la matrice vers les peptides donnant des ions préférentiellement monochargés ($[M+H]^+$). Les ions produits au niveau de la source sont ensuite accélérés par un champ électrique uniforme (19 kV) avec une extraction retardée (200 ns) puis traversent un tube de vol, libre de champ, où ils volent jusqu'à l'analyseur selon une vitesse inversement proportionnelle au rapport masse sur charge (les ions les plus chargés et les plus légers atteignent le détecteur les premiers). Une calibration interne est effectuée en utilisant les fragments d'autolyse de la trypsine $([M+H]^+ a) m/z 842,510$, 1045,546 et 2211,104). L'analyse par spectrométrie de masse MALDI-TOF permet ainsi d'obtenir une liste de masses des différents fragments de la protéine obtenus après la digestion trypsique.

B.4.3. Identification des protéines par recherche dans les banques de données

Dans la mesure où le génome de la drosophile est connu, l'analyse des masses des fragments trypsiques dans des banques de données protéiques et génomiques est facilitée. Ainsi, les valeurs de masse mesurées sont comparées, en utilisant les outils informatiques disponibles sur internet (essentiellement Protein Prospector: prospector.ucsf.edu et Matrix Science: www.matrixscience.com), avec les valeurs calculées contenues dans les banques de données spécialisées. Cette comparaison est associée à une valeur (ou "score") fondée sur le nombre de fragments protéolytiques obtenus dont les masses peuvent correspondre à celles des fragments protéolytiques des banques de données, et par la précision avec laquelle les masses sont déterminées. Les meilleurs scores correspondent aux comparaisons les plus fiables. On peut ainsi caractériser une protéine connue (soit identifiée, soit répertoriée comme produit de gène) uniquement grâce à une mesure de masses. De plus, les logiciels de recherche utilisés permettent l'identification de modifications chimiques (carbamidométhylation par exemple) ou post-traductionnelles (phosphorylation par exemple) sur les peptides analysés.

Les étapes de la stratégie d'identification des protéines issues des gels 2D sont récapitulées sur la Figure 8.

C. Approche par marquage isotopique des protéines

Cette méthode est basée sur le marquage des protéines de nos échantillons par un réactif isotopique ICAT. Un marqueur portant l'isotope léger est accroché sur les protéines d'un premier échantillon (hémolymphe contrôle), tandis que le second (hémolymphe après infection) est marqué par l'isotope lourd. Les deux échantillons sont alors mélangés et traités simultanément jusqu'à une analyse chromatographique couplée à la spectrométrie de masse en tandem (LC-MS/MS). Le rapport d'intensité des ions porteurs du marqueur léger/lourd permet de calculer les concentrations relatives d'une protéine donnée dans les deux échantillons.

Cette approche a été réalisée en collaboration avec le laboratoire du Dr Alan Aderem à l'Institute for Systems Biology (ISB, Seattle). Les expériences décrites ici ont été réalisées par Samuel Donohoe à l'ISB sur des échantillons prélevés à Strasbourg.



Figure 8 : Stratégie d'identification des protéines contenues dans les spots des gels 2D. Le spot d'intérêt est excisé du gel puis digéré par la trypsine. Les fragments peptidiques sont extraits, leurs masses sont mesurées par spectrométrie de masse et utilisées pour les recherches dans les banques de données.

Le réactif ICAT utilisé est un réactif de deuxième génération composé de quatre parties (Figure 9) :

- un groupe réactif (iodoacétamide) qui permet de lier de façon covalente le marqueur à la protéine en alkylant les cystéines,

- une chaîne $C_{10}H_{17}N_3O_3$ qui permet le marquage isotopique en substituant neuf carbones ¹²C par du ¹³C,

- une étiquette biotine qui va permettre de sélectionner les peptides marqués par affinité de la biotine pour l'avidine,

- un site de clivage acide qui permet d'éliminer une partie du marqueur (celle contenant l'étiquette biotine) avant l'analyse par LC-MS/MS.

Cette dernière étape résulte en l'amélioration de l'analyse quantitative par rapport aux réactifs de première génération, notamment grâce à une co-élution des peptides marqués par l'isotope lourd ou léger lors de la chromatographie alors que l'étiquette biotine créait un léger décalage d'élution.

C.1. Marquage et digestion des protéines

L'hémolymphe des drosophiles contrôles ou infectées (environ 50 μ g) est prélevée dans 250 μ L de solution tampon phosphate salin (PBS pour Phosphate Buffered Saline) pH 7,2. Une ultrafiltration sur Microcon 3000 (Millipore) est réalisée jusqu'à réduction du volume à 10 μ L. Les échantillons sont dilués et stockés dans 100 μ L d'une solution composée de 50 mM Tris pH 8,5 et 0,1% SDS. Avant marquage, la solution est ajustée à 6 M d'urée, 5 mM d'acide éthylènediaminetétraacétique (EDTA), 200 mM Tris pH 8,3, 0,05% SDS et de l'eau milliQ qsp 160 μ L (tampon de marquage ICAT). Une étape de réduction des ponts disulfure est effectuée en ajoutant de l'hydrochlorure de Tris (2-CarboxyEthyle) Phosphine (TCEP, Pierce, Rockford, IL) à une concentration finale de 5 mM et en incubant les échantillons 30 min à 37°C. Le marquage des échantillons par le réactif ICAT est effectué selon les instructions du fournisseur (kit de réactifs ICAT-clivable pour marquage protéique, Applied Biosystems, Foster City, CA). 1,225 nmol de marqueurs ICAT légers ou lourds sont ajoutés à la solution d'échantillons contrôles ou infectés, respectivement (volume final 200 μ L). La réaction de marquage est réalisée pendant 1 h à 37°C puis stoppée par



Figure 9 : Structure schématique du réactif clivable ICAT. (Voir texte pour les détails).

l'ajout de DTT en excès à une concentration finale de 10 mM. Les échantillons marqués sont dilués au 1/5 avec de l'eau milliQ (volume final 1 mL) puis mélangés dans un même tube. 1% de trypsine (m/v) sont ajoutés et la digestion est effectuée une nuit à 37°C.

C.2. Prépurification des peptides sur cartouche échangeuse d'ions

L'échantillon est injecté sur une colonne Polysulfoethyl Aspartamide (Polysulfoethyl A^{TM} , 2,1 mm x 200 mm, PolyLC, Columbia, MD). La cartouche est tout d'abord conditionnée par 2 mL de solution A composée de KH₂PO₄ 5 mM et d'ACN 25%, pH 2,7. Le même volume de solution A est ajouté aux échantillons puis ils sont injectés doucement dans la cartouche (1 goutte/seconde). La cartouche est lavée par 1 mL de solution A puis les peptides sont élués par 1 mL de solution B (solution A additionnée de KCl 600 mM). La solution récupérée est entièrement séchée au Speedvac.

C.3. Chromatographie d'affinité et coupure de l'étiquette biotine

Les peptides biotinylés sont purifiés grâce à une chromatographie d'affinité pour l'avidine.

La cartouche d'avidine (4 mm x 15 mm, Applied Biosystems) est activée par injection de 2 mL de PBS (2X), pH 7,2. L'échantillon séché est dissout et neutralisé par 2 mL de PBS (2X). L'échantillon est injecté doucement (1 goutte/seconde) dans la cartouche d'avidine et la solution récupérée est réinjectée une seconde fois. La cartouche est alors lavée successivement avec 1 mL de PBS (1X), 1 mL de NH₄HCO₃ 50 mM dans 20% méthanol puis 1 mL d'eau. Les peptides biotinylés sont élués de la cartouche par 1 mL d'ACN 30% et TFA 0,4%. La solution peptidique récupérée est séchée complètement au Speed-vac.

La coupure de l'étiquette biotine du marqueur ICAT est réalisée suivant les instructions du fournisseur (Applied Biosystems) 2 h à 37°C en diluant l'échantillon dans 5 μ L de solution de clivage du kit. L'échantillon est de nouveau asséché entièrement au Speed-vac.

C.4. Fractionnement des échantillons par chromatographie échangeuse d'ions

La cartouche (PolysulfoethylATM, 2,1 mm x 200 mm, PolyLC) est équilibrée par injection de 1 mL d'acide acétique 0,5% et ACN 5%. L'échantillon est dissout dans 100 μ L d'ACN 5% et injecté sur la colonne. La solution récupérée est injectée une seconde fois. L'échantillon est lavé par 1 mL d'ACN 5%. L'échantillon est fractionné en injectant successivement 100 μ L d'une solution d'ACN 5% additionné de 20 à 500 mM d'acétate d'ammonium par paliers puis d'ACN 30% additionné de 500 mM d'acétate d'ammonium.

Les différentes fractions collectées sont évaporées au Speed-vac, dissoutes dans 6 μ L d'acide acétique 0,2% et transférées dans des tubes de chargement automatique pour l'analyse par LC-MS/MS.

C.5. Chromatographie liquide couplée à la spectrométrie de masse en tandem

L'analyse par LC-MS/MS est réalisée sur un spectromètre de masse à trappe ionique LCQ DECA (Thermo Finnigan, San Jose, CA). Le solvant est délivré grâce à une pompe HP1100 (Agilent, Palo Alto, CA) et les échantillons sont délivrés automatiquement par un système autosampler Famos (LC Packings, San Francisco, CA) dans une précolonne capillaire (2 cm x 100 μ m, Michrom Bioresources, Auburn, CA), constituée de silice fondue et de résine Magic C₁₈AQTM. Les échantillons sont lavés sur la précolonne avec un premier solvant composé d'acide formique 0,1% et ACN 5% puis élués avec un gradient de 10-35% d'ACN dans l'acide formique 0,1% en 128,5 min sur une colonne capillaire (14 cm x 75 μ m), constituée de silice fondue et de résine Magic C₁₈AQTM, connectée à la source d'ionisation (μ ESI, microElectroSpray Ionization) comme décrit par Yi et collaborateurs (Yi et al., 2003). Les échantillons sont alors délivrés dans le spectromètre de masse avec un flux d'approximativement 300 nL/min et analysés par MS/MS comme décrit par Gygi et al. (Gygi et al., 1999).

C.6. Analyse des résultats

Les spectres obtenus sont comparés de manière automatique avec les banques de données grâce à SEQUEST (Thermo Finnigan) et la quantification relative entre les peptides marqués par ICAT légers ou lourds est effectuée par le programme ASAPRatio développé par Li et collaborateurs (Li et al., 2003). Le programme PeptideProphet est utilisé pour valider les identifications (Keller et al., 2002), les résultats sont visualisés et filtrés grâce au programme INTERACT (Han et al., 2001) et les fichiers de données obtenus sont groupés et analysés statistiquement grâce au programme ProteinProphet (Nesvizhskii et al., 2003). Les données peuvent être exportées sous forme de fichiers Excel pour être analysées.

D. La RT-PCR quantitative en temps réel

La RT-PCR (Reverse Transcrition-Polymerase Chain Reaction) quantitative, basée sur la PCR classique, permet de suivre en temps réel l'évolution de la quantité d'ADN synthétisé au cours des cycles, grâce à une sonde spécifique du gène d'intérêt couplée à un fluorophore. La synthèse d'ADN évolue selon une courbe de type exponentiel avec un doublement de la quantité d'ADN produite dans chaque échantillon à chaque cycle. En mesurant la quantité d'ADN produite dès le seuil de sensibilité, il est possible de calculer la quantité initiale d'ADN complémentaires des gènes d'intérêt. Cette mesure est possible grâce aux molécules de SYBRGreen, qui émettent une forte fluorescence en se fixant sur les doubles brins d'ADN. Cette technique possède l'avantage de nécessiter peu de drosophiles en comparaison de la quantité requise pour les Northern blots.

Les PCR quantitatives sont réalisées au laboratoire par Marie-Céline Lafarge.

D.1. Extraction des ARN totaux

L'extraction est réalisée en plaque de 96 puits (Qiagen, Hilden, Allemagne) contenant chacun cinq drosophiles et une bille de tungstène. Les plaques sont stockées à -80°C jusqu'à leur utilisation. Les échantillons sont broyés à l'aide d'un Mixer Mill 300MM (Retch, Haan, Allemagne) 2 x 1 min à 25 Hz. Une extraction d'ARN est alors effectuée grâce au kit NucleoSpin 96 RNA (Macherey-Nagel, Duren, Allemagne) selon les instruction du fournisseur. Le broyat est resuspendu dans 350 μ L d'une solution de lyse (tampon RA du kit / 1% β -mercaptoethanol). Cette solution est riche en sel de guanidine qui inhibe les ARNase et en phénol qui retient les protéines. Puis, une solution favorisant l'adhésion de l'ARN à la membrane de silice est ajoutée au mélange. L'ADN contaminant, ayant adhéré en même temps à la membrane, est dégradé par l'ajout d'une ADNase. Plusieurs étapes de lavages servent ensuite à éliminer les sels, les métabolites, et les débris cellulaires restants. L'ARN est élué dans 100 μ L d'eau MilliQ et stocké à –80°C.

D.2. Transcription inverse

La transcription inverse est réalisée sur 2 μ g d'ARN en suivant les instructions du Kit SuperScript II RNase H reverse Transcriptase (Invitrogen, Carlsbad, CA). Le mélange est composé du tampon de transcription inverse (1X), de dNTPs (0,5mM), d'héxamères (0,25 μ g), de DTT (10mM), de la RNaseOUT Recombinant Ribonuclease Inhibitor (2 U), de transcriptase reverse SuperScript (10 U). La réaction se déroule pendant 10 min à 25°C, 60 min à 42°C et 15 min à 70°C.

D.3. PCR quantitative en temps réel

La PCR est réalisée sur les ADN complémentaires obtenus en utilisant le kit qPCR SybrGreen I-No ROX (Eurogentec). Des amorces spécifiques pour les gènes d'intérêts, ainsi que pour le gène codant la protéine ribosomale 49 (RP49) qui sert de référence, ont été dessinées grâce à un outil disponible en ligne (Primer3, http://frodo.wi.mit.edu) et fabriquées par la société Sigma. Les séquences des amorces utilisées dans cette étude sont reportées en Annexe 1.

La réaction est réalisée grâce à un i-Cycler (BioRad) sur des plaques de 96 puits dans lesquels 2 μ L d'ADN complémentaires de chaque échantillon sont déposés dans un mélange composé des amorces (0,3 μ M), de tampon réactionnel (1 X), de MgCl₂

(5mM), de dNTPs (0,2mM), de fluorescéine (1 nM), de SYBRGreen (solution diluée au 1/75000), d'enzyme Hot Gold Star (0,625U), et d'eau qsp 25 µL. Les échantillons sont déposés au moins en duplicats. Les fragments amplifiés lors de la PCR sont détectés grâce au SYBRGreen (excitation : 494 nm ; émission : 530 nm) ajouté dans le milieu réactionnel. La lecture de la fluorescence émise par le SYBRGreen se fait à la fin de chaque cycle d'élongation. La PCR est effectuée selon le programme suivant : 9 min à 95°C, 15 s à 95°C suivi d'une min à 60°C (40 cycles). La spécificité de la réaction est vérifiée en mesurant la température de fusion du produit en augmentant progressivement par paliers de 0,5°C de 65°C à 95°C. La quantification se fait par rapport à une courbe étalon obtenue avec des dilutions d'ADN de concentration connue. Le logiciel détermine par interpolation le cycle seuil Ct à partir duquel le signal d'amplification est significatif par rapport au bruit de fond. Ce cycle seuil est directement proportionnel à la quantité initiale d'ADN et à l'efficacité de la PCR (selon la formule : $Q_n = Q_0 (1+e)^{Ct}$, où Q_n est la quantité d'ADN au cycle Ct donnant un signal significatif, Q₀ la quantité initiale d'ADN, et e l'efficacité de la PCR). Les efficacités de PCR sont proches de 100% (déterminées sur les courbes standards) et toujours supérieures à 85% pour que la PCR soit validée.

E. Inactivation de gène par injection d'ARN double brin

Cette technique est basée sur un processus post-transcriptionnel, déclenché par l'introduction d'ARN double brin, qui mène à l'inactivation d'un gène d'une manière séquence-spécifique. Cette technique a été mise au point sur la drosophile par Akira Goto au laboratoire dans l'équipe de Jean-Marc Reichhart (Goto et al., 2003).

Des amorces spécifiques des gènes d'intérêts à éteindre sont dessinées grâce à un outil disponible en ligne (Primer3, http://frodo.wi.mit.edu) auxquels est ajoutée, à l'extrémité 5', la séquence promotrice de la polymérase T7 (cf. Annexe 1 pour les séquences). Les amorces sont réalisées par la société Sigma. Les fragments d'ADN correspondant à chaque gène sont amplifiés par une PCR standard sur de l'ADN complémentaire de drosophile. La réaction se déroule pendant 5 min à 95°C, 30 s à 58°C, 45 s à 72°C et 30 s à 95°C. L'élongation finale, après 35 cycles, se fait à 72°C pendant 5 min. Après avoir vérifié l'efficacité de la PCR sur gel d'agarose, les ARNs simple brin sont synthétisés grâce au kit de transcription MEGAscript T7 (Ambion, Austin, TX). Pour cela, 5 μ L des produits de PCR sont utilisés pour la transcription *in vitro* selon les instructions du fournisseur. Cette étape est effectuée en quatre exemplaires (4 x 5 μ L de produits de PCR). Après incubation une nuit à 37°C, l'ADN ayant servi de matrice est éliminé grâce à 1 μ L d'ADNase (sans ARNase). Les quatre tubes sont regroupés et une précipitation à l'isopropanol est réalisée. Le culot est lavé par 400 μ L d'éthanol absolu et suspendu dans 100 μ L de solution TE du kit. Les ARNs simple brin complémentaires sont alors hybridés en plaçant les tubes dans un grand volume d'eau à 95°C et en laissant refroidir lentement. Une extraction au phénol/chloroforme est réalisée suivie d'une précipitation à l'éthanol/acétate d'ammonium. Les ARNs double brin sont suspendus dans du tampon d'injection comprenant 0,1 mM de phosphate de sodium, pH 6,8 et 5 mM de KCl. La concentration en ARN est contrôlée par mesure de l'absorbance à 260 nm et la concentration est ajustée à 3 mg/mL.

32,2 nL de solution d'ARN double brin sont injectés dans le thorax de chaque drosophile, âgée d'au moins 5 jours, à l'aide d'un nanoinjecteur (Drummond). Après 24 heures de récupération, les mouches peuvent être infectées expérimentalement.

RESULTATS

RESULTATS

Chapitre 1 : Recherche de nouvelles molécules impliquées dans la réponse immunitaire de la drosophile : approche par gels d'électrophorèse bidimensionnelle

A. Mise en place de la méthodologie

A.1. Electrophorèse bidimensionnelle

La mise en place d'une approche par électrophorèse bidimensionnelle (2D) est différente pour chaque type d'échantillon étudié. Lorsque je suis arrivée au laboratoire, peu d'études par électrophorèse 2D avaient été réalisées sur la drosophile (Holzl et al., 2000; Santaren, 1996; Santaren et al., 1998) et ces analyses avaient été faites sur des larves. De nombreuses mises au point ont été nécessaires afin de déterminer les meilleures conditions expérimentales pour l'étude de l'hémolymphe de drosophiles adultes.

A.1.1. Optimisation de la solution de solubilisation des protéines et de la gamme de pH

La préparation de l'échantillon est une étape cruciale pour la qualité des cartes 2D. Il est donc essentiel d'optimiser les conditions de solubilisation nécessaires à l'analyse de protéines contenues dans l'hémolymphe de drosophiles par électrophorèse 2D. De nombreuses conditions ont été testées sur des gels 2D de petits formats (8 x 8 cm) à partir d'un mélange en quantité équivalente d'hémolymphe de drosophiles mâles et femelles non infectées (15 μ g de protéines, soit l'hémolymphe d'environ 30 individus).

Les premières expériences, basées sur les recommandations des fournisseurs (plus particulièrement BioRad et Amersham Biosciences) ont été réalisées sur des capillaires par Laurence Sabatier avant mon arrivée au laboratoire. L'efficacité de deux détergents a été étudiée en les introduisant dans la solution de solubilisation l'un après l'autre, puis en combinaison, à différentes concentrations (0,5 à 2% de Triton X-100 ou 2 à 4% de CHAPS). Un mélange composé de 0,5% Triton X-100 et de 4% CHAPS, s'est révélé être le plus performant pour solubiliser les protéines de l'hémolymphe de drosophile. Il a également été montré qu'il est préférable de prélever l'hémolymphe directement dans la solution de solubilisation et à froid.

A mon arrivée au laboratoire, nous avons décidé de continuer les expériences sur des bandelettes à gradient de pH immobilisé (IPG). L'apparition sur le marché de ces bandelettes a permis de s'affranchir des problèmes de déplacement de pH que l'on pouvait rencontrer au cours de la migration lors de l'utilisation des capillaires. Ces bandelettes, qui existent en plusieurs gammes de pH, ont largement contribué à augmenter la reproductibilité des analyses (Gorg et al., 2000). Les premières cartes ont été obtenues grâce à des bandelettes couvrant une large gamme de pH (3-10). J'ai ainsi montré que nous obtenions des cartes assez bien résolues et mis en évidence que la quasi totalité des spots avait des pI supérieurs à 5 (Figure 10, A). Une gamme plus restreinte (pH 5-8) a alors été utilisée afin d'améliorer la résolution dans cette zone (Figure 10, B).

Par ailleurs, j'ai testé l'efficacité d'un mélange urée 7 M / thiourée 2 M à la place des 8 M d'urée. Il est considéré que ces deux agents chaotropes, en association, possèdent un meilleur pouvoir de solubilisation que l'urée seule, particulièrement sur les molécules hydrophobes telles que les protéines membranaires (Rabilloud et Chevallet, 2000). La plupart des études décrites dans la littérature montre une amélioration de la solubilisation des protéines des échantillons en présence de thiourée, cependant dans mon cas, la thiourée est à l'origine d'une dégradation importante du profil des cartes (non illustré). Une étude réalisée par Galvani et collaborateurs (Galvani et al., 2001) explique que l'usage de thiourée dans la solution de solubilisation peut résulter en une mauvaise alkylation des protéines. Ils ont en effet constaté une réactivité importante de l'iodoacétamide, agent utilisé le plus couramment pour l'alkylation des protéines, avec la thiourée, aboutissant à la formation d'un réactif intermédiaire empêchant une alkylation efficace des molécules de l'échantillon. Une mauvaise alkylation peut se traduire par la formation de traînées sur les gels 2D. Nous ne savons pas si ce phénomène est en cause dans nos échantillons, mais l'utilisation de la thiourée n'a pas été poursuivie.



Figure 10 : Comparaison des gels 2D de gammes de pH 3-10 et 5-8. L'hémolymphe de 30 drosophiles (15 mâles et 15 femelles) a été analysée par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 3-10 (A) ou 5-8 (B), la seconde dimension sur un gel SDS-PAGE 12%. Les gels ont été colorés au nitrate d'argent. On constate que la majorité des protéines ont un pI supérieur à 5.

J'ai également tenté de substituer le DTT (20 mM) par de la tributylphosphine (TBP 5 mM) dans la solution de solubilisation. Le TBP est décrit dans la littérature comme alternative au DTT qui possède l'inconvénient de s'ioniser à pH basique et donc d'être rapidement perdu lors de la focalisation isoélectrique. De plus, le TBP participe à l'amélioration de la solubilisation des protéines (Herbert et al., 1998). Une fois encore, la résolution des cartes a été altérée par ce changement (non illustré).

J'ai enfin étudié l'efficacité d'un kit de solubilisation des protéines (ReadyPrep Protein Extraction Kit (Soluble/Insoluble), Biorad). Ce kit est basé sur une stratégie de fractionnement différentiel des molécules de l'échantillon selon leur solubilité. Il permet d'étudier séparément les protéines hydrophiles (solubles) et les protéines hydrophobes (insolubles). La concentration de chacune des fractions peut, par la même occasion, être augmentée, améliorant ainsi l'analyse des protéines présentes en faible abondance. De plus, la solution de solubilisation des protéines du kit contient de l'amidosulfobétaine-14 (ASB-14) qui est un détergent zwitterionique puissant. Cependant, outre la perte d'échantillon occasionnée par les étapes de précipitation des protéines dans la séparation des fractions soluble et insoluble, la résolution des cartes a de nouveau été diminuée. Ceci semble dû à un manque de compatibilité entre nos échantillons et la solution de solubilisation du kit. En effet, en utilisant notre tampon de solubilisation à la place de celui du kit, nous obtenons une meilleure résolution (Figure 11). Il est intéressant de noter que la fraction insoluble ne contenait aucune protéine visible sur gel 2D (non illustré) indiquant que nos échantillons ne possèdent probablement que très peu de protéines hydrophobes.

L'ensemble de mes résultats illustre bien l'importance de la composition de la solution de solubilisation pour la qualité des cartes 2D et la spécificité de cette solution selon l'échantillon étudié.

A.1.2. Choix du sexe des individus

Le profil d'une carte d'hémolymphe de drosophiles mâles a été comparé avec celui de femelles. Une cartographie sur 30 individus de chaque sexe a été réalisée (Figure 12). La différence de profil est frappante et de nombreux spots supplémentaires



Figure 11 : Cartes 2D de la fraction soluble des protéines de l'hémolymphe de drosophile. Les protéines solubles de l'hémolymphe de dosophiles mâles ont été isolées grâce au kit de solubilisation (Soluble/Insoluble) de BioRad. 20 µg de protéines, solubilisées soit dans la solution de solubilisation de notre composition (A), soit dans celle du kit (B), ont été analysés par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 5-8, la seconde dimension sur un gel SDS-PAGE 11%. Les gels ont été colorés au nitrate d'argent.



Figure 12 : Cartes 2D de l'hémolymphe de drosophile mâles et femelles. L'hémolymphe provenant de 30 drosophiles femelles (A) ou mâles (B), soit environ 15 μ g de protéines, a été analysée par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 5-8, la seconde dimension sur un gel SDS-PAGE 11%. Les gels ont été colorés au nitrate d'argent.
(une trentaine environ) sont présents uniquement sur la carte des femelles. A partir de cette étude, j'ai décidé de ne travailler qu'avec des drosophiles mâles afin de m'affranchir des ces protéines additionnelles, impliquées pour la plupart dans l'ovogenèse, pouvant masquer d'autres spots d'intérêt.

L'ensemble des mises au point sur gels de petite dimension a permis d'obtenir, à partir de l'hémolymphe de 30 drosophiles mâles contrôles (15 μ g de protéines), des cartes reproductibles sur lesquelles environ 130 spots peuvent être visualisés après coloration au nitrate d'argent (Figure 13).

A.1.3. Cartes de grande dimension

Les gels de petite dimension (8 x 8 cm) nécessitent peu d'échantillon (15 μ g de protéines) et sont assez rapides à obtenir, ce qui en fait un outil de choix pour effectuer des expériences de mise au point. Cependant la résolution obtenue sur ces gels n'est pas suffisante pour effectuer une analyse qualitative précise. Afin d'augmenter la résolution des cartes et analyser les protéines présentes en plus faible quantité, nous avons étendu l'analyse aux gels de plus grande dimension (20 x 20 cm) sur lesquels de 100 à 300 μ g de protéines, selon le mode de coloration, peuvent être analysés. Après avoir ajusté les conditions de focalisation (8 h à 10000 V au lieu de 5 h à 4000 V), nous avons obtenu des cartes à partir de l'hémolymphe d'environ 200 drosophiles (100 μ g de protéines), sur lesquelles environ 350 spots peuvent être visualisés après coloration à l'argent (Figure 14).

A.2. Mise en place de l'analyse différentielle

L'analyse différentielle consiste à comparer les protéines contenues dans l'hémolymphe de drosophiles avant et après une infection microbienne. Trois inducteurs ont été utilisés lors de ces expériences, une bactérie à Gram-négatif (*Escherichia coli*), une bactérie à Gram-positif (*Micrococcus luteus*) et un champignon (*Beauveria bassiana*).



Figure 13 : Carte 2D de petite dimension de l'hémolymphe de drosophiles contrôles. L'hémolymphe d'environ 30 drosophiles mâles (15 µg de protéines) a été analysée par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 7 cm pH 5-8, la seconde dimension sur un gel SDS-PAGE 12%. Après coloration au nitrate d'argent, on peut visualiser environ 130 spots sur cette carte.



Figure 14 : Carte 2D de grande dimension de l'hémolymphe de drosophiles contrôles. 100 µg de protéines provenant de l'hémolymphe d'environ 200 drosophiles mâles ont été analysés par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 17 cm pH 5-8, la seconde dimension sur un gel SDS-PAGE 10%. Après coloration au nitrate d'argent, on peut visualiser environ 350 spots sur cette carte.

A.2.1. Etude cinétique : choix du temps d'infection

A.2.1.1. Infections bactériennes

Les premières expériences ont été réalisées en inoculant les mouches avec un mélange bactérien, composé d'une bactérie à Gram-positif (*M. luteus*) et d'une bactérie à Gram-négatif (*E. coli*), connu pour déclencher de manière globale et efficace la réponse immunitaire de la drosophile (Uttenweiler-Joseph et al., 1998). La plupart des analyses réalisées au laboratoire, notamment lors de l'étude des peptides antimicrobiens et des DIMs ont été réalisées 6 h après les infections bactériennes. Cependant, certaines protéines impliquées dans la réponse immunitaire de la drosophile font partie de cascades protéolytiques et interviennent donc très tôt après l'infection. Elles peuvent aussi être dégradées rapidement. C'est pourquoi il était important d'étudier des temps très courts. L'étude différentielle a donc été menée selon une cinétique d'induction comprise entre 30 s et 6 h. Chaque expérience a été reproduite de deux à cinq fois sur des gels de petite dimension.

Les cartes obtenues sur les temps courts (30 s à 15 min) n'ont montré aucune variation significative dans l'intensité des spots. En augmentant les temps d'induction à 1 h puis 6 h, nous avons noté que certains spots apparaissent. Les variations maximales ont été obtenues 6 h après infection. Les résultats obtenus sur gels 2D sont en accord avec l'apparition de la majorité des DIMs et peptides antimicrobiens, dont le maximum d'induction est obtenu six heures après l'infection (L. Sabatier, communication personnelle). Nous avons décidé de garder 6 h comme temps d'induction pour les piqûres septiques.

A.2.1.2. Infections fongiques

Pour les champignons, le temps d'infection avant la collecte de l'hémolymphe doit être supérieur à celui défini pour les bactéries car, après l'infection naturelle, il faut du temps pour que les spores de champignon germent et pénètrent dans l'hémocoele de la drosophile. Une étude cinétique sur gels 2D a montré qu'en dessous de 72 heures après l'infection, très peu de variations étaient observées. Ceci est en accord avec le faible niveau d'expression des peptides antimicrobiens observé pour des temps d'infection fongique inférieurs à 72 heures (L. Sabatier, communication personnelle). Nous avons décidé de garder 72 h comme temps d'induction pour les infections fongiques.

A.3. Etude de la reproductibilité des cartes 2D

Afin de déterminer le facteur de confiance à utiliser lors de l'analyse différentielle quantitative, la reproductibilité des cartes contrôles a été étudiée à partir de trois expériences indépendantes. J'ai ainsi observé, après analyse des cartes grâce au logiciel PDQuest, qu'en fixant le facteur de variation à deux, l'intensité de 80% des spots était invariable sur les trois cartes contrôles. En augmentant ce facteur à cinq, 98% des spots observés étaient invariables. Ainsi, en sélectionnant un facteur de variation important, les risques de fausses interprétations sont minimisés, les variations observées ont une très forte probabilité d'être dues aux traitements expérimentaux et non à une variabilité au sein des gels. Nous avons décidé, pour nos analyses quantitatives, que seules les variations d'au moins un facteur cinq étaient à considérer.

B. Analyses différentielles des protéines de l'hémolymphe de drosophiles

B.1. Coloration au nitrate d'argent

Les premières analyses différentielles ont été réalisées sur des gels de grande dimension, à partir de l'hémolymphe de drosophiles contrôles ou infectées par un mélange bactérien. Cette étude a révélé que l'intensité d'une cinquantaine de spots variait. Cette première approche différentielle sur grands gels colorés au nitrate d'argent, nous a conforté dans la faisabilité d'une telle analyse, en terme de résolution et de sensibilité.

Afin d'étudier la spécificité des variations d'abondance des protéines, les infections ont ensuite été réalisées avec un inducteur à la fois. Les drosophiles ont été infectées par une bactérie à Gram-positif (*M. luteus*), une bactérie à Gram-négatif (*E.*



Figure 15 : Cartes 2D colorées au nitrate d'argent des protéines de l'hémolymphe de drosophiles contrôles ou infectées par *B. bassiana*. 100 μ g de protéines (hémolymphe d'environ 200 drosophiles mâles) provenant de drosophiles contrôles (A) ou infectées par *Beauveria bassiana* (B) ont été analysés sur gels 2D. La première dimension a été réalisée sur bandelette IPG pH 5-8, la deuxième dimension sur un gel SDS-PAGE 11%. Les cartes ont été colorées au nitrate d'argent et analysées par PDQuest. Le résultat des analyses quantitatives est reporté dans le tableau II. Les spots annotés sont des exemples de variation d'intensité (i : inductions ; + : surexpression ; s : suppression ; - : répression). La zone encadrée correspond à une zone riche en spots régulés et a été étudiée en détail (cf. B.2).

	B. bassiana	M. luteus	E. coli
Induits	23	22	14
Surexprimés	15	7	8
Total des spots régulés positivement	38	29	22
Supprimés	31	6	10
Réprimés	10	18	10
Total des spots régulés négativement	41	24	20

Tableau II : Récapitulatif du nombre de spots régulés d'au moins un facteur cinq après infection. Résultat de l'analyse quantitative réalisée par le logiciel PDQuest sur les gels 2D colorés au nitrate d'argent obtenus à partir de l'hémolymphe de drosophiles 72 h après infection naturelle par le champignon *B. bassiana* ou 6 h après inoculation expérimentale par une bactérie à Gram-positif (*M. luteus*) ou Gram-négatif (*E. coli*). *coli*) et un champignon filamenteux (*B. bassiana*). La Figure 15 montre en exemple, après coloration à l'argent, les cartes 2D des protéines de l'hémolymphe de drosophiles avant et après infection fongique. Nous observons sur cette carte que l'intensité de nombreux spots varie ce qui a été confirmé après analyse par le logiciel PDQuest. Le tableau II récapitule le nombre de spots dont l'intensité varie d'au moins un facteur cinq avant et après chaque type d'infection. Nous constatons que l'infection fongique semble induire beaucoup plus de changements que les infections bactériennes.

B.2. Gammes étroites de pH

Sur la figure précédente, nous pouvons remarquer qu'une zone est particulièrement dense en variations (Figure 15, zone entourée). Pour étudier plus précisément cette région, j'ai effectué une analyse par gel 2D avec une gamme restreinte de pH, ciblant plus particulièrement cette zone d'intérêt (pH 5,5-6,7). L'utilisation de gammes étroites permet d'augmenter la quantité d'échantillon analysée, d'affiner la résolution et de détecter des spots dont l'intensité était trop faible pour être visualisés sur les cartes 5-8. Une analyse sur gamme étroite a été réalisée à partir de 200 μ g d'hémolymphe de drosophiles avant et après infection par *B. bassiana*. La Figure 16 montre cette carte sur laquelle on peut dénombrer une cinquantaine de spots supplémentaires dans les échantillons contrôles par rapport à cette même zone sur des gels 5-8. L'analyse par le logiciel PDQuest indique que plus de 60 spots varient d'au moins un facteur cinq. Ces spots ont été excisés du gel et analysés par spectrométrie de masse. Parmi le grand nombre de spots induits ou surexprimés après l'infection fongique, certains seront mentionnés plus tard et ont été annotés sur la figure.

B.3. Coloration au bleu de Coomassie colloïdal

Les spots régulés provenant de nos différents gels colorés à l'argent ont été excisés du gel et analysés selon l'approche classique par cartographie massique des produits de digestion trypsique et recherche dans les banques de données. Malheureusement, seule une faible partie des spots analysés a été identifiée (15%). En effet, autant la coloration au nitrate d'argent est sensible et permet la détection d'un



Figure 16 : Cartes 2D de gamme étroite de pH. L'hémolymphe d'environ 400 drosophiles mâles (200 µg de protéines) avant (A) et après (B) infection par le champignon *Beauveria bassiana* a été analysée par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 5,5-6,7, la seconde dimension sur un gel SDS-PAGE 11%. Les gels ont été colorés au nitrate d'argent. L'étude qualitative et quantitative par le logiciel PDQuest a montré que l'intensité d'environ 60 spots varie suite à l'infection. L'analyse par spectrométrie de masse des protéines correspondantes a permis d'identifier, entre autres, la transferrine (spots marqués d'une flèche), neuf fragments de transferrine (spots entourés) ainsi que les protéines correspondant aux spots A à F (cf. tableau III pour les caractéristiques de ces molécules).

grand nombre de spots, autant l'identification des protéines contenues dans ces spots est difficile par manque de signal lors des analyses par spectrométrie de masse. C'est pourquoi nous avons décidé d'augmenter la quantité d'échantillon à analyser et d'utiliser une coloration plus compatible avec l'analyse des spots par spectrométrie de masse. Ainsi, après avoir triplé la quantité d'échantillon à analyser (300 μ g), les gels ont été colorés au bleu de Coomassie (CBB) colloïdal, réputé pour sa grande compatibilité avec l'analyse par spectrométrie de masse, mais nettement moins sensible que la coloration à l'argent. En effet, environ 160 spots ont été détectés sur le gel contrôle coloré au CBB colloïdal, contre 350 sur le gel coloré à l'argent (Figure 17, A).

L'analyse qualitative et quantitative par PDQuest a révélé que l'intensité de 70 spots est modifiée d'au moins un facteur cinq après une infection bactérienne ou fongique (Figure 18). D'après la répartition des spots selon le type d'infection, nous observons que (i) la plupart des variations est liée à l'infection par le champignon, (ii) seuls deux spots sont régulés par tous les types d'infections, (iii) l'intensité de très peu de spots varie suite aux infections par *B. bassiana* en commun avec une des infections bactériennes et (iv) un nombre équivalent de spots est régulé soit de manière spécifique après stimulation par *M. luteus* (9) ou *E. coli* (7) soit par les deux à la fois (6).

L'analyse par spectrométrie de masse des spots excisés a permis d'identifier beaucoup de ces spots (cf. Tableau III). Plusieurs classes de protéines, définies selon les données de Flybase, sont représentées dans ce tableau. Certaines sont des protéines de structure ou des molécules du métabolisme général (glycolyse, transport d'oxygène). D'autres, font partie de la famille des protéases, des inhibiteurs de protéases ou des molécules de reconnaissance du pathogène dont l'implication dans la réponse immunitaire est connue ou très fortement supposée. Nous trouvons également la transferrine ainsi que des fragments de clivage de cette protéine. Enfin, d'autres molécules, telles que des protéines sensorielles, des facteurs de croissance de disque imaginaux, une peptidylglycine α -hydroxylating monooxygenase et une molécule de liaison au phosphatidyléthanolamine, apparaissent comme de nouveaux candidats pour l'étude des mécanismes de la réponse immunitaire de la drosophile. C'est sur l'ensemble de ces molécules que la discussion sera centrée.



Figure 17 : Cartes 2D colorées au bleu de Coomassie colloïdal des protéines de l'hémolymphe de drosophiles contrôles ou infectées. L'hémolymphe d'environ 600 drosophiles mâles (300 µg de protéines) contrôles (A), 72 h après infection par *Beauveria bassiana* (B) ou 6 h après inoculation par *Micrococcus luteus* (C) a été analysée par électrophorèse 2D. La première dimension a été réalisée sur des bandelettes IPG 5-8, la seconde dimension sur un gel SDS-PAGE 11%. Après coloration au bleu de Coomassie colloidal, on peut compter 160 spots sur la carte contrôle. Les spots numérotés correspondent aux spots dont l'intensité varie d'au moins un facteur cinq après l'infection et ont été excisés puis analysés par spectrométrie de masse MALDI-TOF. Les informations correspondantes sont récapitulées dans le tableau III.

Résultats

	Identité du gène	Noms / similitudes ^b	N° des	Taux de variation ^d		
Fonctions ^a			spots ^c	B. bassiana	M. luteus	E. coli
Molécules de liaison au calcium	CG15848 (Scp1)	Sarcoplasmic calcium-binding protein 1	1	i	ND	ND
	CG17876 (Amy-d)	Alpha-amylase distal	2	++ (5)	++ (5)	+
Protéines chaperonnes	CG14207	HSP20-like chaperone	3	+	++ (7)	++ (9)
Immunité / Défense	001057())	NT d	4			
Serpines	CG6687	Necrotic-like	4 D	++ (0) i	= ND	= ND
	20000/	Necrotic-like	Ē	i	ND	ND
Protéases à sérine	CG16705	Prophenoloxydase activating enzyme	5	++ (6)	=	_
	CG1102	Prophenoloxydase activating enzyme	C/6	i	ND	ND
	CG9372	Sérine Protéase à domaine CLIP	F	++ (6)	ND	ND
Molécules de liaison	CG13422	Gram Négative Binding Protein-like	7	i	=	=
ADN	CG10303 (TepTV)	I hiolester containing protein IV	A	1	ND	ND
ADN	CG7760 (Dilaseii)	Deoxyribonuclease II	0	++ (12)	=	=
Engumos		Deoxynoondelease n	,	1	—	_
<u>Enzymes</u>		Pentidylglycine α -hydroxylating				
Amidation	CG3832 (Phm)	monooxygenase	В	i	ND	ND
Métabolisme éporgétique						
Déshydrogenases	CG9042 (Gpdh)	Sn-glycerol 3 phosphate dehydrogenase	10	8	ND	ND
Desny ur ogenuses	(opun)	Sn-glycerol 3 phosphate dehydrogenase	11	s	ND	ND
	CG9244 (Acon)	Aconitase (mitochondrial)	12	ND	i	ND
	CG3752	Aldehyde dehydrogenase	13	(5)	+	+
Glycolyse	CG12055 (Gapdh1)	Glyceraldehyde 3 phosphate	14	+	(11)	_
	CG17654 (Eno)	Enolase	15	s	ND	ND
	001/007(200)	Enolase	16	_	++ (6)	+
	CG6058 (Ald)	Fructose 1,6 bisphosphate aldolase	17	+	=	++ (6)
Kinases	CG4929 (Argk)	Arginine Kinase	18	i	ND	ND
Protéases	CG11459	Cathepsin L	19	i	_	_
Facteurs de croissance	CG4475 (Idgf2)	Imaginal Disc Growth Factor 2	20	S	ND	ND
		Imaginal Disc Growth Factor 2	21	S	ND	ND
	CG4559 (Idgf3)	Imaginal Disc Growth Factor 3	22	=	(10)	-
Transport d'hydrogène	<i>CG11154</i> (ATPsyn-b) <i>CG6030</i> (ATPsyn-d)	ATP synthase b chain ATP synthase subunit d	23 24	=	i ND	i ND
	CG6186 (Tsf1)	Transferrin 1 entire form	25	++(10)	_	_
Métabolisme du fer		Transferrin 1 entire form	26	++(15)	=	=
		Transferrin 1 cleaved form	27	i	ND	ND
		Transferrin 1 cleaved form	28	i	ND	ND
		Transferrin 1 cleaved form	29	1	ND	ND
		Transferrin 1 cleaved form	30 31	1 i	ND ND	ND ND
		Transferrin 1 cleaved form	32	i	ND	ND
		Transferrin 1 cleaved form	33	i	ND	ND
	CG2216 (Fer1HCH)	Ferritin 1 heavy chain homologue	34	(22)	++ (6)	++(11)
		Ferritin 1 heavy chain homologue	35	(17)	=	=
	CG1469 (Fer2LCH)	Ferritin 2 light chain homologue	36	S	+	_
		Ferritin 2 light chain homologue	37	s	+	+
		Ferritin 2 light chain homologue	38	(13)	=	=
		Ferritin 2 light chain homologue	39	+	(5)	-
Molécules sensorielles	CG7584 (Obp 99c)	Odorant-binding protein 99c	40	++ (7)	(5)	_
Autres molécules	CG17919	Phosphatidylethanolamine binding protein	41	ND	i	_

Tableau III : Identité des protéines dont l'abondance varie après infection d'après l'analyse des gels 2D.

^a Catégories de fonction indiquées par Flybase (http://flybase.bio.indiana.edu) ; ^b Le nom de la protéine est indiqué lorsque le produit des gènes correspondants a été identifié, sinon la similitude de fonction est mentionnée ; ^c Annotation des spots : les lettres correspondent aux spots annotés sur la Figure 16 et les numéros à ceux annotés sur la Figure 17 ; ^d Le taux de variation des spots a été déterminé grâce à une analyse par le logiciel PDQuest.

i : protéines induites après infection ; s : protéines supprimées après infection ; ++ (x) : protéines dont l'intensité sur le gel augmente d'au moins un facteur cinq après infection (le taux de variation est indiqué entre parenthèses) ; - (x) : protéines dont l'intensité sur le gel diminue d'au moins un facteur cinq après infection (le taux de variation est indiqué entre parenthèses) ; + : protéines dont l'intensité sur le gel augmente d'un facteur inférieur à cinq après infection ; - : protéines dont l'intensité sur le gel diminue d'un facteur inférieur à cinq après infection ; - : protéines dont l'intensité sur le gel ne varie pas après infection ; ND: protéine non détectée sur le gel.



Figure 18 : Diagramme de Venn représentant la répartition du nombre de spots régulés en fonction du type d'infection. Les gels 2D des protéines de l'hémolymphe de drosophiles, 6 h après inoculation d'*Escherichia coli* ou de *Micrococcus luteus*, ou 72 h après infection par *Beauveria bassiana*, ont été analysés par PDQuest. Cette analyse a permis de définir la répartition des spots dont l'intensité varie d'au moins un facteur cinq suite à une, deux ou trois infections. Les protéines identifiées par spectrométrie de masse sont regroupées dans le tableau III.

L'ensemble des gels 2D réalisés au cours de ma thèse a permis l'identification de beaucoup de protéines de l'hémolymphe de drosophile. Au total, 128 spots ont été identifiés correspondant à 66 protéines différentes. Un gel 2D synthétique regroupant toutes les protéines de l'hémolymphe de drosophiles ayant été détectées au cours de ma thèse figure en Annexe 2. Ces protéines ont des masses moléculaires comprises entre 17 et 114 kDa, des points isoélectriques entre 4,7 et 8,8 et des valeurs d'hydropathie (GRAVY pour Grand Average of hydropathy (Kyte and Doolittle, 1982)) comprises entre -0,739 et 0,177.

C. Etude complémentaire : Analyse de la transferrine

Nous avons mis en évidence que de nombreux spots dont l'intensité varie d'au moins un facteur cinq après infection correspondent à des protéines du métabolisme du fer, notamment la transferrine (TSF). Cette protéine, impliquée dans le transport du fer, est visualisée sur gel 2D sous forme d'une série de spots espacés les uns des autres d'environ 0,2 unités de pH (Figure 19, A). L'abondance relative de ces isoformes est inchangée suite aux infections bactériennes testées (*E. coli* et *M. luteus*) mais toutes sont surexprimées après l'infection fongique. De manière intéressante, neuf autres spots ont été identifiés comme étant des produits de clivage de la TSF. Ces fragments, très conservés sur les différents gels réalisés, sont induits uniquement après infection par le champignon. L'analyse par spectrométrie de masse a révélé que la coupure s'effectue du côté C-terminal aussi bien que du côté N-terminal (Figure 19, B). L'ensemble de ces données nous a amené à nous intéresser de plus près à cette protéine.

C.1. Régulation de la transferrine chez des drosophiles mutantes

Afin d'étudier la régulation de la TSF et vérifier si celle-ci dépend des voies de signalisation Toll ou IMD, l'expression de la TSF dans différents mutants pour ces voies a été étudiée. Une analyse par PCR quantitative a été préférée en première approche car elle est plus rapide à réaliser et nécessite beaucoup moins de drosophiles que l'analyse par gels 2D (cinq par expérience).



Figure 19 : La transferrine et ses fragments de clivage. Neuf fragments de transferrine ont été détectés sur les gels 2D (A). Le recouvrement obtenu par spectrométrie de masse (B) montre que la transferrine peut être clivée aussi bien du côté N-terminal que du côté C-terminal. La transferrine entière est composées de 641 acides aminés dont un peptide signal de 28 acides aminés (en vert). Les spots 6 et 7 ont été distingués grâce aux gels 2D de gamme étroite.

Le niveau d'expression des ARN messagers de TSF a été mesuré dans des mouches mutantes pour la voie Toll (mutants *spaetzle*, *dif*, *perséphone* ou *GNBP3*) ou IMD (mutants *kenny* et *PGRP-LC*). Des infections par *E. coli*, *M. luteus* ou *B. bassiana* ont été réalisées et les expériences répétées de une à six fois. Les niveaux d'expression des peptides antimicrobiens Drosomycine et Diptéricine ont été mesurés en parallèle pour contrôler la validité des résultats.

Ces analyses confirment que les infections bactériennes ne modifient pas le niveau d'expression des ARN messagers de TSF tandis que l'infection fongique entraîne une augmentation de ce niveau. Les mutants de la voie IMD testés se comportent comme des contrôles tandis que les mutants *dif* ne surexpriment plus la TSF après infection par le champignon *B. bassiana*. Comme nous le soupçonnions d'après nos gels 2D, la régulation de la TSF semble être liée à la voie Toll. Dans le but d'affiner l'étude de la régulation de la TSF par la voie Toll, plusieurs mutants affectés à différents niveaux de cette voie ont été testés (*GNBP3, perséphone* ou *spaetzle*). Les résultats indiquent clairement et de manière reproductible que les mutants *spaetzle* synthétisent nettement moins de TSF après infection fongique que des mouches contrôles (Figure 20, A). Lorsqu'on considère des mutants affectés pour des protéines situées plus en amont que *spaetzle* dans la voie Toll (*perséphone* ou *GNBP3*), on constate que l'expression de la TSF est normale voire supérieure aux contrôles. Cependant, une étude sur un double mutant *perséphone / GNBP3* montre une diminution notable du niveau d'expression des transcrits de la TSF.

Afin de vérifier la relation entre le niveau d'expression des ARN messagers de la TSF et le taux d'abondance de la protéine dans l'hémolymphe, une étude par petits gels 2D a été réalisée à partir de l'hémolymphe des drosophiles mutantes. Cette étude a permis de confirmer les résultats obtenus par PCR quantitative pour les mutants *spaetzle*, *GNBP3* et *perséphone*. En effet, on peut voir que les mutants *spaetzle* ne surexpriment plus la TSF après infection fongique tandis que les mutants *GNBP3* et *perséphone* le peuvent encore (Figure 20, B). On notera cependant que chez les mutants *perséphone*, l'induction de la TSF semble être moins importante que ce que laissait supposer le niveau d'expression de ses transcrits. Pour les mutants *dif*, les résultats sont plus nuancés. Chez ces mutants le niveau d'expression de la TSF augmente après infection mais légèrement moins que chez les mouches sauvages.



Figure 20 : Régulation de la transferrine chez des drosophiles mutantes. Le niveau d'expression des transcrits de la transferrine avant et après infection par *Beauveria bassiana* (Bb) a été mesuré par PCR quantitative chez des mutants de la voie Toll (A). Des gels 2D ont été réalisé avec l'hémolymphe d'environ 30 drosophiles mâles contrôles ou mutantes avant et après infection par Bb (B). Seule la portion des gels regroupant les spots correspondants à la transferrine est représentée ici. NI : non infecté ; *spz* : mutant *spaetzle* ; *psh* : mutant *perséphone*.

C.2. Inactivation du gène de la transferrine

Afin d'étudier l'influence de la TSF sur la capacité de défense des drosophiles vis-à-vis de certains pathogènes, nous avons entrepris d'inactiver le gène codant la TSF grâce à une approche par injection d'ARN double brin. L'inhibition a été contrôlée par PCR quantitative en suivant le taux d'ARN messagers de transferrine jour après jour. Les résultats montrent que le taux de TSF diminue d'environ 80% 48 h après l'injection d'ARN double brin et que cette inhibition dure au moins six jours (Figure 21, A). Les mouches ont donc été infectées 48 h après l'injection d'ARN double brin et le taux de survie après différentes infections a été étudié (Figure 21, B). L'efficacité des infections a été vérifiée sur des mutants de la voie Toll ou IMD en fonction du pathogène utilisé. On peut constater que l'inactivation du gène de la TSF n'affecte pas le taux de survie des drosophiles suite aux infections testées.

C.3. Spécificité des fragments de transferrine

Les champignons sécrètent de nombreuses protéases pour pénétrer dans l'hémocoele des drosophiles. Afin de vérifier que la TSF est coupée de manière spécifique et non pas par une protéase sécrétée par le champignon, nous avons effectué un gel 2D sur des drosophiles surexprimant une protéase endogène, Perséphone. Nous retrouvons la plupart des fragments de TSF sur ce gel alors même que les mouches n'ont pas été en contact avec le champignon. Ce résultat nous conforte dans l'hypothèse d'une coupure spécifique de la TSF par un processus endogène de la drosophile.



Figure 21 : Inactivation de la transferrine par injection d'ARN double brin. Le taux d'expression de la transferrine après l'injection de l'ARN double brin a été mesuré par PCR quantitative (A). Ce taux diminue d'environ 80% dès deux jours (2j) et jusqu'à au moins six jours (6j) après l'injection des double brins d'ARN (NI : drosophiles non injectées). Le taux de survie des drosophiles injectées a été étudié après infection par différents germes (B). Noir : mouches sauvages infectées ; jaune : mouches mutantes infectées (contrôle positif) ; orange : mouches injectées par du tampon seul et non infectées ; rouge : mouches injectées par de l'ARN double brin et non infectées ; bleu clair : mouches injectées par de l'ARN double brin et infectées ; bleu clair : mouches injectées par de l'ARN double brin et infectées ; bleu clair : mouches injectées par de l'ARN double brin et infectées.

Temps (h)

Temps (h)

Chapitre 2 : Recherche de nouvelles molécules impliquées dans la réponse immunitaire de la drosophile : approche par marquage isotopique différentiel ICAT

L'analyse par électrophorèse 2D est une technique très puissante pour étudier les échantillons biologiques complexes, néanmoins cette approche n'est pas optimale pour l'étude des molécules de faible masse moléculaire, de masse moléculaire supérieure à 150 kDa, des molécules hydrophobes ou à pI extrêmes. De plus, la quantité d'échantillon analysée ne permet souvent que de visualiser les molécules les plus abondantes. Par ailleurs, les méthodes basées sur des gels d'électrophorèse se heurtent souvent à un problème de reproductibilité inter-gels. Enfin, l'identification des protéines contenues dans les spots est une étape qui est loin d'être parfaite. En effet, au cours de la digestion "in gel", bien que performante, de nombreux peptides trypsiques peuvent être perdus lors de l'extraction.

Pour résoudre ces problèmes, des approches protéomiques sans gel ont été mises en place parmi lesquelles on trouve l'ICAT (Isotope-Coded Affinity Tag). Cette méthodologie utilise le marquage isotopique différentiel des échantillons suivi d'une analyse par chromatographie liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS). En plus de s'affranchir des problèmes liés à l'électrophorèse 2D, elle possède l'avantage de mettre en évidence des variations d'abondance plus faibles.

Nous avons eu la chance d'appliquer cette méthodologie à l'étude des protéines de l'hémolymphe de drosophiles grâce à une collaboration avec les Dr Alan Aderem et Adrian Ozinsky à l'ISB (Institute for Systems Biology, Seattle). J'ai eu l'opportunité de me former, au cours d'un stage de 15 jours à l'ISB, à la compréhension et la maîtrise des outils informatiques permettant l'analyse des données issues de la stratégie ICAT.

A. Description de l'approche ICAT

A.1. Principe méthodologique

Cette technique repose sur la fixation d'un marqueur contenant un isotope léger ou lourd, respectivement sur les protéines de nos échantillons contrôles ou infectés. Le marqueur ICAT est composé d'une étiquette biotine, d'un groupement iodoacétamide et d'une chaîne marquée par un isotope léger (12 C) ou lourd (13 C) (figure 22). Cette chaîne permet de distinguer les deux marqueurs uniquement par la différence de masse de 9 Da résultant de la substitution de neuf ¹²C en neuf ¹³C. Les cystéines d'un premier échantillon sont marquées avec la sonde légère et celles d'un second sont marquées avec la sonde lourde. Les échantillons sont alors mélangés et digérés par la trypsine. Les peptides non marqués (sans cystéine) sont éliminés grâce à une colonne d'affinité et les fractionnés par échantillons sont chromatographie échangeuse d'ions. Les peptides/protéines sont identifiés par LC-MS/MS. Lors de la première étape de spectrométrie de masse, les peptides apparaissent par paires dont la masse varie de 9 Da (légers et lourds). La seconde étape fragmente chaque paire de peptides pour en donner la séquence, permettant l'identification de la protéine. Finalement, le rapport d'intensité des ions porteurs de l'isotope léger / lourd permet de calculer les concentrations relatives des peptides, donc des protéines dans l'échantillon de départ. La stratégie ICAT est récapitulée sur la Figure 22.

A.2. Principe des logiciels d'analyse des données

Avec les avancées des approches protéomiques basées sur la spectrométrie de masse, la capacité de générer des données est freinée par notre capacité de les analyser. En effet, la majorité des données provenant des analyses par MS/MS sont polluées par des bruits de fond ou des contaminants qui n'ont aucune valeur significative et qui compliquent l'analyse des données. Ainsi le développement d'algorithmes pour faciliter l'interprétation des spectres a été d'une grande aide pour les analyses protéomiques. De nos jours, la plupart de ces logiciels sont basés sur des scores associés à un seuil en dessous duquel la probabilité d'identification est considérée comme non significative.



Figure 22 : Représentation schématique de la stratégie ICAT. La structure du réactif ICAT est rappelée dans l'encart bleu. (Voir texte pour les détails).

Ces filtres nécessitent généralement une vérification par un utilisateur, plus ou moins averti, ce qui amène une part supplémentaire de subjectivité à la validation des résultats. Ceci laisse la porte ouverte à de nombreux faux positifs, ainsi qu'à la mise à l'écart de données qui auraient pu être interprétées. Ainsi, outre le fait que l'analyse des données nécessite beaucoup de temps, la notion d'acceptabilité dépend du logiciel utilisé, du laboratoire d'étude, ainsi que de la vérification manuelle. Ces paramètres font qu'il est très difficile de comparer des expériences entre elles et encore plus entre différents laboratoires.

Aujourd'hui, des logiciels sont développés pour apporter une solution à ce problème en incluant des algorithmes statistiques suffisamment stringents pour assigner une valeur de confiance à chaque identification de peptide et protéine. Ceci permet d'estimer le taux d'erreur (faux positifs) d'une analyse à grande échelle et il reste donc à l'utilisateur de choisir le seuil d'acceptabilité de ce taux. Ceci est de loin plus réaliste et plus précis que toutes les méthodes commercialisées et utilisées auparavant. Tandis que la plupart des logiciels assignent une probabilité aux données générées, ils ne peuvent pas encore prendre en compte la qualité de ces données. Le programme PeptideProphet (Keller et al., 2002) développé à l'ISB est basé sur un algorithme EM (Expectation-Maximization) qui permet de discriminer de manière très précise entre les peptides corrects et incorrects en tenant compte, entre autres, de la qualité de la digestion (nombre de résidus terminaux correspondant à une coupure trypsique) et des spectres (différence entre masse mesurée et masse théorique). ProteinProphet (Nesvizhskii et al., 2003) complète PeptideProphet en effectuant un regroupement non aléatoire des peptides selon leur appartenance à une même protéine en incluant des paramètres de pénalisation selon le nombre de peptides générés par la protéine, la longueur des peptides et des protéines, si un peptide correspond à plusieurs matchs dans les banques de données, etc.

J'ai eu la chance de pouvoir utiliser ces logiciels puissants et d'appliquer la technique ICAT à l'étude des molécules de l'hémolymphe de drosophiles avant et après différentes infections expérimentales.

B. Etude différentielle des protéines de l'hémolymphe après infection expérimentale

B.1. Mise en place de la méthodologie

Afin d'étudier la faisabilité de cette approche, une première analyse a été menée sur l'hémolymphe de drosophiles contrôles ou 6 h après infection par un mélange bactérien (*E. coli / M. luteus*). Les hémolymphes de 100 drosophiles contrôles (environ 50 μ g) et 100 drosophiles infectées ont été soumises à une analyse ICAT comme décrit plus haut.

B.1.1. Choix du seuil de sensibilité

Comme nous l'avons évoqué précédemment, l'utilisateur doit choisir la probabilité minimale à considérer afin de trouver le meilleur compromis entre le nombre de faux positifs et celui des identifications correctes. La Figure 23 représente les taux d'erreur (fraction correspondant aux faux positifs) et de sensibilité (fraction correspondant aux identifications correctes) déterminés par ProteinProphet pour cette expérience, en fonction de la probabilité minimale considérée.

On peut voir sur ce graphique que si l'on choisit par exemple une probabilité minimale de 0,99, seules 115 protéines entrent dans l'analyse mais toutes sont correctement identifiées (0,2% de faux positifs en ne tenant compte que de 80,8% des identifications totales). En diminuant la probabilité minimale, on risque d'avoir plus de faux positifs mais ceci est compensé par le fait qu'on met de côté moins de protéines correctement identifiées. Ainsi pour une probabilité minimale de 0,8, 137 molécules entrent dans l'analyse (soit 22 de plus qu'avec une probabilité de 0,99) dont 136 sont correctement identifiées et une seule est un faux positif (1% de faux positif en tenant compte de 95,4% des identifications totales). Si l'on réduit encore la probabilité minimale, le risque d'avoir plus de faux positif n'est plus contrebalancé par le bénéfice de plus de protéines identifiées correctement. Dans cette expérience, j'ai donc choisi une probabilité minimale de 0,8.



Figure 23 : Choix de la probabilité minimale d'identification des protéines par ProteinProphet. (A) Taux d'erreur et de sensibilité calculés par ProteinProphet à partir de l'ensemble des données provenant de l'analyse des molécules de l'hémolymphe de drosophiles avant et après infection par un mélange bactérien *E. coli / M. luteus*. Vert : Taux d'erreur, fraction correspondant aux protéines incorrectes ; Rouge : Taux de sensibilité, fraction correspondant aux protéines correctes. Le tableau (B) présente les valeurs de sensibilité et d'erreur ainsi que le nombre correspondant de protéines correctes et incorrectes.

Pour chaque expérience, l'utilisateur choisit de cette manière le compromis le plus raisonnable entre le nombre de faux positifs et celui des identifications correctes.

B.1.2. Calcul du taux d'abondance relatif des protéines

Afin de calculer automatiquement l'abondance relative des protéines dans chaque échantillon, un algorithme statistique, ASAPRatio (pour Automated Statistical Analysis of Protein abundance Ratio), a été développé à l'ISB (Li et al., 2003).

Avant la création de ce programme, le calcul de l'abondance relative des protéines dans les échantillons était effectué par le programme XPRESS (Han et al., 2001). Pour chaque peptide identifié par MS/MS, XPRESS reconstruit le profil d'élution pour chaque ion et isole les chromatogrammes de l'isotope lourd et léger ; il détermine une aire pour chaque pic de peptides et calcule l'abondance relative par comparaison de ces aires. De nombreux facteurs peuvent cependant compliquer l'analyse quantitative : (1) plusieurs peptides peuvent appartenir à une même protéine, (2) lors de la séparation chromatographique par échange d'ions, un peptide peut être présent dans différentes fractions, (3) un même peptide peut être identifié plusieurs fois dans différents états de charge, (4) un peptide peut porter une modification post-traductionnelle, (5) la qualité des données peptidiques (rapport entre signal/bruit de fond) peut varier. Ainsi, il était indispensable de créer des outils automatiques pouvant prendre ces paramètres en considération pour évaluer de manière fiable l'abondance relative des protéines.

Le programme ASAP, basé sur de nombreux algorithmes numériques et statistiques, reprend le principe de calcul du programme XPRESS mais permet de normaliser les ratios des peptides identifiés plusieurs fois et affecte une valeur unique pour chaque peptide. L'abondance relative d'une protéine est ensuite calculée en regroupant tous les ratios des peptides qui proviennent de cette protéine et en affectant une probabilité pour que ce ratio soit significatif. Toutes ces contraintes font de l'ICAT une méthodologie quantitative fiable et précise nous permettant d'observer des variations plus faibles que par l'approche par gels 2D. En accord avec les recommandations des personnes compétentes à l'ISB et la qualité de nos résultats, nous avons fixé le seuil minimal de variation à deux.

B.1.3. Analyse des protéines de l'hémolymphe de drosophiles

Les molécules de l'hémolymphe d'environ 100 drosophiles (50 µg) avant et après blessure septique ont été marquées différentiellement par le réactif ICAT et analysées par chromatographie liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS). L'abondance relative des protéines dans chaque échantillon a été calculée en utilisant le programme ASAP comme décrit précédemment.

Pour cette expérience, 137 protéines ont été identifiées avec une probabilité minimale de 0,8, parmi lesquelles 40 sont régulées d'au moins un facteur deux. Les protéines détectées ont des masses moléculaires comprises entre 4,3 et 426 kDa, des points isoélectriques entre 4,0 et 9,95 et des valeurs d'hydropathie (GRAVY) comprises entre -1,089 et 0,867. La liste des protéines régulées est reportée en Annexe 3. Plusieurs familles de molécules sont représentées : des protéases, des inhibiteurs de protéase, des molécules de liaison ou de transport, des molécules impliquées dans le métabolisme, etc.

Cette étude préliminaire a montré que la stratégie est adaptée à l'analyse des molécules de l'hémolymphe de drosophiles et que 50 µg de protéines étaient suffisants pour obtenir des résultats exploitables.

B.2. Infections spécifiques

B.2.1. Choix des agents microbiens et des temps d'infection

Afin d'étudier la spécificité des régulations, il était nécessaire d'effectuer l'analyse ICAT sur des drosophiles contrôles et infectées par un seul agent microbien à la fois. Le choix des pathogènes et des temps d'infection a été défini selon les connaissances que nous avions de la réponse immunitaire des drosophiles incluant celles qui ont été acquises au cours des premières années de ma thèse et celles obtenues grâce aux gels 2D. Nous avons constaté que 6 h après infection par les bactéries *E. coli* et *M. luteus*, les drosophiles semblent déclencher une réponse immunitaire relativement discrète par rapport à l'infection fongique. Ainsi, nous avons décidé d'effectuer des

infections avec des bactéries plus pathogènes. D'autre part, l'équipe de Julien Royet au laboratoire venait de montrer que certaines bactéries à Gram-positif déclenchent la voie Toll dépendamment ou indépendamment de PGRP-SA (Bischoff et al., 2004). C'est pourquoi deux bactéries à Gram-positif ont été sélectionnées : une qui active la voie Toll de manière PGRP-SA-dépendante (*Enterococcus faecalis*) et l'autre de manière PGRP-SA-indépendante (*Enterococcus faecalis*) et l'autre de manière PGRP-SA-indépendante (*Staphylococcus saprophyticus*). A ces bactéries à Gram-positif, nous avons ajouté une infection par la bactérie à Gram-négatif *Agrobacterium tumefaciens*. Nous avons gardé l'infection par le champignon entomopathogène *Beauveria bassiana*. Enfin, à la suite des données génétiques et transcriptomiques obtenues au laboratoire par l'équipe de Jean-Luc Imler sur la réponse immunitaire de la drosophile après infection par le virus C de la drosophile (DCV), nous avons décidé d'inclure à notre étude une infection par ce virus.

L'hémolymphe de 100 à 160 mouches (50 à 80 µg de protéines) a été collectée 6 h et 24 h après infections bactériennes, 48 h après infection virale et 72 h après infection fongique. Concernant les infections bactériennes, seuls les échantillons d'hémolymphe 24 h après stimulation ont été analysés à l'heure actuelle. Les échantillons d'hémolymphe de drosophiles 6 h après stimulation par les trois bactéries ont été marqués par le réactif ICAT mais n'ont pas encore été analysés par LC-MS/MS.

B.2.2. Identification des protéines

Pour chaque expérience une probabilité minimale de 0,7 a été choisie car elle représente, pour ces analyses, le meilleur compromis entre le nombre de protéines correctement identifiées et le nombre de faux positifs (en moyenne trois faux positifs pour 150 protéines).

En regroupant les résultats des expériences issues des cinq infections, 262 protéines ont été analysées parmi lesquelles 162 sont régulées d'au moins un facteur deux. La liste des protéines identifiées est donnée en Annexe 4. Ces protéines ont des masses moléculaires comprises entre 4,1 et 515 kDa, des points isoélectriques entre 4,4 et 10,9 et des valeurs GRAVY comprises entre -1,108 et 0,874.

Parmi ces 262 protéines, certaines sont connues pour leur implication dans la réponse immunitaire de la drosophile et nous avons vérifié que leur abondance relative

vis-à-vis des infections est en accord avec les données de la littérature ou celles obtenues par d'autres méthodes. Ainsi, nous constatons par exemple que la protéine Spaetzle (produit du gène *CG6134*), ligand du récepteur Toll, est surexprimée après infection par *E. faecalis* (Gram-positif) mais non par *A. tumefaciens* (Gram-négatif) ni par le DCV ; la Drosomycine (issue du *CG10810*) est surexprimée suite à l'infection fongique et par *E. faecalis* ; PHK-2 (produit du gène *pebIII*, *CG11390*) est surexprimée suite à plusieurs infections mais bien plus fortement par le DCV. Ces données confortent la validité des résultats de l'analyse par ICAT.

Parmi les 262 protéines, la Drosomycine est le seul peptide antimicrobien connu. De même, uniquement six DIMs sur les 33 caractérisées sont détectées (DIMs 3, 10, 12, 13, 23 et 33). Ceci s'explique par le fait que la plupart des peptides antimicrobiens et des DIMs ne possèdent pas de cystéines et ne peuvent pas donc être analysés avec le marqueur utilisé. Il faut noter cependant que certains peptides ne possédant pas de cystéines peuvent tout de même être copurifiés de la colonne d'affinité et entrer dans l'analyse. Ces peptides participent à l'identification de la protéine correspondante mais ne fournissent aucune information quantitative. Ainsi, les DIMs 10, 12 et 13, référencés comme produit du *CG18279*, ne possèdent pas de cystéines mais ils ont été identifiés de manière non ambiguë à partir de trois peptides. La concentration relative de ces DIMs dans l'échantillon concerné n'a cependant pas pu être calculée (notée NA dans le Tableau de l'Annexe 4).

Il faut garder à l'esprit que la quantité de cystéines est différente d'une protéine à une autre. La caractérisation des protéines repose donc sur l'identification d'un nombre variable de peptides. Pour certaines, ce nombre sera élevé, tandis que pour d'autres, l'identification sera effectuée à partir d'un seul peptide. Il est intéressant de noter qu'en moyenne près de 30% des protéines issues de l'analyse ont été identifiées à partir d'un seul peptide (Figure 24).

Afin de tirer profit de la multiplicité des agents infectieux utilisés, nous avons voulu, dans un premier temps, effectuer une analyse comparative de l'abondance des protéines détectées dans l'hémolymphe de drosophiles après chaque type d'infection. Cependant, comme nous l'avons vu plus haut, près de 30% des protéines ont été identifiées à partir d'un seul peptide. Après l'étape chromatographique, si plusieurs



Figure 24 : Nombre de peptides ayant servi à l'identification des protéines. En moyenne, 29% des protéines ont été identifiées à partir d'un seul peptide.



Figure 25 : Nombre de protéines identifiées en commun dans une, deux, trois, quatre ou cinq infection(s) en fonction du nombre moyen de peptides utilisés pour l'identification. La plupart des protéines identifiées en commun dans les cinq infections a été identifiée à partir de plus de cinq peptides, tandis que la plupart des protéines identifiées dans une seule des cinq infections a été identifiée à partir de seule ment un peptide.

Résultats

peptides sont coélués, le spectromètre de masse est programmé pour sélectionner ceux dont l'intensité de signal est la plus forte. Ainsi, il est possible de ne pas sélectionner un peptide minoritaire, donc de ne pas obtenir d'information de séquence, et par conséquent aucune identification de la protéine correspondante. D'ailleurs, lorsqu'on regarde l'histogramme représentant le nombre de protéines identifiées dans chaque infection en fonction du nombre de peptides ayant servi à leur identification (Figure 25), nous constatons que plus le nombre de peptides est élevé et plus la probabilité de retrouver une même protéine dans plusieurs infections est grande. C'est pourquoi le fait de ne pas retrouver une protéine dans un des échantillons ne prouve pas que cette protéine est absente de cet échantillon. L'analyse comparative ne peut donc être menée que sur l'ensemble des protéines détectées et quantifiées pour toutes les infections (54/162 protéines régulées).

La Figure 26 représente la distribution de ces 54 protéines identifiées et quantifiées pour toutes les infections. On constate que la majorité des variations suite aux infections expérimentales sont des surexpressions. De manière intéressante, on peut noter que (i) l'abondance relative de certaines protéines varie spécifiquement, c'est-àdire après une seule des cinq infections : trois protéines semblent régulées spécifiquement après l'infection par *A. tumefaciens*, une par *E. faecalis*, trois par *S. saprophyticus*, deux par le DCV et dix par *B. bassiana* ; (ii) l'abondance relative d'un grand nombre de ces protéines (38/54) varie suite à la stimulation par *B. bassiana*, spécifiquement ou en commun avec d'autres infections ; (iii) seules deux molécules sont régulées par les cinq infections à la fois. Les familles de protéines identifiées ainsi que les taux d'abondance relative sont regroupées dans le Tableau IV (page suivante).

Cette approche comparative permet, au travers de ces 54 protéines, d'avoir un point de vue unique sur le comportement de plusieurs familles de molécules suite à l'infection par cinq germes différents. On remarque, par exemple, que toutes les protéases et tous les inhibiteurs de protéase sont surexprimés après infection fongique en commun ou non avec d'autres stimulations. On notera également que trois "Odorant Binding Proteins" sur quatre sont surexprimées suite à l'infection par le DCV. On peut voir, de plus, que certaines protéines sont surexprimées quel que soit l'inducteur (la protéase à sérine issue du gène *CG2045* et l'amylase proximale, produit du gène

78



Figure 26 : Diagramme d'Euler-Venn représentant la distribution des molécules identifiées par la stratégie ICAT en fonction des infections. Les 54 protéines identifiées et quantifiées dans chaque infection sont regroupées selon leur régulation par les différents inducteurs. Les résultats sont représentés sur un diagramme d'Euler-Venn à cinq entrées. La lecture de ce diagramme se fait, pour chaque infection, en suivant l'ensemble de la couleur correspondante. *B. b., Beauveria bassiana*; *A. t., Agrobacterium tumefaciens*; *DCV, Drosophila C virus*; *S. s., Staphylococcus saprophyicus*; *E. f., Enterococcus faecalis.* Gras : protéines surexprimées ; entre parenthèses : protéines réprimées ; italique : protéines surexprimées ou réprimées selon l'inducteur. *, molécules liées à la réponse immunitaire ; •, protéases ; \circ , inhibiteurs de protéases ; Ø, molécules du métabolisme ; §, Molécules de liaison ou de transport ; •, autres molécules ; Ø, molécules de fonction inconnue.

Famille	Identité du gène	Nom (abreviation)	A.t	B.b	DCV	E.f	S.s
Molécules liées	CG11390	Ejaculatory bulb protein III (PHK-2/PebIII)	0.35	0.40	0.12	0.41	0.60
à l'immunité	CG15066	Drosophila Immune induced molecule 23 (DIM23)	0.35	0.06	1.39	0.11	0.37
(*)	CG5008	Gram-negative bacteria binding protein 3 (GNBP3)	2.20	0.69	0.69	1.02	1.41
Protéases	CG9372		1.88	0.30	2.13	1.25	1.31
(•)	CG2045	(SER7)	0.36	0.13	0.34	0.24	0.51
	CG1102		0.34	0.24	0.93	0.22	0.84
	CG16705		0.64	0.15	0.83	0.23	0.56
	CG10593	Angiotensin-converting enzyme-related (ACER)	1.27	0.26	0.87	1.17	1.35
	CG5390		0.74	0.39	0.57	0.41	1.54
	CG8586		1.05	0.45	0.99	0.92	1.83
	CG31326		1.27	0.44	0.66	0.54	0.75
	CG6806	Larval serum protein 2 (LSP2)	2.76	8.79	0.75	1.72	0.59
Tabibitana da	CG16704		1.21	0.50	0.93	0.91	1.08
innibiteurs de	CG16712		0.75	0.51	0.78	0.58	1.10
proteases	CG16713		0.50	0.20	0.41	0.35	0.84
(0)	CG33103	Papilin (PPN)	0.65	0.38	0.57	0.85	1.66
	CG8050	Cystatin-like (CYS)	0.72	0.36	0.66	0.81	0.00
Malfaulas da	CG11797	Odorant-binding protein 56a (OBP 56a)	1.14	0.23	0.15	0.50	1.01
Notecules de	CG11218	Odorant-binding protein 56d (OBP56d)	0.69	0.49	0.20	0.60	0.90
transport	CG13873	Odorant-binding protein 56g (OBP 56g)	0.35	0.86	0.43	0.63	1.68
uansport	CG7584	Odorant-binding protein 99c (OBP99c)	1.56	1.53	0.94	1.01	2.28
(§)	CG4475	Imaginal disc growth factor 2 (IDGF2)	0.29	0.43	1.36	0.66	1.80
	CG4559	Imaginal disc growth factor 3 (IDGF3)	0.50	0.27	1.19	0.68	1.05
	CG1780	Imaginal disc growth factor 4 (IDGF4)	1.01	0.32	0.56	0.88	1.44
	CG6186	Transferrin 1 (TSF1)	1.27	0.11	0.84	0.44	1.30
	CG2216	Ferritin 1 heavy chain homologue (Fer1HCH)	1.59	4.89	0.95	1.93	6.87
	CG11001	FK506-binding protein 2 (FKBP2)	0.47	0.43	0.64	0.76	0.56
	CG15848	Sarcoplasmic calcium-binding protein 1 (SCP1)	0.64	0.58	1.03	0.63	0.48
	CG17210	SCP-containing protein B (SCPR-B)	0.88	1.44	0.36	0.87	1.17
	CG6426		0.45	0.41	0.79	0.36	0.64
	CG18730	Amylase proximal (AMY-p)	0.20	0.09	0.49	0.19	0.28
	CG11064	Retinoid- and fatty-acid binding protein (RFABP)	1.27	0.75	1.11	0.92	2.10
	CG31150		1.16	0.26	1.08	0.68	0.92
Malándas liáns	CG9244	Aconitase (ACON)	0.60	1.31	0.57	0.42	1.03
au métabolisme	CG32954	Alcohol dehydrogenase (ADH)	0.87	0.29	0.81	0.52	0.60
au metabolisme	CG9042	Glycerol 3 phosphate dehydrogenase (GPDH)	0.61	0.49	0.57	0.50	0.80
(♦)	CG8893	Glyceraldehyde 3 phosphate dehydrogenase 2 (GPDH2)	1.42	1.02	0.83	0.51	2.06
	CG7070	Pyruvate kinase (PYK)	0.53	1.01	0.65	0.68	1.64
	CG7998		0.80	1.06	0.34	0.65	1.01
Autres	CG8913		0.47	0.29	0.35	0.32	0.64
molécules	CG1634	Neuroglian (NRG)	1.23	0.90	1.15	2.04	2.16
(■)	CG5210	Chitinase-like (CHIT)	1.25	0.36	1.01	0.79	1.33
	CG11720	Salivary gland secretion 3 (SGS-3)	0.26	0.34	0.76	1.05	22.05
	CG14629		1.40	0.18	0.43	0.85	1.18
Molécules de	CG15616	(Acp53C14b)	1.59	0.20	0.79	2.22	1.36
fonction	CG32667		1.28	0.35	1.14	0.88	1.86
inconnue	CG17278		0.62	0.43	0.84	0.67	0.22
(Ø)	CG18064	(Met75Cb)	1.27	0.10	0.15	0.08	0.17
	CG30197		0.87	0.48	0.31	0.69	1.42
	CG11315		1.55	2.25	0.94	1.15	2.84
	CG31883		0.91	0.21	1.33	1.42	1.79
	CG32952		0.31	0.78	1.53	1.23	1.04
	CG14933		1.38	0.77	0.52	0.49	1.45
	CG14430		9999 00	1.75	1.18	1.38	1.97

Tableau IV : Identification des protéines quantifiées pour toutes les infections issues de l'analyse ICAT. Les ratios sont calculés en effectuant le rapport d'intensité des ions porteur de l'isotope léger/lourd. Les cases grisées correspondent aux protéines dont l'abondance relative augmente d'au moins un facteur deux (< ou = 0,5) après infection. Les cases noircies correspondent aux protéines dont l'abondance relative diminue d'au moins un facteur deux (> ou = 2) après infection. Les valeurs extrêmes (0 ou 9999) correspondent à des protéines totalement induites ou supprimées après une infection. Les symboles (*, •, \circ , §, •, \blacksquare , Ø) correspondent aux catégories de protéines telles qu'elles apparaissent sur la figure 26. Les analyses ont été réalisées 24 h après infection par les bactéries (A.t : *Agrobacterium tumefaciens* ; E.f : *Enterobacter faecalis* ; S.s : *Staphylococcus saprophyticus*), 48 h après infection par le virus C de la drosophile (DCV) et 72 h après infection par le champignon *Beauveria bassiana* (B.b). *CG18730*) tandis que d'autres ont un profil de régulation plus complexe. En effet, nous observons que la protéase issue du *CG9372* est fortement surexprimée après une infection par *B. bassiana* alors qu'elle est réprimée après stimulation par le DCV. De la même manière, le produit du gène *CG15616*, de fonction inconnue, est surexprimé après l'infection fongique et réprimé après l'infection par *E. faecalis*, alors même que ces deux inducteurs déclenchent la même voie de signalisation (Toll) dans le cadre de la réponse immunitaire. Ces résultats illustrent que certains aspects de la réponse immunitaire des insectes échappent encore à notre compréhension.

Au cours de ma thèse, les deux approches protéomiques (électrophorèse 2D et ICAT) réalisées pour étudier les molécules régulées au cours de la réponse immunitaire de la drosophile ont révélé que plusieurs familles de protéines sont très représentées. Il s'agit de protéases, d'inhibiteurs de protéase, de molécules sensorielles, de facteurs de croissance des disques imaginaux, de molécules antioxydantes, de protéines de liaison au fer, aux lipides et au calcium. La discussion sera centrée sur ces familles de molécules.

Les résultats obtenus au cours de ma thèse par l'approche 2D ont fait l'objet d'une publication dans Molecular and Cellular Proteomics (3, 156-166, 2004). Une seconde publication regroupant les données de l'analyse peptidomique et protéomique est parue dans Biochimie (86, 607-616, 2004). Ces publications se trouvent en Annexe 5. Le travail concernant l'approche ICAT est en cours de rédaction.

DISCUSSION

DISCUSSION

L'analyse protéomique consiste en l'étude qualitative et quantitative de toutes les protéines exprimées dans un organisme, un tissu, une cellule, une organelle, ou encore un fluide biologique, à un certain moment et dans certaines conditions physiologiques. Ce type d'approche, complémentaire aux études génomiques et transcriptomiques, permet de donner un point de vue nouveau sur l'étude d'un processus biologique.

Lors de ma thèse, deux approches ont été utilisées afin d'étudier le profil des protéines de l'hémolymphe de drosophile suite à différentes infections expérimentales. La première est basée sur l'utilisation combinée de l'électrophorèse bidimensionnelle et de la spectrométrie de masse, tandis que la seconde, développée plus récemment, est basée sur le marquage isotopique des protéines couplé à une analyse par chromatographie liquide et spectrométrie de masse en tandem.

NB : Comme nous l'avons vu précédemment, certaines protéines ont été identifiées par l'approche ICAT mais pas quantifiées (notamment par manque de cystéine). Ainsi, il est très important de garder à l'esprit tout au long de cette discussion qu'une protéine <u>détectée</u> n'a pas forcément été <u>quantifiée</u>.

A. Comparaison des deux approches

A.1. Caractéristiques des protéines détectées

L'approche par gels 2D a permis d'identifier 66 protéines différentes dont les masses moléculaires sont comprises entre 17 et 114 kDa, les points isoélectriques entre 4,7 et 8,8 et les valeurs GRAVY entre -0,739 et 0,177. L'approche ICAT a permis d'identifier 262 protéines différentes dont les masses moléculaires sont comprises entre 4,1 et 515 kDa, les points isoélectriques entre 4,4 et 10,9 et les valeurs GRAVY entre - 1,108 et 0,874.

En comparant les caractéristiques des protéines identifiées par chaque approche, on constate que la stratégie ICAT, comme nous l'attendions, couvre une plus large gamme de molécules. En effet, de par leurs masses moléculaires et leurs pI, plus de 50% des protéines détectées par ICAT n'auraient pas pu être détectées sur nos gels 2D. De plus, les protéines très hydrophobes telles que les protéines membranaires ont des valeurs GRAVY aux alentours de 0,6. Aucune protéine de ce type n'a été détectée sur nos gels 2D tandis que plusieurs l'ont été par l'étude ICAT (par exemple, les produits des gènes *CG32952*, *CG15068* et *CG15075*).

Parmi les molécules identifiées par l'approche 2D, 24 n'ont pourtant pas été détectées par la deuxième stratégie. Ces molécules, pour la plupart, ne possèdent pas ou peu de cystéines et ont été écartées de l'analyse ICAT après la chromatographie d'affinité. De plus, certaines de ces protéines n'ont été détectées que sur les gels 2D de gammes étroites de pH, indiquant qu'elles peuvent être présentes en très faible abondance dans l'hémolymphe des drosophiles. Les peptides générés par la digestion trypsique de ces protéines peuvent ne pas avoir été sélectionnés lors de l'analyse par spectrométrie de masse en tandem de par leur faible signal.

Parmi les protéines quantifiées dont l'expression est régulée en réponse aux infections, 18 sont détectées en commun par les deux stratégies : trois ont été détectées mais pas quantifiées par la stratégie ICAT, huit varient dans le même sens, cinq sont invariables selon une stratégie et régulées selon l'autre, et seules deux molécules ne varient pas dans le même sens (produit des gènes *CG9042* et *CG4475*). Les différences du niveau d'expression entre les deux stratégies peuvent provenir du fait que certaines protéines sont présentes dans l'hémolymphe sous plusieurs isoformes. L'approche par gels 2D discrimine entre les protéines ayant subi des modifications post-traductionnelles tandis que l'approche ICAT ne le peut qu'exceptionnellement et seul le ratio moyen des isoformes est calculé.

Enfin, il est important de mentionner ici la différence dans les seuils de variation pris en considération selon la stratégie. D'un côté, une analyse sur la reproductibilité des cartes 2D a montré qu'il était préférable de ne considérer que les variations d'au moins un facteur cinq. De l'autre, l'analyse ICAT, grâce à sa précision quantitative, a permis de prendre en compte les variations à partir d'un facteur deux. Ceci explique pourquoi la seconde stratégie a mis en évidence un nombre de protéines régulées très supérieur à celui issu de l'analyse par gels 2D.

A.2. Nature des variations

Les deux approches ont révélé que les variations d'abondance de la plupart des protéines sont liées à l'infection par le champignon *B. bassiana*. Ceci s'explique notamment par le mode d'infection et le développement du pathogène dans les drosophiles. D'un côté, l'infection par *B. bassiana*, champignon entomopathogène, est effectuée de manière naturelle en couvrant les drosophiles de spores qui germent sur la cuticule en occasionnant des lésions sévères aboutissant à une réponse immunitaire massive (Lemaitre et al., 1996). De l'autre côté, les bactéries sont inoculées localement par piqûre dans le thorax. De plus, ni *Micrococcus luteus* ni *Escherichia coli*, bactéries utilisées dans l'approche par gels 2D, ne sont pathogènes pour la drosophile. Ces microorganismes ont été sélectionnés originellement comme modèle d'infection par *E. coli* et *M. luteus*, est maîtrisée en moins de 12 h par la drosophile (Lemaitre et al., 1996; Lu et al., 2001).

Nous avons regroupé les molécules détectées par les deux approches en familles. Certaines ont un rôle connu ou fortement supposé dans la défense de la drosophile (GNBPs, PGRPs, TEPs, etc.), d'autres sont de nouveaux candidats (molécules sensorielles, antioxydantes, de liaison au calcium ou du métabolisme du fer). Les protéines de fonction inconnues ou du métabolisme général ne seront pas discutées ici.

Il est intéressant de noter que depuis 2000, date du séquençage du génome de la drosophile, plusieurs équipes se sont intéressées à l'étude transcriptomique peptidomique ou protéomique de cet insecte. Ces analyses, complémentaires les unes des autres, ont été publiées dans un laps de temps très rapproché. Nous discuterons également du recouvrement de ces différentes approches en axant sur les protéines détectées lors de ma thèse.
B. Molécules dont le rôle dans la réponse immunitaire est connu ou fortement supposé

B.1. Peptides antimicrobiens et DIMs

Parmi les protéines mises en évidence lors de notre approche protéomique, la Drosomycine est le seul peptide antimicrobien détecté. De même, uniquement six DIMs sont observées (DIMs 3, 10, 12, 13, 23 et 33). Ceci peut s'expliquer par le fait que la plupart des peptides antimicrobiens et des DIMs sont de petites molécules ne possédant pas de cystéines. Elles ne sont donc pas facilement analysées par les stratégies ICAT ou 2D utilisées ici.

La Drosomycine, produit du gène *CG10810* (également appelée Drosomycine A), est détectée dans les échantillons de mouches infectées par *B. bassiana* et *E. faecalis* où elle est fortement surexprimée. L'approche peptidomique a révélé que cette Drosomycine existe également sous une forme glycosylée (Tableau I, DIM 21). Le peptide ayant servi à l'identification de la protéine contient le site potentiel de glycosylation mais la masse du peptide sélectionné pour la fragmentation indique qu'il s'agit de la forme non glycosylée de la Drosomycine.

Les approches biochimiques effectuées jusqu'à présent n'avaient révélé que la Drosomycine A dans l'hémolymphe des mouches alors que le génome comporte sept séquences codant ce peptide antimicrobien. De plus les analyses par puces à ADN ont montré que le niveau des transcrits de deux Drosomycines est surexprimé suite aux infections microbiennes. Il s'agit de la Drosomycine A et de la Drosomycine 5 (produit du gène *CG10812*) surexprimées après infections bactériennes et fongiques (De Gregorio et al., 2001). Il est intéressant de noter que l'analyse par ICAT permet de détecter en plus de la Drosomycine A, non pas la Drosomycine 5, mais la Drosomycine 6 (produit du gène *CG32268*), surexprimée dans l'échantillon de mouches infectées par le champignon. Ceci illustre bien les différences qu'il peut y avoir entre le niveau d'expression des ARN messagers et celui des protéines.

Les deux Drosomycines ont été identifiées de manière non ambiguë par spectrométrie de masse, mais à partir de l'analyse d'un seul peptide. C'est pourquoi, comme nous l'avons vu précédemment, il est possible que ces molécules soient présentes dans les échantillons issus des autres infections mais que le peptide correspondant n'ait pas été sélectionné lors de l'analyse par spectrométrie de masse.

B.2. Protéases

Le taux d'abondance de nombreuses protéases est modifié après infection expérimentale. La plupart d'entre elles sont des protéases à sérine ou des homologues. Ces molécules représentent la seconde famille de gènes la plus représentée dans le génome de la drosophile. Elle compte en effet 204 membres dont 147 protéases à sérine et 57 homologues. Ces protéases sont caractérisées par une séquence signature centrée sur trois acides aminés (Histidine - Acide aspartique - Sérine), formant la triade catalytique, inclus dans des séquences très conservées (TAAHC, DIAL et GDSGGP, respectivement). Le terme "homologues" de protéases à sérine est utilisé car ces molécules possèdent de grandes similarités de séquence avec les protéases à sérine des cascades protéolytiques d'Arthropodes, mais l'activité enzymatique n'est pas conservée car un ou plusieurs résidus de la triade manque(nt) (Ross et al., 2003). Cependant, il a été montré chez certains Invertébrés, que ces homologues peuvent avoir une activité antimicrobienne, par exemple chez le moustique *A. gambiae* (Dimopoulos et al., 1997) ou chez la limule (Kawabata et al., 1996). Ils peuvent également intervenir au niveau de la reconnaissance du pathogène chez l'oursin (Huang et al., 2000).

Parmi les molécules détectées lors de notre étude protéomique, 20 sont des sérine protéases (13) ou homologues (7) dont 17 sont régulées de manière significative après infection. Parmi elles, 12 sont des protéases à domaine clip telle que Perséphone (produit du gène *CG6367*) qui a déjà été décrite pour son rôle dans la cascade d'activation extracellulaire de la voie Toll (Ligoxygakis et al., 2002a).

B.2.1. Protéases à sérine à domaine clip

Le domaine clip a été nommé d'après la forme en trombone qu'il peut adopter. Sa fonction exacte n'est pas connue mais pourrait intervenir dans l'interaction protéineprotéine. Ce domaine est activé par coupure protéolytique entre le domaine clip et le domaine catalytique, les deux domaines restant liés par un pont disulfure (Jiang and

Kanost, 2000). Les protéases à domaine clip régulées au cours de cette analyse protéomique sont, pour la plupart, fortement surexprimées après l'infection fongique en commun ou non avec d'autres infections. Cette famille de protéases fait l'objet d'une thèse en cours au laboratoire, au travers d'une approche de type "gènes candidats", où 28 protéases de drosophiles ont été clonées et surexprimées in vivo dans des lignées de mouches transgéniques. La surexpression de deux d'entre elles induit une activation constitutive de la voie Toll (produits des gènes CG2045 et CG5896) et une troisième présente un phénotype de mélanisation spontanée (produit du gène CG9372) (D. Rabel, communication personnelle). Ces trois protéines sont détectées lors de l'analyse par ICAT. Elles sont surexprimées après infection fongique en commun avec d'autres infections ou non. La protéine issue du gène CG2045 est l'une des deux protéines surexprimées après toutes les infections mais beaucoup plus fortement après l'infection par B. bassiana tandis que le taux d'abondance du produit du gène CG9372 est augmenté après infection fongique et diminué après infection par le DCV. Ce dernier est d'ailleurs trouvé surexprimé sur le gel 2D de gamme étroite après infection par le champignon. Le produit du gène CG5896, détecté par ICAT uniquement après stimulation par B. bassiana, est aussi surexprimé de manière significative. L'étude de ces trois protéases est toujours en cours au laboratoire.

B.2.2. Molécules potentiellement impliquées dans l'activation de la prophénoloxydase

Parmi les protéases détectées lors de l'analyse protéomique, certaines sont des enzymes putatives d'activation de la prophénoloxydase (PPAE). Les PPAEs activent le précurseur prophénoloxydase (PPO) en phénoloxydase (PO) lors de la réaction de mélanisation (Soderhall and Cerenius, 1998). Chez certains Arthropodes, des protéases à sérine à domaine clip, possédant deux cystéines supplémentaires et conservées dans leur domaine catalytique permettant la formation d'un pont disulfure, sont impliquées dans la coupure de la PPO (Jiang and Kanost, 2000). Le génome de la drosophile code 11 protéases à sérine ou "homologues" possédant ces cystéines et qui sont donc des PPAEs putatives (Ross et al., 2003). Une analyse par puces à ADN a révélé que l'expression des transcrits de cinq d'entre elles est augmentée suite à une stimulation

immune (De Gregorio et al., 2001). Parmi ces molécules, quatre sont également détectées lors des analyses protéomiques. Trois sont surexprimées suite à l'infection fongique (produits des gènes *CG16705*, *CG1102* et *CG3066*) en commun ou non avec d'autres infections. La quatrième (issue du *CG3505*), détectée uniquement dans les infections par *A. tumefaciens* et *E. faecalis*, n'est surexprimée de manière significative que suite à l'infection par la bactérie à Gram-négatif. Des lignées mutantes pour certaines de ces protéases ont été testées au laboratoire par l'équipe de Jean-Marc Reichhart. Ainsi, parmi les mutants testés, seule la protéine codée par le gène *CG3066* a été identifiée comme impliquée dans l'activation de la PO (Leclerc et al., soumis).

B.2.3. Métallopeptidases à zinc

Il est intéressant de noter que malgré le grand nombre de protéases régulées au niveau transcriptomique (De Gregorio et al., 2001; Irving et al., 2001), aucun membre de cette famille n'a été observé dans les autres approches. Pourtant, dix protéases détectées au cours de notre analyse protéomique sont des métallopeptidases à zinc. Deux groupes de métallopeptidases à zinc se distinguent chez les Mammifères : les molécules de la famille de la néprilysine (NEP) et celles de l'enzyme de conversion de l'angiotensine (ACE). La drosophile possède des homologues dans les deux catégories de peptidases (Isaac et al., 2000).

Chez les Mammifères, NEP est une glycoprotéine membranaire responsable de l'inactivation métabolique de peptides en circulation ou dans le cerveau. Tous les membres de la famille de NEP ont en commun d'être inhibés par le phosphoramidon et le tiorphan. Cette caractéristique, très conservée au cours de l'évolution, a été très utile lors de la caractérisation des molécules possédant une activité néprilysine chez les Invertébrés. Les spécificités de substrat des NEPs de Vertébrés et Invertébrés sont très conservées, coupant la liaison peptidique adjacente à un groupe de résidus hydrophobes se trouvant du côté N-terminal. Le génome de la drosophile code 24 molécules de la famille des NEPs. Le rôle principal de ces molécules est l'inactivation de neuropeptides et d'hormones (Turner et al., 2001).

L'ACE est une dipeptidyl carboxypeptidase responsable de la conversion de l'angiotensine I en angiotensine II, vasoconstricteur puissant, ainsi que de l'inhibition de la bradykinine, peptide vasodilatateur, par coupure des dipeptides C-terminaux. De manière générale, l'ACE possède une forte affinité pour les acides aminés basiques (His, Arg et Lys). L'ACE est une protéine membranaire fortement glycosylée. Elle est caractérisée par deux domaines possédant chacun un site actif, différant l'un de l'autre notamment par la spécificité du substrat. Chez la drosophile, les homologues de l'ACE sont ANCE (angiotensin converting enzyme) et ACER (angiotensin converting enzyme related). Leurs séquences présentent environ 60% de similitude avec les domaines N- et C-terminaux de l'ACE mais se distinguent de cette dernière par le fait qu'elles sont sécrétées, peu glycosylées et ne possèdent qu'un seul site actif.

Chez les Vertébrés et certains Invertébrés, des molécules possédant une activité dipeptidyl peptidase ont été impliquées dans la modulation de la réponse immunitaire. Ainsi, chez les Mammifères, CD26, une dipeptidyl peptidase IV, est impliquée dans la coupure de cytokines et autres peptides fonctionnels, abrogeant ainsi leur activité (Nemoto et al., 1999). De plus, l'angiotensine II régule la réponse immunitaire inflammatoire des Mammifères notamment en activant la prolifération et le recrutement de cellules immunocompétentes ou en stimulant la production de médiateurs de l'inflammation (Suzuki et al., 2003). Des molécules du système rénine-angiotensine (angiotensine II, récepteur 1 de l'angiotensine et ACE) participent également à la modulation de la réponse immunitaire de la sangsue *Theromyzon tessulatum* et de la moule *Mytilus edulis* (Salzet et al., 2001).

Au cours de notre analyse protéomique, dix molécules appartenant à la famille des métallopeptidases, peptidyl dipeptidases, ou possédant une activité proche (alanyl aminopeptidase et carboxypeptidases), ont été détectées. Parmi ces molécules, le taux d'abondance de six est modifié de manière significative après une ou plusieurs infection(s). Deux sont réprimées (produits des gènes *CG6806* et *CG8358*) et quatre surexprimées (issues des gènes *CG10593*, *CG11951*, *CG11034* et *CG4933*) après infection. On retrouve notamment ACER surexprimée suite à l'infection par *B. bassiana* tandis que la concentration de ANCE ne semble pas être modifiée suite aux stimuli étudiés.

La variation d'abondance d'un grand nombre de peptidases de la famille de NEP ou de l'ACE, en plus de leur implication dans des processus immunitaires chez les Vertébrés, permet de spéculer l'importance des molécules de cette famille dans les mécanismes de défense des insectes.

B.3. Inhibiteurs de protéases

Le taux d'abondance de 14 inhibiteurs de protéase est modifié suite aux infections expérimentales : quatre sont des inhibiteurs de protéases à sérine (serpines), six sont des inhibiteurs de type Kunitz, deux de type Kazal et deux sont des inhibiteurs de protéases à cystéine (de la famille des cystatines).

Les serpines sont des protéines d'environ 350 à 500 acides aminés partageant une structure commune, composée de trois feuillets β , de huit à neuf hélices α , et d'une boucle RCL (Reactive Center Loop), située du côté C-terminal, qui contient le site de coupure reconnu par la protéase cible. Contrairement au mécanisme d'inhibition des autres inhibiteurs de protéase de type Kunitz ou Kazal, formant des complexes réversibles avec leur protéase cible, les serpines participent à un complexe covalent ; ce sont des "inhibiteurs suicides" (pour revue, Silverman et al., 2001). Parmi les serpines régulées, on retrouve Necrotic (NEC), un des composants essentiels de la cascade protéolytique conduisant à l'activation du récepteur Toll. Cette protéine est surexprimée 72 h après infection fongique mais pas 6 h après stimulation par les bactéries M. luteus et E. coli. De manière intéressante, une analyse par Western blot utilisant un anticorps anti-NEC a permis de détecter sur gel 2D après infection fongique un deuxième spot immunopositif. Ce spot, situé à 45 kDa alors que la molécule native possède une masse moléculaire de 54 kDa, a été identifié comme étant une forme tronquée de NEC. Une comparaison des profils de masse de la molécule entière ou clivée indique que la coupure se fait dans la partie N-terminale. Cependant, le site exact de coupure n'est pas encore localisé. Ces résultats sont en accord avec des données récemment obtenues au laboratoire et en cours de publication (Pelte et al., soumis).

Deux autres serpines sont surexprimées après infection fongique, le produit du gène *CG18525* (SPN5) et celui du gène *CG6687*. Ce dernier, présent uniquement sur le gel 2D après infection par *B. bassiana*, semble être également tronqué dans la partie N-terminale. En effet, la masse moléculaire observée est inférieure à celle de la molécule

entière d'environ 10 kDa et les 100 premiers acides aminés n'ont pas été détectés par l'analyse par MS.

Enfin, la protéine issue du gène *CG9453* (SPN4), qui n'est détectée que dans l'échantillon d'hémolymphe de drosophiles infectées par le DCV, est réprimée.

Il est intéressant de noter que la plupart des inhibiteurs de protéases identifiées lors des analyses transcriptomiques sont très surexprimés suite aux infections bactériennes (De Gregorio et al., 2001; Irving et al., 2001). Dans notre cas, ces molécules sont plutôt détectées dans l'hémolymphe de drosophiles après l'infection fongique. Cependant seul le niveau d'expression de la SPN5 a été trouvé régulé dans les deux types d'approches et peut être comparé ; leur niveau de variation semble assez corrélé.

Il est important de souligner ici que seule l'approche par électrophorèse 2D permet la détection de molécules clivées et plus largement de molécules possédant des modifications post-traductionnelles. En effet, la différence observée entre la masse et/ou le pI théorique d'une protéine et ceux observés sur le gel 2D donne une première indication de la modification. De plus, le taux de recouvrement obtenu lors de l'identification par spectrométrie de masse MALDI-TOF, souvent très bon, permet d'étudier les peptides portant la modification ou d'estimer si la coupure à lieu dans la partie N- ou C-terminale. L'approche ICAT se limite à l'analyse des modifications portées par les peptides possédant des cystéines.

L'importance du système protéases/inhibiteurs de protéase dans la réponse immunitaire de la drosophile n'est plus à démontrer. De nombreux composants impliqués dans la cascade d'activation extracellulaire de la voie Toll sont encore à découvrir. Les molécules surexprimées suite aux infections par les bactéries à Grampositif et/ou les champignons sont des candidats intéressants pour lesquels une analyse approfondie devra être menée.

B.4. Molécules de reconnaissance du pathogène

B.4.1. Gram-negative binding proteins

La molécule la plus fortement induite (facteur 70) lors de l'étude par gels 2D correspond à une "GNBP-like" (Gram-Negative Binding Protein-like), produit du gène *CG13422*. Cette protéine ne possède pas de cystéine ; elle n'a donc pas été détectée par l'approche ICAT. Cette protéine appartient à la famille des GNBPs mais diffère des autres membres par manque du domaine β -glucanase. La première GNBP a été isolée chez *Bombyx mori* où son implication dans la liaison aux bactéries à Gram-négatif lui a valu son nom (Lee et al., 1996). Depuis, de nombreuses études ont montré que d'autres protéines de cette famille n'étaient pas impliquées exclusivement dans la reconnaissance des bactéries à Gram-négatif (Kim et al., 2000), voire plus du tout. En effet, dans l'introduction de ce manuscrit, l'implication de deux GNBPs dans la voie Toll a été mentionnée. Une mutation dans le gène *osiris*, codant GNBP1 (produit du gène *CG6895*), bloque l'activation de la voie Toll en réponse aux bactéries à Gram-positif mais pas aux bactéries à Gram-négatif ni aux champignons (Gobert et al., 2003). Une mutation dans le gène *hades (CG5008*), codant GNBP3, a permis de montrer l'implication de GNBP3 dans la reconnaissance des champignons (Gottar et al., 2002).

Des analyses transcriptomiques ont révélé que le niveau d'expression de l'ARN messager du gène *CG13422* est augmenté 6 h après les infections par *E. coli* ou *M. luteus* aussi bien que 72 h après stimulation par *B. bassiana* (De Gregorio et al., 2001; Irving et al., 2001). Encore une fois, nous remarquons qu'au niveau des protéines les résultats divergent. En effet, le taux de variation de cette "GNBP-like" est augmenté suite à l'infection fongique mais est invariable après stimulation par *E. coli* ou *M. luteus*.

Le produit du gène *CG13422*, phylogénétiquement plus proche de GNBP3 (Figure 27), pourrait être un autre membre de la famille des GNBPs impliqué dans la voie de signalisation activée par les champignons.

La surexpression très importante de cette molécule par microarrays (facteur cinq) et surtout par gels 2D (facteur 70) après infection fongique a amené l'équipe de Dominique Ferrandon à initier une étude fonctionnelle de cette "GNBP-like". Un



Figure 27 : Structure schématique et distribution phylogénétique des GNBPs. Représentation schématique de la structure des GNBPs (A), caractérisée par un domaine de liaison au β -1,3 glucane et un domaine catalytique β -glucanase. La distribution phylogénétique du domaine de fixation au β -1,3 glucane de certaines β GRP et GNBPs est représentée en (B). La GNBP-like est encadrée en jaune ; les autres GNBPs de drosophile sont encadrées en vert. L'alignement a été réalisé par Marie Gottar au laboratoire à l'aide du programme MegAlign en utilisant la méthode ClustalW (matrice des séries de Gonnet). Il montre que le produit du gène CG13422 est phylogénétiquement plus proche de GNBP3 que des deux autres GNBPs.

anticorps dirigé contre la protéine a été obtenu et a permis de confirmer par Western blot sa forte surexpression 72 h après infection par *B. bassiana* mais pas 6 h après les infections par *M. luteus* ou *E. coli*. Des lignées de mouches transgéniques invalidées pour le gène CG13422 ont été établies et la survie de ces mutants vis-à-vis des différentes infections est en cours d'étude. Des drosophiles surexprimant le gène CG13422 seront également étudiées.

B.4.2. Peptidoglycan Recognition Proteins

Les PGRPs sont des molécules de reconnaissance du peptidoglycane (PGN). Le PGN est le composant principal de la paroi des bactéries à Gram-positif. On le retrouve également dans la paroi des bactéries à Gram-négatif mais minoritaire (20%) et situé sous une large couche de LPS. Les PGRPs présentent une similitude de séquence avec le lysozyme du bactériophage T7, qui hydrolyse le PGN. Certaines PGRPs sont dépourvues d'activité enzymatique en raison d'une substitution d'acides aminés dans leur domaine catalytique et servent donc uniquement de molécules de reconnaissance (Dziarski, 2004).

La famille des PGRPs chez la drosophile compte 13 membres dont sept sont des PGRPs courtes (PGRP-S (Small)) et six des PGRPs longues (PGRP-L). Les PGRP-S sont caractérisées par la présence d'un domaine PGRP, d'un peptide signal, suggérant qu'il s'agit de molécules extracellulaires, et ne possèdent pas de domaines transmembranaires. La sous-famille des PGRP-L est plus hétérogène, puisqu'elle regroupe des molécules hydrophobes potentiellement transmembranaires, intracytoplasmiques ou sécrétées (Royet, 2004).

Lors d'un crible génétique réalisé au laboratoire, l'équipe de Julien Royet a identifié une lignée mutante, *Semmelweis* (*seml*), dans un gène codant une PGRP courte, PGRP-SA (Michel et al., 2001). Les mutants *PGRP-SA*^{seml} présentent une sensiblité accrue aux infections par des bactéries à Gram-positif et l'expression de la Drosomycine est abolie suite à des infections par ce type de bactéries. Ces individus réagissent cependant comme des mouches sauvages aux infections fongiques ou par des bactéries à Gram-négatif. PGRP-SA est une protéine circulante qui initie une voie de signalisation

spécifique aux infections par des bactéries à Gram-positif aboutissant à l'activation de Toll (Gobert et al., 2003; Michel et al., 2001).

Lors de l'analyse de l'hémolymphe des drosophiles par l'approche ICAT, PGRP-SA a été détectée uniquement dans l'échantillon infecté par le DCV. Sa concentration est invariable suite à cette infection. Comme plus de 30% des protéines détectées par ICAT, PGRP-SA a été identifiée par spectrométrie de masse grâce à l'analyse d'un seul peptide, ce qui peut expliquer le fait de ne pas la détecter dans les autres échantillons.

Un autre membre de la famille des PGRP-S, PGRP-SB1, est mis en évidence par l'approche ICAT. L'abondance relative de cette molécule est augmentée de manière significative suite à la stimulation par les deux bactéries à Gram-positif et très fortement suite à l'inoculation d'*A. tumefaciens* (Gram-négatif). Le niveau d'expression des transcrits est en accord avec celui de la protéine (De Gregorio et al., 2001; Irving et al., 2001). Il a été montré que PGRP-SB1 possède une expression inductible et est exprimée principalement par les cellules du corps gras, organe où est également synthétisée la plupart des peptides antimicrobiens (Steiner, 2004). Son implication dans la réponse immunitaire devra être étudiée plus particulièrement.

B.4.3. Facteurs de croissance des disques imaginaux

Cette famille de molécules putatives de la reconnaissance du pathogène est bien représentée dans cette analyse protéomique. Trois facteurs de croissance des disques imaginaux (IDGFs, Imaginal Disc Growth Factors) voient leur taux d'abondance augmenter suite à l'infection par *B. bassiana* en commun avec *A. tumefaciens* (IDGF2 et 3) ou par *B. bassiana* seule (IDGF4).

Le génome de la drosophile code au moins cinq IDGF dont trois sont organisés en cluster (IDGF1, 2 et 3). Ces molécules dérivent des chitinases pour lesquelles l'activité catalytique aurait été perdue mais la conservation partielle du site permet encore la liaison aux oligosaccharides (Varela et al., 2002). Il est intéressant de noter qu'une molécule de la famille des chitinases (produit du gène *CG5210*) est détectée dans l'analyse par ICAT. Elle est surexprimée spécifiquement suite à l'infection fongique. La surexpression de ces molécules après l'infection par *B. bassiana* suggère leur implication dans les mécanismes de défense contre les champignons. De plus, le fait que les IDGFs, en comparaison avec les chitinases, aient perdu l'activité catalytique mais conservé la liaison aux oligosaccharides en font de bons candidats pour la reconnaissance des champignons.

B.5. Molécules similaires aux protéines du complément

Deux protéines appartenant à la superfamille des C3/ α 2-Macroglobuline, TEP2 et TEP4 (Thiolester Containing Protein 2 et 4), sont détectées lors de cette approche protéomique.

Chez la drosophile, six gènes codent des TEPs (dTEP1 à dTEP6) dont quatre possèdent un site thioester parfaitement conservé (Figure 28). Trois d'entre eux (dTEP1, 2 et 4) voient leur transcription augmentée après infection bactérienne (Boutros et al., 2002; De Gregorio et al., 2001; Irving et al., 2001; Lagueux et al., 2000). Toutes les TEPs de drosophile possèdent un peptide signal et sont secrétées dans l'hémolymphe (Meister and Lagueux, 2003). Chez le moustique A. gambiae, plusieurs membres de la famille des TEPs ont été identifiés (15 gènes). Le gène *aTEP1* code une glycoprotéine spécifique des hémocytes qui agit comme une opsonine et favorise la phagocytose de certaines bactéries à Gram-négatif (Levashina et al., 2001). De plus la liaison de aTEP1 avec la bactérie se fait au niveau du site thioester, aboutissant à la coupure de aTEP1. De manière intéressante, le spot correspondant à TEP4 sur gel 2D, induit après infection par le champignon B. bassiana, représente une forme tronquée de la molécule. En effet, la masse moléculaire observée sur le gel est d'environ 97 kDa alors que la protéine entière possède une masse moléculaire de 168 kDa. Le recouvrement de la molécule observé lors de l'analyse par spectrométrie de masse englobe les résidus 1 jusqu'à 639, indiquant ainsi que ce fragment correspond à la partie N-terminale de la protéine. Nous n'avons pas trouvé le site exact de coupure mais nous savons qu'il se fait après le résidu 639. La protéine TEP4 détectée par l'approche par ICAT, surexprimée suite à l'infection par *E. faecalis*, correspond plus probablement à la forme entière de la molécule puisque les peptides ayant servi à son identification mènent jusqu'au résidu 995.

Une autre étude différentielle protéomique, réalisée par 2D-DIGE (twodimensional difference gel electrophoresis), montre que TEP2 est surexprimée dans l'hémolymphe des larves de drosophiles suite à une stimulation par du LPS (Vierstraete



Figure 28 : Structure schématique de la région thioester des TEPs. La protéine TEP4 (168 kDa) est schématisée. La partie N-terminale observée par spectrométrie de masse est représentée en bleu. Selon la masse moléculaire observée sur le gel 2D (97 kDa), le site de clivage de la protéine se situe après le résidu 639.

et al., 2004). La masse apparente du spot correspondant est également inférieure à la masse de la protéine entière indiquant que TEP2 serait également coupée dans l'hémolymphe des larves. Dans l'approche par ICAT, TEP2 est également détectée et son abondance augmente après infection par les deux bactéries à Gram-positif testées.

L'étude des protéines à domaine thioester suggère qu'elles jouent un rôle important dans les défenses antimicrobiennes. L'équipe de Marie Lagueux au laboratoire a étudié cette famille de molécules. Un anticorps anti-TEP4 a été construit dans le but de rechercher le fragment complémentaire de la protéine contenant le site de liaison. Une analyse différentielle par gels mono et bidimensionnels, couplée à un Western blot, sur des extraits de drosophiles contrôles ou mutantes pour *dTEP4* et infectées par *B. bassiana* a été initiée. Malheureusement, la faible quantité de matériel issue des extraits n'a pas permis d'obtenir des gels exploitables.

C. Nouveaux candidats de la réponse immunitaire

C.1. Molécules sensorielles

L'étude protéomique a permis de mettre en avant de nombreuses molécules de la famille des OBPs (Odorant Binding Proteins). De par leur localisation spécifique dans les tissus sensoriels, ces molécules sont connues pour leur rôle dans la prise en charge de molécules odorantes ou de phéromones. Récemment, certains membres de cette famille ont été détectés dans des tissus non sensoriels où ils agissent comme transporteurs de peptides hydrophobes (Graham and Davies, 2002).

Chez la drosophile, on compte 51 gènes codant des molécules de cette famille (Hekmat-Scafe et al., 2002). Une étude réalisée au laboratoire par Laurence Sabatier et l'équipe de Jean-Luc Imler a permis de mettre en évidence deux phérokines de la famille des OBPs impliquées dans la réponse antibactérienne et antivirale de la drosophile (PHK-2 et PHK-3), suggérant un rôle dans la reconnaissance et/ou la neutralisation de microorganismes (Sabatier et al., 2003). L'une de ces phérokines (PHK-2, produit du gène *CG11390*) a été trouvée par l'approche ICAT comme étant surexprimée dans presque toutes les infections mais bien plus fortement après l'infection virale. D'autres membres de cette famille ont été détectés dans notre analyse (cf. Tableaux III et IV). De

manière intéressante, trois des quatre OBPs détectées sont surexprimées après l'infection virale. En ce qui concerne l'OBP 99c, son taux d'abondance est plutôt réprimé après infection fongique et par *S. saprophyticus*. De telles observations indiquent que cette famille de protéines mérite d'être plus amplement étudiée dans le contexte de la réponse immunitaire.

C.2. Peptidylglycine α-Hydroxylating Monooxygenase

Cette molécule a été détectée comme induite après l'infection fongique dans l'approche par gels 2D. Cette enzyme est impliquée dans l'amidation C-terminale de peptides, modification spécifique et nécessaire de nombreux neuropeptides et hormones peptidiques (Jiang et al., 2000).

Plusieurs peptides antimicrobiens de différentes familles structurales sont amidés à leur extrémité C-terminale (Boulanger et al., 2002; Bulet et al., 1999; Lorenzini et al., 2003; Silva et al., 2000). Chez la drosophile la Cécropine possède cette modification post-traductionnelle et il est supposé que la Diptéricine est également amidée comme chez *Phormia* (Bulet et al., 1995). Plusieurs exemples dans la littérature montrent une augmentation de l'activité antimicrobienne grâce à l'amidation ou au contraire, une diminution si le groupement amide est supprimé (Mor et al., 1994; Subbalakshmi et al., 2001). Un des rôles les plus importants attribué à cette modification C-terminale est une stabilité supérieure des peptides amidés vis-à-vis des protéases notamment les carboxypeptidases. Parmi les DIMs sécrétés dans l'hémolymphe de drosophile après piqûre septique, plusieurs possèdent cette modification (DIMs 1, 2 et 4, Tableau I (Uttenweiler-Joseph et al., 1998)). Notons qu'aucune des études transcriptomiques ne mentionne la variation de cette molécule suite à une infection microbienne.

Ainsi, l'apparition de PHM sur les cartes 2D après l'infection fongique nous permet de proposer un rôle essentiel de cette enzyme dans la maturation de molécules effectrices.

C.3. Molécules de liaison à la phosphatidyléthanolamine

Le produit du gène *CG17919*, détecté uniquement au cours de l'analyse par gels 2D, est induit 6 h après l'infection par *M. luteus*. Cette molécule contient un domaine "phosphatidylethanolamine binding" (PEB). Une autre protéine contenant un domaine PEB, issue du gène *CG18594*, détectée par une analyse protéomique sur les larves de drosophiles, est surexprimée suite à une stimulation par du LPS (Vierstraete et al., 2004).

Les protéines PEB (PEBP) appartiennent à une famille très conservée possédant des homologues chez de nombreux organismes (Simister et al., 2002). Malgré cela, peu de choses sont connues sur le rôle des membres de cette famille et de nombreuses fonctions sont suggérées comme la liaison des lipides (Schoentgen and Jolles, 1995), l'inhibition de sérine protéases (Hengst et al., 2001), ou la liaison de molécules sensorielles (Pikielny et al., 1994). De plus, il a été montré qu'une PEBP de Mammifères (RKIP pour Raf-1 Kinase Inhibitor Protein) inhibe l'activation de NF- κ B, le facteur de transcription nécessaire pour la surexpression d'un grand nombre de gènes impliqués dans la réponse inflammatoire, aux infections bactériennes ou virales et à d'autres stimuli liés au stress (Yeung et al., 2001). Ces données laissent supposer un rôle important des molécules de la famille des PEBP, particulièrement dans la régulation de nombreuses voies de signalisation dépendantes de NF- κ B chez les Vertébrés et les Invertébrés.

Encore une fois, cette famille de molécule n'était pas représentée lors des analyses transcriptomiques. Seules des approches protéomiques ont permis de mettre en évidence la régulation de ce type de protéines.

C.4. Molécules de la famille des immunophilines

Chez les Mammifères, les FKBPs (FK-506 Binding Proteins) forment, avec les cyclophilines (CyPs), la grande famille des immunophilines. Ces deux sous-familles, bien que possédant une structure très différente, partagent deux propriétés communes : (1) une activité rotamase de type peptidyl-prolyl *cis/trans* isomérase (PPIase), catalysant le repliement et le réarrangement de protéines, et (2) la capacité à fixer des

immunosuppresseurs : les FKBPs peuvent fixer le FK506 (également nommé tacrolimus) et la rapamycine tandis que les CyPs sont capables de fixer la cyclosporine A (CsA).

Les deux complexes FKBP/FK506 et CyP/CsA régulent des voies signalétiques effet. qui sont calcium-dépendantes. En la cible pour le complexe immunophiline/immunosuppresseur calcineurine. sérine/thréonine est la une phosphatase qui est calcium- et calmoduline-dépendante. La calcineurine régule la translocation de facteurs nucléaires de cellules T activées, étape essentielle pour l'expression des gènes de cytokines dans les cellules T activées (Matsuda et al., 2000). De plus, des études récentes ont montré que la CsA et le FK506 permettent de bloquer l'activation des voies JNK et p38 (Matsuda and Koyasu, 2003).

Lors de l'analyse protéomique, une FK506-Binding Proteins, FKBP-2 (également nommée FKBP-12), voit son abondance relative augmentée suite à une infection par *A. tumefaciens* et par *B. bassiana*. La cyclophiline 1 est également détectée mais son abondance est invariable à la suite des infections testées.

Une autre molécule de la famille des PPIase est trouvée surexprimée, 24 h après infection par un mélange de bactéries *E. coli* et *M. luteus*, dans une étude transcriptomique sur la sangsue *Theromyzon tessulatum* (Lefebvre et al., 2004). De même, lors d'une étude protéomique récente sur les larves de drosophile, la parvulin-like PPIase est surexprimée après infection naturelle bactérienne (de Morais Guedes et al., 2005). Ces molécules, impliquées dans différents processus cellulaires importants chez de nombreux organismes, pourraient donc participer aux mécanismes de défense des Invertébrés.

C.5. Molécules de liaison au calcium

Lors de notre étude, il est intéressant de noter que plusieurs protéines régulées possèdent un domaine de liaison au calcium. Le calcium est l'un des messagers intracellulaires les plus anciens et les plus utilisés. Peu coûteux énergétiquement, il est impliqué dans de très nombreux processus biologiques.

C.5.1. Sarcoplasmic Calcium-binding Protein et SCP-related protein

La SCP1 est surexprimée spécifiquement après l'infection par *S. saprophyticus*. Les molécules de la famille des SCPs, impliquées dans la morphogenèse musculaire, ont en commun de lier le calcium grâce à un domaine EF-hand (Hermann and Cox, 1995). Une autre molécule issue de l'analyse par ICAT (produit du gène *CG6426*, surexprimée après infection par *A. tumefaciens*, *B. bassiana* et *E. faecalis*) possède ce domaine particulier.

Les domaines EF-hand sont caractérisés par la présence d'un motif héliceboucle-hélice dans leur séquence d'acides aminés. La liaison du calcium au domaine entraîne un changement de conformation conduisant à la régulation d'un large spectre de protéines cibles. Le plus souvent, le déplacement d'une hélice par rapport à l'autre permet de créer une zone hydrophobe à la surface de la protéine ; mais la grande diversité de conformations permet de proposer de nombreuses interfaces favorisant l'interaction avec différentes enzymes (kinases, phosphatases), des récepteurs ou des lipides (Bhattacharya et al., 2004; Yap et al., 1999). Ces caractéristiques suggèrent que les protéines possédant ce type de domaine peuvent jouer un rôle clé notamment dans la régulation des cascades de signalisation.

C.5.2. Alpha-amylase

L'α-amylase de drosophile possède également un domaine de liaison au calcium. Cette enzyme, très répandue dans l'ensemble du monde vivant, est impliquée dans l'hydrolyse des glucides complexes en sucres plus simples.

Chez la drosophile, deux gènes codent une α -amylase : le *CG18730* code l' α amylase proximale (AMY-p) et le *CG17876* code l' α -amylase distale (AMY-d). Les deux protéines sont détectées lors de notre analyse protéomique. AMY-d est surexprimée cinq fois sur les gels 2D suite aux infections par *B. bassiana* et *M. luteus*. AMY-p est l'une des deux protéines surexprimées très fortement après toutes les infections testées lors de l'approche ICAT.

Une étude récente met en avant l'importance des motifs de sucres dans la reconnaissance des pathogènes chez la drosophile. On supposait que chez la drosophile, le LPS des bactéries à Gram-négatif était responsable de l'activation de la voie IMD. Or, il a été montré récemment que le LPS n'induit pas de réponse quand il est injecté dans la drosophile et qu'en fait, le système immunitaire de la drosophile est capable de faire la différence entre le PGN des bactéries à Gram-positif et celui des bactéries à Gramnégatif (Kaneko et al., 2004; Leulier et al., 2003). Le PGN est formé par la répétition d'un motif de deux sucres aminés, la N-acétylglucosamine (GlcNAc) et l'acide Nacétylmuramique (MurNAc) (Figure 29). Le résidu MurNAc porte un tétrapeptide dont le troisième acide aminé diffère chez les bactéries à Gram-positif et les bactéries à Gram-négatif. Dans le PGN des bactéries à Gram-positif, cet acide aminé est une L-Lysine, tandis que dans celui des bactéries à Gram-négatif, on trouve un résidu diaminopimélique (DAP). On parle de Lys-PGN et de DAP-PGN. La présence ou l'absence du résidu DAP dans le PGN apparaît comme un élément majeur de la discrimination entre bactéries à Gram-négatif et bactéries à Gram-positif (Leulier et al., 2003). Il a été montré récemment qu'en plus de ce résidu, des éléments de la chaîne sucrée pourraient participer à cette discrimination. En effet, une équipe du laboratoire de Bruno Lemaitre a mis en évidence que la forme "1,6 anhydro" du MurNAc, présente naturellement à l'extrémité de la chaîne peptidique semble jouer un rôle important dans la reconnaissance du PGN des bactéries à Gram-négatif. En effet, si les polymères de PGN sont dégradés par des enzymes telles que les muramidases, qui coupent la liaison β-1,4 entre les résidus MurNAc et GlcNAc, l'activation de la voie IMD est diminuée (révélée par une diminution du niveau d'expression de la Diptéricine). En revanche, si le PGN est clivé par une transglycosidase de E. coli, SltY, qui génère des muropeptides similaires à ceux obtenus après coupure par les muramidases, mais avec la particularité de former la liaison "1,6 anhydro" à chaque fois, la voie IMD est alors activée efficacement (Stenbak et al., 2004).

Il reste encore à éclaircir comment la voie IMD peut reconnaître un constituant bactérien "enfoui" sous une couche de LPS. On peut imaginer que du PGN est relargué lors des divisions bactériennes, ou que la paroi bactérienne est altérée par d'autres mécanismes provoquant la libération de PGN (Leulier et al., 2003).



Figure 29 : Structure schématique du peptidoglycane. Le peptidoglycane est un polymère complexe composé d'une longue chaîne formée par la répétition de deux sucres aminés : la N-acétylglucosamine et l'acide N-acétylmuramique. Le troisième acide aminé L-Lys de la chaîne latérale est retrouvé dans le PGN des bactéries à Grampositif tandis que le L-DAP est retrouvé surtout dans le PGN des bactéries à Gramnégatif et des bacilles. *D'après Royet 2004*.

Ces résultats permettent d'imaginer l'importance des processus de dégradation des sucres dans la formation de motifs de reconnaissance des pathogènes où l'activité glucanohydrolase de l' α -amylase pourrait trouver sa place. Le fait que cette protéine soit surexprimée suite à un large spectre d'infections est en accord avec cette hypothèse.

D'autres études protéomiques sur la drosophile mettent en évidence des molécules de liaison du calcium : par exemple, une protéine de la famille des régucalcines est libérée dans l'hémolymphe de larves en réponse à une injection par du LPS (Vierstraete et al., 2004). L'α-glucosidase, contenant un domaine EF-hand, et la calmoduline sont également modulées dans la fraction nucléaire de cellules mbn-2, 30 min après traitement par du LPS (Loseva and Engstrom, 2004).

Peu de choses sont connues quant à la dépendance de la réponse immunitaire de la drosophile vis-à-vis du calcium mais le fait de trouver plusieurs protéines de ce groupe suggère un intérêt particulier.

C.6. Molécules antioxydantes et métabolisme des lipides

De manière naturelle, le métabolisme à tendance à produire des dérivés radicalaires lors de divers processus physiologiques (phagocytose, mélanisation, encapsulation mélanotique, réactions énergétiques, etc.). Ces radicaux libres, à dose raisonnable, sont utiles pour l'organisme et peuvent intervenir dans des mécanismes de défense ou comme médiateurs tissulaires. Cependant, l'accumulation ou la production excessive de ces dérivés résulte en une toxicité contre laquelle l'organisme va devoir lutter. En effet, les radicaux libres peuvent causer des dommages cellulaires notamment sur les lipides insaturés, les protéines ou les acides nucléiques. Ainsi, chaque organisme possède une balance antioxydants/prooxydants qui doit être régulée (Le Bourg, 2001).

L'un des processus produisant le plus de radicaux libres est le métabolisme des acides gras. Lors de notre analyse protéomique 12 molécules impliquées dans ce métabolisme ou dans la liaison des lipides sont détectées. Le taux d'abondance de certaines est diminué tandis que d'autres voient leur concentration augmentée suite aux infections testées. D'une part, la répression de molécules du métabolisme des lipides pourrait suggérer que l'organisme agressé tente de limiter la production de dérivés radicalaires libérés lors de ce processus. D'autre part, la surexpression de molécules impliquées dans le transport des lipides pourrait faire partie d'un mécanisme de détoxification par élimination des lipides attaqués par les radicaux libres, renforçant ainsi l'action de molécules endogènes antioxydantes. Parmi les molécules de liaison aux lipides, le taux d'abondance d'une protéine de liaison au rétinoïde et aux acides gras (RFABP pour Retinoid and Fatty Acid Binding Protein, produit du gène *CG11064*) est pourtant diminué suite à l'infection par *S. saprophyticus*. Il a été suggéré que cette protéine, qui appartient à la famille des proapolipophorines, est impliquée dans le transfert d'électrons, participant ainsi à la production de dérivés radicalaires (Duncan et al., 1999). Cette information conforte notre première hypothèse concernant la limitation de la production de molécules impliquées dans des processus libérant des radicaux libres.

Chez la drosophile, de nombreuses protéines sont impliquées dans la détoxification des radicaux libres (catalase, thioredoxines), possèdent une activité antioxydante (superoxide dismutases, peroxidases) ou participent à la réparation de dommages éventuellement causés par ces dérivés (enzymes de réparation de l'ADN). Lors de notre étude par ICAT, au moins dix molécules ayant une activité antioxydante ont été détectées dont six sont surexprimées de manière significative suite à différentes infections. Certaines de ces protéines ont également été trouvées surexprimées dans d'autres analyses protéomiques chez la drosophile (de Morais Guedes et al., 2005; Loseva and Engstrom, 2004; Vierstraete et al., 2004). De manière intéressante, Ha et collaborateurs ont récemment étudié l'importance dans la réponse immune épithéliale d'une molécule antioxydante, qu'ils ont nommés IRC (Immune-Regulated Catalase). En effet, des mouches présentant une expression réduite pour cette protéine, produit du gène CG8913, sont plus sensibles à des infections par ingestion de nourriture contaminée (Ha et al., 2005). IRC, que nous retrouvons surexprimée suite à toutes les infections exceptée par S. saprophyticus, semble donc jouer un rôle majeur dans l'interaction hôte-pathogène et dans la protection de l'infection au niveau de l'intestin de drosophiles.

C.7. Molécules impliquées dans le métabolisme du fer

Le taux d'abondance de deux molécules impliquées dans le métabolisme du fer, la transferrine (TSF) et la ferritine, varie suite à une infection expérimentale. Plusieurs formes modifiées de la transferrine sont visualisées sur gel 2D et apparaissent sous forme d'une série de spots espacés les uns des autres d'environ 0,2 unités de pH (72 kDa, pI 6,8 à 7,5). Le taux d'abondance de ces isoformes est inchangé six heures après les infections bactériennes testées (*E. coli* et *M. luteus*) mais toutes sont fortement surexprimées après l'infection fongique. En revanche, le taux d'abondance des deux sous-unités de la ferritine diminue après infection fongique.

Dans le génome de la drosophile trois gènes codent des transferrines (TSF1, TSF2 et TSF3 encodées par les gènes *CG6186*, *CG10620* et *CG3666*, respectivement). Les pI et masses moléculaires des trois TSF sont compatibles avec les gammes utilisées pour les gels 2D et toutes les trois possèdent des cystéines, cependant seule la TSF1 a été détectée lors de nos différentes expériences. L'analyse par puces à ADN montre pourtant que les transcrits de deux TSF (TSF1 et TSF2) sont surexprimés suite à des stimuli immuns : TSF1 est surexprimée après stimulation fongique et TSF2 est surexprimée uniquement après piqûre septique (De Gregorio et al., 2001). Le fait de ne pas retrouver les autres TSF dans l'hémolymphe des drosophiles peut être dû au fait que la concentration des TSF2 et TSF3 est trop faible pour être détectée par nos approches ou que les temps d'infection étudiés ne sont pas suffisamment longs pour obtenir une surexpression amenant les TSF à un niveau détectable dans l'hémolymphe.

La TSF, composant majeur du sang des Vertébrés, est impliquée dans le transport du fer tandis que la ferritine est impliquée dans son stockage. Chez les Invertebrés, le rôle de ces deux molécules n'est pas très bien connu mais la TSF et la ferritine sembleraient toutes deux être impliquées à la fois dans le stockage et le transport du fer. Concernant les pathogènes, les molécules de liaison au fer participent au processus d'interaction avec l'hôte. Par exemple, il a été montré que des molécules impliquées dans le transport du fer chez *Serratia marcescens* sont essentiels pour sa virulence lors de l'infection du nématode *Caenorhabditis elegans* (Kurz et al., 2003).

Par ailleurs, il a été observé chez la drosophile et le moustique *Aedes*, que la concentration de TSF est modifiée dans certaines conditions de stress ou d'infection. En

effet, chez ces insectes, la TSF est surexprimée suite à une infection bactérienne (Yoshiga et al., 1999; Yoshiga et al., 1997) et également lors de l'encapsulation de vers filaires (Beerntsen et al., 1994). Ceci est en accord avec le rôle bactériostatique de la TSF qui, en séquestrant le fer circulant, le rend indisponible pour la croissance des bactéries infectieuses. De plus, il a été montré que la lactoferrine, molécule cationique de la même famille que la TSF, possède des fonctions antimicrobiennes et prévient la formation de biofilms par *Streptococcus mutans* et *Pseudomonas aeruginosa* (Berlutti et al., 2004; Singh et al., 2002). Les résultats de notre analyse constituent le premier exemple de régulation de la TSF d'insecte suite à une infection fongique, suggérant que cette molécule de liaison au fer pourrait également participer à la réponse immunitaire contre les champignons.

Une analyse par PCR quantitative sur des drosophiles mutantes affectées à différents niveaux des voies de signalisation Toll ou IMD a permis de confirmer que la régulation de la TSF est liée à la voie Toll puisque des mutants *spz* ou *dif* ne surexpriment plus la TSF après infection par *B. bassiana*, tandis que les mutants pour la voie IMD se comportent normalement. De manière surprenante, les drosophiles mutantes pour des protéines en amont de SPZ (*psh* et *GNBP3*) se comportent comme des mouches sauvages pour l'expression de la TSF. Cependant, l'étude des doubles mutants *psh/GNBP3* semble montrer une diminution forte de l'expression de la TSF après infection par le champignon. Ce résultat préliminaire intéressant, qui pourrait mettre en évidence une complémentarité de fonction entre les voies de signalisation des champignons passant par PSH et GNBP3 lors de la régulation de la transferrine, mérite d'être approfondi.

L'inhibition de la traduction de la TSF a été réalisée *in vivo* par injection d'ARN double brin et la survie des drosophiles a été testée après infection par différentes bactéries. Les drosophiles injectées ne semblent pas montrer de susceptibilité accrue aux infections par les pathogènes testés. Cependant, il faut noter que l'inhibition de la TSF par injection d'ARN double brin n'est pas totale (80%) et que les transcrits résiduels peuvent suffire à maintenir l'activité de la TSF.

En plus de la surexpression des différentes formes de la TSF, l'infection fongique pourrait également induire la coupure de cette molécule. En effet, plusieurs

spots induits après infection par *B. bassiana* sont identifiés comme des fragments de TSF. Nous avons étudié la spécificité du clivage de la TSF et montré qu'il n'est pas dû à l'action d'une protéase fongique mais probablement à un processus endogène de la drosophile.

Jusqu'alors, aucune étude mentionnant le rôle de fragments de TSF dans la réponse immunitaire des Invertébrés n'avait été décrite. Chez les Vertébrés, le lien potentiel entre les mécanismes de défense et des fragments de TSF a été évoqué lors de plusieurs publications. Deux études montrent que des fragments de l'ovotransferrine de poulet possèdent des activités antimicrobiennes. Un fragment cationique présentant un motif structural proche des Défensines d'insecte, OTAP-92 (OvoTransferrin Antimicrobial Peptide-92), possède une activité contre des bactéries à Gram-négatif en causant des dommages membranaires (Ibrahim et al., 2000). Deux autres fragments de l'ovotransferrine possèdent une activité antivirale (Giansanti et al., 2005). Une troisième étude, réalisée chez le poisson rouge, indique que des formes tronquées de TSF induisent la production d'oxyde d'azote par des macrophages après stimulation par du LPS, contrairement à la TSF entière (Stafford et al., 2001). Les auteurs présentent un modèle d'activation de la réponse antimicrobienne des macrophages de poissons par une molécules non cytokine (i.e. la TSF) (Stafford and Belosevic, 2003) et indiquent que ce mode de réponse existe également chez les Mammifères (Stafford et al., 2004). Récemment, une étude chez le moustique A. aegypti montre que la TSF de cet insecte est également coupée de manière reproductible chez les larves de quatrième stade et dans les ovaires des adultes après la prise d'un repas de sang (Harizanova et al., 2005). De plus, la TSF de moustique possède en 5' plusieurs sites putatifs de liaison au facteur de transcription NF-kB, médiateur de la réponse immunitaire des insectes. Nous ne savons toujours pas si les fragments de TSF possèdent une activité biologique, mais ces données, en combinaison avec nos résultats, suggèrent fortement un rôle de ces produits de clivage dans la réponse immunitaire des diptères.

CONCLUSION ET PERPECTIVES

CONCLUSION ET PERSPECTIVES

Le but de mon travail de thèse était de réaliser, en complément des approches transcriptomiques et peptidomiques déjà effectuées au laboratoire, une analyse protéomique des molécules de l'hémolymphe de drosophiles avant et après infection. Les analyses transcriptomiques ont permis d'obtenir un grand nombre de gènes candidats dont le niveau d'expression varie suite aux infections. Cependant, il était essentiel d'effectuer une analyse directe du taux d'abondance des protéines dans l'hémolymphe au vu des divergences souvent obtenues entre le niveau d'expression des transcrits et celui des molécules correspondantes. De plus, l'analyse protéomique permet de cibler les protéines dans leur état fonctionnel à un moment donné, impliquant souvent des modifications post-traductionnelles essentielles pour l'activité de ces protéines. L'approche que j'ai mise en place au laboratoire est basée sur deux stratégies complémentaires : l'une repose sur l'électrophorèse bidimensionnelle et l'autre sur le marquage isotopique différentiel des protéines. Nous avons mis en évidence de nombreuses molécules dont la concentration relative varie de manière significative après différents types d'infection expérimentale.

Un problème qui se pose lorsque l'on étudie un processus aussi vaste que la réponse immunitaire est lié à l'exhaustivité de l'analyse. En effet, sans même parler de l'énorme diversité des protéines d'un organisme, chacune des approches utilisées au cours de ma thèse possède ses limites propres, liées notamment au seuil de détection ou à la nature des protéines étudiées. Ainsi, comme nous l'avons évoqué, l'électrophorèse 2D n'est pas adaptée à l'analyse des protéines de faibles masses moléculaires (<15 kDa) ou de masses moléculaires élevées (>150 kDa), des protéines présentes en faible abondance ainsi que des protéines hydrophobes. Ces inconvénients sont partiellement résolus grâce à l'approche ICAT, mais celle que nous avons utilisée ne cible que les protéines possédant des cystéines. Cependant, il semble évident que la détection de certaines molécules échappent à ces deux stratégies (par exemple les protéines hydrophobes et ne possédant pas de cystéines) alors qu'elles peuvent être impliquées dans des mécanismes importants de la réponse immunitaire. La détection d'un

maximum de protéines est donc essentielle pour la compréhension d'une problématique telle que l'étude des molécules impliquées dans la réponse immunitaire de la drosophile.

Un grand nombre des problèmes liés aux deux méthodologies utilisées peuvent être, sinon résolus, largement améliorés. Tout d'abord, un effort peut être porté sur l'étude des protéines présentes en plus faible concentration et généralement considérées comme étant parmi les plus importantes (protéines de régulation, de transduction du signal, etc.). L'utilisation de bandelettes à gamme de pH étroites (jusqu'à une unité de pH) permet de charger une plus grande quantité d'échantillons sur les gels 2D et donc de détecter des protéines de faible concentration. Nous avons d'ailleurs montré que l'utilisation des bandelettes de pH 5,5-6,7 augmente le nombre de spots détectés dans cette même zone par rapport à l'analyse sur bandelette de gamme 5-8. Des gammes complémentaires pourront être étudiées.

Des étapes de préfractionnement peuvent également être inclues avant les analyses augmentant les chances d'observer les protéines faiblement représentées mais également permettant d'étudier des sous-ensembles du protéome. Par exemple, on parlera de l'étude du phosphoprotéome si on analyse sélectivement les molécules phosphorylées. Ceci peut se faire par exemple grâce à une précipitation par immunoaffinité à l'aide d'anticorps reconnaissant spécifiquement certains résidus phosphorylés, ou encore par modifications chimiques des sites de phosphorylation permettant, grâce à la fixation d'une étiquette, la purification par colonne d'affinité. D'autres modifications post-traductionnelles peuvent être ciblées par ce type d'approche.

D'autres méthodologies sans gel, complémentaire à l'approche ICAT, permettent également d'analyser un échantillon complexe en sélectionnant des résidus spécifiques. L'utilisation de marqueurs ciblant d'autres acides aminés que les cystéines (méthionine et tryptophane par exemple) peuvent être utilisés et apporter des informations supplémentaires. On peut citer également la chromatographie diagonale ciblée sur l'analyse unique des peptides N-terminaux (COFRADIC pour COmbined FRActional DIagonal Chromatography). Par le fait de ne sélectionner que le peptide contenant le résidu N-terminal d'une protéine, cette méthodologie permet de simplifier considérablement l'échantillon avant l'analyse par LC-MS/MS et donc d'observer un plus grand nombre de protéines. Cependant cette dernière méthode possède l'inconvénient de ne pas apporter d'information quantitative.

Il est essentiel de retenir que nous ne sommes pas près d'avoir à notre disposition une méthodologie unique permettant d'étudier la totalité des protéines d'un organisme. Seule l'utilisation combinée de différentes stratégies peut nous approcher de ce but. Il est donc important de savoir quelles informations nous souhaitons obtenir par l'étude protéomique. Si le but est d'obtenir la liste la plus complète possible des protéines d'un échantillon d'intérêt, la stratégie basée sur la MS/MS, de par son grand rendement et la possibilité d'automatiser les analyses, est la plus efficace. En revanche, si nous souhaitons étudier des modifications occasionnées par un changement biologique, la méthodologie par gels d'électrophorèse, par le fait qu'elle permette d'observer des changements "inattendus" sur nos gels (modifications de pI et/ou de masse moléculaire traduisant des modifications post-traductionnelles, des coupures protéolytiques, des complexes protéine-protéine), reste à l'heure actuelle sans rivale.

Bien que non exhaustive, la stratégie protéomique mise en place au cours de ma thèse, associant l'analyse par gels 2D et l'ICAT, a tout de même permis de mettre en évidence un très grand nombre de candidats dont l'analyse pourra améliorer nos connaissance des mécanismes de défense de la drosophile. En effet, le taux d'abondance de plus de 170 protéines varie de manière significative dans l'hémolymphe de drosophiles à la suite d'au moins une infection expérimentale. Certaines ont un rôle connu ou fortement supposé dans la défense de la drosophile (GNBPs, PGRPs, TEPs, etc.), d'autres sont de nouveaux candidats (métallopeptidases à zinc, molécules de la famille des immunophilines, protéines sensorielles, antioxydantes, de liaison au calcium ou du métabolisme du fer).

Plusieurs de ces protéines sont en cours d'étude au laboratoire. C'est le cas notamment des protéases à sérine à domaine clip, des molécules de la famille des PPAEs et de la "GNBP-like". Cependant d'autres candidats présentent un intérêt particulier car leur implication dans les mécanismes de la réponse immunitaire des insectes était insoupçonnée jusqu'alors. L'analyse de ces nouveaux candidats pourra être menée en mettant à profit la puissance de la drosophile comme outil génétique. En effet,

109

des approches classiques de biologie moléculaire et de génétique permettront d'étudier l'implication de ces candidats dans la réponse immunitaire, notamment en observant les éventuels phénotypes des drosophiles présentant une perte ou un gain de fonction pour ces molécules.

Enfin, l'approche protéomique, au-delà de la mise en évidence de nouveaux candidats, devrait permettre d'aller plus loin dans l'analyse des mécanismes de défense de la drosophile, notamment grâce à l'étude des interactions protéine-protéine. La recherche de partenaires, au sein de complexes fonctionnels, des différents candidats identifiés au cours de ma thèse devrait ouvrir de nouveaux horizons dans la compréhension de la réponse immunitaire.

REFERENCES BIBLIOGRAPHIQUES

REFERENCES BIBLIOGRAPHIQUES

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., *et al.* (2000). The genome sequence of Drosophila melanogaster. Science 287, 2185-2195.

Anderson, K. V., and Nusslein-Volhard, C. (1984). Information for the dorsal--ventral pattern of the Drosophila embryo is stored as maternal mRNA. Nature *311*, 223-227.

Ashida, M., and Brey, P. T. (1995). Role of the integument in insect defense: prophenol oxidase cascade in the cuticular matrix. Proc Natl Acad Sci U S A 92, 10698-10702.

Baggerman, G., Boonen, K., Verleyen, P., De Loof, A., and Schoofs, L. (2005). Peptidomic analysis of the larval Drosophila melanogaster central nervous system by two-dimensional capillary liquid chromatography quadrupole time-of-flight mass spectrometry. J Mass Spectrom 40, 250-260.

Baggerman, G., Cerstiaens, A., De Loof, A., and Schoofs, L. (2002). Peptidomics of the larval Drosophila melanogaster central nervous system. J Biol Chem 277, 40368-40374.

Beerntsen, B. T., Severson, D. W., and Christensen, B. M. (1994). Aedes aegypti: characterization of a hemolymph polypeptide expressed during melanotic encapsulation of filarial worms. Exp Parasitol *79*, 312-321.

Belvin, M. P., and Anderson, K. V. (1996). A conserved signaling pathway: the Drosophila toll-dorsal pathway. Annu Rev Cell Dev Biol *12*, 393-416.

Berlutti, F., Ajello, M., Bosso, P., Morea, C., Petrucca, A., Antonini, G., and Valenti, P. (2004). Both lactoferrin and iron influence aggregation and biofilm formation in Streptococcus mutans. Biometals *17*, 271-278.

Bhattacharya, S., Bunick, C. G., and Chazin, W. J. (2004). Target selectivity in EF-hand calcium binding proteins. Biochim Biophys Acta *1742*, 69-79.

Bischoff, V., Vignal, C., Boneca, I. G., Michel, T., Hoffmann, J. A., and Royet, J. (2004). Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. Nat Immunol *5*, 1175-1180.

Bjellqvist, B., Ek, K., Righetti, P. G., Gianazza, E., Gorg, A., Westermeier, R., and Postel, W. (1982). Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. J Biochem Biophys Methods *6*, 317-339.

Blandin, S., Shiao, S. H., Moita, L. F., Janse, C. J., Waters, A. P., Kafatos, F. C., and Levashina, E. A. (2004). Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector Anopheles gambiae. Cell *116*, 661-670.

Boulanger, N., Munks, R. J., Hamilton, J. V., Vovelle, F., Brun, R., Lehane, M. J., and Bulet, P. (2002). Epithelial innate immunity. A novel antimicrobial peptide with antiparasitic activity in the blood-sucking insect Stomoxys calcitrans. J Biol Chem 277, 49921-49926.

Boutros, M., Agaisse, H., and Perrimon, N. (2002). Sequential activation of signaling pathways during innate immune responses in Drosophila. Dev Cell *3*, 711-722.

Bulet, P., Hegy, G., Lambert, J., van Dorsselaer, A., Hoffmann, J. A., and Hetru, C. (1995). Insect immunity. The inducible antibacterial peptide diptericin carries two O-glycans necessary for biological activity. Biochemistry *34*, 7394-7400.

Bulet, P., Hetru, C., Dimarcq, J. L., and Hoffmann, D. (1999). Antimicrobial peptides in insects; structure and function. Dev Comp Immunol *23*, 329-344.

 $\underline{\mathbf{C}}$ arton, Y., and Nappi, A. J. (2001). Immunogenetic aspects of the cellular immune response of Drosophilia against parasitoids. Immunogenetics 52, 157-164.

Choe, K. M., Lee, H., and Anderson, K. V. (2005). Drosophila peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor. Proc Natl Acad Sci U S A *102*, 1122-1126.

Chosa, N., Fukumitsu, T., Fujimoto, K., and Ohnishi, E. (1997). Activation of prophenoloxidase A1 by an activating enzyme in Drosophila melanogaster. Insect Biochem Mol Biol 27, 61-68.

Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., and Sanchez, J. C. (2000). The dynamic range of protein expression: a challenge for proteomic research. Electrophoresis *21*, 1104-1115.

De Gregorio, E., Han, S. J., Lee, W. J., Baek, M. J., Osaki, T., Kawabata, S., Lee, B. L., Iwanaga, S., Lemaitre, B., and Brey, P. T. (2002a). An immune-responsive Serpin regulates the melanization cascade in Drosophila. Dev Cell *3*, 581-592.

De Gregorio, E., Spellman, P. T., Rubin, G. M., and Lemaitre, B. (2001). Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc Natl Acad Sci U S A *98*, 12590-12595.

De Gregorio, E., Spellman, P. T., Tzou, P., Rubin, G. M., and Lemaitre, B. (2002b). The Toll and Imd pathways are the major regulators of the immune response in Drosophila. Embo J *21*, 2568-2579.

de Morais Guedes, S., Vitorino, R., Domingues, R., Tomer, K., Correia, A. J., Amado, F., and Domingues, P. (2005). Proteomics of immune-challenged Drosophila melanogaster larvae hemolymph. Biochem Biophys Res Commun *328*, 106-115.

DeLotto, Y., and DeLotto, R. (1998). Proteolytic processing of the Drosophila Spatzle protein by easter generates a dimeric NGF-like molecule with ventralising activity. Mech Dev 72, 141-148.

Dimopoulos, G., Richman, A., Muller, H. M., and Kafatos, F. C. (1997). Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malaria parasites. Proc Natl Acad Sci U S A *94*, 11508-11513.

Dissing, M., Giordano, H., and DeLotto, R. (2001). Autoproteolysis and feedback in a protease cascade directing Drosophila dorsal-ventral cell fate. Embo J *20*, 2387-2393.

Duncan, T., Osawa, Y., Kutty, R. K., Kutty, G., and Wiggert, B. (1999). Heme-binding by Drosophila retinoid- and fatty acid-binding glycoprotein (RFABG), a member of the proapolipophorin gene family. J Lipid Res *40*, 1222-1228.

Dushay, M. S., Asling, B., and Hultmark, D. (1996). Origins of immunity: Relish, a compound Rel-like gene in the antibacterial defense of Drosophila. Proc Natl Acad Sci U S A *93*, 10343-10347.

Dyanov, H. M., and Dzitoeva, S. G. (1995). Method for attachment of microscopic preparations on glass for in situ hybridization, PRINS and in situ PCR studies. Biotechniques *18*, 822-824, 826.

Dziarski, R. (2004). Peptidoglycan recognition proteins (PGRPs). Mol Immunol 40, 877-886.

Ekengren, S., and Hultmark, D. (2001). A family of Turandot-related genes in the humoral stress response of Drosophila. Biochem Biophys Res Commun 284, 998-1003.

Engstrom, Y., Kadalayil, L., Sun, S. C., Samakovlis, C., Hultmark, D., and Faye, I. (1993). kappa B-like motifs regulate the induction of immune genes in Drosophila. J Mol Biol *232*, 327-333.

Herrandon, D., Jung, A. C., Criqui, M., Lemaitre, B., Uttenweiler-Joseph, S., Michaut, L., Reichhart, J., and Hoffmann, J. A. (1998). A drosomycin-GFP reporter transgene

reveals a local immune response in Drosophila that is not dependent on the Toll pathway. Embo J 17, 1217-1227.

Galvani, M., Rovatti, L., Hamdan, M., Herbert, B., and Righetti, P. G. (2001). Protein alkylation in the presence/absence of thiourea in proteome analysis: a matrix assisted laser desorption/ionization-time of flight-mass spectrometry investigation. Electrophoresis *22*, 2066-2074.

Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S., and Mische, S. M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis *20*, 601-605.

Giansanti, F., Massucci, M. T., Giardi, M. F., Nozza, F., Pulsinelli, E., Nicolini, C., Botti, D., and Antonini, G. (2005). Antiviral activity of ovotransferrin derived peptides. Biochem Biophys Res Commun *331*, 69-73.

Gobert, V., Gottar, M., Matskevich, A. A., Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J. A., and Ferrandon, D. (2003). Dual activation of the Drosophila toll pathway by two pattern recognition receptors. Science *302*, 2126-2130.

Gorg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis *21*, 1037-1053.

Goto, A., Blandin, S., Royet, J., Reichhart, J. M., and Levashina, E. A. (2003). Silencing of Toll pathway components by direct injection of double-stranded RNA into Drosophila adult flies. Nucleic Acids Res *31*, 6619-6623.

Gottar, M., Gobert, V., Michel, T., Belvin, M., Duyk, G., Hoffmann, J. A., Ferrandon, D., and Royet, J. (2002). The Drosophila immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. Nature *416*, 640-644.
Graham, L. A., and Davies, P. L. (2002). The odorant-binding proteins of Drosophila melanogaster: annotation and characterization of a divergent gene family. Gene *292*, 43-55.

Grosshans, J., Bergmann, A., Haffter, P., and Nusslein-Volhard, C. (1994). Activation of the kinase Pelle by Tube in the dorsoventral signal transduction pathway of Drosophila embryo. Nature *372*, 563-566.

Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Mol Cell Biol *19*, 1720-1730.

Ha, E. M., Oh, C. T., Ryu, J. H., Bae, Y. S., Kang, S. W., Jang, I. H., Brey, P. T., and Lee, W. J. (2005). An antioxidant system required for host protection against gut infection in Drosophila. Dev Cell *8*, 125-132.

Hall, M., Wang, R., van Antwerpen, R., Sottrup-Jensen, L., and Soderhall, K. (1999). The crayfish plasma clotting protein: a vitellogenin-related protein responsible for clot formation in crustacean blood. Proc Natl Acad Sci U S A *96*, 1965-1970.

Han, D. K., Eng, J., Zhou, H., and Aebersold, R. (2001). Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol *19*, 946-951.

Harizanova, N., Georgieva, T., Dunkov, B. C., Yoshiga, T., and Law, J. H. (2005). Aedes aegypti transferrin. Gene structure, expression pattern, and regulation. Insect Mol Biol *14*, 79-88.

He, Y. W., Li, H., Zhang, J., Hsu, C. L., Lin, E., Zhang, N., Guo, J., Forbush, K. A., and Bevan, M. J. (2004). The extracellular matrix protein mindin is a pattern-recognition molecule for microbial pathogens. Nat Immunol *5*, 88-97.

Hekmat-Scafe, D. S., Scafe, C. R., McKinney, A. J., and Tanouye, M. A. (2002). Genome-wide analysis of the odorant-binding protein gene family in Drosophila melanogaster. Genome Res *12*, 1357-1369.

Hengst, U., Albrecht, H., Hess, D., and Monard, D. (2001). The phosphatidylethanolamine-binding protein is the prototype of a novel family of serine protease inhibitors. J Biol Chem 276, 535-540.

Herbert, B. R., Molloy, M. P., Gooley, A. A., Walsh, B. J., Bryson, W. G., and Williams, K. L. (1998). Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. Electrophoresis *19*, 845-851.

Hermann, A., and Cox, J. A. (1995). Sarcoplasmic calcium-binding protein. Comp Biochem Physiol B Biochem Mol Biol *111*, 337-345.

Hochstrasser, D. F., Harrington, M. G., Hochstrasser, A. C., Miller, M. J., and Merril, C. R. (1988). Methods for increasing the resolution of two-dimensional protein electrophoresis. Anal Biochem *173*, 424-435.

Hoffmann, J. A. (2003). The immune response of Drosophila. Nature 426, 33-38.

Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., and Ezekowitz, R. A. (1999). Phylogenetic perspectives in innate immunity. Science 284, 1313-1318.

Holzl, H., Kapelari, B., Kellermann, J., Seemuller, E., Sumegi, M., Udvardy, A., Medalia, O., Sperling, J., Muller, S. A., Engel, A., and Baumeister, W. (2000). The regulatory complex of Drosophila melanogaster 26S proteasomes. Subunit composition and localization of a deubiquitylating enzyme. J Cell Biol *150*, 119-130.

Horng, T., and Medzhitov, R. (2001). Drosophila MyD88 is an adapter in the Toll signaling pathway. Proc Natl Acad Sci U S A 98, 12654-12658.

Huang, T. S., Wang, H., Lee, S. Y., Johansson, M. W., Soderhall, K., and Cerenius, L. (2000). A cell adhesion protein from the crayfish Pacifastacus leniusculus, a serine proteinase homologue similar to Drosophila masquerade. J Biol Chem 275, 9996-10001.

brahim, H. R., Sugimoto, Y., and Aoki, T. (2000). Ovotransferrin antimicrobial peptide (OTAP-92) kills bacteria through a membrane damage mechanism. Biochim Biophys Acta *1523*, 196-205.

Ip, Y. T., Reach, M., Engstrom, Y., Kadalayil, L., Cai, H., Gonzalez-Crespo, S., Tatei, K., and Levine, M. (1993). Dif, a dorsal-related gene that mediates an immune response in Drosophila. Cell *75*, 753-763.

Irving, P., Troxler, L., Heuer, T. S., Belvin, M., Kopczynski, C., Reichhart, J. M., Hoffmann, J. A., and Hetru, C. (2001). A genome-wide analysis of immune responses in Drosophila. Proc Natl Acad Sci U S A *98*, 15119-15124.

Isaac, R. E., Siviter, R. J., Stancombe, P., Coates, D., and Shirras, A. D. (2000). Conserved roles for peptidases in the processing of invertebrate neuropeptides. Biochem Soc Trans 28, 460-464.

Iwanaga, S., and Kawabata, S. (1998). Evolution and phylogeny of defense molecules associated with innate immunity in horseshoe crab. Front Biosci *3*, D973-984.

Jiang, H., and Kanost, M. R. (2000). The clip-domain family of serine proteinases in arthropods. Insect Biochem Mol Biol *30*, 95-105.

Jiang, H., Wang, Y., and Kanost, M. R. (1998). Pro-phenol oxidase activating proteinase from an insect, Manduca sexta: a bacteria-inducible protein similar to Drosophila easter. Proc Natl Acad Sci U S A 95, 12220-12225.

Jiang, N., Kolhekar, A. S., Jacobs, P. S., Mains, R. E., Eipper, B. A., and Taghert, P. H. (2000). PHM is required for normal developmental transitions and for biosynthesis of secretory peptides in Drosophila. Dev Biol 226, 118-136.

Kappler, C., Meister, M., Lagueux, M., Gateff, E., Hoffmann, J. A., and Reichhart, J. M. (1993). Insect immunity. Two 17 bp repeats nesting a kappa B-related sequence confer inducibility to the diptericin gene and bind a polypeptide in bacteria-challenged Drosophila. Embo J *12*, 1561-1568.

Karlsson, C., Korayem, A. M., Scherfer, C., Loseva, O., Dushay, M. S., and Theopold, U. (2004). Proteomic analysis of the Drosophila larval hemolymph clot. J Biol Chem 279, 52033-52041.

Kawabata, S., Tokunaga, F., Kugi, Y., Motoyama, S., Miura, Y., Hirata, M., and Iwanaga, S. (1996). Limulus factor D, a 43-kDa protein isolated from horseshoe crab hemocytes, is a serine protease homologue with antimicrobial activity. FEBS Lett *398*, 146-150.

Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem *74*, 5383-5392.

Kim, Y. S., Ryu, J. H., Han, S. J., Choi, K. H., Nam, K. B., Jang, I. H., Lemaitre, B., Brey, P. T., and Lee, W. J. (2000). Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in Drosophila melanogaster cells. J Biol Chem 275, 32721-32727.

Kurz, C. L., Chauvet, S., Andres, E., Aurouze, M., Vallet, I., Michel, G. P., Uh, M., Celli, J., Filloux, A., De Bentzmann, S., *et al.* (2003). Virulence factors of the human opportunistic pathogen Serratia marcescens identified by in vivo screening. Embo J *22*, 1451-1460.

Kussmann, M., Lassing, U., Sturmer, C. A., Przybylski, M., and Roepstorff, P. (1997). Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of the neural cell adhesion protein neurolin purified by sodium dodecyl sulfate polyacrylamide gel electrophoresis or acidic precipitation. J Mass Spectrom *32*, 483-493.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *157*, 105-132.

Lagueux, M., Perrodou, E., Levashina, E. A., Capovilla, M., and Hoffmann, J. A. (2000). Constitutive expression of a complement-like protein in toll and JAK gain-of-function mutants of Drosophila. Proc Natl Acad Sci U S A 97, 11427-11432.

Lanot, R., Zachary, D., Holder, F., and Meister, M. (2001). Postembryonic hematopoiesis in Drosophila. Dev Biol 230, 243-257.

Le Bourg, E. (2001). Oxidative stress, aging and longevity in Drosophila melanogaster. FEBS Lett 498, 183-186.

Lee, K. M., Lee, K. Y., Choi, H. W., Cho, M. Y., Kwon, T. H., Kawabata, S., and Lee, B. L. (2000). Activated phenoloxidase from Tenebrio molitor larvae enhances the synthesis of melanin by using a vitellogenin-like protein in the presence of dopamine. Eur J Biochem *267*, 3695-3703.

Lee, W. J., Lee, J. D., Kravchenko, V. V., Ulevitch, R. J., and Brey, P. T. (1996). Purification and molecular cloning of an inducible gram-negative bacteria-binding protein from the silkworm, Bombyx mori. Proc Natl Acad Sci U S A *93*, 7888-7893.

Lefebvre, C., Cocquerelle, C., Vandenbulcke, F., Hot, D., Huot, L., Lemoine, Y., and Salzet, M. (2004). Transcriptomic analysis in the leech Theromyzon tessulatum: involvement of cystatin B in innate immunity. Biochem J *380*, 617-625.

Lemaitre, B., Kromer-Metzger, E., Michaut, L., Nicolas, E., Meister, M., Georgel, P., Reichhart, J. M., and Hoffmann, J. A. (1995). A recessive mutation, immune deficiency

(imd), defines two distinct control pathways in the Drosophila host defense. Proc Natl Acad Sci U S A 92, 9465-9469.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M., and Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell *86*, 973-983.

LeMosy, E. K., Tan, Y. Q., and Hashimoto, C. (2001). Activation of a protease cascade involved in patterning the Drosophila embryo. Proc Natl Acad Sci U S A *98*, 5055-5060.

Leulier, F., Parquet, C., Pili-Floury, S., Ryu, J. H., Caroff, M., Lee, W. J., Mengin-Lecreulx, D., and Lemaitre, B. (2003). The Drosophila immune system detects bacteria through specific peptidoglycan recognition. Nat Immunol *4*, 478-484.

Levashina, E. A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J. A., and Reichhart, J. M. (1999). Constitutive activation of toll-mediated antifungal defense in serpin-deficient Drosophila. Science 285, 1917-1919.

Levashina, E. A., Moita, L. F., Blandin, S., Vriend, G., Lagueux, M., and Kafatos, F. C. (2001). Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, Anopheles gambiae. Cell *104*, 709-718.

Li, X. J., Zhang, H., Ranish, J. A., and Aebersold, R. (2003). Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. Anal Chem 75, 6648-6657.

Ligoxygakis, P., Pelte, N., Hoffmann, J. A., and Reichhart, J. M. (2002a). Activation of Drosophila Toll during fungal infection by a blood serine protease. Science *297*, 114-116.

Ligoxygakis, P., Pelte, N., Ji, C., Leclerc, V., Duvic, B., Belvin, M., Jiang, H., Hoffmann, J. A., and Reichhart, J. M. (2002b). A serpin mutant links Toll activation to melanization in the host defence of Drosophila. Embo J *21*, 6330-6337.

Lorenzini, D. M., da Silva, P. I., Jr., Fogaca, A. C., Bulet, P., and Daffre, S. (2003). Acanthoscurrin: a novel glycine-rich antimicrobial peptide constitutively expressed in the hemocytes of the spider Acanthoscurria gomesiana. Dev Comp Immunol *27*, 781-791.

Loseva, O., and Engstrom, Y. (2004). Analysis of signal-dependent changes in the proteome of Drosophila blood cells during an immune response. Mol Cell Proteomics *3*, 796-808.

Lu, Y., Wu, L. P., and Anderson, K. V. (2001). The antibacterial arm of the drosophila innate immune response requires an IkappaB kinase. Genes Dev *15*, 104-110.

Manfruelli, P., Reichhart, J. M., Steward, R., Hoffmann, J. A., and Lemaitre, B. (1999). A mosaic analysis in Drosophila fat body cells of the control of antimicrobial peptide genes by the Rel proteins Dorsal and DIF. Embo J *18*, 3380-3391.

Matsuda, S., and Koyasu, S. (2003). Regulation of MAPK signaling pathways through immunophilin-ligand complex. Curr Top Med Chem *3*, 1358-1367.

Matsuda, S., Shibasaki, F., Takehana, K., Mori, H., Nishida, E., and Koyasu, S. (2000). Two distinct action mechanisms of immunophilin-ligand complexes for the blockade of T-cell activation. EMBO Rep *1*, 428-434.

Meister, M. (2004). Blood cells of Drosophila: cell lineages and role in host defence. Curr Opin Immunol *16*, 10-15.

Meister, M., Braun, A., Kappler, C., Reichhart, J. M., and Hoffmann, J. A. (1994). Insect immunity. A transgenic analysis in Drosophila defines several functional domains in the diptericin promoter. Embo J *13*, 5958-5966. Meister, M., and Lagueux, M. (2003). Drosophila blood cells. Cell Microbiol 5, 573-580.

Michel, T., Reichhart, J. M., Hoffmann, J. A., and Royet, J. (2001). Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature *414*, 756-759.

Mor, A., Hani, K., and Nicolas, P. (1994). The vertebrate peptide antibiotics dermaseptins have overlapping structural features but target specific microorganisms. J Biol Chem 269, 31635-31641.

Nappi, A. J., Vass, E., Frey, F., and Carton, Y. (1995). Superoxide anion generation in Drosophila during melanotic encapsulation of parasites. Eur J Cell Biol *68*, 450-456.

Nappi, A. J., Vass, E., Frey, F., and Carton, Y. (2000). Nitric oxide involvement in Drosophila immunity. Nitric Oxide *4*, 423-430.

Nemoto, E., Sugawara, S., Takada, H., Shoji, S., and Horiuch, H. (1999). Increase of CD26/dipeptidyl peptidase IV expression on human gingival fibroblasts upon stimulation with cytokines and bacterial components. Infect Immun 67, 6225-6233.

Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem *75*, 4646-4658.

Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis *9*, 255-262.

Park, D. S., Shin, S. W., Hong, S. D., and Park, H. Y. (2000). Immunological detection of serpin in the fall webworm, Hyphantria cunea and its inhibitory activity on the prophenoloxidase system. Mol Cells *10*, 186-192.

Park, J. M., Brady, H., Ruocco, M. G., Sun, H., Williams, D., Lee, S. J., Kato, T., Jr., Richards, N., Chan, K., Mercurio, F., *et al.* (2004). Targeting of TAK1 by the NF-kappa B protein Relish regulates the JNK-mediated immune response in Drosophila. Genes Dev *18*, 584-594.

Pikielny, C. W., Hasan, G., Rouyer, F., and Rosbash, M. (1994). Members of a family of Drosophila putative odorant-binding proteins are expressed in different subsets of olfactory hairs. Neuron *12*, 35-49.

Pili-Floury, S., Leulier, F., Takahashi, K., Saigo, K., Samain, E., Ueda, R., and Lemaitre, B. (2004). In vivo RNA interference analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in Drosophila adults. J Biol Chem 279, 12848-12853.

Rabel, D., Charlet, M., Ehret-Sabatier, L., Cavicchioli, L., Cudic, M., Otvos, L., Jr., and Bulet, P. (2004). Primary structure and in vitro antibacterial properties of the Drosophila melanogaster attacin C Pro-domain. J Biol Chem 279, 14853-14859.

Ramet, M., Manfruelli, P., Pearson, A., Mathey-Prevot, B., and Ezekowitz, R. A. (2002). Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. coli. Nature *416*, 644-648.

Ramet, M., Pearson, A., Manfruelli, P., Li, X., Koziel, H., Gobel, V., Chung, E., Krieger, M., and Ezekowitz, R. A. (2001). Drosophila scavenger receptor CI is a pattern recognition receptor for bacteria. Immunity *15*, 1027-1038.

Ross, J., Jiang, H., Kanost, M. R., and Wang, Y. (2003). Serine proteases and their homologs in the Drosophila melanogaster genome: an initial analysis of sequence conservation and phylogenetic relationships. Gene *304*, 117-131.

Royet, J. (2004). Drosophila melanogaster innate immunity: an emerging role for peptidoglycan recognition proteins in bacteria detection. Cell Mol Life Sci *61*, 537-546.

Rutschmann, S., Jung, A. C., Hetru, C., Reichhart, J. M., Hoffmann, J. A., and Ferrandon, D. (2000a). The Rel protein DIF mediates the antifungal but not the antibacterial host defense in Drosophila. Immunity *12*, 569-580.

Rutschmann, S., Jung, A. C., Zhou, R., Silverman, N., Hoffmann, J. A., and Ferrandon, D. (2000b). Role of Drosophila IKK gamma in a toll-independent antibacterial immune response. Nat Immunol *1*, 342-347.

Sabatier, L., Jouanguy, E., Dostert, C., Zachary, D., Dimarcq, J. L., Bulet, P., and Imler, J. L. (2003). Pherokine-2 and -3. Eur J Biochem 270, 3398-3407.

Salzet, M., Deloffre, L., Breton, C., Vieau, D., and Schoofs, L. (2001). The angiotensin system elements in invertebrates. Brain Res Brain Res Rev *36*, 35-45.

Santaren, J. F. (1996). Two-dimensional gel analysis of proteins in cell lines from the central nervous system of larval Drosophila. In Vitro Cell Dev Biol Anim *32*, 434-440.

Santaren, J. F., Milan, M., and Garcia-Bellido, A. (1998). Two-dimensional gel analysis of proteins in the Drosophila wing imaginal disc mutants fat and lethal (2) giant discs. Exp Cell Res *243*, 199-206.

Schoentgen, F., and Jolles, P. (1995). From structure to function: possible biological roles of a new widespread protein family binding hydrophobic ligands and displaying a nucleotide binding site. FEBS Lett *369*, 22-26.

Schulz-Knappe, P., Zucht, H. D., Heine, G., Jurgens, M., Hess, R., and Schrader, M. (2001). Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. Comb Chem High Throughput Screen *4*, 207-217.

Shai, Y. (1999). Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. Biochim Biophys Acta *1462*, 55-70.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem *68*, 850-858.

Silva, P. I., Jr., Daffre, S., and Bulet, P. (2000). Isolation and characterization of gomesin, an 18-residue cysteine-rich defense peptide from the spider Acanthoscurria gomesiana hemocytes with sequence similarities to horseshoe crab antimicrobial peptides of the tachyplesin family. J Biol Chem 275, 33464-33470.

Silverman, G. A., Bird, P. I., Carrell, R. W., Church, F. C., Coughlin, P. B., Gettins, P. G., Irving, J. A., Lomas, D. A., Luke, C. J., Moyer, R. W., *et al.* (2001). The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. J Biol Chem *276*, 33293-33296.

Silverman, N., and Maniatis, T. (2001). NF-kappaB signaling pathways in mammalian and insect innate immunity. Genes Dev 15, 2321-2342.

Silverman, N., Zhou, R., Stoven, S., Pandey, N., Hultmark, D., and Maniatis, T. (2000). A Drosophila IkappaB kinase complex required for Relish cleavage and antibacterial immunity. Genes Dev *14*, 2461-2471.

Simister, P. C., Banfield, M. J., and Brady, R. L. (2002). The crystal structure of PEBP-2, a homologue of the PEBP/RKIP family. Acta Crystallogr D Biol Crystallogr 58, 1077-1080.

Singh, P. K., Parsek, M. R., Greenberg, E. P., and Welsh, M. J. (2002). A component of innate immunity prevents bacterial biofilm development. Nature *417*, 552-555.

Soderhall, K., and Cerenius, L. (1998). Role of the prophenoloxidase-activating system in invertebrate immunity. Curr Opin Immunol *10*, 23-28.

Stafford, J. L., and Belosevic, M. (2003). Transferrin and the innate immune response of fish: identification of a novel mechanism of macrophage activation. Dev Comp Immunol 27, 539-554.

Stafford, J. L., Neumann, N. F., and Belosevic, M. (2001). Products of proteolytic cleavage of transferrin induce nitric oxide response of goldfish macrophages. Dev Comp Immunol 25, 101-115.

Stafford, J. L., Wilson, E. C., and Belosevic, M. (2004). Recombinant transferrin induces nitric oxide response in goldfish and murine macrophages. Fish Shellfish Immunol *17*, 171-185.

Steiner, H. (2004). Peptidoglycan recognition proteins: on and off switches for innate immunity. Immunol Rev *198*, 83-96.

Stoven, S., Ando, I., Kadalayil, L., Engstrom, Y., and Hultmark, D. (2000). Activation of the Drosophila NF-kappaB factor Relish by rapid endoproteolytic cleavage. EMBO Rep *1*, 347-352.

Subbalakshmi, C., Nagaraj, R., and Sitaram, N. (2001). Biological activities of retro and diastereo analogs of a 13-residue peptide with antimicrobial and hemolytic activities. J Pept Res *57*, 59-67.

Sun, H., Bristow, B. N., Qu, G., and Wasserman, S. A. (2002). A heterotrimeric death domain complex in Toll signaling. Proc Natl Acad Sci U S A *99*, 12871-12876.

Suzuki, Y., Ruiz-Ortega, M., Lorenzo, O., Ruperez, M., Esteban, V., and Egido, J. (2003). Inflammation and angiotensin II. Int J Biochem Cell Biol *35*, 881-900.

Takehana, A., Katsuyama, T., Yano, T., Oshima, Y., Takada, H., Aigaki, T., and Kurata, S. (2002). Overexpression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the

prophenoloxidase cascade in Drosophila larvae. Proc Natl Acad Sci U S A 99, 13705-13710.

Takehana, A., Yano, T., Mita, S., Kotani, A., Oshima, Y., and Kurata, S. (2004). Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in Drosophila immunity. Embo J *23*, 4690-4700.

Tauszig-Delamasure, S., Bilak, H., Capovilla, M., Hoffmann, J. A., and Imler, J. L. (2002). Drosophila MyD88 is required for the response to fungal and Gram-positive bacterial infections. Nat Immunol *3*, 91-97.

Turner, A. J., Isaac, R. E., and Coates, D. (2001). The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. Bioessays *23*, 261-269.

Tzou, P., Ohresser, S., Ferrandon, D., Capovilla, M., Reichhart, J. M., Lemaitre, B., Hoffmann, J. A., and Imler, J. L. (2000). Tissue-specific inducible expression of antimicrobial peptide genes in Drosophila surface epithelia. Immunity *13*, 737-748.

Uttenweiler-Joseph, S., Moniatte, M., Lagueux, M., Van Dorsselaer, A., Hoffmann, J. A., and Bulet, P. (1998). Differential display of peptides induced during the immune response of Drosophila: a matrix-assisted laser desorption ionization time-of-flight mass spectrometry study. Proc Natl Acad Sci U S A *95*, 11342-11347.

Varela, P. F., Llera, A. S., Mariuzza, R. A., and Tormo, J. (2002). Crystal structure of imaginal disc growth factor-2. A member of a new family of growth-promoting glycoproteins from Drosophila melanogaster. J Biol Chem 277, 13229-13236.

Vierstraete, E., Verleyen, P., Baggerman, G., D'Hertog, W., Van den Bergh, G., Arckens, L., De Loof, A., and Schoofs, L. (2004). A proteomic approach for the analysis of instantly released wound and immune proteins in Drosophila melanogaster hemolymph. Proc Natl Acad Sci U S A *101*, 470-475.

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D.F., and Williams, K.L. (1995). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol. Genet. Eng. Rev. *13*, 19-50.

 $\underline{\mathbf{Y}}$ ap, K. L., Ames, J. B., Swindells, M. B., and Ikura, M. (1999). Diversity of conformational states and changes within the EF-hand protein superfamily. Proteins *37*, 499-507.

Yates, J. R., 3rd (1998). Mass spectrometry and the age of the proteome. J Mass Spectrom 33, 1-19.

Yeung, K. C., Rose, D. W., Dhillon, A. S., Yaros, D., Gustafsson, M., Chatterjee, D., McFerran, B., Wyche, J., Kolch, W., and Sedivy, J. M. (2001). Raf kinase inhibitor protein interacts with NF-kappaB-inducing kinase and TAK1 and inhibits NF-kappaB activation. Mol Cell Biol *21*, 7207-7217.

Yi, E. C., Lee, H., Aebersold, R., and Goodlett, D. R. (2003). A microcapillary trap cartridge-microcapillary high-performance liquid chromatography electrospray ionization emitter device capable of peptide tandem mass spectrometry at the attomole level on an ion trap mass spectrometer with automated routine operation. Rapid Commun Mass Spectrom *17*, 2093-2098.

Yoshiga, T., Georgieva, T., Dunkov, B. C., Harizanova, N., Ralchev, K., and Law, J. H. (1999). Drosophila melanogaster transferrin. Cloning, deduced protein sequence, expression during the life cycle, gene localization and up-regulation on bacterial infection. Eur J Biochem *260*, 414-420.

Yoshiga, T., Hernandez, V. P., Fallon, A. M., and Law, J. H. (1997). Mosquito transferrin, an acute-phase protein that is up-regulated upon infection. Proc Natl Acad Sci U S A 94, 12337-12342.

ANNEXES

Annexe 1 - Liste des amorces utilisées

ANNEXE 1 :

Liste des amorces utilisées

1. PCR quantitative

<u>RP 49</u>

Nom :	Fly ribosomal protein 49 Sens 23-mer
Séquence:	GACGCTTCAAGGGACAGTATCTG
Nom :	Fly ribosomal protein 49 Anti-Sens 19-mer
Séquence:	AAACGCGGTTCTGCATGAG

Drosomycine

Nom :	Drosomycine Sens 25-mer
Séquence :	CGTGAGAACCTTTTCCAATATGATG
Nom :	Drosomycine Anti-Sens 18-mer
Séquence:	TCCCAGGACCACCAGCAT

Diptéricine

Nom :	Diptericine Sens 20-mer
Séquence :	GCTGCGCAATCGCTTCTACT
Nom :	Diptericine Anti-Sens 19-mer
Séquence :	TGGTGGAGTGGGCTTCATG

Transferrine (TSF)

Nom :	TSF Sens 20-mer
Séquence :	CGATCCCAAGGTCTATTTGG
Nom :	TSF Anti-Sens 20-mer
Séquence :	CGCTCGATCACATCCTTGTA

2. Injection d'ARN double brin

Transferrine (TSF)

Nom :	TSF Sens 22-mer + boîte polymérase T7
Séquence :	TAATACGACTCACTATAGGGAGACCACCCATTTATCGCCTGTGTGTG
Nom :	TSF Anti-Sens 22-mer + boîte polymérase T7
Séquence :	TAATACGACTCACTATAGGGAGACCACGCCATCCTCGCACAGATATT

Annexe 2 : Gel synthétique de l'ensemble des spots identifiés par approche 2D

ANNEXE 2 : Gel synthétique de l'ensemble des spots identifiés par approche 2D



Représentation synthétique d'un gel 2D rassemblant toutes les protéines de l'hémolymphe de drosophiles ayant été détectées au cours de ma thèse. Ces protéines ont des masses moléculaires comprises entre 17 et 114 kDa, des points isoélectriques entre 4,7 et 8,8 et des valeurs d'hydropathie comprises entre -0,739 et 0,177. L'identité des protéines est regroupée dans le tableau suivant. Les lignes grisées correspondent aux spots absents des cartes contrôles, c'est-à-dire aux spots détectés uniquement après une infection.

	Identité du gène	Abréviation	Nom	Masse moléculaire (Da)	pI	hydrophobi e
1	CG 11661	Nc73EF	Neural conserved at 73EF	113625	6.52	-0.328
2	CG 11661	Nc73EF	Neural conserved at 73EF	113625	6.52	-0.328
3	CG 10363	TepIV	Thiolester containing protein IV	168543	5.51	-0.43
4	CG 9244	Acon	Aconitase	85398	8.49	-0.318
5	CG 5397			70935	6.44	-0.12
6	<i>CG</i> 6186 + <i>CG</i> 5779	Tsf1+Bc	Transferrin 1 + Black cells	71812	6.69	-0.469
7	<i>CG</i> 6186 + <i>CG</i> 5779	Tsf1+Bc	Transferrin 1 + Black cells	71812	6.69	-0.469
8	CG 18135			86748	7.1	-0.344
9	CG 6186	Tsf1	Transferrin 1	71812	6.69	-0.469
10	CG 6186	Tsf1	Transferrin 1	71812	6.69	-0.469
11	CG 6186 + CG 8193	Tsf1	Transferrin 1 + CG8193	71812	6.69	-0.469
12	CG 6186	Tsf1	Transferrin 1	71812	6.69	-0.469
13	CG 12140			65988	6.17	-0.275
14	CG 5165	Pgm	Phosphogluconate mutase	60766	6.25	-0.138
15	CG 15825			58032	6.45	-0.329
16	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
17	CG 4347	UGP	UTP-glucose-1-phosphate uridylyltransferase	58509	7.23	-0.285
18	CG 17337			53164	5.43	-0.326
19	CG 8586			49633	5.66	-0.288
20	CG 6917	Est-6	Esterase 6	61125	5.57	-0.378
21	CG 3752			57019	6.37	-0.214
22	CG 7430			53084	6.41	0.034
23	CG 7430			53084	6.41	0.034
24	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
26	CG 11154	ATPsyn-β	ATP synthase-β	54108	5.14	0.034
27	CG 11154	ATPsyn-β	ATP synthase-β	54108	5.14	0.034
28	CG 5390			44820	5.32	-0.368
29	CG 17876	Amy-d	Amylase distal	53780	5.47	-0.313
30	CG 17876	Amy-d	Amylase distal	53780	5.47	-0.313
31	CG 14935			67014	5.12	-0.4
32	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469

33	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
34	CG 17654	Eno	Enolase	54310	8.68	-0.232
35	CG 17654	Eno	Enolase	54310	8.68	-0.232
36	CG 5162			44427	6.22	0.003
37	CG 4475	Idgf2	Imaginal disc growth factor 2	49159	6.08	-0.309
38	CG 4475	Idgf2	Imaginal disc growth factor 2	49159	6.08	-0.309
39	CG 4475	Idgf2	Imaginal disc growth factor 2	49159	6.08	-0.309
40	CG 17654	Eno	Enolase	54310	8.68	-0.232
41	CG 17654	Eno	Enolase	54310	8.68	-0.232
42	CG 17654	Eno	Enolase	54310	8.68	-0.232
43	CG 1102			43704	6.27	-0.534
44	CG 1102			43704	6.27	-0.534
45	CG 1857	nec	necrotic <tronquée></tronquée>	53727	6.82	-0.337
46	CG 7176	Idh	Isocitrate dehydrogenase	52589	7.16	-0.301
47	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
48	CG 5154	Idgf5	Imaginal disc growth factor 5	50108	6.32	-0.355
49	CG 1106	Gel	Gelsolin cleaved?	98276	5.11	-0.442
50	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
51	CG 6050	EfTuM	Elongation factor Tu mitochondrial	53992	8.24	-0.275
52	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
53	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
54	CG 6687			48775	6.29	-0.198
55	CG 6687			48775	6.29	-0.198
56	CG 1857	nec	necrotic	53727	6.82	-0.337
57	CG 1857	nec	necrotic	53727	6.82	-0.337
58	CG 1106	Gel	Gelsolin	98276	5.11	-0.442
59	CG 3127	Pgk	Phosphoglycerate kinase	43862	7.01	0.049
60	CG 4559	Idgf3	Imaginal disc growth factor 3	49255	7.12	-0.397
61	CG 5178	Act88F	Actin 88F	41700	5.29	-0.172
62	CG 5178	Act88F	Actin 88F	41700	5.29	-0.172
63	CG 9456	Spn1	Serine protease inhibitor 1	41179	5.51	-0.098
64	CG 18525	Spn5	Serine protease inhibitor 5	48476	5.37	-0.314
65	CG 3731			51874	5.67	-0.178
66	CG 2915			51057	5.76	-0.425

67	CG 2915			51057	5.76	-0.425
68	CG 4847			40767	5.75	-0.139
69	CG 11963			54808	8.82	-0.262
70	CG 16705			43947	5.92	-0.135
71	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
73	CG 3832	Phm	Peptidylglycine-α-hydroxylating monooxygenase	40564	5.7	-0.416
74	CG 9372			46121	5.68	-0.347
75	CG 4929	Argk	Arginine kinase	61264	7.12	-0.491
76	CG 15293			36698	7.77	-0.538
77	CG 15293			36698	7.77	-0.538
78	CG 4929	Argk	Arginine kinase	61264	7.12	-0.491
79	CG 4929	Argk	Arginine kinase	61264	7.12	-0.491
80	CG 4929	Argk	Arginine kinase	61264	7.12	-0.491
81	CG 7780	DNaseII	Deoxyribonuclease II	41340	6.06	-0.513
82	CG 7780	DNaseII	Deoxyribonuclease II	41340	6.06	-0.513
83	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
84	CG 10467			39627	6.11	-0.393
85	CG 9453	Spn4	Serine protease inhibitor 4	47666	5.66	-0.111
86	CG 6058	Ald	Aldolase	43243	6.63	-0.235
87	CG 6058	Ald	Aldolase	43243	6.63	-0.235
88	CG 6058	Ald	Aldolase	43243	6.63	-0.235
90	CG 11459			37877	6.55	-0.403
91	CG 9042	Gpdh	Glycerol 3 phosphate dehydrogenase	39684	6.17	0.044
92	CG 9042	Gpdh	Glycerol 3 phosphate dehydrogenase	39684	6.17	0.044
93	CG 9042	Gpdh	Glycerol 3 phosphate dehydrogenase	39684	6.17	0.044
94	CG 9042	Gpdh	Glycerol 3 phosphate dehydrogenase	39684	6.17	0.044
95	CG 9042	Gpdh	Glycerol 3 phosphate dehydrogenase	39684	6.17	0.044
96	CG 9042	Gpdh	Glycerol 3 phosphate dehydrogenase	39684	6.17	0.044
97	CG 6084			40110	8.01	-0.358
98	CG 11876			39351	7.66	0.041
99	CG 6058	Ald	Aldolase	43243	6.63	-0.235
100	CG 18067			26230	5.07	-0.452
101	CG 6058	Ald	Aldolase	43243	6.63	-0.235
102	CG 2216	Fer1HCH	Ferritin 1 heavy chain homologue	23146	5.59	-0.24

103	CG 2216	Fer1HCH	Ferritin 1 heavy chain homologue	23146	5.59	-0.24
104	CG 14207			20842	5.67	-0.621
105	CG 1633	Jafrac1	thioredoxin peroxidase 1	21738	5.52	-0.207
106	CG 17919			22647	5.76	-0.401
107	CG 1469	Fer2LCH	Ferritin 2 light chain homologue	25242	5.89	-0.295
108	CG 1469	Fer2LCH	Ferritin 2 light chain homologue	25242	5.89	-0.295
109	CG 1469	Fer2LCH	Ferritin 2 light chain homologue	25242	5.89	-0.295
110	CG 1469	Fer2LCH	Ferritin 2 light chain homologue	25242	5.89	-0.295
111	CG 12055	Gapdh1	Glyceraldehyde 3 phosphate dehydrogenase 1 35350 8.		8.25	0.03
112	CG 6030	ATPsyn-d	ATP synthase, subunit d	20201	6.1	-0.739
113	CG 6030	ATPsyn-d	ATP synthase, subunit d	20201	6.1	-0.739
114	CG 1469	Fer2LCH	Ferritin 2 light chain homologue	25242	5.89	-0.295
115	CG 1721	Pglym78	Phosphoglyceromutase	28573	6.41	-0.505
116	CG 1721	Pglym78	Phosphoglyceromutase 28573		6.41	-0.505
117	CG 1721	Pglym78	Phosphoglyceromutase	28573	6.41	-0.505
118	CG 6186	Tsf1	Transferrin 1 <tronquée></tronquée>	71812	6.69	-0.469
119	CG 4929	Argk	Arginine kinase	61264	7.12	-0.491
120	CG 3481	Adh	Alcohol dehydrogenase	27761	7.73	0.177
121	CG 15848	Scp1	Sarcoplasmic calcium-binding protein 1	21647	4.71	-0.494
122	CG 7592	Obp99b	Odorant-binding protein 99b	17188	6.1	-0.364
123	CG 7584	Obp99c	Odorant-binding protein 99c	17188	6.1	-0.364
124	CG 11315			17552	7.55	-0.018
125	CG 13422			17131	6.97	-0.546
126	CG 7584	Obp99c	Odorant-binding protein 99c	16999	7.58	-0.381
127	CG 7584	Obp99c	Odorant-binding protein 99c	16999	7.58	-0.381
128	CG 7584	Obp99c	Odorant-binding protein 99c	16999	7.58	-0.381

Annexe 3 : Liste des protéines détectées par ICAT et régulées après infection par un mélange bactérien

L'hémolymphe de drosophiles 6 h après infection par un mélange bactérien (E. coli + M. luteus) a été analysée par la stratégie ICAT. Les taux de varations sont calculés par le programme ASAPRatio en effectuant le rapport d'intensité des ions porteurs de l'isotope léger / lourd. Les cases grisées correspondent aux protéines dont l'abondance relative augmente d'au moins un facteur deux (< ou = 0,5) après infection. Les cases noircies correspondent aux protéines dont l'abondance relative diminue d'au moins un facteur deux (> ou = 2) après infection. Les valeurs extrêmes (0 ou 9999) correspondent à des protéines totalement induites ou absentes après une infection.

Les protéines ont été classées par famille.

	Identité	Nom	Fonction moléculaire	Masse Moléculaire	Point isoélectrique	Valeur d'hydrophobie (GRAVY)	proba	ratio	stdev
Protéases	•								
	CG6692	Cysteine proteinase-1	nucleic acid binding, cathepsin L activity	41601	6.74	-0.483	1	0.3	0.05
	CG3088		serine-type endopeptidase activity; structural constituent of ribosome	26882	7.54	0.131	0.99	0.5	0.22
	CG10593	Angiotensin-converting enzyme- related	peptidyl-dipeptidase A activity, zinc ion binding	73057	5.92	-0.496	1	0.5	0.05
	CG8827	Angiotensin converting enzyme	zinc ion binding; peptidyl-dipeptidase activity; peptidyl-dipeptidase A activity;	70914	5.02	-0.427	1	0.5	0.24
	CG3775		zinc ion binding; metalloendopeptidase activity	73737	6.49	-0.219	0.98	0.4	0.78
Inhibiteurs de	protease								
	CG10031		serine-type endopeptidase inhibitor activity, protease inhibitor activity	14372	9.45	-0.469	0.85	0.5	0.11
	CG33103	Papilin	metallopeptidase activity, serine-type endopeptidase inhibitor activity	312666	4.46	-0.773	1	0.5	0.38
	CG15369		cysteine protease inhibitor activity	13160	5.59	0.106	1	2.7	1.22
	CG1262	Accessory gland peptide 62F	protease inhibitor activity, serine-type endopeptidase inhibitor activity	12570	7.34	0.03	1	0.5	0.13
Molécules de l	<mark>iaison ou de t</mark> ra	ansport							
OBP	CG13873	Odorant-binding protein 56g	odorant binding	14137	6.21	-0.07	0.98	37.2	5.97
	CG7584	Odorant-binding protein 99c	odorant binding	16999	7.58	-0.381	1	0.5	0.34
Molécules de liaison au calcium	CG5730	Annexin IX	actin binding; calcium-dependent phospholipid binding	36058	4.77	-0.258	0.77	3.6	0.55
	CG12070	Saposin-related		105963	4.84	-0.2	0.99	0.2	0.04
	CG6426		calcium ion binding	17312	4.58	0.012	1	0.5	0.27
	CG9297		calmodulin binding; calcium ion binding;	106090	4.09	-0.963	0.72	99999.0	0.00
	CG18730	Amylase proximal	calcium ion binding, alpha-amylase activity	53821	5.47	-0.318	1	0.2	0.10
Autres molécules de	CG1349		RNA binding, nucleic acid binding, catalytic activity	21437	7.72	0.173	0.98	0.5	0.26
liaisons ou transport	CG4264	Heat shock protein cognate 4	ATP binding, ATPase activity, unfolded protein binding	71131	5.36	-0.416	0.99	2.5	1.36
	CG6920	mutagen-sensitive 309	nucleic acid binding; ATP binding; helicase activity; DNA helicase activity; ATP- dependent DNA helicase activity;	166079	8.72	-0.684	0.9	0.5	0.11

	CG4463	Heat shock protein 23	actin binding	20629	5.55	-0.526	1	2.6	1.40	
Molécules du	<mark>métabolisme</mark>	général								
	CG9244	Aconitase	aconitate hydratase activity	85398	8.49	-0.318	1	0.5	0.12	
	<mark>CG9042</mark>	Glycerol 3 phosphate dehydrogenase	glycerol-3-phosphate dehydrogenase (NAD+) activity	39684	6.17	0.044	1	0.4	0.12	
	<mark>CG3127</mark>	Phosphoglycerate kinase	phosphoglycerate kinase activity, carbohydrate kinase activity	43862	7.01	0.049	0.99	17.6	21.02	
	<mark>CG8893</mark>	Glyceraldehyde 3 phosphate dehydrogenase 2	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity	35369	8.56	0.018	1	0.4	0.13	
	CG3001	Hexokinase A	ATP binding; hexokinase activity	59228	5.65	-0.128	0.96	3.2	3.30	
	<mark>CG11876</mark>		pyruvate dehydrogenase (acetyl-transferring) activity	39351	7.66	0.041	0.99	2.0	0.76	
Autres molée	Autres molécules									
	CG1634	Neuroglian		143618	5.31	-0.528	1	0.1	0.00	
	CG1803			35360	6.02	-0.197	0.89	9999.0	0.00	
	CG9027		superoxide dismutase activity, metal ion binding, copper, zinc superoxide dismutase activity	19221	6.02	-0.167	0.97	0.3	0.28	
	CG5397		sterol O-acyltransferase activity, hydrolase activity, acting on ester bonds	70935	6.44	-0.12	1	0.5	0.26	
	CG5178	Actin 88F	structural constituent of cytoskeleton	41700	5.29	-0.172	0.96	4.8	4.04	
	CG4178	Larval serum protein 1 β	nutrient reservoir activity, oxygen transporter activity	95914	5.41	-0.39	1	0.5	0.23	
	CG17927	Myosin heavy chain	ATP binding; cytoskeletal protein binding; structural constituent of muscle; structural constituent of cytoskeleton; ATPase activity, coupled; microfilament motor activity;	224609	5.9	-0.864	1	2.1	0.24	
Molécules de	fonction inco	nnue								
	CG2233		unknown	46509	5.95	-0.053	1	0.3	0.23	
	CG11315			17552	7.55	-0.018	1	23.5	20.36	
	CG14933			8513	8.82	-0.286	1	2.9	0.59	
	CG10799			18736	4.67	-0.726	1	3.9	3.81	
	CG15065			4259	8.68	0.622	1	0.2	0.08	
	CG5778			10632	5.91	-0.206	1	3.4	0.45	
	CG13335			46454	6.74	-0.451	0.99	0.0	0.00	

Annexe 4 : Liste des protéines détectées par ICAT et régulées après infection par différents pathogènes

L'hémolymphe de drosophiles 24 h après infection par *A. tumefaciens*, *E. faecalis* et *S. saprophyticus*, 72 h après infection par *B. bassiana* ou 48 h après infection par le DCV, a été analysée par la stratégie ICAT. Les taux de varations sont calculés par le programme ASAPRatio en effectuant le rapport d'intensité des ions porteurs de l'isotope léger / lourd. Les cases grisées correspondent aux protéines dont l'abondance relative augmente d'au moins un facteur deux (< ou = 0,5) après infection. Les cases noircies correspondent aux protéines dont l'abondance relative deux (> ou = 2) après infection. Les valeurs extrêmes (0 ou 9999) correspondent à des protéines totalement induites ou absentes après une infection.

Les protéines ont été classées par famille.

	Identité du gène	Nom	Fonction moléculaire	Masse Moléculaire	Point isoélectrique	Valeur d'hydrophobie (GRAVY)
Molécules connue	es de la répons	e immunitaire				
	CG10810	Drosomycin A	ion channel inhibitor activity	7752	6.52	0.06
	CG32268	drosomycin-6		8151	8.17	0.033
	CG16844	Drosophila Immune induced molecule 3		4100	8.94	0.874
	CG18279	Drosophila Immune induced molecule 10, 12, 13		29124	10.92	-0.706
	CG15066	Drosophila Immune induced molecule 23		11552	7.6	0.091
	CG31509	Drosophila Immune induced molecule 33 (Turandot A)		14220	4.75	-0.14
	CG31508	Turandot C		14296	5.06	-0.173
	CG6134	spaetzle	Toll binding, morphogen activity, cytokine activity	37447	6.15	-0.703
	CG11709	Peptidoglycan recognition protein SA	peptidoglycan receptor activity; N-acetylmuramoyl-L-alanine amidase activity; pattern recognition receptor activity; peptidoglycan binding;	22260	6.58	-0.028
	CG9681	Peptidoglycan recognition protein SB1	receptor activity, N-acetylmuramoyl-L-alanine amidase activity, peptidoglycan binding	20998	9.4	-0.438
	CG6367	persephone	peptidase activity, trypsin activity, serine-type endopeptidase activity, chymotrypsin activity	42228	5.99	0.089
	CG11390	Ejaculatory bulb protein III	carrier activity	14540	8.26	-0.567
	CG5008	Gram-negative bacteria binding protein 3	Gram-negative bacterial binding, pattern recognition receptor activity	55211	6	-0.517
	CG7052	Thiolester containing protein II	serine-type endopeptidase inhibitor activity; wide-spectrum protease inhibitor activity;	158348	6.18	-0.342
	CG10363	Thiolester containing protein IV	sugar porter activity; protease inhibitor activity; endopeptidase inhibitor activity;	168543	5.51	-0.43
Protéases						
Protéases à sérine à domaine CLIP	CG2045	SER7	trypsin activity, serine-type peptidase activity, chymotrypsin activity	43649	7.06	-0.256
	CG5896		trypsin activity; serine-type endopeptidase activity; chymotrypsin activity;	41511	6.37	-0.259
	CG9372		trypsin activity, serine-type peptidase activity, chymotrypsin activity	46121	5.68	-0.347
PPAE	CG16705		trypsin activity, monophenol monooxygenase activator activity	43947	5.92	-0.135
	CG1102		DNA binding, trypsin activity, monophenol monooxygenase activator activity, serine-type endopeptidase activity, transcription factor activity	43704	6.27	-0.534
	CG3066		trypsin activity, monophenol monooxygenase activator activity, serine carboxypeptidase activity, chymotrypsin activity	43128	6.18	-0.303
	CG3505		serine-type endopeptidase activity, chymotrypsin activity	39592	5.99	-0.206
	CG5390		chymotrypsin activity	44820	5.32	-0.368

Idontitó du		<i>A</i> .	tumefaci	ens	1	B. bassian	ıa	DCV				E. faecali	is	S. saprophyticus		
gène	Nom	proba	ratio	stdev	proba	ratio	stdev	proba	ratio	stdev	proba	ratio	stdev	proba	ratio	stdev
Molécules co	nnues de la réponse immunitaire	1		1	1						1					
CG10810	Drosomycin A				1	0.1	0.06				1	0.2	0.06			
CG32268	drosomycin-6				0.97	0.2	0.04									
CG16844	Drosophila Immune induced molecule 3				1	1.0	0.19				1	0.7	0.34	0	NA	NA
CG18279	Drosophila Immune induced molecule 10, 12, 13				1	NA	NA				0	NA	NA			
CG15066	Drosophila Immune induced molecule 23	1	0.4	0.08	1	0.1	0.05	1	1.4	0.36	1	0.1	0.05	1	0.4	0.13
CG31509	Drosophila Immune induced molecule 33 (Turandot A)							0.98	NA	NA						
CG31508	Turandot C							0.96	NA	NA						
CG6134	spaetzle	0.99	1.2	0.35	NA	NA	NA	0.98	1.8	0.34	0.94	0.3	0.06			
CG11709	Peptidoglycan recognition protein SA							0.79	1.8	2.01						
CG9681	Peptidoglycan recognition protein SB1	0.98	0.1	0.06				1	0.7	0.12	0.88	0.5	0.1	0.96	0.3	0.1
CG6367	persephone				1	0.3	0.04	0.76	0.4	0.22	1	1.3	0.22	1	1.3	0.35
CG11390	Ejaculatory bulb protein III	1	0.4	0.06	1	0.4	0.19	1	0.1	0.04	1	0.4	0.13	1	0.6	0.07
CG5008	Gram-negative bacteria binding protein 3	1	2.2	0.31	1	0.7	0.24	1	0.7	0.19	1	1.0	0.1	1	1.4	1.09
CG7052	Thiolester containing protein II							1	1.0	0.19	1	0.4	0.04	0.99	0.3	0.15
CG10363	Thiolester containing protein IV				0.31	NA	NA	0.99	NA	NA	1	0.3	0.22	0.44	NA	NA
Protéases																
CG2045	SER7	1	0.4	0.19	1	0.1	0.09	0.98	0.3	0.07	1	0.2	0.04	1	0.5	0.3
CG5896					0.92	0.4	0.2									
CG9372		1	1.9	0.37	1	0.3	0.04	0.98	2.1	0.41	1	1.3	0.18	1	1.3	0.34
CG16705		1	0.6	0.16	1	0.2	0.09	1	0.8	0.09	1	0.2	0.09	1	0.6	0.1
CG1102		1	0.3	0.18	1	0.2	0.08	1	0.9	0.38	1	0.2	0.11	1	0.8	0.1
CG3066		1	0.9	0.59	0.97	0.4	0.11	1	0.9	0.25				0	NA	NA
CG3505		0.98	0.4	0.07							0.94	0.6	0.88			
CG5390		1	0.7	0.31	1	0.4	0.07	1	0.6	0.25	1	0.4	0.07	1	1.5	0.3

Autres protéases à sérine	CG8586		chymotrypsin activity	49633	5.66	-0.288
	CG6639		serine-type endopeptidase activity, chymotrypsin activity	53182	6.97	-0.466
	CG6361		trypsin activity; chymotrypsin activity; serine-type endopeptidase activity;	40549	5.26	-0.064
	CG11841		trypsin activity, serine-type endopeptidase activity, chymotrypsin activity	35593	5.39	-0.289
	CG31326		serine-type endopeptidase activity, chymotrypsin activity	57228	6.48	-0.375
	CG40160		serine-type peptidase activity; peptidase activity; enzyme regulator activity; enzyme activator activity; chymotrypsin activity; trypsin activity;	45646	5.46	-0.243
	CG32833		serine-type endopeptidase activity, trypsin activity	297664	8.45	-0.192
	CG9649		elastase activity, chymotrypsin activity, trypsin activity	56574	6.65	-0.404
	CG31217		serine-type endopeptidase activity; low-density lipoprotein receptor activity; scavenger receptor activity; chymotrypsin activity; trypsin activity;	69433	5.44	-0.400
	CG18563		serine-type endopeptidase activity; trypsin activity;	38193	6.54	-0.183
	CG4821	Tequila	chitin binding; ATP binding; serine-type endopeptidase activity; trypsin activity; pancreatic elastase activity; chymotrypsin activity; scavenger receptor activity; hydrogen-transporting ATP synthase activity, rotational mechanism; hydrogen-transporting ATPase activity, rotational mechanism;	308365	6	-0.693
Metallopeptidases à zinc	CG10593	Angiotensin-converting enzyme-related	peptidyl-dipeptidase A activity, zinc ion binding	73057	5.92	-0.496
	CG8827	Angiotensin converting enzyme	zinc ion binding; peptidyl-dipeptidase activity; peptidyl- dipeptidase A activity;	70914	5.02	-0.427
	CG14526		endothelin-converting enzyme activity, zinc ion binding, neprilysin activity	79172	5.69	-0.577
	CG4408		metallocarboxypeptidase activity, carboxypeptidase A activity	54385	5.91	-0.363
	CG11951		zinc ion binding, membrane alanyl aminopeptidase activity	93562	4.91	-0.299
	CG2915		carboxypeptidase A activity	51057	5.76	-0.425
	CG8358		zinc ion binding; neprilysin activity	80825	5.97	-0.403
	CG11034		dipeptidyl-peptidase IV activity	84500	4.94	-0.357
	CG4933		O-sialoglycoprotein endopeptidase activity; zinc ion binding;	38436	6.62	-0.148
	CG6806	Larval serum protein 2	nutrient reservoir activity, oxygen transporter activity, carboxypeptidase A activity	83349	5.87	-0.502
Autres protéases	CG10938	Proteasome α subunit	endopeptidase activity	26883	4.9	-0.124
	CG8392	1(2)05070	endopeptidase activity	24264	5.12	-0.138
	CG6692	Cysteine proteinase-1	nucleic acid binding, cathepsin L activity	41601	6.74	-0.483
	CG4847		cathepsin K activity	44056	6.01	-0.130
Inhibiteurs de pr	otéase					
Serpine	CG9453	Serine protease inhibitor 4	serine-type endopeptidase inhibitor activity	47666	5.66	-0.111
	CG18525	Serine protease inhibitor 5	serine-type endopeptidase inhibitor activity	48476	5.37	-0.314

CG8586		1	1.1	0.16	1	0.5	0.13	1	1.0	0.12	1	0.9	0.19	1	1.8	0.48
CG6639		0.99	0.5	0.08	1	0.1	0.08				0.98	0.1	0.02			
CG6361					0.96	0.1	0.02									
CG11841		1	0.7	0.11	0.73	0.6	0.07	1	0.4	0.14	0.62	NA	NA	1	0.9	0.15
CG31326		1	1.3	0.36	1	0.4	0.22	1	0.7	0.26	1	0.5	0.22	1	0.8	0.28
CG40160		1	0.9	0.34	1	0.9	0.1	1	0.9	0.25	0.3	NA	NA	0.87	1.4	0.36
CG32833		0.98	1.6	0.31	0.98	0.7	0.22	0.98	2.2	0.39						
CG9649		0.96	1.2	1.26	0.97	0.3	0.22	1	1.1	0.11	0.98	0.3	0.08			
CG31217					1	1.4	0.51	0	NA	NA	0.98	1.3	0.41			
CG18563		1			0.84	0.1	0.03									
CG4821	Tequila							1	0.6	0.16						
CG10593	Angiotensin-converting enzyme-related	1	1.3	0.29	1	0.3	0.18	1	0.9	0.29	1	1.2	0.2	1	1.4	0.21
CG8827	Angiotensin converting enzyme				1	0.7	0.2	1	0.8	0.43	1	0.5	0.15	1	1.4	0.23
CG14526		1	1.1	0.71	0	NA	NA	1	0.7	0.13	0.98	1.7	0.3			
CG4408		1	1.1	0.76	1	0.7	0.12	1	0.9	0.32	1	0.8	0.24			
<u>CG11951</u>		0.74	0.7	0.08	0.89	0.5	0.25	0.87	0.4	0.09						
CG2915 CG8358		1						0.81	1.0	0.55	0.98	25	0.58			
CG11034		1									0.85	0.4	0.83			
CG4933		1												1	0.0	0.02
CG6806	Larval serum protein 2	0.99	2.8	2.82	1	8.8	1.64	1	0.8	0.54	0.9	1.7	0.38	0.98	0.6	0.52
CG10938	Proteasome α subunit							0.98	1.8	0.43						
CG8392	1(2)05070							0.98	0.5	0.08				0	NA	NA
CG6692	Cysteine proteinase-1		0.6	0.11	1	0.6	0.12	1	0.7	0.11	0	NA	NA 0.20	0.00	1.0	0.14
			1.0	0.28	1	0.6	0.09	1	0.7	0.22		1.1	0.39	0.99	1.2	0.14
CC0452	Sarina protassa inhibitar 4							0.76	2.4	0.54	l					
CG9453	Serine protease inhibitor 4		0.6	0.12	1	0.2	0.07	0.76	2.4	0.54	1	0.9	0.21			
CG18525	Serine protease inhibitor 5	1	0.6	0.13	1	0.2	0.07	1	1.5	1	1	0.8	0.31			

Kunitz	CG3513		serine-type endopeptidase inhibitor activity	9491	6.53	0.028
	CG16713		serine-type endopeptidase inhibitor activity	9078	6.52	0.132
	CG10031		serine-type endopeptidase inhibitor activity, protease inhibitor activity	14372	9.45	-0.469
	CG33103	Papilin	metallopeptidase activity, serine-type endopeptidase inhibitor activity	312666	4.46	-0.773
	CG16704		transporter activity, serine-type endopeptidase inhibitor activity	9007	8.13	-0.057
	CG31704		serine-type endopeptidase inhibitor activity	7475	8.15	0.316
	CG16712		serine-type endopeptidase inhibitor activity	8751	5.74	0.266
	CG3604		serine-type endopeptidase inhibitor activity	14332	4.5	-0.19
	CG15369		cysteine protease inhibitor activity	13160	5.59	0.106
	CG6953	fat-spondin	structural molecule activity; serine-type endopeptidase inhibitor activity;	84946	6.38	-0.486
Kazal	CG31758		serine-type endopeptidase inhibitor activity	8864	8.16	0.108
Autres inhibiteurs de protéase	CG8050	Cystatin-like	cysteine protease inhibitor activity	13437	4.42	-0.098
	CG1262	Accessory gland peptide 62F	protease inhibitor activity, serine-type endopeptidase inhibitor activity	12570	7.34	0.03
Molécules de liai	ison ou de tra	nsport				
IDGF	CG4472	Imaginal disc growth factor 1	imaginal disc growth factor activity, chitinase activity, hydrolase activity, hydrolyzing N-glycosyl compounds	49376	6.79	-0.202
	CG4475	Imaginal disc growth factor 2	imaginal disc growth factor activity, chitinase activity, hydrolase activity, hydrolyzing N-glycosyl compounds	49159	6.08	-0.309
	CG4559	Imaginal disc growth factor 3	imaginal disc growth factor activity	49255	7.12	-0.397
	CG1780	Imaginal disc growth factor 4	imaginal disc growth factor activity	48604	7.65	-0.307
	CG5154	Imaginal disc growth factor 5	imaginal disc growth factor activity; chitinase activity; hydrolase activity, hydrolyzing N-glycosyl compounds; growth factor activity;	50108	6.32	-0.355
OBP	CG2297	Odorant-binding protein 44a	odorant binding	15860	8.21	-0.280
	CG30473	Odorant-binding protein 51a	odorant binding	13355	8.2	0.085
	CG11797	Odorant-binding protein 56a		15403	5.2	-0.42
	CG11218	Odorant-binding protein 56d	odorant binding	14119	5.31	0.038
	CG8462	Odorant-binding protein 56e	odorant binding	14172	8.5	-0.086
	CG13873	Odorant-binding protein 56g	odorant binding	14137	6.21	-0.07
	CG18111	Odorant-binding protein 99a	odorant binding	15995	6.18	-0.185
	CG7592	Odorant-binding protein 99b	odorant binding	17188	6.1	-0.364
	CG7584	Odorant-binding protein 99c	odorant binding	16999	7.58	-0.381
	CG1668	Pheromone-binding protein-related protein 2	pheromone binding; phenylalkylamine binding;	16782	4.83	-0.429
FKBP	CG11001	FK506-binding protein 2	FK506 binding, peptidyl-prolyl cis-trans isomerase activity	11666	7.85	-0.444
	CG9916	Cyclophilin 1	peptidyl-prolyl cis-trans isomerase activity; cyclin-dependent protein kinase regulator activity;	24666	9.34	-0.122
Molécules de liaison au calcium	CG15848	Sarcoplasmic calcium-binding protein 1	calcium ion binding	21647	4.71	-0.494
	CG17210	SCP-containing protein B		29282	6.11	-0.269

CG3513	1	1.1	0.38	1	1.1	0.57	1	0.9	0.18	1	1.1	0.11			
CG16713	1	0.5	0.08	1	0.2	0.08	1	0.4	0.09	1	0.4	0.12	1	0.8	0.32
CG10031	1	1.0	0.18	1	0.8	0.1	1	0.6	0.16	1	0.7	0.08	1	1.4	0.14
CG33103 Papilin	1	0.7	0.8	1	0.4	0.13	1	0.6	0.41	1	0.9	0.34	1	1.7	0.27
CG16704	1	1.2	0.28	1	0.5	0.13	1	0.9	0.09	1	0.9	0.11	1	1.1	0.52
CG31704	1	0.4	0.2	1	0.1	0.02	1	0.8	0.36	1	0.4	0.07			
CG16712	1	0.8	0.14	1	0.5	0.18	1	0.8	0.15	1	0.6	0.14	1	1.1	0.19
CG3604	0.98	0.5	0.08				0.47	NA	NA	0.64	NA	NA			
CG15369	0.97	0.8	0.13	1	2.0	0.08	0.66	NA	NA	0.94	0.9	0.2	1	1.5	0.17
CG6953 fat-spondin							1	0.3	0.03						
CG31758	1	2.2	0.5	1	1.1	0.38	1	0.5	0.29	1	0.6	0.14			
CG8050 Cystatin-like	1	0.7	0.13	1	0.4	0.14	1	0.7	0.14	0.99	0.8	0.4	1	0.0	0
CG1262 Accessory gland peptide 62F	0.98	1.3	0.2	1	0.8	0.03	1	1.5	0.23	1	1.6	0.22	1	1.6	0.11
Molécules de liaison ou de transport															
CG4472 Imaginal disc growth factor 1	1	1.2	0.16	0.96	0.2	0.04	0.88	1.8	0.31	1	0.7	0.1	0	NA	NA
CG4475 Imaginal disc growth factor 2	1	0.3	0.22	1	0.4	0.12	1	1.4	0.25	1	0.7	0.25	1	1.8	0.32
CG4559 Imaginal disc growth factor 3	1	0.5	0.1	1	0.3	0.02	1	1.2	0.14	1	0.7	0.13	1	1.1	0.14
CG1780 Imaginal disc growth factor 4	1	1.0	0.23	1	0.3	0.08	1	0.6	0.42	1	0.9	0.4	1	1.4	0.09
CG5154 Imaginal disc growth factor 5	1	0.8	0.15				1	1.0	0.18	0.99	0.7	0.09	1	0.6	0.29
CG2297 Odorant-binding protein 44a	1	0.6	0.29				1	1.2	0.43				0	NA	NA
CG30473 Odorant-binding protein 51a				0.94	2.0	0.46									
CG11797 Odorant-binding protein 56a	1	1.1	0.48	1	0.2	0.11	1	0.2	0.05	1	0.5	0.07	0.87	1.0	0.21
CG11218 Odorant-binding protein 56d	1	0.7	0.13	1	0.5	0.1	1	0.2	0.07	1	0.6	0.09	1	0.9	0.18
CG8462 Odorant-binding protein 56e	1	0.9	0.15	1	0.7	0.15	1	0.9	0.1	1	1.6	0.26	1	0.8	0.19
CG13873 Odorant-binding protein 56g	1	0.4	0.09	1	0.9	0.43	1	0.4	0.14	1	0.6	0.25	0.76	1.7	0.25
CG18111 Odorant-binding protein 99a	1	0.9	0.09	0.95	0.9	0.55	1	1.1	0.14	1	1.2	0.13	1	1.9	0.3
CG7592 Odorant-binding protein 99b	1	0.7	0.34	1	1.3	0.19	1	0.6	0.06	1	0.9	0.24	1	1.4	0.23
CG7584 Odorant-binding protein 99c	1	1.6	0.46	1	1.5	0.41	1	0.9	0.16	1	1.0	0.21	1	2.3	0.14
CG1668 Pheromone-binding protein-rela	ated protein 2	0.5		0	NA	NA	0.98	1.3	0.5					0.4	
CG11001 FK506-binding protein 2	1	0.5	0.18	1	0.4	0.04	1	0.6	0.1	1	0.8	0.13	1	0.6	0.13
CG9916 Cyclophilin 1	1	0.7	0.08	0.83	1.2	0.18	1	0.6	0.1				0	NA	NA
CG15848 Sarcoplasmic calcium-binding p	protein 1 1	0.6	0.11	1	0.6	0.22	1	1.0	0.43	1	0.6	0.13	1	0.5	0.53

	CG18730	Amylase proximal	calcium ion binding, alpha-amylase activity	53821	5.47	-0.318
	CG6426		calcium ion binding	17312	4.58	0.012
	CG1803	regucalcin		35360	6.02	-0.197
	CG9138		Notch binding, structural molecule activity, receptor activity, calcium ion binding	369392	5.17	-0.334
	CG5803	Fasciclin 3		55884	5.18	-0.358
Molécules antioxydantes	CG8913		peroxidase activity	78280	6.98	-0.363
	CG11793	Superoxide dismutase	antioxidant activity; copper, zinc superoxide dismutase activity; metal ion binding;	15699	5.67	-0.116
	CG9027		superoxide dismutase activity, metal ion binding, copper, zinc superoxide dismutase activity	19221	6.02	-0.167
	CG12013	PHGPx	glutathione peroxidase activity; peroxidase activity;	26238	8.99	-0.305
	CG2151	Thioredoxin reductase-1	antioxidant activity; NOT glutathione-disulfide reductase activity; thioredoxin-disulfide reductase activity; glutathione-disulfide reductase activity; metal ion binding; oxidoreductase activity, acting on NADH or NADPH, disulfide as acceptor;	64322	8.11	-0.123
	CG6871	Catalase	heme binding; antioxidant activity; catalase activity; peroxidase activity;	57150	8.39	-0.509
	CG10045	Glutathione S transferase D1	glutathione transferase activity	23866	6.75	-0.231
	CG8905	Superoxide dismutase 2 (Mn)	antioxidant activity; superoxide dismutase activity; manganese superoxide dismutase activity; metal ion binding;	24659	7.74	-0.464
	CG5844		oxidoreductase activity; hydro-lyase activity	41619	5.51	-0.415
	CG1633	thioredoxin peroxidase 1	peroxidase activity; antioxidant activity	21738	5.52	-0.207
Liaison / Métabolisme des lipides	CG11064	Retinoid- and fatty-acid binding protein	fatty acid binding, retinoid binding, microtubule binding, heme binding	372677	8.13	-0.311
	CG31150		lipid transporter activity	167843	6.06	-0.29
	CG5162		lipase activity, structural molecule activity	44427	6.22	0.003
	CG12070	Saposin-related		105963	4.84	-0.2
	CG15828		lipid transporter activity, lipid binding	498714	5.91	-0.392
	CG11124	secretory Phospholipase A2	phospholipase A2 activity, calcium ion binding, calcium-dependent phospholipase A2 activity	21087	6.29	-0.355
	CG4600	yippee interacting protein 2	acetyl-CoA C-acyltransferase activity	41610	8.73	0.095
	CG5878	α4GT2	transferase activity, transferring glycosyl groups; transferase activity; galactosyltransferase activity;	15867	10.11	-0.440
	CG9914		structural molecule activity; oxidoreductase activity	34786	7.62	-0.088
	CG14034		phospholipase activity	25115	5.84	-0.024
	CG3523		fatty-acid synthase activity; hydrolase activity, acting on ester bonds; zinc ion binding; alcohol dehydrogenase activity, zinc- dependent; cofactor binding;	266444	5.99	-0.136
	CG33126	Neural Lazarillo	binding; carrier activity; lipid binding	24334	4.88	0.077
Métabolisme du fer	CG2216	Ferritin 1 heavy chain homologue	ferric iron binding, ferrous iron binding	23146	5.59	-0.24

CG18730	Amylase proximal	1	0.2	0.09	1	0.1	0.07	1	0.5	0.14	1	0.2	0.11	1	0.3	0.15
CG6426		1	0.5	0.21	1	0.4	0.13	1	0.8	0.16	1	0.4	0.15	1	0.6	0.22
CG1803	regucalcin	1	1.4	0.2	1	0.2	0.33	1	0.9	0.16	0.98	NA	NA	1	0.9	0.2
CG9138		0.97	2.2	1.16	1	0.6	0.31	1	0.7	0.13	1	0.7	0.25	0.55	NA	NA
CG5803	Fasciclin 3	1	1.2	0.64	1	1.9	0.42	1	1.3	0.14	1	0.9	0.14	0.99	1.8	0.31
CG8913		1	0.5	0.14	1	0.3	0.1	1	0.4	0.16	1	0.3	0.1	0.99	0.6	0.14
CG11793	Superoxide dismutase	1	0.7	0.1				1	0.5	0.06	1	0.6	0.14	1	0.6	0.18
CG9027		0.92	0.8	0.22	0.92	0.9	0.35	0.98	0.8	0.12	0.55	NA	NA	0.86	1.7	0.29
CG12013	PHGPx							1	0.6	0.11	0.25	NA	NA			
CG2151	Thioredoxin reductase-1							1	0.6	0.61						
CG6871	Catalase							1	0.4	0.07						
CG10045	Glutathione S transferase D1]						0.98	0.6	0.26						
CG8905	Superoxide dismutase 2 (Mn)							0.7	0.2	0.03						
CG5844											0.98	0.3	0.08			
CG1633	thioredoxin peroxidase 1				0	NA	NA	0.62	NA	NA	0.92	0.4	0.08			
CG11064	Retinoid- and fatty-acid binding protein	1	1.3	0.86	1	0.8	0.63	1	1.1	0.67	1	0.9	0.38	1	2.1	1.01
CG31150		1	1.2	0.11	1	0.3	0.15	1	1.1	0.13	1	0.7	0.15	1	0.9	0.3
CG5162		1	0.8	0.21	1	1.1	0.2	1	0.7	0.18	1	1.0	0.1	0.99	NA	NA
CG12070	Saposin-related	1	0.6	0.12	0.54	NA	NA	1	0.3	0.18	1	0.7	0.16	0.31	NA	NA
CG15828		0.98	0.4	0.07	1	0.5	0.09	1	1.4	0.22	1	1.2	0.16			
CG11124	secretory Phospholipase A2	0.96	7.5	13.39							0.92	0.0	0			
CG4600	yippee interacting protein 2	0.93	NA	NA												
CG5878	α4GT2	0.75	1.3	0.2												
CG9914]			0.92	1.1	0.22	0.25	NA	NA						
CG14034					0.91	0.5	0.27				0.89	1.0	0.38	0.98	2.4	0.58
CG3523								1	0.8	0.12				0	NA	NA
CG33126	Neural Lazarillo										0.97	0000 0	0			
CG2216	Ferritin 1 heavy chain homologue	1	1.6	0.13	1	4.9	0.75	1	1.0	0.1	1	1.9	1.84	1	6.9	0.81
		J		l	I			l		l	I			l		l

	CG1469	Ferritin 2 light chain homologue	ferric iron binding, ferrous iron binding	25242	5.89	-0.295										
	CG6186	Transferrin 1	ferric iron binding, iron ion transporter activity	71812	6.69	-0.469										
Autres molécules de liaison ou transport	CG31839		receptor binding, structural molecule activity	45759	8.13	-0.44										
	CG7002	Hemolectin	chitin binding, sugar binding structural, molecule activity	426245	5.62	-0.383										
	CG1656	lectin-46Ca	galactose binding	34782	5.08	-0.508										
	CG10344		metal ion binding	32167	9.01	0.339										
	CG4178	Larval serum protein 1 β	nutrient reservoir activity, oxygen transporter activity	95914	5.41	-0.39										
	CG33208	MICAL	actin binding; centromeric DNA binding; zinc ion binding;	301484	5.83	-0.667										
	CG33149	Muscle LIM protein at 60A	protein binding; zinc ion binding;	9958	8.59	-0.558										
	CG3153		receptor binding	17284	8.86	-0.034										
	CG6342	Iron regulatory protein 1B	mRNA binding; iron ion binding; aconitate hydratase activity;	98578	5.66	-0.128										
	CG30084	tungus	actin binding; structural constituent of cytoskeleton; zinc ion binding;	274437	6.13	-0.741										
	CG3536		intracellular cyclic nucleotide activated cation channel activity; receptor activity; potassium channel activity;	134013	9.03	-0.543										
	CG10236	Laminin A	RNA binding; receptor binding; structural molecule activity; serine-type peptidase activity;	411130	5.06	-0.465										
	CG31622	Gustatory receptor 39a	taste receptor activity;	44517	9.77	0.408										
	CG33173		chitin binding; ATP binding; ATPase activity, coupled to transmembrane movement of substances; transporter activity; nucleotide binding;	206458	7.92	0.102										
Molécules du mé	étabolisme géné	ral														
	CG9042	Glycerol 3 phosphate dehydrogenase	glycerol-3-phosphate dehydrogenase (NAD+) activity	39684	6.17	0.044										
	CG7997		alpha-galactosidase activity	47404	5.18	-0.371										
	CG12055	Glyceraldehyde 3 phosphate dehydrogenase 1	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity	35350	8.25	0.03										
	<mark>CG10120</mark>	Malic enzyme	carboxy-lyase activity, malate dehydrogenase (oxaloacetate- decarboxylating) (NADP+) activity	84626	6.18	-0.345										
	CG32954	Alcohol dehydrogenase	RNA binding, GTP binding, alcohol dehydrogenase activity	30274	4.86	-0.026										
	CG32031	Arginine kinase	arginine kinase activity	61264	7.12	-0.491										
	CG2171	Triose phosphate isomerase	triose-phosphate isomerase activity, intramolecular transferase activity	37939	7.55	-0.167										
	CG7070	Pyruvate kinase	pyruvate kinase activity, carbohydrate kinase activity	57440	7.13	-0.042										
	CG7430		dihydrolipoyl dehydrogenase activity, metal ion binding	53084	6.41	0.034										
	CG7998		L-malate dehydrogenase activity	35318	9.2	0.099										
	CG11314			17451	7.5	0.049										
	CG6084		aldehyde reductase activity	40110	8.01	-0.358										
	CG9244	Aconitase	aconitate hydratase activity	85398	8.49	-0.318										
	CG4347		UTP-glucose-1-phosphate uridylyltransferase activity	58509	7.23	-0.285										
	CG6058	Aldolase	fructose-bisphosphate aldolase activity	43243	6.63	-0.235										
	CG3861	l(1)G0030	citrate (Si)-synthase activity, lyase activity	58217	9.1	-0.334										
CG1469	Ferritin 2 light chain homologue	1	1.1	0.29	1	2.0	1.49	1	0.7	0.29	1	2.1	2.04	1	NA	NA
--------------	--	------	-----	------	------	-----	------	------	---------	------	------	------	------	------	-----	------
CG6186	Transferrin 1	1	1.3	0.35	1	0.1	0.07	1	0.8	0.12	1	0.4	0.09	1	1.3	0.26
CG31839		1	0.8	0.29				1	1.0	0.29	1	0.9	0.12	1	0.7	0.26
CG7002	Hemolectin	1	1.3	0.34				1	0.4	0.15	0.98	NA	NA			
CG1656	lectin-46Ca	1	1.1	0.15				1	1.5	0.27				1	1.6	0.69
CG10344		0.95	1.1	0.27												
CG4178	Larval serum protein 1 β				1	1.4	0.29	0.33	NA	NA				0	NA	NA
CG33208	MICAL				0.72	0.8	0.09							0	NA	NA
CG33149	Muscle LIM protein at 60A							1	99999.0	0	0	NA	NA			
CG3153								0.98	0.3	0.1						
CG6342	Iron regulatory protein 1B							0.98	0.0	0						
CG30084	tungus							0.96	1.6	0.19						
CG3536								0.89	0.8	0.1						
CG10236	Laminin A										0.8	0.9	0.29			
CG31622	Gustatory receptor 39a				0	NA	NA	0	NA	NA	0.77	23.9	4.66			
CG33173														0.89	1.0	0.6
Molécules du	ı métabolisme général															
CG9042	Glycerol 3 phosphate dehydrogenase	1	0.6	0.17	1	0.5	0.04	1	0.6	0.32	1	0.5	0.3	1	0.8	0.12
CG7997		1	1.9	0.3	1	0.8	0.12	1	0.9	0.12	1	1.3	0.33	1	1.7	0.33
CG12055	Glyceraldehyde 3 phosphate dehydrogenase 1	1	1.7	0.27	1	1.3	0.3	1	1.0	0.34	1	0.9	0.11			
CG10120	Malic enzyme	1	1.0	0.19	1	1.0	0.21	1	0.7	0.2	1	1.2	0.32	1	1.6	0.22
CG32954	Alcohol dehydrogenase	1	0.9	0.4	1	0.3	0.02	1	0.8	0.05	1	0.5	0.09	1	0.6	0.09
CG32031	Arginine kinase	1	0.6	0.16	1	0.7	0.37	1	0.9	0.3	1	1.2	0.28	1	1.5	0.81
CG2171	Triose phosphate isomerase	1	1.5	0.35	1	0.8	0.26	1	1.1	0.3	1	1.0	0.1	0.99	1.8	0.3
CG7070	Pyruvate kinase	1	0.5	0.1	1	1.0	0.21	1	0.7	0.09	1	0.7	0.12	1	1.6	0.75
CG7430		1	NA	NA				0	NA	NA						
CG7998		1	0.8	0.1	0.97	1.1	0.59	1	0.3	0.05	1	0.7	0.07	0.86	1.0	0.26
CG11314		1	2.0	0.23	1	0.6	0.1	1	0.6	0.04	1	1.2	0.12	1	1.8	0.39
CG6084		1	1.3	0.56	1	0.4	0.04	1	0.6	0.14				0.95	1.1	0.23
CG9244	Aconitase	1	0.6	0.07	0.98	1.3	0.29	1	0.6	0.1	1	0.4	0.13	1	1.0	0.34
CG4347		1	1.0	0.13				1	1.0	0.19	0.96	1.2	0.26	0	NA	NA
CG6058	Aldolase	1	NA	NA				0.95	NA	NA	1	0.4	0.05	0	NA	NA
CG3861	l(1)G0030	1	0.4	0.04	0.56	NA	NA	1	0.4	0.09	1	0.9	0.09	0.49	NA	NA

	CG1721	Phosphoglyceromutase	phosphoglycerate mutase activity	28573	6.41	-0.505
	CG3127	Phosphoglycerate kinase	phosphoglycerate kinase activity, carbohydrate kinase activity	43862	7.01	0.049
	CG5320	Glutamate dehydrogenase	glutamate dehydrogenase [NAD(P)+] activity	62536	8.44	-0.351
	CG5165	Phosphogluconate mutase	phosphoglucomutase activity	60766	6.25	-0.138
	CG11321		ubiquitin-protein ligase activity, zinc ion binding	288535	5.18	-0.791
	CG5178	Actin 88F	structural constituent of cytoskeleton	41700	5.29	-0.172
	<mark>CG8893</mark>	Glyceraldehyde 3 phosphate dehydrogenase 2	glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) activity	35369	8.56	0.018
	CG17797	Accessory gland-specific peptide 29AB	galactose binding; hormone activity;	27176	7.66	-0.432
	CG11089		IMP cyclohydrolase activity, phosphoribosylaminoimidazolecarboxamide formyltransferase activity	63323	7.97	-0.035
	<mark>CG4716</mark>		methylenetetrahydrofolate dehydrogenase activity	24828	4.77	-0.87
	<mark>CG17146</mark>	Adenylate kinase-1	ATP binding; adenylate kinase activity; uridine kinase activity;	47185	6.42	-0.255
	CG5362		L-malate dehydrogenase activity	36059	6.91	0.054
	CG11661	Neural conserved at 73EF	oxoglutarate dehydrogenase (succinyl-transferring) activity	113625	6.52	-0.328
	CG11853	takeout		27344	7.61	-0.008
	CG7176	Isocitrate dehydrogenase	isocitrate dehydrogenase (NADP+) activity	52589	7.16	-0.301
	CG17011	lectin-30A	galactose binding	26296	8.5	-0.401
	<mark>CG9364</mark>	Trehalase	alpha,alpha-trehalase activity; hydrolase activity, hydrolyzing N- glycosyl compounds;	67689	4.87	-0.460
	CG4094	l(1)G0255	fumarate hydratase activity	53576	8.61	-0.16
	<mark>CG5889</mark>	malate dehydrogenase	malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) activity	68594	6.26	-0.152
	CG2827	transaldolase	transaldolase activity	36734	7.55	-0.034
	CG8251	phosphogluconate dehydrogenase	phosphogluconate dehydrogenase (decarboxylating) activity; glucose-6-phosphate isomerase activity;	62339	6.63	-0.237
	CG5214		acetyltransferase activity; dihydrolipoyllysine-residue succinyltransferase activity;	49924	9.51	-0.046
	CG6206		hydrolase activity, hydrolyzing N-glycosyl compounds	122685	5.54	-0.446
	CG3011		glycine hydroxymethyltransferase activity	58801	9.14	-0.287
	CG4706		aconitate hydratase activity	85054	6.69	-0.313
Autres mole	écules					
	CG1179	Lysozyme B	lysozyme activity	15612	5.47	-0.14
	CG7780	Deoxyribonuclease II	nucleic acid binding, deoxyribonuclease II activity	41340	6.06	-0.513
	CG11051	Neuropeptide-like precursor 2	neuropeptide hormone activity	9412	5.36	0.062
	CG1634	Neuroglian		143618	5.31	-0.528
	CG14235		cytochrome-c oxidase activity	11424	8.93	-1.089
	CG5397		sterol O-acyltransferase activity, hydrolase activity, acting on ester bonds	70935	6.44	-0.12
	CG3665	Fasciclin 2		96926	5.76	-0.374
	CG9538	Antigen 5-related		28067	6.91	-0.191

CG1721	Phosphoglyceromutase	1	NA	NA	0.76	NA	NA	1	NA	NA	0.99	NA	NA	1	NA	NA
CG3127	Phosphoglycerate kinase	0.99	0.6	0.14				1	1.0	0.18	0.98	0.8	0.17	0.99	NA	NA
CG5320	Glutamate dehydrogenase	0.94	0.8	0.12				1	0.6	0.26	1	0.7	0.16	0	NA	NA
CG5165	Phosphogluconate mutase	0.94	1.6	1.04				1	0.9	0.15						
CG11321		0.89	0.4	0.08	0	NA	NA									
CG5178	Actin 88F	0.87	1.3	0.31	1	0.6	0.1	1	0.2	0.07	0.92	0.9	0.12	0.97	NA	NA
CG8893	Glyceraldehyde 3 phosphate dehydrogenase 2	0.81	1.4	0.24	1	1.0	0.16	1	0.8	0.26	1	0.5	0.09	1	2.1	0.31
CG17797	Accessory gland-specific peptide 29AB	0.78	2.0	1.46							0.98	1.9	0.29			
CG11089					1	0.4	0.16	1	2.3	1.89	0.9	0.7	0.18			
CG4716					1	0.9	0.16	1	0.3	0.11				0.62	NA	NA
CG17146	Adenylate kinase-1				0.98	0.5	0.09	0.95	0.6	0.18						
CG5362								1	0.4	0.16						
CG11661	Neural conserved at 73EF							1	0.6	0.19						
CG11853	takeout							1	0.5	0.19						
CG7176	Isocitrate dehydrogenase							1	0.5	0.06						
CG17011	lectin-30A							0.97	1.3	0.29						
CG9364	Trehalase							0.95	0.9	0.49	0	NA	NA	1	1.1	0.51
CG4094	l(1)G0255							0.95	4.9	0.85	1	0.6	0.11			
CG5889	malate dehydrogenase]			0	NA	NA	0.94	0.7	0.4						
CG2827	transaldolase							0.94	0.5	0.06						
CG8251	phosphogluconate dehydrogenase							0.89	1.2	0.22						
CG5214]									0.99	0.8	0.11			
CG6206											0.88	1.2	0.6	1	2.5	0.91
CG3011														0.85	0.2	0.12
CG4706														0.7	0.0	0
Autres moléo	cules															
CG1179	Lysozyme B	1	1.1	0.19	0.98	0.0	0	1	1.0	0.11	1	1.2	0.19	0.4	NA	NA
CG7780	Deoxyribonuclease II	1	0.7	0.16	1	0.2	0.05	0.97	1.1	0.17	1	0.2	0.03			
CG11051	Neuropeptide-like precursor 2	1	NA	NA	1	NA	NA	0.95	NA	NA	1	NA	NA			
CG1634	Neuroglian	1	1.2	0.16	1	0.9	0.13	1	1.2	0.47	1	2.0	0.69	1	2.2	1.09
CG14235		1	0.7	0.25				1	1.5	0.34	0.96	0.5	0.08	1	0.5	0.17
CG5397		1	0.8	0.24	1	1.3	0.36	1	1.3	0.21						
CG3665	Fasciclin 2	1	1.5	0.19				0.96	2.5	0.47				0.95	1.4	0.38
CG9538	Antigen 5-related	1	0.3	0.04				1	0.5	0.36						

CG5210	Chitinase-like	chitinase activity, hydrolase activity, hydrolyzing N-glycosyl compounds, growth factor activity	50354	7.99	-0.312
CG6435			14141	5.5	-0.271
CG4757		carboxylesterase activity	62950	7.25	-0.25
CG18087	Salivary gland secretion 7	structural molecule activity	7920	8.1	0.382
CG1349	di-1ß	RNA binding, nucleic acid binding, catalytic activity	21437	7.72	0.173
CG2238	Elongation factor 2b	GTP binding; translation elongation factor activity;	94459	6.17	-0.270
CG4264	Heat shock protein cognate 4	ATP binding, ATPase activity, unfolded protein binding	71131	5.36	-0.416
CG8063	yellow-f2	dopachrome isomerase activity; receptor binding; structural molecule activity; intramolecular oxidoreductase activity;	51029	7.06	-0.392
CG13780	PDGF- and VEGF-related factor 2	vascular endothelial growth factor receptor binding, growth factor activity	46946	8.88	-0.846
CG11720	Salivary gland secretion 3	structural molecule activity	32196	9.95	-0.871
CG9696	domino	DNA binding, nucleic acid binding	350239	9.05	-0.525
CG3723	Dynein heavy chain at 93AB	motor activity, structural constituent of cytoskeleton, ATPase activity, coupled, microtubule motor activity, cysteine-type endopeptidase activity	515016	5.33	-0.219
CG2210	abnormal wing discs	GTP binding, microtubule binding, ATP binding, nucleoside- diphosphate kinase activity, nucleotide kinase activity	19208	8.66	0.06
CG7966		selenium binding	54620	5.31	-0.443
CG8622	Accessory gland-specific peptide 53Ea	hormone activity	13607	9.3	0.148
CG6917	Esterase 6	carboxylesterase activity, receptor binding	61125	5.57	-0.378
CG10181	Multiple drug resistance 65	ATP binding; ATPase activity, coupled to transmembrane movement of substances; ATPase activity; nucleotide binding; nucleoside-triphosphatase activity; multidrug transporter activity;	143784	6.09	0.057
CG2926		DNA binding; ubiquitin-protein ligase activity; zinc ion binding;	253239	9.66	-1.108
CG18734	Furin 2	ATP binding; furin activity; transmembrane receptor protein tyrosine kinase activity; subtilase activity;	183684	6.12	-0.517
CG8280	Elongation factor 1α48D	GTP binding; translation elongation factor activity; translation release factor activity; GTPase activity;	50307	9.14	-0.256
CG4123	Multiple inositol polyphosphate phosphatase 1	phosphoprotein phosphatase activity; acid phosphatase activity; inositol or phosphatidylinositol phosphatase activity;	53571	6.28	-0.321
CG3989	ade5	phosphoribosylaminoimidazole carboxylase activity; phosphoribosylaminoimidazolesuccinocarboxamide synthase activity;	47281	8.3	-0.093
CG1915	sallimus	structural constituent of cytoskeleton; myosin-light-chain kinase activity;	2058217	4.96	-0.805
CG9976	Galactose-specific C-type lectin	galactose binding; sugar binding; receptor activity;	21505	5.96	-0.428
CG7431		amine receptor activity; tyramine receptor activity;	68196	9.21	0.047
CG6028		isomerase activity; lyase activity; carboxy-lyase activity;	32293	6.13	-0.216
CG4463	Heat shock protein 23	actin binding	20629	5.55	-0.526

CG5210	Chitinase-like	1	1.3	0.09	1	0.4	0.09	1	1.0	0.12	1	0.8	0.07	1	1.3	0.31
CG6435		0.99	0.7	0.1	1	0.1	0.01	1	0.4	0.15	1	0.5	0.08			
CG4757		0.99	0.3	0.08	1	0.1	0.03				1	0.2	0.16	0.72	0.7	0.11
CG18087	Salivary gland secretion 7	0.99	0.2	0.06			1	1	0.7	0.29			1			
CG1349	dj-1β	0.99	1.5	0.34			1			1	1	0.4	0.1			
CG2238	Elongation factor 2b	0.98	0.7	0.15	0.24	NA	NA	0	NA	NA			1			
CG4264	Heat shock protein cognate 4	0.98	1.0	0.19	1	0.7	0.16	1	1.1	0.12				1	1.5	0.26
CG8063	yellow-f2	0.96	1.9	2.14	1	1.3	0.54	1	1.3	0.45	0.98	1.4	0.18			
CG13780	PDGF- and VEGF-related factor 2	0.95	99999.0	0												
CG11720	Salivary gland secretion 3	0.95	0.3	0.11	1	0.3	0.06	1	0.8	0.13	1	1.1	0.17	0.99	22.1	34.91
CG9696	domino	0.88	NA	NA]	0	NA	NA]			
CG3723	Dynein heavy chain at 93AB	0.79	3.0	0.52	0	NA	NA									
CG2210	abnormal wing discs	0.77	0.6	0.13	0.89	0.5	0.06	0.91	0.6	0.1	0.77	0.5	0.07	0.34	NA	NA
CG7966		0.75	2.3	0.3				0.47	NA	NA						
CG8622	Accessory gland-specific peptide 53Ea				1	0.4	0.06			1	0.96	1.8	0.33	0.92	1.4	0.22
CG6917	Esterase 6				1	0.3	0.16	1	9999.0	0			1			
CG10181	Multiple drug resistance 65				0.95	2.1	0.42				0	NA	NA			
CG2926					0.92	NA	NA									
CG18734	Furin 2				0.84	0.4	0.09									
CG8280	Elongation factor 1α48D				0.4	NA	NA	1	0.7	0.04						
CG4123	Multiple inositol polyphosphate phosphatase 1							0.98	0.7	0.09	0.98	0.9	0.18			
CG3989	ade5							0.98	0.5	0.11						
CG1915	sallimus	1		1	0	NA	NA	0.98	0.5	0.1			1	0	NA	NA
CG9976	Galactose-specific C-type lectin]]	0.98	0.8	0.28			1]
CG7431]]	0.98	9999.0	0]			
CG6028]]	0.98	6.3	1.11]			
CG4463	Heat shock protein 23]				0.97	0.8	0.11]			

CG8029		hydrogen:potassium-exchanging ATPase activity; hydrogen- transporting ATP synthase activity, rotational mechanism; hydrogen- transporting ATPase activity, rotational mechanism;	41124	4.98	0.087
CG8094	Hexokinase C	ATP binding; hexokinase activity;	50738	6.14	-0.148
CG9108	Regulator of G-protein signaling 7	small GTPase regulatory/interacting protein activity; regulator of G- protein signaling activity;	69824	8.76	-0.506
CG2841	proximal to raf		132905	8.8	-0.736
CG1784	13	flavin-linked sulfhydryl oxidase activity; glucosidase activity; electron transporter activity;	62776	6.2	-0.315
CG9149		acetyl-CoA C-acetyltransferase activity	41162	6.56	0.007
CG5212	Pellino	kinase regulator activity	46770	6.64	-0.380
CG1757	75		32479	5.08	-0.3274
CG1885	5	uroporphyrinogen-III synthase activity; hydrolase activity	30895	5.48	-0.268
CG3162	defective transmitter release		166592	4.58	-0.77
Molécules de fonction in	connue				
CG1813	35		86748	7.1	-0.344
CG1140	00		21946	8.82	-0.597
CG6503			6557	8.37	0.316
CG1462	<u>9</u>		35032	5.21	-0.049
CG8626	Acp53C14a		13499	7.65	0.456
CG1561	6 Acp53C14b		14762	5.85	0.18
CG3266	57		9557	8.08	0.134
CG1727	78		8789	6.12	-0.02
CG1806	Met75Cb		5681	5.98	0.102
CG3292	20		19896	8.88	0.031
CG3019	07		11633	8.84	-0.136
CG1137	78		33992	5.36	-0.109
CG2233			46509	5.95	-0.053
CG8369			10163	8.51	-0.119
CG1131	15		17552	7.55	-0.018
CG3188	33		32588	8.61	-0.925
CG3295	52		5282	7.68	0.681
CG1493	33		8513	8.82	-0.286
CG9917	7		33817	5.16	-0.109
CG1443	30		16833	5.81	-0.123
CG1506	57		15823	9.46	-0.595
CG9336	<u>í</u>		15887	7.96	0.341
CG1043	33		13860	5.5	-0.015
CG1281	1		22821	5.95	-0.397
CG1810)7		4701	7.82	0.867
CG3017	71		70778	5.83	-0.420
CG5597	7		29649	4.51	-0.213
CG5791			10209	7.64	-0.088

CG8029CG8094Hexokinase CCG9108Regulator of G-protein signaCG2841proximal to rafCG17843CG17843CG9149CG5212CG17575CG1885CG1623defective transmitter releaseMaléoulos de fonction incompare	ling 7				0	NA	NA	0.96 0.96 0.85 0.7 0 0.52 0.56	1.0 1.5 99999.0 0.6 NA NA NA NA	0.14 0.28 0 0.12 NA NA NA	0.98 0.84 0.71 0.71 0.71	1.4 0.9 1.0 4.5 NA	0.2 0.16 0.21 0.96 NA	0.77	9999.0	0
Inforcenties are forcefort incontrace CG18135 CG11400 CG6503 CG14629 CG8626 Acp53C14a CG15616 Acp53C14b CG32667 CG18064 Met75Cb CG32920 CG30197 CG30197 CG13378 CG2233 CG13155 CG32952 CG14933 CG19917 CG15067 CG3366 CG10433 CG12811 CG18107 CG30171		1 1 1 1 1 1 1 1 1 1 1 1 1 1	1.2 1.9 1.9 1.4 2.1 1.6 1.3 0.6 1.3 0.9 0.9 0.9 1.1 1.3 1.0 1.6 0.9 0.9 1.1 1.3 1.0 1.6 0.9 0.9 0.9 1.1 1.3 1.0 1.6 0.9 0.9 0.9 0.9 0.9 1.1 1.3 1.0 0.6 0.9 0.9 0.9 0.9 0.9 0.9 0.9 0.9	0.49 1.44 0.15 0.65 0.8 0.54 0.35 0.32 0.24 0.07 0.13 0.13 0.47 0.18 0.65 0.31 0.52 0.36 0.43 0 NA 0.07 0.25 0.67 1.32 0.16	$ \begin{array}{c} 1\\ 1\\ 1\\ 1\\ 1\\ 1\\ 1\\ 1\\ 1\\ 1\\ 0.92\\ 1\\ 1\\ 1\\ 0.94\\ 1\\ 0.98\\ \end{array} $	1.1 1.0 0.7 0.2 0.4 0.2 0.4 0.2 0.4 0.5 0.7 0.6 2.3 0.2 0.8 0.8 0.8 0.9 0.9	$\begin{array}{c} 0.15\\ 0.16\\ 0.36\\ 0.14\\ 0.16\\ 0.09\\ 0.11\\ 0.07\\ 0.12\\ 0.08\\ 0.15\\ 0.39\\ 0.09\\ 0.49\\ 0.02\\ 0.21\\ 0.1\\ 0.31\\ 0.21\\ 0\\ 0.09\\ 0.32\\ \end{array}$	1 1 1 1 1 1 1 1 1 1 1 1 1 1	0.8 1.3 1.1 0.4 0.9 0.8 1.1 0.8 0.2 0.6 0.3 0.9 1.1 1.3 0.9 1.3 1.5 0.5 0.7 1.2 0.4 NA 1.3 0.9 0.4	0.13 0.18 0.14 0.39 0.18 0.16 0.3 0.15 0.23 0.12 0.03 0.12 0.03 0.12 0.39 0.25 0.2 0.39 0.23 0.2 0.12 0.14 0.13 NA 0.23 0.36 0.1	1 0.48 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1.7 NA 1.3 0.9 2.2 0.9 0.7 0.1 0.7 0.8 0.7 0.8 0.7 1.2 1.4 1.3 1.4 0.3 1.0 NA	0.23 NA 0.17 0.09 0.25 0.14 0.2 0.25 0.24 0.05 0.24 0.04 0.15 0.32 0.22 0.22 0.22 0.22 0.22 0.22 0.22	$ \begin{array}{c} 1\\ 0\\ 1\\ 1\\ 1\\ 0.86\\ 1\\ 1\\ 1\\ 0\\ 1\\ 1\\ 0.99\\ 1\\ 0\\ 1\\ 0\\ 1 \end{array} $	1.2 NA 0.8 1.2 0.9 1.4 1.9 0.2 0.10 1.4 NA 0.7 2.8 1.8 1.0 1.5 2.0 0.6 NA 1.1	0.23 NA 0.5 0.89 0.2 0.55 0.23 0.16 0.18 0.26 NA 0.39 0.22 0.66 0.24 0.29 1.49 0.12 NA 0.19

CG15068		4190	7.81	0.742
CG10799		18736	4.67	-0.726
CG9691		12918	5.69	-0.268
CG15065		4259	8.68	0.622
CG4182	yellow-c	49246	5.13	-0.34
CG9572		46802	7.26	-0.341
CG31997		16456	5.11	-0.226
CG6040		190205	8.27	-0.546
CG5778		10632	5.91	-0.206
CG9796		27692	7.01	-0.086
CG16836		8090	6.68	0.503
CG15201		12558	8.08	0.103
CG15199		13034	4.67	-0.242
CG13822		24721	6.87	-0.16
CG9338		15743	6.15	0.337
CG3624		20467	4.46	0.122
CG2444		12122	7.57	0.098
CG14967		259622	8.42	-0.320
CG13032		75260	7.58	-0.838
CG15431		109051	5.9	-0.477
CG15917		13203	5.18	-0.286
CG14050		29209	7.72	-0.904

CG15068		0.78	1.7	0.85	1	1.0	0.14	1	0.7	0.22	1	0.6	0.09	1	1.1	0.4
CG10799					1	0.7	0.13	1	1.1	0.72						
CG9691					1	0.5	0.03	1	1.0	0.26	0	NA	NA	1	NA	NA
CG15065					1	0.2	0.03				1	0.1	0.02			
CG4182	yellow-c				1	0.6	0.11	0.98	0.3	0.06	0.68	NA	NA	1	0.4	0.05
CG9572					0.97	0.4	0.08									
CG31997					0.96	0.3	0.05	1	0.4	0.13						
CG6040					0.85	NA	NA	0	NA	NA	0	NA	NA	0	NA	NA
CG5778					0.84	1.4	0.32				0.95	0.3	0.07			
CG9796								1	NA	NA						
CG16836								1	0.4	0.03	0.98	0.2	0.07			
CG15201								0.99	0.9	0.17						
CG15199								0.91	0.9	0.59						
CG13822								0.91	0.4	0.06						
CG9338								0.88	0.0	0	0.79	NA	NA			
CG3624								0.86	0.4	0.07						
CG2444								0.7	2.4	0.3						
CG14967								0.7	1.5	0.41						
CG13032											0.96	0.4	0.16			
CG15431														0.89	NA	NA
CG15917														0.74	1.1	0.22
CG14050														0.7	9999.0	0

Annexe 5 : Publications



[Signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]

Proteomic Analysis of the Systemic Immune Response of Drosophila

Francine Levy, Philippe Bulet, and Laurence Ehret-Sabatier

Molecular & Cellular Proteomics, 2004, Vol. 3, Pages 156-166

Pages 156 à 166 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Il est possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: <u>peb.sciences@scd-ulp.u-strasbg.fr</u>



[Signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]

Peptidomic and proteomic analyses of the systemic immune response of Drosophila

Francine Levy, David Rabel, Maurice Charlet, Philippe Bulet, Jules A. Hoffmann and Laurence Ehret-Sabatier

Biochimie, 2004, Volume 86, N° 9-10 , Pages 607-616

Pages 607 à 616 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur :

http://dx.doi.org/10.1016/j.biochi.2004.07.007

Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: <u>peb.sciences@scd-ulp.u-strasbg.fr</u>