



**HAL**  
open science

# **Listeria monocytogenes : Caractérisation fonctionnelle d'un mutant ferritine. Etude de la biodiversité par une approche protéomique**

Emilie Dumas

► **To cite this version:**

Emilie Dumas. *Listeria monocytogenes : Caractérisation fonctionnelle d'un mutant ferritine. Etude de la biodiversité par une approche protéomique.* Ingénierie des aliments. Université Blaise Pascal - Clermont-Ferrand II; Université d'Auvergne - Clermont-Ferrand I, 2007. Français. NNT : 2007CLF21758 . tel-00718111

**HAL Id: tel-00718111**

**<https://theses.hal.science/tel-00718111>**

Submitted on 16 Jul 2012

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

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

UNIVERSITE D'Auvergne  
N° DU : 1758

UNIVERSITE BLAISE PASCAL  
Année 2007

**ECOLE DOCTORALE  
DES SCIENCES DE LA VIE ET DE LA SANTE**  
N° d'ordre : 460

**Thèse**  
pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITE BLAISE PASCAL**  
**Discipline : Sciences des Aliments**  
Présentée et soutenue publiquement par

**Emilie Dumas**

le 4 Juillet 2007

---

*Listeria monocytogenes* :  
Caractérisation fonctionnelle d'un mutant ferritine.  
Etude de la biodiversité par une approche  
protéomique

---

**Rapporteurs** Mme C. Buchrieser - Habilitée à Diriger des Recherches, Institut Pasteur, Paris  
M J. Guzzo - Professeur, ENSBANA, Université de Bourgogne, Dijon

**Membres** M P. Peyret - Professeur, Université d'Auvergne, Clermont-Ferrand  
M M. Federighi - Professeur, Ecole Nationale Vétérinaire, Nantes  
M Y. Briand - Professeur, Université Blaise Pascal, Clermont-Ferrand

**Directeur de thèse** M M. Hébraud - Directeur de Recherche, INRA, Clermont-Ferrand Theix

Unité de Recherche 454 Microbiologie - Equipe Qualité et Sécurité des Aliments  
Institut National de la Recherche Agronomique - Clermont-Ferrand / Theix



## Remerciements

J'adresse mes sincères remerciements à Monsieur Yves Briand, professeur à l'université Blaise Pascal de Clermont-Ferrand qui a accepté de présider le jury de cette thèse, ainsi qu'à Carmen Buchrieser, chargée de Recherche à l'institut Pasteur de Paris et Jean Guzzo, professeur, à l'université de Bourgogne de Dijon, d'avoir accepté d'être rapporteurs de ce travail. Je remercie également Pierre Peyret, professeur à l'université d'Auvergne de Clermont-Ferrand et Michel Federighi, professeur à l'école nationale vétérinaire de Nantes d'avoir accepté d'être membres du jury.

Ce travail a été réalisé à l'INRA de Clermont-Ferrand/Theix dans l'unité Microbiologie, au sein de l'équipe Qualité et Sécurité des aliments, dirigée par Régine Talon que je remercie pour son accueil.

Je tiens à remercier tout particulièrement Michel Hébraud en tant que directeur de thèse pour son aide et pour m'avoir conseillé et encouragé au cours de ces années de thèse.

Je tiens également à adresser mes vifs remerciements à Mickaël Desvaux pour son implication dans ma thèse, ses nombreux conseils et pour m'avoir fait partager ses connaissances, notamment concernant les systèmes de sécrétion et la biologie moléculaire.

Un grand merci à Bruno Meunier, assistant ingénieur à l'UR1213 et à Jean-Louis Berdague, directeur de l'UR370 de m'avoir accordé de leur temps pour l'analyse statistique de mes résultats.

Je voudrais aussi remercier Christophe Chambon, ingénieur d'étude à la plate-forme protéomique, pour sa disponibilité et l'aide apportée au cours de l'identification des protéines par spectrométrie de masse.

Je remercie Ingrid Chafsey d'avoir su transmettre ses connaissances acquises lors de ses formations en protéomiques et plus particulièrement concernant Image Master 2D Platinum.

Je souhaite remercier Thierry Said, pour ses conseils judicieux qui m'ont été d'un grand secours lors de la mise au point de l'isoélectrofocalisation avec l'IEF-cell.

J'adresse également mes remerciements à Pierre Peyret d'avoir accepté d'être mon tuteur dans le cadre du module "enseignement" de l'école doctorale. Merci de m'avoir ainsi permis de réaliser mes premières heures d'enseignement et d'effectuer par la suite des vacances à l'IUT d'Aurillac. Je souhaite également remercier Philippe Veysseire et Françoise Leriche de m'avoir confié des heures d'enseignement.

Mes remerciements s'adressent également à Olivier Camares, Roland, Solange et Abdel pour leur accueil à Aurillac et leur gentillesse.

Un grand merci à Brigitte pour ses nombreux conseils et dépannages en informatique, toujours effectués dans la bonne humeur.

Je tiens à remercier Jean-Paul et Nicole, les deux sympathiques bricoleurs du laboratoire, pour avoir toujours su trouver une astuce pour donner un dernier souffle à du matériel en fin de vie et ne pas se retrouver ainsi coincée lors d'une expérience.

Je souhaite remercier toute l'équipe du laboratoire où régnait une agréable bonne humeur, tellement appréciable. Merci à Sabine et Isabelle pour tous leurs conseils. Merci à Jean Labadie, Yvette, Marie-Claire, Karine, Géraud, Jérôme, Patrick et tous ceux que j'oublie. Et bien sur, merci à toutes celles qui ont vécu leurs années de thèse en même temps que moi : Emilie, Stella et Stéphanie.

Enfin, je voudrais remercier mes proches. Merci à mes amis et plus particulièrement Corinne, Fahima, Isabelle et Sophie, amies de longue date, pour leur gentillesse et les moments partagés. Je remercie ma famille qui m'a toujours soutenue dans mes choix et cru en mes capacités. Un merci tout particulier à Richard, pour tous ces instants de bonheur. Merci pour ton soutien au quotidien, tes encouragements et ta patience durant ces années de thèse

<b>INTRODUCTION .....</b>	<b>9</b>
<b>SYNTHESE BIBLIOGRAPHIQUE.....</b>	<b>15</b>
- CHAPITRE 1 - <i>LISTERIA MONOCYTOGENES</i> .....	17
1 - <i>Historique</i> .....	17
2 - <i>Le genre Listeria</i> .....	17
3 - <i>Sous-typage de l'espèce L. monocytogenes</i> .....	18
3-1. Sérotypie .....	18
3-2. Autres techniques .....	20
3-3. Sous-typage et linéages .....	22
4 - <i>Physiologie de L. monocytogenes</i> .....	22
4-1. Caractères bactériologiques .....	23
4-2. Conditions de survie et de multiplication.....	24
5 - <i>Environnement</i> .....	25
5-1. Niches écologiques.....	25
5-2. <i>Listeria monocytogenes</i> dans l'industrie agroalimentaire.....	26
6 - <i>Pathologie</i> .....	26
6-1. Pathologie animale .....	26
6-2. Pathologie humaine .....	27
7 - <i>Epidémiologie</i> .....	28
7-1. Système de surveillance en France .....	28
7-2. Caractéristiques épidémiologiques en France.....	29
8 - <i>Relation dose-réponse</i> .....	32
9 - <i>Portage asymptomatique</i> .....	33
10 - <i>Pouvoir pathogène et facteurs de virulence</i> .....	33
10-1. Voies d'infection chez l'homme .....	33
10-2. Facteurs de virulence.....	35
10-3. Régulation des facteurs de virulence par PrfA.....	45
11 - <i>Les tests de virulence</i> .....	48
11-1. Tests in vivo .....	49
11-2. Tests in vitro.....	50
- CHAPITRE 2 - BIODIVERSITE DES <i>LISTERIA MONOCYTOGENES</i> .....	51
1 - <i>Biodiversité et environnement</i> .....	52
1-1. Répartition des souches en fonction de leur sous-type .....	52

1-2. Biofilms.....	53
2 - Biodiversité et génomes.....	55
2-1. Génomes séquencés .....	55
2-2. Comparaison de souches par hybridation d'ADN .....	56
3 - Biodiversité et protéomes.....	57
4 - Biodiversité et virulence.....	59
4-1. Différences de niveau de virulence entre les souches.....	59
4-2. Sérovar et virulence.....	60
4-3. Gènes de virulence .....	61
5 - Conclusion.....	63
- CHAPITRE 3 - ADAPTATION AUX STRESS ENVIRONNEMENTAUX.....	65
1 - Principaux stress et mécanismes de survie.....	65
1-1. Survie au stress acide .....	65
1-2. Survie au stress osmotique.....	66
1-3. Survie au stress thermique froid.....	67
2 - La ferritine .....	69
2-1. Le fer, élément indispensable à la vie .....	70
2-2. Importance du fer chez <i>L. monocytogenes</i> .....	71
2-3. Les ferritines bactériennes.....	72
<b>RESULTATS .....</b>	<b>77</b>
PARTIE 1 : CARACTERISATION DU MUTANT FERRITINE DE <i>L. MONOCYTOGENES</i> EGDE 79	
1 - Problématique.....	80
2 - Stratégie .....	80
3 - Principaux résultats .....	81
<b>Article n°1 : <i>Listeria monocytogenes</i> ferritin protects against multiple stresses and is required for virulence .....</b>	<b>83</b>
4 - Discussion .....	93
PARTIE 2 : ETUDE DE LA BIODIVERSITE DE <i>L. MONOCYTOGENES</i> PAR UNE APPROCHE PROTEOMIQUE .....	97
1 - Problématique.....	98
2 - Stratégie et méthodologie utilisée .....	99
2-1. Stratégie générale .....	99
2-2. Analyse de la biodiversité de <i>L. monocytogenes</i> .....	100
2-3. Analyse des protéines sécrétées par <i>L. monocytogenes</i> .....	105

<i>3 - Principaux résultats</i> .....	106
3-1. Analyse de la biodiversité au sein de <i>L. monocytogenes</i> .....	106
3-2. Analyse de la biodiversité au sein des souches de sérovar 4b .....	107
3-3. Analyse des protéines sécrétées par <i>L. monocytogenes</i> .....	108
<b>Article n°2 : Comparative analysis of extracellular and intracellular proteomes of twelve <i>Listeria monocytogenes</i> strains</b> .....	<b>111</b>
<b>Article n°3 : Subproteome comparison of four <i>Listeria monocytogenes</i> 4b strains from different origins</b> .....	<b>145</b>
<b>Article n°4 : Proteomic analysis of <i>Listeria monocytogenes</i> extracellular proteins</b> .....	<b>161</b>
<i>4 - Discussion</i> .....	201
<b>CONCLUSION ET PERSPECTIVES</b> .....	<b>209</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b> .....	<b>215</b>
<b>PUBLICATIONS ET COMMUNICATIONS</b> .....	<b>239</b>



## LISTE DES FIGURES

Figure 1 : Composition des acides téichoïques chez différents sérotypes de <i>L. monocytogenes</i> (selon Promadej <i>et al.</i> , 1999) .....	20
Figure 2 : Incidence de la listériose (cas / million d'habitant) entre 1999 et 2003 en France.....	29
Figure 3 : Nombre de décès dus à la listériose entre 1999 et 2003 .....	30
Figure 4 : Nombre de cas de listériose en fonction des mois en 2001, 2002 et 2003 .....	30
Figure 5 : Incidence de la listériose (cas / million d'habitants par an) selon l'âge en 2004 .....	31
Figure 6 : Sérovars à l'origine des cas de listériose en France sur la période 1987-2003 .....	31
Figure 7 : Voies d'infection de <i>L. monocytogenes</i> chez l'homme (Vazquez-Boland <i>et al.</i> , 2001).....	35
Figure 8 : Cycle d'infection intracellulaire de <i>L. monocytogenes</i> (d'après (Vazquez-Boland <i>et al.</i> , 2001) .....	36
Figure 9 : Modèle de l'entrée de <i>L. monocytogenes</i> dans les cellules épithéliales grâce à l'internaline InlA (Sousa <i>et al.</i> , 2004).....	38
Figure 10 : Modèle illustrant la synergie entre les domaines N-ter et C-ter de InlB afin d'induire le signal permettant l'invasion des cellules de l'hôte (Jonquieres <i>et al.</i> , 2001).....	40
Figure 11 : Modèle de l'assemblage de l'actine, induit par ActA (Cossart & Bierne, 2001) .....	45
Figure 12 : Organisation de l'îlot de pathogénicité LIPI-1 chez les différentes espèces de <i>Listeria</i> .....	46
Figure 13 : Modèle du mécanisme d'autorégulation induit par PrfA (Vega <i>et al.</i> , 1998) .....	48
Figure 14 : Arbre phylogénique du genre <i>Listeria</i> , basé sur les ARNr 16S et 23S, <i>iap</i> , <i>prs</i> , <i>vclB</i> et <i>ldh</i> . Les souches pathogènes sont indiquées en rouge et les non pathogènes en bleu .....	51
Figure 15 : Organisation structurale d'InlA chez la souche EGDe (A) et LO28 (B). LRRs=leucine-rich repeats (Jonquieres <i>et al.</i> , 1998) .....	62
Figure 16 : Effet de la température de croissance sur la composition des acides gras majeurs de <i>L. monocytogenes</i> 10403 (a) et SLCC53 (b). Les bactéries ont été cultivées en milieu « soy broth » jusqu'en milieu de phase exponentielle à la température indiquée. Symboles : rond (a- C <sub>15:0</sub> ), carré (a-C <sub>17:0</sub> ), triangle (i-C <sub>15:0</sub> ) (Annous <i>et al.</i> , 1997) .....	68
Figure 17 : Structures de la bacterioferritine (24 sous-unités) et de Dps (12 sous-unités) chez <i>E. coli</i> .....	72
Figure 18 : Courbe de progression de l'incorporation du fer par <i>L. innocua</i> en fonction de la concentration en Fe(II) (Stefanini <i>et al.</i> , 1999).....	74
Figure 19 : Méthode d'obtention des protéines sécrétées et intracellulaires.....	100

Figure 20 : l'EBD consiste à effectuer dans un premier temps une séparation des protéines en fonction de leur charge (isoélectrofocalisation) dans un gradient de pH pré-établi, puis une deuxième séparation, à 90° par rapport à la première, en fonction de la masse moléculaire. ....	101
Figure 21 : scatter-plot .....	102
Figure 22 : Principe de l'appariement des spots protéiques entre les gels. ....	102
Figure 23 : Appariement des spots protéiques entre les gels et par rapport à un gel de référence réalisé à partir d'un échantillon protéique provenant d'un mélange en quantités égales des protéines des 12 souches . ....	103
Figure 24 : classification des souches de sérovar 1/2a en fonction de leur profil protéique (protéines sécrétées). ....	107
Figure 25 : Classification des souches de sérovar 1/2a en fonction de leur profil protéique (protéines sécrétées). ....	108

### LISTE DES TABLEAUX

Tableau 1 : Sérovars du genre <i>Listeria</i> .....	19
Tableau 2 : Distribution des sérovars parmi quelques espèces de <i>Listeria</i> .....	19
Tableau 3 : Caractères bactériologiques différenciant les espèces de <i>Listeria</i> (+ : positif ; - : négatif) (Rocourt <i>et al.</i> , 2000).....	23
Tableau 4 : Caractères différenciant les listeria de genres bactériens présentant des caractères phénotypiques voisins (Rocourt <i>et al.</i> , 2000).....	24
Tableau 5 : Anadémies survenues depuis 1980 en France et dans d'autres pays (au 01/07/00) (Rapport AFSSA).....	32
Tableau 6 : Distribution des souches en fonction de leur sérovar dans différents environnements alimentaires.....	53
Tableau 7 : Principales caractéristiques des souches de <i>Listeria</i> séquencées (Glaser <i>et al.</i> , 2001 ; Nelson <i>et al.</i> , 2004) .....	55
Tableau 8 : Répartition (en %) des principaux sérovars de 603 souches de <i>L. monocytogenes</i> à l'origine des cas humains diagnostiqués de 2001 à 2003, selon la forme clinique des cas (source : CNR des <i>listeria</i> ) .....	61



# INTRODUCTION



*Listeria monocytogenes* est une bactérie ubiquitaire qui peut être trouvée dans le sol, les réseaux hydriques, sur les végétaux ainsi que dans l'environnement industriel des ateliers de production et de transformation des aliments. Cette bactérie peut également être présente dans le tractus intestinal des animaux et de l'homme sous forme de portage asymptomatique. Elle est l'agent étiologique de la listériose, une infection grave d'origine alimentaire affectant essentiellement les personnes immunodéprimées, les personnes âgées et les femmes enceintes. Bien que la fréquence des cas de listériose ait diminué ces dernières années grâce aux efforts des industriels de la filière alimentaire et la mise en place d'un système de surveillance efficace, le risque sanitaire reste important. Ainsi, au cours des années 1999 à 2003, 200 à 270 cas /an ont été signalés en France (InVS) avec un taux de mortalité de 20 à 30%. Le risque sanitaire lié à *L. monocytogenes* concerne des secteurs variés de l'industrie agroalimentaire avec, par exemple, une incidence de 3,0% à 14,8% dans les produits de charcuterie (Salvat *et al.*, 1997) et de 1,2% dans la filière lait (Farber *et al.*, 1996). Par ailleurs, les conséquences économiques liées à une crise alimentaire de listériose peuvent être dramatiques. En effet, outre le manque à gagner direct dû au retrait et à la destruction du produit incriminé, les pertes indirectes liées à la détérioration de l'image de marque du produit peuvent être fatales à l'entreprise.

Dans l'état actuel des connaissances et d'un point de vue de santé publique, toutes les souches de *L. monocytogenes* sont considérées comme potentiellement pathogènes pour l'homme. En effet, les gènes de virulence connus à ce jour sont présents dans les 13 sérovars de l'espèce et aucune des nombreuses méthodes de sous-typage de *L. monocytogenes* n'est capable de prédire la capacité d'une souche à entraîner la maladie. Cependant, la littérature montre que la virulence expérimentale est variable selon les souches, ce qui laisserait penser qu'une biodiversité importante existe et que seules certaines d'entre elles constituent un risque sanitaire réel. De plus, il est étonnant de constater que 98% des cas de listériose humaine ne sont dus qu'aux sérovars 1/2a, 1/2b et 4b et que le sérovar 4b, prépondérant dans les cas épidémiques de listériose, est beaucoup moins fréquemment isolé que les deux autres sérovars dans les aliments ou dans l'environnement industriel. La capacité d'une souche à être pathogène peut également être liée à sa capacité à résister aux nombreux stress rencontrés dans l'environnement, les aliments et chez son hôte. En effet, elle doit être capable de survivre dans les aliments qui peuvent être acides ou contenir une forte concentration de sel, et doit résister aux faibles températures rencontrées dans les industries agroalimentaires ou dans le réfrigérateur du consommateur. Une fois chez l'hôte, *L. monocytogenes* doit notamment survivre au stress acide rencontré lors du passage gastrique chez l'homme et lors

de l'invasion intracellulaire. Elle doit également résister au stress oxydatif et à la carence en fer rencontrée chez l'hôte.

Dans un contexte d'analyse et de gestion du risque sanitaire lié aux aliments, il est indispensable de mieux comprendre la virulence de *L. monocytogenes* et de caractériser les risques sanitaires auxquels s'exposent les consommateurs vis-à-vis de produits alimentaires faiblement contaminés.

Mon travail de thèse s'est inscrit dans cet objectif, avec dans un premier temps, l'étude du produit du gène de la ferritine (*fri*) qui est impliqué dans différentes fonctions telles que la résistance à des stress environnementaux ou chimiques, et dans un deuxième temps, une analyse protéomique d'un panel de souches qui vise à mieux caractériser le risque sanitaire représenté par *L. monocytogenes*.

Les ferritines sont connues pour être des protéines de stockage du fer. Elles peuvent être utilisées pour améliorer la croissance quand il y a peu de fer disponible dans le milieu extracellulaire. Elles permettent également de protéger les bactéries contre la toxicité du fer en séquestrant ce métal. Chez *L. monocytogenes*, cette protéine est surexprimée lors de chocs thermiques. De plus, chez certaines bactéries, il a été montré que la ferritine contribuait à la virulence de la bactérie. Il semble donc que cette protéine puisse avoir un rôle plus important que le simple stockage du fer. La première partie de mon travail de thèse s'est inscrite dans une étude réalisée en collaboration avec l'Institut Pasteur, qui avait pour objectif de mieux comprendre l'importance de la ferritine de *L. monocytogenes* dans l'adaptation aux stress environnementaux et dans la virulence.

La deuxième partie de mon travail de thèse s'est insérée dans un programme de recherche, financé par la DGAL, intitulé «Approches génomiques et post-génomiques pour le typage des souches de *Listeria monocytogenes* et la prédiction du risque». Les objectifs visés par ce programme sont la caractérisation d'un panel important de souches de *L. monocytogenes* d'origine humaine de portage (isolées de fèces d'individus asymptomatiques pour la listériose), d'origine alimentaire (isolées à partir d'aliments ou d'ateliers de fabrication) et épidémiques, par l'analyse du génome, du transcriptome, du protéome et de la virulence par des tests *in vitro* et *in vivo*. L'analyse génomique permet d'avoir une approche globale et une connaissance plus approfondie de la distribution et de la diversité du contenu génétique chez *L. monocytogenes*. Les analyses post-génomiques permettent d'étudier l'expression des gènes au niveau transcriptionnel et traductionnel et de caractériser les protéines fonctionnelles.

Ce projet vise à identifier des gènes (ou protéines) marqueurs, qui pourraient être utilisés pour mieux évaluer et hiérarchiser les risques sanitaires auxquels s'exposent les consommateurs vis-à-vis de produits alimentaires faiblement contaminés.

Nos travaux se sont inscrits dans la partie post-génomique de ce projet, et ont été plus particulièrement centrés sur une analyse protéomique. Dans le cadre de ce projet, nous avons comparé et caractérisé les sous-protéomes (protéines cytosolubles et sécrétées) de 12 souches de différents sérovars et différentes origines (épidémiques, portage asymptomatique et environnement agro-alimentaire) par électrophorèse bidimensionnelle (EBD).

La première partie de ce mémoire est une étude bibliographique qui fait une synthèse des connaissances sur la virulence, la biodiversité et la résistance aux stress de *L. monocytogenes*. La deuxième partie présente les résultats sous la forme de 4 publications scindées en 2 chapitres. Le premier porte sur l'étude du gène de la ferritine et de son implication dans la virulence et dans la résistance à divers stress (article 1). Le deuxième chapitre regroupe l'ensemble des résultats sur la biodiversité de *L. monocytogenes* obtenus par l'analyse comparative de deux sous-protéomes de 12 souches. Il se décline en trois publications qui présentent respectivement (i) la comparaison des sous-protéomes des 12 souches de sérovar 1/2a, 1/2b et 4b (article 2) (ii) l'analyse comparative des protéomes des souches 4b en fonction de leur origine (article 3) et (iii) le sécrétome prédit et expérimental de *L. monocytogenes* (article 4).





# **SYNTHESE BIBLIOGRAPHIQUE**



# - CHAPITRE 1 -

## **LISTERIA MONOCYTOGENES**

### **1 - HISTORIQUE**

En Mai 1924, les lapins de laboratoire du département de pathologie de l'université de Cambridge furent touchés par une épidémie (Murray, 1926) qui entraînait une perte de poids et une mort subite. L'autopsie permit de découvrir entre autre des nécroses du foie et une monocytose. EGD Murray et ses associés en conclurent que cette maladie était due à une bactérie à Gram positif isolée des lapins malades. Cette bactérie fut tout d'abord nommée *Bacterium monocytogenes*, à cause des symptômes observés. En 1927, J. Harvey Pirie découvrit une bactérie à l'origine de la mort inhabituelle de gerbilles en Afrique du Sud et l'appela *Listerella hepatolytica*. Les souches de Murray et Pirie étant identiques, la bactérie fut renommée *Listerella monocytogenes*. Le terme listerellose étant déjà utilisé pour décrire certaines pathologies, le nom *Listeria monocytogenes* fut proposé par Pirie (1940) et définitivement accepté. Durant de nombreuses années, les listérioses furent principalement considérées comme des maladies des animaux même si des cas sporadiques et parfois dramatiques étaient décrits chez l'homme. L'intérêt pour *L. monocytogenes* grandit quand la transmission alimentaire fut prouvée en 1983 (Schlech *et al.*, 1983) Depuis cette date, *L. monocytogenes* fait l'objet de nombreuses études et de surveillances accrues.

### **2 - LE GENRE LISTERIA**

Le genre *Listeria* appartient au phylum des firmicutes, à la classe des bacilli et à l'ordre des bacillales. Les listeria sont regroupées avec les Brochotrix dans la famille des listeriaceae. Cependant, *Listeria* a longtemps été classé dans la famille des *Corynebacteriaceae*. Cette position au sein des bactéries à Gram positif est devenue définitive\_avec les résultats de taxonomie numérique, de séquençage de l'ARN 16S et par son faible G+C % (37-41%).

Actuellement, ce genre compte 6 espèces organisées en deux branches génomiques distinctes :

- *L. monocytogenes*, *L. innocua*, *L. welshimeri*, *L. seeligeri*, *L. ivanovii*.

- *L. grayi* et *L. murrayi* qui présentent de fortes similitudes génomiques. En 1992, Rocourt propose de réunir ces deux taxons en une espèce unique, *L. grayi* composée de deux sous-espèces, *L. grayi* subsp. *grayi* et *L. grayi* subsp. *murrayi*.

Parmi ces six espèces, seule *L. monocytogenes* est reconnue comme étant l'agent responsable d'infections cliniques bien que *L. ivanovii* soit aussi pathogène, mais plus particulièrement impliquée dans l'avortement des ruminants (O'Driscoll *et al.*, 1996).

### **3 - SOUS-TYPAGE DE L'ESPECE *L. MONOCYTOGENES***

Différentes méthodes sont aujourd'hui disponibles pour différencier les souches de *L. monocytogenes*. Certaines sont anciennes mais encore très largement utilisées comme la sérotypie et la lysotypie. Les méthodes plus récentes reposent sur la caractérisation du génome et s'appliquent à de nombreuses espèces bactériennes ; elles ont été particulièrement bien développées et rapidement appliquées à *L. monocytogenes* en raison de l'importance du problème épidémiologique posé par la listériose.

#### **3-1. Sérotypie**

Le sérotypage a été la première méthode permettant de discriminer les souches de *L. monocytogenes*. Cette technique exploite les réactions sérologiques de 14 antigènes somatiques (O) et 4 antigènes flagellaires (H) de *L. monocytogenes* avec une série d'antisérum. L'apparition ou non d'une réaction d'agglutination permet de révéler la présence ou l'absence de ces différents antigènes, et de définir 13 sérotypes (Seeliger & Höhne, 1979) chez *L. monocytogenes* : 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e et 7 (Tableau 1).

Tableau 1 : Sérovars du genre *Listeria*

Espèce	Sérovar	Antigène O (somatique)	Antigène H (flagellaire)
<i>L. monocytogenes</i>	1/2a	I ; II ; III	A ; B
	1/2b		A ; B ; C
	1/2c		B ; D
	3a	II ; III ; IV	A, B
	3b	II ; III ; IV ; XII ; XIII	A, B, C
	3c		B, D
	4a		A ; B ; C
	4ab	III ; V ; VI ; VII ; IX ; X	
	4b	III, V, VI	
	4c	III ; V ; VII	
4d	III ; V ; VI ; VIII		
4e	III ; V ; VI ; VIII ; IX		
7	III ; XII, XIII		
<i>L. ivanovii</i>	5	III ; V ; VI ; VIII ; X	
<i>L. innocua</i>	6a	III ; V ; VI ; VII ; IX ; XV	A ; B ; C
	6b	III ; V ; VI ; VII ; IX ; X ; XI	
<i>L. grayi</i>		III ; XII, XIV	E

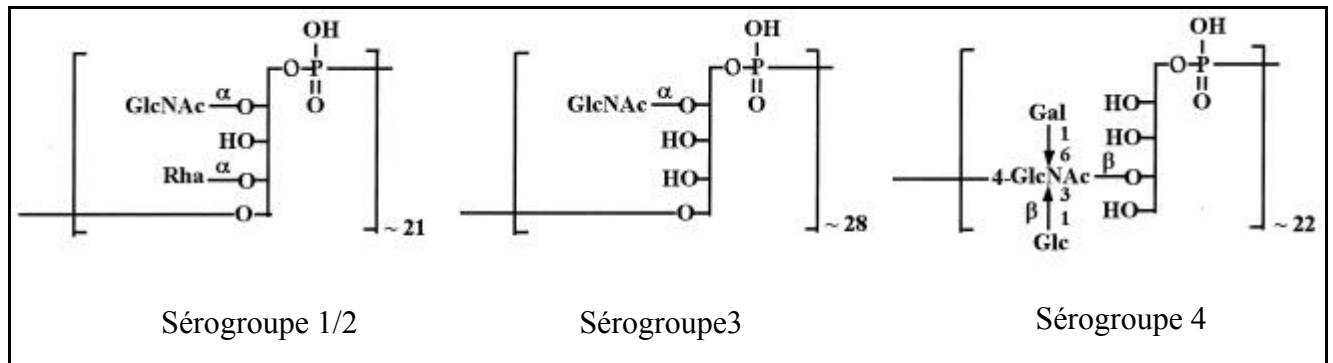
Au total, il existe 17 sérovars au sein du genre *Listeria*. A l'exception des souches du sérovar 5, qui semblent toutes appartenir à l'espèce *L. ivanovii*, il n'y a pas de corrélation entre les sérovars et les espèces (Tableau 2).

Tableau 2 : Distribution des sérovars parmi quelques espèces de *Listeria*

Espèce	Sérovars
<i>L. monocytogenes</i>	1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e et 7
<i>L. innocua</i>	3, 6a, 6b, 4ab
<i>L. ivanovii</i>	5
<i>L. seeligeri</i>	1/2a, 1/2b, 1/2c, 4b, 4c, 4d, 6b
<i>L. welshimeri</i>	1/2a, 4c, 6a, 6b

Les composants somatiques correspondent aux acides téichoïques (polyribitol phosphate lié de manière covalente au peptidoglycane) présents au niveau de la paroi. Des substitutions glycosidiques des unités ribitol-phosphate rend ce composant variable au niveau structural et antigénique. Les acides téichoïques des sérogroupe 1/2 et 3 sont des polyribitol phosphates avec des substitués N-acetylglucosamine (Glc-NAc) et rhamnose (dans le cas du sérogroupe 1/2) sur le ribitol. Au contraire, chez les souches du sérogroupe 4, le Glc-NAc est incorporé dans les chaînes d'acide téichoïque et porte selon le sérovar (4a, 4b...) des substitués de

galactose et/ou de glucose. Seules les souches de sérovar 4b ont des substitués galactose et glucose sur les Glc-NAc des acides téichoïques (Figure 1).



**Figure 1 : Composition des acides téichoïques chez différents sérotypes de *L. monocytogenes* (selon Promadej *et al.*, 1999)**

Cette caractérisation est internationalement utilisée et reconnue. Bien que son pouvoir discriminant soit faible, puisque la plupart des isolats humains de *L. monocytogenes* appartiennent aux sérovars 1/2a, 1/2b et 4b, le sérotypage est une première approche indispensable dans la différenciation de l'espèce.

## 3-2. Autres techniques

### 3-2.1. La lysotypie

Cette technique est fondée sur la sensibilité aux bactériophages et permet de subdiviser les souches d'un même sérovar (Audurier *et al.*, 1977 ; Rocourt *et al.*, 1985). En 1981, un atelier de travail international a permis de sélectionner 29 phages (appartenant aux familles des *Myoviridae* et des *Styloviridae*) isolés de souches de *Listeria* sp. ou du milieu extérieur. De nombreuses souches, notamment des souches présentes dans les aliments ou dans le milieu extérieur, ne peuvent pas être typées mais, 93% des souches du sérovar 1/2 et 99% des souches du sérogroupe 4 ont pu être typées dans une étude effectuée par (Loessner & Busse, 1990 ; Loessner, 1991) La lysotypie est très utile pour les enquêtes épidémiologiques et, depuis 1991, la procédure dite "reverse" a simplifié la mise en œuvre de cette technique. Dans la procédure "reverse", les suspensions de phages sont déposées sur une gélose tryptose qui est mise à sécher puisensemencée avec une culture en phase exponentielle de la souche à typer. La lecture est effectuée après 12 h et 36 h d'incubation. Les boîtes de gélose sur lesquelles sont absorbées les phages peuvent être préparées à l'avance et elles se conservent 6

semaines à + 4°C, ce qui facilite la réalisation de la technique et permet de gagner beaucoup de temps.

### **3-2.2. Multilocus Enzyme Electrophoresis : MEE**

L'analyse du polymorphisme électrophorétique des enzymes (ou «Multilocus Enzyme Electrophoresis», [MEE]) permet de différencier les isolats selon la mobilité électrophorétique d'un certain nombre d'enzymes métaboliques. Cette analyse a été appliquée dès 1989 à *L. monocytogenes* (Piffaretti *et al.*, 1989 ; Bibb *et al.*, 1990).

### **3-2.3. Restriction Enzyme Analysis (REA) et Restriction Fragment Length Polymorphism Pattern (RFLP)**

L'analyse des profils de restriction de l'Acide DésoxyriboNucléique (ADN) chromosomique total (ou «Restriction Enzyme Analysis», [REA]) permet une discrimination assez bonne, cependant les profils d'ADN demeurent difficiles à analyser du fait d'un grand nombre de fragments générés. Ceci a été partiellement résolu par la technique «Restriction Fragment Length Polymorphism Pattern» (RFLP) qui consiste à analyser le polymorphisme électrophorétique de certains gènes chromosomiques à partir d'une électrophorèse d'ADN total couplée à un transfert et une hybridation avec certains fragments d'ADN. La discrimination apportée par cette technique est intéressante pour les souches de sérotype 1, mais plus limitée avec les souches du sérotype 4.

### **3-2.4. Le ribotypage**

C'est une technique fondée sur le même principe que la méthode RFLP, la sonde utilisée étant une séquence d'ADN s'hybridant avec les gènes codant pour l'ARN (Acide RiboNucléique) ribosomique. Cette approche est très largement utilisée pour la caractérisation de nombreuses espèces bactériennes (Grimont & Grimont, 1986). Son pouvoir discriminant varie en fonction des sérovars, mais son inconvénient majeur est la lourdeur et la longueur des travaux qu'elle impose. Depuis peu, une automatisation de la technique a été développée. Ce système, appelé Riboprinter™ (E.I. Dupont) permet une bonne standardisation de la méthode accompagnée d'une base de données de profils et devient accessible à un plus grand nombre de laboratoires. Cependant, le coût total du ribotypage reste élevé et limite son utilisation aux laboratoires de grande capacité.



### 3-2.5. Pulsed Field Gel Electrophoresis: PFGE

La technique de détermination du profil de restriction de l'ADN total après électrophorèse en champ pulsé (ou «Pulsed Field Gel Electrophoresis», [PFGE]) est fondée sur l'électrophorèse de gros fragments d'ADN générés après l'action d'une enzyme de restriction à faible fréquence de coupure. Cette électrophorèse s'effectue dans un champ électrique pulsé qui permet la migration de gros fragments d'ADN. Elle a été largement utilisée ces dernières années dans un but épidémiologique (Franciosa *et al.*, 1998), pour comparer les souches isolées d'aliments avec celles isolées de patients (Boerlin *et al.*, 1997), afin d'analyser la traçabilité de contaminants microbiens au sein de sites de production industriels (Giovannacci *et al.*, 1999). La PFGE paraît reproductible et a été utilisée dans le cadre d'études multicentriques interlaboratoires. D'après Brosch *et al.* (1994), cette technique est l'une des plus discriminantes pour le typage de *L. monocytogenes*.

### 3-2.6. Random Amplified Polymorphism DNA » (RAPD)

D'autres méthodes fondées sur l'amplification génique ont été développées. La technique «Random Amplified Polymorphism DNA» (RAPD) (Boerlin *et al.*, 1995) est fondée sur l'utilisation d'amorces nucléotidiques choisies au hasard. Cette méthode est plus rapide à mettre en oeuvre mais manque parfois de reproductibilité.

## 3-3. Sous-typage et lignages

Ce sous-typage des souches de *L. monocytogenes* est important pour les études épidémiologiques lors de cas groupés de listériose (comparaison des souches cliniques et isolées des aliments) et pour l'étude d'environnements agroalimentaires, afin d'identifier la source de contamination.

Sur la base des nombreuses méthodes de typage moléculaire (ribotypage, PFGE,...), il a été montré que les souches de *L. monocytogenes* se répartissaient en deux divisions majeures, appelées lignage I et II (Brosch *et al.*, 1994 ; Zhang *et al.*, 2003), (Graves *et al.*, 1994) (Piffaretti *et al.*, 1989) (Ripabelli *et al.*, 2000) :

Etudes	Lignage I	Lignage II	Méthode
(Brosch <i>et al.</i> , 1994)	1/2a, 1/2c, 3a, et 3c	1/2b, 3b, 4b, 4d, et 4e	PFGE
(Graves <i>et al.</i> , 1994)	1/2a, 1/2c, 3a	1/2b, 3b, 4b, 4ab	Ribotype
(Zhang <i>et al.</i> , 2003)	1/2a	1/2b, 4b	Hybridation ADN
(Piffaretti <i>et al.</i> , 1989)	1/2a, 1/2c	4b, 1/2b, 4a	polymorphisme
(Ripabelli <i>et al.</i> , 2000)	1/2a, 1/2c	1/2b, 3b, 4b	AFLP

Il n'existe à ce jour aucune nomenclature officielle, le nom de ces lignages variant d'une étude à l'autre. De plus, certaines études décrivent l'existence d'un troisième lignage

qui semble être associé de matière prédominante aux cas de listériose animale (Jeffers *et al.*, 2001) (Wiedmann *et al.*, 1997).

## 4 - PHYSIOLOGIE DE *L. MONOCYTOGENES*

### 4-1. Caractères bactériologiques

Le genre *Listeria* fermente sans gaz de nombreux glucides : glucose, fructose, mannose, amygdaline, saliciline, cellobiose, maltose, trehalose, gentobiose, D-arabitol. Les germes de ce genre sont catalase-positif, oxydase-négative et hydrolysent l'esculine.

*L. monocytogenes* est un bacille à Gram positif se présentant sous forme de bâtonnets réguliers de 0,5 à 2 µm de longueur sur 0,4 à 0,5 µm de diamètre, aux extrémités arrondies, associés parallèlement en courtes chaînes ou en paires sous forme de V. Dans les cultures âgées ou carencées, des filaments de plusieurs micromètres peuvent apparaître.

La bactérie n'est ni sporulée ni capsulée. Elle est mobile par des flagelles péritriches lorsqu'elle est cultivée entre 20 et 25°C, et immobile ou très faiblement mobile à 37°C.

Elle est capable de se développer en atmosphère aérobie ou anaérobie (Seeliger & Jones, 1987). Sur gélose nutritive, elle forme en 24-48 h à 37°C des colonies de 0,5 à 1,5 mm de diamètre, translucides à reflets bleutés en lumière oblique. A partir du glucose, cette bactérie produit essentiellement de l'acide L(+)-lactique.

Les caractères biochimiques principaux permettant de réaliser la différenciation entre les différentes espèces de *Listeria* sont présentés dans le Tableau 3.

**Tableau 3 : Caractères bactériologiques différenciant les espèces de *Listeria* (+ : positif ; - : négatif) (Rocourt *et al.*, 2000)**

	Hémolyse	CAMP test <i>S. aureus</i> *	CAMP test <i>R. equi</i> **	D-xylose	L-rhamnose	α méthyl D-mannosi de	Ribose	Mannitol
<i>L. monocytogenes</i>	+	+	-	-	+	+	-	-
<i>L. ivanovii</i> ( <i>subsp ivanovii</i> )	+	-	-	+	-	-	+	-
<i>L. ivanovii</i> ( <i>subsp londoniensis</i> )	+	-	+	+	-	variable	-	-
<i>L. innocua</i>	-	-	-	-	variable	+	-	-
<i>L. welshimeri</i>	-	-	-	+	variable	+	-	-
<i>L. seeligeri</i>	+	+	-	+	-	-	-	-
<i>L. grayi</i>	-	-	-	-	-	non défini	-	+

\* *Staphylococcus aureus* ; \*\* *Rhodococcus equi*

Il peut également être utile de différencier *L. monocytogenes* et les autres *Listeria* de certains genres pouvant donner des colonies d'aspect voisin, en particulier sur gélose au sang (Tableau 4).

**Tableau 4 : Caractères différenciant les listeria de genres bactériens présentant des caractères phénotypiques voisins (Rocourt *et al.*, 2000)**

	<i>Listeria</i>	<i>Erysipelothrix</i>	<i>Corynebacterium</i>	<i>Enterococcus</i> <i>Streptococcus</i>
Mobilité à 25°C	+	-	-	-
	(peu mobile à 37°C)			
Catalase	+	-	+	-
Esculine	+	-	-	variable
	(positivité en 2-3 heures)			
H <sub>2</sub> S	-	+	-	-
Inclusions métachromatiques	-	-	+	-

## **4-2. Conditions de survie et de multiplication**

*L. monocytogenes* est une bactérie ubiquiste, très largement répandue dans l'environnement. Ce n'est pas un germe très exigeant et sa culture est obtenue sur les milieux nutritifs classiques. Cependant, cette bactérie ne peut pas se développer dans des milieux carencés en fer. En effet, l'apport de citrate de fer et d'esculine stimule sa croissance.

*L. monocytogenes* possède un certain nombre de caractéristiques lui permettant de survivre et de croître dans les aliments mais également de coloniser de nombreuses niches écologiques au sein des industries agroalimentaires. Parmi ses caractéristiques, sa capacité à croître dans une large gamme de températures et notamment aux températures de réfrigération est déterminante. En effet, sa croissance est démontrée expérimentalement entre - 2°C et +45°C (Augustin *et al.*, 1999). Son optimum de croissance est situé à 30-37°C. Par contre, *L. monocytogenes* est rapidement détruite à 60°C (Breand *et al.*, 1998) et n'est donc pas considérée comme un germe thermorésistant. Toutefois, certains prétraitements thermiques peuvent influencer sa thermotolérance (Pagan *et al.*, 1997). Ceci pourrait contribuer à la survie de *L. monocytogenes* dans les aliments subissant un préchauffage avant pasteurisation.

Une autre propriété importante est sa capacité à se développer en présence d'une forte concentration en sel (jusqu'à 10% NaCl), ce qui explique sa persistance dans les aliments salés. Il a même été montré que certaines souches pouvaient survivre dans des saumures de fromagerie contenant de 13 à 14% de NaCl (Farber *et al.*, 1992). Son  $a_w$  limite est de 0.93 (Nolan *et al.*, 1992), le germe restant viable mais sans multiplication pour des valeurs d' $a_w$  plus faibles.

*L. monocytogenes* se multiplie entre pH 4.6 et 9.6 (Pearson & Marth, 1990) avec un optimum de croissance à pH 7.1. Elle peut toutefois survivre pendant de très longues périodes à des pH proches de pH 4 comme c'est le cas dans les ensilages de maïs.

La capacité de *L. monocytogenes* à se multiplier dans une large gamme de température, pH,  $a_w$ , mais également ses exigences nutritives modérées expliquent la capacité de cette bactérie à croître dans des conditions très défavorables et à coloniser de nombreuses niches écologiques.

Ses capacités de résistance et d'adaptation constituent donc un risque sérieux pour les industries agroalimentaires et classe cette bactérie parmi les agents pathogènes les plus surveillés.

## 5 - ENVIRONNEMENT

### 5-1. Niches écologiques

Malgré l'absence de spores, les capacités de résistance de *Listeria* dans le milieu extérieur sont remarquables. Elle est notamment très largement répandue dans l'environnement (sols, végétaux, pâturages, eaux douces, eaux de mer, vase, eaux d'égouts), dans les locaux d'élevage (litière, sol, parois, fenêtres, mangeoires, abreuvoirs...) et dans les locaux d'habitation (torchons, serpillières, périphérie des conduites d'évacuation, réfrigérateurs, brosses à dents) (Beumer *et al.*, 1996). L'environnement est essentiellement contaminé par les fèces d'animaux sains ou malades. En effet, 10 à 30% des bovins, ovins, porcins et poulets hébergent naturellement cette bactérie dans leur tube digestif, ainsi que 1 à 20% des humains (chapitre 1 § 9 -). Il est connu que les troupeaux de ruminants sont essentiellement contaminés par une alimentation à base d'ensilages (Wilesmith & Gitter, 1986).

La listériose humaine est également majoritairement transmise par l'alimentation. Toutes les grandes catégories d'aliments peuvent être contaminés, qu'il s'agisse des produits laitiers, des produits carnés, des végétaux ou encore des poissons ou crustacés, mais avec des fréquences et des taux variables.

La listériose humaine est présente dans les pays industrialisés mais elle est quasiment absente des pays en voie de développement. Outre les différences existant dans les moyens de diagnostic et de surveillance sanitaire, cette répartition géographique s'expliquerait paradoxalement par une meilleure hygiène et par la généralisation de la chaîne du froid dans

les pays développés. En effet, *L. monocytogenes* pouvant survivre et se développer à 4°C, la réfrigération des aliments permettrait une sélection de cette bactérie.

La contamination peut survenir à tous les stades de la fabrication et de la distribution mais elle peut aussi se produire chez le consommateur.

## **5-2. *Listeria monocytogenes* dans l'industrie agroalimentaire**

Les études microbiologiques menées dans l'environnement des élevages ont montré une faible contamination des filières porcines et avicoles. Par contre, différentes enquêtes menées dans les filières de production de viande montre une amplification entre l'élevage et l'abattoir puis la découpe. Lors de cette étape, les couteaux, les tapis et autres machines sont fréquemment contaminés par *L. monocytogenes* et sont à l'origine de la contamination. De fortes contaminations des surfaces de travail ont également été montrées dans les salaisons.

L'existence de souches «résidentes» a été confirmée par des comparaisons génotypiques. Il semble donc que des carences en opérations de nettoyage et désinfection constituent une source potentielle de *L. monocytogenes* dans les entreprises agroalimentaires.

Les fréquences de contamination du lait cru sont variables selon les études effectuées dans différents pays (0.3 à 7%) mais sont plus faibles que les fréquences de contamination des viandes (2 à 60%). Concernant les fromages, les fréquences de contamination sont très variables d'un type de fromage à un autre. Ainsi les fromages frais qui présentent un pH acide, inférieur ou égal à pH 4.5 ne permettent pas la croissance et la survie de la bactérie (fromages de chèvre, parmesan...). D'autres fromages permettent la survie de *L. monocytogenes* mais inhibent sa croissance (fromages à pâte pressée, bleus). Enfin, les fromages à pâte molle, affinés, à croûte fleurie (Camembert, Brie...) ou à croûte lavée (Munster, Maroilles...) permettent la croissance. Concernant les yaourts, le pH acide atteint en fin de fabrication permet d'inhiber la croissance de *L. monocytogenes*. Les crèmes et le beurre sont rarement contaminés.

# **6 - PATHOLOGIE**

## **6-1. Pathologie animale**

La listériose est une infection essentiellement animale, accidentellement humaine. Elle sévit de façon sporadique chez les animaux, mais peut évoluer de façon endémique dans certains élevages en fonction des techniques d'élevage. Elle semble être présente surtout dans les zones tempérées, mais on peut la rencontrer dans le monde entier. La contamination des

animaux s'effectue le plus généralement par ingestion des végétaux. Ce sont donc les herbivores qui sont principalement atteints. *L. monocytogenes* est reconnue comme pathogène pour les animaux et surtout chez l'homme. *L. ivanovii* peut être pathogène pour les petits ruminants, particulièrement chez les ovins (Chand & Sadana, 1999). *L. innocua*, bien que fréquemment isolée chez des animaux présentant des signes cliniques, n'est généralement pas considérée comme pathogène. Dans certaines exploitations, son rôle pathogène est évoqué et mériterait une réelle expertise (Walker *et al.*, 1994). Les principales formes cliniques de la maladie sont des avortements, des entérites, des sépticémies et des formes nerveuses. Les formes nerveuses sont essentiellement décrites sous l'appellation de «circling disease» car les animaux atteints se déplacent difficilement, titubent et tournent en rond toujours dans le même sens avant de tomber.

## **6-2. Pathologie humaine**

*L. monocytogenes* est la seule espèce du genre *Listeria* à être considérée comme pathogène chez l'homme. Cependant, bien que *L. ivanovii* est considérée comme non pathogène pour l'homme, il est à noter que 8 cas de listériose dus à cette espèce ont été rapportés (Cummins *et al.*, 1994 ; Lessing *et al.*, 1994 ; Snapir *et al.*, 2006).

*L. monocytogenes* est à l'origine d'une maladie grave, la listériose. La plupart des adénies de listériose ont été observées en Europe et en Amérique du nord. Cette infection est rare comparée aux autres maladies transmises par les aliments, puisqu'elle a une incidence de 2 à 10 cas par million d'habitants par an. Cependant, cette maladie entraîne une hospitalisation quasi systématique et un nombre élevé de décès. En effet, environ 25-30% des personnes atteintes de listériose décèdent, ce qui en fait une des plus fréquente cause de décès dû à une contamination alimentaire. Pour cette catégorie de maladie, elle est la deuxième cause de mortalité juste après la salmonellose en France (Vaillant *et al.*, 2005) et aux Etats-Unis (28% des décès dus à une infection alimentaire (Mead *et al.*, 1999)). Malgré les efforts des industriels de la filière alimentaire et la mise en place d'un système de surveillance efficace en France qui a permis de diminuer les cas de listériose humaine, le risque sanitaire reste donc important. De plus, alors que les cas de listériose ont diminués en France ces dernières années, il est important de noter que cette maladie a progressé dans certains pays comme l'Allemagne avec 510 cas de listérioses recensés en 2005 contre seulement 217 en 2001.

La listériose affecte essentiellement les personnes dont le système immunitaire est perturbé. Il s'agit notamment des femmes enceintes et de leur fœtus ou de leur nouveau-né,

des personnes âgées et des personnes immunodéprimées (séropositivité au VIH, cancer, transplantés, diabétiques...).

On peut schématiquement diviser les listérioses humaines en deux grands groupes d'infection :

- Les formes foeto-maternelles et néonatales.
- Les formes de l'adulte.

Chez la femme enceinte, la listériose se traduit par un épisode fébrile d'allure pseudo grippale, et n'entraîne quasiment jamais d'infections du système nerveux. Elle peut entraîner la contamination du fœtus (listériose foeto-maternelle) ou de l'enfant lors de l'accouchement (listériose néonatale). *L. monocytogenes* infectant le fœtus peut entraîner la mort *in utero*, un avortement ou un accouchement prématuré selon le stade de la grossesse. La listériose affectant le nouveau-né se traduit sous forme septicémique qui peut évoluer vers une forme méningée.

En ce qui concerne la forme adulte, elle se manifeste essentiellement par une bactériémie ou une atteinte du système nerveux central (méningites, méningo-encéphalites, encéphalites). Il a également été montré que la listériose pouvait parfois être à l'origine de cas de gastroentérites graves (Dalton *et al.*, 1997 ; Sim *et al.*, 2002), et que cette forme touchait essentiellement des personnes en bonne santé.

En plus d'une mortalité élevée, les séquelles neurologiques suite à une listériose sont fréquentes en dépit de l'antibiothérapie.

## 7 - EPIDEMIOLOGIE

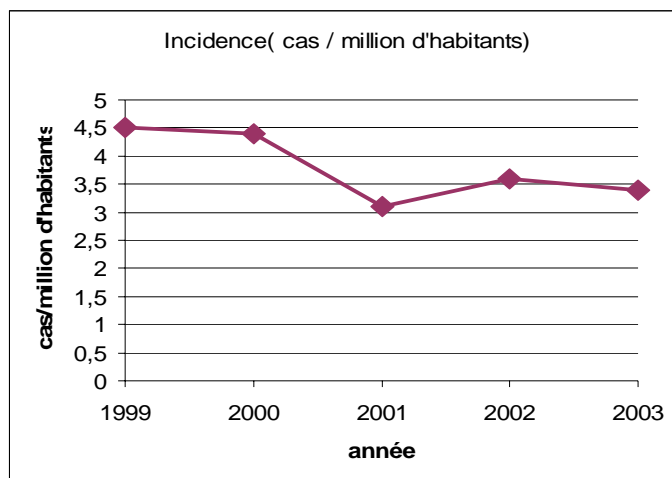
### 7-1. Système de surveillance en France

La surveillance de la listériose en France est réalisée par l'intermédiaire de la déclaration obligatoire (DO) depuis 1999 et du Centre national de référence (CNR) situé à l'institut Pasteur depuis 1990 qui centralise et caractérise les souches de *L. monocytogenes* provenant des laboratoires de microbiologie. La DO permet au médecin de la Direction Départementale des Affaires Sanitaires et Sociales (DDASS) de recueillir de façon précoce des informations sur la consommation alimentaire des patients. La DO a également comme objectif de disposer d'informations sur les patients (forme clinique, âge, terrain à risque...), afin de suivre les tendances évolutives de cette maladie.

Lorsque le CNR repère un groupement dans le temps de souches "identiques" (non différenciables par la méthode de typage utilisée), il en informe les membres de la "cellule Listeria" chargée de la coordination des investigations et des actions, constituée de représentants de la Direction Générale de la santé (DGS), de l'Institut de Veille Sanitaire (InVS), de la Direction Générale de l'Alimentation (DGAI), de la Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes (DGCCRF) et du CNR. L'InVS analyse les informations concernant ces patients (fiches DO et questionnaires alimentaires) et décide ou non, en fonction de cette analyse, de passer en phase d'alerte. Lorsqu'il y a alerte, les différents partenaires de la cellule de coordination décident des investigations à entreprendre afin d'identifier un éventuel aliment commun à l'origine de ces cas et d'éviter de nouveaux cas par la mise en œuvre de mesures appropriées.

## **7-2. Caractéristiques épidémiologiques en France**

Sur la période 1999-2003, on observe une diminution de l'incidence en 2001, qui se stabilise en 2002 et 2003 (Figure 2). En effet, l'incidence des cas de listériose est passée de 4.5 cas / million d'habitants en 1999 à 3.4 cas / million d'habitants en 2003 ([www.invs.fr](http://www.invs.fr)).



**Figure 2 : Incidence de la listériose (cas / million d'habitant) entre 1999 et 2003 en France**

Malgré un nombre de cas moins important, le nombre de décès suite à une listériose est resté stable entre 1999 et 2002 (Figure 3), avec une moyenne de 60 décès par an (létalité comprise entre 21 et 31 % selon les années). Seule l'année 2003 montre une baisse de la létalité avec 39 décès, qui représentent 18.6% des patients.



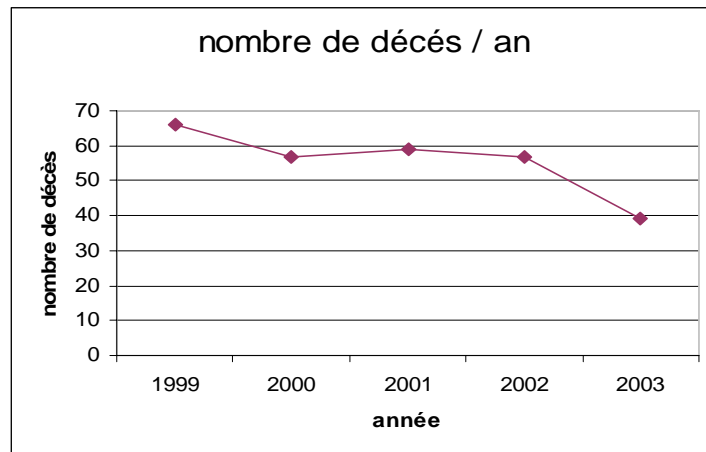


Figure 3 : Nombre de décès dus à la listériose entre 1999 et 2003

De 2001 à 2003, on observe une variation saisonnière, avec une augmentation estivale notable des cas (période de mai à août) (Figure 4).

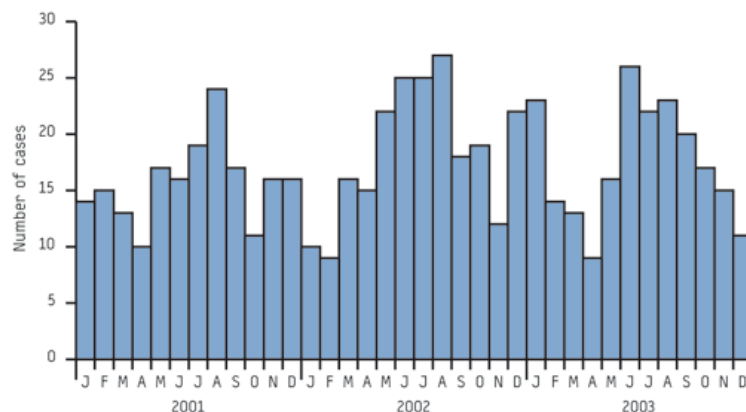
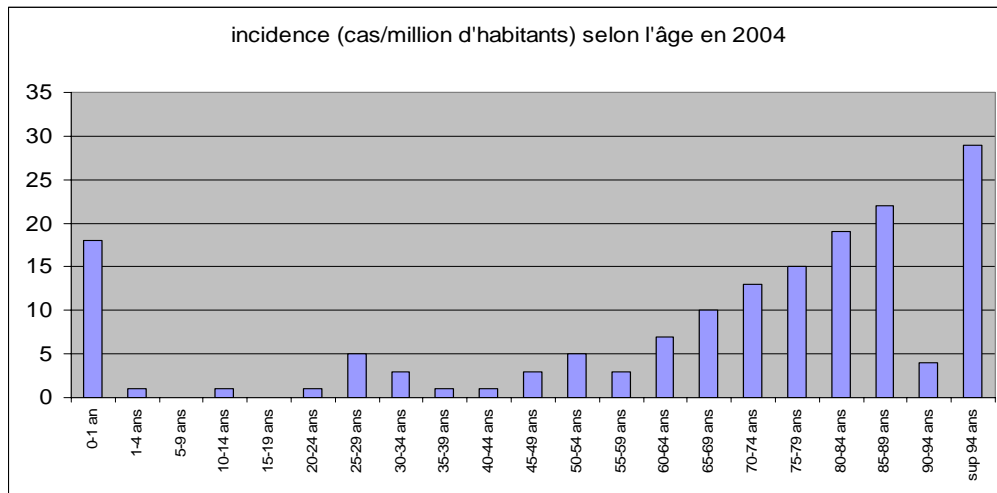


Figure 4 : Nombre de cas de listériose en fonction des mois en 2001, 2002 et 2003

**Les formes materno-néonatales** représentent 24% des cas de listériose et la mortalité concerne uniquement le nouveau-né et le fœtus, les femmes enceintes présentant généralement des manifestations pseudo-grippales ou aucune manifestation. La létalité est élevée avec une moyenne de 32.5% de décès entre 1999 et 2003.

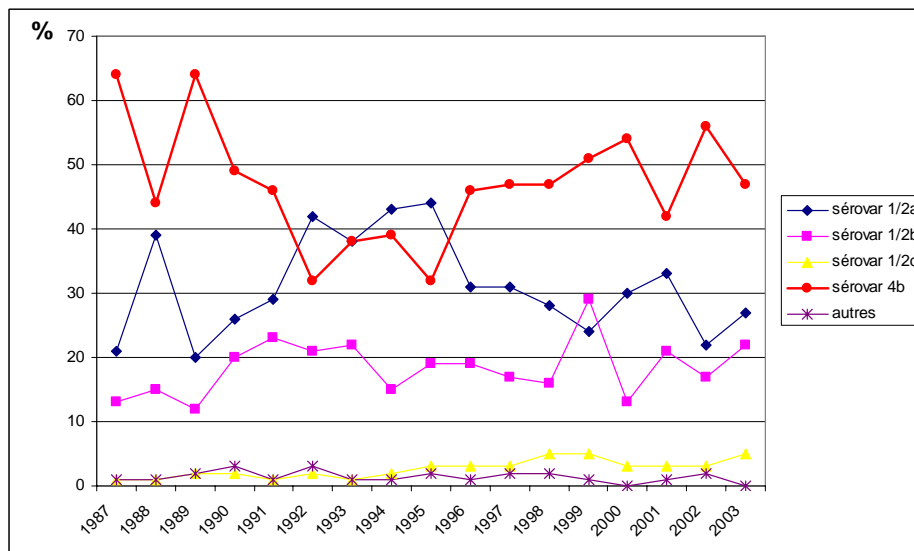
**Les formes non materno-néonatales** surviennent majoritairement chez les hommes (environ 60% des cas). Parmi les patients touchés par la listériose entre 1999 et 2003, 69.6% avaient un terrain à risque, 12.4% avaient une autre pathologie et 15 % n'avaient aucune pathologie connue au moment de l'hospitalisation. Pour 3% des patients, il n'a pas pu être précisé s'il existait un terrain à risque ou non. De plus, l'incidence augmente avec l'âge à

partir de 60 ans (Figure 5). La létalité est de 21.5 % (oscillant entre 16% et 32% selon les années).



**Figure 5 : Incidence de la listériose (cas / million d'habitants par an) selon l'âge en 2004**

Les souches à l'origine des cas de listériose humaine appartiennent majoritairement aux sérovars 4b, 1/2a et 1/2b. La prédominance des souches de sérovar 4b est observée depuis 1987, à l'exception de la période 1992 à 1995 où les souches de sérovar 1/2a étaient prépondérantes (Figure 6).



**Figure 6 : Sérovars à l'origine des cas de listériose en France sur la période 1987-2003**

Au sein des souches à l'origine des formes non materno-néonatales, le sérovar 4b est plus fréquemment retrouvé dans les infections du système nerveux central (56% des cas) que lors d'une bactériémie (41%).

De même, le serovar 4b est majoritairement à l'origine des cas groupés de listériose avec 15 des 25 cas signalés par le CNR entre 2001 et 2003 (60%).

## 8 - RELATION DOSE-REPONSE

Malgré une contamination importante des aliments (Chapitre 1 § 5-2) et donc une exposition régulière du consommateur, l'incidence de la listériose chez l'homme est faible. Notermans *et al.* (1998) ont montré par des tests sur souris que ce décalage ne pouvait pas s'expliquer par une atténuation de la virulence de certaines souches. Cette observation pourrait plus facilement s'expliquer par le fait que la listériose n'apparaît que suite à une dose importante de bactéries ingérées. Cette dose infectante ne peut évidemment pas être déterminée, pour des raisons éthiques, à l'aide de volontaires humains, comme cela a été fait pour une infection à Rotavirus (Ward *et al.*, 1986). De par une période d'incubation longue et une possible multiplication de *L. monocytogenes* aux températures de réfrigération, la densité bactérienne de l'aliment incriminé est difficile à établir. Cependant, les résultats des enquêtes réalisées sur les produits incriminés dans les anadémies font état de niveaux de contamination supérieurs à  $10^2$  *L. monocytogenes*/g d'aliment (Tableau 5).

**Tableau 5 : Anadémies survenues depuis 1980 en France et dans d'autres pays (au 01/07/00) (Rapport AFSSA)**

Pays	Année	Nombre de patients	Véhicule	Niveau de contamination (UFC/g)*	Référence
Canada	1980	41	Coleslaw (chou)		Schlech <i>et al.</i> 1983
USA	1983	49	lait pasteurisé		Fleming <i>et al.</i> 1985
USA	1985	> 142	Fromage à pâte molle « mexican style »	$10^3$ - $10^4$ D	Linnan <i>et al.</i> 1988
Suisse	1983-87	122	Fromage à pâte molle « Vacherin »	$10^4$ - $10^6$ D	Bille 1990
Royaume Uni, Irlande	1987-89	> 350	Pâté	$10^2$ - $10^6$ D	McLauchlin <i>et al.</i> 1991
USA	1989	10	Crevettes		Riedo <i>et al.</i> 1994
Australie	1990	9	Pâté	$10^3$ D et C	Kittson 1992
France	1992	279	Langue de porc en gelée Charcuterie à la coupe	$10^4$ - $10^6$ D, < 100	Goulet <i>et al.</i> 1993 Rocourt <i>et al.</i> 1993
Nouvelle Zélande	1992	2	Moules fumées		Brett <i>et al.</i> 1998
France	1993	38	Rillettes	$10^1$ - $10^3$ D $10^4$ C	Goulet <i>et al.</i> 1998a Goulet <i>et al.</i> 1998b
Suède	1994-95	9	Truite « gravad »	$10^2$ - $10^6$ D	Ericsson <i>et al.</i> 1997
France	1995	36	Fromage à pâte molle « Brie »		Goulet <i>et al.</i> 1995 Vaillant <i>et al.</i> 1998
France	1997	14	Fromage à pâte molle « Pont l'Evêque », « Livarot »		Jacquet <i>et al.</i> 1998
France	1999	3	Fromage à pâte molle type « Epoisses »		Non publié
USA	1998-99	101	Hot dog, « deli meat »	< 0,3 D et C	Mead 1999
Finlande	1999	18	Beurre	$10^1$ - $10^4$ D et C	Lyytikäinen <i>et al.</i> 1999
France	1999	10	Rillettes	<10 P $10^3$ C et $10^6$ C	De Valk <i>et al.</i> 2000
France	1999-2000	32	Langue de porc en gelée Charcuterie		Non publié

Chen *et al.* (2006) ont montré que la dose-réponse était apparemment dépendante du sérovar de la souche. De plus, il est à noter que cette dose infectante est certainement variable en fonction du statut immunitaire du patient.

## 9 - PORTAGE ASYMPTOMATIQUE

Etant donné la faible incidence de la listériose, on peut également se demander s'il existe un portage de *L. monocytogenes* ne provoquant pas de maladie. Les publications sur ce sujet montrent une faible fréquence de portage asymptomatique, puisque *L. monocytogenes* a été trouvé dans 0.12% à 4.7% des échantillons fécaux prélevés sur des adultes en bonne santé (Berger & Pietsch, 1975 ; MacGowan *et al.*, 1994 ; Cobb *et al.*, 1996 ; Grif *et al.*, 2001 ; Sauders *et al.*, 2005). Cependant, cette fréquence est élevée par rapport au nombre de cas de listériose, puisque seulement 0.00034% de la population est touchée par la listériose chaque année.

La faible fréquence de portage fécal de *L. monocytogenes* pourrait s'expliquer par un effet protecteur de l'acide gastrique contre le passage d'organismes pathogènes. En effet, Cobb (1996) a montré une importante augmentation de la fréquence de *L. monocytogenes* (20% de portage) chez les patients recevant un traitement supprimeur de l'acide gastrique. Cependant, aucun de ces patients n'a développé de listériose. La plupart des souches isolées de ces échantillons fécaux appartenaient aux mêmes sérovares que ceux principalement retrouvés dans les cas de listériose, c'est-à-dire les sérovares 1/2a, 1/2b et 4b.

Grif *et al.* (2001) ont suivi trois volontaires adultes en bonne santé pendant un an. En moyenne, deux épisodes de portage de *L. monocytogenes* ont été observés par personne et le portage fécal ne durait pas plus de 4 jours.

## 10 - POUVOIR PATHOGENE ET FACTEURS DE VIRULENCE

### 10-1. Voies d'infection chez l'homme

La gravité des infections à *L. monocytogenes* est liée au pouvoir invasif de ce pathogène qui est capable de traverser la barrière intestinale ou placentaire et de pénétrer le système nerveux central. De plus, *L. monocytogenes* est une bactérie intracellulaire qui peut notamment envahir les cellules épithéliales et les macrophages et s'y multiplier, la mettant à l'abri du système immunitaire.

La porte d'entrée de l'infection chez l'homme est le plus souvent le tube digestif à la suite de l'absorption d'aliments contaminés (Figure 7). Après avoir traversé la barrière intestinale, *L. monocytogenes* se retrouve dans les cellules phagocytaires de la *lamina propria*. La dissémination se fait par voie lymphatique et sanguine et permet à la bactérie d'atteindre le foie et la rate. Il a été montré chez la souris que *L. monocytogenes* est alors rapidement éliminée par les macrophages résidents (Ebe *et al.*, 1999). Les bactéries survivantes vont infecter les hépatocytes. La lyse des hépatocytes par les neutrophiles libère les bactéries qui peuvent être phagocytées par les neutrophiles ou les macrophages. Chez les personnes possédant un bon système immunitaire, *L. monocytogenes* est rapidement éliminée. Ceci est certainement l'événement se produisant le plus couramment suite à une infection à *L. monocytogenes*, ce qui expliquerait la rareté des cas de listériose comparée à la fréquence des aliments contaminés. De plus il a été montré que des lymphocytes T dirigés contre des antigènes de *Listeria* étaient couramment retrouvés chez des personnes en bonne santé (Munk & Kaufmann, 1988), ce qui pourrait s'expliquer par une stimulation courante du système immunitaire par *L. monocytogenes* fréquemment présente dans les aliments. Il semblerait donc que chez les individus en bonne santé, l'ingestion d'une faible dose de *L. monocytogenes* n'aurait d'autres effets que la stimulation du système immunitaire.

Chez les personnes immunodéprimées ou si la dose ingérée est trop élevée, l'infection peut ne pas être stoppée par le système immunitaire et entraîner une prolifération importante dans le foie, entraînant le largage de bactéries dans la circulation sanguine. Ceci peut causer une septicémie, la colonisation du système nerveux central ou du placenta.

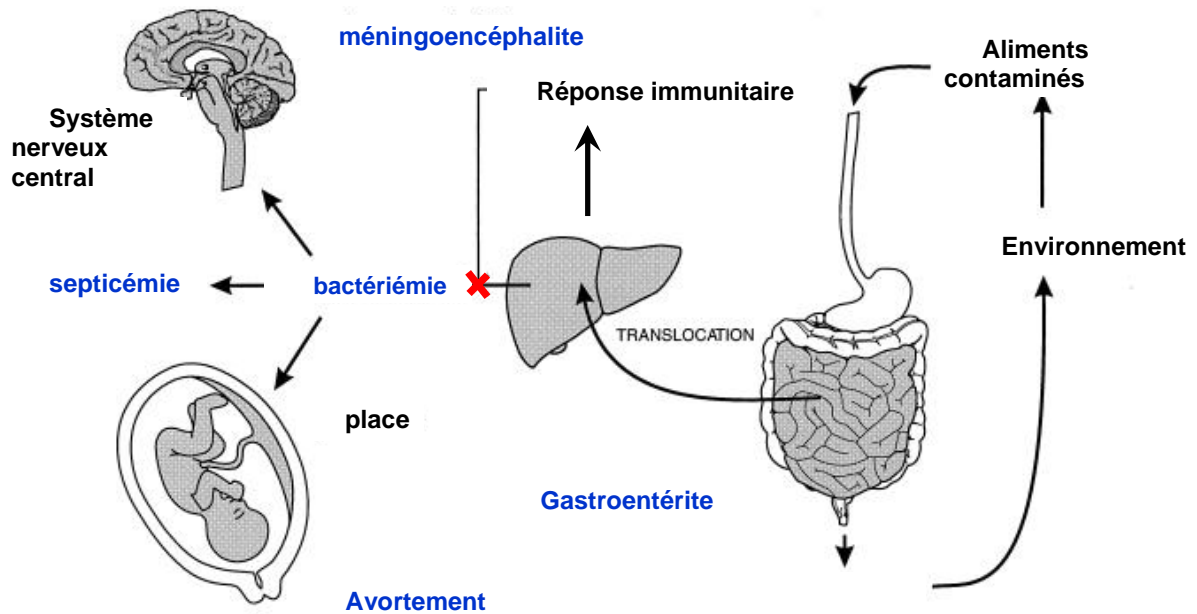


Figure 7 : Voies d'infection de *L. monocytogenes* chez l'homme (Vazquez-Boland *et al.*, 2001)

## 10-2. Facteurs de virulence

En plus d'envahir les macrophages, *L. monocytogenes* peut entraîner sa propre internalisation dans de nombreuses cellules dont les cellules épithéliales, les fibroblastes, les hépatocytes, les cellules endothéliales et des cellules nerveuses. Après pénétration dans ces différentes cellules, la bactérie réalise son cycle infectieux intracellulaire qui se décompose en plusieurs étapes (**Erreur ! Source du renvoi introuvable.**) : internalisation de la bactérie dans la cellule hôte, multiplication intra cytoplasmique, propulsion jusqu'à la cellule voisine et formation d'une protrusion\_endocytée par une autre cellule, lyse de la vacuole à double membrane et début d'un autre cycle. L'étude de ce processus a conduit à la mise en évidence de plusieurs facteurs de virulence.

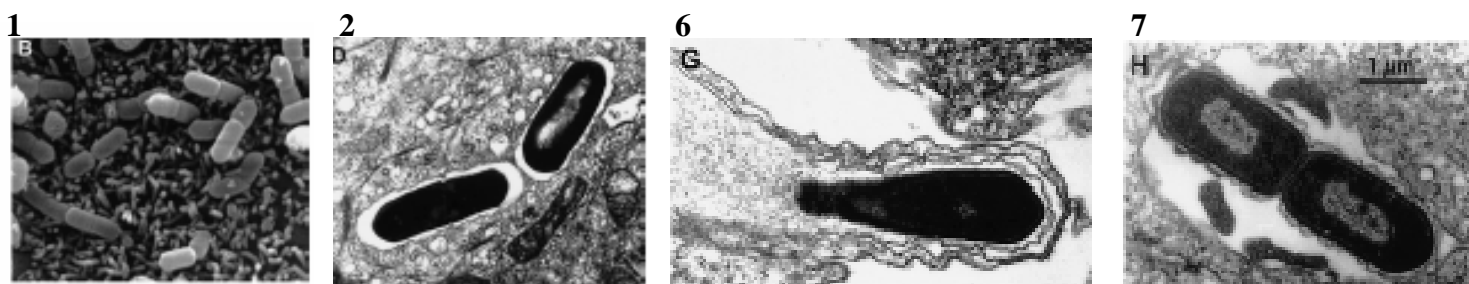
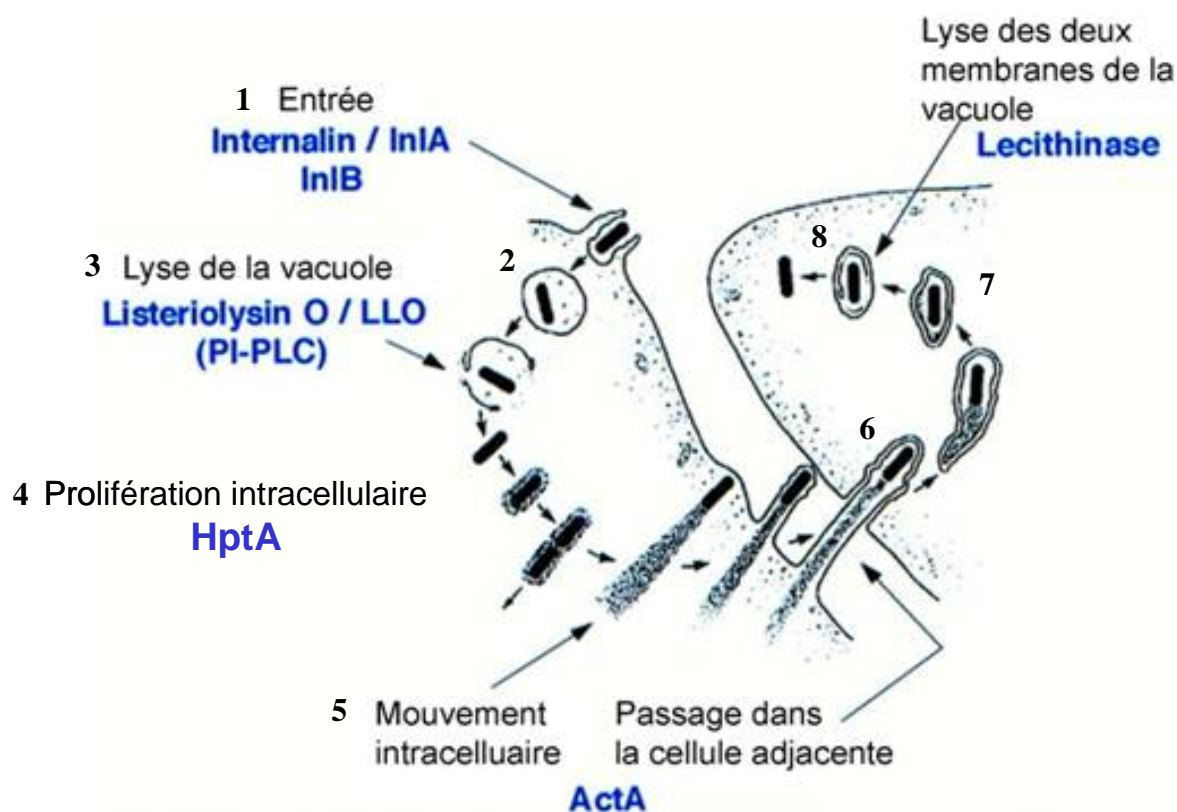


Figure 8 : Cycle d'infection intracellulaire de *L. monocytogenes* (d'après (Vazquez-Boland *et al.*, 2001))

La bactérie induit sa propre internalisation à l'intérieur de la cellule (1) grâce principalement aux internalines codées par la bactérie (et notamment InlA et InlB). Après la phagocytose, la bactérie est internalisée dans une vacuole (2). *L. monocytogenes* va s'échapper de la vacuole. La destruction de la vacuole est principalement due à la listeriolysine O et à la phospholipase C (3). Une fois dans le cytoplasme, *L. monocytogenes* va se multiplier, notamment grâce à HptA qui permet de capturer les hexoses phosphates utilisés comme source de carbone (4). Après multiplication, la bactérie va se déplacer dans le cytoplasme grâce à la formation d'une queue d'actine (5) jusqu'à atteindre la membrane cytoplasmique. La formation de cette queue d'actine est sous la dépendance du gène *actA*. En

contact avec la membrane cytoplasmique, la bactérie va induire la formation de protubérances cellulaires (6) qui seront phagocytées par la cellule adjacente pour donner naissance à des vacuoles à deux membranes (7). Après lyse de ces vacuoles (8) grâce à l'action combinée de deux phospholipases C (codées par *plcA* et *plcB*) et de la listériolysine O, la bactérie peut initier un nouveau cycle infectieux.

### 10-2.1. Adhésion et entrée dans la cellule hôte

L'entrée de la bactérie dans les cellules phagocytaires professionnelles se fait par phagocytose. Pour l'invasion des cellules non phagocytaires, plusieurs protéines ont été identifiées comme étant impliquées dans l'adhésion et l'invasion des cellules eucaryotes.

#### 10-2.1.1. Les internalines

Les internalines sont une famille de protéines caractérisées par un domaine en N-terminal contenant plusieurs régions de 22 acides aminés, riches en leucine (LRR = Leucine Rich Region). Elles possèdent également un peptide signal qui leur permet d'être exportées à la surface ou dans le milieu extracellulaire (Cabanès *et al.*, 2002). Le séquençage du génome de *L. monocytogenes* EGDe a permis de révéler l'existence d'un nombre important d'internalines (Glaser *et al.*, 2001). Cette famille peut être divisée en trois classes :

- Les internalines possédant un motif LPXTG en région C-terminale qui permet leur ancrage à la surface de la bactérie. InlA est la protéine la mieux caractérisée de cette classe qui contient 19 membres au total
- La deuxième classe ne comprend qu'une seule internaline chez EGDe, InlB, qui contient en C-terminal une région répétée de 80 acides aminés commençant par le motif GW. Ce motif permet une liaison non covalente d'InlB à la surface bactérienne.
- La dernière classe ne possède pas de motif d'ancrage à la surface et les internalines de cette famille sont donc sécrétées dans le milieu extracellulaire.

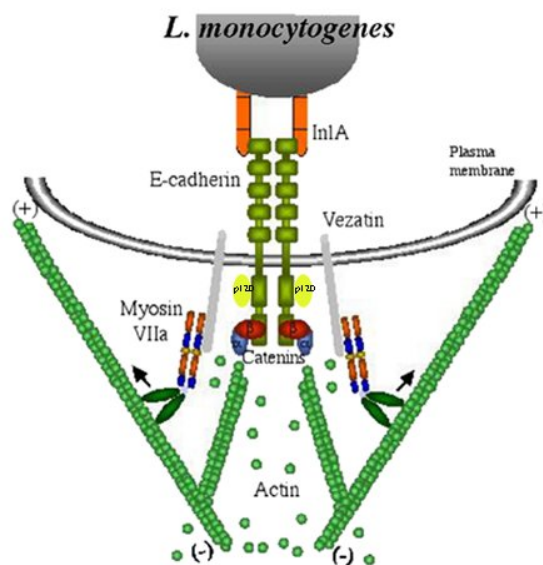
Les premières internalines découvertes, InlA et InlB, sont les deux principales protéines permettant l'entrée de *L. monocytogenes* dans les cellules non phagocytaires. Ces deux protéines permettent l'entrée dans des cellules bien spécifiques, certainement due à des récepteurs au niveau de la cellule hôte, bien différents entre les deux protéines. En effet, InlA permet d'entrer dans les cellules exprimant son récepteur, la E-cadhérine humaine, comme les cellules épithéliales, et dans certaines lignées hépatocytaires comme HepG-2 ou les cellules placentaires. InlB induit l'invasion dans de nombreux types cellulaires, dont les cellules



épithéliales (HeLa, Hep-2, Vero), hépatocytiques (HepG2 et Tib73), endothéliales (HUVEC), endothéliales microvasculaires du cerveau (HBMEC) et les fibroblastes.

- InlA

InlA est le facteur majeur permettant l'entrée de *L. monocytogenes* dans les cellules épithéliales (Gaillard *et al.*, 1991). Cette protéine est suffisante pour l'entrée dans les cellules non phagocytaires exprimant la E-cadhérine. Le domaine extracellulaire de la E-cadhérine, qui est abondamment présente à la surface de nombreuses cellules dont les cellules épithéliales, interagit avec le domaine LRR de InlA. Cette interaction InlA-E-cadhérine est espèce-spécifique puisqu' InlA reconnaît la E-cadhérine humaine mais pas la E-cadhérine de souris. Le domaine LRR de InlA forme une cavité permettant d'englober la région N-terminale de la E-cadhérine humaine (Figure 9).



**Figure 9 : Modèle de l'entrée de *L. monocytogenes* dans les cellules épithéliales grâce à l'internaline InlA (Sousa *et al.*, 2004)**

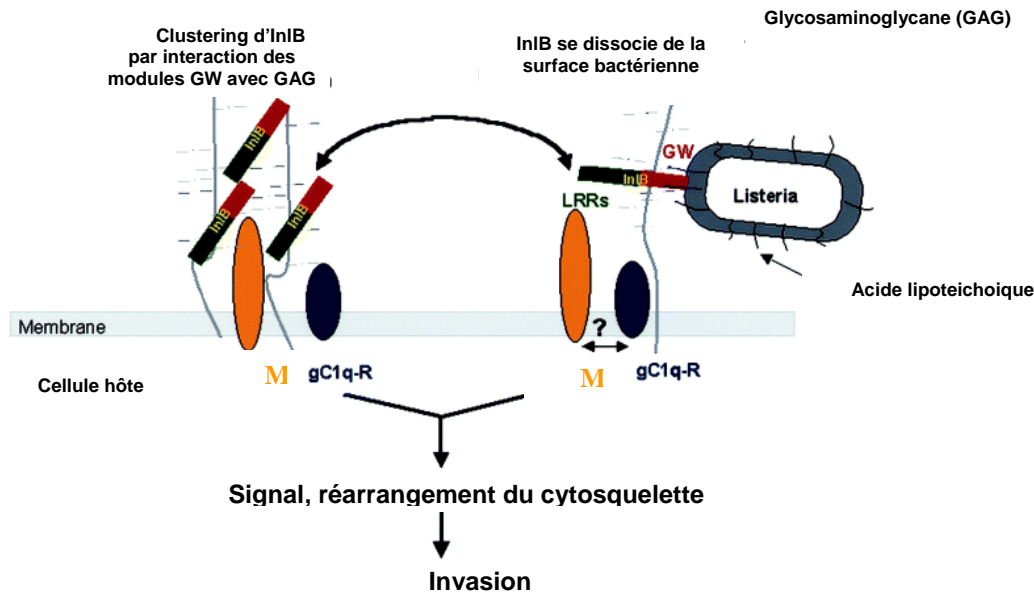
L'accès de la E-cadhérine à cette cavité est bloqué quand la proline en position 16 (E-cadhérine humaine) est remplacée par l'acide glutamique (E-cadhérine murine) (Lecuit *et al.*, 1999). Le rôle d'InlA a été évalué dans des souris transgéniques chez qui la E-cadhérine humaine était exprimée au niveau intestinal. L'infection de ces souris transgéniques avec la souche sauvage de *L. monocytogenes* ou une souche mutée au niveau d'InlA a montré l'importance d'InlA pour traverser la barrière intestinale (Lecuit *et al.*, 2001). Chez les cobayes, qui possèdent une E-cadhérine reconnue par InlA, l'infection du foie, de la rate, de la

lymphe et du petit intestin était fortement diminuée après l'inoculation orale d'une souche mutée au niveau d'InlA par rapport à la souche sauvage. Ces résultats montrent donc l'importance d'InlA pour traverser la barrière intestinale chez les espèces possédant une E-cadhérine permissive. InlA est également importante pour traverser la barrière placentaire (Lecuit *et al.*, 2004) Il est intéressant de noter que la barrière hémato-méningée est composée d'endothélium microvasculaire et d'épithélium des plexus choroïdes exprimant la E-cadhérine. Il est donc possible que l'interaction entre InlA et la E-cadhérine soit à l'origine de l'invasion du système nerveux central par *L. monocytogenes*.

- InlB

Le gène InlB a un rôle dans la virulence, principalement pour la colonisation hépatique, comme montré dans un modèle murin après une infection intraveineuse (Gaillard *et al.*, 1996). InlB est une protéine indispensable pour la colonisation de nombreux types cellulaires pour lesquels InlA n'a aucun rôle, comme les hépatocytes, les cellules endothéliales et les fibroblastes. InlB possède un domaine LRR, comme toutes les internalines et une série de répétitions commençant avec le dipeptide GW en C-terminal. Ces deux domaines sont impliqués dans la reconnaissance du récepteur cellulaire de l'hôte. GW permet l'attachement de manière non covalente de la protéine à la surface de la bactérie. InlB peut également être sécrétée dans le milieu extracellulaire et agir comme une molécule soluble (Jonquieres *et al.*, 1999). InlB peut également être sécrétée dans le milieu extracellulaire et agir comme une molécule soluble (Jonquieres *et al.*, 1999). InlB interagit avec de nombreux récepteurs cellulaires, Gc1qR (Braun *et al.*, 2000), Met (Shen *et al.*, 2000) et GAG (Jonquieres *et al.*, 2001), ce qui permet l'entrée de la bactérie. Comme pour InlA, InlB est espèce-spécifique. Il est capable de promouvoir l'entrée dans un certain nombre de cellules grâce à son interaction avec Met chez les cellules humaines et murines, mais pas dans les cellules de cobayes et de lapins, conséquence probable de différences entre les espèces au niveau de Met.

Le module GW interagit avec gC1qR et GAGs et le domaine LRR se lie à Met (Figure 10). Chacune de ces régions (module GW et le domaine LRR) participent à l'activation de Met et à l'entrée dépendant d'InlB. L'interaction d'InlB avec GAGs à travers ses modules GW conduit à son détachement de la surface bactérienne et son attachement à Met par son domaine LRR, ce qui favorise l'activation de la voie de signalisation de Met (Jonquieres *et al.*, 2001). Les voies de signalisation activées par InlB conduisent au réarrangement du cytosquelette et à l'entrée de *L. monocytogenes* dans la cellule hôte.



**Figure 10 : Modèle illustrant la synergie entre les domaines N-ter et C-ter de InlB afin d'induire le signal permettant l'invasion des cellules de l'hôte (Jonquieres *et al.*, 2001)**

InlA et InlB sont donc deux protéines majeures de la virulence de *L. monocytogenes* puisqu'elles sont impliquées dans l'invasion des cellules de l'hôte. Ces deux protéines sont absentes de *L. ivanovii*, souche pathogène chez les animaux, ce qui suggère que les internalines sont critiques pour les interactions des *Listeria* avec leurs hôtes.

#### 10-2.1.2. Ami

Ami est une amidase de 917 acides aminés qui a une activité lytique sur la paroi de *Listeria*. Cette protéine semble avoir un rôle dans la virulence puisque des mutants *ami* ont une virulence atténuée. Ami semble avoir un rôle dans l'adhésion aux cellules hôte.

#### 10-2.1.3. P60

p60, une protéine de 60kDa codée par le gène *iap*, contient une séquence signal classique, suivie d'une région qui possède un domaine SH3 dont la fonction est inconnue. La région centrale contient une série de répétitions thréonine-arginine. La partie C-terminale possède un domaine avec une activité muréine hydrolase, nécessaire à la formation normale du septum et qui est essentielle pour la viabilité de la cellule (Wuenscher *et al.*, 1993). P60 a un rôle dans l'invasion intestinale et dans la survie *in vivo* (Hess *et al.*, 1995). Il a également été mis en évidence que p60 se liait spécifiquement aux cellules intestinales Caco-2 (Park *et al.*, 2000)

#### 10-2.1.4. FbpA

FbpA est une protéine de 570 acides aminés qui possède une forte homologie avec les protéines liant la fibronectine, PavA de *Streptococcus pneumoniae*, Fbp54 de *S. pyogenes* et FbpA de *S. gordonii*. Alors que FbpA de *L. monocytogenes* ne possède pas de peptide signal classique, elle est présente à la surface de la bactérie (Dramsi *et al.*, 2004). Des résultats récents montrent que cette protéine est un substrat pour le système de sécrétion SecA2 (Dramsi *et al.*, 2004). Cette protéine purifiée se lie à de la fibronectine humaine immobilisée dans les puits d'une plaque de microtitration. Elle contribue à l'adhésion des cellules. En effet, l'adhésion de fibronectine exogène à des lignées cellulaires Hep-2 ne synthétisant pas de fibronectine augmente l'adhésion de la bactérie sauvage mais pas de celle mutée au niveau de *fbpA*. Il est intéressant de constater que FbpA semble moduler le taux de listériolysine O et d'InlB au niveau post-transcriptionnel, ce qui suggère que cette protéine pourrait également agir comme une molécule chaperonne empêchant la dégradation de certaines protéines de virulence.

#### 10-2.2. Survie et multiplication dans la cellule hôte

Une fois entrée dans la cellule hôte, *L. monocytogenes* se retrouve dans le compartiment phagosomal qui est rapidement acidifié par des pompes à protons. La bactérie s'échappe alors rapidement de cette vacuole grâce à l'action combinée de la listériolysine O et de deux phospholipases. Une fois dans le cytoplasme de l'hôte, la bactérie se multiplie puis envahit les cellules adjacentes.

##### 10-2.2.1. La listériolysine O (LLO)

La listériolysine O (ou LLO) a été le premier facteur de virulence identifié chez *L. monocytogenes* (Gaillard *et al.*, 1986). Cette protéine est une hémolysine de 60 kDa qui appartient à la famille des cytolysines cholestérol dépendantes. La LLO est codée par le gène *hly*. Elle a été identifiée comme un facteur absolument nécessaire à la virulence par l'analyse de mutants non hémolytiques obtenus par mutagenèse (Gaillard *et al.*, 1986 ; Kathariou *et al.*, 1987). La virulence du mutant non hémolytique était fortement réduite chez le modèle murin avec une différence de 4 log de la dose létale 50 après une inoculation intraveineuse. Les mutants non hémolytiques sont également affectés dans des modèles d'infection et de cellules épithéliales et phagocytaires, puisque ces mutants ne pouvaient pas s'échapper du phagolysosome. De plus, l'expression hétérologue du gène *hly* chez *Bacillus subtilis* permet à cet organisme de lyser la vacuole (Bielecki *et al.*, 1990).

La LLO est indispensable à la bactérie pour s'échapper de la vacuole dans laquelle elle se retrouve une fois entrée dans la cellule hôte. Le potentiel pathogène de *Listeria* a été relié à sa capacité d'induire l'hémolyse et donc à la production de LLO. Cependant, malgré son rôle fondamental dans la virulence, la LLO en elle-même n'est pas suffisante pour conférer un potentiel pathogénique à des bactéries à Gram positif non pathogènes (Portnoy *et al.*, 1992). De plus, *L. seeligeri*, qui a une activité hémolytique faible, n'est pas pathogène pour l'homme, ce qui confirme que l'hémolysine n'est pas le seul facteur permettant la virulence de *L. monocytogenes* (Seeliger & Jones, 1987).

L'activité hémolytique optimale de la LLO est observée à pH acide, situation rencontrée par *L. monocytogenes* dans le phagolysosome. Cette observation suggère que la LLO est active à l'intérieur du phagosome et inactive dans le cytosol afin d'éviter d'endommager la cellule hôte.

Une séquence PEST-like a été détectée chez la LLO. Ce motif pourrait permettre la dégradation de la LLO quand la bactérie se retrouve dans le cytoplasme afin de limiter son action à la vacuole phagocytaire et éviter ainsi la lyse de la cellule hôte. En effet, Decatur et Portnoy (2000) ont montré que des mutants délétés de la région PEST s'accumulaient dans le cytosol de la cellule hôte, suggérant que cette région permet la dégradation de la LLO. Par contre, Lety *et al.* (2003) ont montré que les bactéries sécrétant une LLO délétée de PEST, étaient incapables de s'échapper du compartiment phagosomal. La différence de résultat entre ces deux études pourrait s'expliquer par l'utilisation de lignées cellulaires différentes et donc signifier que ce motif a un rôle différent en fonction du type cellulaire. Par contre, les deux études ont montré que la délétion de ce motif n'affectait pas la sécrétion ou l'activité hémolytique de LLO mais abolissait la virulence de la bactérie.

### 10-2.2.2. Phospholipases et metaloproteases

*L. monocytogenes* sécrète deux phospholipases, PlcA et PlcB, qui permettent d'endommager la membrane, en synergie avec la LLO, et donc de s'échapper du phagolysosome. Ces deux phospholipases sont aussi importantes l'une que l'autre pour la virulence, puisque des mutants *plcA* ou *plcB* sont atténués. De plus, un double mutant délété des deux phospholipases (Smith *et al.*, 1995) est 500 fois moins virulent qu'un simple mutant, ce qui montre l'importance et la complémentarité de ces deux facteurs dans la listériose.

- PlcA

La phospholipase PlcA est une protéine sécrétée de 33 kDa qui agit en synergie avec la LLO et PlcB pour permettre l'échappement de la bactérie de la vacuole phagocytaire. Ceci a

été montré en étudiant des simples ou des doubles mutants (Camilli *et al.*, 1993 ; Smith *et al.*, 1995 ; Marquis *et al.*, 1995). Cette protéine n'a aucun rôle dans la propagation de cellules en cellules, suggérant qu'elle n'est pas impliquée dans l'échappement de la vacuole à double membrane (Smith *et al.*, 1995). Des mutants *plcA* sont légèrement moins virulents chez la souris après une inoculation intraveineuse et ne peuvent plus coloniser le foie. Par contre, ces mutants sont toujours capables de coloniser la rate (Camilli *et al.*, 1993).

- PlcB

La phospholipase PlcB est une protéine de 29 kDa possédant une activité phosphatidylcholine phosphohydrolase. Elle a une activité optimale à pH 5.5-7.0, en accord avec son rôle pour l'échappement des phagosomes où le pH est acide. PlcB est codée par le gène *plcB* et est exprimée comme un précurseur inactif. Elle peut être activée par un clivage protéolytique impliquant une metalloprotéase zinc-dépendante, Mpl codée par le gène *mpl*, mais également par des protéases cellulaires.

*In vitro*, le précurseur de PlcB et son activateur Mpl sont sécrétés et faiblement associés à la paroi bactérienne de manière non-covalente (Snyder & Marquis, 2003). Durant la croissance intracellulaire, la translocation de PlcB est inefficace et la protéine reste à l'interface membrane/paroi. Lors de l'acidification des phagosomes, des lots de PlcB inactifs sont transloqués à travers la paroi bactérienne (Marquis *et al.*, 1997 ; Snyder & Marquis, 2003). L'activation de PlcB requiert le clivage d'un propeptide. Le clivage par Mpl de PlcB est corrélé avec sa translocation à travers la paroi. (Yeung *et al.*, 2005).

Les mutants déficients pour PlcB, suite à une délétion de *plcB* ou de *mpl*, sont moins virulents chez la souris après injection intraveineuse (Raveneau *et al.*, 1992). PlcB a également un rôle important durant une listériose avec des symptômes cérébraux chez la souris, puisque les mutants *plcB* sont fortement atténués dans le modèle d'infection intracrânien (Schluter *et al.*, 1998). PlcB est impliquée dans la lyse de la double membrane de la vacuole phagocytaire et donc dans la propagation de cellules en cellules (Vazquez-Boland *et al.*, 1992). Dans certaines cellules, comme HenLe 407 et Hela, mais pas dans les macrophages, PlcB lyse la membrane des phagosomes primaires qui contiennent *L. monocytogenes*, en l'absence de LLO (Grundling *et al.*, 2003). Ceci suggère qu'en fonction du type cellulaire, PlcB pourrait soit agir seul soit avec l'aide de LLO pour lyser la vacuole primaire.

### 10-2.2.3. ActA

Après la lyse de la vacuole, *L. monocytogenes* se retrouve dans le cytoplasme, où elle va induire la polymérisation des filaments d'actine. Ceci va entraîner le mouvement intracellulaire de la bactérie et permettre sa dissémination de cellules en cellules. Ce phénomène est dû à la protéine ActA. Le gène *actA* code une protéine de 639 résidus, qui est sécrétée et la protéine mature de 611 acides aminés est ancrée dans la membrane bactérienne grâce à son motif transmembranaire présent en C-terminal. Dans les cellules de mammifères, ActA est phosphorylée mais le rôle de cette phosphorylation n'a pas été déterminé (Brundage *et al.*, 1993).

ActA est distribuée de manière polaire à la surface bactérienne. Elle est nécessaire et suffisante pour permettre la polymérisation de l'actine. Son rôle a été montré par expression de la protéine dans l'espèce non pathogène *L. innocua*. Chez cette bactérie, ActA est capable d'induire l'assemblage d'actine et la propulsion de la bactérie dans des extraits de cytoplasme d'œufs de *Xenopus* (Kocks *et al.*, 1995).

La partie N-terminale d'ActA a un rôle primordial pour le mouvement. En effet, les bactéries exprimant un chimère composé de la partie N-terminale d'ActA et du fragment w de LacZ sont mobiles dans des extraits cellulaires (Lasa *et al.*, 1995). Cette région est similaire au niveau structural et fonctionnel à la famille protéique eucaryote WASP/N-WASP (Wiskott-Aldrich syndrome protein). Elle lie et active Arp2/3, un complexe protéique de l'hôte qui induit la polymérisation de l'actine et la génération de filaments d'actine, conduisant à la mobilité de *L. monocytogenes*. Cependant, l'activation d'Arp2/3 seul n'est pas suffisant pour permettre *in vitro* la mobilité due à l'actine (Loisel *et al.*, 1999). En plus de l'actine et de l'activation du complexe Arp2/3, le facteur de dépolymérisation de l'actine (ADF ou cofiline) et une protéine Cap (CapZ) sont également nécessaires à la mobilité (Figure 11). Le mouvement est encore plus efficace quand la profiline, l'actinine et VASP sont présents (Cossart & Bierne, 2001).

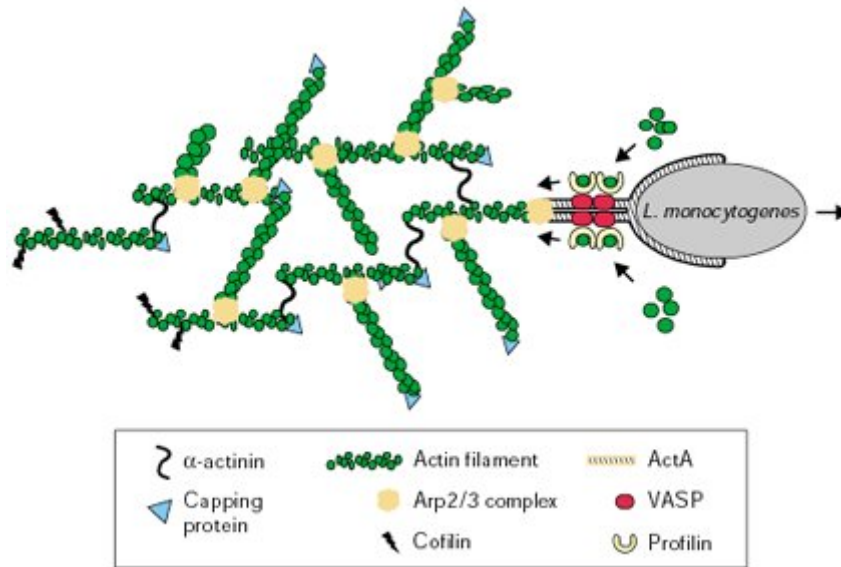


Figure 11 : Modèle de l'assemblage de l'actine, induit par ActA (Cossart & Bierne, 2001)

#### 10-2.2.4. HptA

Alors que la réplication dans la cellule est essentielle pour que la bactérie colonise son hôte, peu de choses sont connues sur les mécanismes mis en œuvre par *L. monocytogenes* pour obtenir des nutriments dans le cytosol de l'hôte. Cependant, il a été montré que *L. monocytogenes* exploitait les hexoses phosphates présents dans le cytoplasme de l'hôte comme source de carbone et d'énergie pour permettre une croissance intracellulaire rapide. La capture des hexoses phosphates est réalisée par Hpt, un homologue bactérien de la translocase mammifère qui transporte le glucose-6-phosphate du cytosol au réticulum endoplasmique dans la phase finale de la gluconéogenèse et de la glucogénolyse. La perte de Hpt entraîne une perte de la prolifération intracellulaire et une diminution de la virulence chez la souris (Chico-Calero *et al.*, 2002)

### 10-3. Régulation des facteurs de virulence par PrfA

#### 10-3.1. PrfA, régulateur majeur de la virulence

Les gènes de virulence les plus importants de *L. monocytogenes* (*prfA*, *plcA*, *hly*, *mpl*, *actA*, *plcB*) sont physiquement liés dans l'îlot de pathogénicité 1 (LIPI-1) de 9 kb situé sur le chromosome, contrôlé positivement par l'activateur transcriptionnel PrfA (Figure 12).



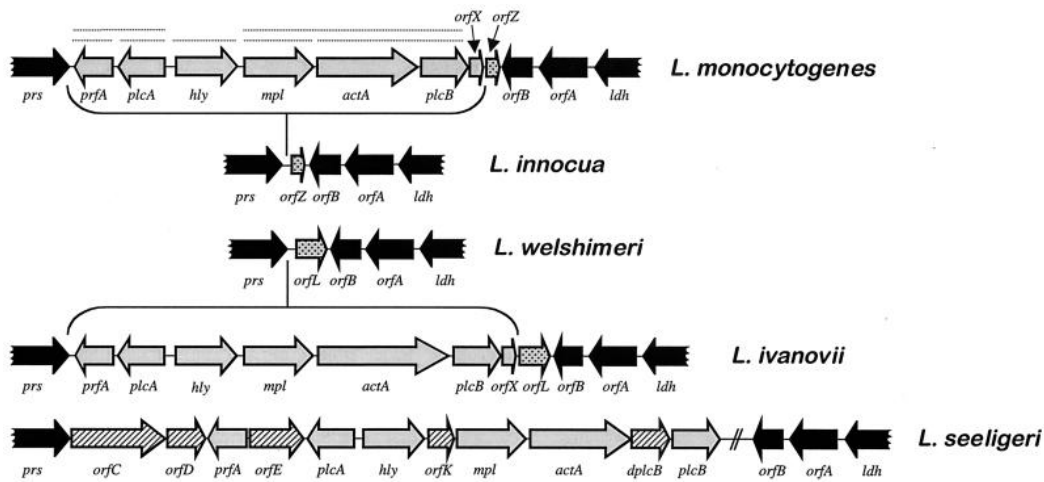


Figure 12 : Organisation de l'îlot de pathogénicité LIPI-1 chez les différentes espèces de *Listeria*

LIPI-1 est absent des espèces non pathogéniques à l'exception de *L. seeligeri*. Cependant, dans cette espèce, les gènes ne sont pas exprimés correctement à cause d'une insertion entre *prfA* et *plcA*. *InlA* et *inlB* sont regroupés sur un autre opéron, appelé *inlAB* et n'est que partiellement contrôlé par cet activateur. L'expression de certains gènes de virulence, tel que le locus *inlGHE*, est totalement indépendante de PrfA. PrfA régule également négativement certains gènes tels que *clpC* et les gènes associés à la motilité *motA* et *flaA*.

### 10-3.2. Contrôle environnemental de l'expression des gènes de virulence

Les facteurs de virulence sont essentiellement impliqués dans la colonisation des tissus animaux. Leur synthèse en dehors de l'hôte peut compromettre la capacité de *L. monocytogenes* à survivre dans l'environnement. De nombreuses études semblent montrer que *L. monocytogenes* a développé des mécanismes lui permettant d'exploiter les signaux délivrés par son environnement afin d'adapter l'expression des gènes de virulence pour avoir un comportement saprophyte ou parasite. La température est un de ces signaux environnementaux. La transcription des gènes régulés par PrfA est faible en dessous de 30°C, mais est induite à 37°C (température du corps des animaux à sang chaud). Cependant, une augmentation de la température n'est pas suffisante pour atteindre le niveau maximum de transcription de *prfA* et du régulon qui en dépend. Les gènes régulés par PrfA sont faiblement exprimés en milieu riche (BHI) à 37°C, mais leur transcription est fortement activée quand les bactéries sont cultivées en BHI supplémenté avec du charbon actif (Rouquette *et al.*, 1996). Le rôle du charbon n'est pas clairement défini. Une induction similaire à celle observée avec le charbon a lieu lorsque les bactéries cultivées en BHI sont transférées en milieu minimum

essentiel (MEM) (Bohne *et al.*, 1994). Ce milieu ne permet pas une croissance normale des bactéries, il n'est donc pas clair si l'activation des gènes de virulence est due à un mécanisme de senseur impliquant des signaux chimiques, à une réponse à un stress nutritionnel, ou bien aux deux.

Il a été démontré que *L. monocytogenes* répondait à de nombreux stress comme l'exposition à de hautes températures (42°C), à des oxydants ou à l'entrée en phase stationnaire, en augmentant l'expression de gènes dont *hly* grâce à l'induction de *prfA* (Sokolovic *et al.*, 1990 ; Sokolovic *et al.*, 1993).

### 10-3.3. Régulation des gènes de virulence par la source en carbone

Les carbohydrates fermentescibles causent une forte répression des gènes de virulence de *L. monocytogenes* (Milenbachs *et al.*, 1997 ; Behari & Youngman, 1998a). Cette régulation négative se passe seulement si les carbohydrates sont ajoutés en quantité suffisante pour permettre la croissance, suggérant que cette régulation est liée à la répression catabolique (RC). Cependant, un mutant délété de *ccpA*, un médiateur central de la RC, ne présente pas d'altération de la régulation des gènes de virulence par la source de carbone (Behari & Youngman, 1998b). La répression due aux sucres affecte PlcA, LLO et PlcB.

Contrairement au glucose et aux autres carbohydrates fermentescibles communs, le glucose-1-phosphate (et en général tous les hexoses phosphate) stimule efficacement la croissance de *L. monocytogenes* sans entraîner la répression des gènes de virulence (Chico-Calero *et al.*, 2002).

Les gènes PrfA-dépendants sont sélectivement activés dans le cytoplasme des cellules de l'hôte. Dans ce compartiment, le G1P et le G6P sont disponibles en tant que sources de carbone exogènes pour la croissance bactérienne (résultat du métabolisme du glycogène).

### 10-3.4. Mécanisme de régulation de PrfA

Le modèle de régulation le plus couramment proposé pour PrfA est décrit en figure 13. PrfA a deux états de fonctionnement, actif et inactif, et passe d'une forme à l'autre en se liant à un cofacteur de faible poids moléculaire qui transmet le signal environnemental à la machinerie transcriptionnelle *prfA* (Vazquez-Boland *et al.*, 2001). PrfA est toujours présent dans le cytoplasme bactérien, même en quantité minimale.

*prfA* est exprimé selon deux voies (Figure 13) :

- A) Sous le contrôle d'un promoteur dans la région intergénique *plcA-prfA*, qui donne un transcrite monocistronique à partir duquel PrfA est synthétisée en quantité faible mais détectable. Ceci se passe même si la bactérie est exposée à des conditions n'activant pas le régulon *prfA*.
- B) Sous le contrôle d'un promoteur dépendant de PrfA qui se situe en amont du gène *plcA*. Un transcrite *plcA-prfA* est généré, créant une boucle d'autorégulation, qui est critique pour l'activation adéquate de *prfA*.

Si *L. monocytogenes* ressent une combinaison de signaux environnementaux (ex : 37°C et environnement cytoplasmique), la concentration intracellulaire du cofacteur augmente, entraînant l'activation de la protéine PrfA (a), qui se lie avec une forte affinité aux promoteurs régulés par PrfA (PrfA-boxes) (b), qui entraîne la synthèse accrue de PrfA (dans la conformation active) par la boucle d'autorégulation positive réalisée par le transcrite bicistronique PrfA dépendant *plcA-prfA*, (c) puis qui induit la transcription de tous les gènes PrfA dépendant.

Ceci se déroule aussi longtemps que le cofacteur est présent dans le cytoplasme bactérien, et s'arrête rapidement si le signal d'activation environnemental cesse (par conséquent si le cofacteur diminue).

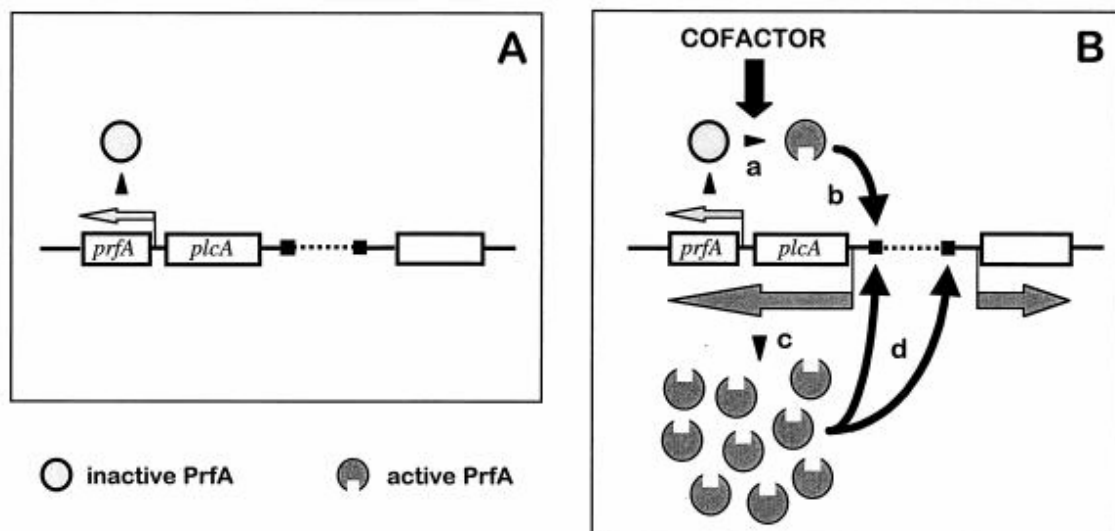


Figure 13 : Modèle du mécanisme d'autorégulation induit par PrfA (Vega *et al.*, 1998)

## 11 - LES TESTS DE VIRULENCE

Une des premières questions qui se pose lorsque l'on fait le parallèle entre la faible incidence des cas de listériose et la forte contamination des aliments est : toutes les souches de *L. monocytogenes* ont-elles le même niveau de virulence ? Afin d'apporter des éléments de

réponse à cette question, de nombreuses méthodes ont été utilisées pour évaluer la virulence des souches.

### **11-1. Tests in vivo**

Comme nous l'avons vu précédemment, l'infection est un processus comportant de nombreuses étapes et de nombreuses variables autant au niveau de l'hôte que de l'agent pathogène. Pour ces raisons, la compréhension de la physiopathologie d'une infection humaine nécessite un modèle animal dans lequel l'agent infectieux a les mêmes effets directs et entraîne des dommages immunopathologiques identiques à ceux produit chez l'homme. Comme la listériose animale existe, les espèces animales naturellement infectées par *L. monocytogenes* peuvent être utilisées comme modèles pour étudier la physiopathologie de la listériose humaine. Cependant, ceci pose problème car les animaux développant une listériose proche de celle des hommes sont des animaux de fermes tels que les moutons ou les vaches, et non pas des animaux de laboratoire tels que des souris ou des rats. Les souris sont généralement de bons modèles de laboratoire car elles sont petites et résistantes, elles s'adaptent bien à la vie en captivité, se reproduisent rapidement, ont des périodes de gestation courtes et une physiologie et une pathophysiologie comparables aux humains. Cependant, certains pathogènes de l'homme n'entraînent pas de maladie chez la souris ou entraînent des symptômes qui peuvent être différents. La souris peut être considérée comme un modèle non satisfaisant pour l'étude de la listériose en raison de la différence au niveau du récepteur d'InlA (la E-cadhérine) qui rend inefficace le passage à travers la barrière intestinale. Aucun cas naturel de listériose chez la souris n'a été rapporté dans la littérature et les essais expérimentaux d'infection par voie orale ont montré que cet animal était difficilement infecté par *L. monocytogenes*. La plupart des études utilisant des souris ou des rats ont montré qu'il était difficile d'induire une infection létale reproductible suite à une infection orale.

Pour contourner ce problème, d'autres routes d'infection ont été étudiées bien qu'elles ne correspondent pas à la route naturelle d'infection. Les voies d'infections les plus utilisées sont les voies intraveineuses et intrapéritonéales. L'infection de la souris par voie intraveineuse a permis de produire une infection létale et de produire un bon modèle pour étudier les facteurs de virulence impliqués dans la vie intracellulaire de *L. monocytogenes*.

Chez le cobaye et la souris transgénique exprimant la E-cadhérine humaine, les bactéries peuvent traverser efficacement la barrière épithéliale de l'intestin (Lecuit *et al.*, 2001). Ces 2 modèles semblent donc plus appropriés que le modèle murin classique pour l'étude de la listériose transmise par la voie orale.

Certaines études ont utilisé un modèle d'œufs embryonnés inoculés par voie chorioallantoïque pour caractériser la pathogénicité de *L. monocytogenes* (Terplan & Steinmeyer, 1989 ; Notermans *et al.*, 1998 ; Olier *et al.*, 2002). Ce modèle a l'avantage d'être sensible et de prendre en compte le rôle d'InlA car la E-cadhérine est identique à la E-cadhérine humaine. Cependant, cette méthode est lourde à mettre en œuvre et ne permet pas de tests d'infection des organes comme le permet le modèle murin.

## **11-2. Tests in vitro**

Les tests sur cultures cellulaires ont été développés comme alternative moins coûteuse que les tests sur souris. Ces méthodes mesurent la capacité de *L. monocytogenes* à causer des effets cytopathogènes sur les lignées cellulaires entérocytiques Caco-2 (Pine *et al.*, 1991) ou à former des plaques sur les lignées cellulaires humaines d'adénocarcinome HT-29 (Roche *et al.*, 2005). De nombreuses autres lignées cellulaires (cellules hépatocytaires Hep-G2, de macrophage J774, épithéliales Henle 407 et L2) ont également été utilisées pour tester la capacité de *L. monocytogenes* à adhérer, envahir, s'échapper des vacuoles, se multiplier à l'intérieur des cellules et se propager aux cellules voisines.

L'avantage principal de ces tests *in vitro* est leur faible coût et leur facilité d'utilisation. Cependant, ces tests prennent beaucoup de temps et connaissent des problèmes de reproductibilité, freinant leur utilisation en laboratoire pour déterminer la virulence et le potentiel pathogène des souches.

Il est également possible d'étudier les activités enzymatiques impliquées dans la virulence, c'est-à-dire de rechercher les activités hémolytiques et phospholipasiques (PI-PLC et PC-PLC). Sur une gélose contenant 5% de sang de mouton, de cheval, de lapin ou d'homme, les colonies de *L. ivanovii*, *L. monocytogenes* et *L. seeligeri* sont bêta-hémolytiques. L'hémolyse est généralement importante avec *L. ivanovii* (le rayon de la zone d'hémolyse peut atteindre 1 cm après 3 j d'incubation) et les souches peuvent s'entourer de plusieurs zones d'hémolyse. En revanche, l'hémolyse est moins importante pour *L. monocytogenes* ou *L. seeligeri* et elle n'est parfois visible que sous la colonie. L'activité PC-PLC est caractérisée par le développement de zones opaques dues à la dégradation de la lécithine de jaune d'oeuf présent dans le milieu de culture. L'activité PI-PLC peut être testée en recouvrant les colonies bactériennes développées sur une gélose nutritive avec un substrat appelé L-ft-phosphatidylinositol. Les colonies possédant une activité PI-PLC présentent un halo trouble autour de la colonie, dû au dégagement de diacylglycerol insoluble issu du substrat.

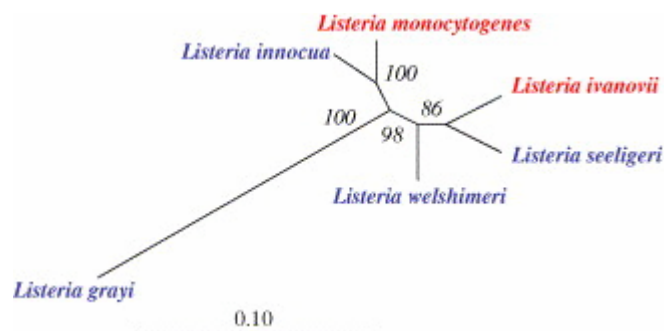
## - CHAPITRE 2 -

# BIODIVERSITE DES *LISTERIA*

## *MONOCYTOGENES*

Le genre *Listeria* contient 2 espèces pathogènes (*L. monocytogenes* et *L. ivanovii*) et 4 espèces non pathogènes (chapitre 1 § 2 -). La pathogénicité de ces deux espèces est notamment possible grâce à un cluster de gènes de 9 kb appelé LIPI-1 (chapitre 1 § 10-3), qui est également présent chez *L. seeligeri* mais est non fonctionnel à cause d'une insertion entre *prfA* et *plcA*.

Des analyses phylogéniques basées sur les gènes codant les ARNr 16S et 23S, et sur les gènes *prs*, *ldh*, *vclA*, *vclB* (gènes flanquant le cluster de gènes de virulence) et *iap*, indiquent que 5 des 6 espèces de *Listeria* sont divisées en 2 groupes, *L. monocytogenes* et *L. innocua* formant un groupe tandis que le second groupe inclus *L. welshimeri*, *L. ivanovii*, et *L. seeligeri*. *L. grayi* forme la branche la plus éloignée de ce genre (Figure 14) (Schmid *et al.*, 2005).



**Figure 14 :** Arbre phylogénique du genre *Listeria*, basé sur les ARNr 16S et 23S, *iap*, *prs*, *vclB* et *ldh*. Les souches pathogènes sont indiquées en rouge et les non pathogènes en bleu

Aucune trace de transfert horizontal de gène n'étant détectable, le scénario le plus probable est que le cluster de gènes de virulence était présent chez l'ancêtre des différentes espèces actuelles de *Listeria*. Ceci laisse supposer qu'en conséquence de la délétion de ce cluster de

gènes de virulence, l'ancêtre *Listeria* a cessé d'être seulement un pathogène intracellulaire et a adopté une vie extracellulaire en acquérant de nouveaux gènes permettant la diversité des *Listeria* et la colonisation de nouvelles niches écologiques.

Au sein de l'espèce pathogène pour l'homme *L. monocytogenes*, il semble également exister une grande diversité. Alors que *L. monocytogenes* est divisée en 13 sérovars, les sérovars 1/2a, 1/2b et 4b sont à l'origine de 96% des cas de listériose humaine. La prédominance de certaines souches à l'origine de la listériose reste encore un mystère, la faible fréquence des autres sérovars ne semblant pas due à une absence des gènes de virulence. Il semble donc évident qu'il existe une corrélation entre la composition antigénique et le pouvoir pathogène de la souche.

L'ingestion de *L. monocytogenes* est un événement très commun, étant donné la distribution ubiquiste de cette bactérie dans les aliments et les industries agroalimentaires. Cependant, l'incidence des cas de listériose humaine est très faible. Cette bactérie semble donc avoir un potentiel pathogène plus faible que d'autres pathogènes transmis par les aliments. Cette observation est en corrélation avec une dose létale 50% élevée (Audurier *et al.*, 1980 ; Okamoto *et al.*, 1994). Une hétérogénéité de la virulence a été observée entre de nombreuses souches de *L. monocytogenes* sans que celle-ci puisse être clairement corrélée à l'origine de la souche ou à des caractéristiques particulières (Van Langendonck *et al.*, 1998).

A ce jour, malgré une bonne compréhension des mécanismes du parasitisme de *L. monocytogenes*, de nombreuses questions restent sans réponse. Pourquoi certaines souches sont plus virulentes que d'autres ? Pourquoi seuls certains sérovars sont à l'origine des cas de listériose ? Le but de ce 2<sup>ème</sup> chapitre de la partie bibliographique est de dresser un bilan des connaissances sur la biodiversité des souches de *L. monocytogenes*, concernant leur répartition dans l'environnement, leur patrimoine génétique, mais également concernant la biodiversité au niveau du protéome et de la virulence.

## **1 - BIODIVERSITE ET ENVIRONNEMENT**

### **1-1. Répartition des souches en fonction de leur sous-type**

#### **1-1.1. Sérotypes**

Comme nous l'avons déjà vu précédemment, 96% des cas de listériose humaine sont dus à seulement 3 des 13 sérovars de *L. monocytogenes*. En France entre 2001 et 2003,

49% des souches à l'origine d'une listériose étaient de sérovar 4b, 27% de sérovar 1/2a et 20% de sérovar 1/2b (source InVS).

La rareté des cas de listériose dus aux autres sérovats peut s'expliquer par leur faible présence dans l'environnement des ateliers agroalimentaires et sur les aliments.

Curieusement, alors que le sérovar 4b est prépondérant dans les cas de listériose, il est beaucoup moins fréquemment isolé que les autres sérovats dans les aliments ou dans l'environnement des ateliers de production et de transformation des produits alimentaires. Par exemple, des prélèvements réalisés dans des usines de fabrication de saucissons secs ont montré que la majorité des isolats de *L. monocytogenes* étaient de sérovar 1/2 (49.5% de 1/2a, 13% de 1/2b et 19% de 1/2c) alors que seulement 8% des isolats étaient de sérovar 4b (Thevenot *et al.*, 2005). Le Tableau 6 dresse un bilan de différentes études sur la répartition des sérovats trouvée dans l'environnement des industries agro-alimentaires.

Ces observations suggèrent que les souches de sérovar 1/2a et 1/2b sont mieux adaptées pour survivre dans l'environnement agro-alimentaire et les aliments, entraînant une ingestion fréquente qui expliquerait la part importante de ces sérovats dans les cas de listériose. Les souches de sérovar 4b étant trouvées de façon minoritaire dans les aliments, et de façon majoritaire dans les cas de listériose, les souches de sérovar 4b semblent avoir un potentiel épidémique plus élevé que les autres souches.

**Tableau 6 : Distribution des souches en fonction de leur sérovar dans différents environnements alimentaires**

Sérovar 1/2a	Sérovar 1/2b	Sérovar 1/2c	Sérovar 4b	Isolement	Référence
45%	31%	0%	17%	Réservoirs de lait (laiteries USA)	(Van Kessel <i>et al.</i> , 2004)
49,5%	13%	19,5%	8%	Saucissons /environnement industrie	(Thevenot <i>et al.</i> , 2005)
34%	1,5%	63%	1,5%	Viande	(Nicolas <i>et al.</i> , 1989)
44%	5%	28%	23%	Viande	(Fantelli & Stephan, 2001)
54%	12%	12%	19,5%	Aliments divers	(Gilot <i>et al.</i> , 1996)
33,5%	58,5%	5%	3%	Fromages	(Loncarevic <i>et al.</i> , 1995)
76%				Fromages/viandes	(Schonberg <i>et al.</i> , 1989)

### 1-1.2. Autres répartitions

De nombreuses méthodes de typage moléculaire (ribotypage, PFGE, séquençage de gènes de virulence...) ont montré que les souches de *L. monocytogenes* se répartissaient en deux divisions majeures (Brosch *et al.*, 1994), (Zhang *et al.*, 2003), (Graves *et al.*, 1994), (Piffaretti *et al.*, 1989), (Ripabelli *et al.*, 2000). Les souches de sérovar 1/2b, 3b, 4b, 4d, 4e, 4ab et 4a sont groupées dans un même lignage que nous appellerons lignage I. Les souches de sérovar 1/2a, 1/2c 3c et 3a appartiennent à un autre lignage, que nous appellerons lignage II.



D'autres études ont montré l'existence d'un troisième linéage (Jeffers *et al.*, 2001) (Jacquet *et al.*, 1993 ; Wiedmann *et al.*, 1997). Ces études suggèrent que le linéage I est associé de manière spécifique aux cas de listériose humaine, puisque la majorité est due aux souches de sérovar 4b, tandis que les souches du linéage II sont plus fréquemment trouvées dans les aliments. De plus, il a été montré que les souches du linéage I étaient plus fréquemment à l'origine des cas de listériose humaine que des cas de listériose animale, à l'inverse des souches de linéage III (Gray *et al.*, 2004 ; Jeffers *et al.*, 2001).

## **1-2. Biofilms**

*L. monocytogenes* est capable de former des biofilms, ce qui permet à ce pathogène de persister dans les ateliers de production d'aliments. Un biofilm consiste en des agrégats ou des tapis mono- ou pluri-couches de cellules immobilisées sur une surface biotique ou abiotique, fréquemment englobées dans une matrice de polymères organiques d'origine microbienne. Plusieurs travaux ont été réalisés afin de comparer la capacité de différentes souches de *L. monocytogenes* à adhérer à des surfaces abiotiques et à former des biofilms. Lunden (2000), Borucki (2003) et Norwood (1999) ont montré que les souches qui avaient préalablement persisté dans l'environnement d'industries agroalimentaires avaient une meilleure capacité à former des biofilms que des souches non persistantes. Certains auteurs ont montré une relation entre le sérotype des souches et leur capacité à former des biofilms. Norwood et Gilmour (1999) ont suggéré que les souches de sérovar 1/2c adhéraient mieux que les souches de sérovar 1/2a et 4b et Borucki *et al.* (2003) ont montré une meilleure capacité à former des biofilms pour les souches appartenant au linéage I (1/2a et 1/2c). Au contraire, Djordjevic *et al.* (2002) ont montré que les souches appartenant au linéage II (sérovars 1/2b et 4b) formaient plus facilement des biofilms sur des microplaques de PVC, et a émis l'hypothèse que la capacité plus élevée de ces souches à former des biofilms pourrait expliquer leur dominance dans les cas de listériose humaine. Cependant, d'autres études ont montré qu'il n'y avait aucune relation entre la capacité à former des biofilms et (i) le fait qu'une souche soit persistante ou non dans l'environnement, (ii) l'origine de la souche (aliments ou souche clinique), et (iii) le sérovar ou le linéage de la souche (Kalmokoff *et al.*, 2001). Les divergences dans les résultats observés par ces auteurs pourraient s'expliquer par l'utilisation de conditions expérimentales différentes.

Alors que des travaux chez d'autres pathogènes comme *Salmonella* ou *Enterococcus* ont suggéré que la croissance en biofilm pouvait augmenter la virulence, aucune relation de cette nature n'a été mise en évidence à ce jour chez *L. monocytogenes*.

## 2 - BIODIVERSITE ET GENOMES

### 2-1. Génomes séquencés

En 2001, le premier génome d'une souche de *L. monocytogenes* a été séquencé par Glaser *et al* (2001). Une souche de *L. innocua* a également été séquencée par ce groupe, afin de comparer les génomes d'une espèce pathogène et d'une espèce non pathogène de *Listeria*. Les souches sélectionnées étaient *L. monocytogenes* EGDe (sérovar 1/2a) et *L. innocua* CLIP 11262 (sérovar 6a). En 2004, les génomes de 3 autres souches de *L. monocytogenes* ont été publiés (Nelson *et al.*, 2004). Ces souches séquencées étaient à l'origine de cas cliniques de listériose aux Etats-Unis dont la source de contamination avait été identifiée. Il s'agit de *L. monocytogenes* F2365 (sérovar 4b, isolée de fromage), F6854 (sérovar 1/2a, isolée de saucisse de francfort) et H7858 (sérovar 4b, isolée de viande). Les principales caractéristiques de ces 5 souches sont décrites dans le Tableau 7.

**Tableau 7 : Principales caractéristiques des souches de *Listeria* séquencées (Glaser *et al.*, 2001 ; Nelson *et al.*, 2004)**

	<i>L. monocytogenes</i>				<i>L. innocua</i>
	EGDe (1/2a)	F6854 (1/2a)	F2365 (4b)	H7858 (4b)	CLIP 11262 (6a)
Taille du chromosome (pb)	2.944.528	2.953.211	2.905.310	2.893.921	3.011.209
G+C (%)	39%	37,8%	38%	38%	37%
Nombre de gènes codant des protéines	2853	2973	2847	3024	2973
Plasmides	0	0	0	1 (94 gènes, 82 kpb)	1 (79 gènes, 82 kpb)
Nombre de gènes spécifiques à la souche	61	97	51	69	/

La comparaison du génome de *L. monocytogenes* EGDe et *L. innocua* CLIP11262 (Glaser *et al.*, 2001) a montré une organisation conservée du génome et un nombre important de gènes orthologues. Cette comparaison a également permis de mettre en évidence 270 gènes (9.5%) spécifiques à *L. monocytogenes* EGDe et 149 gènes (5%) spécifiques à *L. innocua* CLIP11262 sans prendre en compte le plasmide présent chez cette dernière. Les gènes spécifiques à *L. monocytogenes* sont distribués dans 100 fragments ADN répartis sur tout le génome, et les gènes spécifiques à *L. innocua* CLIP11262 sont clustérisés en 63 régions contenant de 1 à 7 gènes (Glaser *et al.*, 2001). Cette organisation particulière de petites régions dans les génomes de *Listeria* suggère que de nombreux événements d'acquisition / délétion ont conduits aux génomes actuels.

La comparaison des 5 génomes séquencés de *Listeria* (Nelson *et al.*, 2004) au niveau nucléaire et des protéines prédites a permis de mettre en évidence des différences qui

pourraient être liées aux caractéristiques de survie, de croissance et de pathogénicité des souches et des sérotypes. 83 gènes sont spécifiques aux souches de sérovar 1/2a et 51 gènes sont spécifiques aux souches de sérovar 4b. 44% des gènes spécifiques au sérovar 1/2a et 65% des gènes spécifiques au sérovar 4b sont des protéines hypothétiques pour lesquelles il n'y a aucune information sur la nature exacte et la fonction.

Parmi les gènes spécifiques au sérovar 1/2a, il existe des gènes qui codent des voies de transport et du métabolisme de carbohydrates dont le ribose et un pentose non identifié. De même, les souches de sérovar 1/2a possèdent un opéron qui code les voies de biosynthèse du rhamnose, (antigène décorant) déterminant antigénique lié aux polymères d'acides téichoïques de la paroi bactérienne chez les souches de sérovar 1/2a.

Au total, 51, 97, 69 et 61 gènes spécifiques ont été identifiés chez les souches F2365, F6854, H7858 et EGDe, respectivement. La plupart de ces gènes spécifiques ont une composition atypique, suggérant que certains ont été acquis par transfert.

Cette comparaison de différents génomes a permis d'identifier des séquences protéiques non connues auparavant qui ont des caractéristiques de facteurs de virulence. C'est notamment le cas des protéines possédant un motif LPXTG de liaison aux peptidoglycanes de la paroi cellulaire, des protéines liant la choline, des lipoprotéines et des internalines. Ces protéines existent en nombre variable au sein des 4 souches séquencées de *L. monocytogenes*. De même, il a été identifié des gènes impliqués dans la biosynthèse de la paroi et des acides téichoïques qui sont certainement à l'origine des différences d'antigènes somatiques et impliqués dans la virulence et l'immunogénicité des *Listeria*.

## **2-2. Comparaison de souches par hybridation d'ADN**

La comparaison de génomes séquencés de *Listeria* a permis de mettre en évidence un certain nombre de différences entre l'espèce pathogène *L. monocytogenes* et l'espèce non pathogène *L. innocua* (Glaser *et al.*, 2001), mais également des différences entre les souches ou les différents sérovares de *L. monocytogenes* (Nelson *et al.*, 2004).

Afin d'élargir cette comparaison des souches séquencées à une large collection de souches, des études utilisant la technique d'hybridation de l'ADN ont été mises en oeuvre. Doumith *et al.* (2004b) ont utilisé une membrane contenant 409 sondes, dont 262 étaient spécifiques de *L. monocytogenes* EGDe par rapport à *L. innocua* CLIP 11262, 94 étaient spécifiques de *L. innocua* CLIP 11262 par rapport à *L. monocytogenes* EGDe, et 53 étaient spécifiques de *L. monocytogenes* CLIP80459, une souche de sérovar 4b.

L'analyse de 93 souches de *L. monocytogenes* grâce à cette puce ADN a permis de définir 3 lignées identiques à ceux trouvés par les autres méthodes moléculaires (chapitre 2 § 2-2). La distribution des gènes codant des protéines de surface était très hétérogène. Chaque lignée de *L. monocytogenes* était caractérisée par la présence de gènes codant des protéines de surface bien spécifiques et seul un nombre infime de protéines de surface étaient présentes chez toutes les souches étudiées. Les différences entre les gènes codant des systèmes phosphotransférase (PTS) ou des protéines de régulation étaient moins prononcées. Cette analyse a permis d'identifier de nombreux gènes spécifiques aux lignées et aux sérovars. 19 gènes étaient spécifiquement associés au lignée I (sérovar 1/2a, 3a, 1/2c, 3c) et 5 gènes étaient marqueurs du lignée II (sérovars 4b, 4d, 4e, 1/2b, 3b). Puisque les souches de sérovar 4b sont les principales responsables des cas de listériose humaine, il était également important d'identifier des marqueurs de ce sérovar. Trois gènes étaient spécifiquement présents dans les 27 souches de sérovar 4b étudiées. Ces gènes codaient une protéine sécrétée (ORF2110), une protéine précurseur C de l'acide téichoïque (ORF2372) et une protéine de fonction inconnue (ORF0799). Aucun gène spécifique du lignée III (sérovars 4a et 4c) n'a été caractérisé puisqu'aucun gène représentatif de ce lignée n'était déposé sur la membrane. Par contre, 37% des gènes de *L. monocytogenes* EGDe étaient absents des souches appartenant au lignée III. Il est à noter que ces gènes étaient également absents des souches de *L. innocua*.

A partir de ces résultats, 4 gènes marqueurs des sérovars ont été sélectionnés afin de mettre au point une PCR multiplex permettant la différenciation des principaux sérovars de *L. monocytogenes* (Doumith *et al.*, 2004a).

Une autre approche utilisant une puce à ADN Zhang *et al.* (2003) ont montré des résultats similaires en permettant d'identifier des gènes spécifiques aux sérovars. L'étude de 113 souches de *L. monocytogenes*, basée sur la séquence des gènes appartenant au cluster contrôlé par PrfA, a également confirmé l'existence de trois lignées (Ward *et al.*, 2004).

Ces approches génomiques récentes permettent d'approfondir les connaissances sur la distribution et de la diversité du contenu génétique chez *L. monocytogenes*.

### 3 - BIODIVERSITE ET PROTEOMES

La génomique a pour objectif de dresser l'inventaire exhaustif des gènes. Après avoir déchiffré le génome, l'un des enjeux les plus importants est la détermination de la nature et de la fonction de chaque produit de gène. En effet, l'identification de tous les gènes représente la connaissance de la partie émergée d'un iceberg, indispensable mais insuffisante pour

appréhender la partie immergée constituée des transcrits et des protéines. Alors que le génome est constant dans toutes les cellules d'un organisme, le transcriptome et le protéome peuvent varier considérablement dans le temps en fonction des conditions environnementales et de l'état physiologique. De plus, un gène peut donner plus d'un transcrit et chacun de ceux-ci peut coder une protéine susceptible de subir des modifications post-traductionnelles résultant en des fonctionnalités différentes. Il est donc indispensable de s'intéresser non seulement au génome mais également au transcriptome et au protéome d'un organisme. La protéomique, focalisée sur le produit final de l'expression des gènes, c'est-à-dire la protéine, permet d'avoir une vision globale des molécules fonctionnelles dans une cellule.

Concernant *Listeria*, plusieurs études protéomiques ont été réalisées afin d'appréhender, par exemple, la réponse de la bactérie à différents stress (Phan-Thanh & Gormon, 1995 ; Phan-Thanh & Gormon, 1997 ; Hébraud & Guzzo, 2000 ; Duche *et al.*, 2002b ; Helloin *et al.*, 2003), la formation de biofilm (Tremoulet *et al.*, 2002), la modifications au cours des différents stades de croissance (Folio *et al.*, 2004). Par contre, peu d'études ont abordé la biodiversité des souches par une approche protéomique. Dans ces quelques travaux, comme au niveau génomique, des différences ont pu être observées entre les espèces de *Listeria* mais également entre les souches et les différents sérovars de *L. monocytogenes*.

Les protéomes de *L. monocytogenes* et de l'espèce non pathogène *L. innocua* ont été comparés par EBD ou par 2-D LC-MS/MS. La comparaison des protéines cytosolubles montrait une forte similarité. La diversité entre les espèces était plus importante au niveau des protéines sécrétées dans le milieu extracellulaire ou des protéines de surface. Les deux espèces sécrètent une quantité et un nombre de protéines équivalent (Trost *et al.*, 2005). Cependant, parmi la centaine de protéines identifiée pour chaque espèce, seule la moitié était présente chez les deux espèces. Seize protéines présentes dans le surnageant de *L. monocytogenes* n'avaient pas d'orthologue chez *L. innocua*. Ces protéines ont un intérêt tout particulier puisque tous les gènes de virulence connus à ce jour appartiennent à ce groupe et les autres protéines ont des fonctions putatives dans la pathogénicité (InlH), dans l'adhésion (Lmo1666) ou dans l'infection (Lmo1076). Un système à deux composants a également été trouvé. Des différences reflétant le mode de vie spécifique de ces deux espèces ont également été mises en évidence. Par exemple, *L. innocua* exprimait plus de systèmes PTS que *L. monocytogenes* et les protéines du flagelle étaient exprimées à 37°C alors qu'elles ne le sont qu'à une température plus faible chez *L. monocytogenes*. Par ailleurs, *L. monocytogenes* exprimait plus de protéines impliquées dans l'adaptation à des conditions atypiques.

Au niveau de la paroi (Calvo *et al.*, 2005), 19 protéines et 11 protéines ont été identifiées par 2-D LC-MS/MS chez *L. monocytogenes* et *L. innocua*, respectivement. La plupart de ces protéines possédaient un motif LPXTG de liaison covalente à la paroi cellulaire.

Des études ont également été réalisées afin d'étudier la biodiversité des protéomes des souches de *L. monocytogenes*.

Deux études effectuées sur un nombre important de souches de *L. monocytogenes* par SDS-PAGE (Tabouret *et al.*, 1992) ou par EBD (Gormon & Phan-Thanh, 1995) ont montré que leurs protéomes étaient très différents de ceux des autres espèces de *Listeria* (*L. innocua*, *L. welshimeri*, *L. seeligeri* et *L. ivanovii*). Dans ces deux études, il a été montré une forte homogénéité de profil protéique pour des souches de *L. monocytogenes* de même sérovar. Les profils protéiques de souches de sérovars différents présentaient des différences plus importantes. Cependant, alors que Gormon & Phan-Than (1995) ont montré que les souches se répartissaient en deux groupes, le groupe I comprenant notamment les souches de sérovar 4b et 1/2b, et le groupe II les souches de sérovar 1/2a, Tabouret *et al.* ont, quant à eux, montré que les souches de sérovar 4b se regroupaient indépendamment des souches de sérovar 1/2a et 1/2b. Les protéines différenciant les différents clusters n'ont pas été identifiées.

Ramnath *et al.* (2003) ont également comparé par EBD la souche séquencée EGDe (sérovar 1/2a, isolée d'une listériose animale) avec 3 autres souches de sérovar 1/2a et une souche de sérovar 1/2b, isolées des aliments. En moyenne, 13% des spots protéiques majeurs et 28% des spots protéiques mineurs des souches isolées des aliments n'étaient pas retrouvés chez la souche EGDe, montrant une importante variation au niveau protéique selon l'origine des souches.

## 4 - BIODIVERSITE ET VIRULENCE

### 4-1. Différences de niveau de virulence entre les souches

Afin d'étudier la virulence de *L. monocytogenes*, des infections expérimentales d'animaux de laboratoire (inoculation intraveineuse ou intrapéritonéale de souris) et l'invasion de cellules de mammifère cultivées *in vitro* ont été utilisées (chapitre 1 § 11 -). Par ces techniques, la majorité des cultures de *L. monocytogenes* ont été trouvées comme virulentes, c'est-à-dire capables de tuer les souris ou de survivre dans l'estomac ou le foie, ou bien capables d'envahir des cellules de mammifère *in vitro*. Brosch *et al.* (1993) et Conner *et al.* (1989) ont réalisé l'analyse d'une collection importante de souches (125 et 218 souches respectivement) en inoculant des souris par voie intrapéritonéale (Conner *et al.*, 1989) ou

intraveineuse (Brosch *et al.*, 1993). Chacune de ces études a évalué une sélection de souches d'origines diverses (cliniques ou issues d'aliments) et appartenant à différents sérovars. Brosch *et al.* (1993) ont également établi des différences entre les souches étudiées en utilisant des techniques de sous-typage : lysotypage, ribotypage, PFGE. Les deux études montraient que la majorité des souches testées (> 93%) était capable de tuer les souris ou de persister dans l'estomac ou le foie. Par contre, aucune corrélation entre l'origine ou le sérovar et la capacité de la souche à causer la maladie n'a pu être établie. Des variations au niveau de la persistance ou du taux de mortalité ont été observées mais il n'y avait pas non plus de corrélation entre l'origine, le sérotype, le phagovar, le ribotype, ou le profil PFGE. Par exemple, parmi les souches testées par Conner *et al.* (1989), des souches avirulentes étaient trouvées aussi bien dans des sérovars rarement (3a et 3b) que fréquemment (4b) associés à des cas de listérioses humaines. De plus, Brosch *et al.* (1993) n'ont pas trouvé de différences de virulence entre les souches isolées de cas épidémiques et celles isolées d'aliments.

## **4-2. Sérovar et virulence**

Des différences de virulence ont été observées entre les souches, mais il était difficile d'établir une discrimination claire entre le niveau de virulence et l'origine des souches. Par contre, il semble exister une évidente association entre la composition antigénique et la pathogénicité de la souche. En effet, comme nous l'avons déjà vu, les gènes de virulence connus à ce jour semblent présents chez les 13 sérovars de l'espèce *L. monocytogenes* (Doumith *et al.*, 2004b) mais seuls 3 sérovars (1/2a, 1/2b et 4b) sont principalement responsables des cas de listériose aussi bien humains (chapitre 1 § 7 -) qu'animaux. Les souches de sérovar 4b sont majoritaires aussi bien dans les cas foeto-maternels et néonataux (MN) que chez l'adulte. Bien que les souches de sérovar 4b soient prédominantes dans toutes les formes cliniques, il semble exister des différences en fonction des symptômes observés (Tableau 8). D'après les résultats collectés en France entre 2001 et 2003 ([www.invs.fr](http://www.invs.fr)), le sérovar 1/2b est plus souvent impliqué dans les formes MN que dans les formes non MN de la listériose [26.9 % versus 17.7 %] et le sérovar 1/2c est plus fréquemment impliqué dans les formes non MN que dans les formes MN de la listériose [4.5 % versus 0.7 %]. Au sein des souches à l'origine des formes non MN, le sérovar 4b est plus fréquemment retrouvé dans les infections du système nerveux central que dans les bactériémies [56 % versus 41 %]. Ce dernier résultat a déjà été observé pour les cas de listériose diagnostiqués en 1995.

**Tableau 8 : Répartition (en %) des principaux sérovars de 603 souches de *L. monocytogenes* à l'origine des cas humains diagnostiqués de 2001 à 2003, selon la forme clinique des cas (source : CNR des *listeria*)**

	1/2a	1/2b	1/2c	4b
Total (n=462)	29	18	5	48
Non MN				
Infection Système Nerveux Central (n=165)	27	14	3	56
Bactériémie (n=270)	31	21	6	41
MN (n=141)	21	27	< 1	52
Total( n=603)	27	20	4	49

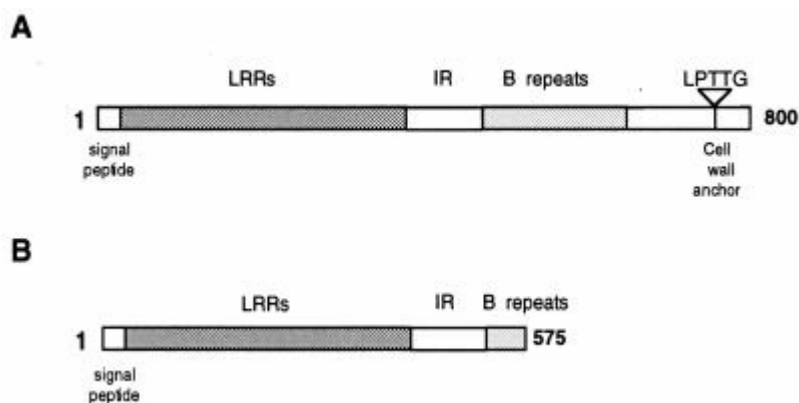
### 4-3. Gènes de virulence

Le cycle de virulence de *L. monocytogenes* est bien connu à ce jour (Vazquez-Boland *et al.*, 2001), les différentes étapes et les gènes impliqués sont décrits dans le chapitre 1 § 10-2.

L'étude de Doumith *et al* (2004b) par hybridation sur une puce à ADN d'une collection de 93 souches de *L. monocytogenes* a montré que toutes les souches, quel que soit leur sérovar, possédaient les gènes de virulence appartenant à LIPI-1 (*prfA*, *plcA*, *hly*, *mpl*, *actA*, et *plcB*), ainsi que les internalines *inlA* and *inlB*. La différence de virulence entre les souches ne semble donc pas due à une délétion d'un ou plusieurs de ces gènes impliqués dans la virulence.

Par contre, Il apparaît chez certaines souches des différences dans la séquence de certains gènes, entraînant la synthèse d'une protéine non fonctionnelle qui pourrait être à l'origine d'une virulence atténuée. Il a été observé que certaines souches produisaient une InlA tronquée. Ceci a été observé pour la première fois chez la souche LO28, dont la virulence est atténuée par rapport à celle d'EGDe. Par des expériences de Southern blot et PCR, le gène *inlA* semblait être identique entre ces deux souches (Jonquieres *et al.*, 1998). En fait, le séquençage du gène a permis de constater qu'il existait 4 mutations silencieuses, déjà trouvées lors d'une étude sur le polymorphisme de *inlA*, ainsi que 10 autres mutations et la délétion d'une adénine en position 1637. Cette délétion entraîne la création d'un codon non sens en position 1729 et d'une protéine tronquée de 63 kDa ne possédant pas le motif LPXTG d'ancrage à la paroi (Figure 15).





**Figure 15 : Organisation structurale d'InIA chez la souche EGDe (A) et LO28 (B).  
LRRs=leucine-rich repeats (Jonquieres *et al.*, 1998)**

De ce fait, la protéine n'était retrouvée que dans le milieu extracellulaire et n'était pas présente à la surface de la bactérie. De même, Olier *et al* (2003) lors de l'étude de souches issues de portage asymptomatique, ont montré que celles présentant une adhésion et une entrée diminuées dans les cellules Caco-2 produisaient des formes tronquées d'InIA, avec des masses moléculaires apparentes de 47 kDa. Les souches à l'origine de cas de listériose produisaient une protéine fonctionnelle de 80 kDa. Jonquieres *et al* (1998) ont également étudié 14 souches cliniques et 12 souches isolées des aliments. Quatre de ces souches, dont une clinique, possédaient une InIA tronquée, de masse moléculaire de 40, 45, 47 et 75 kDa. Ces observations ont mené Jacquet *et al* (2004) à effectuer une étude épidémiologique afin (i) d'examiner l'expression et la taille de InIA dans un panel de souches isolées de patients ou issues de programmes de surveillance alimentaire et (ii) de chercher si le polymorphisme de InIA pouvait être corrélé avec les manifestations de la maladie. Il ressort que l'expression d'une InIA entière est fortement associée aux souches cliniques puisque 96% d'entre elles possédaient une protéine entière, alors que ce n'était le cas que pour 65% des souches d'origine alimentaire. L'observation de l'expression de InIA au sein des différents sérovars révélait que toutes les souches de sérovar 4b et 1/2b possédaient une protéine entière. Par contre, aucune souche de sérovar 1/2c n'exprimait une internaline A intacte. La rareté des souches de sérovar 1/2c dans les cas cliniques pourrait donc s'expliquer par la production d'une InIA tronquée, non fonctionnelle. 83% des souches de sérovar 1/2a possédaient une internaline A entière et 16% une protéine tronquée. Dans ce sérovar, l'expression d'une InIA intacte était plus souvent associée avec les souches cliniques (127 / 131 souches) qu'avec les souches isolées d'aliments (56 /89 souches).

Jacquet *et al.* (2002) ont également étudié le polymorphisme des facteurs de virulence ActA, Ami, InlB et listériolysine O (LLO). Toutes les souches exprimaient LLO, InlB et ActA. Ces résultats sont en accord avec ceux obtenus par Doumith *et al.* (2004b). Par contre, Ami était absent de 17 des 300 souches cliniques étudiées et de 1 des 150 souches issues des aliments. Toutes les souches où Ami n'a pas été détecté appartenaient au sérovar 4b. Aucun polymorphisme n'était observé pour LLO et InlB. Par contre, en se basant sur le poids moléculaire des protéines détectées, les souches pouvaient être divisées en deux groupes pour Ami (75% Ami1 et 21% Ami2) et en 4 groupes pour ActA (52% des souches ActA1, 18% ActA2, 30% ActA3 et 1 souche ActA4). Les souches isolées des aliments appartenaient plus fréquemment au groupe ActA3 que les souches humaines. Les souches isolées de cas non MN et à l'origine de bactériémie étaient de manière préférentielle associées au groupe Ami1. Les souches à l'origine d'infections du système nerveux central étaient quant à elles associées aux groupes ActA2 et ActA3.

Ces résultats montrent qu'il existe des différences claires dans la structure des protéines entre les souches de différents sérovars, entre des souches qui ont des origines différentes ou qui sont la cause de différentes formes cliniques de listériose. Cependant, ces différences ne peuvent constituer des marqueurs pour déterminer sans ambiguïté si une souche trouvée dans les aliments est virulente ou non.

Concernant le niveau d'expression de ces gènes, Werbrouck *et al.* (2006) ont montré que les taux d'ARNm de *inlA* et *inlB* étaient significativement plus faibles chez les souches cliniques que chez les souches non cliniques. Par contre, aucune différence n'était observée pour *ami*.

## 5 - CONCLUSION

Cette partie bibliographique montre donc qu'il existe une grande biodiversité au sein de l'espèce *L. monocytogenes*, aussi bien au niveau génomique, protéomique qu'au niveau de la distribution environnementale ou de la virulence. Cependant, cette biodiversité observée ne permet pas d'expliquer clairement pourquoi des souches sont plus virulentes que d'autres et la prédominance de certaines souches dans l'environnement des industries agroalimentaires ou dans les cas de listériose humaine. La prédominance des souches 1/2 dans l'environnement des industries agroalimentaires et des souches 4b dans les cas de listériose humaine pourrait refléter une spécificité de niche écologique dont les caractéristiques de colonisation ne sont

pas élucidées. La virulence, ainsi que la capacité à coloniser une niche spécifique semblent nécessiter des mécanismes complexes.

A l'heure actuelle, il faut considérer toutes les souches de *L. monocytogenes* comme étant pathogènes. En effet, bien que la grande majorité des cas de listériose soit dus aux sérovars 1/2a, 1/2b et 4b, quelques rares cas de listérioses humaines dues à des souches d'autres sérovars sont recensés chaque année. De plus, bien que différentes techniques de sous-typage soient disponibles pour subdiviser les souches de *L. monocytogenes*, elles ne permettent pas de préciser la capacité d'une souche à entraîner la maladie. Enfin, toutes les souches de *L. monocytogenes* possèdent les gènes de virulence connus à ce jour. Cependant, la présence seule de ces gènes n'est pas suffisante pour rendre compte de la virulence. Des mutations ponctuelles peuvent avoir des conséquences importantes pour la virulence. Par exemple, l'absence de virulence de certaines souches peut s'expliquer par la synthèse d'une InlA tronquée non fonctionnelle. Cependant cette caractéristique n'est pas spécifique de toutes les souches avirulentes et il n'existe à ce jour aucun caractère permettant de déterminer clairement la virulence d'une souche.

# - CHAPITRE 3 -

## ADAPTATION AUX STRESS ENVIRONNEMENTAUX

Lors de son cycle de virulence, *L. monocytogenes* est confrontée à de nombreux stress. Elle doit être capable de survivre dans les aliments qui peuvent être acides ou contenir une forte concentration de sel, et doit résister aux faibles températures rencontrées dans les industries agroalimentaires ou dans le réfrigérateur du consommateur. Une fois chez l'hôte, *L. monocytogenes* doit notamment survivre au stress acide rencontré lors du passage gastrique chez l'homme et lors de l'invasion intracellulaire. La capacité d'adaptation de ce pathogène à ces nombreux changements environnementaux est essentielle pour sa croissance et sa survie.

Dans ce chapitre, nous ferons dans un premier temps le point sur certains mécanismes mis en œuvre par *L. monocytogenes* pour résister aux différents stress rencontrés dans l'environnement et chez l'hôte. Nous nous attarderons ensuite sur la ferritine, protéine connue pour être impliquée dans l'adaptation à une carence en fer, et qui semble avoir un rôle plus important dans l'adaptation à divers stress.

### 1 - PRINCIPAUX STRESS ET MECANISMES DE SURVIE

#### 1-1. Survie au stress acide

*L. monocytogenes* est fréquemment confrontée à des stress acides : le faible pH de certains aliments, le passage gastrique chez l'homme et dans le phagosome lors de l'invasion cellulaire. Afin de survivre dans ces environnements à faible pH, ce pathogène utilise des mécanismes d'adaptation.

Phan-Thanh & Mahouin (1999) ont étudié l'expression protéique de *L. monocytogenes* exposées à des pH acides létaux (stress acide) et à des pH acide non létaux (adaptation acide). La protéine GroEL, une ATP synthase et de nombreux régulateurs transcriptionnels étaient

plus abondants lors de stress acide. Les bactéries soumises au préalable à une adaptation à l'acide (pH 5.2 pendant 2 h) ont une résistance accrue au choc chaud (52°C) et aux chocs osmotiques (25-30 % NaCl), ce qui suggère que l'adaptation acide permet également une protection croisée contre d'autres facteurs de stress (Phan-Thanh *et al.*, 2000).

De plus, *L. monocytogenes* possède un système appelé GAD (Glutamate decarboxylate system) qui lui permet de résister à de faibles pH (Cotter *et al.*, 2001). La bactérie possède deux transporteurs capables de transporter le glutamate dans la cellule et 3 enzymes GAD (GadA, GadB, GadC) qui entraînent la décarboxylation du glutamate. Cette réaction convertit le glutamate en GABA (gamma-amino-butyric acid) qui est transporté hors de la cellule en échange de glutamate. Cette réaction entraîne la consommation de protons dans le cytoplasme bactérien et une augmentation du pH dans le cytoplasme. Ce système ne fonctionnant qu'en présence de glutamate dans le milieu extérieur, les aliments contenant de grandes quantités de glutamate pourraient augmenter la tolérance à l'acide de *L. monocytogenes* dans l'estomac.

Egalement, il semblerait que la survie de *L. monocytogenes* exposée à un stress acide est sous la dépendance de l'expression de protéines *sigB*-dépendantes. En effet, Wiedmann *et al.* (1998) ont montré qu'un mutant *sigB* était moins résistant au stress acide que la souche sauvage. Ferreira *et al.* (2003) ont montré que *L. monocytogenes*, en phase exponentielle, avait une survie et une résistance accrues aux fluides gastriques après une exposition à des conditions légèrement acides, et qu'elle était partiellement dépendante de  $\sigma^B$ .

## **1-2. Survie au stress osmotique**

L'ajout de sels pour diminuer l' $a_w$  des aliments est une méthode ancestrale de conservation qui est encore utilisée aujourd'hui dans l'industrie agroalimentaire. Cependant, *L. monocytogenes* peut survivre à de fortes concentrations de sels et rend donc difficile le contrôle de ce pathogène dans les aliments. De plus, le lumen du tractus gastro-intestinal est une région avec une forte salinité (0.3 M NaCl). Le passage de la bactérie de l'estomac au petit intestin représente également un choc osmotique.

*L. monocytogenes* possède des systèmes de transport d'osmolytes lui permettant d'augmenter son osmotolérance. Ces systèmes transportent des osmolytes dans la cellule bactérienne afin de contrebalancer la pression exercée et de prévenir la perte d'eau du cytoplasme. Les transporteurs membranaires BetL et Gbu (Ko & Smith, 1999) sont les principaux systèmes permettant la capture de la glycine et de la bétaine. OpuC est également capable de transporter la glycine bétaine mais ses fonctions primaires sont le transport de la carnitine, un osmolyte associé aux tissus animaux. Il a été montré que OpuC est nécessaire

pour une virulence maximale chez la souris (Sleator *et al.*, 2001). Par contre, la mutation de *gbu* et *betL* n'a pas d'impact sur la virulence (Wemekamp-Kamphuis *et al.*, 2002). Cependant, d'autres études ont montré que Gbu et BetL étaient essentiels pour la survie dans les aliments, et spécialement ceux d'origine végétale. Il semble donc que tous ces systèmes sont importants pour que *L. monocytogenes* puisse réaliser son cycle de virulence de l'aliment à son hôte.

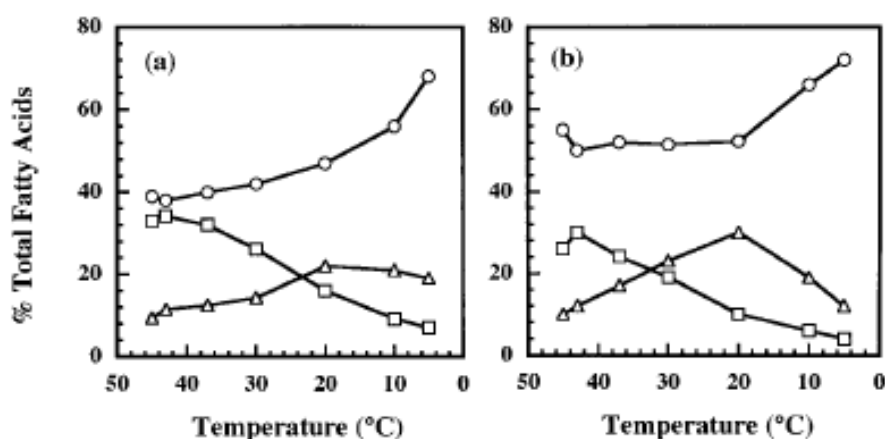
Un autre mécanisme utilisé pour survivre au stress salin est la variation de l'expression génétique, entraînant l'augmentation ou la diminution de la synthèse de nombreuses protéines. Duché *et al* (2002a) ont étudié le protéome de *L. monocytogenes* soumise à un stress salin, par électrophorèse bidimensionnelle. 12 protéines induites par le stress salin ont été repérées et pouvaient être classées en deux catégories. Certaines protéines, appelées protéines de stress salin (Ssp), étaient induites rapidement mais surexprimées sur une courte période. D'autres protéines, qualifiées de protéines d'acclimatation au stress (Sap), étaient rapidement induites également, mais continuaient à l'être plusieurs heures après la fin du stress. Deux protéines de stress (DnaK et Ctc) ont été identifiées parmi les Ssp. Parmi les 11 Saps identifiées, GbuA qui fonctionne comme un transporteur de l'osmoprotectant glycine bétaine était induit en réponse au stress salin. Puisque la protéine Ctc était induite par le stress salin, Gardan *et al.* (2003) ont étudié le rôle de cette protéine dans la résistance au stress osmotique. Ils ont montré que le gène *ctc* était impliqué dans la résistance de *L. monocytogenes* aux fortes osmolarité en l'absence des osmoprotecteurs glycine bétaine et de la carnitine. L'expression de ce gène *ctc* est sous dépendance de *sigB*. Ce facteur de stress est également important pour l'utilisation de bétaine et carnitine comme osmoprotectants (Becker *et al.*, 1998).

### **1-3. Survie au stress thermique froid**

*L. monocytogenes* est capable de croître dans une large gamme de température (2-45°C). Sa survie et sa croissance aux températures de réfrigération font partie des aptitudes de ce pathogène qui rendent difficile son contrôle dans les aliments. La réfrigération puis la conservation au froid sont les principales techniques de conservation des aliments utilisées de nos jours. De ce fait la compréhension des mécanismes de survie et de croissance à faible température est importante pour apporter des informations qui pourraient permettre de mieux contrôler ce pathogène.

Un des mécanismes utilisés par *L. monocytogenes* pour survivre au froid est l'adaptation de la composition lipidique membranaire. A température élevée, la membrane est fluide, les

chaînes acylées sont dans un état de désordre. Chez de nombreuses bactéries, lorsque la température baisse, les chaînes acylées des phospholipides s'organisent en un réseau ordonné et les membranes se gélifient. Lorsque la membrane est gélifiée, aucun échange n'est possible. Pour que la membrane puisse jouer son rôle dans les échanges de solutés et permettre l'assemblage des protéines de transport et le fonctionnement des enzymes membranaires, elle doit être fluide. Chez *L. monocytogenes*, les changements de température entraînent une altération dans la composition lipidique de la membrane, permettant le maintien de la fluidité membranaire idéale pour une activité enzymatique correcte et le transport de solutés à travers la membrane. Un des principaux changements est une forte augmentation de la proportion d'anteiso-C<sub>15:0</sub> en dessous de 20°C au détriment d'anteiso-C<sub>17:0</sub> et une légère diminution d'iso-C<sub>15:0</sub>, effet qui s'accroît avec la diminution de température (Figure 16) (Annous *et al.*, 1997). Le changement de température de croissance de 20 à 5°C entraîne donc une diminution de la longueur des acides gras (une diminution de C<sub>17:0</sub>) et un passage des branchements iso à anteiso (i-C<sub>15:0</sub> à a-C<sub>17:0</sub>) (Figure 16).



**Figure 16 : Effet de la température de croissance sur la composition des acides gras majeurs de *L. monocytogenes* 10403 (a) et SLCC53 (b). Les bactéries ont été cultivées en milieu « soy broth » jusqu'en milieu de phase exponentielle à la température indiquée. Symboles : rond (a-C<sub>15:0</sub>), carré (a-C<sub>17:0</sub>), triangle (i-C<sub>15:0</sub>) (Annous *et al.*, 1997)**

La diminution de la longueur des chaînes d'acides gras diminue les interactions carbone-carbone entre les chaînes voisines de la membrane, aidant à maintenir le degré optimum de la fluidité membranaire pour la croissance à basses températures (Beales, 2004).

De plus, *L. monocytogenes* est capable d'accumuler des solutés, comme la glycine betaine et la carnitine, qui peuvent agir comme des cryoprotectants (Angelidis & Smith, 2003). Ceci est possible grâce aux transporteurs BetL (glycine betaine porter I), Gbu (glycine betaine

porter II) et Opu (transporteur de la carnitine). Sigma B qui est sur-exprimé lors d'un stress froid semble être impliqué dans l'accumulation de ces cryoprotectants puisque le mutant *sigB* est incapable de les accumuler.

Liu *et al* (2002) ont identifié des ARN<sub>m</sub> dont la synthèse augmentait en réponse à un choc froid de 37°C à 10°C. L'augmentation de l'expression en ARNm de protéases chaperonne comme GroEL, ClpP et ClpB indique que ces enzymes pourraient être impliquées dans la dégradation de polypeptides endommagés lors du choc froid.

Bayles *et al* (1996) se sont intéressés aux protéines induites par un choc froid, en comparant par électrophorèse bidimensionnelle, le profil protéique de *L. monocytogenes* cultivée à 37°C ou 5°C ou transférées de 37°C à 5°C. Ils ont montré que *L. monocytogenes* produisait des protéines de choc froid (Csps) en réponse à la chute de température et des protéines d'acclimatation au froid (Caps) pendant la croissance à faible température. Douze protéines (Csps) étaient induites par le choc froid dont 4 n'étaient présentes que dans les cellules cultivées à 5°C et pouvaient ainsi être considérées comme des protéines d'acclimatation au froid.

Phan-Thanh and Gormon (1995) ont, quant à eux, montré la surexpression de 38 protéines Csps lors d'un choc froid de 25 à 4°C. Il est étonnant de remarquer que 6 de ces Csps étaient également surexprimées lors d'un choc chaud de 25 à 49°C. Parmi ces dernières, une protéine de 17.6 kDa était aussi bien induites par un stress thermique chaud que froid. Cette protéine également mise en évidence comme une csp dans d'autres études sur la réponse aux chocs froids (Bayles *et al.*, 1996 ; Hébraud & Guzzo, 2000). Elle a été identifiée comme appartenant à la famille des ferritines (Hébraud & Guzzo, 2000) et correspondait au produit du gène *fri* de *L. monocytogenes*. Les ferritines sont connues pour être des protéines de stockage du fer. Elles peuvent être utilisées pour améliorer la croissance quand il y a peu de fer disponible dans le milieu extracellulaire. Elles permettent également de protéger les bactéries contre la toxicité du fer en séquestrant ce métal. Cependant, cette protéine étant surexprimée lors de chocs thermiques, il est possible qu'elle ait un rôle plus important dans l'adaptation aux stress environnementaux.

## 2 - LA FERRITINE

Lors de son ingestion par l'homme, *L. monocytogenes* est, comme nous l'avons vu précédemment, soumis à de nombreux stress. Les conditions environnementales dans l'intestin mais également lors de son passage dans la lymphe ou le sang sont peu favorables à



son développement. L'une des carences auxquelles est notamment confronté le microorganisme est l'absence de fer à l'état libre. La ferritine (Fri), protéine capable de stocker le fer est impliquée dans l'adaptation à cette carence. La bibliographie montre que le rôle de Fri serait beaucoup plus large, avec une participation dans l'adaptation à divers stress, et indirectement à la virulence. Cette partie bibliographie propose de faire le point sur le rôle complexe de la ferritine, avec une première partie rappelant le rôle primordial du fer chez *L. monocytogenes*

## **2-1. Le fer, élément indispensable à la vie**

Le fer existe sous deux formes : la forme ferreuse  $Fe^{2+}$  (forme réduite) et la forme ferrique  $Fe^{3+}$  (forme oxydée). A quelques exceptions près, le fer est devenu indispensable à toute forme de vie. Il est impliqué dans de nombreux processus biologiques, comme la photosynthèse, la fixation du  $N_2$ , la production et la consommation d' $H_2$ , la respiration, le cycle TCA, le transport de l'oxygène, la régulation de gène et la biosynthèse de l'ADN.

Le fer est un métal abondant, puisque c'est le quatrième élément le plus fréquent à la surface de la terre. Pour cette raison, le fer est considéré comme ayant été un choix idéal pour l'incorporation dans les protéines lors de l'apparition de la vie. La vie est apparue dans un environnement anaérobie et le fer, sous forme ferreuse hydrosoluble, était biodisponible. Lorsque les microorganismes photosynthétiques émergèrent, transformant l'eau en oxygène, le fer devenant ferrique ( $Fe^{III}$ ) précipita irréversiblement sous formes d'oxydes et d'hydroxydes. Les réactions entre le fer et l'oxygène engendrèrent des radicaux hautement toxiques (superoxyde, hydroxyle). L'arrivée de l' $O_2$  fut donc une catastrophe à laquelle peu d'organismes vivants survécurent. Ils durent développer des mécanismes protecteurs contre les radicaux de l'oxygène et des mécanismes pour solubiliser le fer qui leur était nécessaire. Le «paradoxe du fer» réside entre son rôle vital, donc sa nécessité, et sa forte toxicité. Son homéostasie doit donc être très précisément régulée par les organismes vivants, carence et surcharge constituant des problèmes majeurs.

Les bactéries utilisent essentiellement 5 stratégies pour assurer l'homéostasie du fer :

1. Des transporteurs à haute affinité pour le fer, permettant sa capture sous formes variées, dans l'environnement de la bactérie.
2. Possession de systèmes pour le stocker, afin d'apporter une source de fer quand la bactérie se retrouve dans un environnement carencé en fer.
3. Utilisation de systèmes de résistance aux stress redox.

4. Contrôle de la consommation de fer en diminuant l'expression de protéines nécessitant du fer, lorsque celui-ci est faiblement disponible.
5. Une régulation sensible au fer qui coordonne l'expression des différents systèmes cités ci-dessus selon la disponibilité en fer dans le milieu environnant

## **2-2. Importance du fer chez *L. monocytogenes***

L'acquisition de fer en quantité assez importante est un problème auquel doivent plus particulièrement faire face les bactéries pathogènes. En effet, une des défenses de l'hôte contre les microorganismes consiste à limiter la quantité de fer disponible. Les mammifères utilisent des protéines se liant au fer (transferrine, lactoferrine) afin de réduire la quantité de fer extracellulaire libre à environ  $10^{-18}$  M, quantité insuffisante pour la croissance des bactéries. De plus, l'hôte produit des protéines qui se lient à l'hème et à l'hémoglobine (haemopexine et haptoglobine) ce qui limite la disponibilité de l'hème comme source de fer pour les bactéries pathogènes. Cependant, ces bactéries ont développé des systèmes leur permettant de récupérer ce fer séquestré.

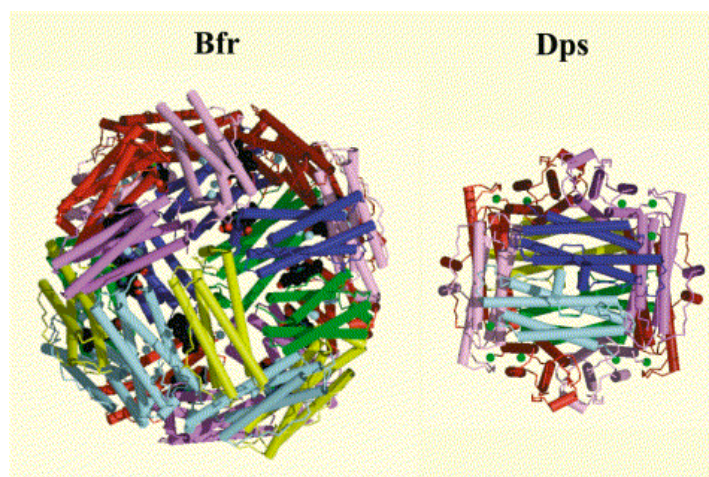
*L. monocytogenes* ne peut pas se développer dans des milieux carencés en fer. Le rôle direct du fer dans le pouvoir pathogène expérimental de *L. monocytogenes* a été décrit initialement par Sword (1966). Cet auteur a montré que la LD50 de *L. monocytogenes* était plus faible pour des souris surchargées en fer que pour des animaux non traités.

Il a également été montré que le fer influençait l'expression d'importants gènes de virulence impliqués dans l'invasion (internalines), l'échappement des vacuoles (listériolysine O) ou dans le mouvement intracellulaire chez l'hôte (ActA) (Coward, 1981 ; Conte *et al.*, 1996 ; Conte *et al.*, 2000). La présence de quantités importantes de fer facilitait l'invasion des cellules et la transcription du cluster *inlAB*, tandis qu'une augmentation de l'expression d'Hly et d'ActA était observée dans des conditions de restriction en fer. La différence de régulation d'*inlAB*, *hly* et *actA* reflète une expression fine, temporelle et fer-dépendante de ces gènes. En effet, lorsque *L. monocytogenes* est dans le milieu extracellulaire, où du fer est disponible, il est important que la bactérie exprime *inlAB* afin de s'attacher et de pénétrer dans les cellules de l'hôte et réprime *hly* qui pourrait endommager les cellules de l'hôte. Par contre une fois dans la cellule, la bactérie a besoin d'Hly pour s'échapper de la vacuole et d'ActA pour être mobile. Le fer pourrait donc être un signal permettant une régulation appropriée de ces gènes.

## 2-3. Les ferritines bactériennes

### 2-3.1. Ferritines, bacterioferritines et protéines Dps (DNA-Binding-protein)

Le fer extracellulaire n'est pas la seule source disponible en fer, de nombreuses bactéries possèdent des réserves intracellulaires de fer. Des protéines de stockage du fer peuvent être utilisées pour améliorer la croissance quand il y a peu de fer disponible dans le milieu extracellulaire. Elles permettent également de protéger les bactéries contre la toxicité du fer en séquestrant ce métal. Trois types de protéines de stockage du fer ont été identifiées chez les bactéries : les ferritines, également présentes chez les eucaryotes ; les bacterioferritines trouvées uniquement chez les eubactéries et les protéines Dps présentes seulement chez les procaryotes (revue : (Andrews *et al.*, 2003). Il est intéressant de constater que ces 3 types de protéines peuvent être présentes en même temps dans un même organisme. Bien que ces trois formes de stockage du fer forment des familles évolutives différentes, elles ont de nombreuses structures et fonctions similaires. L'architecture de ces protéines est la clef de leur capacité à stocker le fer. Elles sont composées de 24 (ferritines et bacterioferritines) ou de 12 (protéines Dps) sous-unités similaires qui constituent une protéine en forme de coquille sphérique entourant une cavité centrale et agissant comme un réservoir de stockage du fer (Figure 17).



**Figure 17 : Structures de la bacterioferritine (24 sous-unités) et de Dps (12 sous-unités) chez *E. coli***

Les ferritines et les bacterioferritines peuvent stocker 2000 à 3000 atomes de fer par 24-mer, tandis que les protéines Dps, plus petites, ont une capacité de stockage plus faible, de l'ordre de 500 atomes de fer par 12-mer.

Ces polymères lient le fer présent sous forme ferreuse, mais le stockent dans la cavité centrale sous sa forme oxydée ferrique. Le processus de stockage nécessite donc une étape de

ferroxydation qui est catalysée par des sites spécifiques présents dans les protéines de stockage. Pour les ferritines et les bactérioferritines, ce site s'appelle le centre ferroxidase et est localisé dans la région centrale de chaque sous-unité. Ces résidus sont fortement conservés et agissent comme des ligands pour la liaison de deux ions ferreux, qui sont ensuite oxydés par l'oxygène, résultant en la formation d'un intermédiaire  $(\text{Fe}^{3+})_2\text{O}$ . Les ions ferriques ainsi formés vont ensuite migrer vers la cavité centrale.

Les résidus ferroxidase ne sont pas conservés chez les protéines Dps. Ces protéines lient et oxydent le fer ferreux à un site totalement différent. Les résidus impliqués dans la liaison du fer à ce site sont bien conservés chez les protéines Dps, mais pas chez les ferritines et les bactérioferritines.

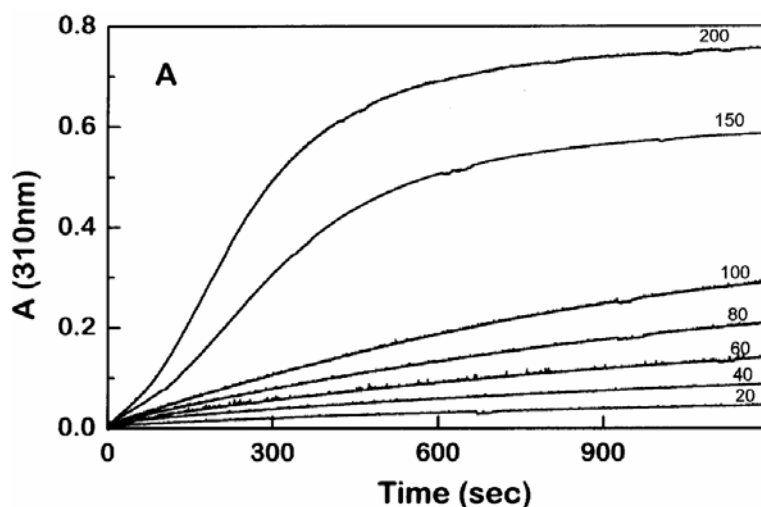
### 2-3.2. La ferritine chez *Listeria*

Une protéine capable de lier et incorporer le fer chez *L. innocua* a été découverte en 1997 (Bozzi *et al.*, 1997). Cette découverte est intéressante puisque c'est le premier exemple d'une ferritine authentique qui possédait une grande similarité de séquence avec les protéines Dps. Les DNA-binding proteins from starved cells (Dps) sont une classe de protéines exprimées par des bactéries sous des conditions de stress oxydatif ou nutritionnel qui lient l'ADN sans spécificité de séquence et le protège contre les dommages oxydatifs. Les premières protéines Dps ont été découvertes chez *E. coli*. Cette protéine était induite en phase stationnaire, par le facteur SigB et se liait à l'ADN de manière non spécifique, afin de protéger l'ADN des stress redox (Almiron *et al.*, 1992). Des études récentes ont montré que cette protéine Dps chez *E. coli* pouvait aussi stocker le fer, mais préférant l' $\text{H}_2\text{O}_2$  comme oxydant, l' $\text{O}_2$ , généralement utilisé par les ferritines, n'étant qu'une alternative (Zhao *et al.*, 2002).

Contrairement à la protéine Dps d'*E. coli*, la protéine Dps de *L. innocua* ne peut pas se lier à l'ADN, certainement en raison de l'absence de résidus lysine en N-terminal, motif conservé chez les protéines Dps (Bozzi *et al.*, 1997). Par contre, le produit final de la réduction du dioxygène chez cette ferritine est l'eau (Yang *et al.*, 2000), et non le peroxyde d'hydrogène comme chez les ferritines des mammifères. Comme chez *E. coli*, il semblerait donc que l' $\text{H}_2\text{O}_2$  soit un meilleur oxydant du  $\text{Fe}^{2+}$  puisque l' $\text{H}_2\text{O}_2$  n'est pas observé comme produit final. L'effet protecteur de cette protéine contre la combinaison toxique  $\text{Fe}^{2+} + \text{H}_2\text{O}_2$  a été montré par Su *et al.* (2005).

Les protéines Dpr de *Streptococcus mutans* (Yamamoto *et al.*, 2002), Dps de *H. pylori* (NapA) (Cooksley *et al.*, 2003) et Dps de *Campylobacter jejuni* (Ishikawa *et al.*, 2003) permettaient également de stocker du fer et de résister au peroxyde.

La ferritine de *L. innocua* est une protéine de 240 000 Da composée de 12 sous-unités de 18 000 Da. Chaque dodecamer peut oxyder jusqu'à 500 atomes de fer. Au delà, on observe une précipitation du fer. Stefanini *et al* (1999) se sont intéressés à la cinétique d'incorporation du fer par cette protéine Dps. En dessous de 150 atomes de Fe(II) par dodecamer, le temps d'oxydation du fer par *L. innocua* est hyperbolique. Lorsqu'il y a 150 atomes de Fe(II) par dodecamer ou plus, le temps d'oxydation est sigmoïdal (Figure 18).



**Figure 18 : Courbe de progression de l'incorporation du fer par *L. innocua* en fonction de la concentration en Fe(II) (Stefanini *et al.*, 1999)**

De plus, lorsque le pH augmente, l'oxydation du fer augmente (Stefanini *et al.*, 1999). Cette protéine Dps est également présente chez *L. monocytogenes* et est codée par le gène *fri* (Glaser *et al.*, 2001). L'expression de ce gène augmente légèrement lors de l'entrée en phase stationnaire aussi bien chez *L. innocua* que chez *L. monocytogenes*. Il est fortement induit dans un milieu pauvre en fer (Polidoro *et al.*, 2002), contrairement aux ferritines classiques. Il semble donc qu'en plus de son rôle de stockage du fer et de protection contre l' $H_2O_2$ , Fri pourrait être impliquée dans l'adaptation à d'autres stress environnementaux comme un milieu carencé en fer ou le passage de la phase exponentielle à la phase stationnaire. Par ailleurs, cette protéine était sur-exprimée lors de nombreux stress tels que les chocs thermiques comme nous l'avons vu précédemment (Phan-Thanh & Gormon, 1995 ; Bayles *et al.*, 1996 ; Hébraud & Guzzo, 2000), mais également lors de stress dus à la présence de SDS, de deoxycholate ou d'éthanol (Phan-Thanh & Gormon, 1997).

Il est intéressant de noter que les protéines Dps de nombreux autres organismes, capables de lier le fer, sont surexprimées lors de stress. Par exemple la protéine Dpr de *S. mutans* (Yamamoto *et al.*, 2002) est synthétisée en plus grande quantité lorsqu'elle est exposée à l'air et un mutant *dpr* est incapable de former des colonies sur un milieu gélosé quand la croissance a lieu en présence d'air, la protéine Dps de *H. pylori* (NapA) (Cooksley *et al.*, 2003), est, quant à elle, sur-exprimée en phase stationnaire de croissance, comme chez *Listeria*.

Enfin, il a été montré chez *Porphyromonas gingivalis* qu'une ferritine dodécamerique contribuait à sa survie dans des cellules HUVEC (Ueshima *et al.*, 2003) et que la protéine Dps de *Salmonella enterica* sérovar *typhimurium* contribuait à sa survie dans les macrophages chez la souris et à la virulence de la bactérie (Halsey *et al.*, 2004).

Ces protéines Dps, capables de lier le fer, ont donc un rôle plus large que celui de stockage du fer. Elles semblent jouer un rôle essentiel dans la survie des cellules face à de nombreux stress et peuvent avoir un rôle dans la virulence des espèces pathogènes.



# RESULTATS





## Partie 1

---

Caracterisation du mutant ferritine ( $\Delta fri$ ) de *Listeria*  
*monocytogenes* EGDe  
(Article 1)

---

## 1 - PROBLÉMATIQUE

Comme nous l'avons évoqué dans la partie bibliographique, *L. monocytogenes* fait partie des bactéries pathogènes pour l'homme. Elle est responsable chaque année de graves infections alimentaires, mortelles dans un quart des cas. En plus de sa pathogénicité, *L. monocytogenes* se caractérise par son importante capacité de résistance et d'adaptation à des conditions environnementales hostiles. Elle est notamment capable de survivre et de croître aux températures de réfrigération, ce qui fait partie des multiples éléments rendant difficile son contrôle dans les aliments. Différentes études protéomiques ont montré la synthèse de protéines de choc froid (Csps pour Cold shock proteins) en réponse à la chute de température et de protéines d'acclimatation au froid (Caps) pendant la croissance à basse température (Bayles *et al.*, 1996). Il est surprenant de constater que certaines Csps étaient également surexprimées lors d'un choc chaud (Phan-Thanh & Gormon, 1995). Parmi ces protéines aussi bien induites par un stress thermique chaud ou froid, une protéine de 18 kDa montrait une très forte surexpression. Des travaux réalisés au laboratoire ont permis de caractériser cette protéine majeure de choc froid (Hébraud & Guzzo, 2000). Il a été montré que cette protéine de 18 kDa et de *pI* 5,1 avait 98% d'identité de séquence avec la ferritine de *L. innocua* (Bozzi *et al.*, 1997). Cette protéine était également sur-exprimée lors de différents autres stress (Phan-Thanh & Gormon, 1997)

Les ferritines sont connues pour être des protéines de stockage du fer. Elles peuvent être utilisées pour améliorer la croissance quand il y a peu de fer disponible dans le milieu extracellulaire. Elles permettent également de protéger les bactéries contre la toxicité du fer en séquestrant ce métal. De plus, il a été montré que des ferritines dodécamériques contribuaient à la virulence de la bactérie (Ueshima *et al.*, 2003 ; Halsey *et al.*, 2004).

La ferritine de *L. monocytogenes* étant surexprimée lors de différents stress, il semble qu'elle puisse jouer un rôle plus important que le simple stockage du fer. L'objectif de cette première étude était de mieux comprendre l'importance de la ferritine dans l'adaptation de la bactérie aux stress environnementaux et dans la virulence.

## 2 - STRATEGIE

Afin d'étudier le rôle de la ferritine (codée par le gène *fri*) de *L. monocytogenes*, un mutant  $\Delta$ *fri* a été construit par échange allélique (Dussurget O., Institut Pasteur de Paris) chez la

souche EGDe dont le génome a été séquencé (Glaser *et al.*, 2001). La caractérisation de ce mutant a été effectuée en collaboration avec l'UIBC de l'institut Pasteur de Paris. Pour notre part, nous avons cultivé la souche sauvage et le mutant en présence de différentes concentrations en fer, afin de déterminer l'importance de cet élément sur la croissance. Enfin, le protéome de la souche sauvage et du mutant  $\Delta fri$  ont été comparés par électrophorèse bidimensionnelle.

En parallèle, les études menées à l'institut Pasteur ont consisté à comparer la survie de la souche sauvage et du mutant  $\Delta fri$  chez la souris, après injection intraveineuse d'une population non létale de bactéries, et de comparer la résistance des souches à un stress oxydatif.

### 3 - PRINCIPAUX RESULTATS

La caractérisation fonctionnelle du mutant  $\Delta fri$  a permis de montrer que Fri avait un rôle dans la virulence de *L. monocytogenes*. En effet, 48h après l'infection de souris BALBc par injection intraveineuse, les souris infectées par le mutant présentent 10 fois moins de bactéries dans la rate et 100 fois moins dans le foie que celles infectées par la souche sauvage de *L. monocytogenes*. De plus, 72H après l'infection, aucune souche du mutant  $\Delta fri$  n'était détectable dans le cerveau, alors que la souche sauvage était encore présente.

Suite à une préculture dans un milieu minimum sans fer, la souche sauvage et la souche  $\Delta fri$  avaient une densité optique finale identiques quelle que soit la concentration en fer lors de la culture. Par contre, la phase de latence de  $\Delta fri$  était nettement plus longue que celle du sauvage (16,9 +/- 1,7h versus 6,3 +/- 1,2h) Lorsque la préculture était réalisée en présence de fer, la croissance du mutant était inchangée. Par contre, la souche sauvage atteignait une densité optique plus élevée, pour des concentrations en fer égales ou supérieures à 1  $\mu\text{g/ml}$ . Les phases de latence des deux souches étaient équivalentes à celles observées pour des précultures sans fer. Il semble donc que le fer stocké par la ferritine puisse être libéré afin d'améliorer la croissance en condition limitée en fer.

Les expériences ont également pu montrer que le mutant était hypersensible au stress oxydatif et qu'il atteignait une DO légèrement moins élevée que la souche sauvage après un choc thermique chaud ou froid. Les résultats suggèrent que la ferritine de *L. monocytogenes* a un rôle protecteur contre divers stress. Le rôle protecteur de la ferritine pourrait s'expliquer en grande partie par la séquestration du fer. De plus, Fri pourrait avoir un rôle indirect dans la régulation de l'expression et/ou la stabilité de protéines impliquées dans l'adaptation au

stress. La comparaison des protéomes du mutant *fri* avec la souche sauvage a permis de mettre en évidence plusieurs protéines présentant des différences d'abondance, dont 15 ont pu être identifiées. Il est intéressant d'observer que certaines de ces protéines sont impliquées dans la réponse au stress ou dans la virulence.

Le rôle de la ferritine dans la virulence de *L. monocytogenes* est donc probablement dû à la combinaison de sa capacité à stocker le fer, à réguler des protéines de stress et de virulence et à protéger la bactérie contre de nombreux stress.

## Article n°1

O. Dussurget, E. Dumas, C. Archambaud, I. Chafsey, C. Chambon, M. Hébraud, P. Cossart

*Listeria monocytogenes* ferritin protects against multiple stresses and is  
required for virulence

FEMS Microbiology Letters 250(2): 253-261 (2005)

## *Listeria monocytogenes* ferritin protects against multiple stresses and is required for virulence

Olivier Dussurget<sup>a,1</sup>, Emilie Dumas<sup>b,1</sup>, Cristel Archambaud<sup>a</sup>, Ingrid Chafsey<sup>b</sup>,  
Christophe Chambon<sup>c</sup>, Michel Hébraud<sup>b,c,\*</sup>, Pascale Cossart<sup>a,\*</sup>

<sup>a</sup> *Unité des Interactions Bactéries-Cellules, INSERM U604, INRA USC2020, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France*

<sup>b</sup> *QuaPA-Microbiologie, INRA de Theix, 63122 Saint-Genès Champanelle, France*

<sup>c</sup> *Plate-forme protéomique, INRA de Theix, 63122 Saint-Genès Champanelle, France*

Received 30 April 2005; received in revised form 13 July 2005; accepted 13 July 2005

First published online 10 August 2005

Edited by M. Mitsuyama

### Abstract

In this study, the role of *Listeria monocytogenes* ferritin was investigated. The *fri* gene encoding the ferritin was deleted and the phenotype of the mutant was analyzed demonstrating that ferritin is necessary for optimal growth in minimal medium in both presence and absence of iron, as well as after cold- and heat-shock. We also showed that ferritin provides protection against reactive oxygen species and is essential for full virulence of *L. monocytogenes*. A comparative proteomic analysis revealed an effect of the *fri* deletion on the levels of listeriolysin O and several stress proteins. Together, our study demonstrates that *fri* has multiple roles that contribute to *Listeria* virulence.

© 2005 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

**Keywords:** *Listeria*; Ferritin; Dps; Proteome; Stress; Pathogenesis

### 1. Introduction

Iron is an essential nutrient for most prokaryotes, including pathogenic bacteria. However, iron is highly insoluble at physiological pH in aerobic environments and can react with oxygen intermediates to produce deleterious free radicals. Bacteria have developed several mechanisms to acquire iron while maintaining it in a nontoxic state. These include secretion of siderophores

for iron uptake and synthesis of ferritins for iron storage [1]. Three types of ferritins are produced by bacteria: dodecameric ferritins, tetracosameric ferritins and tetracosameric heme-containing bacterioferritins [2]. Dodecameric ferritins, including Dps (DNA-binding protein from starved cells) and Dps-like proteins, are formed of 12 identical subunits ( $\approx 20$  kDa) assembling into a spherical shell which can accommodate approximately 500 iron atoms. Tetracosameric ferritins, found in both prokaryotes and eukaryotes, and bacterioferritins, only present in bacteria and having heme as a cofactor, are 24-mer spheric structures containing 2000–4500 iron atoms. Apart from their conserved function of iron binding and storage, some ferritins can protect against reactive oxygen species and iron overload, others can also bind and protect DNA [2].

\* Corresponding authors. Tel.: +33 4 73 62 46 70; fax: +33 4 73 62 42 68 (M. Hébraud), Tel.: +33 1 40 61 30 32; fax: +33 1 45 68 87 06 (P. Cossart).

E-mail addresses: [hebraud@clermont.inra.fr](mailto:hebraud@clermont.inra.fr) (M. Hébraud), [pcossart@pasteur.fr](mailto:pcossart@pasteur.fr) (P. Cossart).

<sup>1</sup> Contributed equally to this work.

*Listeria monocytogenes* is a pathogenic Gram-positive bacillus widely distributed in the environment and encodes a single ferritin gene, *fri* [3]. *L. monocytogenes* ferritin is a major cold shock protein which is also strongly overexpressed after heat shock or chemical stress [4–6]. Ferritins from *L. monocytogenes* and the non-pathogenic bacterium *L. innocua* have 98% identity at the amino acid level [3]. *L. innocua* ferritin belongs to the dodecameric ferritin subfamily [7]; however, it cannot bind DNA, probably because it lacks N-terminal lysine residues conserved in Dps proteins [7]. The expression of *fri* slightly increases upon entry into stationary phase of both *L. innocua* and *L. monocytogenes* [8] and is strongly induced in low iron medium [8]. Interestingly, major virulence factors of *L. monocytogenes*, such as listeriolysin O (LLO) and the protein responsible for actin-based motility ActA, are overproduced in low iron conditions, a signal mimicking that encountered in the host during the infectious process [9,10]. Here, we inactivated *fri* to analyze the functions of the ferritin in *L. monocytogenes*.

## 2. Materials and methods

### 2.1. Bacterial strains and growth conditions

*L. monocytogenes* EGDe (BUG1600) and *L. monocytogenes* EGDe  $\Delta$ *fri* (BUG1962) strains were grown at 37 °C in brain heart infusion (BHI) medium (Difco). When indicated, *L. monocytogenes* strains were grown in a chemically defined improved minimal medium (IMM) [11]. When required, chloramphenicol was added at 7 µg/ml and erythromycin at 5 µg/ml. *Escherichia coli* TOP10 strains were grown in Luria–Bertani (LB) medium (Difco) at 37 °C. When required, antibiotics were included at the following concentrations: ampicillin, 100 µg/ml; kanamycin, 20 µg/ml.

For iron experiments, *L. monocytogenes* strains were first cultured for 24 h at 37 °C in IMM containing 20 µg/ml of FeSO<sub>4</sub>. These precultures were used to inoculate IMM containing 20 µg/ml of FeSO<sub>4</sub> or no iron at the initial OD<sub>600nm</sub> of 0.1 and cultures were incubated 24 h at 37 °C. After centrifugation at 5000g for 2 min, bacterial pellets were washed twice in IMM and resuspended in IMM. Growth under various iron concentrations was measured with a Microbiology reader bioscreen (Labsystems) in sterile 100 well-microplates containing 300 µl of culture in each well. Experiments were performed at least twice and were repeated twice independently.

For cold- and heat-shock experiments, *L. monocytogenes* grown in BHI to stationary phase was used to inoculate BHI at an OD<sub>600nm</sub> of 0.1. Growth was monitored at 600 nm using a spectrophotometer. When

OD<sub>600nm</sub> was between 0.3 and 0.5, cultures were transferred to a waterbath at 4, 8 or 45 °C. Growth was monitored until stationary phase. Experiments were performed at least in duplicate and were repeated twice independently.

### 2.2. Mutagenesis

DNA extractions were performed as described by Archambaud et al. [12]. A 1002-bp DNA fragment containing the *fri* upstream sequence was generated by PCR using oligonucleotides U175 (5'-AAGGATCCAATCATATAAGCAATACC-3') and L176 (5'-AAACGCGTTCTCCTATTTTATAATG-3'). The fragment was cloned into the *Bam*HI/*Mlu*I-digested thermosensitive shuttle plasmid pMAD [13], constructing pOD62. A 987-bp DNA fragment containing the *fri* downstream sequence was generated by PCR using oligonucleotides U177 (5'-AAACGCGTAGAGTAATAAAAACGTG-3') and L178 (5'-AAAGATCTTGTAGCTATTACTTATC-3'). The fragment was cloned into *Mlu*I/*Bgl*II-digested pOD62, constructing pOD64. To achieve allelic exchange, pOD64 was electroporated into *L. monocytogenes* EGDe at 2500 V, 200 Ω and 25 µF. Homologous recombinations and deletion of the target gene *fri* were carried out as described by Arnaud et al. [13]. Mutants were identified by PCR on colonies using oligonucleotides U189 (5'-AGATAAATAATGAGAAAATGG-3') and L190 (5'-ATGAAGATAATATACCTCTTTC-3'). Deletion of the *fri* gene in the *fri* mutant strain (BUG1962) was verified by PCR and Southern blot.

### 2.3. Complementation

A 1.5-kbp DNA fragment containing the *fri* gene and its promoter was amplified by PCR using oligonucleotides U175 and L226 (5'-AACTGCAGGAATATTGTACTATAACAG-3'). The PCR product was digested with *Bam*HI and *Pst*I and cloned into the replicative plasmid pMK4 [14], constructing pOD85. The plasmid pOD85 and pMK4 were electroporated in the *fri* mutant at 2500 V, 200 Ω and 25 µF, creating strains BUG2218 and BUG2219, respectively.

### 2.4. Determination of growth inhibition by zone diffusion assay

*L. monocytogenes* were grown in BHI at 37 °C to OD<sub>600nm</sub> = 0.8. After spreading 500 µl of culture on BHI plates, sterile paper disks (diameter, 6 mm) were placed on the agar. Ten microliters of the following reagents were applied to each disk: 25 mM plumbagin (Sigma), 500 mM *S*-nitroso-*N*-acetyl-penicillamine (Sigma) and 500 mM diamide (Sigma). Diameter of the zone of growth inhibition was measured in mm after overnight incubation at 37 °C.



## 2.5. Animal studies

*L. monocytogenes* growth in vivo was studied as reported elsewhere [12]. Briefly,  $5 \times 10^3$  cfu were injected intravenously to 8-week-old specific pathogen-free female BALB/c mice (Charles River). At 24, 48 and 72 h after infection, liver, spleen and brain were dissected under aseptic conditions and the number of cfu was determined by plating serial dilutions of organ homogenates on BHI agar medium.

## 2.6. Two-dimensional gel electrophoresis (2-DE), identification of proteins by mass spectrometry and phosphoprotein gel staining

*L. monocytogenes* strains were grown in BHI to stationary phase at 37 °C. These precultures were used to inoculate BHI at an OD<sub>600nm</sub> of 0.1. Cultures were incubated at 37 °C with shaking at 150 rpm during 24 h. Extraction of soluble proteins and 2-DE were carried out as previously described [15]. For each *L. monocytogenes* strain, at least four gels were analyzed for each pH gradient and from two independent bacterial cultures. After image analysis and comparison, only spots detected in all the gels but one or in all the gels were taken into account. Proteins separated by 2-DE and stained with Coomassie blue were excised from the gel and identified by mass spectrometry as described by Folio et al. [15]. Detection of phosphoproteins using Pro-Q Diamond detection kit (Molecular Probes) was performed according to the manufacturer's recommendations.

## 3. Results

### 3.1. Growth of the *L. monocytogenes* *fri* deletion mutant is affected in both high iron and low iron minimal media

The *fri* gene of *L. monocytogenes* (*lmo0943*) was inactivated by allelic exchange using the thermosensitive plasmid pMAD [13]. The wild-type *fri* mutant strains, had identical growth characteristics in BHI at 37 °C (data not shown). In minimal medium (IMM), the *fri* mutant showed an increased lag phase ( $16.9 \pm 1.7$  h) compared to that of the wild-type strain ( $6.3 \pm 1.2$  h) independently of the iron concentration (Fig. 1), suggesting that Fri is important for adaptation to nutritional stress. The *fri* mutant reached a lower optical density than the wild-type strain when grown in presence of 1 µg/ml iron or more (Fig. 1). In absence of iron, the wild-type and mutant strains reached the same optical density. The *fri* mutant had a shorter doubling time than the wild-type strain when grown in iron-free IMM and IMM containing 0.5 µg/ml iron (Fig. 1). The wild-type strain reached a lower optical density and had a

higher doubling time after preculture in iron-free IMM compared to preculture in IMM with iron (Fig. 1). In contrast, growth of the *fri* mutant was not affected by the conditions of preculture. These results indicate that Fri is able to store and restore iron taken up from iron-rich medium to promote maximal growth in iron-limited medium.

### 3.2. Growth of the *L. monocytogenes* *fri* deletion mutant is affected by heat and cold stresses

To investigate the protective effect of Fri independently of its role in iron storage, growth of the wild-type and *fri* mutant strains was measured after temperature shifts to 4, 8 or 45 °C. Both strains had a similar growth after a shift to 8 °C (Fig. 2). The mutant strain could not reach wild-type maximal optical density at 4 and 45 °C in BHI medium, which suggests that ferritin is important for full resistance to heat and cold shocks.

### 3.3. The *L. monocytogenes* *fri* deletion mutant is hypersensitive to oxidative stress

Since bacterial response to heat shock overlaps other stress responses, we tested the hypothesis of a role for Fri in response to other stresses. The sensitivity of the wild-type and *fri* mutant strains to plumbagin, a superoxide-generating agent [16], was determined using a disk diffusion assay. The sensitivity of the wild-type strain ( $6.25 \pm 0.50$  mm) was significantly lower ( $p < 0.0001$ ) than that of the *fri* mutant strain ( $11.50 \pm 0.58$  mm). Complementation of the mutant strain with the *fri* gene restored wild-type sensitivity to plumbagin ( $6.25 \pm 0.50$  mm). Transformation of the *fri* mutant with the vector alone did not change its sensitivity significantly ( $11.75 \pm 0.50$  mm). These results indicate that Fri protects *L. monocytogenes* against superoxide-induced damage. Using the same disk diffusion assay, the mutant and wild-type strains showed similar sensitivities to *S*-nitroso-*N*-acetyl-penicillamine, a NO-generating agent, and to diamide, a sulfhydryl-oxidizing agent (data not shown).

### 3.4. *L. monocytogenes* Fri is required for virulence in mice

The survival of wild-type and *fri* mutant strains was determined in BALB/c mice after intravenous injection of a sublethal number of bacteria. As soon as 48 h post-infection, there was approximately 10 times fewer mutant bacteria than wild-type bacteria in the spleen (Fig. 3). Approximately 100 times fewer mutant bacteria were recovered from the liver 48 h post-infection. A similar control of the multiplication of the mutant strain was observed at 72 h post-infection in both spleen and liver. In the brain, the multiplication of the *fri* mutant was slightly impaired compared to that of the wild-type

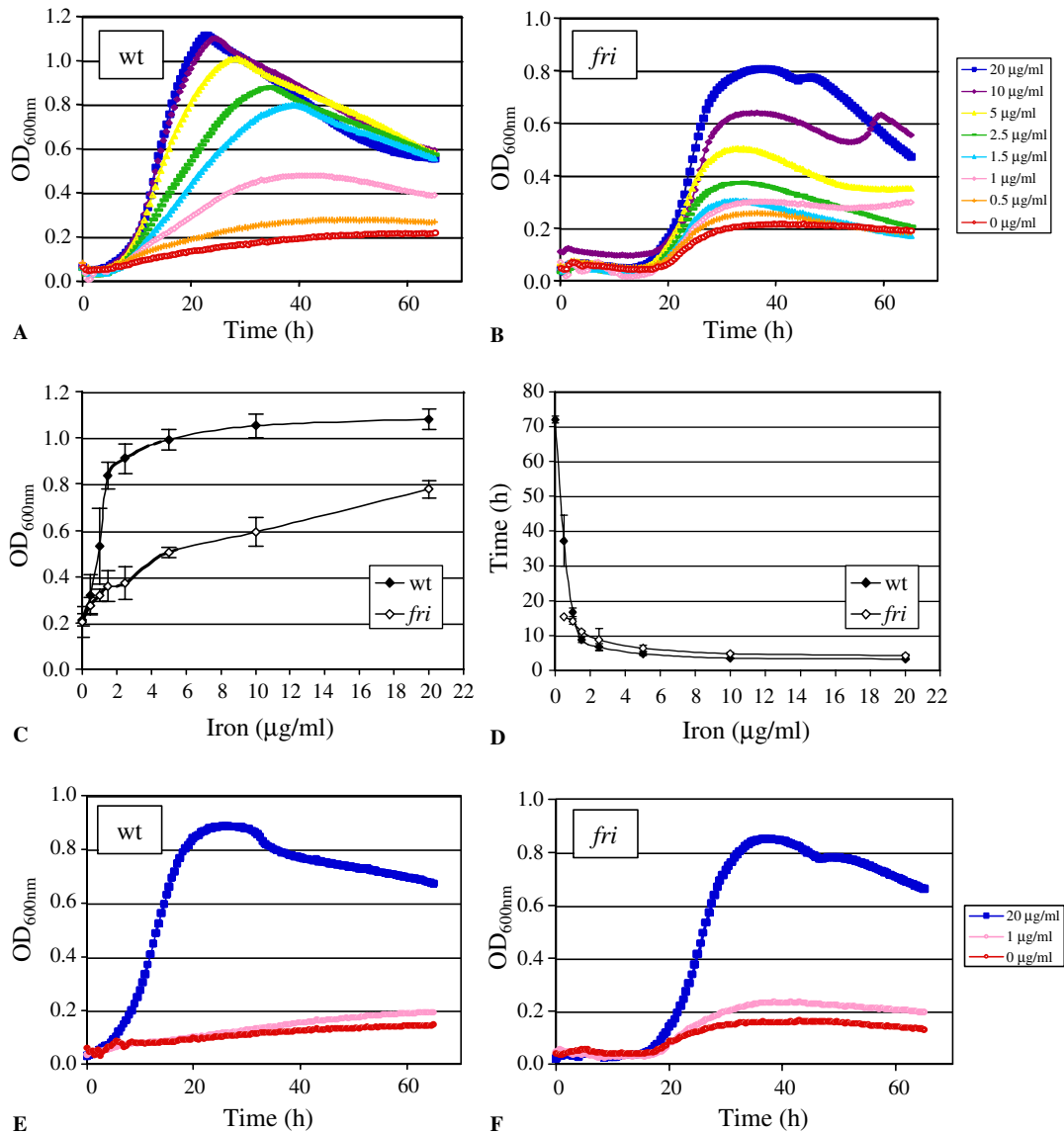


Fig. 1. Effect of the *fri* deletion on *Listeria monocytogenes* growth. Growth of the wild-type EGDe (A) and *fri* mutant (B) strains was analyzed at 37 °C in minimal medium containing different concentrations of iron by measuring bacterial density after preculture in presence of iron. Mean values of maximal OD<sub>600nm</sub> (C) and doubling times (D) are indicated with standard deviations. Growth of the wild-type EGDe (E) and *fri* mutant (F) strains was analyzed at 37 °C in minimal medium containing different concentrations of iron by measuring bacterial density after preculture in absence of iron.

strain at 48 h (<1 log). At 72 h, the mutant was cleared whereas 10<sup>4</sup> wild-type *L. monocytogenes* were recovered from the brain. These results suggest that Fri contributes to virulence of *L. monocytogenes*.

### 3.5. Inactivation of *L. monocytogenes* *fri* affects production of the major virulence factor LLO and of multiple stress proteins

Protein expression patterns of *L. monocytogenes* wild-type and *fri* mutant strains grown in BHI at 37 °C were analyzed by 2-DE. Forty two protein spots were differentially expressed by the wild-type and *fri*

mutant strains (Fig. 4), and 26 spots corresponding to 15 different proteins could be identified (Table 1). Among these, 4 proteins were produced at lower level and 11 proteins to higher level in the *fri* mutant compared to the wild-type strain. The different spots corresponding to Fri [15], were not detected in the mutant strain. As already observed, LLO migrated as several spots [15]. Interestingly, one of the 3 spots of LLO, was produced to lower level in the mutant (spot 4a, Fig. 4). In order to investigate whether the different spots of LLO could correspond to various levels of phosphorylation, crude protein extracts of the wild-type and *fri* mutant strains were submitted to

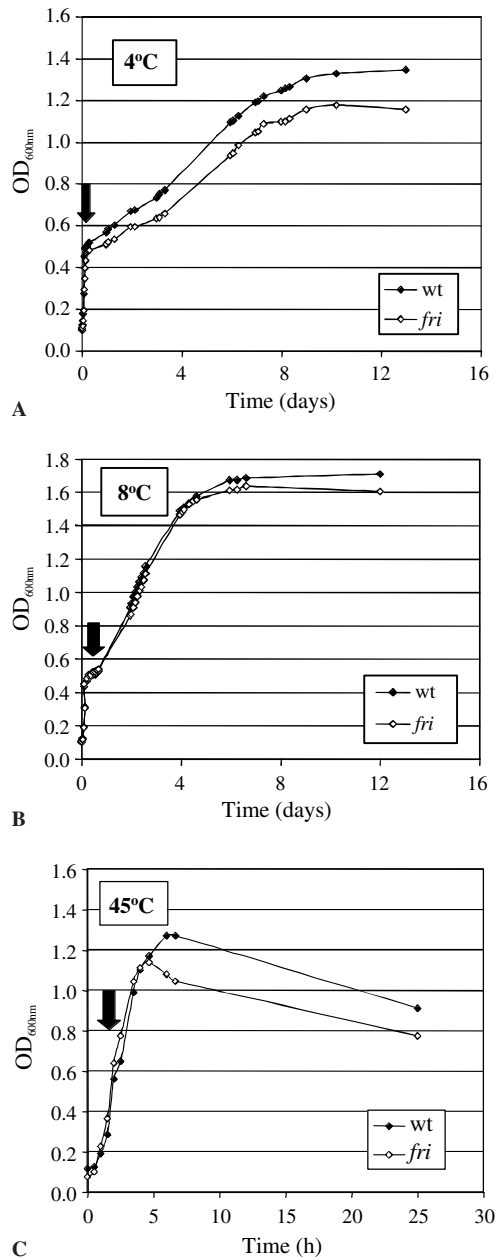


Fig. 2. Effect of the *fri* deletion on *Listeria monocytogenes* growth after a thermal shock. Growth of the wild-type EGDe and *fri* mutant strains was analyzed in BHI by measuring bacterial density. Bacteria were grown at 37 °C until OD<sub>600nm</sub> = 0.3–0.5 and then grown at 4 °C (A), 8 °C (B) or 45 °C (C). The arrow indicates the time of the temperature shift.

SDS–PAGE. Phosphorylation of LLO could not be detected by Pro-Q gel staining (data not shown).

Out of the 13 remaining differentially expressed proteins, 6 were related to stress response. Two of them were expressed to a lower level by the mutant, i.e., the catabolite control protein A CcpA (Lmo1599) and a homolog of the stress protein OsmC (Lmo2199). The four proteins related to stress overexpressed in the mutant were the sigma B regulator RsbW (Lmo0894), the

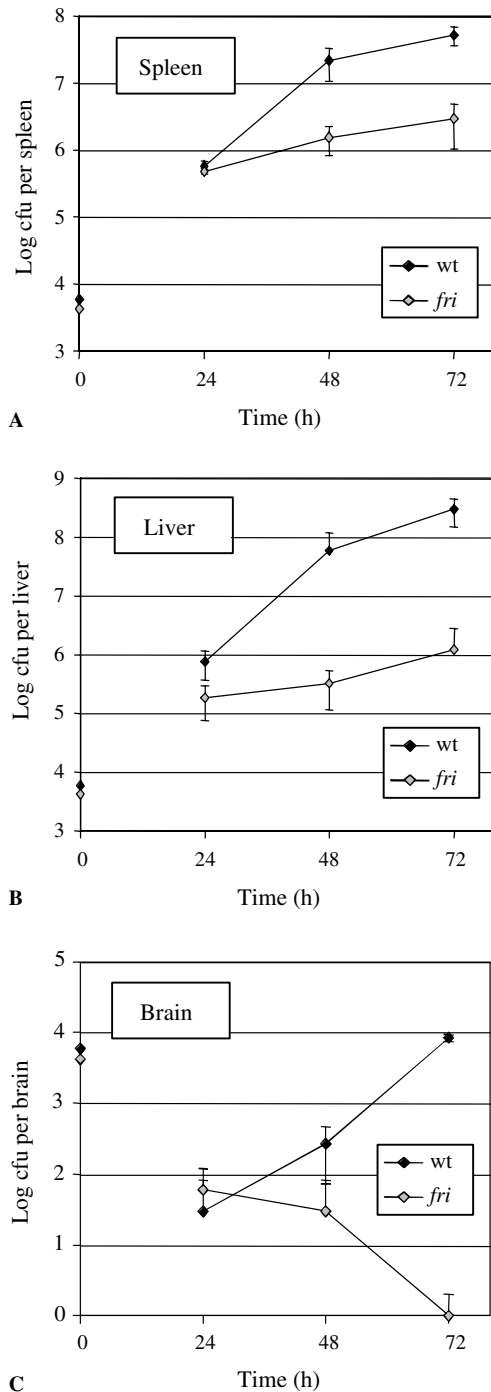


Fig. 3. Effect of the *fri* deletion on survival of *Listeria monocytogenes* in BALB/c mice. Multiplication of the wild-type EGDe and *fri* mutant strains in the spleen (A), liver (B) and brain (C) of BALB/c mice infected intravenously with  $5 \times 10^3$  bacteria. For each strain, growth was determined in groups of four mice at 24, 48 and 72 h after infection.

pyridoxine biosynthesis protein Lmo2101, the stress protein GSP26 homolog Lmo2748 and the mannose specific phosphotransferase system component IIB protein Lmo0783. The 7 other proteins produced at higher levels in the mutant were the deoxyribose-phosphate

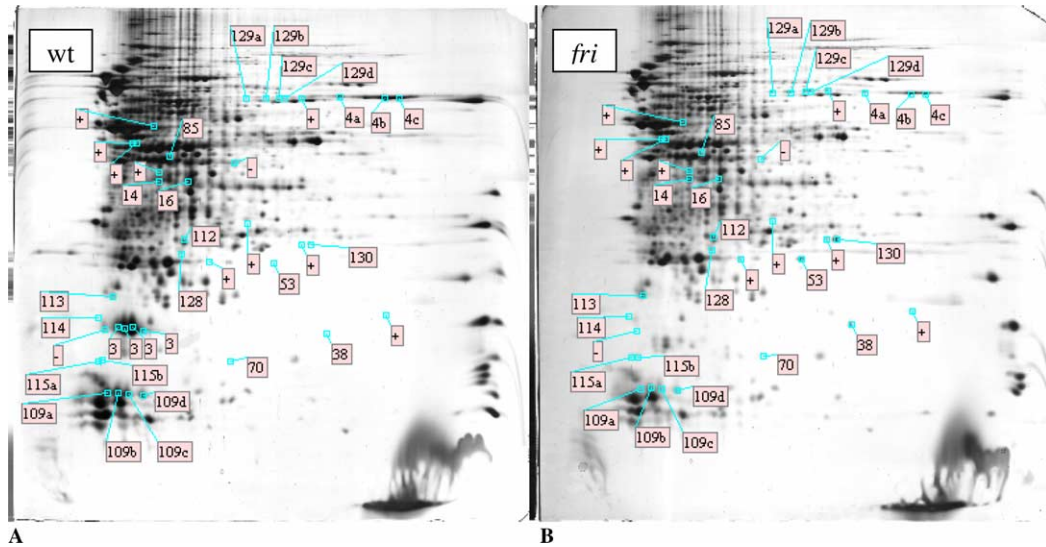


Fig. 4. 2-DE maps of the soluble protein fractions of *Listeria monocytogenes* wild-type EGDe (A) and *fri* mutant (B) strains. Proteins were separated in pH 3–10 nonlinear immobilized pH gradient gel. Numbers point out identified proteins listed in Table 1.

Table 1  
Proteins differentially expressed in the *Listeria monocytogenes* EGDe and *fri* mutant strains

Spot	Gene name	List number	Description	Mr (kDa)		pI		Ratio
				Obs	Calc	Obs	Calc	
<i>Proteins underexpressed in the fri mutant strain</i>								
3	<i>fri</i>	Lmo0943	Non-heme iron-binding ferritin	19.6	18.0	4.95 4.95 4.90 4.75	4.57	Absent
4a	<i>hly</i>	Lmo0202	Listeriolysin O	61.2	58.7	6.39	8.23	0.5
85	<i>ccpA</i>	Lmo1599	Catabolite control protein A	42.5	36.9	5.20	5.05	0.3
109	<i>lmo2199</i>	Lmo2199	Similar to OsmC/Ohr family proteins	11.0 11.1 11.4 11.1	14.5	5.00 4.85 4.75 4.65	4.52	0.1 0.6 0.5 0.7
<i>Proteins overexpressed in the fri mutant strain</i>								
14	<i>lmo2101</i>	Lmo2101	Pyridoxine synthesis protein Pdx1	37.8	31.7	5.10	5.12	3.0
38	<i>lmo0783</i>	Lmo0783	Mannose-specific PTS component IIB	18.6	17.9	6.30	6.55	56.0
53	<i>lmo1830</i>	Lmo1830	Conserved hypothetical dehydrogenase	26.8	20.9	6.05	5.97	Present
70	<i>lmo1468</i>	Lmo1468	Protein of unknown function	14.8	16.5	5.80	6.04	5.7
112	<i>lmo0794</i>	Lmo0794	Similar to <i>Bacillus subtilis</i> YwnB protein	29.0	23.4	5.30	5.11	4.7
113	<i>lmo0796</i>	Lmo0796	Ycel-like family protein	22.3	19.3	4.70	4.45	4.6
114	<i>rsbW</i>	Lmo0894	Anti-sigma B factor RsbW	20.6	17.5	4.60	4.22	56.0
115	<i>lmo2748</i>	Lmo2748	Similar to <i>Bacillus subtilis</i> GSP26 protein	15.7 15.7	15.7	4.60 4.50	4.31	Present
128	<i>dra</i>	Lmo1995	Deoxyribose-phosphate aldolase	27.6	23.5	5.35	5.02	4.8
129	<i>lmo0913</i>	Lmo0913	Similar to succinate semialdehyde dehydrogenase	60.6	53.2	5.90 6.00 6.05 6.10	5.87	9.0 11.0 5.0 3.7
130	<i>lmo2391</i>	Lmo2391	Similar to <i>Bacillus subtilis</i> YhfK protein	28.5	22.7	6.20	6.30	17.0

aldolase Dra (Lmo1995), the succinate semialdehyde dehydrogenase Lmo0913, the short chain alcohol dehydrogenase Lmo1830, the putative NAD-dependent

dehydrogenase Lmo0794, the putative reductase Lmo2391 and the proteins of unknown function Lmo1468 and Lmo0796.

#### 4. Discussion

In this study, analysis of a *fri* deletion mutant provided evidence that *L. monocytogenes* ferritin promotes adaptation to nutritional and thermal shifts, protects from oxidative damage, and contributes to virulence.

Although the molecular mechanism of iron uptake is largely unknown in *L. monocytogenes*, it has been proposed that bacteria could bind heterologous siderophores, transferrin, ferric citrate or catecholamines, and could then use ferrous iron released from these iron sources after a step of reduction by a ferric reductase [17–20]. Genome analysis revealed that *L. monocytogenes* EGDe could encode several proteins involved in iron metabolism including putative ferrichrome-binding proteins, ferrichrome transport permeases, a ferrous iron transport protein, a ferric citrate transport permease and a protein of the dodecameric ferritin family [3]. This latter protein is encoded by *fri*.

In presence of iron in minimal medium, our *fri* mutant reached a lower bacterial density than the wild-type EGDe strain. It is most likely that iron sequestration in Fri prevents iron toxicity. In *L. innocua* ferritin, reactive ferrous iron is first oxidized to its ferric form and mineralized as insoluble ferrihydrite [21]. Ferroxidation in *L. innocua* dodecameric ferritin has recently been demonstrated to play a critical role in decreasing the production of hydroxyl radical OH<sup>•</sup> by the Fenton reaction and in DNA protection from oxidative damage [21,22]. Preculture of bacteria with iron allowed improved growth of the wild-type EGDe strain in conditions of iron restriction, while it had only minor effects on the growth of the *fri* mutant. It has been shown that iron storage in *L. innocua* dodecameric ferritin increases when the iron concentration of the medium increases [8]. Together, these observations suggest that iron stored in the ferritin can be released to support growth under condition of low iron availability.

The hypersensitivity of the *L. monocytogenes* *fri* mutant to oxidative stress and to a lower extent to thermal shock provided other evidence of the protective role of ferritin against stresses. All three types of bacterial ferritins have been shown to protect against oxidative stress in various species [2]. In dodecameric ferritins, protection against reactive oxygen species mediated toxicity is provided by both DNA-binding ferritins [23–26], and those which do not bind DNA [27–31]. Dodecameric ferritins detoxify oxidative agents chemically through their iron-binding property and ferroxidase activity and, for those that can bind DNA like Dps, physically by forming nonspecific stable nucleoprotein complexes with the chromosome called biocrystals [23,32]. *L. monocytogenes* ferritin has not been reported to bind DNA and *L. innocua* ferritin, its closest orthologue, failed to bind DNA [7]. Therefore, one can hypothesize that the protective role of *L. monocytogenes*

ferritin against multiple stresses is mainly related to its iron-binding activity.

Another role of Fri could be indirect regulation of expression and/or stability of protein involved in stress adaptation. When the proteomes of the wild-type and *fri* mutant strains were compared, 15 proteins were found to be expressed differently and identified. Interestingly, several of the differentially expressed proteins are involved in stress response or virulence. CcpA, a central mediator of catabolite repression is under-expressed in the *fri* mutant. CcpA is overproduced after osmotic stress in *L. monocytogenes* [33] and cold shock in *Lactococcus lactis* [34]. Disruption of *L. monocytogenes* *ccpA* results in pleiotropic effects on growth but not in virulence [35]. Thus, the role of ferritin in virulence is probably not mediated by CcpA. The level of Lmo2199, a homolog of OsmC/Ohr family proteins, was decreased in the *fri* mutant. OsmC proteins are induced by ethanol and osmotic stresses [36]. Ohr proteins are induced by organic peroxide and are involved in organic hydroperoxide detoxification [36]. Thus, it is possible that Lmo2199 is involved in detoxification. A protein similar to component IIB of a mannose-specific phosphoenolpyruvate sugar phosphotransferase system (PTS), Lmo0783, was present in higher amounts in the *fri* mutant. In *L. monocytogenes*, upregulation of a mannose specific PTS system by osmotic stress has been reported [33]. The pyridoxine biosynthesis protein Pdx1 (Lmo2101) and a homolog of the general stress protein of *Bacillus subtilis* GSP26 which contains a pyridoxamine 5'-phosphate oxidase domain (Lmo2748), were induced in the *fri* mutant. Pyridoxine and vitamin B6 compounds are reactive oxygen species quenchers protecting against oxidative stress [37,38]. Overexpression of these proteins in the *fri* mutant reinforces the hypothesis of their role against oxidative stress. Strikingly, RsbW, an anti-sigma B factor with a kinase activity whose substrate is the anti-anti-sigma factor RsbV, was undetectable in the wild-type strain but present in the *fri* mutant. Sigma B controls the general stress response. In *B. subtilis*, RsbW binds and inactivates sigma B during exponential growth. In stress conditions, RsbV is dephosphorylated and forms a complex with RsbW thus releasing sigma B [39,40]. Interestingly a sigma B recognition sequence upstream the distal transcription start point of the ferritin gene had been identified in *L. monocytogenes* [8] and *fri* has recently been shown to be regulated by both sigma B and the hydrogen peroxide regulator PerR [31]. At present, how RsbW expression is modulated by Fri remains elusive.

LLO migrated as several spots in 2-DE, as previously reported [15]. One of the 3 spots corresponding to LLO was significantly under-expressed in the *fri* mutant, suggesting that ferritin plays a direct or indirect role on the production and/or stability of LLO. The decrease in LLO could contribute to the attenuation of the *fri* mutant strain in the mouse model of infection.

In conclusion, few studies were conducted to investigate the role of ferritins from pathogenic bacteria. *Helicobacter pylori* canonical ferritin has been reported to be necessary for colonization of gerbil stomach [41]. Dodecameric ferritin has been shown to contribute to survival of *Porphyromonas gingivalis* in HUVEC cells [26]. *Salmonella enterica* serovar Typhimurium Dps has been shown to contribute to survival in murine peritoneal macrophages and to virulence in mice [42]. It has been recently shown that *L. monocytogenes* Fri promotes intracellular multiplication in J774 macrophage-like cells [31]. The important role of *L. monocytogenes* ferritin in virulence most probably comes from a combination of its properties of iron storage, stress protection and regulation of virulence factors.

## Acknowledgments

This work was supported by the Pasteur Institute (GPH9), INSERM, the French Ministry of Research (Programme de Microbiologie Fondamentale et Appliquée, and the French Ministry of Agriculture (Programme AQS A03/02)). Cristel Archambaud and Emilie Dumas received a PhD grant from the French Ministry of Research. Pascale Cossart is an international scholar from the Howard Hughes Medical Institute.

## References

- Andrews, S.C., Robinson, A.K. and Rodriguez-Quinones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol. Rev.* 27, 215–237.
- Smith, J.L. (2004) The physiological role of ferritin-like compounds in bacteria. *Crit. Rev. Microbiol.* 30, 173–185.
- Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K.D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L.M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J.M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J.C., Purcell, R., Rimmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehland, J. and Cossart, P. (2001) Comparative genomics of *Listeria* species. *Science* 294, 849–852.
- Phan-Thanh, L. and Gormon, T. (1995) Analysis of heat and cold shock proteins in *Listeria* by two-dimensional electrophoresis. *Electrophoresis* 16, 444–450.
- Phan-Thanh, L. and Gormon, T. (1997) Stress proteins in *Listeria monocytogenes*. *Electrophoresis* 18, 1464–1471.
- Hebraud, M. and Guzzo, J. (2000) The main cold shock protein of *Listeria monocytogenes* belongs to the family of ferritin-like proteins. *FEMS Microbiol. Lett.* 190, 29–34.
- Bozzi, M., Mignogna, G., Stefanini, S., Barra, D., Longhi, C., Valenti, P. and Chiancone, E. (1997) A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in *Listeria innocua*. *J. Biol. Chem.* 272, 3259–3265.
- Polidoro, M., De Biase, D., Montagnini, B., Guarrera, L., Cavallo, S., Valenti, P., Stefanini, S. and Chiancone, E. (2002) The expression of the dodecameric ferritin in *Listeria* spp. is induced by iron limitation and stationary growth phase. *Gene* 296, 121–128.
- Berche, P., Gaillard, J.L. and Richard, S. (1988) Invasiveness and intracellular growth of *Listeria monocytogenes*. *Infection* 16 (Suppl. 2), S145–S148.
- Conte, M.P., Longhi, C., Petrone, G., Polidoro, M., Valenti, P. and Seganti, L. (2000) Modulation of *actA* gene expression in *Listeria monocytogenes* by iron. *J. Med. Microbiol.* 49, 681–683.
- Phan-Thanh, L. and Gormon, T. (1997) A chemically defined minimal medium for the optimal culture of *Listeria*. *Int. J. Food Microbiol.* 35, 91–95.
- Archambaud, C., Gouin, E., Pizarro-Cerda, J., Cossart, P. and Dussurget, O. (2005) Translation elongation factor EF-Tu is a target for Stp, a serine-threonine phosphatase involved in virulence of *Listeria monocytogenes*. *Mol. Microbiol.* 56, 383–396.
- Arnaud, M., Chastanet, A. and Debarbouille, M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl. Environ. Microbiol.* 70, 6887–6891.
- Sullivan, M.A., Yasbin, R.E. and Young, F.E. (1984) New shuttle vectors for *Bacillus subtilis* and *Escherichia coli* which allow rapid detection of inserted fragments. *Gene* 29, 21–26.
- Folio, P., Chavant, P., Chafsey, I., Belkorchia, A., Chambon, C. and Hebraud, M. (2004) Two-dimensional electrophoresis database of *Listeria monocytogenes* EGDe proteome and proteomic analysis of mid-log and stationary growth phase cells. *Proteomics* 4, 3187–3201.
- Farr, S.B., Natvig, D.O. and Kogoma, T. (1985) Toxicity and mutagenicity of plumbagin and the induction of a possible new DNA repair pathway in *Escherichia coli*. *J. Bacteriol.* 164, 1309–1316.
- Coulanges, V., Andre, P., Ziegler, O., Buchheit, L. and Vidon, D.J. (1997) Utilization of iron-catecholamine complexes involving ferric reductase activity in *Listeria monocytogenes*. *Infect. Immun.* 65, 2778–2785.
- Hartford, T., O'Brien, S., Andrew, P.W., Jones, D. and Roberts, I.S. (1993) Utilization of transferrin-bound iron by *Listeria monocytogenes*. *FEMS Microbiol. Lett.* 108, 311–318.
- Deneer, H.G. and Boychuk, I. (1993) Reduction of ferric iron by *Listeria monocytogenes* and other species of *Listeria*. *Can. J. Microbiol.* 39, 480–485.
- Deneer, H.G., Healey, V. and Boychuk, I. (1995) Reduction of exogenous ferric iron by a surface-associated ferric reductase of *Listeria* spp. *Microbiology* 141, 1985–1992.
- Su, M., Cavallo, S., Stefanini, S., Chiancone, E. and Chasteen, N.D. (2005) The so-called *Listeria innocua* ferritin is a Dps protein. Iron incorporation, detoxification, and DNA protection properties. *Biochemistry* 44, 5572–5578.
- Ilari, A., Latella, M.C., Ceci, P., Ribacchi, F., Su, M., Giangiacomo, L., Stefanini, S., Chasteen, N.D. and Chiancone, E. (2005) The unusual intersubunit ferroxidase center of *Listeria innocua* Dps is required for hydrogen peroxide detoxification but not for iron uptake. A study with site-specific mutants. *Biochemistry* 44, 5579–5587.
- Ilari, A., Ceci, P., Ferrari, D., Rossi, G.L. and Chiancone, E. (2002) Iron incorporation into *Escherichia coli* Dps gives rise to a ferritin-like microcrystalline core. *J. Biol. Chem.* 277, 37619–37623.
- Cooksley, C., Jenks, P.J., Green, A., Cockayne, A., Logan, R.P. and Hardie, K.R. (2003) NapA protects *Helicobacter pylori* from oxidative stress damage, and its production is influenced by the ferric uptake regulator. *J. Med. Microbiol.* 52, 461–469.

- [25] Gupta, S. and Chatterji, D. (2003) Bimodal protection of DNA by *Mycobacterium smegmatis* DNA-binding protein from stationary phase cells. *J. Biol. Chem.* 278, 5235–5241.
- [26] Ueshima, J., Shoji, M., Ratnayake, D.B., Abe, K., Yoshida, S., Yamamoto, K. and Nakayama, K. (2003) Purification, gene cloning, gene expression, and mutants of Dps from the obligate anaerobe *Porphyromonas gingivalis*. *Infect. Immun.* 71, 1170–1178.
- [27] Ishikawa, T., Mizunoe, Y., Kawabata, S., Takade, A., Harada, M., Wai, S.N. and Yoshida, S. (2003) The iron-binding protein Dps confers hydrogen peroxide stress resistance to *Campylobacter jejuni*. *J. Bacteriol.* 185, 1010–1017.
- [28] Ceci, P., Ilari, A., Falvo, E. and Chiancone, E. (2003) The Dps protein of *Agrobacterium tumefaciens* does not bind to DNA but protects it toward oxidative cleavage: X-ray crystal structure, iron binding, and hydroxyl-radical scavenging properties. *J. Biol. Chem.* 278, 20319–20326.
- [29] Yamamoto, Y., Higuchi, M., Poole, L.B. and Kamio, Y. (2000) Role of the *dpr* product in oxygen tolerance in *Streptococcus mutans*. *J. Bacteriol.* 182, 3740–3747.
- [30] Pulliainen, A.T., Haataja, S., Kahkonen, S. and Finne, J. (2003) Molecular basis of H<sub>2</sub>O<sub>2</sub> resistance mediated by streptococcal Dpr. Demonstration of the functional involvement of the putative ferroxidase center by site-directed mutagenesis in *Streptococcus suis*. *J. Biol. Chem.* 278, 7996–8005.
- [31] Olsen, K.N., Larsen, M.H., Gahan, C.G., Kallipolitis, B., Wolf, X.A., Rea, R., Hill, C. and Ingmer, H. (2005) The Dps-like protein Fri of *Listeria monocytogenes* promotes stress tolerance and intracellular multiplication in macrophage-like cells. *Microbiology* 151, 925–933.
- [32] Frenkiel-Krispin, D., Levin-Zaidman, S., Shimoni, E., Wolf, S.G., Wachtel, E.J., Arad, T., Finkel, S.E., Kolter, R. and Minsky, A. (2001) Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection. *EMBO J.* 20, 1184–1191.
- [33] Duche, O., Tremoulet, F., Glaser, P. and Labadie, J. (2002) Salt stress proteins induced in *Listeria monocytogenes*. *Appl. Environ. Microbiol.* 68, 1491–1498.
- [34] Wouters, J.A., Kamphuis, H.H., Hugenholtz, J., Kuipers, O.P., de Vos, W.M. and Abee, T. (2000) Changes in glycolytic activity of *Lactococcus lactis* induced by low temperature. *Appl. Environ. Microbiol.* 66, 3686–3691.
- [35] Behari, J. and Youngman, P. (1998) A homolog of CcpA mediates catabolite control in *Listeria monocytogenes* but not carbon source regulation of virulence genes. *J. Bacteriol.* 180, 6316–6324.
- [36] Atichartpongkul, S., Loprasert, S., Vattanaviboon, P., Whang-suk, W., Helmann, J.D. and Mongkolsuk, S. (2001) Bacterial Ohr and OsmC paralogues define two protein families with distinct functions and patterns of expression. *Microbiology* 147, 1775–1782.
- [37] Bilski, P., Li, M.Y., Ehrenshaft, M., Daub, M.E. and Chignell, C.F. (2000) Vitamin B6 (pyridoxine) and its derivatives are efficient singlet oxygen quenchers and potential fungal antioxidants. *Photochem. Photobiol.* 71, 129–134.
- [38] Chumnantana, R., Yokochi, N. and Yagi, T. (2005) Vitamin B6 compounds prevent the death of yeast cells due to menadione, a reactive oxygen generator. *Biochim. Biophys. Acta* 1722, 84–91.
- [39] Dufour, A., Voelker, U., Voelker, A. and Haldenwang, W.G. (1996) Relative levels and fractionation properties of *Bacillus subtilis* sigma B and its regulators during balanced growth and stress. *J. Bacteriol.* 178, 3701–3709.
- [40] Delumeau, O., Lewis, R.J. and Yudkin, M.D. (2002) Protein-protein interactions that regulate the energy stress activation of sigma B in *Bacillus subtilis*. *J. Bacteriol.* 184, 5583–5589.
- [41] Waidner, B., Greiner, S., Odenbreit, S., Kavermann, H., Velayudhan, J., Stahler, F., Guhl, J., Bisse, E., van Vliet, A.H., Andrews, S.C., Kusters, J.G., Kelly, D.J., Haas, R., Kist, M. and Bereswill, S. (2002) Essential role of ferritin Pfr in *Helicobacter pylori* iron metabolism and gastric colonization. *Infect. Immun.* 70, 3923–3929.
- [42] Halsey, T.A., Vazquez-Torres, A., Gravidahl, D.J., Fang, F.C. and Libby, S.J. (2004) The ferritin-like Dps protein is required for *Salmonella enterica* serovar Typhimurium oxidative stress resistance and virulence. *Infect. Immun.* 72, 1155–1158.

## 4 - DISCUSSION

Les ferritines sont connues pour être des protéines de stockage du fer. La ferritine de *Listeria* est différente des ferritines classiques puisqu'elle appartient à une famille de protéines appelées DPS (DNA-binding proteins), exprimées dans des conditions de stress oxydatif ou nutritionnel. Cette dénomination vient du fait que la première protéine DPS, découverte chez *E. coli* (Almiron *et al.*, 1992), se liait à l'ADN de manière non spécifique afin de le protéger du stress oxydatif. Contrairement à la protéine Dps d'*E. coli*, la protéine Dps de *L. innocua* ne peut pas lier l'ADN (Bozzi *et al.*, 1997). Par contre, il a été montré récemment que toutes les protéines Dps étudiées pouvaient lier le fer (Yamamoto *et al.*, 2002 ; Ishikawa *et al.*, 2003 ; Cooksley *et al.*, 2003). Les travaux que nous avons réalisés en collaboration avec l'Institut Pasteur de Paris ont montré que le mutant  $\Delta fri$  était hypersensible au stress oxydatif produit par la plumbagine, molécule capable de générer des radicaux superoxyde  $O_2^-$ , qui entraînent la formation de radical hydroxyle ( $HO^\cdot$ ) et de peroxyde d'hydrogène ( $H_2O_2$ ). Par contre, le mutant et la souche sauvage montraient une sensibilité similaire au S-nitroso-N-acétyl-pénicillamine, un agent entraînant la formation de monoxyde d'azote ( $NO^\cdot$ ). Il semble donc que la ferritine ait un effet protecteur contre l'effet toxique du peroxyde d'hydrogène, mais pas contre le monoxyde d'azote. Les études de Mohamed *et al.* (2006) et Olsen *et al.* (2005) ont également montré que la protéine Fri de *L. monocytogenes* permettait de résister au stress oxydatif. Il a été démontré que la protéine Dps chez *E. coli*, pour stocker le fer, utilisait l' $H_2O_2$  comme oxydant, l' $O_2$  généralement utilisé par les ferritines, n'étant qu'une alternative (Zhao *et al.*, 2002). L'effet protecteur de la ferritine de *L. innocua* contre la combinaison toxique  $Fe^{2+} + H_2O_2$  a été montré par Su *et al.* (2005). La séquence de Fri de *L. monocytogenes* ayant 98% d'identité avec la protéine Fri de *L. innocua* (Bozzi *et al.*, 1997), ce qui laisse supposer un rôle équivalent à celui décrit chez cette dernière bactérie, c'est-à-dire un rôle de protection basé sur la séquestration du fer. Cependant, la séquestration du fer par la protéine Fri de *L. monocytogenes* n'a pas été prouvé expérimentalement et pour affirmer ce rôle, des expériences supplémentaires seraient nécessaires. Les expériences de croissance de la souche sauvage et du mutant en présence de différentes concentrations en fer après préculture avec ou sans fer tendent à montrer que le fer peut effectivement être stocké par la ferritine et libéré en condition de carence en fer, ce qui permet d'améliorer la croissance.



De plus, la phase de latence plus longue du le mutant *fri* par rapport à la souche sauvage lors de la croissance en milieu minimum, suggère un rôle de Fri dans l'adaptation à un stress nutritionnel, puisque cet effet n'était pas observé lors d'une culture en milieu riche (BHI). Après des chocs thermiques de 37°C à 4°C ou de 37°C à 45°C, la souche mutée atteignait une densité optique maximale inférieure à celle de la souche sauvage. Plusieurs études ont montré une sur-expression de la ferritine aussi bien au niveau transcriptionnel que traductionnel après un choc thermique ou une croissance à faible ou haute température (Phan-Thanh & Gormon, 1997 ; Hébraud & Guzzo, 2000 ; Mohamed *et al.*, 2006). Il semble donc que Fri ait un rôle dans l'adaptation aux stress thermiques, mais son implication n'est pas clairement définie. Il a également été montré que cette protéine était induite lors d'autres stress comme un stress éthanol (Phan-Thanh & Gormon, 1997) ou lors de la transition en phase stationnaire (Polidoro *et al.*, 2002). Olsen et al (2005) ont montré que l'expression de Fri était *sigB* dépendante. Il est connu que SigB est impliquée dans la résistance à de nombreux stress, dont les chocs thermiques et le stress éthanol (van Schaik & Abee, 2005), ce qui expliquerait la sur-expression de Fri dans ces conditions de stress. La comparaison du protéome de la souche sauvage et du mutant a montré que certaines des protéines présentant des différences d'abondance entre les deux souches étaient des protéines de stress et de virulence. La ferritine pourrait donc jouer un rôle direct ou indirect dans la régulation de l'expression et/ou la stabilité de protéines impliquées dans l'adaptation au stress. Les protéomes des deux souches ont été étudiés en BHI à 37°C. Il pourrait être intéressant d'étudier le profil protéique de ces 2 souches suite à divers stress comme des stress thermiques ou un stress éthanol. Cette étude permettrait de déterminer les protéines surexprimées en condition de stress chez la souche sauvage par rapport à la souche mutée, afin de mettre en évidence le rôle régulateur de Fri en condition de stress.

L'étude protéomique a permis d'observer que la ferritine était présente sous plusieurs spots. Des études réalisées précédemment au laboratoire avaient montré que l'abondance de ces différents spots variait en fonction de la phase de croissance de la bactérie (Folio *et al.*, 2004). Ces spots ayant une masse identique mais des *pI* différents, il est probable qu'ils soient dûs à des modification post-traductionnelles de la ferritine et plus spécifiquement à différents niveaux de phosphorylations. Il serait intéressant de caractériser la nature et les sites de modification post-traductionnelles. Plusieurs techniques peuvent être mises en œuvre pour révéler la présence de protéines phosphorylées : (i) leur marquage radioactif *in vivo* en cultivant les bactéries en présence de <sup>32</sup>P ou <sup>33</sup>P puis leur séparation par EBD et leur révélation par autoradiographie, ou (ii) leur révélation directe après EBD par une coloration

spécifique (kit Pro-Q Diamond, Molecular Probes) ou indirecte après Western blot avec des anticorps spécifiques aux phosphosérines, aux phosphotyrosines ou aux phosphothréonines. Des techniques de spectrométrie de masse (MS/MS) peuvent également être utilisées pour mettre en évidence et localiser les sites de phosphorylations. Chacune de ces méthodes présentent des avantages et des inconvénients (Reinders & Sickmann, 2005). L'abondance de ces différents spots variant en fonction de la phase de croissance de la bactérie (Folio *et al.*, 2004), il est donc envisageable que des signaux environnementaux entraînent des cascades de phosphorylations afin de réguler la protéine Fri. En effet, il est connu que les modifications post traductionnelles peuvent avoir des conséquences notamment sur le repliement des protéines, leur stabilité, leur activité et donc sur la fonction des protéines. De plus, de nombreuses études ont montré l'importance de la phosphorylation de certaines protéines chez *L. monocytogenes* (Archambaud *et al.*, 2005, Deutscher *et al.*, 2005, Herro *et al.*, 2005, Archambaud *et al.*, 2006, Schnupf *et al.*, 2006). Il serait donc intéressant d'étudier la variation de ces formes en fonction de différentes conditions environnementales (avec / sans fer, stress thermiques...) mais également d'étudier le rôle de ces différentes formes.

Comme pour d'autres bactéries (Ueshima *et al.*, 2003 ; Halsey *et al.*, 2004), cette étude a permis de montrer une implication de la ferritine dans la virulence de la bactérie. En effet, sa capacité à stocker le fer lui permet de survivre dans des environnements carencés en fer (cellules hôtes, sang, lymphocyte) ou en présence de peroxyde d'hydrogène qui est toxique pour la bactérie. De plus, la ferritine semble avoir une incidence sur la listériolysine puisque l'un des 3 spots protéiques de cette protéine majeure de virulence est sous exprimée chez le mutant.

Le rôle important de Fri dans la virulence de *L. monocytogenes* vient donc probablement de la combinaison de ses propriétés de stockage du fer, de protection contre de nombreux stress et de son influence directe ou indirecte sur la régulation de facteurs de virulence.



## Partie 2

---

Etude de la biodiversité de *L. monocytogenes* par une approche  
protéomique  
(Articles 2-3-4)

---

## 1 - PROBLEMATIQUE

Les efforts des industriels de la filière alimentaire et la mise en place d'un système de surveillance efficace en France ont permis de diminuer de manière significative les cas de listériose. Cependant, sur la période 1999 à 2003, 200 à 270 cas par an ont encore été signalés en France, avec un taux de mortalité de 20 à 30%. De plus, alors que les cas de listériose ont diminué en France ces dernières années, il est important de noter que cette maladie a progressé dans certains pays comme l'Allemagne avec 510 cas de listérioses recensés en 2005 contre seulement 217 en 2001. Outre l'importance de cette maladie en terme de santé publique, les conséquences économiques liées à une épidémie de listériose peuvent être dramatiques pour une industrie agroalimentaire. En effet, le retrait et la destruction du produit incriminé entraînent un coût financier très conséquent, aggravé par la détérioration de l'image de marque du produit auprès du consommateur, qui peut être fatale à l'entreprise.

Dans ce contexte, il est important de mieux caractériser le risque lié à *L. monocytogenes*. Grâce à un effort de recherche soutenu, les principaux mécanismes moléculaires et cellulaires impliqués dans l'infection par cette bactérie ont été élucidés et un certain nombre de facteurs de virulence ont été découverts. Cependant, des niveaux de virulence différents ont été mis en évidence entre les souches et 96% des cas de listériose ne sont dus qu'à 3 des 13 sérovars de *L. monocytogenes*, alors que toutes les souches possèdent l'ensemble des gènes de virulence décrits à ce jour. La présence seule de ces gènes de virulence n'est donc pas suffisante pour rendre compte de la virulence. C'est pourquoi il est important de mieux caractériser le risque par l'étude d'un ensemble de souches représentatif des divers habitats de *L. monocytogenes*. Les travaux effectués au cours de ma thèse se sont insérés dans un programme de recherche, financé par la DGAL, intitulé «Approches génomiques et post-génomiques pour le typage des souches de *Listeria monocytogenes* et la prédiction du risque». Les objectifs visés par ce programme étaient la caractérisation d'un panel important de souches de *L. monocytogenes* d'origine (i) humaine de portage (isolées de fèces d'individus asymptomatiques pour la listériose), (ii) alimentaire (isolées à partir d'aliments ou d'ateliers de fabrication) et (iii) épidémiques. Cette caractérisation a été effectuée par l'analyse du génome, du transcriptome, du protéome et de la virulence par des tests *in vitro* et *in vivo*. L'analyse génomique permet d'avoir une approche globale et une connaissance plus approfondie de la distribution et de la diversité du contenu génétique chez *L. monocytogenes*. Les analyses post-génomiques

permettent d'étudier l'expression des gènes au niveau transcriptionnel et traductionnel et de caractériser les protéines fonctionnelles.

Ce projet vise à identifier des gènes (ou protéines) marqueurs, qui pourraient être utilisés pour mieux évaluer et hiérarchiser les risques sanitaires auxquels s'exposent les consommateurs vis-à-vis de produits alimentaires faiblement contaminés.

## **2 - STRATEGIE ET METHODOLOGIE UTILISEE**

### **2-1. Stratégie générale**

Nos travaux se sont inscrits dans la partie post-génomique de ce projet, et ont porté plus précisément sur l'analyse protéomique comparative de 12 souches de *L. monocytogenes*. Nous avons caractérisé par électrophorèse bidimensionnelle (EBD) deux sous-protéomes, les protéines cytosolubles et les protéines sécrétées (ou sécrétome) de 4 souches isolées de cas cliniques humains, 4 souches de portage asymptomatique chez l'homme et 4 souches isolées d'environnement agro-alimentaire. Ces souches appartenaient aux 3 sérovars majoritairement impliqués dans les cas de listériose humaine (1/2a, 1/2b et 4b).

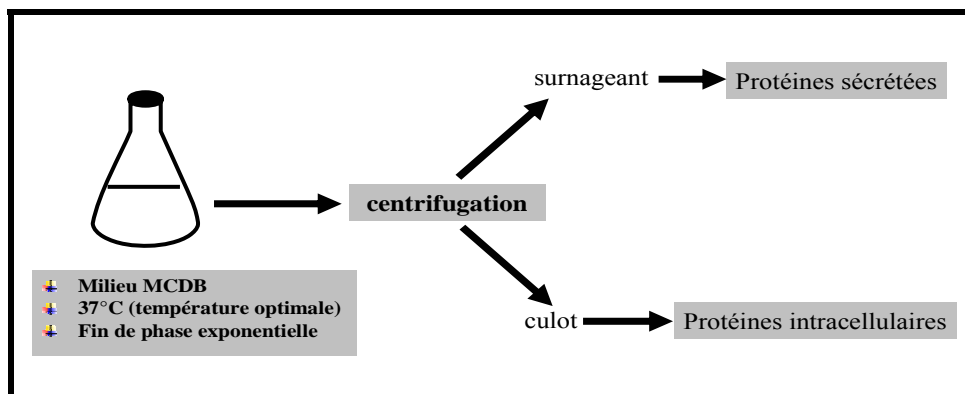
Depuis quelques années, l'équipe Qualité et Sécurité des Aliment (QuaSA) de l'Unité de Recherche de Microbiologie dispose de cartes protéomiques de référence de la souche EGDe de *L. monocytogenes* dans différents gradients de pH, où la nature protéique de plus de 200 spots est identifiée. Ces gels d'EBD ont été établis à partir d'échantillons de protéines intracellulaires obtenus dans diverses conditions de culture (milieu BHI ou MCDB) à 20°C ou 37°C, en phase exponentielle de croissance ou en phase stationnaire (Folio *et al.*, 2004). Ces cartes protéomiques de référence ont été utilisées pour l'identification de protéines d'intérêt dans le cadre de notre étude. Par contre, au début du programme de recherche, aucune carte protéomique des protéines sécrétées n'était disponible au sein de l'équipe QuaSA ou dans la littérature. Or, il est connu que la pathogénicité des bactéries dépend fortement de leur capacité à sécréter des facteurs de virulence soit en surface de la bactérie, soit dans le milieu extracellulaire ou encore qui sont injectés directement dans la cellule hôte. La majorité des facteurs de virulence de *L. monocytogenes* connus à ce jour sont des protéines sécrétées. De plus, de nombreuses protéines sont sécrétées afin d'assurer la colonisation de l'environnement. Il était donc important d'inclure la caractérisation de ce sous-protéome dans notre étude. La première partie de ce travail a donc consisté à mettre au point la technique

d'étude des protéines sécrétées de *L. monocytogenes*. Dans un deuxième temps, le protéome intracellulaire et le sécrétome de 12 souches d'origine, de serovar et de niveaux de virulence différents ont été comparés afin d'étudier la biodiversité au sein de l'espèce *L. monocytogenes* et mieux caractériser le risque sanitaire (article 2). Puis, nous avons étudié plus spécifiquement les données de protéomique comparative obtenues pour les 4 souches de serovar 4b en fonction de leur origine (article 3). Enfin, dans la dernière partie du travail, une analyse bioinformatique a été réalisée afin de caractériser les protéines identifiées dans le milieu extracellulaire des souches de *L. monocytogenes* (article 4).

## **2-2. Analyse de la biodiversité de *L. monocytogenes***

### Préparation des échantillons protéiques

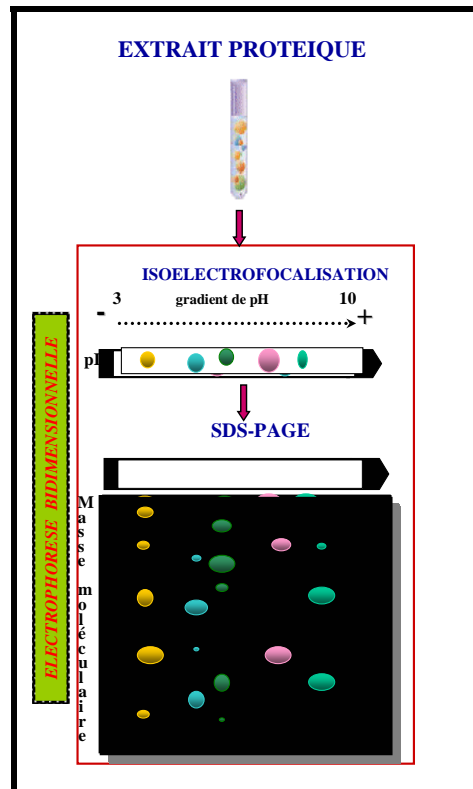
Les cellules de *L. monocytogenes* ont été cultivées jusqu'en fin de phase exponentielle dans un milieu chimiquement défini, le milieu MCDB202, à 37°C (température optimale de croissance). Les cultures ont alors été centrifugées et le surnageant a été précipité afin de récupérer les protéines sécrétées. Les protéines intracellulaires ont été extraites par sonication du culot cellulaire (Figure 19).



**Figure 19** : Méthode d'obtention des protéines sécrétées et intracellulaires.

### Electrophorèse bidimensionnelle (EBD)

L'électrophorèse bidimensionnelle (Figure 2) consiste à séparer les protéines selon leur point isoélectrique ( $pI$ ) dans la première dimension puis selon leur masse moléculaire dans la deuxième dimension. L'EBD a été choisie car elle permet non seulement de séparer plusieurs centaines de protéines en même temps mais aussi de réaliser des études qualitatives et quantitatives comparatives de l'expression des protéines.



**Figure 20 : l'EBD consiste à effectuer dans un premier temps une séparation des protéines en fonction de leur charge (isoélectrofocalisation) dans un gradient de pH pré-établi, puis une deuxième séparation, à 90° par rapport à la première, en fonction de la masse moléculaire.**

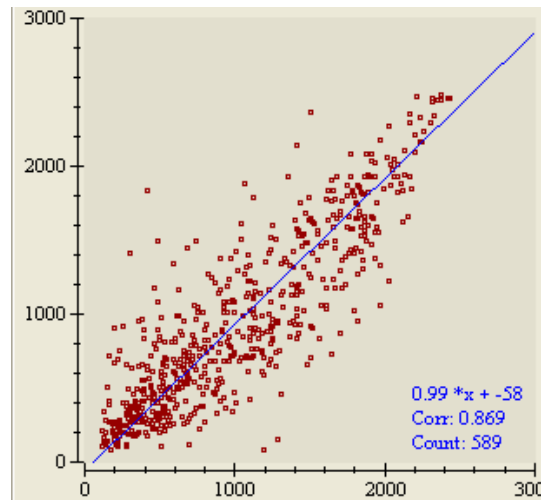
Pour les protéines intracellulaires, un gradient de pH 4-7 a été utilisé lors de l'isoélectrofocalisation car la majorité des protéines ont un  $pI$  compris entre 4 et 7 et cette gamme de pH permet une bonne séparation et résolution des spots protéiques. Pour les protéines sécrétées dans le milieu extracellulaire, un gradient de pH large (3-10 non linéaire) a été utilisé afin de séparer un maximum de protéines. En effet, la proportion de protéines avec un  $pI$  compris entre 4 et 7 est en théorie moins élevée que pour les protéines intracellulaires et de nombreuses protéines sécrétées ont un  $pI$  basique.

### Analyse d'image

Sur le plan expérimental, un minimum de 6 gels d'EBD avec des échantillons protéiques issus de ceux cultures bactériennes ont été réalisés par sous-protéome et pour chacune des 12 souches. Pour chaque souche et chaque sous-protéome, les gels d'EBD ont tout d'abord été comparés entre eux afin de sélectionner les 5 gels les plus reproductibles. Ce choix a été



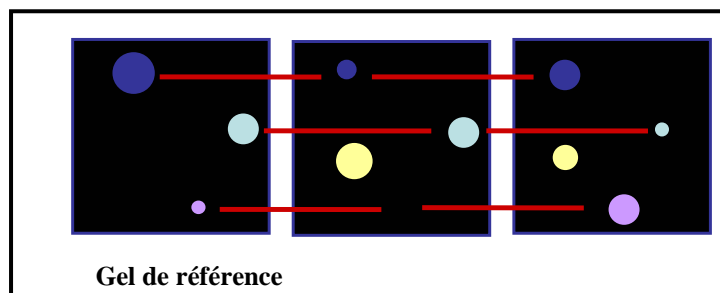
effectué grâce à l’outil scatter-plot (Figure 21), disponible avec le logiciel d’analyse d’image « Image Master 2D Platinum ».



**Figure 21 : scatter-plot**

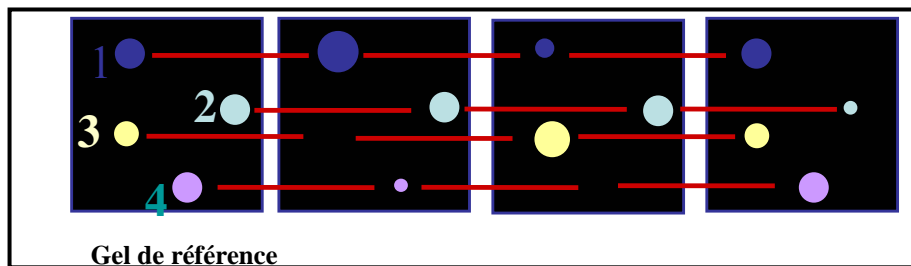
Cet outil permet de connaître la relation entre les valeurs des spots de deux gels en cherchant la dépendance linéaire entre les valeurs des spots d’un gel en comparaison à un autre gel. Un coefficient de corrélation proche de 1 indique un bon ajustement, c’est-à-dire que la valeur du spot d’un gel peut être prédit par la valeur du spot correspondant dans l’autre gel. Nous avons sélectionné uniquement les gels avec une corrélation supérieure à 0.8. Cette opération permet de sélectionner des gels reproductibles au sein d’une même souche et ainsi d’éliminer des gels pouvant présenter d’importantes différences dues aux conditions expérimentales (extraction, EBD ou coloration).

Une fois les 5 gels sélectionnés chez les 12 souches, les spots des 60 gels ont été appariés. L’appariement est une opération clef dans l’analyse d’image de gels d’EBD. Un algorithme compare les gels afin d’apparier tous les spots qui ont migré au même endroit. Cet appariement est réalisé par rapport à un gel de référence et les spots considérés comme identiques dans les différents gels forment ainsi un groupe (Figure 22).



**Figure 22 : Principe de l’appariement des spots protéiques entre les gels.**

Lors de cette étape, le choix du gel de référence est primordial. Dans l'exemple ci-dessus, le gel de référence choisi est celui de gauche. Le spot protéique jaune étant absent de ce gel, il sera écarté de l'analyse. Les douze souches étudiées lors de notre analyse présentaient toutes des profils protéiques différents, et il n'était donc pas possible de choisir l'un de ces gels comme référence sans entraîner inévitablement des biais importants. Aussi, afin de limiter le nombre de spots protéiques non appariés, un gel de référence a été réalisé avec un échantillon protéique provenant d'un mélange en quantités égales des protéines des 12 souches (Figure 5). Ce gel contenait en théorie l'ensemble des spots protéiques contenus dans les 12 souches.



**Figure 23 : Appariement des spots protéiques entre les gels et par rapport à un gel de référence réalisé à partir d'un échantillon protéique provenant d'un mélange en quantités égales des protéines des 12 souches .**

A partir de cet appariement, une matrice de données a été réalisée pour chaque sous-protéome comprenant pour chaque spot apparié le pourcentage de volume du spot par rapport à l'ensemble des spots du gel considéré. Ainsi, les matrices de données pour les protéines sécrétées et pour les protéines intracellulaires contenaient 599 et 600 spots, respectivement.

Il faut noter que les conditions expérimentales peuvent affecter l'expression protéique. De plus, des erreurs peuvent être faites lors de l'étape d'appariement et des mésappariements peuvent même subsister après correction. De ce fait, les matrices de données comportent parfois des valeurs manquantes qui peuvent perturber voire empêcher les analyses statistiques et la classification hiérarchique réalisées par la suite. Afin de résoudre ce problème, nous avons décidé, à partir des deux matrices initiales, de générer deux matrices secondaires ne contenant que des spots protéiques fiables. Un spot protéique était considéré comme fiable quand il était présent ou absent dans 4 des 5 gels d'une même souche. Dans le cas d'un spot protéique qualifié de "présent fiable" mais avec une valeur manquante, cette dernière était remplacée par la moyenne des valeurs existantes pour cette souche et pour ce spot. Si le spot protéique était qualifié d'"absent fiable" mais avec une valeur significative dans l'un des 5 gels de la souche, les valeurs manquantes étaient remplacées par la plus petite valeur de spot

trouvée dans chacun des 4 autres gels considérés. C'est à partir de ces nouvelles matrices de données que les analyses statistiques ont été réalisées.

### Classification hiérarchique et analyse statistique

Nous avons tout d'abord cherché à classer les individus selon leur profil protéique et grouper les spots ayant un profil d'expression similaire afin de mettre en évidence des spots protéiques spécifiques du sérovar, de l'origine ou de la virulence. Dans ce but, une stratégie de classification hiérarchique a été développée. Les données de protéomique ont pour caractéristique de se distribuer sur une large gamme de valeurs. De ce fait, quelques spots protéiques très abondants peuvent peser bien plus fortement sur les résultats de la classification que la majorité des spots protéiques faiblement abondants. Afin de résoudre ce problème, nous avons décidé d'utiliser une approche basée sur le ratio et une transformation logarithmique. Pour cela, la valeur de chaque spot protéique de la matrice était divisée par la moyenne des valeurs existantes pour ce spot protéique. Cette matrice de données sous forme de ratio était ensuite soumise à une transformation logarithmique. Cette approche permet de comparer chaque gel à un gel virtuel moyen. Des exemples de valeur sont présentés dans le tableau ci-dessous.

<b>valeur</b>	<b>signification</b>
<b>0</b>	Pas de différence entre la valeur du spot étudié et la valeur moyenne de ce spot.
<b>- 1</b>	La valeur du spot étudié est deux fois moins importante que la valeur moyenne.
<b>+1</b>	La valeur du spot étudié est deux fois plus importante que la valeur moyenne.

Ainsi les matrices de données étaient soumises à deux étapes de normalisation. La première, le passage en % de volume permet de comparer les gels, et la deuxième permet de comparer les protéines.

Les matrices de données normalisées ont alors été soumises à une classification hiérarchique réalisée avec le logiciel PermutMatrix. Ce logiciel permet une représentation en couleur de la matrice de données, où l'intensité de couleur est proportionnelle à la valeur du spot protéique. Il permet également de visualiser le résultat de la classification avec un dendrogramme des souches et un dendrogramme pour les protéines.

La mise au point de la méthode de classification a fait l'objet d'une publication (Meunier *et al.*, 2007) jointe en annexe.

Afin de trouver des spots protéiques spécifiques pour le sérovar, l'origine ou la virulence, la variance relative estimée de chaque spot pour chacun des facteurs étudiés a été calculée. Plus cette variance relative estimée est élevée, plus le facteur (sérovar, origine, virulence) explique la variation du spot. Une analyse factorielle discriminante a ensuite été réalisée avec les spots les plus significatifs.

Les protéines d'intérêt ont finalement été soumises à identification par spectrométrie de masse MALDI-ToF à partir de leur empreinte peptidique massique.

### **2-3. Analyse des protéines sécrétées par *L. monocytogenes***

L'analyse de la biodiversité de *L. monocytogenes* a permis de cibler de nombreuses protéines d'intérêt dans le milieu extracellulaire des différentes souches analysées. En parallèle et afin d'avoir une connaissance plus générale des protéines sécrétées par les souches de *L. monocytogenes*, d'autres protéines communes aux différentes souches ont été prélevées et identifiées. Chez *Listeria*, les protéines peuvent être transloquées par différents systèmes de sécrétion (Desvaux et Hébraud, 2006): (i) Sec, (ii) Tat (iii) FPE (Fimbrillin-protein exporter), (iv) FEA (Flagella Export Apparatus), (v) les holins et (v) Wss (WXG100 secretion system). Afin de caractériser les protéines identifiées dans le milieu extracellulaire de *L. monocytogenes*, des analyses de séquences ont été réalisées avec différents outils bioinformatiques. Tout d'abord, la présence d'un peptide signal a été recherchée. Cependant, l'identification d'un peptide signal, s'il signifie que la protéine est transloquée, ne renseigne pas sur la voie de sécrétion utilisée. D'autres analyses ont alors été effectuées afin de déterminer si les protéines étaient prédites comme sécrétées par le système Sec, Tat ou FPE. Par ailleurs, les protéines ne possédant pas de peptide signal ne sont pas toutes cytoplasmiques. Elles peuvent être sécrétées par le système FEA, Wss ou par les holins. Enfin, des recherches bibliographiques et l'utilisation de SecretomeP ont permis de déterminer si les protéines restantes pouvaient être prédites comme sécrétées par un système inconnu.

La présence d'un peptide signal indique que la protéine est ciblée vers la membrane cytoplasmique mais ne garantit pas sa sécrétion dans le milieu extracellulaire. C'est pourquoi des études complémentaires ont été réalisées afin de déterminer si les protéines étaient prédites :

- associées à la membrane cytoplasmique.
- associées à la paroi.
- sécrétées dans le milieu extracellulaire.

Ces résultats, combinés à des logiciels de prédiction de localisation cellulaire, ont permis de conclure sur la localisation prédite de ces protéines.

### 3 - PRINCIPAUX RESULTATS

#### **3-1. Analyse de la biodiversité au sein de *L. monocytogenes***

Il existe une large biodiversité au sein de l'espèce *L. monocytogenes*. En effet, des différences marquantes de profil protéique étaient observées entre les souches étudiées. Il est notamment intéressant d'observer que des différences d'expression des facteurs de virulence existe entre les souches. En effet, la protéine PlcB était séparées sous de multiples spots, présents ou absents de manière différente selon la souche. Egalement, la souche CLIP90602 (souche clinique, sérovar 1/2b, virulence A) n'exprimait pas les facteurs de virulence ActA, listériolysine O, PlcA et PlcB ou à un faible niveau alors que les 11 autres souches exprimaient toutes ces facteurs de virulence dans nos conditions expérimentales.

La classification hiérarchique des 12 souches montre deux classes principales, la classe I regroupant les 6 souches de serovar 1/2a et la classe II regroupant les souches de sérovar 4b et 1/2b. Les souches de sérovar 4b se classent indépendamment des souches de sérovar 1/2b à partir des données du protéome intracellulaire mais pas à partir de celles du sécrétome. En effet, excepté les deux grandes classes, les sous-groupes sont différents entre les deux sous-protéomes. Cette classification hiérarchique a permis également de grouper les spots protéiques qui ont un profil d'expression similaire entre les souches.

Afin de déterminer si certains spots protéiques étaient spécifiques de la virulence ou de l'origine, la variance relative estimée (VRE) de chaque spot a été calculée en fonction de ces facteurs. Puis une analyse discriminante a été réalisée afin de déterminer si les spots protéiques qui ont un poids significatif par l'analyse de la VRE permettaient une discrimination des souches en fonction du facteur étudié.

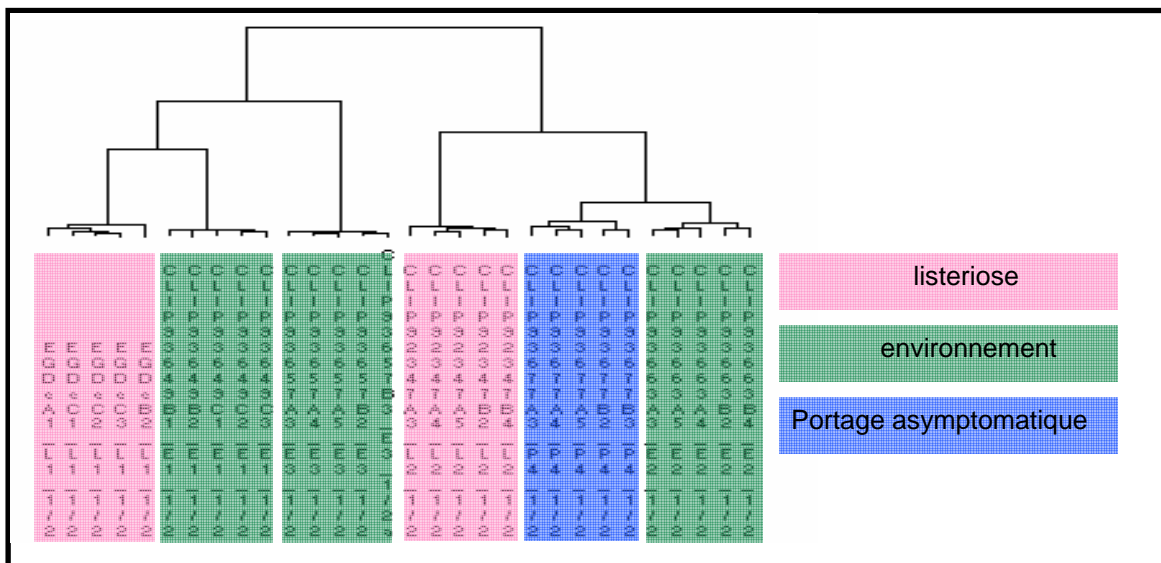
Bien que notre analyse ait permis une bonne discrimination des souches de virulence A (100% de mortalité des œufs embryonnés en moins de 3 jours) par rapport aux souches de virulence B (80% de mortalité des œufs embryonnés en plus de 3 jours) en prenant en considération un ensemble de protéines, aucune protéine prise individuellement n'a montré un

niveau d'expression homogène au sein du groupe de virulence A ou B. Concernant le facteur origine des souches, il s'est avéré impossible de séparer clairement les souches isolées de cas cliniques de celles isolées de portage ou de l'environnement d'industries agroalimentaires. Par contre, cette analyse a confirmé que l'abondance de nombreux spots protéiques était en relation avec le facteur sérovar. Une analyse factorielle discriminante réalisée avec les spots les plus représentatifs pour le facteur sérovar montre une excellente discrimination, et apparaissent donc comme de bons marqueurs du sérovar.

La classification hiérarchique et le calcul de la VRE ont ainsi contribué à sélectionner un certain nombre de spots protéiques significatifs pour la discrimination du facteur sérovar. Ces spots d'intérêt ont été soumis à une identification par spectrométrie de masse MALDI-TOF.

### **3-2. Analyse de la biodiversité au sein des souches de sérovar 4b**

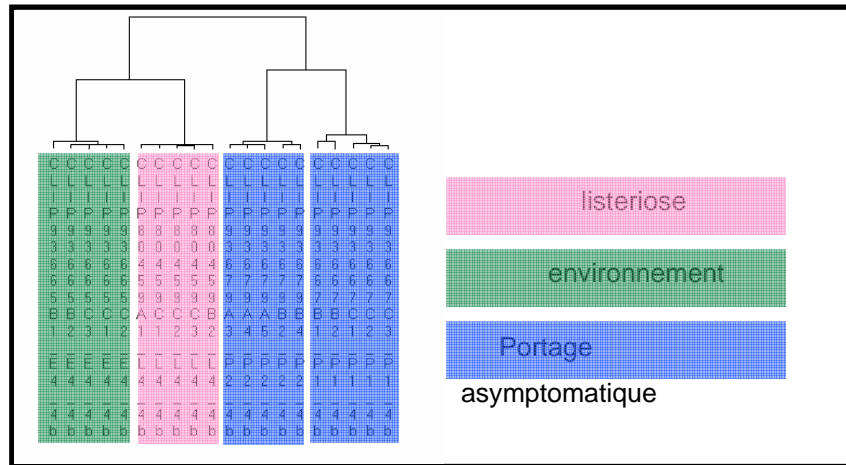
L'analyse par classification hiérarchique des 6 souches de sérovar 1/2a, indépendamment des 6 autres souches, n'a pas révélé de relation entre l'origine des souches et leur profil protéique pour les deux sous-protéomes étudiés (Figure 24).



**Figure 24 : classification des souches de sérovar 1/2a en fonction de leur profil protéique (protéines sécrétées).**

Par contre, une analyse équivalente des 4 souches de sérovar 4b, indépendamment des 8 autres souches, a révélé que les 2 souches de portage asymptotique étaient classées ensemble, tandis que la souche issue de l'environnement et la souche clinique étaient

regroupées dans une autre classe (Figure 25). Ce résultat était observé aussi bien pour les protéines sécrétées que pour les protéines intracellulaires.



**Figure 25 : Classification des souches de sérovar 1/2a en fonction de leur profil protéique (protéines sécrétées).**

De ce fait, des spots protéiques spécifiques des 2 souches de portage asymptomatique ont pu être mis en évidence. De même, des spots protéiques étaient spécifiquement présents chez la souche clinique et chez la souche environnementale. Il est évident que ces spots protéiques ne peuvent pas être pris en considération de façon définitive en tant que marqueurs de l'origine des souches 4b en raison de l'échantillonnage beaucoup trop faible de souches analysées. Ils constituent cependant un ensemble significatif de pistes à explorer en élargissant l'analyse à un nombre beaucoup plus conséquent de souches 4b pour vérifier un lien possible entre leur origine et leur profil protéique.

### **3-3. Analyse des protéines sécrétées par *L. monocytogenes***

Au total, 189 spots protéiques ont été identifiés dans le milieu extracellulaire de *L. monocytogenes*. De nombreuses protéines étaient présentes sous plusieurs spots sur les gels d'EBD, ce qui peut être dû à des modifications post-traductionnelles. L'élimination des fragments protéiques et des protéines redondantes a montré que ces 189 spots protéiques correspondaient à 77 protéines. Certaines de ces protéines étaient des orthologues, et le total des spots identifiés correspondaient finalement à 61 protéines.

La majorité des protéines identifiées étaient prédites comme sécrétées. En effet, le bilan de cette analyse montre que :

- 19 protéines étaient prédites comme cytoplasmiques dont 18 également prédites comme étant extracellulaires.
- 18 protéines étaient prédites comme associées à la membrane (5 lipoprotéines et 13 protéines membranaires intégrales) dont 4 étaient également prédites comme sécrétées dans le milieu extracellulaire.
- 15 protéines étaient prédites comme localisées au niveau de la paroi dont 13 étaient aussi prédites comme sécrétées dans le milieu extracellulaire.
- 25 protéines sécrétées par la voie Sec, étaient prédites comme relarguées dans le milieu extracellulaire.

Au total, 35 protéines étaient prédites comme ayant des localisation cellulaires multiples. Dans l'ensemble, 60 des 77 protéines étaient clairement prédites comme extracellulaires.





## Article n°2

Dumas, E., Meunier, B., Berdagué, J-L., Chambon, C., Desvaux, M. and Hébraud M.

Comparative analysis of extracellular and intracellular proteomes of twelve  
*Listeria monocytogenes* strains

## **Comparative analysis of extracellular and intracellular proteomes of twelve**

### ***Listeria monocytogenes* strains**

Dumas, E.<sup>1</sup>, Meunier, B.<sup>2</sup>, Berdagué, J.-L.<sup>3</sup>, Chambon, C.<sup>4</sup>, Desvaux, M.<sup>1</sup> and Hébraud M.<sup>1,4\*</sup>

Institut National de la Recherche Agronomique, Centre de Recherche Clermont-Ferrand/Theix, UR454 Microbiologie Equipe Qualité et Sécurité des Aliments<sup>1</sup>, UR1213 Herbivores Equipe Croissance et Métabolisme du Muscle<sup>2</sup>, UR370 Qualité des Produits Animaux Equipe Typicité Aromatique et Authentification<sup>3</sup>, Plate-Forme de Protéomique<sup>4</sup>, F-63122 Saint-Genès Champanelle, France.

Keywords: *Listeria monocytogenes*, intracellular proteome, extracellular proteome, MALDI-TOF mass spectrometry, serovar, clustering, biodiversity.

\*Corresponding author: Tel.: +33 4 73 62 46 70; E-mail address: [hebraud@clermont.inra.fr](mailto:hebraud@clermont.inra.fr)

## Abstract

*Listeria monocytogenes*, the etiologic agent of listeriosis, remains a serious public health concern with its frequent occurrence in food coupled with a high mortality rate. Among the 13 serovars, human listeriosis is mostly associated with serovar 4b strains. To investigate the diversity of *L. monocytogenes*, the intracellular and extracellular proteins of 12 strains belonging to different serovars (4b, 1/2a, 1/2b), origin and level of virulence were analysed by two-dimensional gel electrophoresis. The clustering of the strains in 2 groups, based on proteomic patterns, is in agreement with the *L. monocytogenes* defined lineages. Statistical analysis did not allow to highlight proteins specific to the origin or the virulence level of the strains, but 26 and 21 protein spots were shown to be significantly overexpressed and underexpressed, respectively, in the 6 strains of serovar 1/2a (lineage I). These protein spots, considered as potential markers of interest, were identified through their peptide mass fingerprint by using MALDI-TOF mass spectrometry.

## 1 - INTRODUCTION

*Listeria monocytogenes* is the causative agent of listeriosis, a severe disease that mainly affects pregnant women, newborn children, the elderly and immuno-compromised persons. Clinical features include severe gastroenteritis, mother-to-child infection and central nervous system infections (Vazquez-Boland *et al.*, 2001). Listeriosis is a relatively rare disease compare to other foodborne illnesses, with an incidence between 2 to 10 reported cases per million population per year. However, because of its high mortality rate (20-30%), listeriosis ranks among the most frequent causes of foodborne death since it ranks second just after salmonellosis in the USA and France (Mead *et al.*, 1999; Vaillant *et al.*, 2005) *L. monocytogenes* is problematic in the food industry since beyond its ability to form biofilm, it can tolerate a wide range of temperature, salt concentration and pH. Despite significant efforts worldwide by research organizations and the food industry to reduce the incidence of listeriosis, this ubiquitous Gram-positive pathogenic bacterium remains a critical threat to the safety of food supply.

The infectious cycle of *L. monocytogenes* is now well known (Vazquez-Boland *et al.*, 2001). First, internalins A (InlA) and B permits adhesion to the surface of the eukaryote cell and penetration into the host cell via phagocytosis. Then, bacteria escape the phagocytic vacuole by the production of a hemolysin called listeriolysin O (LLO) in combination with phospholipases A (PlcA) and B. Once free from the vacuole the bacteria multiply, become motile, gaining propulsion from the rapid polar polymerisation of host actin performed indirectly by a bacterial protein ActA. The bacteria can spread to neighboring cells

by coming into contact with the membrane and pushing it, which leads to the formation of pseudopods, which are engulfed by phagocytosis. That results in the formation of a secondary phagosome delimited by a double membrane, and the beginning of another cycle.

Although these known virulence factors are present in the 13 different serovars of *L. monocytogenes* (Doumith *et al.*, 2004), only three serovars, *i.e.* 4b, 1/2a and 1/2b, are responsible for 96% of the human listeriosis cases (respectively 49%, 27% and 20% in France between 2001 and 2003) (Goulet *et al.*, 2006). Whereas serogroup 1/2 strains are often found in the environment, the 4b serovar strains are rarely isolated from food or from the industrial environment. For example, in dried sausage processing plants, the majority of the *L. monocytogenes* isolates were of serogroup 1/2 (49.5% of serovar 1/2a, 13.0% of serovar 1/2b and 19.0% of serovar 1/2c) whereas only 8.0% of the isolates were of serovar 4b (Thevenot *et al.*, 2005). Similar observations were made on *Listeria* strains isolated from soft and semi-soft cheeses (Loncarevic *et al.*, 1995). Thus, it seems that serovar 4b strains have a higher epidemic potential and that the serovar 1/2 strains are better adapted to survive in food plants.

The current revolution in microbial investigations due to the sequencing of complete genomes has revealed new insights into the genetic structures of a number of bacterial species. In *L. monocytogenes*, whole genome comparisons of 2 strains of serovar 4b and 2 strains of serovar 1/2a (Nelson *et al.*, 2004) showed that 51 and 83 genes were respectively specific to serovar 4b and serovar 1/2a. Using DNA microarrays, genomic comparison of 93 *L. monocytogenes* strains revealed distribution into 3 distinct lineages, with serovars 4b, 1/2b and 3b found predominantly in lineage II, and with serovars 1/2a, 1/2c and 3c found predominantly in lineage I (Doumith *et al.*, 2004), which corroborated previous molecular subtyping data (Brosch *et al.*, 1994; Wiedmann *et al.*, 1997; Gray *et al.*, 2004; Nightingale *et al.*, 2005).

Although a large number of discriminatory techniques exist to subtype *L. monocytogenes*, none of them are able to predict the ability of a strain to cause disease. It should also be stressed that discrepancies exist between data at the gene level (*i.e.* coded information) and data at the protein level (functional entity) since gene expression can be regulated at different levels and led to different phenotypes. Thus, comparison of proteins content of different strains could bring new insight into survival mechanisms of this bacterium in the environment and food as well as in bacterial virulence. Proteomic analysis on cellular proteins of *L. monocytogenes* strains of different serovars indicated a large intra species variability as only 75 protein spots were found to be common to all strains (Gormon & Phan-Thanh, 1995). This analysis permitted to apprehend for the first time the biodiversity of *L. monocytogenes* by a proteomic approach and allowed to classify strains into two clusters identical to the lineages previously found by molecular biology techniques. This investigation, though, focused on whole cellular protein contents and did not take into account the nature of proteins with different levels of expression. To provide complementary information, we analyzed the extracellular and intracellular proteins of 12 strains of *L. monocytogenes* by two-dimensional gel electrophoresis (2-DE).

## 2 - MATERIALS AND METHODS

### 2-1. Strains and bacterial culture

The *L. monocytogenes* provided by Institut Pasteur (Paris) are listed in Table 1. After two precultures in brain-heart infusion (BHI), cells in stationary phase were transferred to MCDB202, a chemically defined medium (CryoBiosystem, L'Aigle, France), supplemented with 0.36% glucose and adjusted at pH 7.3. The MCDB medium was supplemented before use with 0.1% (v/v) of element traces and 1% (v/v) of a solution of 10% Yeast nitrogen base. The 250 ml-cultures were adjusted to an  $OD_{600nm} = 0.1$  and incubated at 37°C with shaking (150 rpm) until late exponential phase ( $OD_{600nm} = 0.9$ , 3 to 4 h-cultures).. The cells were harvested by centrifugation (15 min, 7500 g, 4°C).

### 2-2. Extraction of cytosoluble proteins

Cell pellets were washed twice with 1 ml TE buffer (20 mM TrisHCl; pH 7.5; 5 mM EDTA; 5 mM  $MgCl_2$ ) then resuspended in TE buffer, pH 9.0 and stored at  $-20^\circ C$  until their treatment. Cells were sonicated with a Vibracell sonifier (Bioblock Scientific, Ilkirch, France) ten times for 2 min, at power level 5 and 50% of active cycle. The samples were then treated with 1% DNase I and 1% RNase A for 30 min at room temperature. A solution was added at a final concentration of 4 M urea, 2 M thiourea, 2% w/v CHAPS and 2 mM TBP and the incubation was allowed to continue for 30 min on ice with intermittent agitation. The cells were eliminated by centrifugation (10 min, 4 500 g, 4°C) and the supernatant was ultra-centrifugated (30 min, 200 000 g, 4°C). The supernatant was collected and proteins were quantified by the method of (Bradford, 1976). Bovine serum albumin was use as a standard. The protein samples were precipitated with three volumes of cold acetone at  $-20^\circ C$  overnight, and then pelleted by centrifugation (13 000g, 40 min). Pellets were resuspended in isoelectric focusing (IEF) buffer (5 M urea, 2 M thiourea, 2% w/v CHAPS, 10 mM TrisHCl, in 50% TFE and trace of bromophenol blue) at a concentration of 5  $\mu g/\mu l$  proteins and stored at  $-20^\circ C$ .

### 2-3. Precipitation of extracellular proteins

The supernatants of the cultures were filtered (0.2  $\mu m$ ) and 0.2 mM PMSF was added to inhibit proteases activity. 0.2 mg/ml Na deoxycholate was added to the solution and was incubated 30 min on ice. Na deoxycholate supports proteins precipitation that was carried out by addition of 10% TCA, and incubation overnight at 4°C. After centrifugation (20 300 g, 30 min, 4°C), the precipitate was washed with ice cold acetone and solubilized in IEF buffer.

---

#### 2-4. 2-DE

For IEF, precast IPG strips with pH 3-10 non-linear gradient (extracellular proteins) or pH 4-7 linear gradient (intracellular proteins) were passively rehydrated during 17 h 30 in a reswelling tray with 400  $\mu$ l of IEF buffer containing 0.3% v/v ampholytes 3-10 (extracellular proteins) or 0.15% v/v ampholytes 4-6 and 0.15% v/v ampholytes 5-7, 2 mM TBP and 50  $\mu$ g of proteins. Proteins were first subjected to isoelectric focusing (IEF) for a total of 66 450 Vhs. (7 h at 50 V, 2 h at 200 V, linear gradient to 1 000 V in 2 h, 1 h at 1 000 V, linear gradient to 8 000 V in 5 h and 8 000 V till the end). The strips were equilibrated twice for 15 min in an equilibration solution (50 mM TrisHCl, pH 8.8; 6 M urea; 2% w/v SDS; 30% v/v glycerol) containing 2 mM TBP for the first step and 2.5% w/v iodoacetamide and trace of bromophenol blue for the second step. The second dimension (SDS-PAGE) was carried out with 12.5% acrylamide gel in a Multicell Protean II XL system (Bio-Rad, Marnes-la-Coquette, France). Gels were silver stained according to the method described by Rabilloud.

#### 2-5. Image analysis

In order to limit variations due to manipulation, two extractions from two different cultures were carried out and at least three 2-DE gels per sample were run per subproteome. Stained 2-DE gels were scanned by a GS-800 imaging densitometer (Bio-Rad Marnes-la-Coquette, France) and image analysis was performed using Image Master 2D Platinum software v.5.0 (GE Healthcare, Uppsala, Sweden). Under these conditions, very similar protein patterns were obtained for a given strain. For each strain, a scatter plot was performed with all the gels. The scatter plot in ImageMaster 2D Platinum gives information about the relationship between the spot values from two 2-DE gels by searching for the linear dependence between the values from one gel and the values from another gel. For each subproteome, the five most reproducible gels from two cultures for a strain were selected, and included for image analysis. These selected gels were then matched with the reference gel and a number was assigned to each group of spots matched. The reference gel has been carried out from a mixture of an equivalent quantity of proteins of the 12 strains, in order to match the maximum of spots. Relative spot volume (% vol), *i.e.* digitized staining intensity integrated over the area (volume) of the individual spot divided by the sum of volume of all spots and multiplied by 100 were used for spot quantification. A matrix was obtained with the % volume of 600 spots and 599 spots for the 60 gels of the intracellular and extracellular proteomes, respectively.

#### 2-6. Cluster analysis

For cluster analysis, only reliable protein spots were taken into account. A spot was considered reliable when it was present or absent in at least  $n-1$  gels out of the  $n = 5$  gels per strain. When the spot was

present in 4 out of 5 gels per strain, the absent value was replaced by the mean of the 4 other values. When the protein spot was reliable absent, the absent values were replaced by the lowest value present in the gel. The whole of the values were transformed in log<sub>2</sub> Ratio % Volume (Meunier *et al.*, 2007). In other words, each protein spot volume was divided by the average value of all the existing values for this protein spot in all the gels and then submitted to the base 2 logarithmic transformation.

The hierarchical clustering analysis was carried out with PermutMatrix (Caraux & Pinloche, 2005) (<http://www.lirmm.fr/~caraux/PermutMatrix/>). Clustering results were calculated using an algorithm resulting from the combination of the Ward's aggregation method and the Pearson-based distance metrics for the clustering algorithm. A preliminary study has shown that these parameters were the ones which gave the most reliable clusters (Meunier *et al.*, 2007). PermutMatrix allows clustering result visualization with dendrogram of the samples and dendrogram of the protein spots.

### 2-7. Calculation of the Estimated Relative Variance (ERV)

For the two reliable protein spots matrix, the estimated relative variance (ERV) for the virulence, the origin and the serovar were calculated ( $ERV = [\text{factor sum of square} / \text{total sum of square}] \times 100$ ).

Higher is the value, higher the expression of the spot is dependant of the studied factor. We next compute the mean square factor and mean square error by dividing the sum of squares by the appropriate degrees of freedom and then, at the last step, we divide mean square factor by mean square error to obtain the test F statistic. The p-value indicates the significance of the variation compared with the error variation. A factorial discriminant analysis (FDA) was then carried out using spots with the greatest ERV.

To discriminate the strains according to their virulence (A or B), a stepwise discriminant analysis was performed. The probability to the introduction and the rejection of the variables was fixed to  $p < 10^{-3}$ . That allowed selecting a subset of independent variables among the 159 and 219 spots for the extracellular and intracellular proteomes respectively, and to calculate a discriminant function to have an effective linear separation of virulence A and B strains.

Calculations were carried out with the software STATISTICA<sup>®</sup> version 6.1 (StatSoft, France).

### 2-8. Identification of proteins by mass spectrometry

Semi-preparative 2-DE gels containing 600 µg of proteins were run for protein spots identification. Protein spots were stained with colloidal Coomassie blue and spots of interest were excised in the gel, destained and submitted to tryptic digestion. First, spots were washed with the destaining solution (25 mM ammonium bicarbonate/5% acetonitrile), then two times with ammonium bicarbonate (25 mM)/acetonitrile (50%). Spots were then dehydrated with 100% acetonitrile. The dried gels were reswelled in a solution containing 20 µg trypsin and proteins in the gel were digested at least 5 h at 37°C. Resulting



peptides were extracted with 100% acetonitrile. After 15 min at 37°C, 1 µl of each sample and 1 µl saturated  $\alpha$ -cyano-4-hydroxycinnamic acid matrix were mixed onto the MALDI-TOF target. Positive ion MALDI mass spectra were recorded in the reflectron mode of a MALDI-TOF MS (Voyager DE-Pro, PerSeptive BioSystems, Farmingham, MA, USA) using Voyager software for data collection and analysis. The MS was calibrated with a standard peptide solution (Proteomix, LaserBio Labs, Sophia-Antipolis, France). Internal calibration of samples was achieved by using trypsin autolysis peptides. Monoisotopic peptide masses were assigned and used for NCBI database searches with the Mascot software. The following parameters were considered for the searches: a maximum fragment ion mass tolerance of +/- 25 ppm, modification of cysteines by carbamidomethylation and possible oxidation of methionine.

## 3 - RESULTS

### 3-1. General features of proteome patterns

Serovar, origin and virulence of the 12 strains analysed in this study are summarised in Table 1. Investigated *L. monocytogenes* strains belong to one of the three serovars responsible for 96% of the human listeriosis cases, *i.e.* 4b, 1/2a or 1/2b. These strains originated either of clinical samples, asymptomatical carriages or food plant environments. Two levels of virulence, *i.e.* A and B, were considered for each *L. monocytogenes* isolates and were based on mortality of infected chick embryos (Olier *et al.*, 2003). Virulence A means infection of chick embryos resulted in 100% mortality within 3 days. It can be noticed that all the strains of serovar 4b belonged to this category. Strains were of virulence B when they resulted in 80% mortality in more than 3 days (Table 1). No strain was found as avirulent in this study. Growth curves of all strains were also compared but did not shown any significant differences (data not shown).

Following growth at 37°C until the late exponential phase, the intracellular and extracellular protein fractions were isolated for comparative proteome analyses. Because the majority of intracellular protein spots of *L. monocytogenes* are comprised between pH 4 and 7 (Folio *et al.*, 2004), linear pH 4-7 IPG were used for this subproteome. As some abundant extracellular proteins were present in the basic zone, we decided to use non-linear pH 3-10 IPG to study the extracellular proteome. For each subproteome, a 2-DE gel was carried out from a mixture of an equivalent quantity of proteins of the 12 strains and was used as reference gel for the image analysis to match the maximum of spots. These 2-DE reference gels were further used to visualize all the identified protein spots (Figure 1).

As indicated by proteins patterns on the 2-DE gels from intracellular extracts, many spots were common to the 12 *L. monocytogenes* strains. In contrast, the extracellular proteomes revealed more

pronounced differences in the expression patterns. Nevertheless, some spots were specific to some strains in both subproteomes. In order to determine whether these variations were specific to the virulence, origin or serovar of a strain, further statistical analyses were performed.

### 3-2. Statistical results

#### **Hierarchical clustering**

Data were first scrutinized by hierarchical clustering analysis developed by Meunier *et al.* (Meunier *et al.*, 2007). This allowed classifying data according to their global expression profile and to group spots with similar expression profiles along the samples in order to highlight biomarkers of a specific group. Missing values, due to a real absence of protein spots or a misdetection by the image analysis software, are problematic for clustering algorithms. Thus, only reliable present or absent spots were taken into account. As a result, 219 spots for the intracellular fraction and 159 spots for the extracellular fraction were considered as reliable. These protein spots constitute the data matrix from which the hierarchical clustering was performed (Figure 2). First, good intra-class reproducibility could be observed as all gels of one strain are clustered together. Second, a clear division in two major clusters can be observed. In the two subproteomes, lineage I serovar 1/2a strain formed a significant cluster clearly distinguished from lineage II serovar 4b and 1/2b strains. These results indicated that the major differences observed between the strains was related to their lineage. The clustering of the protein spots with PermutMatrix permitted to highlight spots specific to each lineage. It is also interesting to observe that differences without any relationship with this factor exist between the strains and reflects the biodiversity of *L. monocytogenes*. The specificity of some strains is more pronounced in one subproteome than in the other one. For example, the strain EGDe (1/2a, clinical strain, virulence A) present multiple spots only in the intracellular proteome, whereas CLIP90602 (1/2b, clinical strain, virulence A) present marked differences in the extracellular proteome. In addition, the two major clusters are identical in the two subproteomes, whereas subgroups are different between the extracellular and intracellular proteomes. Only the subcluster grouping the 2 strains of serovar 4b, *i.e.* CLIP 93679 and CLIP 93667, was found in the two subproteomes.

#### **Estimated relative variance**

Data were further analyzed by calculating the Estimated Relative Variance (ERV) for the 3 factors, *i.e.*, virulence, origin and serovar, based on the 2 matrix containing reliable spots. The higher is the ERV, the higher the factor (virulence, origin or serovar) explains the variation of the spot.

For the virulence, 9 spots from extracellular proteome and 11 spots from intracellular proteome had an ERV higher than 50% ( $p < 10^6$ ). As the 12 strains presented only two different levels of virulence, a

stepwise discriminant analysis was carried out allowing to select a subset of independent variables among the two subproteomes spots from which a discriminant function was calculated. In this way, an effective linear separation was obtained for the virulence A and B strains with 6 and 7 protein spots from the extracellular and intracellular proteomes, respectively (Figure 1 S).

For the origin, only 5 spots from extracellular proteome and 9 spots from intracellular proteome had an ERV higher than 40% ( $p < 10^{-5}$ ) and no spot had an ERV higher than 59%. As there were more than two classes, a Factorial Discriminant Analysis (FDA) was performed with the most significant spots. This analysis did not allow discriminating strains according to their origin (Figure 2S).

For the two proteomes, the variation of a large number of spots was explained solely by the serovar. Indeed, 69 and 61 spots for the extracellular and intracellular proteomes respectively, had an ERV higher than 50% ( $p < 10^{-6}$ ) (Figure 3). An FDA performed with spots that had an ERV higher than 80%, *i.e.* 21 and 23 protein spots for the extracellular and intracellular proteomes, respectively, showed a very good discrimination of the strains according to their serovar (data not shown). The discrimination between the three serovars studied was always very clear when only 4 or 5 selected protein spots from extracellular or intracellular proteomes respectively were used for the FDA (Figure 3).

### 3-3. Identification of specific protein spots

The previous statistical analyses allowed to highlight protein spots differently expressed between the 12 strains, *i.e.* 199 spots, among which 159 were identified and correspond to 79 proteins. These proteins identified by MALDI-TOF mass spectrometry are listed in Table 2 (serovar specific), Table 3 (strains specific) and Table 4 (other differences). A good discrimination of the two levels of virulence could be observed when a set of spots was taken into account. However none of these spots, taken individually, presented a homogeneous level of expression according to the group A or B of virulence. Nevertheless, it can be noted that 3 spots of the extracellular proteome were overexpressed in all strains of virulence B and underexpressed in all strains of virulence A, excepted CLIP 92347 (serovar 1/2a, clinical strain, virulence A). Two of these spots were identified as Lmo0443, a protein similar to *Bacillus subtilis* transcription regulator LytR.

The statistical analyses of the data matrix did not allow to point out some pertinent protein spots allowing to significantly discriminate the strains according to their origin. On the contrary, 26 protein spots (11 extracellular and 15 intracellular proteins) were specifically overexpressed in the 6 strains of serovar 1/2a (lineage I). An aldo keto reductase (Lmo2700), a toxic ion resistance protein (Lmo1967) and an aminopeptidase (Lmo1611) were only identified in this serovar. Twenty protein spots (9 extracellular and 11 intracellular proteins) were significantly underexpressed in serovar 1/2a strains. An ATP synthase subunit (Lmof2365\_2505), a secreted phosphoglyceromutase (Lmof2365\_2429), a secreted catalase

(Lmof2365\_2776) and secreted proteins of unknown function (Lmof2365\_0970 and Lmof2365\_2053) were not detected in serovar 1/2a strains. The other differences were not due to the absence of the protein but to a difference of *Mr* and *pI* of the protein between the lineage I (serovar 1/2a) and the lineage II (serovar 4b and 1/2b). The purine nucleoside phosphorylase (deoD – Lmo1856) the phosphoglycerate mutase (Lmo2456) and the pyruvate dehydrogenase (E1 alpha subunit) pdhA (Lmo1052) were present in the 2 lineages, but respectively, only 1 out of the 2 , 3 out of the 4 and 2 out of the 4 spots present in serovar 1/2 a strains were present in lineage II strains. The nucleotide diphosphate kinase and FtsZ (Lmo2032) was present in two forms in lineage II but in only one form in lineage I. Also, a zinc metallopeptidase (Lmo2504), a LysM domain protein (Lmo2522), a penicillin-binding protein (Lmo0540) and a protein similar to *B. subtilis* YhfK protein were present in the two lineages but at a position clearly different on the 2-DE gels, in the same or in a different number of spots.

Less spots were differently expressed in strains of serovars 4b or 1/2b compared to the 2 other ones. Indeed, only 2 protein spots identified as a serine protease (LmoF2365\_1900) were present in the 4 strains of serovar 4b and absent in the 8 other serovars. Two spots, identified as the putative 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase 2 (ispD2 (Lmo1086) and a two-component response regulator (Lmo0287), were specifically under-expressed in serovar 4b strains. Only one protein, corresponding to a penicillin-binding protein (Lmo2754), was overexpressed in serovar 1/2b.

## 4 - DISCUSSION

This analysis showed a large biodiversity among *L. monocytogenes*. Indeed, no strain had the same protein pattern and lot of strain specific spots have been highlighted. The extracellular proteomes revealed more pronounced differences in the expression profiles than the intracellular proteomes. From previous genomic study on 93 *L. monocytogenes* strains, it appeared that they possessed all known virulence genes (Doumith *et al.*, 2004). From the present investigation, however, some differences could be observed at the proteomic level for some virulence factors. Indeed, PlcB was present in multiple spots with different *pI* and *Mr*. These spots were present or absent in different manners between the strains. Moreover and contrary to the 11 other strains, *L. monocytogenes* strain CLIP90602 (clinical strain, serovar 1/2b, virulence A) did not express, or at a much lower level, the virulence factors ActA, LLO, PlcA and PlcB. Lower expression of virulence genes had been previously reported in some strains responsible for listeriosis (Ripio *et al.*, 1996; Vega *et al.*, 1998; Milohanic *et al.*, 2003). While it can be argued that expression of virulence factors might be induced upon host cell infection, the 2 other clinical strains investigated here, *i.e.* CLIP92347 and CLIP80459, expressed these genes. This would indicate that environmental conditions and stimuli required for expression of listerial virulence factors (as well as the expression levels) would vary from one strain to another.

Whereas the two level of virulence could be discriminate when a set of spots was taken into account, individually, none of these spots was marker of the group of virulence. However, it could be noted that 2 spots, identified as the protein Lmo0443 and found in the extracellular proteome, were overexpressed in all strains of virulence B and underexpressed in all strains of virulence A, excepted CLIP 92347 (serovar 1/2a, clinical strain, virulence A). This protein belongs to the LytR/CpsA/Psr protein family, a group of proteins with cell-wall maintenance properties, possibly through autolysin regulation (Chatfield *et al.*, 2005). Difference of expression levels suggests distinct cell-wall activities in the two groups under the studied conditions.

Statistical analyses of results failed to discriminate the investigated strains according to their origin (*i.e.* from human listeriosis cases, asymptomatic carriage or food plant environment). Culture of bacterial strains under optimal growth conditions (pH 7.0, 37°C) might have masked specificity of the ecological niche. Inversely, comparison of the strains cultured in conditions that could mimic those encountered in the host (low pH environment of stomach, antimicrobial environment of mucosal surfaces, presence of bile and enzymes in the gastric fluid...) or in food plant environment (low temperature, high concentration of NaCl...) could have highlighted differences.

Carriage strains with truncated internalin A have been previously reported (Olier *et al.*, 2003). However, the 4 carriage strains here analyzed, *i.e.* CLIP93667 (H2), CLIP93679 (H12), CLIP93672 (H28) and CLIP93677 (H38), clearly possess entire InlA (Olier *et al.*, 2002 ; Olier *et al.*, 2003). Still, mutation of other essential virulence factors involved in listerial pathogenesis cannot be excluded. Despite the great care taken in the choice of the strains according to their origin, this information are hindered by numerous parameters. For example, it cannot be completely excluded that asymptomatic patients could not have developed listeriosis later on, or food plant isolates were not previously involved in human listeriosis. In listeriosis cases, the ingested dose of *L. monocytogenes*, the immune state or therapy applied of patients were unknown. Patients on long-term gastric acid suppressive therapy have increased risks of faecal carriage of *L. monocytogenes* but without developing listeriosis (Cobb *et al.*, 1996). Such informations have obviously an impact on statistical analysis of data but unfortunately could not have been taken into considerations.

Still, statistical analysis permitted a good discrimination of the strains according to their serovar and hierarchical clustering shown that investigated strains could be discriminated into two clusters, *i.e.* (i) cluster I including strains of serovars 4b and 1/2b, and (ii) cluster II including serovar 1/2a strains. Antigenic properties were one of the first criteria used to classify *L. monocytogenes* strains following a scheme developed by Seeliger and Höne describing 13 distinct serovars (1979). This scheme is still in use and numerous studies have shown strong correlation between phenotypes or genotypes and specific serovar (Brosch *et al.*, 1994; Wiedmann *et al.*, 1997; Gray *et al.*, 2004; Nightingale *et al.*, 2005). From there, serovars were divided in 3 distinct lineages, where strains of serovars 1/2b and 4b are grouped into

lineage II, and serovar 1/2a strains belong to lineage I. Based on the extracellular and intracellular proteome analyses, our results corroborate such a division and further revealed a strong correlation between the extracellular and intracellular protein content and serovars. This result is of great significance as presence of gene does not reflect expression of an encoded protein, which in the end is the functional entity. Regulation of expression at different levels, *i.e.* transcriptional, or translational, for example by a non functional promoter or a repression of transcription/translation could explain observed differences in these experimental conditions.

Beyond expression level of proteins, this investigation also provided more subtle information. Indeed, absence of a spot on a gel did not necessary reflected the absence of the protein. Indeed, some orthologues with different *Mr* and/or *pI* and/or posttranslational modifications migrated at different positions on the 2-DE gels. It further appeared that presence of such protein spots were serovar- or lineage-dependent and could result in differences in protein functionality. For example, numerous studies have shown the importance of protein phosphorylation on protein activities in *L. monocytogenes* (Herro *et al.*, 2005; Deutscher *et al.*, 2005; Archambaud *et al.*, 2005; Schnupf *et al.*, 2006; Archambaud *et al.*, 2006).

The statistical analyses according to the factor serovar resulted in identifying a total of 26 protein spots specific to lineage I (or serovar 1/2a), including 11 extracellular and 15 intracellular spots and 20 protein spots specific to lineage II (serovar 1/2b and 4b), including 9 extracellular and 11 intracellular proteins. One of the protein present only in lineage II strain, is the catalase (Lmof2365\_2776), primarily predicted as cytoplasmic but was found in the extracellular milieu. Multicellular localization of proteins has been previously observed in Gram-positive bacteria and this catalase is indeed predicted as extracellular by SecretomeP, with a score of 0.58 (Bendtsen *et al.*, 2004). In *B. subtilis*, extracellular catalase is essential to ensure complete protection of bacterial cell against oxidative stress. Whereas lot of spots allowed to discriminate serovar 1/2a strains (lineage I) from the two other serovar (lineage II), much less protein spots were specific to serovar 4b strains or serovar1/2b strains. From analysis of extracellular proteome, 2 protein spots identified as a serine protease (Lmof2365\_1900) were present in the 4 strains of serovar 4b but absent from all the other 8 strains. By demonstrating protein expression, the present proteomic analysis corroborates previous microarray data where gene encoding this serine protease was shown as specific to strains of serovar 4b (Doumith *et al.*, 2004). Serine proteases are known to be important for stress resistance and virulence of *L. monocytogenes*. For example, HtrA is essential for survival after exposure to many types of environmental and cellular stresses and necessary for efficient colonization of spleens (Wonderling *et al.*, 2004; Stack *et al.*, 2005; Wilson *et al.*, 2006). ClpP of *L. monocytogenes* acts as a serine protease and prevents the accumulation of altered proteins that might be toxic for the bacteria under stress conditions (Gaillot *et al.*, 2000; Gaillot *et al.*, 2001). Moreover, in the absence of ClpP, the secretion of functional LLO is reduced (Gaillot *et al.*, 2000). Only one protein was specific to strains of

serovar 1/2b, *i.e.* penicillin-binding protein (PBP) Lmo2754. PBPs are transpeptidases involved in different aspects of the cell wall synthesis in bacteria and contribute in various degrees to  $\beta$ -lactam resistance, cell morphology, or the virulence potential (Guinane *et al.*, 2006).

This investigation highlighted the large biodiversity of *L. monocytogenes* strains based on analyses of two subproteomes, namely the extracellular and intracellular proteomes. It further suggests that differential regulations of protein expression in *L. monocytogenes* may have impact on the clinical behaviour and niche adaptation. While this study was based on the most common serovars involved in human listeriosis cases, a similar approach should be widened to other serovars as well as avirulent strains to have a clearer overview of the biodiversity and specificity of expression profiles in *L. monocytogenes* strains. Ultimately, this would allow targeting protein function of interest or important protein markers and thus developing strategies to face listerial risk.

## 5 - REFERENCES

**Archambaud, C., Gouin, E., Pizarro-Cerda, J., Cossart, P. & Dussurget, O.** (2005). Translation elongation factor EF-Tu is a target for Stp, a serine-threonine phosphatase involved in virulence of *Listeria monocytogenes*. *Mol Microbiol.* 56, 383-396.

**Archambaud, C., Nahori, M. A., Pizarro-Cerda, J., Cossart, P. & Dussurget, O.** (2006). Control of *Listeria* superoxide dismutase by phosphorylation. *J Biol Chem.* 281, 31812-31822.

**Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S.** (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol.* 340, 783-95.

**Bradford, M. M.** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 72, 248-54.

**Brosch, R., Chen, J. & Luchansky, J. B.** (1994). Pulsed-field fingerprinting of listeriae: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. *Appl Environ Microbiol.* 60, 2584-2592.

**Caraux, G. & Pinloche, S.** (2005). PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics.* 21, 1280-1.

**Chatfield, C. H., Koo, H. & Quivey, R. G. Jr.** (2005). The putative autolysin regulator LytR in

*Streptococcus mutans* plays a role in cell division and is growth-phase regulated. *Microbiology*. 151, 625-31.

**Cobb, C. A., Curtis, G. D., Bansi, D. S., Slade, E., Mehal, W., Mitchell, R. G. & Chapman, R. W.** (1996). Increased prevalence of *Listeria monocytogenes* in the faeces of patients receiving long-term H2-antagonists. *Eur J Gastroenterol Hepatol*. 8, 1071-1074.

**Deutscher, J., Herro, R., Bourand, A., Mijakovic, I. & Poncet, S.** (2005). P-Ser-HPr--a link between carbon metabolism and the virulence of some pathogenic bacteria. *Biochim Biophys Acta*. 1754, 118-125.

**Doumith, M., Cazalet, C., Simoes, N., Frangeul, L., Jacquet, C., Kunst, F., Martin, P., Cossart, P., Glaser, P. & Buchrieser, C.** (2004). New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. *Infect Immun*. 72, 1072-1083.

**Folio, P., Chavant, P., Chafsey, I., Belkorchia, A., Chambon, C. & Hebraud, M.** (2004). Two-dimensional electrophoresis database of *Listeria monocytogenes* EGDe proteome and proteomic analysis of mid-log and stationary growth phase cells. *Proteomics*. 4, 3187-3201.

**Gaillot, O., Bregenholt, S., Jaubert, F., Di Santo, J. P. & Berche, P.** (2001). Stress-induced ClpP serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin O-dependent protective immunity. *Infect Immun*. 69, 4938-43 .

**Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P.** (2000). The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol Microbiol*. 35, 1286-1294.

**Gormon, T. & Phan-Thanh, L.** (1995). Identification and classification of *Listeria* by two-dimensional protein mapping. *Res Microbiol*. 146, 143-154.

**Goulet, V., Jacquet, C., Martin, P., Vaillant, V., Laurent, E. & De Valk, H.** (2006). Surveillance of human listeriosis in France, 2001-2003. *Euro Surveill*. 11.

**Gray, M. J., Zadoks, R. N., Fortes, E. D., Dogan, B., Cai, S., Chen, Y., Scott, V. N., Gombas, D. E., Boor, K. J. & Wiedmann, M.** (2004). *Listeria monocytogenes* isolates from foods and humans form distinct but overlapping populations. *Appl Environ Microbiol*. 70, 5833-41.

**Guinane, C. M., Cotter, P. D., Ross, R. P. & Hill, C.** (2006). Contribution of penicillin-binding protein homologs to antibiotic resistance, cell morphology, and virulence of *Listeria monocytogenes*



---

EGDe. *Antimicrob Agents Chemother.* 50, 2824-8.

**Herro, R., Poncet, S., Cossart, P., Buchrieser, C., Gouin, E., Glaser, P. & Deutscher, J. (2005).** How seryl-phosphorylated HPr inhibits PrfA, a transcription activator of *Listeria monocytogenes* virulence genes. *J Mol Microbiol Biotechnol.* 9, 224-234.

**Loncarevic, S., Danielsson-Tham, M. L. & Tham, W. (1995).** Occurrence of *Listeria monocytogenes* in soft and semi-soft cheeses in retail outlets in Sweden. *Int J Food Microbiol.* 26, 245-250.

**Mead, P. S., Slutsker, L., Dietz, V. , McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. & Tauxe, R. V. (1999).** Food-related illness and death in the United States. *Emerg Infect Dis.* 5, 607-625.

**Meunier, B., Dumas, E., Picc, I., Bechet, D., Hebraud, M. & Hocquette, J. F. (2007).** Assessment of Hierarchical Clustering Methodologies for Proteomic Data Mining *J. Proteome Res.* 2007, 6, 358-366.

**Milohanic, E., Glaser, P., Coppee, J. Y., Frangeul, L., Vega, Y., Vazquez-Boland, J. A., Kunst, F., Cossart, P. & Buchrieser, C. (2003).** Transcriptome analysis of *Listeria monocytogenes* identifies three groups of genes differently regulated by PrfA. *Mol Microbiol.* 47, 1613-1625.

**Nelson, K. E., Fouts, D. E., Mongodin, E. F., Ravel, J., DeBoy, R. T., Kolonay, J. F., Rasko, D. A., Angiuoli, S. V., Gill, S. R., Paulsen, I. T., Peterson, J., White, O., Nelson, W. C., Nierman, W. , Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Haft, D. H., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L. D., Uhlich, G. A., Bayles, D. O., Luchansky, J. B. & Fraser, C. M. (2004).** Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res.* 32, 2386-2395.

**Nightingale, K. K., Windham, K. & Wiedmann, M. (2005).** Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. *J Bacteriol.* 187, 5537-5551.

**Olier, M., Pierre, F., Lemaitre, J. P., Divies, C., Rousset, A. & Guzzo, J. (2002).** Assessment of the pathogenic potential of two *Listeria monocytogenes* human faecal carriage isolates. *Microbiology.* 148, 1855-1862.

**Olier, M., Pierre, F., Rousseaux, S. , Lemaitre, J. P., Rousset, A., Piveteau, P. & Guzzo, J. (2003).** Expression of truncated Internalin A is involved in impaired internalization of some *Listeria monocytogenes* isolates carried asymptotically by humans. *Infect Immun.* 71, 1217-1224.

**Ripio, M. T., Dominguez-Bernal, G., Suarez, M., Brehm, K., Berche, P. & Vazquez-Boland, J. A.** (1996). Transcriptional activation of virulence genes in wild-type strains of *Listeria monocytogenes* in response to a change in the extracellular medium composition. *Res Microbiol.* 147, 371-384.

**Schnupf, P., Portnoy, D. A. & Decatur, A. L.** (2006). Phosphorylation, ubiquitination and degradation of listeriolysin O in mammalian cells: role of the PEST-like sequence. *Cell Microbiol.* 8, 353-364.

**Seeliger, H. P. R. & Höhne, K.** (1979). Serotyping of *Listeria monocytogenes* and related species. In *Methods in microbiology*, Academic press. pp.31-49.

**Stack, H. M., Sleator, R. D., Bowers, M., Hill, C. & Gahan, C. G.** (2005). Role for HtrA in stress induction and virulence potential in *Listeria monocytogenes*. *Appl Environ Microbiol.* 71, 4241-4247.

**Thevenot, D., Delignette-Muller, M. L., Christieans, S. & Vernozy-Rozand, C.** (2005). Prevalence of *Listeria monocytogenes* in 13 dried sausage processing plants and their products. *Int J Food Microbiol.* 102, 85-94.

**Vaillant, V., de Valk, H., Baron, E., Ancelle, T., Colin, P., Delmas, M. C., Dufour, B., Pouillot, R., Le Strat, Y., Weinbreck, P., Jouglu, E. & Desenclos, J. C.** (2005). Foodborne infections in France. *Foodborne Pathog Dis.* 2, 221-232.

**Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J. & Kreft, J.** (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clin Microbiol Rev.* 14, 584-640.

**Vega, Y., Dickneite, C., Ripio, M. T., Bockmann, R., Gonzalez-Zorn, B., Novella, S., Dominguez-Bernal, G., Goebel, W. & Vazquez-Boland, J. A.** (1998). Functional similarities between the *Listeria monocytogenes* virulence regulator PrfA and cyclic AMP receptor protein: the PrfA\* (Gly145Ser) mutation increases binding affinity for target DNA. *J Bacteriol.* 180, 6655-6660.

**Wiedmann, M., Bruce, J. L., Keating, C., Johnson, A. E., McDonough, P. L. & Batt, C. A.** (1997). Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. *Infect Immun.* 65, 2707-16.

**Wilson, R. L., Brown, L. L., Kirkwood-Watts, D., Warren, T. K., Lund, S. A., King, D. S., Jones, K. F. & Hruby, D. E.** (2006). *Listeria monocytogenes* 10403S HtrA is necessary for resistance to cellular stress and virulence. *Infect Immun.* 74, 765-8.

**Wonderling, L. D., Wilkinson, B. J. & Bayles, D. O.** (2004). The htrA (degP) gene of *Listeria monocytogenes* 10403S is essential for optimal growth under stress conditions. *Appl Environ Microbiol.* 70, 1935-43.

**Figure 1: Reference gel of intracellular (a) and extracellular (b) proteins from *L. monocytogenes*.**

Reference gels for extracellular (A) and intracellular (B) proteins was carried out with a mixture of an equivalent quantity of proteins of the 12 strains, to obtain a gel with the maximum of protein spots. Proteins were separated by 2-DE, using pH 3-10 non-linear IPG strips and pH 4-7 linear IPG strips for extracellular proteins and intracellular proteins, respectively. The second dimension were performed with 12.5% acrylamide gel. The labels indicate the protein spot numbers and correspond to the spot numbers listed in Table 2

**Figure 2: Hierarchical clustering.** Clustering of the 60 gels from the 12 *Listeria monocytogenes* strains and of the 219 (intracellular fraction[a]) or 159 (extracellular fraction [b]) reliable protein spots, created with PermutMatrix. Increasing intensities of red and green indicate increasing and decreasing protein expression, respectively.

**Figure 3: Estimated Relative Variance (ERV) and Factorial Discriminant Analysis (FDA). for the serovars.**

For the intracellular and extracellular proteome reliable matrix, the ERV was calculated for the 3 serovars. The results are presented as histogram with the value of the ERV (%) for each spot. An FDA was carried out with (a) the 5 most representative spots for the intracellular proteome (among which the translation elongation factor Ts (LMOF2365\_1678), the adenylate kinase (LMOF2365\_2584) and the 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase 2 (ispD2)) and (b) the 4 most representative spots for the extracellular proteome (among which Lmo2522, a peptidase M48 Family (LMOF2365\_2477) and a serine protease (LMOF2365\_1900)). That allowed a very good discrimination of the serovars.

pH 3-10 NL

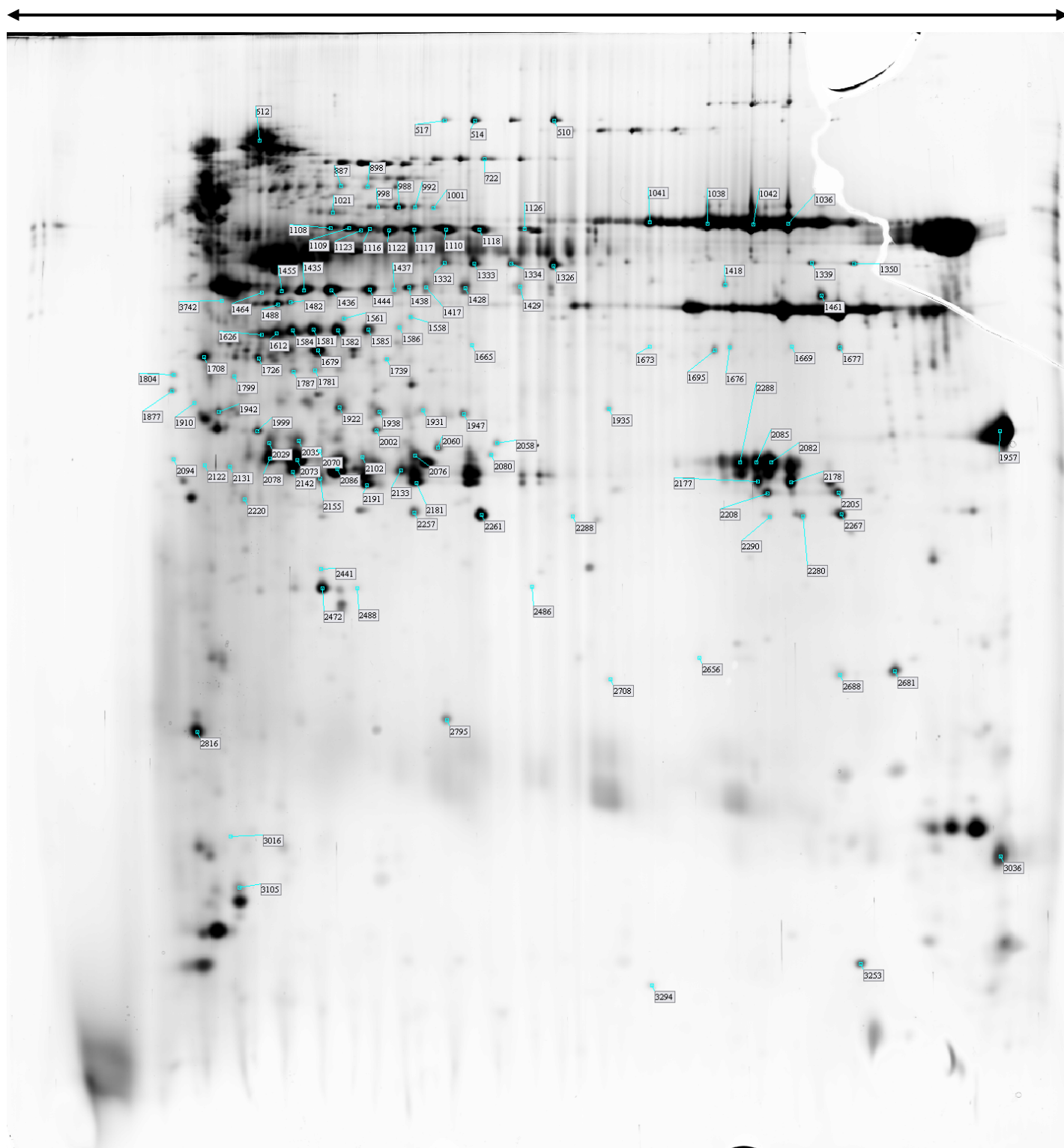


Figure 1A

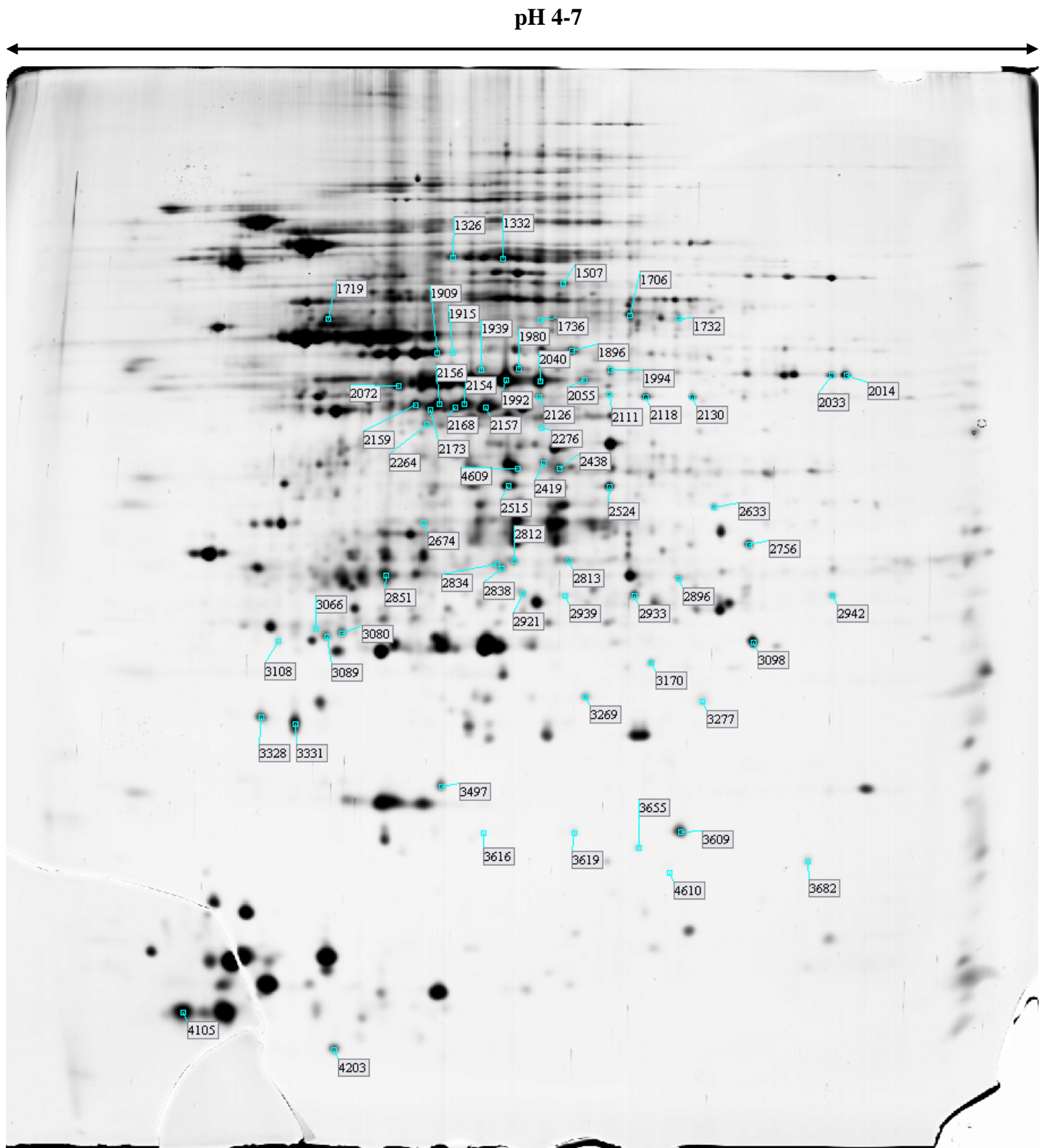


Figure 1B

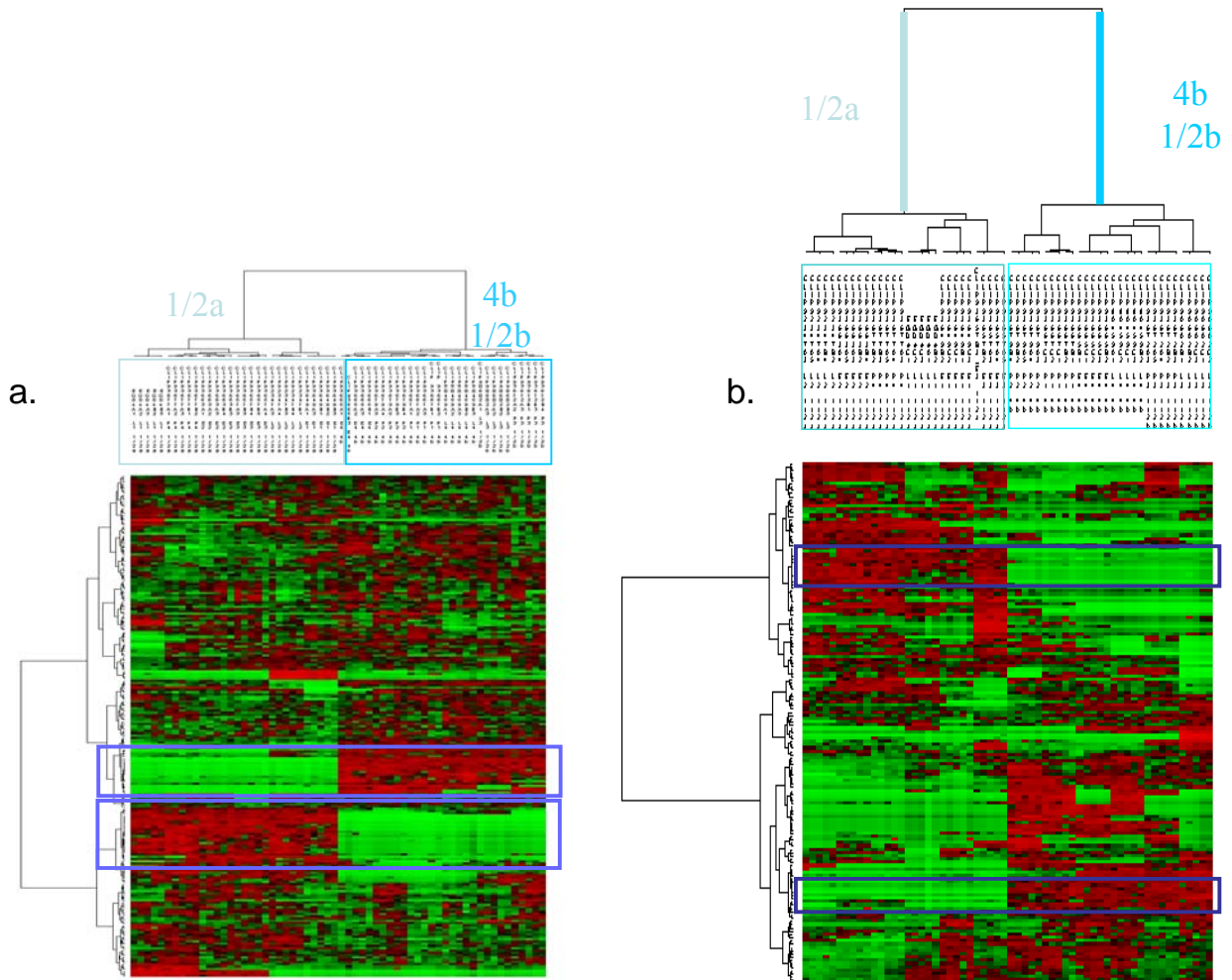


Figure 2

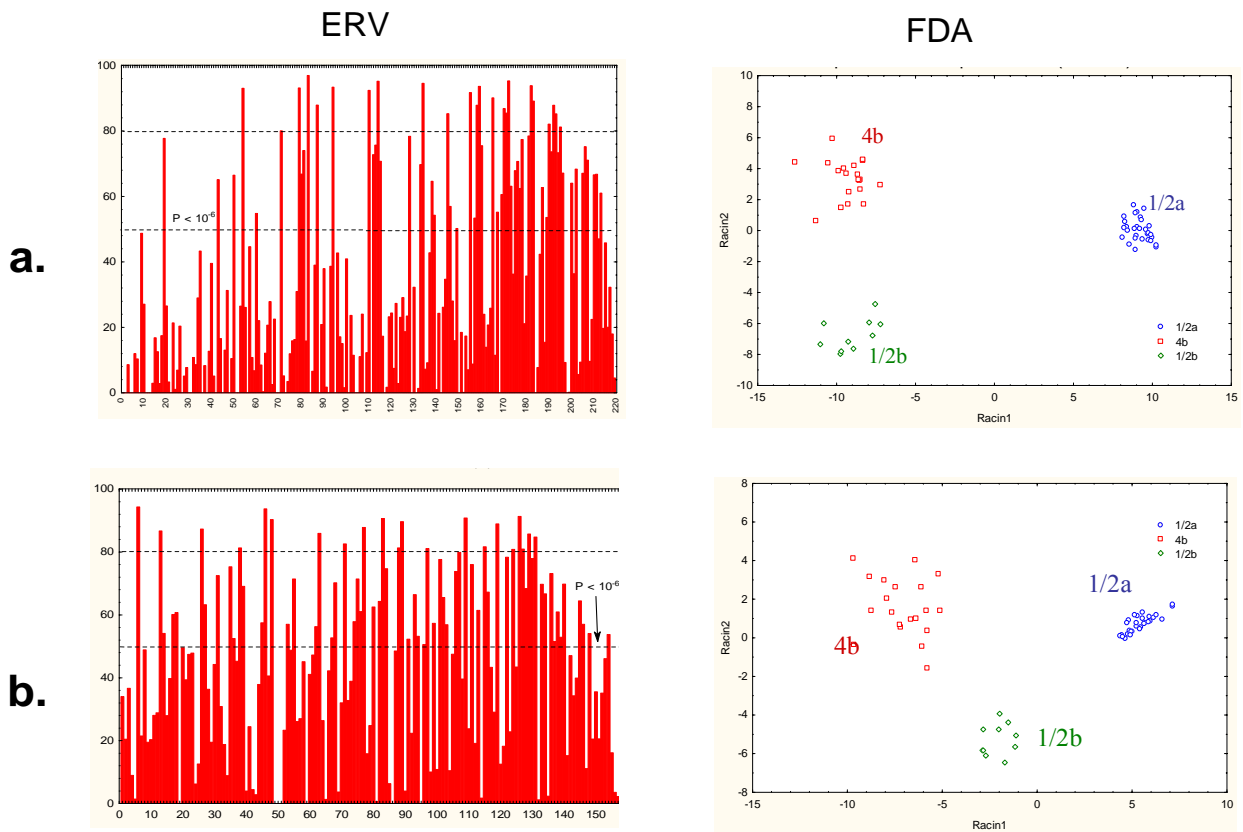


Figure 3

Origin	Strain name	Serovar	Virulence
Human Clinical cases of listeriosis	CLIP80459	4b	A
	CLIP90602	1/2b	A
	CLIP92347	1/2a	A
Animal cases of listeriosis (sequenced strain)	EGDe	1/2a	A
Asymptomatical carriage	CLIP93667	4b	A
	CLIP93679	4b	A
	CLIP93672	1/2b	B
	CLIP93677	1/2a	B
Food plant environment	CLIP93665	4b	A
	CLIP93649	1/2a	A
	CLIP93663	1/2a	B
	CLIP93657	1/2a	B

**Table 1: 12 strains of *L. monocytogenes* of different serovars, origin and virulence were used for this study.** The level of virulence has been assessed in 14 days old chick embryos as previously described (Olier *et al.*, 2003) and results for the 12 strains used are either published (Olier *et al.*, 2003) or personal communication. Virulence A means infection of chick embryos resulted in 100% mortality within 3 days. Virulence B means infection of chick embryos resulted in 80% mortality in more than 3 days.



**Table 2: MALDI-TOF MS identification of serovar specific protein spots**

Expression	localization	Group ID	Swissprot accession number	Nebi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (Da)		pI		Peptide matching	Sequence coverage (%)	VRE serovar (%)
										Theoretical	Observed	Theoretical	Observed			
spots over-expressed in serovar 1/2a strains	intracellular proteome	1736	Q8Y5T8	gi 16804006	lmo1967	lmo1967	similar to toxic ion resistance proteins	4.2	96	45464	50500	5.32	5.30	5	14	85
		2118	Q8Y6S2	gi 16803651	lmo1611	lmo1611	similar to aminopeptidase	2.2	128	38795	36900	5.57	5.50	14	53	
		2633	Q8Y3X9	gi 16804737	lmo2700	lmo2700	similar to aldo/keto reductase	2.1.1	115	32592	30400	5.77	5.80	11	41	88
		2168	Q8Y6M7	gi 16411093	tsf	tsf	translation elongation factor	3.5.3	103	32618	36500	5.11	5.10	13	50	93
		2173	Q8Y6M7	gi 16411093	tsf	tsf	translation elongation factor	3.5.3	180	32618	36500	5.11	5.00	18	64	93
		2157	Q8Y6M7	gi 16411093	tsf	tsf	translation elongation factor	3.5.3	133	32618	36500	5.11	5.20	14	50	78
		2942	Q8Y4P4	gi/16804429	lmo2391	lmo2391	conserved hypothetical protein similar to B. subtilis YhfK protein	5.2	reference map	22696	28600	6.30	6.15			71
		3080	Q8Y644	gi/16803896	deoD (lmo1856)	deoD (lmo1856)	purine nucleoside phosphorylase	2.3	reference map	29500	27700	4.86	4.80			92
		1326	Q8Y4I4	gi 16411944	pgm	pgm	pgm	2.1.2	184	56104	62000	5.10	5.10	19	49	85
		2014	Q8Y865	gi/16803092	pdhA (lmo1052)	pdhA (lmo1052)	highly similar to pyruvate dehydrogenase (E1 alpha subunit)	2.1.2	reference map	41267	40500	6.42	6.2			77
		2033	Q8Y865	gi/16803092	pdhA (lmo1052)	pdhA (lmo1052)	highly similar to pyruvate dehydrogenase (E1 alpha subunit)	2.1.2	reference map	41267	40500	6.42	6.15			73
		3066	Q8Y859	gi/16803099	lmo1059	lmo1059	lmo1059	6	reference map	19927	27800	4.54	4.70			90
		2834	Q8Y449	gi 16412099	adk	adk	adenylate kinase	2.3	84	24248	28700	5.08	5.20	13	50	92
	2812	Q8Y971	gi 16410051	thiD	thiD	highly similar to phosphomethylpyrimidine kinase thiD	2.5	72	28801	28800	5.25	5.25	5	26		
	1507	Q8Y4C0	gi/16804569	atpA (lmo2531)	atpA (lmo2531)	highly similar to H <sup>+</sup> -transporting ATP synthase chain alpha	1.4	reference map	55086	56700	5.17	5.30			65	
	extracellular proteome	2133	Q8Y4C8	gi 16412010	lmo2522	lmo2522	lysM domain protein	1.1	95	29040	26283	5.70	4.95	13	59	70
		2142	Q8Y4C8	gi 16412010	lmo2522	lmo2522	lysM domain protein	1,1	88	29040	26000	5.70	4.85	9	54	87
		2086	Q8Y4C8	gi 16412010	lmo2522	lmo2522	lysM domain protein	1.1	96	29040	26500	5.70	4.95	10	49	94
		1333	Q8Y4E2	gi 16804542	lmo2504	lmo2504	peptidase, M48 family	1.8	74	47021	51000	6.02	5.50	9	23	90
		514	Q8Y4E2	gi 16804542	lmo2504	lmo2504	peptidase, M48 family	1.8	1,9 (Profound)	47030	77000	6.00	5.50	7	23	
1126		Q8Y464	gi 16804542	lmo2504	lmo2504	peptidase, M48 family	1.8	335	57004	57000	6.11	5.80	28	55	61	
1695		Q8Y9I8	gi 16802583	lmo0540	lmo0540	similar to penicillin-binding protein	1.1	78	44506	35000	9.23	6.80	10	30	78	

Reference map: <http://www2.clermont.inra.fr/proteome/>, \*according to the classification of listlist accessible at <http://genolist.pasteur.fr/ListiList/>. SwissProt accession numbers are accessible at <http://expasy.org/sprot/>. Mascot scores were obtained against Firmicutes database where scores greater than or equal to 69 are significant (p<0.05). Spot ID514 could only be identified with Profound.

Expression	localization	Group ID	Swiss prot	Ncbi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (DA)		pI		Peptide matching	Sequence coverage (%)	VRE serovar (%)
										Theoretical	Observed	Theoretical	Observed			
spots under-expressed in serovar 1/2a strains	extracellular proteome	510	Q71WS4	gi 46908125	lmof2365_1921	lmo1892	similar to penicillin-binding protein 2A	1.8	94	90766	77500	8.62	6.00	15	16	90
		1334	Q71WS4	gi 16804542	lmof2365_2477	lmo2504	peptidase, M48 family	1.8	69	47030	51100	6.31	5.70	8	25	86
		2708	P21171	gi 16802625	iap	iap	invasion-associated protein p60	1.1	71	20388	18000	8.67	6.40	8	46	81
		2073	Q71WQ6	gi 46881966	lmof2365_2495	lmo2522	lysM domain protein	1,1	117	29025	27600	5.50	4.90	14	61	87
		1006	Q71VX5	gi 46908975	kaT (lmof2365_2776 )	kaT	catalase	4.2	371	55794	60300	5.38	4.95	33	66	83
		998	Q71VX5	gi 46908975	kaT (lmof2365_2776 )	kaT	catalase	4.2		55794	60300	5.38	4.90			
		1021	Q71WX0	gi 46908629	gpml (lmof2365_2429)	gpml	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	2.1	147	56130	59300	5.18	4.90	17	40	81
		2060	Q721L6	gi 46907182	lmof2365_0970	lmo0950	hypothetical protein	5.2	88	32466	29500	5.84	5.20	9	31	88
		2816	Q71XZ2	gi 46908257	lmof2365_2053	lmo0462	hypothetical protein	6	87	15491	15300	4.48	4.45	6	58	81
	intracellular proteome	2939	Q71X34	gi 46908565	lmof2365_2365	lmo2391	hypothetical protein	5.2	142	22809	28600	5.51	5.30	11	60	74
		3655	Q4EE05	gi 46881432	lmoh7858_2056	ndk	nucleoside diphosphate kinase	2.3	117	16438	15700	5.58	5.50	12	61	89
		1719	Q8Y5M5	gi 16411502	ftsZ (lmo2032)	ftsZ	highly similar to cell-division initiation protein FtsZ	1.7	70	41325	51000	4.76	4.75	10	30	66
		3089	Q721A7	gi 46880558	lmof2365_1080	lmo1059	conserved hypothetical protein	6	69	19897	27000	4.82	4.75	7	36	
		2813	Q722N3	gi 46906912	thiD-2 (lmof2365_0696)	thiD	phosphomethylpyrimidine kinase	2.5	120	28844	28800	5.36	5.30	11	53	85
		2838	Q71WG7	gi 46882056	lmof2365_2584	adk	adenylate kinase	2.3	74	24217	28700	5.08	5.20	9	36	94
		2156	Q71Z12	gi 46881155	lmof2365_1678	tsf	translation elongation factor Ts	3.5.3	124	32676	36600	5.06	5.05	16	53	97
		2159	Q71Z12	gi 46881155	tsf (lmof2365_1678)	tsf	translation elongation factor Ts	3.5.3	248	32676	36600	5.06	5.00	19	72	65
		2154	Q71Z12	gi 46881155	lmof2365_1678	tsf	translation elongation factor Ts	3.5.3	83	32676	36700	5.06	5.10	10	35	93
		3277	Q71Z67	gi 46881100	lmof2365_1622	lmo1601	similar to general stress protein	4,1	52	18435	19100	6.45	5.80	5	25	87
3170	Q71WP6	gi 46881976	atpH (lmof2365_2505)	atpH	ATP synthase F1, delta subunit	1.4	101	20249	24000	5.57	5.50	9	44	94		
spots under-expressed in serovar 4b strains	intracellular proteome	2756	Q8Y832	gi 16803126	ispD2 (lmo1086)	ispD2	putative 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase 2	2.1	92	26695	28800	5.71	6.00	10	30	48
		2674	Q8YA72	gi 16802333	lmo0287	lmo0287	similar to two-component response regulator	4,1	125	27271	29100	4.97	5.00	12	75	78
spots over-expressed in serovar 4b strains	extracellular proteome	1350	Q71YE5	gi 46908103	lmof2365_1900	none	serine protease, putative	4.1	70	48745	51200	8.69	7.90	7	18	89
		1339	Q71YE5	gi 46908103	lmof2365_1900	none	serine protease, putative	4.1	72	48745	51200	8.69	7.50	9	22	91
		1935	Q721A2	gi 46880563	lmof2365_1085	lmo1068	conserved hypothetical protein	6	88	30562	33000	8.45	6.40	6	37	91
		1326	Q71WS4	gi 47092790	lmof2365_2477	lmo2504	peptidase, M48 family	1.8	80	47030	51000	6.31	6.00	9	33	94
spot over-expressed in serovar 1/2b strains	extracellular proteome	1437	Q9ZIC4	gi 4104145	lmo2754	lmo2754	penicillin-binding protein homolog		100	46744	47300	5.67	5	12	34	

**Table 3: MALDI-TOF MS identification of strain specific protein spots.**

Expression	localization	Group ID	Swissprot accession number	Ncbi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (DA)		pI		Peptide matching	Sequence coverage (%)
										Theoretical	Observed	Theoretical	Observed		
spots over-expressed in CLIP80459 strain	extracellular proteome	1482	Q4EEP3	gi 47094183	pgk (Imoh7858_2607)	pgk	phosphoglycerate kinase	2.1.2	78	42050	45800	4.97	4.80	9	36
	intracellular proteome	1915	Q71WW8	gi 46908631	Imof2365_2431	pgk	phosphoglycerate kinase	2.1.2	108	42049	44400	5.02	5.10	14	44
		1909	Q71WW8	gi 46908631	Imof2365_2431	pgk	phosphoglycerate kinase	2.1.2	136	42049	44400	5.02	5.05	14	46
spot under-expressed in CLIP80459 strain	intracellular proteome	3497	Q8Y897	gi 16803057	lmo1017	lmo1017	similar to phosphotransferase system glucose-specific enzyme IIA	1.2	reference map	17671	18300	5.02	5.05		
spots over-expressed in CLIP90602 strain	extracellular proteome	887	Q4ELE1	gi 47091854	Imoh7858_2319	lmo2185	hypothetical protein	1.8	151	63322	64300	6,07	4.95	21	47
		1665	Q723B0	gi 46906785	Imof2365_0569	lmo0540	penicillin-binding protein, putative	1.1	92	44364	35600	9.24	5.50	7	21
	intracellular proteome	2524	Q8YAC3	gi 46906455	Imof2365_0234	cysK (lmo0223)	cysK protein	2.2	134	32184	32184	5.32	5.32	13	51
spots under-expressed in CLIP90602 strain	intracellular proteome	2515	Q8YAC3	gi 16802269	cysK (lmo0223)	cysK (lmo0223)	cysK protein	2.2	reference map	32217	32100	5.08	5.20		
	extracellular proteome	612	P33379	gi 16802250	actA (lmo0204)	actA (lmo0204)	actin-assembly inducing protein precursor	1.8	162	70306	73000	5.01	4.70	17	33
		1036	Q6E8U6	gi 16802248	hly (lmo0202)	hly (lmo0202)	listeriolysin O precursor	4.5	193	58652	57400	7.63	7.20	19	39
		1038	Q6E8U6	gi 16802248	hly (lmo0202)	hly (lmo0202)	listeriolysin O precursor	4.5	211	58652	57400	7.63	6.80	23	49
		1041	Q6E8U6	gi 16802248	hly (lmo0202)	hly (lmo0202)	listeriolysin O precursor	4.5	234	58652	57400	7.63	6.60	22	45
		1461	Q6E8U6	gi 16802248	hly (lmo0202)	hly (lmo0202)	listeriolysin O precursor	4.5	94	58666	46700	7.63	7.60	14	32
		1418	Q6E8U7	gi 16802248	hly (lmo0202)	hly (lmo0202)	listeriolysin O precursor	4.5	95	58666	48100	7.63	6.80	11	29
		1957	P34024	gi 16802247	plcA (lmo0201)	plcA (lmo0201)	PI-PLC	2.4	130	36295	32100	9.63	9.90	14	41
spots over-expressed in EGDe strain	intracellular proteome	3616	Q8Y5X4	gi 16803968	ndk (lmo1929)	ndk (lmo1929)	nucleoside diphosphate kinase	2.3	64	16438	16400	5.58	5.20	8	50
		2419	P33380	gi 16802256	ldh1 (lmo0210)	ldh1 (lmo0210)	L-lactate dehydrogenase 1	2.1.1	reference map	34191	34200	5.05	5.30		
		3619	Q8Y6B6	gi 16803815	purE	purE	hosphoribosylaminoimidazole carboxylase I	2.3	45	16977	16500	5.36	5.30	6	41
		3609	Q8Y6B6	gi 16803815	purE (lmo1775)	purE (lmo1775)	hosphoribosylaminoimidazole carboxylase I	2.3	134	16977	16500	5.36	5.70	11	61
		2896	Q8Y571	gi 16804244	gpmA (lmo2205)	gpmA (lmo2205)	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	2.1.2	reference map	26417	28600	5.68	5.70		

Expression	localization	Group ID	Swissprot accession number	Nebi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (DA)		pI		Peptide matching	Sequence coverage (%)
										Theoretical	Observed	Theoretical	Observed		
spots under-expressed in EGDe strain	intracellular proteome	3098	Q8Y669	gi 16803870	lmo1830	lmo1830	similar to conserved hypothetical protein	5.2	reference map	20917	27000	5.97	6.00		
		3328	P64277	gi 16803536	greA (lmo1496)	greA (lmo1496)	transcription elongation factor greA	3.5.3	reference map	17484	19200	4.31	4.60		
		3331	Q8Y8U6	gi 16802838	lmo0796	lmo0796	conserved hypothetical protein	5.2	reference map	19337	19200	4.45	4.70		
		4203	Q92CZ4	gi 16803068	lmo1028	lmo1028	similar to <i>B. subtilis</i> YkzG protein	5.2	reference map	8292	8400	4.52	4.70		
		2851	Q8Y4I3	gi 16411945	tpi	tpi	triosephosphate isomerase	2.1.2	80	26862	28700	4.73	4.90	8	40
spots over-expressed in CLIP93665 strain	extracellular proteome	1676	Q723B0	gi 46906785	lmof2365_0569	lmo0540	penicillin-binding protein, putative	1.1	96	44364	35600	9.24	6.90	11	34
	intracellular proteome	2126	Q8Y6T3	gi 16803639	ccpA (lmo1599)	ccpA (lmo1599)	catabolite control protein A	3.5.2	reference map	36983	36900	5.05	5.30		
spots under-expressed in CLIP93649 strain	intracellular proteome	1706	Q71X14	gi 46881857	lmof2365_2385	lmo2414	FeS assembly protein SufD	2.2	96	47587	51100	5.61	5.40	13	37
		3269	Q929B9	gi 16804295	lmo2256	lmo2256	Lmo2256 protein	5.2	70	19061	19400	5.22	5.30	5	34
		2921	Q722C3	gi 46880291	lmof2365_0811	lmo0794	hypothetical protein	5.2	55	23328	28600	5.23	5.30	4	21
spots over-expressed in CLIP93657 strain	extracellular proteome	2122	P33379	gi 16802250	actA (lmo0204)	ActA	ActA	1.8	53	19897	27000	4.00	4.50	4	32
	extracellular proteome	1739	Q71YD3	gi 46908115	lmof2365_1912	lmo1883	chitinase	2.1.1	104	37907	35000	5.21	5.20	11	43
		1435	Q8Y707	gi 16803561	lmo1521	Lmo1521	Lmo1521 protein	1.1	181	46043	47400	5.53	4.90	16	47
		1438	Q8Y707	gi 16803561	lmo1521	Lmo1521	Lmo1521 protein	1.1	95	46043	47600	5.53	5.35	7	21
		1444	Q8Y707	gi 16803561	lmo1521	Lmo1521	Lmo1521 protein	1.1	204	46043	47400	5.53	5.20	19	53
		2029	Q8Y6D5	gi 16803561	lmo1752	Lmo1752	Lmo1521 protein	6	58	28697	30000	5.10	4.8	7	27
		2295	Q48754	gi 16410817	tcsA	tcsA	CD4+ T cell-stimulating antigen, lipoprotein	1.2	78	38391	22600	5.02	4.95	8	25
	1799	Q4EGS3	gi 16410817	lmoh7858_1481	tcsA	CD4+ T cell-stimulating antigen, lipoprotein	1.2	93	38391	34300	5.02	4.60	9	33	
intracellular proteome	1896	Q8Y6V0	gi 16411010	ackA	ackA	acetate kinase	2.1.1	131	43822	44500	5.32	5.30	17	48	
spots under-expressed in CLIP93657 strain	extracellular proteome	2155	Q8Y6A8	gi 16803826	InlC (lmo1786)	InlC (lmo1786)	internalin C	1.9	63	33058	25100	6.12	4.90	7	26
	intracellular proteome	4609	Q8Y5G2	gi 16411571	lmo2101	lmo2102	protein required for pyridoxine synthesis	2.5	133	31700	33600	5.35	5.25	15	46
spots over-expressed in CLIP93672 strain	extracellular proteome	2290	P33378	gi 16802251	plcB ( lmo0205 )	plcB (lmo0205)	phospholipase C [Precursor]	2.4	92	33256	22600	8.37	7.00	13	43

**Table 4: MALDI-TOF MS identification of proteins spots with difference of abundance between the 12 studied strains.**

Expression	Localization	Group ID	Swissprot accession number	Nebi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (Da)		pI		Peptide matching	Sequence coverage (%)
										Theoretical	Observed	Theoretical	Observed		
absent CLIP80459 and CLIP90602	extracellular proteome	2280	P33378	gi 16802251	plcB	plcB	phospholipase C	2.4	87	33172	22600	8.37	7.4	13	41
absent CLIP80459, CLIP90602, CLIP93667, CLIP93665, CLIP93657	extracellular proteome	2181	P71451	gi 16803826	InlC	InlC	InlC protein [Precursor]	1.9	157	33058	24500	6.12	5.00	15	50
Absent 1/2a and CLIP80459	intracellular proteome	2933	Q71X34	gi 46881837	lmof2365_2365	lmo2391	hypothetical protein	5.2	102	22809	28600	5.51	5.45	9	47
absent 1/2a and CLIP90602	extracellular proteome	2488	Q71ZN1	gi 46880935	lmof2365_1458	sod	superoxide dismutase, Mn	4.2	93	22601	22600	5.23	4.95	9	47
absent 4b and EGDe	extracellular proteome	1947	Q4EML9	gi 47097689	lmof6854_0015	lmo0013 (qoxA)	quinol oxidase AA3, subunit II	1.4	73	41557	32900	6.42	5.40		
		2080	Q4ERV2	gi 16803108	lmo1068	lmo1068	lipoprotein, putative	6	74	30782		5.86		6	31
absent CLIP80459 and CLIP93665	extracellular proteome	1118	Q8Y464	gi 16804629	lmo2591	lmo2591	surface protein (GW repeat) similar to N-acetylmuramidase	1.8	266	57004	56700	6.11	5.50	27	50
absent CLIP80459, CLIP90602, CLIP93667, CLIP93672, CLIP93665	extracellular proteome	2267	P33378	gi 16802251	plcB ( lmo0205 )	plcB (lmo0205)	phospholipase C [Precursor]	2.4	92	33256	22600	8.37	7.80	12	34
absent CLIP90602 and CLIP93679	extracellular proteome	2177	P33378	gi 16802251	lmo0205	lmo0205	phospholipase C [Precursor]	2.4	119	33256	24700	8.37	7.00	14	39
absent CLIP90602, EGDe, CLIP92347, CLIP93677	extracellular proteome	2441	Q71ZN1	gi 46880935	lmof2365_1458	sod	superoxide dismutase, Mn (sod)	4.2	168	22601	22600	5.23	4.95	13	92
absent CLIP93667 and CLIP93679	intracellular proteome	3682	Q71ZK2	gi 46880965	lmof2365_1487	lmo1468	GatB/Yqey domain protein	5.2	73	16824	15000	5.96	6.10	4	33
		2035	Q8Y436	16804675	lmo2637	lmo2637	Lmo2637 protein	1.8	35	32721		6.25			
absent CLIP93677 and CLIP93663	extracellular proteome	1726	Q8Y572	gi 16804242	lmo2203	lmo2203	similar to N-actylmuramoyl-L-alanine amidase and to internalin B	1.8	179	41704	35100	5.28	4.70	14	47
absent CLIP93679 and CLIP93672	intracellular proteome	1989	Q9L723	gi 16803438	recA (lmo1398 )	recA (lmo1398)	protein recA	3.3	reference map	37993	41500	4.81	5.20		
absent EGDe, CLIP93677, CLIP93649, CLIP93663, CLIP93657	extracellular proteome	2795	Q71W91	gi 46882132	lmof2365_2660	kdpC	potassium-transporting ATPase C chain	1.2	48	20855	16000	6.84	5.30	6	47

Expression	Localization	Group ID	Swissprot accession number	Nebi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (Da)		pI		Peptide matching	Sequence coverage (%)
										Theoretical	Observed	Theoretical	Observed		
absent EGDe, CLIP93677, CLIP93663, CLIP93657	intracellular proteome	3108	Q8Y4Q8	gi 16804414	lmo2376	lmo2376	similar to peptidyl-prolyl cis-trans isomerase	3.9	69	21400	27000	4.60	4.70	6	38
under-expressed CLIP92347 and CLIP93649	intracellular proteome	1992	Q8Y4I1	gi 16804497	gap ( lmo2459 )	gap (lmo2459)	glyceraldehyde 3-phosphate dehydrogenase	2.1.2	reference map	36286	39500	5.00	5.20		
		2040	Q8Y4I1	gi 16804497	gap ( lmo2459 )	gap (lmo2459)	glyceraldehyde 3-phosphate dehydrogenase	2.1.2	reference map	36286	39000	5.00	5.30		
under-expressed EGDe and CLIP90602	intracellular proteome	4105	P0A358	gi 16804055	cspB (lmo2016)	cspB (lmo2016)	similar to major cold-shock protein	4.1	reference map	7298	9500	4.15	4.40		
over-expressed CLIP90602, CLIP93672, CLIP93657	extracellular proteome	1417	Q71W09	gi 46882214	lmof2365_2742	lmo2754	D-alanyl-D-alanine carboxypeptidase	1.1	100	46744	47700	5,67	5.10	12	34
over-expressed CLIP80459, CLIP90602, CLIP93672, CLIP93665	extracellular proteome	1123	Q71WI7	gi 46882035	lmof2365_2564	lmo2591	N-acetylmuramoyl-L-alanine amidase, family 4	1.8	258	57025	56500	5,80	4.95	23	51
over-expressed CLIP80459, CLIP93667, CLIP93679	extracellular proteome	1669	Q723B0	gi 46906785	lmof2365_0569	lmo0540	penicillin-binding protein, putative	1.1	71	44364	35600	9.24	7.20	8	18
		3253	Q4EHT4	gi 47093027	lmoh7858_0018	none	autolysin, putative		40	12608	9600	9.13	8.00	4	18
over-expressed CLIP90602 and CLIP93657	extracellular proteome	1428	Q8Y707	gi 16803561	lmo1521	lmo1521	Lmo1521 protein	1.1	96	46043	47600	5.53	5.40	7	21
over-expressed CLIP92347 and CLIP93649	intracellular proteome	2264	Q8Y9I9	gi 16409915	lmo0539	lmo0539	similar to tagatose-1,6-diphosphate aldolase	2.1.1	77	37660	35700	4.94	5.00	4	13
over-expressed CLIP92347, CLIP93667 and CLIP93663	extracellular proteome	2220	Q844E1	gi 16802250	actA (lmo0204)	actA	actin-assembly inducing protein precursor	1.8	53	22035	22500	9.03	4.70	5	40
over-expressed CLIP93667 and CLIP93679	intracellular proteome	2111	Q71Z56	gi 46881111	lmof2365_1633	lmo1611	putative glutamyl-aminopeptidase	2.2	118	38738	37200	5.69	5.3	9	30
		4610	Q71ZK2	gi 46907696	lmof2365_1487	lmo1468	GatB/Yqey domain protein	5.2	38	16840	14590	6.00	5.6	6	46
over-expressed EGDe and CLIP93657	intracellular proteome	2438	Q8Y5G2	gi 16804140	pdxS (lmo2101)	pdxS (lmo2101)	pyridoxal biosynthesis lyase pdxS	2.5	133	31700	33600	5.35	5.30	15	46

Expression	Localization	Group ID	Swissprot accession number	Nebi accession number	Gene name	EGDe orthologue	Protein description	Functional category*	Mascot score	Mr (DA)		pI		Peptide matching	Sequence coverage (%)
										Theoretical	Observed	Theoretical	Observed		
present CLIP92347, EGDe, CLIP93677, CLIP93663	extracellular proteome	2102	Q83Z11	gi 16803826	InlC	InlC	InlC [Fragment]	1.9	75	33058	28000	6.12	4.95	8	30
present 1/2a and CLIP80459	intracellular proteome	1980	Q8Y6V2	gi 16803619	lmo1579	lmo1579	Lmo1579 protein	2.2	170	33000	41500	4.95	5.30	16	51
present CLIP80459 and CLIP93665	intracellular proteome	2130	Q71Z56	gi 46907842	lmof2365_1633	lmo1611	glutamyl-aminopeptidase, putative	2.2	251	38738	37000	5.69	5.70	22	75
present CLIP80459, CLIP90602, EGDe, CLIP93672, CLIP93665	extracellular proteome	1938	Q725F6	gi 16802061	qoxA (lmo0013)	qoxA	quinol oxidase AA3, subunit II	1.4	64	41558	41800	6.04	5.20	7	23
present CLIP80459, CLIP93679, CLIP93677	extracellular proteome	1677	Q723B0	gi 46906785	lmof2365_0569	lmo0540	penicillin-binding protein, putative	1.1	86	44364	35700	9.24	7.80	12	24
present CLIP80459, CLIP93667, CLIP93679, CLIP93672, CLIP93665	intracellular proteome	4614	Q4ELZ8	gi 47091401	lmoh7858_0296	lmo0265	peptidase, M20/M25/M40 family	2.2	69	41985	42000	4.98	5	9	40
present CLIP92347 and CLIP93649	extracellular proteome	2094	Q8Y9E5	gi 16802628	lmo0585	lmo0586	putative secreted protein	1.8	74	24350	27600	4.57	4.30	6	37
	intracellular proteome	2072	Q8Y4I1	gi 16411947	gap ( lmo2459 )	gap (lmo2459)	Gap protein	2.1.2	82	36264	38700	5.20	5.00	11	33
		1732	Q8Y4M3	gi 16411902	lmo2414	lmo2414	similar to aminotransferase	2.2	91	48072	50800	5.61	5.70	6	17
present CLIP92347, EGDe, CLIP93677, CLIP93649, CLIP93663	extracellular proteome	1999	Q8Y6D5	gi 16803792	lmo1752	lmo1752	lmo1752	6	131	28697	32000	5.10	4.70	10	32
present 93667, CLIP93679, CLIP93665, CLIP93672	intracellular proteome	1994	Q71Z88	gi 46907810	ald (lmof2365_1601)	lmo1579	alanine dehydrogenase	2.2	170	39572	41000	5.24	5.30	16	51
present EGDE, CLIP93677, CLIP93663, CLIP93657	intracellular proteome	2276	Q8Y4I1	gi 16411947	gap (lmo2459)	gap (lmo2459)	glyceraldehyde 3-phosphate dehydrogenase	2.1.2	81	36263	35600	5.20	5.30	11	40
		1626	Q844J0	gi 16802250	actA	actA	ActA [Fragment]	1.8	82	70306	36100	5.01	4.70	10	18

significant variation between the 12 strains	extracellular proteome	2002	Q8YAC3	gi 16802269	cysK (lmo0223)	cysK (lmo0223)	CysK protein	2.2	108	32198	32200	5.32	4.95	13	52
		1582	Q8Y4I1	gi 16804497	gap ( lmo2459 )	gap (lmo2459)	Gap protein	2.1.2	77	36284	36200	5.20	4.95	12	32
		1581	Q8Y4I1	gi 16804497	gap ( lmo2459 )	gap (lmo2459)	Gap protein	2.1.2	139	36264	36500	5.20	4.90	13	42
		1585	Q8Y4I1	gi 16804497	gap ( lmo2459 )	gap (lmo2459)	Gap protein	2.1.2	93	36264	36900	5.20	4.95	12	30
		2191	P71451	gi 16803826	InlC	InlC	InlC protein [Precursor]	1.9	172	33058	24200	6.12	4.95	13	44
		2076	Q83Z11	gi 16803826	inlC	inlC	InlC [Fragment]	1.9	92	33058	28200	6.12	5.05	10	38
		2131	Q8Y9V8	gi 16802456	lmo0412	lmo0412	Lmo0412 protein	6	69	31274	26800	5.04	4.60	9	35
		2257	Q8Y9T0	gi 16802487	lmo0443	lmo0443	Lmo0443 ("similar to <i>B. subtilis</i> transcription regulator LytR")	3.5.2	101	34124	22600	9.05	5.05	13	32
		2261	Q8Y9T0	gi 16802487	lmo0443	lmo0443	Lmo0443 protein	3.5.2	141	34124	22600	9.05	5.50	13	39
		1109	Q8Y862	gi 16803095	lmo1055	LMO1055	PdhD	2.1.2	75	49457	56500	5.24	4.95	7	17
		2058	Q4ERV2	16803108	lmof6854_1121	lmo1068	lipoprotein, putative	6	72	30782	30100	5.86	5.60	9	39
		1436	Q8Y707	gi 16803561	Lmo1521	Lmo1521	Lmo1521 protein	5.2	179	46043	47400	5.53	4.95	17	49
		1455	Q8Y707	gi 16803561	lmo1521	lmo1521	Lmo1521 protein	1.1	153	46043	47400	5.53	4.80	14	44
		1942	Q8Y4D2	gi 16804556	lmo2518	lmo2518	Lmo2518 protein	3.5.2	85	39094	33000	5.57	4.60	9	25
		2078	Q71WQ6	gi 46881967	lmof2365_2495	lmo2522	lysM domain protein	1.1	86	29025	27800	5.50	4.80	9	47
		1117	Q71WI7	gi 46908763	lmof2365_2564	lmo2591	N-acetylmuramoyl-L-alanine amidase, family 4	1.8	159	57025	56500	5.80	5.05	16	36
		1122	Q71WI7	gi 46908763	lmof2365_2564	lmo2591	N-acetylmuramoyl-L-alanine amidase, family 4	1.8	170	57025	56500	5.80	4.95	16	37
		1110	Q8Y464	gi 16804629	lmo2591	lmo2591	surface protein (GW repeat) similar to N-acetylmuramidase	1.8	158	57004	56900	6.11	5.3	14	27
		1108	Q721B1	gi 46880554	lmof2365_1076	PdhD	dihydrolipoamide dehydrogenase	2.1.2	70	49460	57000	5.20	4.95	7	17
		1116	Q721B1	gi 46880554	lmof2365_1076	PdhD	dihydrolipoamide dehydrogenase	2.1.2	56	49451	57000	5.24	5.00	5	13
		2208	Q4EHE7	gi 16802251	lmoh7858_0225 (plcB)	plcB	plcB	2.4	84	33256	23200	8.37	7.00	11	29
		1922	Q725F6	gi 46906239	lmof2365_0016	qoxA	quinol oxidase AA3, subunit II	1.4	89	41544	33200	6.04	4.95	13	56
		1931	Q4EML9	gi 47097689	qoxA (lmof6854_0015)	qoxA	quinol oxidase AA3, subunit II	1.4	104	41557	33250	6.42	5.10	6	20
	2472	Q71ZN1	gi 46880935	lmof2365_1458	sod	superoxide dismutase, Mn	4.2	160	22601	22600	5.23	4.95	13	70	
	intracellular proteome	2055	Q8Y4I1	gi/16804497	gap (lmo2459)	gap (lmo2459)	glyceraldehyde 3-phosphate dehydrogenase	2.1.2	reference map	36286	39400	5.00	5.30		
		1332	Q71WX0	gi 46881901	lmof2365_2429	pgm	phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	2.1.2	179	56130	62000	5.18	5.20	23	51



**Figure 1S: separation of the virulence A and virulence B strains.** To discriminate the strains according to their virulence (A or B), a stepwise discriminant analysis was performed from the totality of the spots. The probability to the introduction and the rejection of the variables was fixed to  $p < 0.001$ . That allowed selecting a subset of independent variables among the 159 (secretome) and the 219 (intracellular proteome) spots to calculate a discriminant function to have an effective linear separation of the virulence A and virulence B strains.

Extracellular proteome: discriminant function =  $0.74 (\text{spot1485})_{\text{CR}} + 0.45 (\text{spot3173})_{\text{CR}} + 2.22 (\text{spot2261})_{\text{CR}} + 0.73 (\text{spot2133})_{\text{CR}} - 1.84 (\text{spot2795})_{\text{CR}} - 1.72 (\text{spot1476})_{\text{CR}}$

Intracellular proteome: discriminant function =  $-1.12 (\text{spot2851})_{\text{CR}} - 0.85 (\text{spot2053})_{\text{CR}} - 0.96 (\text{spot1316})_{\text{CR}} + 1.27 (\text{spot2966})_{\text{CR}} - 0.77 (\text{spot2868})_{\text{CR}} - 1.43 (\text{spot3682})_{\text{CR}} + 1.54 (\text{spot1732})_{\text{CR}}$ .

CR = Center Reduced.

**Figure 2S: Estimated Relative Variance (ERV) and factorial discriminant analysis (FDA) for the origins.** For the (a) intracellular and (b) extracellular proteome reliable matrix, the ERV was calculated for the 3 serovars. The results are presented as histogram with the value of the ERV (%) for each spot. An FDA was carried out.

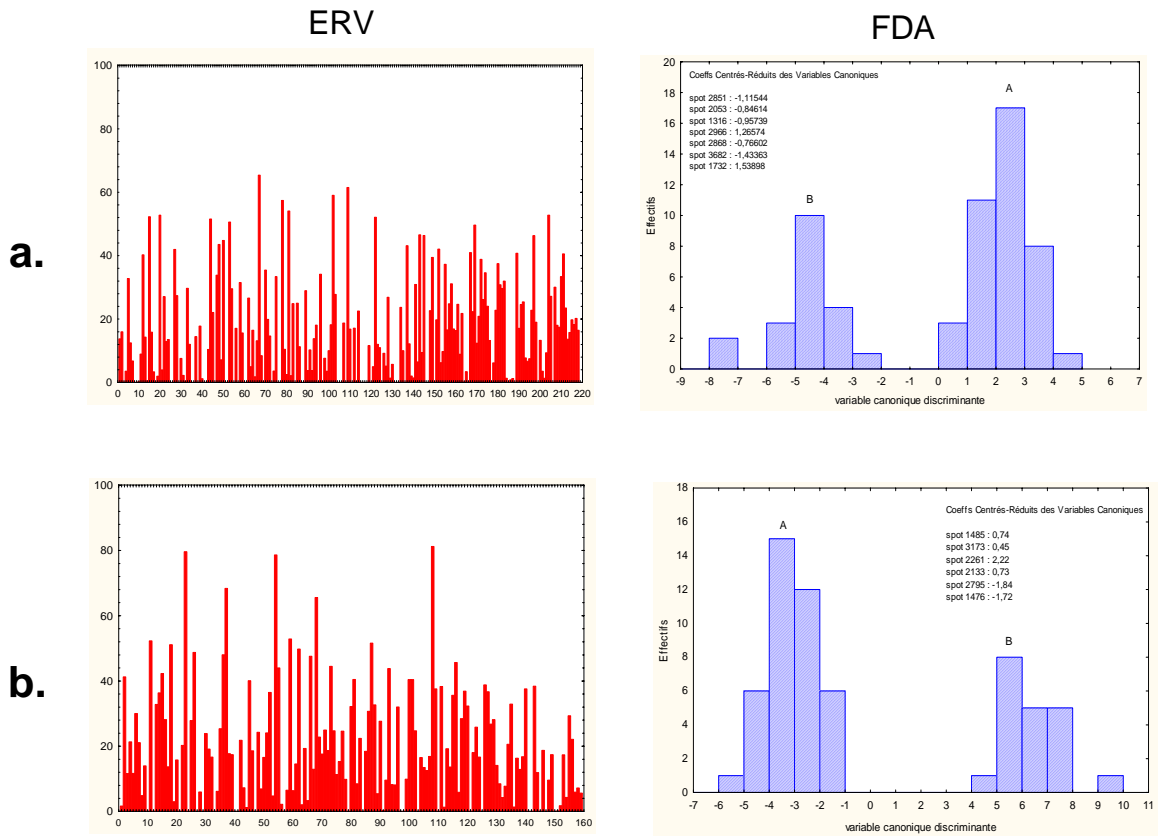


Figure 1S

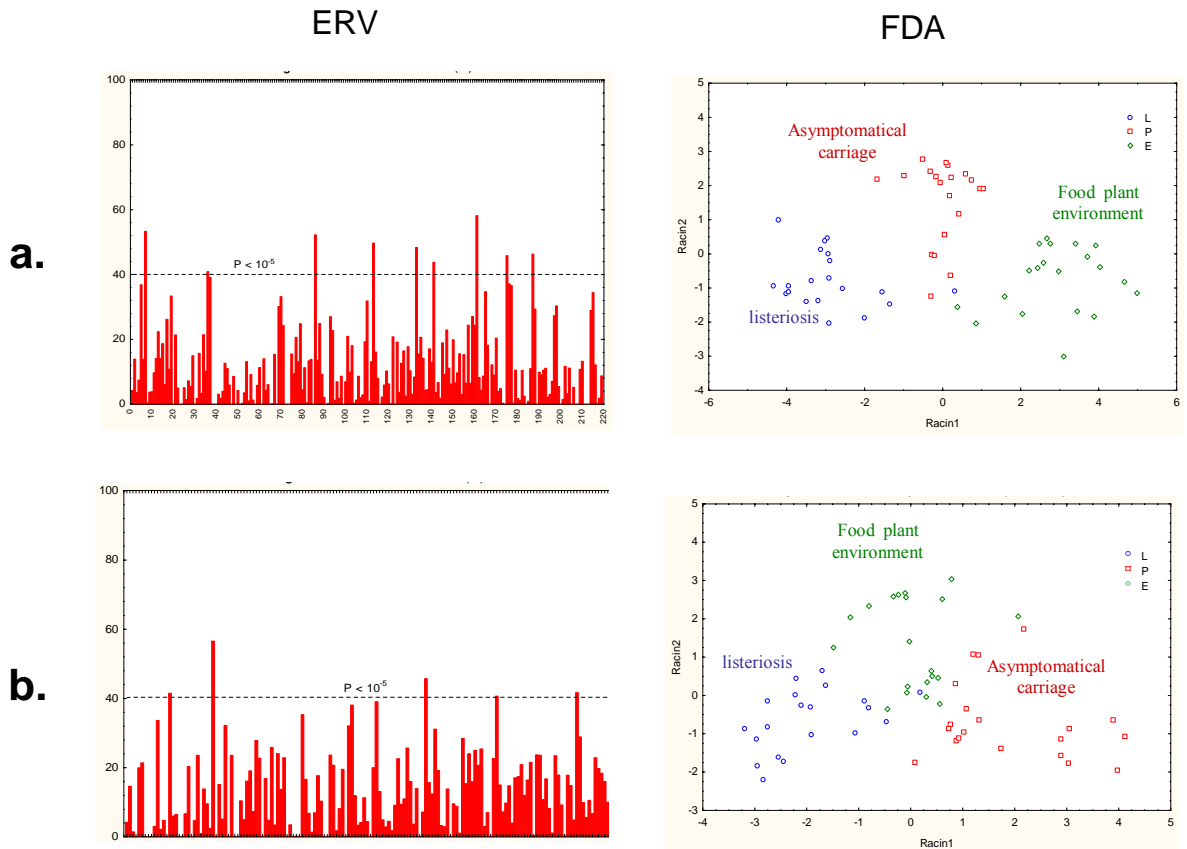


Figure 2S



## **Article n°3**

**(short note)**

Dumas, E., Meunier, B., Berdagué, J.-L., Chambon, C., Desvaux, M. and Hébraud M.

Subproteome comparison of four *Listeria monocytogenes* 4b strains from  
different origins

## **Subproteome comparison of four *Listeria monocytogenes* 4b strains from different origins**

Dumas, E.<sup>1</sup>, Meunier, B.<sup>2</sup>, Berdagué, J.-L.<sup>3</sup>, Chambon, C.<sup>4</sup>, Desvaux, M.<sup>1</sup> and Hébraud M.<sup>1,4\*</sup>

Institut National de la Recherche Agronomique, Centre de Recherche de Clermont-Ferrand/Theix, UR454 Microbiologie - Equipe Qualité et Sécurité des Aliments<sup>1</sup>, UR1213 Herbivores - Equipe Croissance et Métabolisme du Muscle<sup>2</sup>, UR370 Qualité des Produits Animaux - Equipe Typicité Aromatique et Authentification<sup>3</sup>, Plate-Forme de Protéomique<sup>4</sup>, F-63122 Saint-Genès Champanelle, France.

Keywords: *Listeria monocytogenes*, serovar 4b, intracellular proteome, extracellular proteome, MALDI-TOF mass spectrometry, biodiversity.

\*Corresponding author: Tel.: +33 4 73 62 46 70; E-mail address: [hebraud@clermont.inra.fr](mailto:hebraud@clermont.inra.fr)

### **Abstract**

Four *Listeria monocytogenes* strains isolated from food plant environment, asymptomatic carriage and human listeriosis cases and belonging to serovars 4b were analysed by two-dimensional gel electrophoresis. Based on extracellular and intracellular proteomic patterns, the 2 asymptomatic carriage strains were discriminated in a different group of the two others strains. Protein spots which differentiated the subproteome patterns were identified through their peptide mass fingerprint by using MALDI-TOF mass spectrometry.

*Listeria monocytogenes* is an ubiquitous foodborne Gram-positive bacterium, which is of great concern in the food industry since it can tolerate refrigerate temperature, relatively acid pH, high salt concentration, and form biofilms on biotic or abiotic surfaces. This pathogen is the causative agent of listeriosis, a severe disease that mainly affects pregnant women, newborn children, the elderly and immuno-compromised persons. Clinical features include severe gastroenteritis, mother-to-child infection and central nervous system infections (Vazquez-Boland *et al.*, 2001). Listeriosis is a relatively rare disease compare to other foodborne illnesses, with an incidence between 2 to 10 reported cases per million populations per year. However, because of its high mortality rate (20 to 30%), listeriosis ranks among the most frequent causes of foodborne death since it ranks second only after salmonellosis in the USA and France (Vaillant *et al.*, 2005), (Mead *et al.*, 1999). Despite appreciate efforts worldwide by research organizations and the food industry to reduce the incidence of listeriosis, this pathogen remains a critical threat to the safety of our food supply.

It is surprising that among the 13 *L. monocytogenes*, only three, namely serovars 4b, 1/2b and 1/2a represent 96% of the human listeriosis cases. Even though strains of serogroup 1/2 are the most frequently isolated from foodstuffs and food plant environments sources, all major foodborne outbreaks and a majority of sporadic cases of listeriosis are caused by serovar 4b strains, suggesting putative greater potential of adaptation or implantation for the first and a greater pathogenic potential of the 4b strain. However, some serovar 4b strains have been isolated from human in good health. The difference between the carriage strains and the epidemic ones could find an explanation in the difference of the ingested dose of *L. monocytogenes* or in the immune status of the patient. Differences in the genome or in gene expression among the genes belonging to the core genome of *L. monocytogenes* may also contribute to explain in part different clinical behavior of different *L. monocytogenes* strains. So we compared 4 strains of *L. monocytogenes* 4b strains from different origin (food plant environment, human listeriosis cases, asymptomatic carriage) to determine whether differences exist between different serovar 4b strains, which could be link to disease potential. The studied serovar 4b strains were the clinical strain *L. monocytogenes* CLIP80459 responsible for 32 cases of illness during an epidemic of listeriosis in France from November 1999 to February 2000, CLIP93667 and CLIP93679 isolated in France from healthy humans and CLIP93665 isolated from a French cheese plant (environmental strain).

Strains were first grown overnight in BHI medium (Difco), then inoculated at an OD<sub>600</sub> of 0.1 in 250 ml of the chemically defined medium MCDB 202 (CryoBioSystem, L'Aigle, France) supplemented with 1% glucose. Cells were grown at 37°C until late exponential

phase ( $OD_{600} = 0.9$ ) and then harvested by centrifugation. The intracellular proteins were extracted from the cells pellet as previously described (Folio et al., 2004) while the extracellular proteins were obtained from the filtered supernatant. On 0.2  $\mu\text{m}$  membrane. 0.2 mM PMSF and 0.2 mg/ml Na deoxycholate were added to the solution and was incubated 30 min on ice. Na deoxycholate supports protein precipitation that was carried out by addition of 10% TCA and overnight incubation at 4°C. After centrifugation (20 300 g, 30 min, 4°C), the precipitate was washed with ice cold acetone and solubilized in isoelectric focusing (IEF) buffer. For IEF, non-linear pH 3-10 IPG strips or linear pH 4-7 IPG strips for extracellular and intracellular proteins, respectively, were passively rehydrated during 17h30 in a reswelling tray with 400  $\mu\text{l}$  of IEF buffer containing 50  $\mu\text{g}$  of proteins, 2 mM TBP and 0.3 % v/v ampholytes 3-10 or 0.15% v/v ampholytes 4-6 and 0.15% v/v ampholytes 5-7, respectively. Proteins were first subjected to IEF (IEF cell, Bio-Rad, Marnes-la-Coquette, France) for a total of 66 450 Vhs.(7 h at 50 V, 2 h at 200 V, linear gradient to 1 000 V in 2 h, 1 h at 1 000V, linear gradient to 8 000 V in 5 h and 8 000 V till the end). The strips were equilibrated twice for 15 min in an equilibration solution (50 mM TrisHCl, pH 8.8; 6 M urea; 2% w/v SDS; 30% v/v glycerol) containing 2 mM TBP for the first step and 2.5% w/v iodoacétamide and trace of bromophénol blue for the second step. The second dimension (SDS-PAGE) was carried out with 12.5% acrylamide gel in a Multicell Protean II XL system (Bio-Rad). Gels were silver stained according to the method described by Rabilloud (1992)

Two independent cultures and proteins extraction were made for the 4 strains and at least 3 2-DE gels per protein sample were performed for each sub-proteome. Stained 2-DE gels were scanned with a GS-800 imaging densitometer (Bio-Rad) and image analysis was performed using Image Master 2D Platinum (GE Healthcare, Uppsala, Sweden).

For each sub-proteome, the five most reproducible gels for one strain were selected by a scatter plot analysis (Image Master 2D Platinum). The spots of the 20 remaining gels were then matched together. Relative spot volumes (% vol), i.e. the digitized staining intensity integrated over the area (volume) of the individual spot divided by the sum of volume of all spots and multiplied by 100, were used for spot quantification. We obtained a data set per subproteome, containing the % volume for the groups of spots for the 20 gels. Only reliable spots (present or absent in at least  $n-1$  gels out of the  $n = 5$  gels of a strain) were taken into account. The values were transformed in  $\log_2$  Ratio % Volume. A clustering analysis was realised with PermutMatrix, which group gels with similar expression profiles and protein spots with similar behaviour. The distance metrics based on Pearson and the Ward's method were used for the clustering algorithm (Meunier *et al.*, 2007).

Figure 1 shows the hierarchical clustering for the extracellular and intracellular proteomes, based on the abundance of the 221 and 301 reliable spots, respectively. First, we can observe a good intra-class reproducibility as all the gels of one strain clustered together. Second, we can observe a clear division into two major clusters. The two asymptomatic carriage strains formed a clearly distinguished cluster from the listeriosis and the environmental strains as well in the extracellular as in the intracellular proteome. So, the asymptomatic carriage strains seem to have different protein expression than the two other strains. PermutMatrix also permit to cluster protein spots according to similarity of expression patterns between the strains and showed clusters of protein spots specific to the environmental, carriage or clinical strains.

The estimated relative variance was calculated for each spot as described before (cf publication 2). Forty seven and 23 protein spots in the intracellular and extracellular proteomes, respectively, had an ERV higher than 50% (Figure 1 A, B). Factorial discriminant analysis with the most representative spots showed a good discrimination of the strains according to their origin (Figure 1 C, D).

From these analyses 15, 7 and 9 spots appeared specific of clinical, environmental and carriage strains respectively. The spots of interest were excised in 2-DE gels and identified from their peptide mass fingerprints, obtained by MALDI-TOF mass spectrometry (Folio *et al.*, 2004). Database identification was done with the search engine MASCOT. Globally, we successfully identified 14 distinct gene products from 28 protein spots (Table 1) selected in the two sub-proteomes.

Among the 7 protein spots characterising the CLIP93665 environmental strain was the catabolite control protein A (CcpA). This protein is found in Gram-positive bacteria involved in repressing carbohydrate utilization genes and in activating genes involved in transporting excess carbon from the cell. It has been demonstrated that CcpA is overproduced after osmotic stress in *L. monocytogenes* (Duche *et al.*, 2002a) and cold shock in *Lactococcus lactis* (Wouters *et al.*, 2000). Disruption of *L. monocytogenes ccpA* results in pleiotropic effects on growth but not in virulence (Behari & Youngman, 1998). The autolysin LMOh7858\_0018 (Figure 3a) was absent from this environmental strain. To date, five *L. monocytogenes* autolysins have been identified: P60, P45, Ami, MurA and Auto. These enzymes are involved in numerous cellular processes including cell growth, cell wall turnover, peptidoglycan maturation, cell division and separation, formation of flagella, chemotaxis and biofilm formation, genetic competence, protein secretion and pathogenicity. The gene coding this autolysin has been only found in LMOh7858, a serovar 4b strain and



was absent from the 3 other sequenced strains (Nelson *et al.*, 2004), (Glaser *et al.*, 2001). So, it would be of particular interest to study this new autolysin.

Penicillin Binding Proteins (PBPs) are transpeptidases involved in different aspects of the cell wall synthesis in bacteria and their differences of expression suggests distinct cell wall activities according to the origin. In our experimental condition, Lmo1438 is over-expressed in the environmental strain (Figure 3b). It has been demonstrated that disruptions of *lmo1438* greatly attenuated virulence (Guinane *et al.*, 2006). On the other hand, the pbp *lmo0540* (Figure 3c) was absent in the 2-DE patterns of environmental strain. In addition to its role in mouse brain colonization following intravenous inoculation, Lmo0540 also contributes to virulence as determined by colonization of the spleen (Guinane *et al.*, 2006). Another spot of the pbp Lmo0540 was also absent from this environmental strain and has only be found in the CLIP80459 listeriosis associated strain.

In addition to the pbp Lmo0540 spot characterising CLIP80459 strain, 2 spots of phosphoglycerate mutase (one in each subproteome) were over-expressed (Figure 3d and 3e). That is an enzyme that catalyses the formation of ATP to ADP and vice versa. In the second step of the second phase in glycolysis, 1,3-diphosphoglycerate is converted to 3-phosphoglycerate, forming one molecule of ATP. If the reverse were to occur, one molecule of ADP would be formed. This reaction is essential in most cells for the generation of ATP. A phosphotransferase system (PTS) glucose-specific enzyme IIA (Figure 3f) and *lmoh7858\_2319*, a protein which posses 3 NEAT (NEAr Transporter) domains, were under-expressed in CLIP80459 strain. The PTS catalyses the phosphorylation of incoming sugar substrates and coupled with translocation across the cell membrane, makes the PTS a link between the uptake and metabolism of sugars. NEAT domains are found exclusively in a family of cell surface proteins in Gram-positive bacteria. Many NEAT domain-containing proteins, are implicated in haem binding. Overexpression of *IsdA* in *S. aureus*, a NEAT domain-containing protein, enhances growth and a mutant *isdA* has a growth defect, compared with wild type, when grown in media containing haem as the sole iron source. Furthermore, the haem-binding property of *IsdA* is contained within the NEAT domain (Grigg *et al.*, 2007). Also, this strain has 3 proteins with Mr and pI different of other strains. Among them, the alanine dehydrogenase (Figure 3g) is a protein that catalyzes the NAD-dependent reversible reductive amination of pyruvate into alanine. This enzyme can be induced by stress such as saline stress as observed in *L. monocytogenes* (Duche *et al.*, 2002b). Also the epimerase LMO<sub>f</sub>2365\_2365 (Figure 3h), a family of proteins that use NAD as a

cofactor, had a different localization. The proteins in this family use nucleotide-sugar substrates for a variety of chemical reactions.

The two asymptomatic carriage strains had common protein spots different from the two other strains. The majority of these differences are also due to Mr and pI differences. Among them, we can note the superoxide dismutase A (SodA), an enzyme that protects organisms against superoxides and reactive oxygen species (ROS) produced during their active metabolism. ROS are major mediators of phagocytes microbicidal activity. A *sod* deletion mutant is impaired in survival within macrophages and is dramatically attenuated in mice (Archambaud *et al.*, 2006). A glutamylaminopeptidase was overexpressed in the 2 asymptomatic carriage strains. This peptidase is a co-catalytic metallopeptidase, typically binding two atoms of zinc or cobalt.

It is interesting to observe that PlcB (Figure 3i), an important virulence factor present under several spots, presents different localization between the different strains.

A first study of strains from different origins, serovars and virulence were divided in two clusters, one cluster grouping the strains of serovar 1/2a and the other cluster grouping the strains of serovar 4b and 1/2b. This study showed that the major differences between the strains were related to their serovar, and numerous protein spots were specific to each group. No spots specific to all the strains of the same origin of isolation could be identified. The study of the 6 serovar 1/2a strains independently of the other strains showed that the strains did not cluster according to their origin. On the other hand, when analysing only the four serovar 4b strains independently of the other strains, we could observe that the two carriage strains clustered independently of the two other strains. Specific protein spots from the different origin could be highlighted. Obviously, we can not consider these proteins as markers of the origin, which could explain niche adaptation, due to the small number of strains. However, this study allows to observe that whereas strains of serovar 4b are very similar, differences exist between the strains. Also, it is interesting to observe that multiple spots are specific to the two asymptomatic strains, whereas no spots are specific to two other strains. It is well known that serovar 4b strains are responsible for the majority of the human listeriosis cases throughout the world, while they are rarely found in food products and in the food plant environments, suggesting a higher pathogenic potential. The reason why no disease was detected after ingestion of these serovar 4b strains could find different explanations: (i) the immune state of the patient, (ii) a little ingested dose or (iii) differences in gene content and/or gene expression of the strains. This preliminary study highlighted differences specific to the asymptomatic carriage strains that could have a link with the pathogenic potential of the

strains. To confirm or not this first result, It would be necessary to extend this study to a more significant number of strains from different origin.

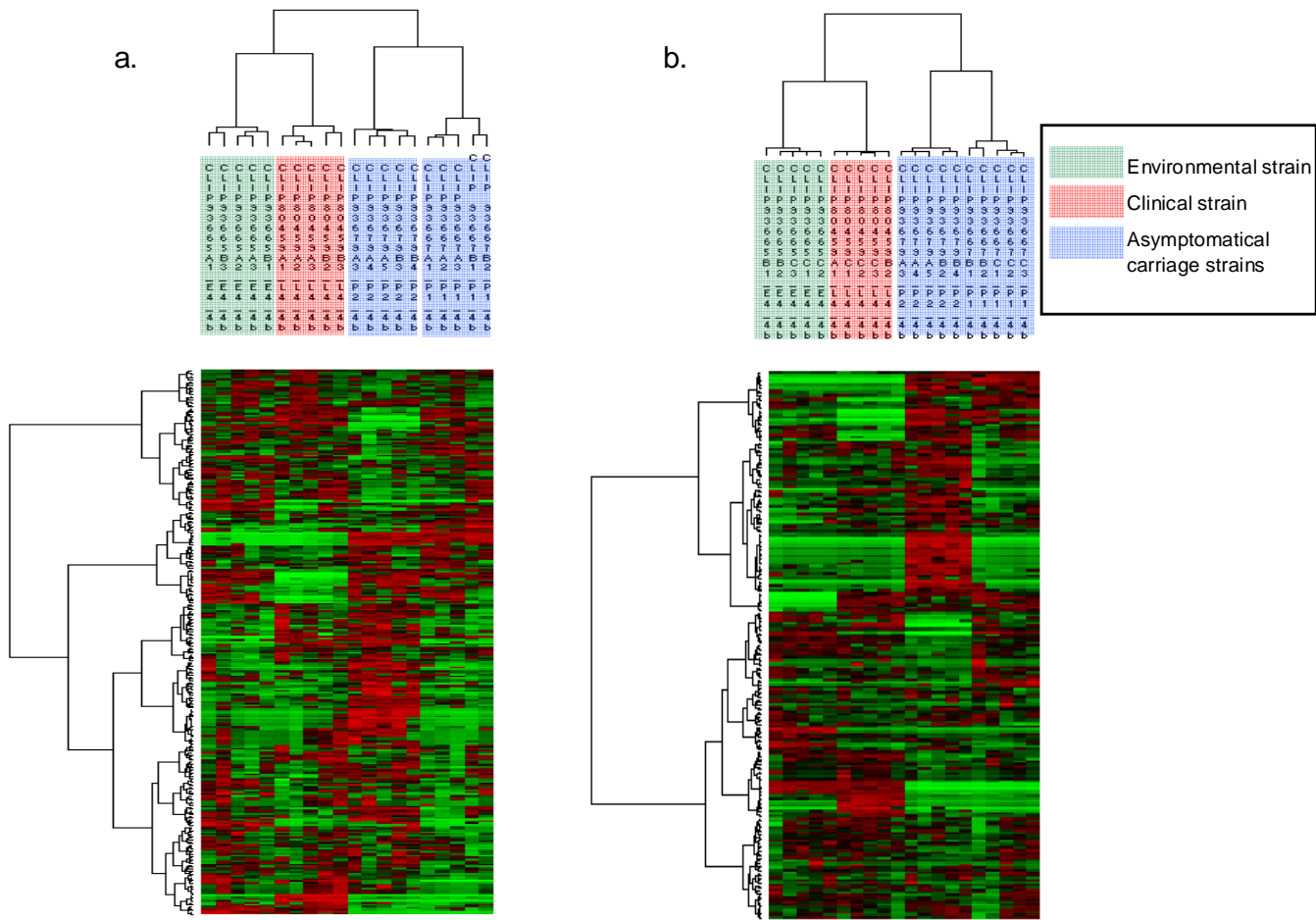
## References

1. Archambaud, C., Nahori, M. A., Pizarro-Cerda, J., Cossart, P. & Dussurget, O. (2006). Control of *Listeria* superoxide dismutase by phosphorylation. *J Biol Chem.* 281, 31812-31822.
2. Behari, J. & Youngman, P. (1998). A homolog of CcpA mediates catabolite control in *Listeria monocytogenes* but not carbon source regulation of virulence genes. *J Bacteriol.* 180, 6316-6324.
3. Duche, O., Tremoulet, F., Glaser, P. & Labadie, J. (2002a). Salt stress proteins induced in *Listeria monocytogenes*. *Appl Environ Microbiol.* 68, 1491-1498.
4. Duche, O., Tremoulet, F., Namane, A. & Labadie, J. (2002b). A proteomic analysis of the salt stress response of *Listeria monocytogenes*. *FEMS Microbiol Lett.* 215, 183-188.
5. Folio, P. , Chavant, P., Chafsey, I., Belkorchia, A., Chambon, C. & Hebraud, M. (2004). Two-dimensional electrophoresis database of *Listeria monocytogenes* EGDe proteome and proteomic analysis of mid-log and stationary growth phase cells. *Proteomics.* 4, 3187-3201.
6. Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J. C., Purcell, R., Rimmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J.

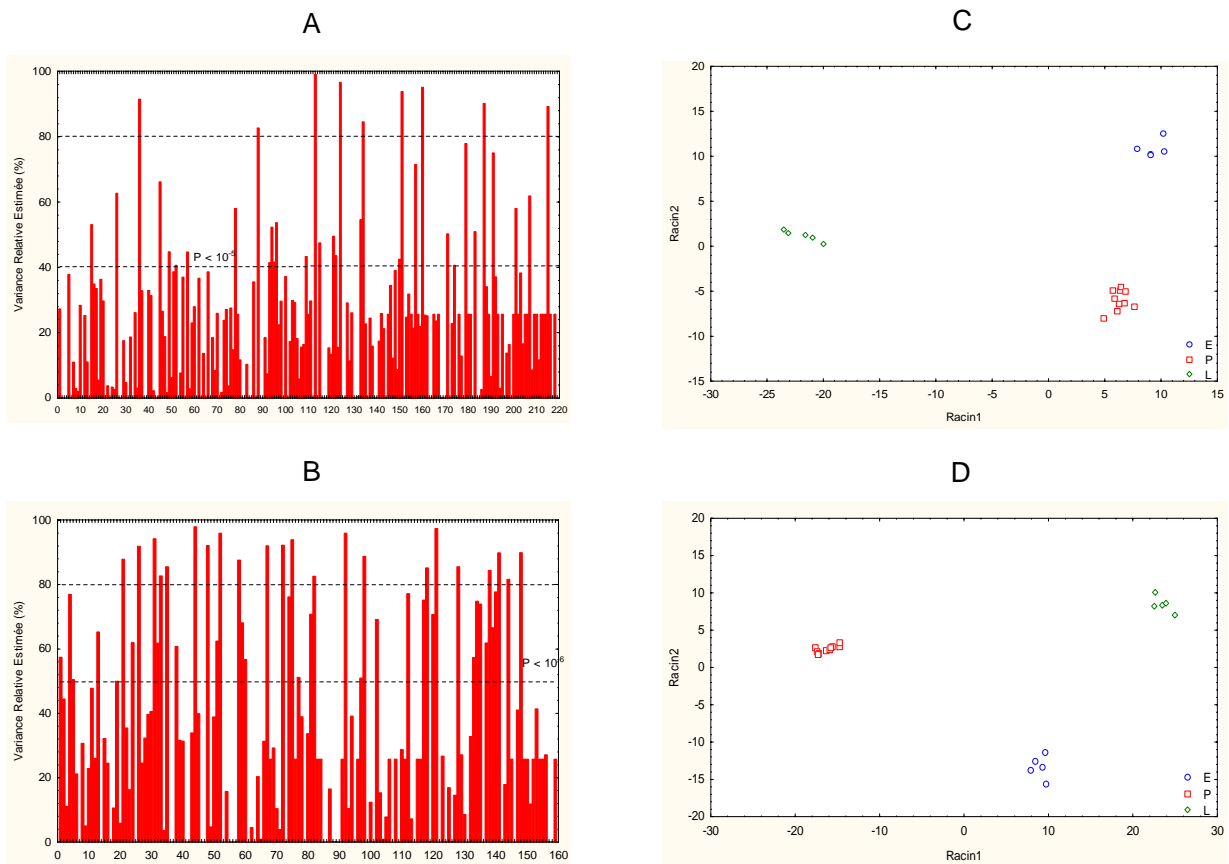
- & Cossart, P.** (2001). Comparative genomics of *Listeria* species. *Science*. 294, 849-852.
7. **Grigg, J. C., Vermeiren, C. L., Heinrichs, D. E. & Murphy, M. E.** (2007). Haem recognition by a *Staphylococcus aureus* NEAT domain. *Mol Microbiol*. 63, 139-149.
  8. **Guinane, C. M., Cotter, P. D., Ross, R. P. & Hill, C.** (2006). Contribution of penicillin-binding protein homologs to antibiotic resistance, cell morphology, and virulence of *Listeria monocytogenes* EGDe. *Antimicrob Agents Chemother*. 50, 2824-8.
  9. **Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. & Tauxe, R. V.** (1999). Food-related illness and death in the United States. *Emerg Infect Dis*. 5, 607-625.
  10. **Meunier, B., Dumas, E., Picc, I., Bechet, D., Hebraud, M. & Hocquette, J. F.** (2007). Assessment of Hierarchical Clustering Methodologies for Proteomic Data Mining. *J Proteome Res.* 2007, 6 (1), 358-366.
  11. **Nelson, K. E., Fouts, D. E., Mongodin, E. F., Ravel, J., DeBoy, R. T., Kolonay, J. F., Rasko, D. A., Angiuoli, S. V., Gill, S. R., Paulsen, I. T., Peterson, J., White, O., Nelson, W. C., Nierman, W., Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Haft, D. H., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L. D., Uhlich, G. A., Bayles, D. O., Luchansky, J. B. & Fraser, C. M.** (2004). Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res*. 32, 2386-2395.
  12. **Rabilloud, T.** (1992). A comparison between low background silver diammine and silver nitrate protein stains. *Electrophoresis*. 13, 429-439.
  13. **Vaillant, V., de Valk, H., Baron, E., Ancelle, T., Colin, P., Delmas, M. C., Dufour, B., Pouillot, R., Le Strat, Y., Weinbreck, P., Jouglu, E. & Desenclos, J. C.** (2005). Foodborne infections in France. *Foodborne Pathog Dis*. 2, 221-232.
  14. **Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J. & Kreft, J.** (2001). *Listeria*

pathogenesis and molecular virulence determinants. *Clin Microbiol Rev.* 14, 584-640.

15. **Wouters, J. A., Kamphuis, H. H., Hugenholtz, J., Kuipers, O. P., de Vos, W. M. & Abee, T.** (2000). Changes in glycolytic activity of *Lactococcus lactis* induced by low temperature. *Appl Environ Microbiol.* 66, 3686-3691.



**Figure 1 : Hierarchical clustering.** Clustering of the 20 gels from the 4 serovar 4b *Listeria monocytogenes* strains and of the intracellular (a) and extracellular (b) reliable protein spots, created with Permut Matrix. Red or green of increasing intensity indicate increasing or decreasing protein expression, respectively.



**Figure 2: Estimated Relative Variances (ERV) for the origin:** The ERV were calculated for the origin factor from the reliable data matrix resulting from the analysis of intracellular (A) and extracellular proteomes (B). The histograms give the value of the ERV (%) for each spot.

**Factorial discriminant analysis (FDA) for the intracellular and extracellular proteomes.** A FDA was carried out with the the 9 most representative spots for the intracellular proteome (C) and the 10 most representative spots for the extracellular proteome (D). Blue circles: environmental strain; red squares: asymptomatic carriage strains; green diamonds: clinical strains).

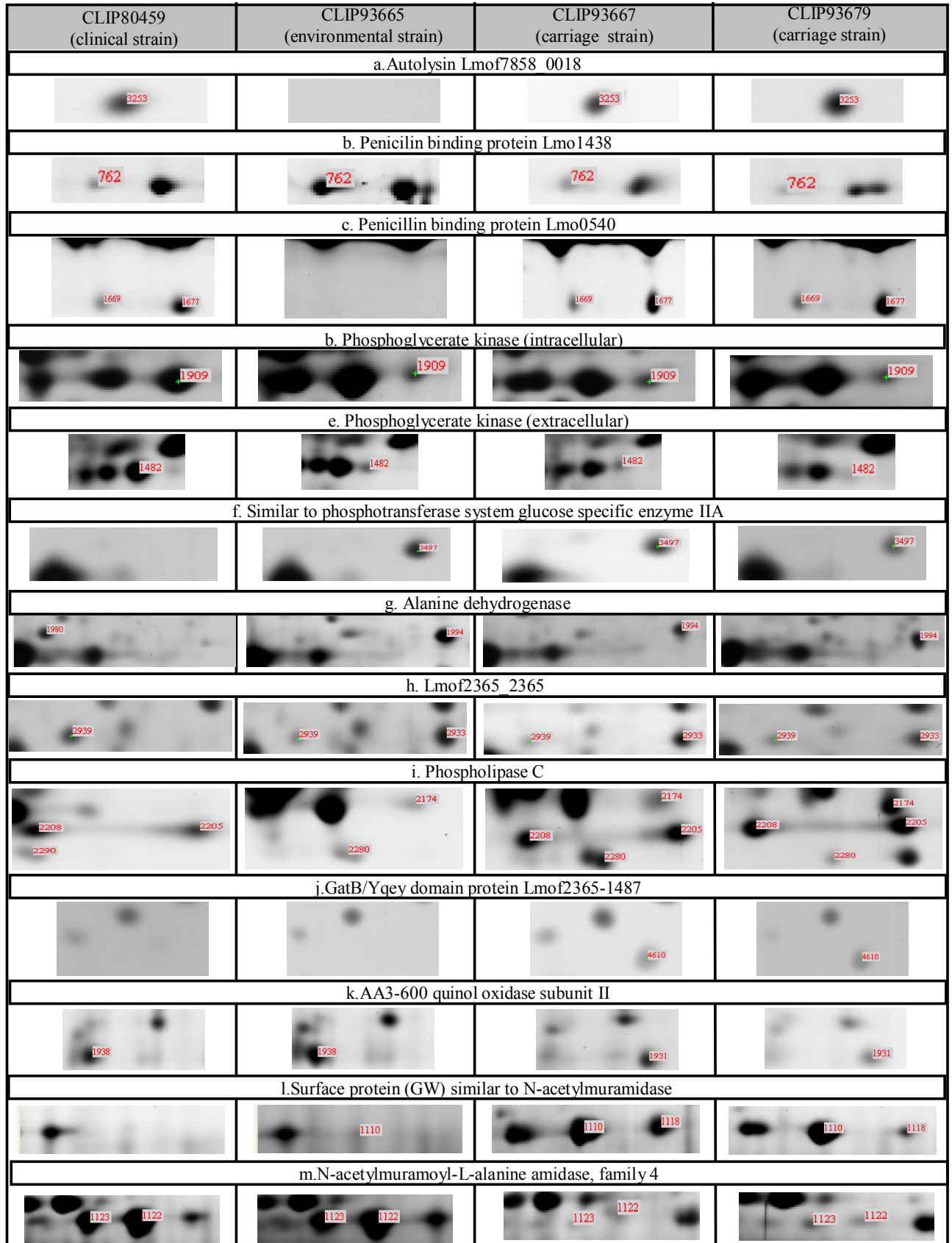


Figure 3: zooms of some differentially expressed proteins in 2-DE patterns.



**Table 1 : Proteins identified by MALDI-TOF mass spectrometry**

Group ID	Localization	Vre (%)	Swissprot accession number	Ncbi accession number	Gene name	Description	Strain	Mascot score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)
									Theoretical	Observed	Theoretical	Observed		
<b>spots over-expressed in clinical strain</b>														
1980	intracellular	99	Q8Y6V2	gi 16803619	lmo1579	alanine dehydrogenase	[ <i>Listeria monocytogenes</i> ]	170	39572	41500	5.24	5.3	16	51
1909		97	Q71WW8	gi 46908631	lmof2365_2431	phosphoglycerate kinase	[ <i>Listeria monocytogenes</i> str. 4b F2365]	136	42049	44500	5.02	5	11	39
2939		75	Q71X34	gi 46881837	lmof2365_2365	hypothetical protein (epimerase)	[ <i>Listeria monocytogenes</i> str. 4b F2365]	142	22809	28580	5.51	5.3	11	60
2290	extracellular	85	P33378	gi 16802251	plcB (lmo0205)	phospholipase C	[ <i>Listeria monocytogenes</i> EGD-e]	92	33256	22600	8.37	7	13	43
1482		96	Q4EEP3	gi 47094183	lmoh7858_2607	phosphoglycerate kinase	[ <i>Listeria monocytogenes</i> str. 4b H7858]	78	42050	45000	4.97	4.8	9	36
1665		90	Q4EIF4	gi 46906785	lmof2365_0569	penicillin-binding protein, putative	[ <i>Listeria monocytogenes</i> str. 4b H7858]	92	44364	35600	9.24	5.5	7	21
<b>spots under-expressed in clinical strain</b>														
2933	intracellular	95	Q71X34	gi 46881837	lmof2365_2365	hypothetical protein (epimerase)	[ <i>Listeria monocytogenes</i> str. 4b F2365]	102	22809	28600	5.51	5.5	9	47
1994		90	Q71Z88	gi 46881079	lmof2365_1601	alanine dehydrogenase	[ <i>Listeria monocytogenes</i> str. 4b F2365]	170	39572	41500	5.24	5.3	16	51
2117		85	Q8Y897	gi 16803057	lmo1017	similar to phosphotransferase system glucose-specific enzyme IIA	[ <i>Listeria monocytogenes</i> EGD-e]	reference map	17671		5.02			
3497		83	Q8Y897	gi 16803057	lmo1017	similar to phosphotransferase system glucose-specific enzyme IIA	[ <i>Listeria monocytogenes</i> EGD-e]	reference map	17671	18000	5.02	5		
2174	extracellular	62	P33378	gi 16409570	plcB (lmo0205)	phospholipase C	[ <i>Listeria monocytogenes</i> ]	72	33256	24700	8.37	7.7	9	36
2280		57	P33378	gi 16802251	lmo0205	phospholipase C	[ <i>Listeria monocytogenes</i> ]	87	33172	22600	8.37	7.4	13	41
887		75	Q4ELE1	gi 47091854	lmoh7858_2319	conserved hypothetical protein	[ <i>Listeria monocytogenes</i> str. 4b H7858]	151	63322	64100	6.07	5	21	47
<b>spots over-expressed in environmental strain</b>														
2126	intracellular	94	Q8Y6T3	gi 16803639	ccpA (lmo1599)	Catabolite control protein A	[ <i>Listeria monocytogenes</i> EGD-e]	reference map	36983	37000	5.05	5.2		
762	extracellular	92	Q8Y763	gi 16803478	lmo1438	similar to penicillin-binding protein	[ <i>Listeria monocytogenes</i> EGD-e]	155	79867	69100	7.53	5	23	34

Group id	Localization	Vre (%)	Swissprot	Ncbi accession number	Gene name	Description	Strain	Mascot score	Mr (Da)		pI		Number of mass matched	Sequence coverage
									Theoretical	Observed	Theoretical	Observed		
<b>spots under-expressed in environmental strain</b>														
2208	extracellular	88	P33378	gi 16802251	plcB (lmo0205)	phospholipase C	[ <i>Listeria monocytogenes</i> str. 4b H7858]	84	33256	23300	8.37	7	11	29
3253		69	Q4EHT4	gi 47093027	lmoh7858_0018	autolysin, putative	[ <i>Listeria monocytogenes</i> str. 4b H7858]	40	12608	9600	9.13	8	4	18
1677		86	Q4EIF4	gi 46906785	lmof2365_0569	penicillin-binding protein, putative	[ <i>Listeria monocytogenes</i> str. 4b F2365]	86	44364	35600	9.24	7.8	12	24
1669		84	Q4EIF4	gi 46906785	lmof2365_0569	penicillin-binding protein, putative	[ <i>Listeria monocytogenes</i> str. 4b F2365]	71	44364	35600	9.24	7.2	9	20
<b>spots over-expressed in asymptomatic carriage strains</b>														
4610	intracellular	89	Q71ZK2	gi 46907696	lmof2365_1487	GatB/Yqey domain protein	[ <i>Listeria monocytogenes</i> str. 4b F2365]	1,7	16840	14600	6	5.6	6	46
2111				Q71Z56	gi 46881111	lmof2365_1633	putative glutamyl-aminopeptidase	<i>Listeria monocytogenes</i> str. 4b F2365	118	38738	37200	5.69	5.3	9
1931	extracellular	89	Q4EML9	gi 47097689	lmof6854_0015	quinol oxidase AA3, subunit II	<i>Listeria monocytogenes</i> str. 1/2a F6854	104	41557	33000	6.42	5.4	6	20
1118		86	Q8Y464	gi 16804629	lmo2591	surface protein (GW repeat) similar to N-acetylmuramidase	[ <i>Listeria monocytogenes</i> EGD-e]	266	57004	57000	6.11	5.7	27	50
1110		83	Q8Y464	gi 16804629	lmo2591	surface protein (GW repeat) similar to N-acetylmuramidase	[ <i>Listeria monocytogenes</i> EGD-e]	158	57004	56900	6.11	5.5	14	27
<b>spots under-expressed in asymptomatic carriage strains</b>														
3682	intracellular	72	Q71ZK2	gi 46880965	lmof2365_1487	GatB/Yqey domain protein	[ <i>Listeria monocytogenes</i> str. 4b F2365]	73	16824	15200	5.96	6	4	33
1938	extracellular	92	Q8YAV0	gi 16802061	qoxA (lmo0013)	AA3-600 quinol oxidase subunit II	<i>Listeria monocytogenes</i>	64	41558	33000	6.04	5.2	7	23
1122		94	Q71WI7	gi 46908763	lmof2365_2564	N-acetylmuramoyl-L-alanine amidase, family 4	[ <i>Listeria monocytogenes</i> str. 4b F2365]	170	57025	56800	5.8	5.2	16	37
1123		94	Q71WI7	gi 46908763	lmof2365_2564	N-acetylmuramoyl-L-alanine amidase, family 4	[ <i>Listeria monocytogenes</i> str. 4b F2365]	258	57025	56400	5.8	5.1	23	51



## Article n°4

Dumas, E., Chambon, C., Desvaux, M. and Hébraud M.

Proteomic analysis of *Listeria monocytogenes* extracellular proteins

## **Proteomic analysis of *Listeria monocytogenes* extracellular proteins**

Dumas, E.<sup>1</sup>, Desvaux, M.<sup>1</sup>, Chambon, C.<sup>2</sup>, and Hébraud M.<sup>1,2\*</sup>

Institut National de la Recherche Agronomique, Centre de Recherche de Clermont-Ferrand/Theix, UR454 Microbiologie - Equipe Qualité et Sécurité des Aliments<sup>1</sup>, Plate-Forme de Protéomique<sup>2</sup>, F-63122 Saint-Genès Champanelle, France.

Keywords: *Listeria monocytogenes*, extracellular proteome, MALDI-TOF mass spectrometry, bioinformatic analysis, biodiversity.

\*Corresponding author: Tel.: +33 4 73 62 46 70; E-mail address: [hebraud@clermont.inra.fr](mailto:hebraud@clermont.inra.fr)

## **Abstract**

*Listeria monocytogenes* is a Gram-positive facultative intracellular foodborne pathogen at the origin of a rare but serious disease called listeriosis. The capacity to secrete proteins in the extracellular environment is of crucial importance for bacteria in general and more particularly for the pathogenesis process of *L. monocytogenes*. To get an overview of the protein set secreted by this bacteria, extracellular protein samples from different strains of *L. monocytogenes* were separated by two-dimensional gel electrophoresis (2-DE) and identified by MALDI-TOF mass spectrometry. A total of 167 protein spots corresponded to 77 distinct proteins were identified. A bioinformatic approach was performed on these proteins to analyse the presence of motifs giving their putative localization. 98.7% of the identified proteins were predicted to be translocated and 78% to be secreted in the extracellular milieu. Protein secretion would rely essentially on the Sec (Secretion) pathway. On the other hand, we did not identify proteins secreted by the twin-arginine translocation pathway, the holins and the WXG100 secretion system. However, 22 out of the 77 proteins seemed to be secreted by unknown system, among which 18 proteins were also predicted to be cytoplasmic. Indeed, 44% of the proteins were predicted to have multicellular localization, and may have different functions according to their localization.

## **1 - INTRODUCTION**

*Listeria monocytogenes* is a Gram-positive bacterium that can be isolated from a large diversity of biotopes including soil, water, vegetables, human and animal faeces, foodstuffs and food industrial environments. It is one of the major foodborne pathogenic bacteria and the causative agent of listeriosis, a very serious localized infection which can be lethal, particularly for high-risk groups such as elderly, pregnant women, neonates, and immunocompromised individuals. A series of epidemic outbreaks in the last two decades has initiated research efforts to determine the role of foods and food processing in the epidemiology. The raw materials are extensively contaminated and, consequently, so are the food production plants and foodstuffs. Obviously, these bacteria are able to survive hostile environments and stress conditions such as those encountered in food-processing technologies. They tolerate high salt concentrations (up to 15%), a wide range of pH (from 4.5 to 9.6) and temperatures (from 0 to 45°C), and low water availability ( $A_w$  down to 0.91–

0.93). Therefore, it is very important to understand the mechanisms of survival and adaptation of *L. monocytogenes*, as well as the molecular mechanisms underlying its virulence.

Bacterial pathogenicity depends greatly on the ability to secrete virulence factors which are displayed on the bacterial cell surface, secreted into the extracellular milieu or injected directly into the host cell. The major steps of the intracellular parasitism of *L. monocytogenes* are known and the pathogenicity of *L. monocytogenes* is associated with the production of defined virulence factors, most of which are either surface-associated or secreted in the extracellular milieu. Among the proteins involved in the infection cycle, the cell wall proteins internalin A and internalinB permits adhesion to the surface of the eukaryote cell and penetration into the host cell via phagocytosis. The secreted listeriolysin O and phospholipases permit escape from the phagocytic vacuole and the membrane anchored ActA protein permits actin-based motility and cell to cell spread (for review see (Vazquez-Boland *et al.*, 2001)).

Also, bacteria secrete numerous proteins which are essential for environment colonization. Whereas the importance of protein secretion, very little is known about the secretion system and the protein secreted by *L. monocytogenes*. Trost (Trost *et al.*, 2005) has realized a bioinformatic study of the *L. monocytogenes* EGDe genome and has predicted that 121 proteins were secreted in the extracellular milieu. This study is based on the presence of a signal peptide and the absence of a retention domain to the surface. So this is only a prediction of the extracellular proteins secreted via the major secretion apparatus in *L. monocytogenes* : the Sec pathway. The availability of the complete genome sequences of *L. monocytogenes* 1/2a EGDe, 4b F2365 and *L. innocua* CLIP11262 ((Glaser *et al.*, 2001 ; Nelson *et al.*, 2004)), as well as the unfinished genome sequences of *L. monocytogenes* 1/2a F6854 and H7858 allowed to Desvaux and Hébraud ((Desvaux & Hébraud, 2006)) to realize a most complete summary of the protein secretion systems present by combining bibliographical and genomic analyses. Finally, five other protein secretion systems could be found in *Listeria*. In addition to the Sec system, a Tat pathway, a flagella export apparatus (FEA), a fimbriin-Protein export (FPE), the holins, and a Wss, all of which permit protein translocation across the cytoplasm membrane could be identified. Twenty three proteins have been predicted to be substrates of these five other secretion systems in *L. monocytogenes* EGDe. After secretion, these proteins can be anchored in the cytoplasmic membrane, to the cell wall or secreted into the extracellular milieu.

Extracellular proteomic analyses of the proteins found in the extracellular milieu of *L. monocytogenes* EGDe (Trost *et al.*, 2005) permitted the identification of 105 extracellular

proteins, among which only 54 belonged to the 117 extracellular proteins predicted with a signal peptide from bioinformatic analyses. Diversity between the species *L. monocytogenes* and *L. innocua* was important in the secretome. It is well known that genome differences do not only exist between different species of *Listeria* and important diversity seems to exist inside the species *L. monocytogenes*, ((Doumith *et al.*, 2004) (Nelson *et al.*, 2004)). So it is important to study other strains than the reference strain EGDe. In this paper, we report and analyze the extracellular proteins identified during study of the biodiversity of *L. monocytogenes*, where 12 strains of different serovars have been compared.

## 2 - MATERIALS AND METHODS

### 2-1. Strains and bacterial culture

*L. monocytogenes* strains analyzed in this study were obtained from Pasteur Institute (Paris) and are listed in Table 1.

After two precultures in brain-heart infusion, cells in stationary phase were transferred to MCDB202, a chemically defined medium (CryoBiosystem, L'Aigle, France), supplemented with 0.36% glucose and adjusted at pH 7.3. Just before use, 0.1% (v/v) of elements traces and 1% (v/v) of a solution of 10% yeast nitrogen base were added to the MCDB medium. Two hundred and fifty ml cultures were adjusted to an  $OD_{600} = 0.1$  and incubated at 37°C with shaking (150 rpm). The bacterial cells were harvested by centrifugation (15 min, 7 500 g, 4°C) at late exponential phase ( $OD_{600} = 0.9$ , 3 to 4 h).

### 2-2. Precipitation of extracellular proteins

The supernatants were filtered on 0.2 µm membranes and 0.2 mM PMSF was added to inhibit proteases activity. 0.2 mg/ml Na deoxycholate was added to the solution which was incubated 30 min on ice. Na deoxycholate supports proteins precipitation that was carried out by addition of 10% TCA, and incubation overnight at 4°C. After centrifugation (20 300 g, 30 min, 4°C), the precipitate was washed with ice cold acetone and solubilized in isoelectric focusing (IEF) buffer (5 M urea, 2 M thiourea, 2% CHAPS, 10Mm TrisHCl, in 50% TFE and trace of bromophenol blue).



### *2-3. Two-dimensional gel electrophoresis (2-DE)*

For isoelectric focusing, non-linear pH 3-10 IPG strips were passively rehydrated during 17h30 in a reswelling tray with 400 µl of IEF buffer containing 0.3% v/v ampholytes 3-10, 2 mM TBP and 500 µg of proteins. Proteins were first subjected to isoelectric focusing (IEF) for a total of 66 450 Vhs (7 h at 50 V, 2 h at 200 V, linear gradient to 1 000 V in 2 h, 1 h at 1 000 V, linear gradient to 8 000 V in 5 h and 8 000 V till the end). The strips were equilibrated twice for 15 min in an equilibration solution (50 mM TrisHCl, pH 8.8; 6 M urea; 2% w/v SDS; 30% v/v glycerol) containing 2 mM TBP for the first step and 2.5% w/v iodoacétamide and trace of bromophénol blue for the second step. The second dimension (SDS-PAGE) was carried out with 12.5% acrylamide gel in a Multicell Protean II XL system (Bio-Rad, Marnes-la-Coquette, France). Gels were stained with Coomassie blue, scanned by a GS-800 imaging densitometer (Bio-Rad) and analyze by using Image Master 2D Platinum software v.5.0 (GE Healthcare, Uppsala, Sweden).

### *2-4. Identification of proteins by MALDI-TOF mass spectrometry*

Protein spots separated by 2-DE stained with Coomassie blue were excised. The gel pieces were destained and submitted to tryptic digestion. First, spots were washed with the destaining solution (25 mM ammonium bicarbonate / 5% acétonitrile) and two times with ammonium bicarbonate (25 mM) / 50% acétonitrile). Spots were then dehydrated with 100% acétonitrile. The dried gels were reswelled in a solution containing 20 µg trypsin and proteins in the gel were digested at least 5 h at 37°C. Resulting peptides were extracted with 100% acétonitrile. After 15 min at 37°C, 1 µl of each sample and 1 µl of saturated  $\alpha$ -cyano-4-hydroxycinnamic acid matrix were mixed onto the MALDI-TOF target. Positive ion MALDI mass spectra were recorded in the reflectron mode of a MALDI-TOF MS (Voyager DE-Pro, PerSeptive BioSystems, Farmingham, MA, USA) using voyager software for data collection and analysis. The MS was calibrated with a standard peptide solution (Proteomix, LaserBio Labs, Sophia-Antipolis, France). Internal calibration of samples was achieved by using trypsin autolysis peptides. Monoisotopic peptide masses were assigned and used for NCBI database searches with the “Mascot” software. The following parameters were considered for the searches: a maximum fragment ion mass tolerance of +/- 25 ppm, modification of cysteines by carbamidomethylation and possible oxidation of methionine.

## 2-5. Bioinformatic analyses

Bioinformatic analyses were performed from web-based servers or under Unix environment and Sun Grid Engine (SGE) from Topaze server homed at MIG (Mathématiques Informatique et Génomes) Research Unit (INRA, Jouy-en-Josas, France).

N-terminal signal peptide were predicted combining results from (i) SPScan, an implementation of von Heijne's weight matrix approach with McGeoch criteria where prompted parameters and optional parameter `-ADJustscores` were used (von Heijne, 1986) (McGeoch, 1985), (ii) SignalP v2.0 and v3.0 using both neural network (NN) and hidden Markov model (HMM) (Bendtsen *et al.*, 2004), (iii) Phobius based on a HMM combining transmembrane  $\alpha$ -helices topology and signal peptide prediction (Kall *et al.*, 2004), (iv) SOSUISignal based on tripartite signal peptide structure analysis (Gomi *et al.*, 2004), (v) PSORTb a support vector machine (SVM) (Cai *et al.*, 2003 ; Gardy *et al.*, 2005) and (vi) PrediSi (Hiller *et al.*, 2004). These tools were trained on bacteria with truncation disabled. Tat signal peptide prediction was performed from TatP v1.0 (Kuroda & Sekiguchi, 1991) and TATFIND v1.4 (Rose *et al.*, 2002). Pseudopilin-like signal peptides were searched with ScanProsite syntax (de Castro *et al.*, 2006) for consensus motif [AG]F[TS]LX[EF] located between the N- and H-domains as recently described in *Listeria* (Desvaux & Hebraud, 2006). Prediction of non-classical secreted proteins, *i.e.* lacking a signal peptide, was performed from SecretomeP v2.0 trained on bacteria (Bendtsen *et al.*, 2004).

For the identification of lipoproteins, sequences were submitted to DOLOP (Madan Babu & Sankaran, 2002), LipoP v1.0 (Juncker *et al.*, 2003), SPEPLip (Fariselli *et al.*, 2003), LipPred (Taylor *et al.*, 2006) and scanned by ScanProsite (de Castro *et al.*, 2006) with both PS51257 profile and G+LPP pattern, *i.e.* [MV]-X(0,13)-[RK]-{DERKQ}(6,20)-[LIVMFESTAG]-[LVIAM]-[IVMSTAFG]-[AG]-C (Sutcliffe & Harrington, 2002).

Modular architecture of proteins was analysed from COG v1.0 (Tatusov *et al.*, 2001) using RPS-BLAST (Reverse Position-Specific BLAST) (Altschul *et al.*, 1997) and using hidden Markov model (HMM) (Eddy, 1996) from InterPro v4.3 (Mulder *et al.*, 2007), Pfam v21.0 (Bateman *et al.*, 2004), SMART v5.1 (Schultz *et al.*, 1998), TIGRFam v6.0 (Selengut *et al.*, 2007), SCOP v1.71 (Murzin *et al.*, 1995) (Wilson *et al.*, 2007), PIRSF (Wu *et al.*, 2004) and Prosite v20.7 (Hulo *et al.*, 2006) databases. When needed, *e.g.* for no significant or inconclusive matches, PSI-BLAST (Position-Specific Iterated BLAST) (Altschul *et al.*, 1997) searches were executed against UniProt until convergence (Wu *et al.*, 2006). All searches were performed with *E*-value cutoff set at  $10^{-3}$ .

Transmembrane  $\alpha$ -helices were predicted combining (i) HMMTOP v2.0 (Tusnady & Simon, 2001), (ii) TMHMM v2.0 (Sonnhammer *et al.*, 1998), (iii) SOSUI v1.1 (Hirokawa *et al.*, 1998), (iv) THUMBUP v1.0 (Zhou & Zhou, 2003), (v) UMDHMMTMHP v1.0 (Zhou & Zhou, 2003) and (vi) MEMSAT v3.4 (Jones, 2007).

These analyses were completed with results from SVMs predicting subcellular localization of proteins, *i.e.* SubLoc v1.0 (Hua & Sun, 2001) and LocTree (Nair & Rost, 2005) trained on prokaryotes, CELLO v2.5 (Yu *et al.*, 2006) and PSORTb v2.0.4 trained on Gram-positive bacteria (Gardy *et al.*, 2005).

### 3 - RESULTS

#### *3-1. Hypothetical 2-DE-map and experimental 2-DE gels of L. monocytogenes EGDe extracellular proteins*

In *Listeria*, a protein can be translocated across the cytoplasmic membrane by different secretion systems, namely. Sec, Tat, FPE, FEA, holins or Wss (Desvaux & Hebraud, 2006); However, protein translocation via one of these systems is no guarantee a protein is released into the extracellular milieu. Excluding proteins bearing cell wall binding motifs, transmembrane domains, lipoproteins, pseudopilins and flagellar components, an hypothetical 2-DE map was obtained using theoretical masses and pI of amino acid sequences of proteins predicted to be secreted by these different secretion pathways (Trost *et al.*, 2005 ; Desvaux & Hebraud, 2006) (Figure 1); post-translational modifications of secreted proteins, essentially cleavage of N-terminal signal peptide, were carefully taken into consideration. This virtual map showed that 77% of the proteins are localized between Mr 10 and 60 kDa. It further appeared that more than 95% of extracellular proteins have a pI comprised between 3.0 and 10.0 grouped into two clusters, with almost any proteins found between pI 7.0 and pI 8.0. Similar repartition was also observed in the hypothetical 2-DE map of whole-cell proteins content (Folio *et al.*, 2004). Depending on the protein secretion pathway involved, different distribution pattern could be observed. While the proteins secreted by the Sec-dependent proteins covers a wide range pI and Mr, *i.e.* from 3.7 to 10.7 and 3 to 212 kDa respectively, proteins secreted by other pathways are localized in part of the gel. The two proteins secreted by the Tat pathway are of pI 5.5 ( $\pm$  0.3) and of Mr 40 kDa ( $\pm$  2). Proteins secreted by holins are basic and of approximately 30 kDa—The two proteins predicted as secreted by the Wss pathway are small proteins of 11 kDa and pI 4.6 and 9.2.

This subproteome was subsequently experimentally analyzed by 2-DE. As 95% of proteins predicted as secreted by *L. monocytogenes* EGDe have a *pI* comprised between 3 and 10, a non linear pH 3-10 IPG strips was used for IEF separation in the first dimension. Proteins present in culture supernatant originated from bacteria grown until late exponential phase in chemically defined MCDB medium at 37°C. As expected from hypothetical 2-DE map, two clusters of proteins could be distinguished. The majority of the protein spots were localized in a region between pH 4.0 and 6.0, but some very abundant spots were present in the basic region. This observation was the same for all strains, although some differences exist between the strain patterns.

### 3-2. The experimental extracellular proteome of other *L. monocytogenes* strains

The diversity of extracellular proteomes of *L. monocytogenes* strains was investigated by analysing a panel of 11 additional strains (Table 1). The strains chosen belonged to the three most representative serovars involved in human listeriosis cases (1/2a, 1/2b, 4b) and were isolated from different origins (food plant environment, human asymptomatic carriage, clinical cases). Their level of virulence, based on mortality of infected chick embryos (Olier *et al.*, 2003), revealed virulent strains (level A, 100% mortality within 3 days) and slightly attenuated strains (level B, 80% mortality in more than 3 days).

A total of 167 spots, numbered on a reference 2-DE gel (Figure 2) carried out from a mixture of an equivalent quantity of proteins of the 12 strains, were thus identified by MALDI-TOF MS after Mascot interrogation against Firmicutes database (NCBIInr). Several proteins were resolved as two or more spots on the 2-DE, which could be caused by post-translational modifications. After manual inspections of the output to remove protein fragments and redundant proteins (*i.e.* amino acid sequences 100% identical), the 167 protein spots corresponded to 77 distinct proteins. As listed in Table 2, it further appeared that some of these proteins were orthologues and/or homeomorphic, and corresponding finally to 61 non-homologous proteins. In order to categorise the proteins here identified in culture supernatants of these different *L. monocytogenes* strains, bioinformatic analyses were carried out. The rationale behind this approach is presented in Figure 3, and summary of predictions are given in supplementary materials (Supplementary Table 1S in supporting information).

### **Signal peptide**

First, proteins were scanned for the presence of N-terminal signal peptides using a large variety of bioinformatic tools, namely SPScan, SignalP, Phobius, SOSUISignal, PSORTb and PrediSi as described in Material and Methods, which led to 16 individual predictions. 63 proteins were predicted as possessing a signal peptide by at least one tool, including 14 proteins predicted as such by all predictors. If at least half of the prediction tools gave a significant positive result, a protein was predicted as possessing a signal peptide, *i.e.* 46 proteins. When between 1 and 7 out of 16 predictors identified a signal peptide, *i.e.* 17 proteins, proteins were further checked for subcellular localization and function. Finally, 54 proteins were predicted as possessing a N-terminal signal peptide.

### **Secretion system**

While a protein bearing a N-terminal signal peptide is targeted to the membrane, it is not systematically translocated via the Sec translocon. Further *in silico* analyses were carried out in order to find and discriminate different types of signal peptide.

TatP and TATFIND searches failed to identify Tat signal peptide and thus potential protein translocated by this secretion system. Similarly, ScanProsite search for class 3 signal peptide motif as recently described in *Listeria* (Desvaux & Hebraud, 2006) could not identify type 4 prepilins potentially translocated by the FPE. The remaining proteins lacking a signal peptide were checked as substrates for alternative secretion systems. Using a database generated from recent review on protein secretion in *Listeria* (Desvaux & Hebraud, 2006), PSI-BLAST searches failed to identify protein substrates of FEA, holins and Wss. However, following SecretomeP search and literature survey, 22 additional proteins were predicted as secreted by unknown secretion system. Finally, 54 proteins predicted with a N-terminal signal peptide were most certainly translocated *via* the Sec translocon and 22 proteins lacking a signal sequence were predicted as secreted *via* unknown secretion system.

### **Cell envelope binding motif**

Presence of N-terminal signal sequence (*i.e.* containing N- and H-domains) unmistakably indicate a protein is targeted to cytoplasmic membrane but, contrary to what is sometimes wrongly assumed in the literature, it is no guarantee a protein is secreted into the extracellular milieu. Indeed, such a protein can either (i) remain associated to the cytoplasmic membrane, (ii) remain associated to the cell wall, or (iii) be released into the extracellular milieu. Therefore, proteins predicted as Sec-dependent were further scanned for the presence of class

2 signal peptide, *i.e.* presence of lipobox which includes invariably a cysteine residue located just after the signal peptide C-domain. Following DOLOP, LipoP, SPEPLip, LipPred and ScanProsite searches, a protein was identified as lipoprotein if at least 3 out of 6 individual predictions gave a significant positive result, *i.e.* 8 proteins.

Motif searches were also carried out from InterPro, Pfam, Smart, TIRGfam, SuperFamily, PIRsf and Prosite in order to identify cell wall retention motifs (Desvaux *et al.*, 2006). Two proteins were predicted with a C-terminal LPXTG domain (IPR001899, PF00746, TIGR01167;  $2.4 \times 10^{-7} \leq E\text{-values} \leq 9.7 \times 10^{-4}$ ) and thus would be covalently anchored to bacterial cell wall by sortases (Marraffini *et al.*, 2006). No YSIRK motif (IPR005877, PF04650, TIGR01168) could be identified within Sec-dependent signal peptides of these proteins; When present, this motif is systematically associated with a LPXTG motif and is required for efficient protein secretion (Bae & Schneewind, 2003). Four proteins were predicted with GW module with copies ranging between 1 and 4 (SSF82057;  $3.0 \times 10^{-23} \leq E\text{-values} \leq 1.7 \times 10^{-12}$ ). This motif allows non-covalent interaction with lipoteichoic acids, where the higher is the number of GW modules, the stronger is the attachment (Jonquieres *et al.*, 1999). A serine protease (47093253) bears 3 copies of a ChW motif (IPR006637, SM00728, PF07538;  $2.5 \times 10^{-16} \leq E\text{-values} \leq 5.6 \times 10^{-11}$ ). As GW, ChW contains a highly conserved Gly-Trp dipeptide motif originally described in *Clostridium acetobutylicum* where it was suggested to allow protein cell surface anchoring (Desvaux *et al.*, 2005) and to be involved in a novel extracellular macromolecular system. Five proteins were predicted with LysM domain found in 2 or 4 copies (IPR002482, PF01476, SM00257, SSF54106;  $3.7 \times 10^{-21} \leq E\text{-values} \leq 2.0 \times 10^{-12}$ ). LysM domains bind directly to peptidoglycan and binding is hindered by secondary cell wall polymers, supposedly lipoteichoic acids (Steen *et al.*, 2003). Two of them also exhibits homology with a Src homology-3 (SH3) domain (IPR003646, SSF50044, SM00287;  $4.1 \times 10^{-15} \leq E\text{-values} \leq 1.3 \times 10^{-14}$ ), which belongs to the SH3 clan (CL0010); it further appeared these domains were more specifically bacterial SH3 of Type 3 (IPR013247, PF08239;  $3.1 \times 10^{-20} \leq E\text{-values} \leq 2.1 \times 10^{-19}$ ). All proteins bearing cell wall binding domain(s) were predicted as secreted in a Sec-dependent manner.

### Cell wall degradation proteins

In addition to autolysin LMOh7858\_0018, a protein originally annotated as hypothetical, *i.e.* Lmo1521, would also be a N-acetylmuramoyl-L-alanine amidase (COG0860;  $E\text{-values} = 1.0 \times 10^{-48}$ ; IPR002508, PF01520;  $E\text{-values} = 3.6 \times 10^{-64}$ ), which further exhibits three SH3 domains of Type 3 ( $2.1 \times 10^{-24} \leq E\text{-values} \leq 1.8 \times 10^{-8}$ ). In addition, 4 muramidases involved

in flagellar assembly were uncovered (COG1705;  $3.0 \times 10^{-52} \leq E\text{-values} \leq 9.0 \times 10^{-41}$ ) (Popowska, 2004); 3 of them were GW proteins and one of them a LysM protein. Two orthologues exhibit a 3D domain (IPR010611, PF06725;  $E\text{-values} = 1.5 \times 10^{-35}$ ), which is a peptidase-like active site containing three conserved aspartate residues and involved in peptidoglycan turnover; These enzymes were also predicted as LysM proteins. Including Iap (also called P60 or CwhA), four proteins exhibited a NlpC/P60 domain (IPR000064, PF00877;  $3.5 \times 10^{-53} \leq E\text{-values} \leq 6.2 \times 10^{-51}$ ) characteristic of cell wall peptidases (Anantharaman & Aravind, 2003); Two of them have LysM domains. As active sites of all these cell wall degradation enzymes have affinity for cell wall components, they are most likely present within this subcellular compartment. Contrary to what is sometimes assumed though (Dussurget *et al.*, 2004), presence of such domains should not be considered as cell wall binding motifs *per se* since primary function of these enzymes are to cleave cell wall components following what they either find a new cleavage site or are released into the extracellular milieu. Such proteins were then predicted as localized in two final subcellular compartments, *i.e.* cell wall and extracellular milieu. All cell wall degradation enzymes were primarily predicted as secreted *via* Sec.

### Research of transmembrane $\alpha$ -helices

All proteins were further analysed for presence of transmembrane  $\alpha$ -helices using HMMTOP, TMHMM, SOSUI, THUMBUP, UMDHMMTMHP and MEMSAT. A protein was predicted as possessing transmembrane domain (TMD) if at least 3 out of 6 tools gave a significant positive result. In addition, presence of motifs (*e.g.* LPXTG), signal peptides and positions of TMD were carefully taken into consideration for prediction of integral membrane protein (IMP). On the first instance, 53 proteins were predicted as IMPs but excluding signal peptide H-domain and LPXTG region, IMPs were reestimated down to 13. Only 4 of them did not display Sec-dependent N-terminal signal peptide. Seven proteins displayed only one putative  $\alpha$ -helical TMD and the remaining ones were predicted as possessing at least 2 transmembrane  $\alpha$ -helices, and up to 5 in one of them, *i.e.* Lmo0927. Excluding their class 2 signal peptide H-domain, 3 out of the 8 predicted lipoproteins exhibited 2 TMDs, *i.e.* AA3-600 quinol oxidase subunit II (QoxA) orthologues, and would thus be IMP covalently attached to membrane long chain fatty acid.

### **Final predicted localization**

All remaining proteins were further submitted to 4 SVMs dedicated to bacterial subcellular localization. Four localizations were considered, *i.e.* cytoplasmic, membrane, cell wall and extracellular. Protein was predicted in a given subcellular compartment if at least 2 out of 4 tools predicted it. These results were combined to previous analyses for final prediction of subcellular. Following recommendations of Gene Ontology (GO) Consortium to unambiguously defined protein localization in a Gram-positive bacterium (Harris *et al.*, 2004), GO numbers were given, *i.e.* cytoplasm (GO:0005737), cytoplasmic membrane (GO:0005886), cell wall (GO:0005618) and extracellular milieu (GO:0005576). In final, 19 proteins were predicted as cytoplasmic with 18 also predicted as extracellular. Eighteen proteins were predicted as membrane associated and more exactly as intrinsic to membrane (GO:0031226) including 13 IMPs, *i.e.* integral to membrane (GO:0005887), and 5 lipoproteins without TMDs, *i.e.* anchored to membrane (GO:0046658). Fifteen proteins were predicted as cell wall localized with 13 also predicted in the extracellular milieu. Twenty five proteins were predicted as secreted *via* the Sec pathway and finally released into the extracellular milieu. Thirty five proteins were predicted with multiple final subcellular localization. Altogether 60 out of 77 proteins were clearly predicted as extracellular.

## **4 - DISCUSSION**

A large majority of the proteins that were identified in the extracellular medium were indeed predicted as translocated. In fact, only one protein, a ribosomal subunit interface protein was predicted as only cytoplasmic. Fifty four of the 77 proteins were predicted as possessing a N-terminal signal peptide. They were all predicted to be secreted via the Sec translocon. None of the proteins lacking a signal peptide were identified as substrate of the FEA, holins or Wss secretion systems ((Desvaux & Hebraud, 2006)). Whereas no proteins potentially secreted via holins could be identified in this study, it should be noted that the autolysin Lmo0129 from *L. monocytogenes* was identified in culture supernatant in another study ((Trost *et al.*, 2005)) and predicted as substrate of this secretion pathway ((Desvaux & Hebraud, 2006)). The different experimental procedures, in particular concerning the culture conditions and mass spectrometry technologies, could explain why this protein was not identified in our study.



As revealed by SecretomeP and literature surveys, 22 proteins were predicted as secreted by unknown secretion systems. Surprisingly enough, a protein homologous to stage V sporulation protein G (Matsuno & Sonenshein, 1999) was identified, knowing that *L. monocytogenes* do not sporulate. However, it has been already demonstrated that *L. monocytogenes* contains 17 supposedly "spore specific genes" by comparison with *Bacillus subtilis*, among which the sequence of SpoVG was the most similar between the 2 strains (Onyenwoke *et al.*, 2004). Moreover, SpoVG has already been identified in the intracellular proteome of *L. monocytogenes* (Folio *et al.*, 2004). This protein could have a role in stress resistance as its expression is induced by heat shock (Periago *et al.*, 2002) and repressed by cold shock or salt stress in *B. subtilis* (Graumann *et al.*, 1996). Numerous proteins predicted to be translocated by unknown secretion systems were proteins implicated in stress resistance. For example, 2 proteins previously described as cytoplasmic proteins involved in resistance to oxidative stress, the catalase and the superoxide dismutase (SOD), were identified. The catalase has also been shown to be secreted in *B. subtilis* to ensure complete protection of the bacterial pathogen in the host. Recently, the cytoplasmic *L. monocytogenes* SOD was shown to be secreted by the secA2 pathway in culture supernatant and in infected cells, where it becomes phosphorylated ((Archambaud *et al.*, 2006)). Also, the enolase and the chaperones DnaK and GroEL seems to be specifically transported by the secA2 pathway ((Lenz & Portnoy, 2002 ; Lenz *et al.*, 2003)). These cytoplasmic chaperones, but also GroES are known to participate to bacterial protein secretion and consequently interact with Sec system (Desvaux *et al.*, 2004), and have been identified in the extracellular environment of several bacteria (Bergonzelli *et al.*, 2006). Some cytoplasmic proteins involved in carbohydrate metabolism such as glyceraldehydes-3-phosphate-dehydrogenase or enolase are moonlighting outside the cell of *L. monocytogenes* as they exhibit plasmin(ogen)-binding activity ((Schaumburg *et al.*, 2004))

It was hypothesized that some surface-associated pneumococcal proteins usually considered as cytoplasmic, such as enolase or glyceraldehydes-3-phosphate-dehydrogenase could be released from lysed cells and scavenged by *S. pneumoniae* competent cells ((Bergmann *et al.*, 2001), Guiral). This phenomenon, which involves several cell wall hydrolases, was named allolysis. It was suggested to stabilize symbiosis with the host, exacerbate infection, contribute to biofilm accretion and ensure survival of subpopulation. As allolysis has not yet been studied in *L. monocytogenes*, this mechanism and its involvement in cytoplasmic protein release/cell surface display, would require further investigation. In total, among the 23 proteins that did not possess a signal peptide, identified in the extracellular

milieu, 4 proteins were predicted to be translocated to the membrane and 18 to the extracellular milieu. So, it seems that uncovered secretion systems exist in *L. monocytogenes*.

As sometimes wrongly assumed, presence of N-terminal signal sequence does not guarantee that the protein is secreted into the extracellular milieu. Searching for cell envelope binding motif allowed identifying proteins with GW motifs LysM motif or LPXTG motif (Desvaux *et al.*, 2006). The 4 proteins with a GW motif and the 5 proteins with LysM motif were implicated in cell wall degradation. As primary function of these enzymes are to cleave cell wall components following what they either find a new cleavage site or are released into the extracellular medium, it is normal to find them in the supernatant. GW is a motif allowing noncovalent cell surface anchoring identified for the first time in *L. monocytogenes* InlB. It was demonstrated that InlB could also be secreted in the extracellular milieu to interact with host receptor ((Beales, 2004)). This motif is found in multicopy and the higher is the number of GW modules, the stronger is the attachment to bacterial cell wall. Proteins exhibiting only one GW motif would not be surface attached. In our study, an autolysin (LMOh7858\_0018) that has only one GW motif was identified. The coding gene of this protein was only identified in the sequenced strain *L. monocytogenes* H7858.

It is interesting to note that a serine protease bears 3 copies of a ChW motif (Clostridial hydrophobic domain with a conserved W residue), a motif recently hypothesized as involved in protein cell wall anchoring (Nolling *et al.*, 2001). This is the first time a protein expressed in *L. monocytogenes* is reported as bearing such repeats. This protein was only present in serovar 4b strains and this result corroborates genomic observations ((Doumith *et al.*, 2004)). Also a protein have SH3 domain. These domains resemble GW domains but are unlikely to have the same function since proline-binding sites are blocked or destroyed in GW domains (Marino *et al.*, 2002).

Eighteen proteins were found to be membrane associated. This category gathered proteins predicted as membrane integrated proteins and lipoproteins. Seventy eight percent of membranes proteins identified in the supernatant had a signal peptide. Eight proteins were identified as lipoproteins. Lipoproteins are characterized by a specific signal peptide of class 2 exhibiting a conserved lipobox sequence; following translocation across the cytoplasmic membrane, such proteins are subjected to several post-translational modifications leading to protein anchoring to membrane long chain fatty acid. Despite being homologous to lipoprotein Lmo1068, LMOh2365\_1085 does not possess a cystein at position +1 respective to signal peptide cleavage site and thus could not be predicted as a lipoprotein. This could further explain presence of this orthologue in culture supernatant. While QoxA was not

predicted by new lipoprotein HMM of Baumgärtner *et al.* (2007), its surface retention is clearly not impaired by  $\Delta lgt$  mutation in *L. monocytogenes* (Baumgartner *et al.*, 2007). Present bioinformatic analysis corroborates this finding since QoxA was predicted with 2 TMDs in its mature part, making of this protein an IMP, *i.e.* integral to membrane (GO:0005887), rather than a protein solely anchored to membrane (GO:0046658). Interestingly, shedding of QoxA into culture supernatant was previously observed and occurs independently of lipidation by Lgt (prolipoprotein diacylglyceryl transferase) in *B. subtilis* (Antelmann *et al.*, 2001).

## 5 - CONCLUSION

This study represents a complete analysis of proteins found in the extracellular milieu of *L. monocytogenes*, with (i) searching of signal peptides, (ii) exploration of the different secretion pathways, (iii) searching of cell wall binding motif, (iv) looking for transmembrane  $\alpha$ -helices (v) and subcellular localization and function. This analysis predicted that 98.7% of the proteins identified are translocated and that 78% are secreted into the extracellular medium. 44% of the proteins were predicted with multiple final subcellular localizations. Indeed, some proteins reported as located in the cytoplasmic compartment, the membrane or anchored to the cell wall, were also predicted and identified in the extracellular milieu. Their potential moonlighting role at different localization remains to be elucidated. The pathway by which the proteins without peptide signal were secreted is unknown, and no proteins in our study were substrate of these known secretion pathways : (i) the Tat pathway, (ii) the FEA, (iii) the Wss, (iv) the holins. So, whereas these pathways are present in *L. monocytogenes* (Desvaux & Hebraud, 2006), most investigations would be necessary to prove that they are functional and to identify experimentally their protein substrates. Further investigation would be necessary to discover possible unknown secretion pathways and to better understand the role of the second SecA paralogue, *i.e.* SecA2.

## 6 - REFERENCE

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein

database search programs. *Nucleic Acids Res.* 25, 3389-3402.

**Anantharaman, V. & Aravind, L.** (2003). Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome Biol.* 4, R11.

**Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J. M. & Hecker, M.** (2001). A proteomic view on genome-based signal peptide predictions. *Genome Res.* 11, 1484-1502.

**Archambaud, C., Nahori, M. A., Pizarro-Cerda, J., Cossart, P. & Dussurget, O.** (2006). Control of *Listeria* superoxide dismutase by phosphorylation. *J Biol Chem.* 281, 31812-31822.

**Bae, T. & Schneewind, O.** (2003). The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing. *J Bacteriol.* 185, 2910-2919.

**Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C. & Eddy, S. R.** (2004). The Pfam protein families database. *Nucleic Acids Res.* 32, D138-141.

**Baumgartner, M., Karst, U., Gerstel, B., Loessner, M., Wehland, J. & Jansch, L.** (2007). Inactivation of Lgt allows systematic characterization of lipoproteins from *Listeria monocytogenes*. *J Bacteriol.* 189, 313-324.

**Beales, N.** (2004). Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: a review. *Comp. Rev. Food Sci. Food Safety* 3, 1, 120.

**Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S.** (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol.* 340, 783-795.

**Bergmann, S., Rohde, M., Chhatwal, G. S. & Hammerschmidt, S.** (2001). alpha-Enolase of *Streptococcus pneumoniae* is a plasmin(ogen)-binding protein displayed on the bacterial cell surface. *Mol Microbiol.* 40, 1273-1287.

**Bergonzelli, G. E., Granato, D., Pridmore, R. D., Marvin-Guy, L. F., Donnicola, D. & Corthesy-Theulaz, I. E.** (2006). GroEL of *Lactobacillus johnsonii* La1 (NCC 533) is cell surface associated: potential role in interactions with the host and the gastric pathogen

*Helicobacter pylori*. *Infect Immun*. 74, 425-434.

**Cai, Y. D., Lin, S. L. & Chou, K. C.** (2003). Support vector machines for prediction of protein signal sequences and their cleavage sites. *Peptides*. 24, 159-161.

**de Castro, E., Sigrist, C. J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P. S., Gasteiger, E., Bairoch, A. & Hulo, N.** (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res*. 34, W362-365.

**Desvaux, M., Dumas, E., Chafsey, I. & Hebraud, M.** (2006). Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiol Lett*. 256, 1-15.

**Desvaux, M. & Hebraud, M.** (2006). The protein secretion systems in *Listeria*: inside out bacterial virulence. *FEMS Microbiol Rev*. 30, 774-805.

**Desvaux, M., Khan, A., Scott-Tucker, A., Chaudhuri, R. R., Pallen, M. J. & Henderson, I. R.** (2005). Genomic analysis of the protein secretion systems in *Clostridium acetobutylicum* ATCC 824. *Biochim Biophys Acta*. 1745, 223-253.

**Desvaux, M., Parham, N. J., Scott-Tucker, A. & Henderson, I. R.** (2004). The general secretory pathway: a general misnomer? *Trends Microbiol*. 12, 306-309.

**Doumith, M., Cazalet, C., Simoes, N., Frangeul, L., Jacquet, C., Kunst, F., Martin, P., Cossart, P., Glaser, P. & Buchrieser, C.** (2004). New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. *Infect Immun*. 72, 1072-1083.

**Dussurget, O., Pizarro-Cerda, J. & Cossart, P.** (2004). Molecular determinants of *Listeria monocytogenes* virulence. *Annu Rev Microbiol*. 58, 587-610.

**Eddy, S. R.** (1996). Hidden Markov models. *Curr Opin Struct Biol*. 6, 361-365.

**Fariselli, P., Finocchiaro, G. & Casadio, R.** (2003). SPEPlip: the detection of signal peptide and lipoprotein cleavage sites. *Bioinformatics*. 19, 2498-2499.

**Folio, P., Chavant, P., Chafsey, I., Belkorchia, A., Chambon, C. & Hebraud, M.**

(2004). Two-dimensional electrophoresis database of *Listeria monocytogenes* EGDe proteome and proteomic analysis of mid-log and stationary growth phase cells. *Proteomics*. 4, 3187-3201.

**Gardy, J. L., Laird, M. R., Chen, F., Rey, S., Walsh, C. J., Ester, M. & Brinkman, F. S.** (2005). PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics*. 21, 617-623.

**Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Ndjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J. C., Purcell, R., Rimmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. & Cossart, P.** (2001). Comparative genomics of *Listeria* species. *Science*. 294, 849-852.

**Gomi, M., Sonoyama, M. & Mitaku, S.** (2004). High performance system for signal peptide prediction : SOSUISignal. *Chem-Bio Informatics journal*. 4, 142-147.

**Graumann, P., Schroder, K., Schmid, R. & Marahiel, M. A.** (1996). Cold shock stress-induced proteins in *Bacillus subtilis*. *J Bacteriol*. 178, 4611-4619.

**Harris, M. A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C., Richter, J., Rubin, G. M., Blake, J. A., Bult, C., Dolan, M., Drabkin, H., Eppig, J. T., Hill, D. P., Ni, L., Ringwald, M., Balakrishnan, R., Cherry, J. M., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S., Fisk, D. G., Hirschman, J. E., Hong, E. L., Nash, R. S., Sethuraman, A., Theesfeld, C. L., Botstein, D., Dolinski, K., Feierbach, B., Berardini, T., Mundodi, S., Rhee, S. Y., Apweiler, R., Barrell, D., Camon, E., Dimmer, E., Lee, V., Chisholm, R., Gaudet, P., Kibbe, W., Kishore, R., Schwarz, E. M., Sternberg, P., Gwinn, M., Hannick, L., Wortman, J., Berriman, M., Wood, V., de la Cruz, N., Tonellato, P., Jaiswal, P., Seigfried, T. & White, R.** (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 32, D258-261.

**Hiller, K., Grote, A., Scheer, M., Munch, R. & Jahn, D.** (2004). PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* 32, W375-379.

**Hirokawa, T., Boon-Chieng, S. & Mitaku, S.** (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics.* 14, 378-379.

**Hua, S. & Sun, Z.** (2001). Support vector machine approach for protein subcellular localization prediction. *Bioinformatics.* 17, 721-728.

**Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., Pagni, M. & Sigrist, C. J.** (2006). The PROSITE database. *Nucleic Acids Res.* 34, D227-230.

**Jones, D. T.** (2007). Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics.* 23, 538-544.

**Jonquieres, R., Bierne, H., Fiedler, F., Gounon, P. & Cossart, P.** (1999). Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of gram-positive bacteria. *Mol Microbiol.* 34, 902-914.

**Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. & Krogh, A.** (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* 12, 1652-1662.

**Kall, L., Krogh, A. & Sonnhammer, E. L.** (2004). A combined transmembrane topology and signal peptide prediction method. *J Mol Biol.* 338, 1027-1036.

**Kuroda, A. & Sekiguchi, J.** (1991). Molecular cloning and sequencing of a major *Bacillus subtilis* autolysin gene. *J Bacteriol.* 173, 7304-7312.

**Lenz, L. L., Mohammadi, S., Geissler, A. & Portnoy, D. A.** (2003). SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. *Proc Natl Acad Sci U S A.* 100, 12432-12437.

**Lenz, L. L. & Portnoy, D. A.** (2002). Identification of a second *Listeria* secA gene associated with protein secretion and the rough phenotype. *Mol Microbiol.* 45, 1043-1056.

**Madan Babu, M. & Sankaran, K.** (2002). DOLOP--database of bacterial lipoproteins.

*Bioinformatics*. 18, 641-643.

**Marino, M., Banerjee, M., Jonquieres, R., Cossart, P. & Ghosh, P.** (2002 ). GW domains of the *Listeria monocytogenes* invasion protein InlB are SH3-like and mediate binding to host ligands. *EMBO J.* 21, 5623-5634.

**Marraffini, L. A., Dedent, A. C. & Schneewind, O.** (2006). Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol Mol Biol Rev.* 70, 192-221.

**Matsuno, K. & Sonenshein, A. L.** (1999). Role of SpoVG in asymmetric septation in *Bacillus subtilis*. *J Bacteriol.* 181, 3392-3401.

**McGeoch, D. J.** (1985). On the predictive recognition of signal peptide sequences. *Virus Res.* 3, 271-286.

**Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Buillard, V., Cerutti, L., Copley, R., Courcelle, E., Das, U., Daugherty, L., Dibley, M., Finn, R., Fleischmann, W., Gough, J., Haft, D., Hulo, N., Hunter, S., Kahn, D., Kanapin, A., Kejariwal, A., Labarga, A., Langendijk-Genevaux, P. S., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Nikolskaya, A. N., Orchard, S., Orengo, C., Petryszak, R., Selengut, J. D., Sigrist, C. J., Thomas, P. D., Valentin, F., Wilson, D., Wu, C. H. & Yeats, C.** (2007). New developments in the InterPro database. *Nucleic Acids Res.* 35, D224-228.

**Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C.** (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol.* 247, 536-540.

**Nair, R. & Rost, B.** (2005). Mimicking cellular sorting improves prediction of subcellular localization. *J Mol Biol.* 348, 85-100.

**Nelson, K. E., Fouts, D. E., Mongodin, E. F., Ravel, J., DeBoy, R. T., Kolonay, J. F., Rasko, D. A., Angiuoli, S. V., Gill, S. R., Paulsen, I. T., Peterson, J., White, O., Nelson, W. C., Nierman, W. , Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Haft, D. H., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L. D., Uhlich, G. A.,**



**Bayles, D. O., Luchansky, J. B. & Fraser, C. M.** (2004). Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res.* 32, 2386-2395.

**Nolling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y. I., Tatusov, R. L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M. J., Bennett, G. N., Koonin, E. V. & Smith, D. R.** (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol.* 183, 4823-4838.

**Olier, M., Pierre, F., Rousseaux, S., Lemaitre, J. P., Rousset, A., Piveteau, P. & Guzzo, J.** (2003). Expression of truncated Internalin A is involved in impaired internalization of some *Listeria monocytogenes* isolates carried asymptotically by humans. *Infect Immun.* 71, 1217-1224.

**Onyenwoke, R. U., Brill, J. A., Farahi, K. & Wiegel, J.** (2004). Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch ( Firmicutes). *Arch Microbiol.* 182, 182-192.

**Periago, P. M., van Schaik, W., Abee, T. & Wouters, J. A.** (2002). Identification of proteins involved in the heat stress response of *Bacillus cereus* ATCC 14579. *Appl Environ Microbiol.* 68, 3486-3495.

**Popowska, M.** (2004). Analysis of the peptidoglycan hydrolases of *Listeria monocytogenes*: multiple enzymes with multiple functions. *Pol J Microbiol.* 53 Suppl, 29-34.

**Rose, R. W., Bruser, T., Kissinger, J. C. & Pohlschroder, M.** (2002). Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol.* 45, 943-950.

**Schaumburg, J., Diekmann, O., Hagendorff, P., Bergmann, S., Rohde, M., Hammerschmidt, S., Jansch, L., Wehland, J. & Karst, U.** (2004). The cell wall subproteome of *Listeria monocytogenes*. *Proteomics.* 4, 2991-3006.

**Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P.** (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A.* 95, 5857-5864.

**Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., Richter, A. R. & White, O.** (2007). TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res.* 35, D260-264.

**Sonnhammer, E. L., von Heijne, G. & Krogh, A.** (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol.* 6, 175-182.

**Steen, A., Buist, G., Leenhouts, K. J., El Khattabi, M., Grijpstra, F., Zomer, A. L., Venema, G., Kuipers, O. P. & Kok, J.** (2003). Cell wall attachment of a widely distributed peptidoglycan binding domain is hindered by cell wall constituents. *J Biol Chem.* 278, 23874-23881.

**Sutcliffe, I. C. & Harrington, D. J.** (2002). Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiology.* 148, 2065-2077.

**Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., Kiryutin, B., Galperin, M. Y., Fedorova, N. D. & Koonin, E. V.** (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 29, 22-28.

**Taylor, P., Toseland, CP., Attwood, TK. & Flower, D.** (2006). LIPPred : a web server for accurate prediction of lipoprotein signal sequences and cleavage sites. *Bioinformatics.* 1, 176-179.

**Trost, M., Wehmhoner, D., Karst, U., Dieterich, G., Wehland, J. & Jansch, L.** (2005). Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species. *Proteomics.* 5, 1544-1557.

**Tusnady, G. E. & Simon, I.** (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics.* 17, 849-850.

**Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J. & Kreft, J.** (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clin Microbiol Rev.* 14, 584-640.

**von Heijne, G.** (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14, 4683-4690.

**Wilson, D., Madera, M., Vogel, C., Chothia, C. & Gough, J.** (2007). The SUPERFAMILY database in 2007: families and functions. *Nucleic Acids Res.* 35, D308-313.

**Wu, C. H., Apweiler, R., Bairoch, A., Natale, D. A., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Mazumder, R., O'Donovan, C., Redaschi, N. & Suzek, B.** (2006). The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* 34, D187-191.

**Wu, C. H., Nikolskaya, A., Huang, H., Yeh, L. S., Natale, D. A., Vinayaka, C. R., Hu, Z. Z., Mazumder, R., Kumar, S., Kourtesis, P., Ledley, R. S., Suzek, B. E., Arminski, L., Chen, Y., Zhang, J., Cardenas, J. L., Chung, S., Castro-Alvear, J., Dinkov, G. & Barker, W. C.** (2004). PIRSF: family classification system at the Protein Information Resource. *Nucleic Acids Res.* 32, D112-114.

**Yu, C. S., Chen, Y. C., Lu, C. H. & Hwang, J. K.** (2006). Prediction of protein subcellular localization. *Proteins.* 64, 643-651.

**Zhou, H. & Zhou, Y.** (2003). Predicting the topology of transmembrane helical proteins using mean burial propensity and a hidden-Markov-model-based method. *Protein Sci.* 12, 1547-1555.

Origin	Strain name	Serovar	Virulence
Human Clinical cases of listeriosis	CLIP80459	4b	A
	CLIP90602	1/2b	A
	CLIP92347	1/2a	A
Animal cases of listeriosis (sequenced strain)	EGDe	1/2a	A
Asymptomatical carriage	CLIP93667	4b	A
	CLIP93679	4b	A
	CLIP93672	1/2b	B
	CLIP93677	1/2a	B
Food plant environment	CLIP93665	4b	A
	CLIP93649	1/2a	A
	CLIP93663	1/2a	B
	CLIP93657	1/2a	B

Table 1: **strains of *L. monocytogenes* used in this study.** The level of virulence has been assessed in 14 days old chick embryos as previously described (Olier *et al.*, 2003). The results are either published (Olier *et al.*, 2003) or personal communication. Virulence A means infection of chick embryos resulted in 100% mortality within 3 days. Virulence B means infection of chick embryos resulted in 80% mortality in more than 3 days.

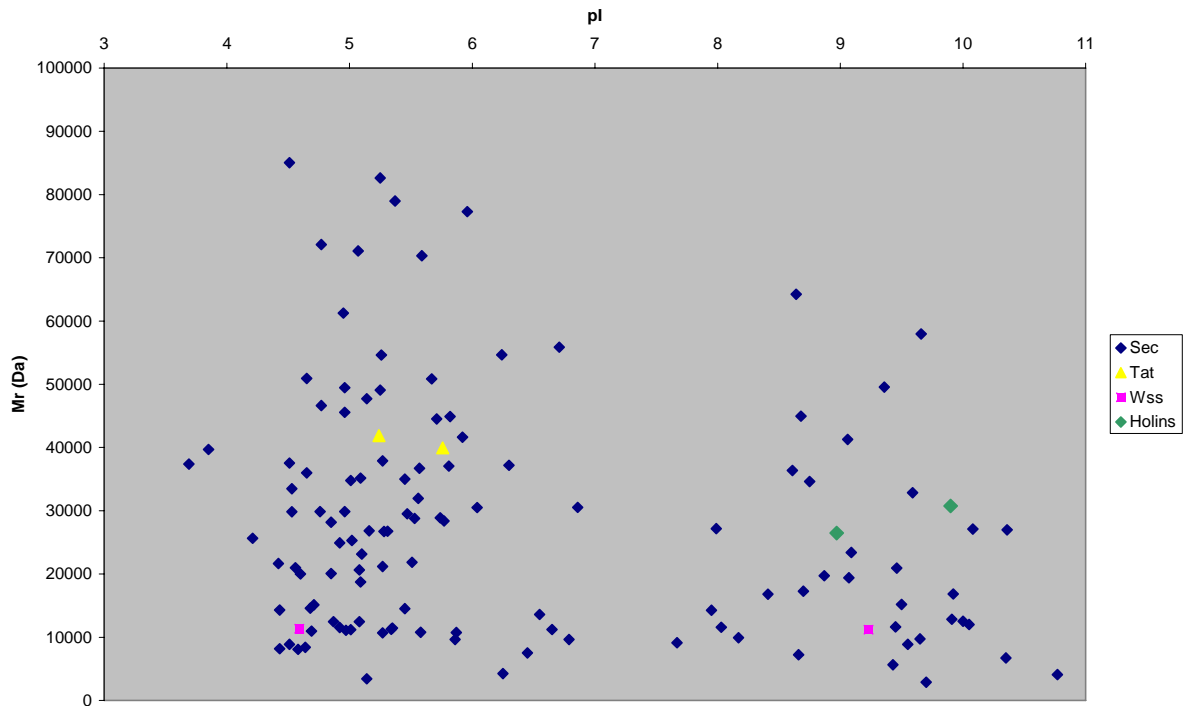


Figure 1: Hypothetical 2-DE map of *L. monocytogenes* EGDe proteins secreted in the extracellular milieu by the Sec pathway, the Twin-arginine translocation system (Tat), the holins and Wss.

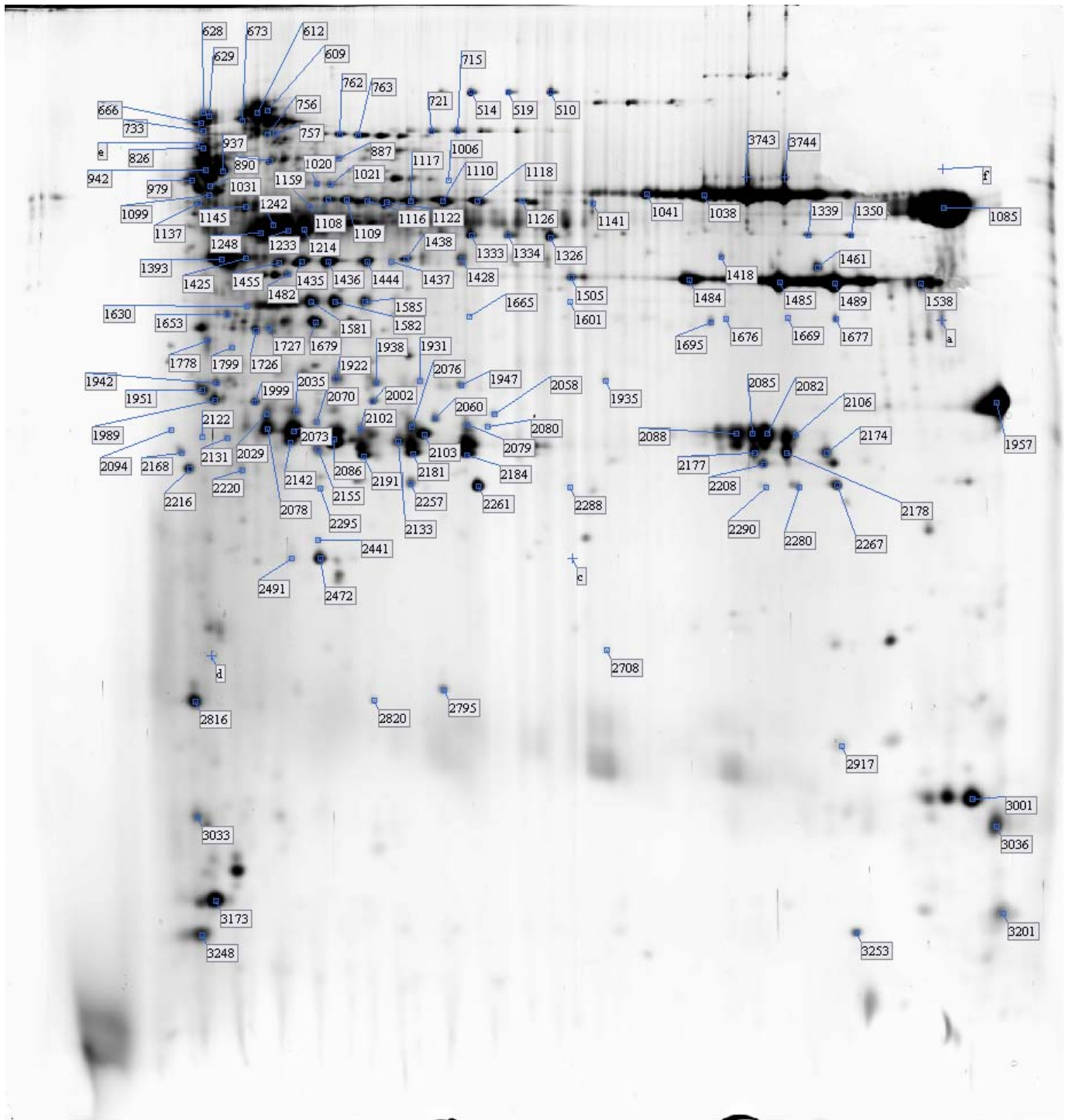


Figure 2: 2-DE pattern of a protein mixture composed with an equivalent quantity of extracellular proteins of the 12 strains studied. The spot numbers are listed in Table 2.

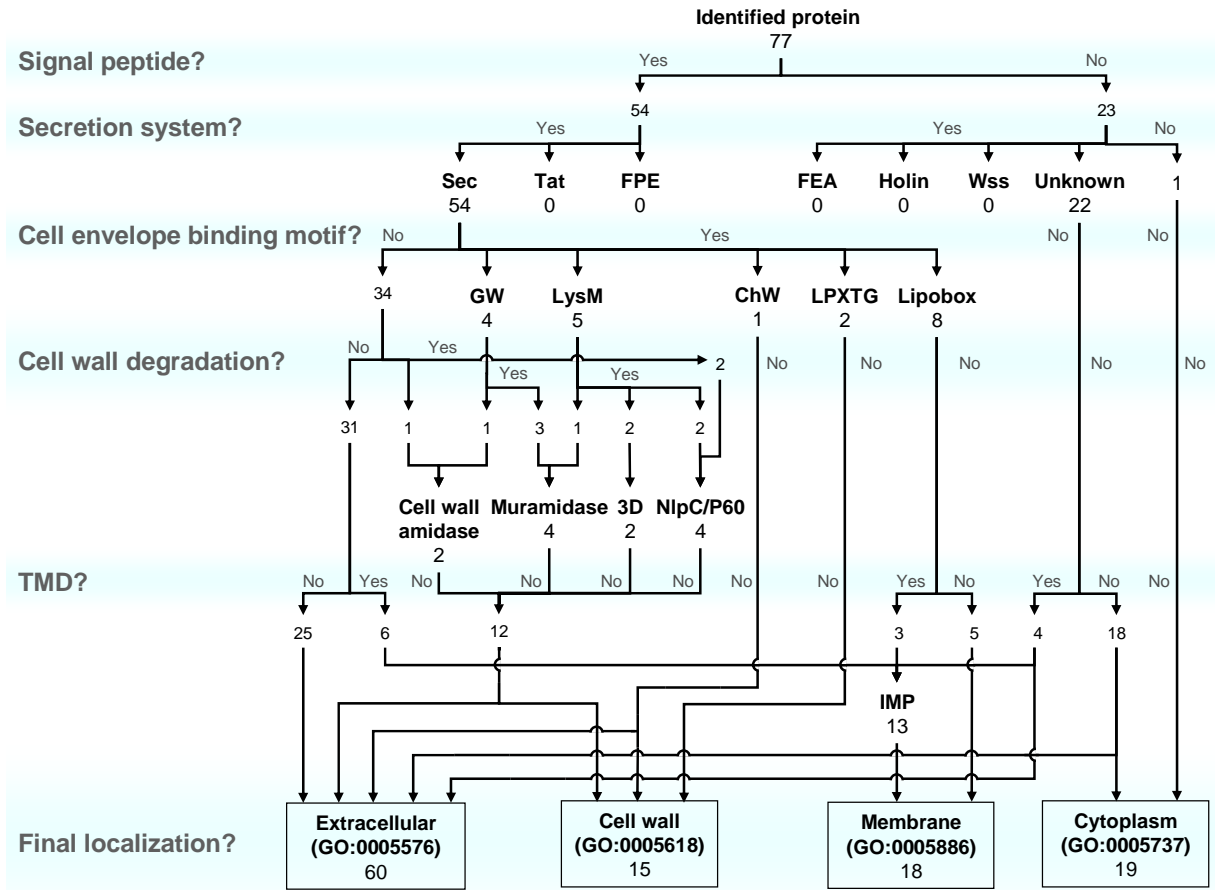


Figure 3: Schematic representation of the bioinformatic analyses carried out to categorise the 77 proteins identified in the extracellular milieu of *L. monocytogenes* strains.

Table 2: Proteins identified in the extracellular milieu by MALDI-TOF mass spectrometry  
**Proteins predicted in the extracellular milieu**

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
serine protease, putative	gi 47093253	lmof2365_1900	Q71YE5	1339	72	48745	51200	8,69	7,5	9	22	4,1	CLIP93667
				1350	70	48745	51200	8,69	7,9	7	18		CLIP93667
K+-transporting ATPase, C subunit	gi 46908859	kdpC (lmof2365_2660)	Q71W91	2790	73	20855		6,84		6	35	1.2	EGDe
				2795	48	20855	16000	6.84	5,3	6	47		CLIP80459
penicillin binding protein lmo2039 (pbp2B)	gi 46908275	lmof2365_2071	Q71XX4	715	1.24 (P)	81920	69600	6,6	5,6		13	1.1	CLIP80459
				721	79	81871	69600	6.63	5,4	13	25	1.1	CLIP90602
similar to penicillin-binding protein lmo0540	gi 46906785	lmof2365_0569	Q4EIF4	1669	71	44364	35600	9,24	7,2	8	18	1.1	CLIP80459
				1677	86	44364	35600	9,24	7,8	12	24		CLIP80459
				1665	92	44364	35600	9.24	5,5	7	21		CLIP90602
				1676	96	44364	35600	9,24	6,9	11	34		CLIP93665
	gi 16802583	lmo0540	Q8Y9I8	1695	78	44506	35000	9,23	6,8	10	30		EGDe
similar to penicillin-binding protein lmo1438	gi 16803478	lmo1438	Q8Y763	762	155	79867	68800	7,53	5	23	34	1.1	EGDe
				763	146	79867	68800	7,53	5,1	23	36		EGDe
				772	45	79867	68800	7.53	5,2	6	10	1.1	EGDe
putative secreted protein	gi 46906830	lmof2365_0614	Q722W5	2168	0,75 (P)	24310	24100	4,6	4,4	4	19	5.2	CLIP80459
	gi 16802628	lmo0585	Q8Y9E5	2094	74	24350	27600	4,57	4,3	6	37		EGDe
similar to penicillin-binding protein 2A	gi 46908125	lmof2365_1921	Q71YC3	510	94	90766	77500	8,62	6	15	16	1.1	CLIP93667
				519	44	90766	77700	8,62	5,8	6	19		CLIP80459
listeriolysin O precursor	gi 16802248	hly (lmo0202)	P13128	1461	94	58666	46700	7.63	7,6	14	32	4,5	CLIP93657
				1418	95	58666	48100	7.63	6,80	11	29		EGDe
				1038	211	58652	57400	7.63	6,8	23	49		EGDe
				1041	234	58652	57400	7,63	6,60	22	45		EGDe
				3743	161	58652	60700	7.63	6,9	18	40		EGDe
				3744	193	58652	60700	7.63	7,2	19	39		EGDe



protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain			
						theoretical	observed	theoretical	observed							
internalin C	gi 16803826	inlC (lmo1786)	Q8Y6A8	2155	63	33058	25100	6.12	4,9	7	26	1,9	EGDe			
				2191	172	33058	24200	6,12	6.12	13	44		EGDe			
				2181	157	33058	24500	6,12	5	15	50		EGDe			
				2184	44	33058		6,12		6	18		EGDe			
				2102	75	33058	28000	6.12	4,95	8	30		EGDe			
				2076	92	33058	28200	6.12	5,05	10	38		EGDe			
				2079	115	33058	24500	6,12	5,6	13	46		EGDe			
similar to B. subtilis transcription regulator LytR	gi 16802487	lmo0443	Q8Y9T0"	2257	101	34124	22600	9.05	5,05	13	32	3,5,2	CLIP93663			
				2261	141	34124	22600	9.05	5,5	13	39		EGDe			
similar to D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 5)	gi 4104145	lmo2754	Q9ZIC4	1437	100	46744	47300	5,67	5	12	34	1.1	CLIP90602			
hypothetical protein	gi 46907182	lmof2365_0970	Q721L6	2060	88	32466	29500	5,84	5,2	9	31	5.2	CLIP93667			
lipoprotein, putative	gi 46880563	lmof2365_1085	Q721A2	1935	88	30562	33000	8.45	6,4	6	37		CLIP80459			
similar to B. subtilis YqzC protein	gi 16803373	lmo1333	Q8Y7E9	3001	100	17083		8.81		11	59	5.2	CLIP90602			
				2917	96	17884		9,11		9	47		EGDe			
hypothetical protein lmo1752	gi 47092905	lmoh7858_1877	Q4EI55	2070	60	28495		5,47		6	16	6	CLIP93665			
				gi 16803792	lmo1752	Q8Y6D5	1999	131	28697	32000	5,1		4,7	10	32	EGDe
							2029	58	28697	30000	5.10		4,80	7	27	CLIP93657
similar to chitinase	gi 46908115	lmof2365_1912	Q71YD3	2237	66	32598		5.33		6	26	2.1.1	CLIP93657			
				1739	104	37907	35000	5,21	5,2	11	43		CLIP93667			
	gi 16803922	lmo1883	Q8Y6I9	1679	94	37897		5,21		8	30		EGDe			
Hypothetical protein	gi 16804195	lmo2156	Q8Y5B1	3201	80	13076	9900	9,6	10	6	38	5.1	EGDe			

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
peptidase, M48 family	gi 46908676	lmo2365_2477	Q71WS4	1334	69	47030	51100	6,31	5,7	8	25	1.8	CLIP93665
				1326	80	47004	51000	6.31	6	9	33		CLIP80459
	gi 16804542	lmo2504	Q8Y4E2	1333	74	47021	51000	6,02	5,5	9	23		CLIP93649
				514	1,9 (P)	47030	77000	6	5,5	7	23		EGDe
similar to B. subtilis putative transcriptional regulator LytR	gi 16804556	lmo2518	Q8Y4D2	1942	85	39094	33000	5.57	4,6	9	25	3.5.2	CLIP93663
				1989	65	39094	32200	5.57	4,6	9	24		EGDe
				1942	85	39094	33000	5.57	4,6	9	25		EGDe
phosphotidyl inositol-specific phospholipase C	gi 16802247	plcA (lmo0201)	P34024	1957	130	36295	32100	9.63	9,9	14	41	2.4	EGDe
phospholipase C	gi 16802251	plcB (lmo0205)	P33378	ID2280	87	33172	22600	8.37	7,4	13	41	2,4	CLIP93657
				2290	92	33256	22600	8,37	7	13	43		EGDe
				2267	92	33256	22600	8,37	7,8	12	34		EGDe
				2085	138	33256	27000	8.37	7	13	32		EGDe
				2106	84	33256	26800	8.37	7,3	11	29		EGDe
				2082	54	33256	27100	8.37	7	5	26		EGDe
				2177	119	33256	24700	8.37	7	14	39		EGDe
				2178	107	33256	24600	8.37	7	10	39		EGDe
				2088	178	33256	27000	8.37	6,9	15	50		EGDe
				2174	72	33256	24700	8.37	7,7	9	36		EGDe
				2208	84	33256	23200	8.37	7	11	29		CLIP80459

**Proteins predicted to be present in the extracellular milieu and in the cell wall**

protein	ncbi protein accession	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
autolysin, putative	gi 47093027	lmoh7858_0018	Q4EHT4	3253	40	12608	9600	9.13	8	4	18		CLIP80459
similar to N-acetylmuramoyl-L-alanine amidase and to internalin B	gi 16804242	lmo2203	Q8Y572	1726	179	41704	35100	5,28	4,7	14	47	1.8	EGDe
				1727	78	41704	35000	5,28	4,8				EGDe
P60 extracellular protein, invasion associated protein Iap	gi 16802625	iap (lmo0582)		a	68	47785	35300	9.16	9,2	10	28	1.1	CLIP80459
				2708	71	20388	18000	8.67	6.4	8	46		CLIP90602
				b	86	49808	34800	9.21	9,1	11	27		CLIP80459
				1085	147	50311	53000	9.32	9,3	14	38		EGDe
				f	73	50311	61000	9.32	9,3	11	24		EGDe
similar to N-acetylmuramoyl-L-alanine amidase	gi 16803561	lmo1521	Q8Y707	1436	179	46043	47400	5,53	4,95	17	49	1,1	EGDe
				1435	181	46043	47400	5,53	4,9	16	47		EGDe
				1455	153	46043	47400	5,53	4,8	14	44		EGDe
				2103	189	46043	27400	5,53	5,4	17	49		EGDe
				1444	204	46043	47400	5,53	5,2	19	53		EGDe
				ID1428	96	46043	47600	5.53	5,4	7	21		CLIP93657
				ID1438	95	46043	47600	5.53	5,35				CLIP93657
similar to hypothetical cell wall binding protein from B. subtilis with LysM domains	gi 46881966	lmof2365_2495	Q71WQ6	2073	117	29025	27600	5.50	4,9	14	61	1,1	CLIP80459
				2078	86	29025	27800	5,5	4,8	9	47		CLIP93665
	gi 16804560	lmo2522	Q8Y4C8	2133	95	29040	26283	5.70	4,95	13	59		CLIP93663
				2086	96	29040	26500	5,7	4,95	10	49		EGDe
				2142	88	29040	26000	5,7	4,85	9	54		EGDe

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
surface protein (GW repeat) similar to N-acetylmuramidase	gi 46908763	lmo2365_2564	Q71WI7	1122	170	57025	56500	5,8	4,95	16	37	1,8	CLIP80459
				1117	159	57025	56500	5,8	5,05	16	36		CLIP80459
				1123	258	57025	56500	5,8	4,95	23	51		CLIP93667
	gi 16804629	lmo2591	Q8Y464	1126	335	57004	57000	6,11	5,8	28	55		CLIEGDe
				1118	266	57004	56700	6,11	5,5	27	50		CLIEGDe
				1141	220	57004	56300	6,11	6,3				CLIEGDe
				1110	158	57004	56900	6,11	5,3	14	27		CLIP93649
similar to autolysin, N-acetylmuramidase	gi 16804728	lmo2691	Q8Y3Y8	1601	44	63534	38000	9,73	6,1	6	8	1.1	EGDe
peptidoglycan lytic protein P45	gi 46908677	lmo2365_2478	Q71WS3	2820	1,68 (P)	24340	15500	5,5	5,2	6	32	1,1	CLIP80459
				1485	179	42686	44000	8,56	7,15	15	38		EGDe
	gi 16804543	spl (lmo2505)	Q9RE04	1489	176	42686	44000	8,56	7,75	15	38		EGDe
				1484	117	42686	44000	8,56	6,7	13	34		EGDe
				1505	89	42686	45000	8,56	6,15	7	24		EGDe

#### Proteins predicted to be present in the extracellular milieu and in the membrane

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
cysK	gi 16802269	lmo0223	Q8YAC3	2002	108	32198	32200	5,32	4,95	13	52	2,2	EGDe
similar to B. subtilis YabE protein	gi 46906418	lmo2365_0197	Q724M7	c	70	44210	22600	8,9	6,2	9	21	5,2	CLIP80459
	gi 16802232	lmo0186	Q8YAE4	1538	100	44448	44400	9,00	9	10	28		EGDe
sulfatase	gi 16802967	lmo0927	Q8Y8H6	1214	73	74661	52600	6,01	4,9	11	14	5,2	EGDe
				1242	123	74661	53000	6,01	4,8	14	19		EGDe
				1233	174	74661	53000	6,01	4,8	14	19		EGDe
				1248	143	74661	53000	6,01	4,75	18	32		EGDe

**Proteins predicted to be present in the extracellular milieu and in the cytoplasm.**

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
conserved hypothetical protein	gi 47091514	lmoh7858_0416	Q4EMB8	3036	21	9622	11000	8.76	9,9	2	45	6	CLIP90602
general stress protein Ctc	gi 16802257	lmo0211	Q8YAD3	2216	73	22641	22600	4.44	4,5	6	43	4,1	EGDe
cold-shock domain family protein	gi 46907590	lmof2365_1381	Q71ZV8	3248	67	7261	9500	4.45	4,5	4	74	4.1	CLIP80459
catalase	gi 47092675	lmoh7858_3052	Q4EIV2	1006	371	55794	60300	5,38	4,95	33	66	4.2	CLIP80459
similar to B. subtilis SpoVG protein	gi 46879692	lmof2365_0207	Q724L7	3033	78	11250	11200	4.46	4,5	9	56	1.7	CLIP80459
glyceraldehyde-3-phosphate dehydrogenase, type I	gi 16804497	lmo2459	Q8Y4I1	1581	139	36264	36500	5,2	4,9	13	42	2,1,2	EGDe
				1585	93	36264	36900	5,2	4,95	12	30		EGDe
				1582	77	36284	36200	5,2	4,95	12	32		CLIP93667
highly similar to dihydrolipoamide dehydrogenase, E3 subunit of pyruvate dehydrogenase complex	gi 46880554	lmof2365_1076	Q721B1	1116	56	49451	57000	5,24	5	5	13	2.1.2	CLIP80459
				1108	1.79 (P)	49460	57000	5,2	4,95	7	17		CLIP80459
	gi 16803095	PdhD (lmo1055)	Q8Y862	1109	75	49457	56500	5,24	4,95	7	17		EGDe
class I heat-shock protein (chaperonin) GroES	gi 16804108	lmo2069	Q9AGE7	3173	105	10058	10000	4,6	4,6	8	82	3,9	EGDe

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
chaperonin GroEL	gi 16804107	lmo2068	Q71XU6	937	74	57332	60000	4,72	4,6	14	28	3.9	CLIP80459
				733	161	57332	69000	4,72	4,6	21	39		CLIP90602
chaperone protein DnaK	gi 46907701	lmof2365_1492	Q71ZJ7	826	71	66118	66000	4,57	4,6	9	24	3.9	CLIP80459
	gi 16803513	lmo1473	Q9S5A4	e	163	66104	68000	4,57	4,45	22	44		EGDe
				666	2,2 (P)	66190	71000	4,6	4,5	9	17		CLIP90602
enolase	gi 46908628	lmof2365_2428	Q71WX1	3745	215	46458		4,7		21	54	2,1,2	EGDe
conserved hypothetical protein	gi 47093198	lmof2365_2053	Q71XZ2	2816	87	15491	15400	4,48	4,4	6	58	6	CLIP93667
YceI like family protein	gi 46907027	lmof2365_0813	Q722C1	d	81	19312	15300	4,69	4,45	7	56	5.2	CLIP80459
phosphoglycerate kinase	gi 47094183	lmoh7858_2607	Q4EEP3	1482	78	42050	45800	4,97	4,8	9	36	2.1.2	CLIP80459
phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	gi 46908629	lmof2365_2429	Q71WX0	1021	147	56130	59300	5,18	4,9	17	40	2.1.2	CLIP93665
				1020	78	56150	15300	5,2	4,45	7	19		CLIP80459
superoxide dismutase, Mn	gi 46880935	lmof2365_1458	Q71ZN1	2488	93	22601	22600	5,23	4,95	9	47	4.2	CLIP80459
				2472	160	22601	22600	5,23	4,95	13	70		CLIP80459
				2491	102	22601	22600	5,23	4,8	7	44		CLIP80459
				2441	168	22601	22600	5,23	4,95	13	92		CLIP80459

**Protein predicted to be present in the cytoplasm.**

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
ribosomal subunit interface protein	gi 46908683	lmof2365_2484	Q71WR7	2602	70	21620	19900	5,25	5,05	6	50	5.2	CLIP80459

Proteins predicted to be present in the membrane

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
hypothetical protein lmo0412	gi 16802456	lmo0412	Q8Y9V8	2131	69	31274	26800	5,04	4,6	9	35	6	EGDe
conserved hypothetical protein	gi 16802619	lmo0576	Q8Y9F2	1778	1,63 (P)	45660	35000	5,0	4,6	6	13	1.8	CLIP80459
actin-assembly inducing protein precursor	gi 46252401	actA (lmo0204)	Q6E9R2	1031	86	69950	59300	4,99	4,6	14	30	1,8	CLIP80459
				673	134	66256	59200	5,08	4,7	15	36		CLIP93667
	gi 29651052	(ActA)	Q844E1	2220	53	22035	22500	9,03	4,7	5	40		EGDe
	gi 9050037	ActA	Q9JMZ5	2122	53	19897	27000	4,00	4,5	4	32		CLIP93657
	gi 9050055	actA (lmo0204)	P33379	609	1,53 (P)	69710	74000	5	4,8	6	14		EGDe
				757	2,22 (P)	69710	69000	5	4,8	10	23		EGDe
	gi 16802250	actA (lmo0204)	P33379	1159	70	70306	55800	5,01	4,9	10	17		EGDe
				979	89	70306	60200	5,01	4,4	11	18		EGDe
				1099	82	70306	57500	5,01	4,6	10	18		EGDe
				756	127	70306	69100	5,01	4,8	18	34		EGDe
				942	71	70306	62000	5,01	4,6	10	18		EGDe
				612	162	70306	70306	5,01	4,7	17	33		EGDe
				1630	50	70306	36000	5,01	4,7	6	10		EGDe
				1137	82	70306	56200	5,01	4,5	10	18		EGDe
AA3-600 quinol oxidase subunit II	gi 47097689	qoxA (lmof6854_0015)	Q4EML9	1931	104	41557	33250	6,42	5,1	6	0	1,4	CLIP93663
				1947	73	41557	32800	6,42	5,60				CLIP93649
	gi 46906239	qoxA (lmof2365_0016)	Q725F6	1922	89	41544	33200	6,04	4,95	13	56		CLIP80459
	gi 16802061	qoxA (lmo0013)	Q8YAV0	1938	64	41558	41800	6,04	5,2	7	23		EGDe
				1947	59	41558	32800	6,04	5,6	8	28		EGDe

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
sulfatase	gi 16802686	lmo0644	Q8Y989	1393	156	69218	48000	5,42	4,6	23	41	5.2	EGDe
				1425	151	69218	48000	5,42	4,7	20	38		EGDe
lipoprotein, putative	gi 16803108	lmo1068	Q8Y850	2058	72	30782	30100	5.86	5,6	9	39		CLIP90602
				2080	74	30782	28200	5.86	5,70	6	31		CLIP90602
conserved hypothetical protein with 3 domains NEAT	gi 47091854	lmoh7858_2319	Q4ELE1	887	151	63322	64300	6,07	4,95	21	47	1.8	CLIP90602
similar to pheromone ABC transporter (binding protein)	gi 16804235	lmo2196 (OppA)	Q9LAT7	890	72	62549	63500	5.30	4,8	11	28	1.2	EGDe
conserved lipoprotein	gi 16804675	lmo2637	Q8Y436	2035	35	32721	30600	6.25	4,9			1.8	CLIP93649
CD4+ T cell-stimulating antigen, lipoprotein	gi 16410817	tcsA(lmo1388)	Q48754	2295	78	38391	22600	5.02	4,95	8	25	1.2	CLIP93657
				1799	93	38391	34300	5.02	4,6	9	33		CLIP93657
				1653	109	38391	35800	5.02	4,6	12	35		EGDe

### Proteins predicted to be present in the cell wall

protein description	ncbi protein accession number	gene name	Swiss-Prot protein accession	group ID	Mascot Score	Mr (Da)		pI		Number of mass matched	Sequence coverage (%)	functional category according to listlist	identified in strain
						theoretical	observed	theoretical	observed				
Ser/Thr protein phosphatase family protein	gi 46906370	lmof2365_0148	Q724S5	629	159	82213	73000	4.87	4,6	22	41	1.8	CLIP80459
				628	85	82213	74000	4.87	4,6	15	25		CLIP90602
peptidoglycan anchored protein	gi 16412214	lmo2714	Q8Y3W5	1951	50	34648	33000	5.03	4,6	6	26	1.8	EGDe

(P) : identified with profoundA



Table 1S : Summary of predictions

GI	Database	Annotation	<i>L.monocytogenes strain</i>	Signal Peptide	predicted as secreted by SecretomeP?	Binding motif		Localization										
								Extracellular	Cell Wall	Membrane	Cytoplasmic	number of predicted localization	Integral Membrane protein?	number of TMD	Lipoprotein?	Translocated?	Extracytoplasmic?	
47093253	ref	serine protease, putative	H7858	Y	N		ChW	Ext					1	N		N	Y	Y
47093027	ref	autolysin, putative	H7858	Y	N	GW		CWamidase	Ext	CW			2	N		N	Y	Y
47091514	ref	conserved hypothetical protein	H7858	N	Y				Ext			Cyto	2	N		N	Y	Y
16802257	ref	general stress protein Ctc	EGDe	N	Y				Ext			Cyto	2	N		N	Y	Y
46907590	ref	cold-shock domain family protein	F2365	N	Y				Ext			Cyto	2	N		N	Y	Y
47092675	ref	catalase	H7858	N	Y				Ext			Cyto	2	N		N	Y	Y
46908859	ref	K <sup>+</sup> -transporting ATPase, C subunit	F2365	Y	N				Ext				1	N		N	Y	Y
46906370	ref	Ser/Thr protein phosphatase family protein	F2365	Y	N	LPXTG				CW			1	N		N	Y	N
46879692	gb	stage V sporulation protein G	F2365	N	Y				Ext			Cyto	2	N		N	Y	Y
16802269	ref	hypothetical protein lmo0223	EGDe	N	Y	IMP			Ext		Mb		2	Y	1	N	Y	Y
16802456	ref	hypothetical protein lmo0412	EGDe	Y	N	IMP					Mb		1	Y	1	N	Y	N
46908275	ref	penicillin-binding protein	F2365	Y	N				Ext				1	N		N	Y	Y
16802583	ref	hypothetical protein lmo0540	EGDe	Y	N				Ext				1	N		N	Y	Y
46906785	ref	penicillin-binding protein, putative	F2365	Y	N				Ext				1	N		N	Y	Y
16803478	ref	hypothetical protein lmo1438	EGDe	Y	N				Ext				1	N		N	Y	Y
16802619	ref	hypothetical protein lmo0576	EGDe	Y	N	IMP					Mb		1	Y	2	N	Y	N
16802628	ref	putative secreted protein	EGDe	Y	N				Ext				1	N		N	Y	Y
46906830	ref	hypothetical protein LMOF2365_0614	F2365	Y	N				Ext				1	N		N	Y	Y
46908125	ref	penicillin-binding protein	F2365	Y	N				Ext				1	N		N	Y	Y
16804242	ref	hypothetical protein lmo2203	EGDe	Y	N	GW		COG1705	Ext	CW			2	N		N	Y	Y
16804497	ref	glyceraldehyde-3-phosphate dehydrogenase	EGDe	N	Y				Ext			Cyto	2	N		N	Y	Y
46908683	ref	ribosomal subunit interface protein	F2365	N	N							Cyto	1	N		N	N	N
16803095	ref	dihydrolipoamide dehydrogenase	EGDe	N	Y				Ext			Cyto	2	N		N	Y	Y
46880554	gb	dihydrolipoamide dehydrogenase	F2365	N	Y				Ext			Cyto	2	N		N	Y	Y
16802250	ref	actin-assembly inducing protein precursor	EGDe	Y	N	IMP					Mb		1	Y	1	N	Y	N
46252401	gb	ActA	NRRL33221	Y	N	IMP					Mb		1	Y	1	N	Y	N
16804108	ref	co-chaperonin GroES	EGDe	N	Y				Ext			Cyto	2	N		N	Y	Y
16804107	ref	chaperonin GroEL	EGDe	N	Y				Ext			Cyto	2	N		N	Y	Y
16803513	ref	class I heat-shock protein (molecular chaperone) DnaK	EGDe	N	Y				Ext			Cyto	2	N		N	Y	Y

46907701	ref	chaperone protein DnaK	F2365	N	Y					Ext			Cyto	2	N		N	Y	Y
46908628	ref	enolase	F2365	N	Y					Ext			Cyto	2	N		N	Y	Y
16802248	ref	listeriolysin O precursor	EGDe	Y	N					Ext				1	N		N	Y	Y
16802625	ref	P60 extracellular protein, invasion associated protein Iap	EGDe	Y	N	LysM		NplC/P60		Ext	CW			2	N		N	Y	Y
46906827	ref	protein P60	F2365	Y	N	LysM		NplC/P60		Ext	CW			2	N		N	Y	Y
16803826	ref	internalin C	EGDe	Y	N					Ext				1	N		N	Y	Y
16409372	emb	AA3-600 quinol oxidase subunit II	EGDe	Y	N	LipoproteinIMP						Mb		1	Y	2	Y	Y	N
47097689	ref	quinol oxidase AA3, subunit II	F6854	Y	N	LipoproteinIMP						Mb		1	Y	2	Y	Y	N
46906239	ref	quinol oxidase AA3, subunit II	F2365	Y	N	LipoproteinIMP						Mb		1	Y	2	Y	Y	N
16802183	ref	hypothetical protein lmo0135	EGDe	Y	N	Lipoprotein						Mb		1	N		Y	Y	N
16802232	ref	hypothetical protein lmo0186	EGDe	N	Y	IMP				Ext		Mb		2	Y	1	N	Y	Y
46906418	ref	hypothetical protein LMOF2365_0197	F2365	N	Y	IMP				Ext		Mb		2	Y	1	N	Y	Y
16802487	ref	hypothetical protein lmo0443	EGDe	Y	N					Ext				1	N		N	Y	Y
47093198	ref	conserved hypothetical protein	H7858	N	Y					Ext			Cyto	2	N		N	Y	Y
4104145	gb	penicillin-binding protein homolog	NCTC10527	Y	N					Ext				1	N		N	Y	Y
16802686	ref	hypothetical protein lmo0644	EGDe	Y	N	IMP						Mb		1	Y	3-4	N	Y	N
46907027	ref	YceI like family protein	F2365	N	Y					Ext			Cyto	2	N		N	Y	Y
16802967	ref	hypothetical protein lmo0927	EGDe	N	Y	IMP				Ext		Mb		2	Y	5	N	Y	Y
46907182	ref	hypothetical protein LMOF2365_0970	F2365	Y	N					Ext				1	N		N	Y	Y
46880563	gb	conserved hypothetical protein	F2365	Y	N					Ext				1	N		N	Y	Y
16803108	ref	hypothetical protein lmo1068	EGDe	Y	N	Lipoprotein						Mb		1	N		Y	Y	N
16803373	ref	hypothetical protein lmo1333	EGDe	Y	N					Ext				1	N		N	Y	Y
16803561	ref	hypothetical protein lmo1521	EGDe	Y	N		SH3b	CWamidase		Ext	CW			2	N		N	Y	Y
16803792	ref	hypothetical protein lmo1752	EGDe	Y	N					Ext				1	N		N	Y	Y
47092905	ref	conserved hypothetical protein	H7858	Y	N					Ext				1	N		N	Y	Y
16803922	ref	hypothetical protein lmo1883	EGDe	Y	N					Ext				1	N		N	Y	Y
46908115	ref	chitinase	F2365	Y	N					Ext				1	N		N	Y	Y
16804195	ref	hypothetical protein lmo2156	EGDe	Y	N					Ext				1	N		N	Y	Y
47091854	ref	conserved hypothetical protein	H7858	Y	N	IMP						Mb		1	Y	1	N	Y	N
16411666	emb	lmo2196	EGDe	Y	N	Lipoprotein						Mb		1	N		Y	Y	N
16804542	ref	hypothetical protein lmo2504	EGDe	Y	N					Ext				1	N		N	Y	Y
46908676	ref	peptidase, M48 family	F2365	Y	N					Ext				1	N		N	Y	Y
16412006	emb	lmo2518	EGDe	Y	N					Ext				1	N		N	Y	Y
16804560	ref	hypothetical protein lmo2522	EGDe	Y	N	LysM		3D		Ext	CW			2	N		N	Y	Y
46881966	gb	LysM domain protein	F2365	Y	N	LysM		3D		Ext	CW			2	N		N	Y	Y
16804629	ref	surface protein (GW repeat) similar to N-acetylmuramidase	EGDe	Y	N	GW		COG1705		Ext	CW			2	N		N	Y	Y

46908763	ref	N-acetylmuramoyl-L-alanine amidase, family 4	F2365	Y	N	GW		COG1705	Ext	CW			2	N		N	Y	Y
16804728	ref	autolysin	EGDe	Y	N	LysM		COG1705	Ext	CW			2	N		N	Y	Y
16804675	ref	hypothetical protein lmo2637	EGDe	Y	N	Lipoprotein					Mb		1	N		Y	Y	N
16412214	emb	peptidoglycan anchored protein (LPXTG motif)	EGDe	Y	N	LPXTG				CW			1	N		N	Y	N
47094183	ref	phosphoglycerate kinase	H7858	N	Y				Ext			Cyto	2	N		N	Y	Y
46908629	ref	phosphoglyceromutase	F2365	N	Y				Ext			Cyto	2	N		N	Y	Y
16802247	ref	phosphatidylinositol-specific phospholipase c	EGDe	Y	N				Ext				1	N		N	Y	Y
16802251	ref	phospholipase C	EGDe	Y	N				Ext				1	N		N	Y	Y
46880935	gb	superoxide dismutase, Mn	F2365	N	Y				Ext			Cyto	2	N		N	Y	Y
16804543	ref	peptidoglycan lytic protein P45	EGDe	Y	N			NplC/P60	Ext	CW			2	N		N	Y	Y
46908677	ref	NLP/P60 family protein	F2365	Y	N			NplC/P60	Ext	CW			2	N		N	Y	Y
16410817	emb	CD4+ T cell-stimulating antigen, lipoprotein	EGDe	Y	N	Lipoprotein					Mb		1	N		Y	Y	Y

Y : yes, N : no, Ext : extracellular, CW : cell wall, Mb : membrane, Cyto : Cytoplasmic.

## 4 - DISCUSSION

A l'heure actuelle des connaissances, de nombreuses protéines impliquées dans la virulence de *L. monocytogenes* ont été mises en évidence. InlA et InlB permettent l'adhésion à la surface de la cellule eucaryote et la pénétration dans la cellule hôte par phagocytose. La production de LLO en combinaison avec PlcA et PlcB permettent l'échappement de la vacuole phagocytaire. La protéine ActA permet alors le déplacement des bactéries qui pourront coloniser les cellules adjacentes. Alors que des différences de virulence ont pu être observées entre différentes souches de *L. monocytogenes*, elles ne semblent pas dues à l'absence de l'un ou l'autre de ces gènes puisqu'il a été montré qu'ils étaient présents dans toutes les souches de *L. monocytogenes*. De plus, il est étonnant de constater que 98% des cas de listériose humaine ne sont dus qu'à 3 des 13 sérovars existants : 1/2a, 1/2b et 4b et que le sérovar 4b, prépondérant dans les cas épidémiques de listériose, est beaucoup moins fréquemment isolé que les deux autres sérovars dans les aliments ou dans l'environnement industriel. La classification en sérovars développée par Seeliger et Höne (1979) en fonction de propriétés antigéniques a été l'une des premières méthodes permettant de discriminer les souches. Depuis, de nombreuses méthodes de typage moléculaire ont montré que les 13 sérovars de *L. monocytogenes* étaient distribués en 3 lignées distincts, avec notamment les souches de sérovars 4b, 1/2b et 3b regroupées dans le lignage II et les souches de sérovars 1/2a, 1/2c et 3a trouvées de manière dominante dans le lignage I (Doumith *et al.*, 2004b, Brosch *et al.*, 1994, Wiedmann *et al.*, 1997, Gray *et al.*, 2004, Nightingale *et al.*, 2005). Nous avons montré qu'il existait une discrimination semblable en étudiant les protéines intracellulaires et extracellulaires de *L. monocytogenes*. En effet, une classification hiérarchique des souches en fonction de leur profil protéique a mis en évidence deux groupes, les souches de sérovar 1/2a appartenant à un groupe et les souches de sérovar 1/2b et 4b étant clairement groupées dans un autre groupe. Des études utilisant la technique d'hybridation de l'ADN (Doumith *et al.*, 2004c) avaient permis de mettre en évidence 19 gènes spécifiquement associés au lignage II ( sérovars 4b, 4d, 4e, 1/2b, 3b) et 5 gènes marqueurs du lignage I (sérovars 1/2a, 3a, 1/2c, 3c). Aucun de ces produits de gènes n'ont été mis en évidence lors de notre étude. Ceci peut être dû au fait que (i) ces protéines ne sont pas ou très faiblement exprimées dans nos conditions expérimentales ou (ii) ont un *pI* et/ou une *Mr* en dehors de la gamme étudiée ou encore (iii) sont liées à la membrane ou à la paroi, compartiments cellulaires que nous n'avons pas étudié. Afin d'acquérir une connaissance plus importante de la diversité protéomique de *L. monocytogenes*, il serait intéressant d'élargir cette étude

comparative à ces compartiments. L'analyse des protéines extracellulaires et intracellulaires a cependant permis de mettre en évidence 20 spots protéiques spécifiques du linéage II (sérovir 1/2b et 4b) et 26 spots protéiques spécifiques du linéage I (sérovir 1/2a).

L'analyse protéomique, en complément de l'analyse génétique, apporte des informations extrêmement importante, car la présence d'un gène ne reflète pas forcément l'expression de la protéine, qui est l'entité fonctionnelle. L'expression des protéines peut être régulée à différents niveaux, transcriptionnel ou traductionnel, avec l'intervention de promoteur(s) et/ou répresseur(s) qui peuvent être fonctionnels ou non selon les conditions expérimentales utilisées. L'étude protéomique a permis de mettre en évidence des différences significatives d'expression de protéines au sein de chaque linéage, différences qui ne semblent pas dues à l'absence des gènes chez les deux souches séquencées EGDe et CLIP80459 de notre panel.

Il faut également noter que l'absence d'un spot protéique à l'endroit attendu dans le gel d'EBD ne reflète pas toujours l'absence de la protéine. Ainsi, certaines protéines sont présentes sous plusieurs spots résultant de différents niveaux de modifications post-traductionnelles. Les modifications post-traductionnelles peuvent modifier la fonction des protéines. Il a notamment été montré que la phosphorylation avait un rôle important dans la fonctionnalité de nombreuses protéines de *L. monocytogenes* (Archambaud *et al.*, 2005, Herro *et al.*, 2005; Deutscher *et al.*, 2005, Archambaud *et al.*, 2006, Schnupf *et al.*, 2006). Par exemple, Schnupf *et al.* ont montré qu'avant que la listériolysine O intracellulaire soit dégradée, elle était ubiquitinée et phosphorylée au niveau de la séquence PEST (Schnupf *et al.*, 2006). Certains de ces spots protéiques étaient donc spécifiques de chaque linéage et appartenaient à différentes catégories (métabolisme, protéines de surface, division cellulaire...). L'absence d'un spot à l'endroit attendu peut dans certains cas provenir de différences dans la séquence en acides aminés qui se traduisent par des modifications plus ou moins importantes de *pI* ou de *Mr*. Un alignement multiple des séquences en acides aminés de ces protéines chez les souches séquencées (Glaser *et al.*, 2001 ; Nelson *et al.*, 2004) montre qu'elles sont identiques au sein des souches de même sérovir mais présentent des différences entre les sérovirs. Il semble que les séquences protéiques des souches d'un même linéage soient plus proches que celles de l'autre linéage. Un séquençage de ces protéines chez toutes les souches étudiées serait nécessaire pour confirmer cette observation. L'essentiel des spots sérovir-spécifiques étaient surexprimés ou sous exprimés chez les souches de sérovir 1/2a, ce qui a clairement permis de classer ces souches dans un groupe différent des souches de sérovir 4b et 1/2b. Seul un spot s'est révélé spécifique des souches de sérovir 1/2b. Concernant les souches de sérovir 4b à l'origine de la majorité des cas de listériose humaine,

3 spots protéiques étaient sous exprimés et correspondaient à 2 protéines. Deux de ces spots protéiques correspondaient à une sérine protéase (LMOF2365\_1900), et étaient présents dans les 4 souches de sérovar 4b étudiés mais absents chez les souches de sérovar 1/2b et 1/2a. Cette analyse protéomique confirme des résultats d'hybridation ADN (Doumith *et al.*, 2004b) où le gène codant cette protéine était présent dans toutes les souches de sérovar 4b étudiées mais absent chez les autres sérovares, et prouve que la protéine est bien synthétisée. Deux sérine protéases de *L. monocytogenes* ont été montrées comme impliquées dans la résistance au stress et dans la virulence de cette bactérie. Par exemple, HtrA est essentielle pour la survie après une exposition à divers stress environnementaux et cellulaires et est nécessaire pour une colonisation efficace de la rate (Wonderling *et al.*, 2004 ; Stack *et al.*, 2005 ; Wilson *et al.*, 2006).

De nombreux spots protéiques présentaient également une différence d'expression entre les souches sans que ceci ne soit lié au sérovar. Ceci concernait différentes catégories de protéines (impliquées dans le métabolisme, protéines de surfaces, protéines de stress...). Ainsi, certaines protéines impliquées dans la virulence présentaient des différences d'abondance. PlcB était présente sous de multiples spots possédant des *pI* et des *Mr* différents. Ces spots présentaient d'importantes différences entre les souches. De plus, la souche CLIP90602 (souche clinique, sérovar 1/2b, virulence A) n'exprimait pas ou à un niveau très faible les facteurs de virulence ActA, LLO, PlcA et PlcB. Au niveau du génome, l'analyse de 93 souches (Doumith *et al.*, 2004c) a permis de montrer que les principaux gènes de virulence étaient présents chez toutes les souches. Par contre, des différences de séquences existent entre les souches pour ces gènes. Il a notamment été observé que certaines souches produisaient une InlA tronquée entraînant la synthèse d'une protéine non fonctionnelle qui pourrait être à l'origine d'une virulence atténuée. Olier *et al* (2003) ont montré que des souches issues de portage asymptomatique produisaient des formes tronquées d'InlA, avec des masses moléculaires apparentes de 47 kDa, alors que les souches à l'origine de cas de listériose produisaient une protéine fonctionnelle de 80 kDa. Jacquet *et al* (2004) ont également montré qu'une InlA entière était fortement associée aux souches cliniques puisque 96% d'entre elles possédaient une protéine entière, alors que ce n'était le cas que pour 65% des souches d'origine alimentaire. Les souches de portage que nous avons étudié possédaient toutes une InlA entière (Olier *et al.*, 2003). De plus, aucune protéine n'a été montrée comme spécifique des 4 souches de portage asymptomatique, ou d'une des 3 origines étudiées.

Malgré un choix minutieux des souches, le facteur origine doit être pris avec précaution. Par exemple, nous ne pouvons pas écarter la possibilité que ces souches de portage puissent

être à l'origine d'une listériose chez une personne ayant un système immunitaire affaibli, ou si une dose plus importante de bactéries avait été ingérée. Également, il n'est pas exclu que les souches isolées de l'environnement puissent être à l'origine d'une listériose humaine. Nos expériences ont été effectuées dans des conditions optimales de croissance. Il serait intéressant d'effectuer les mêmes expériences dans des conditions de culture ressemblant à celles rencontrées lors d'une infection humaine (le faible pH de l'estomac, la présence de bile et d'enzymes dans le fluide gastrique) ou dans des industries agroalimentaires (basses températures, fortes concentrations de sel), qui pourraient permettre d'apporter des éléments de réponse sur la colonisation des différentes niches écologiques. Par contre, l'étude de chaque sérovar indépendamment des deux autres a permis de mettre en évidence, chez les souches de sérovar 4b uniquement, des spécificités en fonction de leur origine. Malgré le faible nombre de souches prises en compte, il est intéressant d'observer que les deux souches isolées de portage asymptomatique ont de nombreux spots en commun, différents des deux autres souches de sérovar 4b. Un panel plus important de souches de sérovar 4b permettrait de vérifier cette observation.

Lors de notre étude, nous avons également analysé le facteur de virulence afin de déterminer si des différences de profils protéiques pouvaient être corrélées avec le niveau de virulence des souches. Les 2 niveaux de virulence déterminés par les tests sur œufs embryonnés pouvaient être discriminés en prenant en compte un ensemble de spots protéiques. Cependant, aucun de ces spots pris individuellement n'est apparu comme un marqueur significatif de son groupe de virulence. Des études précédentes n'avaient pas non plus montré de corrélation entre le niveau de virulence et l'origine, le sérovar, le phago-var, le ribotype ou le profil PFGE. Il faut ajouter que la détermination du niveau de virulence des souches peut donner des résultats variables en fonction de la nature du test mis en œuvre.

L'analyse de la biodiversité de *L. monocytogenes* a permis de cibler de nombreuses protéines d'intérêt dans le milieu extracellulaire des différentes souches analysées. L'étude de ce sous-protéome est importante car il est connu que plusieurs facteurs de virulence sont sécrétés dans le milieu extracellulaire. Or, peu d'études ont été réalisées sur les protéines sécrétées et les systèmes permettant cette sécrétion chez *L. monocytogenes*. Une analyse bioinformatique a cependant permis de prédire l'existence de 6 systèmes de sécrétion chez cette bactérie (Desvaux & Hebraud, 2006) dont 4 permettraient la translocation des protéines dans le milieu extracellulaire : le système Sec, le système Tat (Twin-arginine translocation), les holins et Wss [WXG100 (protéines avec un motif WXG d'~100 résidus) secretion system]. Cependant, cette analyse reste prédictive et des études expérimentales sont nécessaires afin de

déterminer si ces différents systèmes de sécrétion sont bien fonctionnels et quelles protéines sont sécrétées par chacun d'eux. Jusqu'à présent, au sein du genre *Listeria*, seules les protéines sécrétées par *L. monocytogenes* EGDe et *L. innocua* ont été étudiées expérimentalement (Trost *et al.*, 2005). Cette étude a permis de mettre en évidence d'importantes différences entre ces deux espèces. L'étude comparative des protéomes de diverses souches de *L. monocytogenes* que nous avons réalisée a également permis de mettre en évidence une diversité importante concernant les protéines sécrétées au sein même cette espèce. Afin d'élargir notre connaissance des protéines sécrétées par *L. monocytogenes*, des protéines ont été prélevées et identifiées chez différentes souches. Une analyse bioinformatique a ensuite été réalisée afin de caractériser les protéines identifiées dans le milieu extracellulaire. 70% d'entre elles possédaient un peptide signal, ce qui signifie que ces protéines peuvent être sécrétées par le système Sec ou Tat. Cependant, aucune n'a été prédite comme sécrétée par le système Tat. L'absence d'un peptide signal ne signifie pas que la protéine est cytoplasmique. En effet, la sécrétion par les holins et Wss ne nécessite pas de peptide signal. Aucune des protéines prédites comme sécrétée par ces systèmes (Desvaux & Hebraud, 2006) n'a été identifiée dans nos conditions expérimentales. Trost *et al.* (Trost *et al.*, 2005) avaient identifiés Lmo0129 dans le surnageant de *L. monocytogenes* EGDe et Lin0176 chez *L. innocua*, qui sont des protéines prédites comme sécrétées par les holins. Les auteurs avaient alors prédit ces deux protéines comme cytoplasmiques car ils n'avaient pris en considération que la présence d'un peptide signal comme critère pour déterminer qu'une protéine était sécrétée. Il semblerait donc que le système de sécrétion par les holins soit fonctionnel. Concernant Wss, des études ont démontré que le gène codant WXG100-A était transcrit (Way & Wilson, 2005). Par contre, la sécrétion de protéines par ce système n'a pas encore été démontrée. Des études complémentaires sont nécessaires pour déterminer si ces systèmes de sécrétion sont fonctionnels chez *L. monocytogenes* et quel sont leur(s) rôle(s). Les autres protéines ne possédant pas de peptide signal sont généralement considérées comme des protéines cytoplasmiques. La détection de ces protéines dans l'environnement extracellulaire peut être attribuée à de la lyse cellulaire ayant eu lieu durant l'expérience. Cependant, certaines de ces protéines ont été détectées dans le milieu extracellulaire de diverses bactéries, laissant supposer qu'elles sont transloquées par des systèmes de sécrétion inconnus chez des cellules intactes. Il a été montré que certaines de ces protéines, qui avaient un rôle dans le cytoplasme, avaient également un rôle dans le milieu extracellulaire, ce rôle pouvant être différent selon la localisation. Ces protéines ont été appelées "moonlighting" (Jeffery, 1999). Bendtsen *et al.* (2002), suite à une analyse bibliographique, ont développé le



logiciel SecretomeP, qui permet d'identifier les protéines sécrétées par un système inconnu. Parmi les 23 protéines ne possédant pas de peptide signal que nous avons identifié dans le milieu extracellulaire de *L. monocytogenes*, 22 étaient prédites par SecretomeP comme sécrétées. Il semble donc que des systèmes de sécrétion inconnus existent chez *L. monocytogenes*. Par exemple, la superoxide dismutase (SodA), une protéine connue comme cytoplasmique et qui protège les organismes contre le stress oxydatif, a été montrée comme extracellulaire dans nos expériences. SodA a déjà été montrée comme extracellulaire chez *L. monocytogenes* (Lenz *et al.*, 2003 ; Archambaud *et al.*, 2006) et sécrétée par le système SecA2 ; elle a également été identifiée comme protéine extracellulaire chez d'autres bactéries. De même, l'énolase et les protéines chaperones DnaK et GroEL semblent être sécrétées par ce système. Alors que ces protéines ont un rôle clair dans le cytoplasme, le rôle de la plupart de ces protéines dans le milieu extracellulaire reste à déterminer. Certaines protéines que nous avons identifiées dans le milieu extracellulaire et qui sont connues pour leur implication dans une voie métabolique, comme c'est le cas de la glyceraldehyde-3-phosphate-dehydrogenase ou de l'énolase dans le métabolisme des carbohydrates, ont été montrées comme étant des protéines "moonlighting", puisqu'elles sont également présentes en surface où elles possèdent une activité de liaison au plasminogène. Comme un peptide signal ne garantit pas que la protéine soit sécrétée dans le milieu extracellulaire, des études complémentaires ont été effectuées. Vingt des 54 protéines qui possédaient un peptide signal possédaient également un motif de liaison à la surface. Les protéines avec un motif GW ou LPXTG étaient toutes des enzymes impliquées dans la dégradation de la paroi, ce qui explique leur présence dans le milieu extracellulaire. Il est intéressant de noter que la sérine protéase qui n'a été identifiée que chez les souches de sérovar 4b possédait 3 copies du motif ChW (Clostridial hydrophobic domain with a conserved W residue), un motif supposé impliqué dans l'ancrage à la paroi (Nolling *et al.*, 2001). Seule cette protéine de *L. monocytogenes* a été rapportée comme possédant un tel motif. Dix huit protéines ont été trouvées comme membranaires, ce qui rassemble les protéines dites "intégrales à la membrane" et les lipoprotéines.

Au final, 98.7% des protéines identifiées ont été prédites comme transloquées et 78% comme sécrétées dans le milieu extracellulaire. 44% des protéines étaient prédites comme ayant une localisation cellulaire multiple. L'étude de la fonction de ces différentes protéines, qui peut être différente selon leur localisation cellulaire permettrait d'avoir une meilleure connaissance de ces protéines dites "moonlighting". Il serait également intéressant d'étudier la fonction des autres protéines sécrétées dont la fonction est inconnue et plus spécifiquement celles présentant des différences entre les souches de *L. monocytogenes*. L'étude des systèmes

de sécrétion chez cette bactérie serait également nécessaire pour avancer dans la compréhension de ces mécanismes chez cette bactérie et de manière plus large chez les bactéries à Gram positif, qui sont peu étudiées sur ce sujet. Ces connaissances permettraient de mieux comprendre l'éventuelle implication de ces systèmes de sécrétion et des protéines sécrétées dans la pathogénicité et dans la colonisation de l'environnement. Pour cela, de nombreux points doivent être abordés. Tout d'abord il est important de déterminer si les différents systèmes prédits comme présents sont fonctionnels chez *L. monocytogenes*, et quelles protéines sont réellement sécrétées par ces différents systèmes. Ensuite, il serait intéressant d'approfondir les connaissances sur le système SecA2, notamment afin de comprendre pourquoi il y a deux systèmes Sec chez les bactéries à Gram positif et pas chez celles à Gram négatif. Enfin, pour de nombreuses protéines connues comme cytoplasmiques et également identifiées dans le milieu extracellulaire, il faudrait rechercher si des systèmes de sécrétion inconnus à ce jour existent chez *L. monocytogenes*.



# **CONCLUSION ET PERSPECTIVES**



---

## Conclusion et Perspectives

*Listeria monocytogenes* est l'agent étiologique de la listériose, une infection présentant un risque sanitaire important, puisque malgré une diminution de l'incidence dans la fin des années 1990, le nombre de cas se maintient autour des 250 par an avec un taux de mortalité toujours élevé entre 20 et 30%. De nombreuses études ont permis d'élucider les principaux mécanismes moléculaires et cellulaires impliqués dans l'infection par cette bactérie et un certain nombre de facteurs de virulence ont été découverts. Cependant, des différences de niveau de virulence sont observées entre les souches et 96% des cas de listériose ne sont dus qu'à 3 des 13 sérovars de *L. monocytogenes*. Ceci ne peut pas s'expliquer par l'absence d'un ou plusieurs gènes de virulence puisque ceux-ci sont présents dans la quasi-totalité des souches (Doumith *et al.*, 2004c). Le pouvoir pathogène d'une souche peut également être lié à sa capacité à résister aux nombreux stress rencontrés dans l'environnement, lors des procédés de production et transformation des aliments et chez son hôte. A ce jour, toutes les souches de *L. monocytogenes* sont considérées comme potentiellement pathogènes pour l'homme. Cependant, dans un contexte d'analyse et de gestion du risque sanitaire lié aux aliments, il est indispensable de mieux comprendre la virulence de *L. monocytogenes* et de caractériser les risques sanitaires auxquels s'exposent les consommateurs vis-à-vis de produits alimentaires faiblement contaminés.

Mon travail de thèse s'est inscrit dans cet objectif, avec dans un premier temps, l'étude du produit du gène de la ferritine (*fri*), qui a permis de montrer l'importance de cette protéine dans la virulence et la résistance à de nombreux stress, tels que les stress thermiques, les conditions de carence en fer et le stress oxydatif. Le rôle majeur de cette ferritine provient certainement de sa capacité à séquestrer le fer pour éviter son action toxique et pour permettre la croissance de la bactérie dans des conditions de carence en fer. Cependant, la séquestration du fer par la protéine Fri n'a été montrée que chez *L. innocua* et des expériences supplémentaires sont nécessaires afin de prouver expérimentalement ce rôle chez *L. monocytogenes*. Il serait intéressant d'étudier le protéome de la souche sauvage et du mutant, suite à divers stress, par électrophorèse bidimensionnelle afin de mieux comprendre le rôle régulateur de cette protéine. Des études réalisées précédemment au laboratoire ayant montré que l'abondance de ces différents spots variait en fonction de la phase de croissance de la bactérie, il serait intéressant de déterminer si ces différentes formes, sûrement dues à des modifications post-traductionnelles, varient en fonction des différents stress rencontrés. La

détermination de la nature et des sites des modifications apporterait également des informations sur le(s) rôle(s) et le mode d'action de cette protéine. Des techniques de spectrométrie de masse (Q-TOF, Trappe Ionique) devraient fournir les données structurales et différentes constructions génétiques par mutagenèse dirigée devraient contribuer à mieux définir son mode d'action.

La deuxième partie de mon travail a consisté à étudier la biodiversité de *L. monocytogenes* par une approche protéomique. Cette étude s'est insérée dans un programme de recherche, financé par la DGAL, intitulé «Approches génomiques et post-génomiques pour le typage des souches de *Listeria monocytogenes* et la prédiction du risque » qui avait pour but de caractériser un ensemble de souches de diverses origines et sérovars par l'étude du génome, du transcriptome, du protéome et de la virulence. Ce programme de recherche avait comme objectif de déterminer s'il existe une corrélation entre le sérovar, le niveau de virulence ou le caractère épidémiologique d'un isolat de *L. monocytogenes* et sa signature génétique. L'identification de gènes ou produits de gènes conditionnant le pouvoir pathogène de *L. monocytogenes* permettrait d'aider à évaluer et à hiérarchiser les risques sanitaires auxquels s'exposent les consommateurs vis-à-vis de produits alimentaires faiblement contaminés. Les 3 approches mises en œuvre par les différents laboratoires permettent d'avoir des données complémentaires. Les études protéomiques ont permis de montrer une discrimination des souches semblables aux résultats obtenus par l'étude du génome, c'est-à-dire une classification des souches de sérovar 1/2a, indépendante des souches de sérovar 4b et 1/2b. Cependant, contrairement aux études génomiques qui se basent sur la présence ou l'absence d'un gène, les résultats de protéomique sont plus complexes. En effet, les différences observées, et qui permettent notamment de discriminer les souches en deux lignées, peuvent être dues à une absence d'expression de la protéine mais également à des différences de *pI* et de *Mr*, qui, pour certaines, sont dues à des modifications post-traductionnelles. Des études complémentaires seraient nécessaires pour déterminer si certaines des différences observées résultent effectivement de modifications posttraductionnelles et quel est l'impact sur la fonction de la protéine. Parmi les différences observées au niveau protéomique, la présence de la sérine protéase LMOF2365\_1900 uniquement chez les souches de sérovar 4b, sérovar majoritairement impliqués dans les cas de listériose, vient confirmer les données génomiques. Il serait donc intéressant, entre autre, de caractériser la fonction de cette protéine. Il a été montré que certaines sérines protéases jouaient un rôle important dans la résistance au stress et dans la virulence de *L. monocytogenes* (Gaillot *et al.*, 2000 ; Gaillot *et al.*, 2001 ; Wonderling *et al.*, 2004). Il pourrait donc être envisagé dans un premier temps de comparer la

croissance dans différentes conditions de stress de la souche sauvage et d'une souche mutée pour ce gène. Dans un deuxième temps, des tests de virulence et de survie dans les organes de la souche mutante permettraient de déterminer si ce gène est impliqué dans la pathogénicité de *L. monocytogenes*. Enfin, cette protéine est la seule identifiée chez *Listeria* qui contient des motifs ChW décrits chez *Clostridium acetobutylicum* (Nolling *et al.*, 2001a) comme permettant l'ancrage à la surface. Dans notre étude, nous avons identifié cette protéine dans le milieu extracellulaire. Des études de délétion des motifs ChW pourraient être réalisées pour déterminer le rôle de ces motifs chez *L. monocytogenes*.

L'étude protéomique a porté sur l'étude des protéines intracellulaires et extracellulaires. Il serait intéressant d'étudier également les protéines de surface, qui sont fréquemment impliquées dans l'invasion de l'hôte ou dans la colonisation (formation de biofilms) et la perception de l'environnement. A ce jour, peu d'études ont été effectuées sur les systèmes de sécrétion de *L. monocytogenes*. Une meilleure connaissance dans ce domaine pourrait apporter des pistes nouvelles sur la pathogénicité ainsi que la colonisation de et l'adaptation à l'environnement. Pour cela, il serait tout d'abord nécessaire de confirmer si les différents systèmes de sécrétion déterminés chez *L. monocytogenes* par une étude bioinformatique (Desvaux & Hebraud, 2006) sont bien fonctionnels. Par exemple, pour vérifier la fonctionnalité du système de sécrétion Tat, une approche similaire à celle réalisée par Meissner *et al.* (2007) pourrait être envisageable. Dans cette étude, une protéine de fusion entre un peptide signal Tat et la GFP (Green Fluorescent Protein) a été utilisée comme modèle pour déterminer si le système Tat était fonctionnel chez 3 espèces bactériennes à Gram positif. Une analyse fonctionnelle par délétion des gènes impliqués dans ces différents systèmes de sécrétion permettrait de mieux comprendre leur importance. Il serait ensuite nécessaire de déterminer expérimentalement les protéines qui sont effectivement sécrétées par chaque système.

Une banque de données de gels d'EBD des protéines intracellulaires de *L. monocytogenes* a été créée au laboratoire et est accessible sur internet (<http://www.clermont.inra.fr/proteome>). Les gels d'EBD annotés avec les protéines sécrétées identifiées dans notre étude viendront prochainement s'ajouter à cette base de données. L'étude des protéines de l'enveloppe cellulaire et des protéines basiques permettraient de compléter la couverture du protéome de *L. monocytogenes*.

Les souches analysées étaient virulentes ou légèrement atténuées. La comparaison avec des souches non virulentes permettrait une exploration plus pertinente dans le but de caractériser de nouveaux facteurs directement ou indirectement impliqués dans la



pathogénicité. L'étude d'autres sérovars, rarement trouvés dans les cas de listériose humaine, contribuerait également à avoir une meilleure connaissance de la diversité au sein des souches de *L. monocytogenes*. Notre étude ayant été effectuée dans des conditions optimales de croissance pour toutes les souches, les éventuelles spécificités de leur niche écologique d'origine ont pu être atténuées voire même supprimées. La comparaison des souches cultivées dans des conditions proches de celles rencontrées chez l'hôte ou dans les ateliers agroalimentaire pourraient éventuellement mettre en évidence des particularités d'adaptation. La capacité de *L. monocytogenes* à former des biofilms étant une caractéristique importante pour la colonisation de l'environnement, la comparaison de l'aptitude des différentes souches à former des biofilms pourrait être intéressante. Cette analyse pourrait être effectuée avec le système de criblage BioFilm Ring Test® (Chavant *et al.*, 2007), qui est un outil reproductible et rapide développé par la société BioFilm Control SAS hébergée au sein de l'UR454 de Microbiologie à l'INRA de Theix. Il est également connu que les mécanismes moléculaires sont différents entre des bactéries en biofilms et des cellules planctoniques. La comparaison des différentes souches cultivées en biofilm pourrait donc apporter des informations complémentaires.

L'étude génomique, transcriptomique et protéomique a permis de mettre en évidence une diversité importante au sein des souches de *L. monocytogenes*. L'études des profils d'expression protéique de ces souches dans différentes conditions de croissance devrait permettre d'améliorer encore la connaissance de cette biodiversité et de mieux caractériser les spécificités des souches en fonction de leur caractère épidémiologique. Enfin, l'analyse fonctionnelle de mutants pour des gènes ciblés par les différentes approches citées permettraient d'avancer dans la compréhension des mécanismes de virulence et de colonisation de l'environnement par *L. monocytogenes*.

# **REFERENCES BIBLIOGRAPHIQUES**



**Almiron, M., Link, A. J., Furlong, D. & Kolter, R.** (1992). A novel DNA-binding protein with regulatory and protective roles in starved *Escherichia coli*. *Genes Dev.* 6, 2646-2654.

**Andrews, S. C., Robinson, A. K. & Rodriguez-Quinones, F.** (2003). Bacterial iron homeostasis. *FEMS Microbiol Rev.* 27, 215-237.

**Angelidis, A. S. & Smith, G. M.** (2003). Role of the glycine betaine and carnitine transporters in adaptation of *Listeria monocytogenes* to chill stress in defined medium. *Appl Environ Microbiol.* 69, 7492-7498.

**Annous, B. A., Becker, L. A., Bayles, D. O., Labeda, D. P. & Wilkinson, B. J.** (1997). Critical role of anteiso-C15:0 fatty acid in the growth of *Listeria monocytogenes* at low temperatures. *Appl Environ Microbiol.* 63, 3887-3894.

**Archambaud, C., Gouin, E., Pizarro-Cerda, J., Cossart, P. & Dussurget, O.** (2005). Translation elongation factor EF-Tu is a target for Stp, a serine-threonine phosphatase involved in virulence of *Listeria monocytogenes*. *Mol Microbiol.* 56, 383-396.

**Archambaud, C., Nahori, M. A., Pizarro-Cerda, J., Cossart, P. & Dussurget, O.** (2006). Control of *Listeria* superoxide dismutase by phosphorylation. *J Biol Chem.* 281, 31812-31822.

**Audurier, A., Pardon, P., Marly, J. & Lantier, F.** (1980). Experimental infection of mice with *Listeria monocytogenes* and *L. innocua*. *Ann Microbiol (Paris).* 131B, 47-57.

**Audurier, A., Rocourt, J. & Courtieu, A. L.** (1977). [Isolation and characterization of "*Listeria monocytogenes*" bacteriophages (author's transl)]. *Ann Microbiol (Paris).* 128, 185-198.

**Augustin, J. C., Rosso, L. & Carlier, V.** (1999). Estimation of temperature dependent growth rate and lag time of *Listeria monocytogenes* by optical density measurements. *J Microbiol Methods.* 38, 137-146.

**Bayles, D. O., Annous, B. A. & Wilkinson, B. J.** (1996). Cold stress proteins induced in *Listeria monocytogenes* in response to temperature downshock and growth at low temperatures. *Appl Environ Microbiol.* 62, 1116-1119.

**Beales, N.** (2004). Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: a review. *Comp. Rev. Food Sci. Food Safety* 3. 1, 120.

**Becker, L. A., Cetin, M. S., Hutkins, R. W. & Benson, A. K.** (1998). Identification of the gene encoding the alternative sigma factor sigmaB from *Listeria monocytogenes* and its role in osmotolerance. *J Bacteriol.* 180, 4547-4554.

**Behari, J. & Youngman, P.** (1998a). Regulation of hly expression in *Listeria monocytogenes* by carbon sources and pH occurs through separate mechanisms mediated by PrfA. *Infect Immun.* 66, 3635-3642.

**Behari, J. & Youngman, P.** (1998b). A homolog of CcpA mediates catabolite control in *Listeria monocytogenes* but not carbon source regulation of virulence genes. *J Bacteriol.* 180, 6316-6324.

**Berger, U. & Pietsch, U.** (1975). The rate of healthy carriers of *Listeria monocytogenes*. *Med Microbiol Immunol (Berl).* 161, 63-71.

**Beumer, R. R., te Giffel, M. C., Spoorenberg, E. & Rombouts, F. M.** (1996). *Listeria* species in domestic environments. *Epidemiol Infect.* 117, 437-442.

**Bibb, W. F., Gellin, B. G., Weaver, R., Schwartz, B., Plikaytis, B. D., Reeves, M. W., Pinner, R. W. & Broome, C. V.** (1990). Analysis of clinical and food-borne isolates of *Listeria monocytogenes* in the United States by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. *Appl Environ Microbiol.* 56, 2133-2141.

**Bielecki, J., Youngman, P., Connelly, P. & Portnoy, D. A.** (1990). *Bacillus subtilis* expressing a haemolysin gene from *Listeria monocytogenes* can grow in mammalian cells. *Nature.* 345, 175-176.

**Boerlin, P., Bannerman, E., Ischer, F., Rocourt, J. & Bille, J.** (1995). Typing *Listeria monocytogenes*: a comparison of random amplification of polymorphic DNA with 5 other methods. *Res Microbiol.* 146, 35-49.

**Boerlin, P., Boerlin-Petzold, F., Bannerman, E., Bille, J. & Jemmi, T.** (1997). Typing

*Listeria monocytogenes* isolates from fish products and human listeriosis cases. *Appl Environ Microbiol.* 63, 1338-1343.

**Bohne, J., Sokolovic, Z. & Goebel, W.** (1994). Transcriptional regulation of *prfA* and *PrfA*-regulated virulence genes in *Listeria monocytogenes*. *Mol Microbiol.* 11, 1141-1150.

**Borucki, M. K., Peppin, J. D., White, D., Loge, F. & Call, D. R.** (2003). Variation in biofilm formation among strains of *Listeria monocytogenes*. *Appl Environ Microbiol.* 69, 7336-7342.

**Bozzi, M., Mignogna, G., Stefanini, S., Barra, D., Longhi, C., Valenti, P. & Chiancone, E.** (1997). A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in *Listeria innocua*. *J Biol Chem.* 272, 3259-3265.

**Braun, L., Ghebrehiwet, B. & Cossart, P.** (2000). gC1q-R/p32, a C1q-binding protein, is a receptor for the InlB invasion protein of *Listeria monocytogenes*. *EMBO J.* 19, 1458-1466.

**Breand, S., Fardel, G., Flandrois, J. P., Rosso, L. & Tomassone, R.** (1998). Model of the influence of time and mild temperature on *Listeria monocytogenes* nonlinear survival curves. *Int J Food Microbiol.* 40, 185-195.

**Brosch, R., Catimel, B., Millon, G., Buchrieser, C., Vindel, E. & Rocourt, J.** (1993). Virulence heterogeneity of *Listeria monocytogenes* strains from various sources (food, human, animal) in immunocompetent mice and its association with typing characteristics. *J Food Prot.* 56, 296-301.

**Brosch, R., Chen, J. & Luchansky, J. B.** (1994). Pulsed-field fingerprinting of listeriae: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. *Appl Environ Microbiol.* 60, 2584-2592.

**Brundage, R. A., Smith, G. A., Camilli, A., Theriot, J. A. & Portnoy, D. A.** (1993). Expression and phosphorylation of the *Listeria monocytogenes* ActA protein in mammalian cells. *Proc Natl Acad Sci U S A.* 90, 11890-11894.

**Cabanes, D., Dehoux, P., Dussurget, O., Frangeul, L. & Cossart, P.** (2002). Surface proteins and the pathogenic potential of *Listeria monocytogenes*. *Trends Microbiol.* 10, 238-

245.

**Calvo, E., Pucciarelli, M. G., Bierne, H., Cossart, P., Pablo Albar, J. & Garcia-Del Portillo, F.** (2005). Analysis of the *Listeria* cell wall proteome by two-dimensional nanoliquid chromatography coupled to mass spectrometry. *Proteomics*. 5, 433-443.

**Camilli, A., Tilney, L. G. & Portnoy, D. A.** (1993). Dual roles of plcA in *Listeria monocytogenes* pathogenesis. *Mol Microbiol*. 8, 143-157.

**Chand, P. & Sadana, J. R.** (1999). Outbreak of *Listeria ivanovii* abortion in sheep in India. *Vet Rec*. 145, 83-84.

**Chavant, P., Gaillard-Martinie, B., Talon, R., Hebraud, M. & Bernardi, T.** (2007). A new device for rapid evaluation of biofilm formation potential by bacteria. *J Microbiol Methods*. 68, 605-612.

**Chen, Y., Ross, W. H., Gray, M. J., Wiedmann, M., Whiting, R. C. & Scott, V. N.** (2006). Attributing risk to *Listeria monocytogenes* subgroups: dose response in relation to genetic lineages. *J Food Prot*. 69, 335-344.

**Chico-Calero, I., Suarez, M., Gonzalez-Zorn, B., Scortti, M., Slaghuis, J., Goebel, W. & Vazquez-Boland, J. A.** (2002). Hpt, a bacterial homolog of the microsomal glucose- 6-phosphate translocase, mediates rapid intracellular proliferation in *Listeria*. *Proc Natl Acad Sci USA*. 99, 431-436.

**Cobb, C. A., Curtis, G. D., Bansi, D. S., Slade, E., Mehal, W., Mitchell, R. G. & Chapman, R. W.** (1996). Increased prevalence of *Listeria monocytogenes* in the faeces of patients receiving long-term H2-antagonists. *Eur J Gastroenterol Hepatol*. 8, 1071-1074.

**Conner, D. E., Scott, V. N., Summer, S. S. & Bernard, D. T.** (1989). Pathogenicity of foodborne, environmental and clinical isolates of *Listeria monocytogenes* in mice. *J Food Sci*. 54, 1553-1556.

**Conte, M. P., Longhi, C., Petrone, G., Polidoro, M., Valenti, P. & Seganti, L.** (2000). Modulation of actA gene expression in *Listeria monocytogenes* by iron. *J Med Microbiol*. 49, 681-683.

**Conte, M. P., Longhi, C., Polidoro, M., Petrone, G., Buonfiglio, V., Di Santo, S., Papi,**

**E., Seganti, L., Visca, P. & Valenti, P.** (1996). Iron availability affects entry of *Listeria monocytogenes* into the enterocytelike cell line Caco-2. *Infect Immun.* 64, 3925-3929.

**Cooksley, C., Jenks, P. J., Green, A., Cockayne, A., Logan, R. P. & Hardie, K. R.** (2003). NapA protects *Helicobacter pylori* from oxidative stress damage, and its production is influenced by the ferric uptake regulator. *J Med Microbiol.* 52, 461-469.

**Cossart, P. & Bierne, H.** (2001). The use of host cell machinery in the pathogenesis of *Listeria monocytogenes*. *Curr Opin Immunol.* 13, 96-103.

**Cotter, P. D., O'Reilly, K. & Hill, C.** (2001). Role of the glutamate decarboxylase acid resistance system in the survival of *Listeria monocytogenes* LO28 in low pH foods. *J Food Prot.* 64, 1362-1368.

**Cowart, R. E. a. B. G. F.** (1981). The role of iron in the production of haemolysin by *L. monocytogenes*. *Curr Microbiol.*, 287-290.

**Cummins, A. J., Fielding, A. K. & McLauchlin, J.** (1994). *Listeria ivanovii* infection in a patient with AIDS. *J Infect.* 28, 89-91.

**Dalton, C. B., Austin, C. C., Sobel, J., Hayes, P. S., Bibb, W. F., Graves, L. M., Swaminathan, B., Proctor, M. E. & Griffin, P. M.** (1997). An outbreak of gastroenteritis and fever due to *Listeria monocytogenes* in milk. *N Engl J Med.* 336, 100-105.

**Decatur, A. L. & Portnoy, D. A.** (2000). A PEST-like sequence in listeriolysin O essential for *Listeria monocytogenes* pathogenicity. *Science.* 290, 992-995.

**Desvaux, M. & Hebraud, M.** (2006a). The protein secretion systems in *Listeria*: inside out bacterial virulence. *FEMS Microbiol Rev.* 30, 774-805.

**Desvaux, M. & Hebraud, M.** (2006b). The protein secretion systems in *Listeria*: inside out bacterial virulence. *FEMS Microbiol Rev.* 30, 774-805.

**Deutscher, J., Herro, R., Bourand, A., Mijakovic, I. & Poncet, S.** (2005a). P-Ser-HPr--a link between carbon metabolism and the virulence of some pathogenic bacteria. *Biochim Biophys Acta.* 1754, 118-125.

**Djordjevic, D., Wiedmann, M. & McLandsborough, L. A.** (2002). Microtiter plate



assay for assessment of *Listeria monocytogenes* biofilm formation. *Appl Environ Microbiol.* 68, 2950-2958.

**Doumith, M., Buchrieser, C., Glaser, P., Jacquet, C. & Martin, P.** (2004a). Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. *J Clin Microbiol.* 42, 3819-3822 .

**Doumith, M., Cazalet, C., Simoes, N., Frangeul, L., Jacquet, C., Kunst, F., Martin, P., Cossart, P., Glaser, P. & Buchrieser, C.** (2004b). New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. *Infect Immun.* 72, 1072-1083.

**Dramsi, S., Bourdichon, F., Cabanes, D., Lecuit, M., Fsihi, H. & Cossart, P.** (2004). FbpA, a novel multifunctional *Listeria monocytogenes* virulence factor. *Mol Microbiol.* 53, 639-649.

**Duche, O., Tremoulet, F., Glaser, P. & Labadie, J.** (2002a). Salt stress proteins induced in *Listeria monocytogenes*. *Appl Environ Microbiol.* 68, 1491-1498.

**Duche, O., Tremoulet, F., Namane, A. & Labadie, J.** (2002b). A proteomic analysis of the salt stress response of *Listeria monocytogenes*. *FEMS Microbiol Lett.* 215, 183-188.

**Ebe, Y., Hasegawa, G., Takatsuka, H. , Umezu, H., Mitsuyama, M., Arakawa, M., Mukaida, N. & Naito, M.** (1999 ). The role of Kupffer cells and regulation of neutrophil migration into the liver by macrophage inflammatory protein-2 in primary listeriosis in mice. *Pathol Int.* 49, 519-532.

**Fantelli, K. & Stephan, R.** (2001). Prevalence and characteristics of shigatoxin-producing *Escherichia coli* and *Listeria monocytogenes* strains isolated from minced meat in Switzerland. *Int J Food Microbiol.* 70, 63-69.

**Farber, J. M., Coates, F. & Daley, E.** (1992). Minimum water activity requirements for the growth of *Listeria monocytogenes*. *Letters in Applied Microbiology.* 15, 103-105.

**Farber, J. M., Ross, W. H. & Harwig, J.** (1996). Health risk assessment of *Listeria monocytogenes* in Canada. *Int J Food Microbiol.* 30, 145-156.

**Ferreira, A., Sue, D., O'Byrne, C. P. & Boor, K. J.** (2003). Role of *Listeria*

*monocytogenes* sigma(B) in survival of lethal acidic conditions and in the acquired acid tolerance response. *Appl Environ Microbiol.* 69, 2692-2698.

**Folio, P., Chavant, P., Chafsey, I., Belkorchia, A., Chambon, C. & Hebraud, M.** (2004). Two-dimensional electrophoresis database of *Listeria monocytogenes* EGDe proteome and proteomic analysis of mid-log and stationary growth phase cells. *Proteomics.* 4, 3187-3201.

**Franciosa, G., Pourshaban, M., Gianfranceschi, M. & Aureli, P.** (1998). Genetic typing of human and food isolates of *Listeria monocytogenes* from episodes of listeriosis. *Eur J Epidemiol.* 14, 205-210.

**Gaillard, J. L., Berche, P., Frehel, C., Guin, E. & Cossart, P.** (1991). Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. *Cell.* 65, 1127-1141.

**Gaillard, J. L., Berche, P. & Sansonetti, P.** (1986). Transposon mutagenesis as a tool to study the role of hemolysin in the virulence of *Listeria monocytogenes*. *Infect Immun.* 52, 50-55.

**Gaillard, J. L., Jaubert, F. & Berche, P.** (1996). The *inlAB* locus mediates the entry of *Listeria monocytogenes* into hepatocytes *in vivo*. *J Exp Med.* 183, 359-369.

**Gaillot, O., Bregenholt, S., Jaubert, F., Di Santo, J. P. & Berche, P.** (2001). Stress-induced ClpP serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin O-dependent protective immunity. *Infect Immun.* 69, 4938-4943 .

**Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P.** (2000). The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol Microbiol.* 35, 1286-1294.

**Gardan, R., Duche, O., Leroy-Setrin, S. & Labadie, J.** (2003). Role of *ctc* from *Listeria monocytogenes* in osmotolerance. *Appl Environ Microbiol.* 69, 154-161.

**Gilot, P., Genicot, A. & Andre, P.** (1996). Serotyping and esterase typing for analysis of *Listeria monocytogenes* populations recovered from foodstuffs and from human patients with listeriosis in Belgium. *J Clin Microbiol.* 34, 1007-1010.

**Giovannacci, I., Ragimbeau, C., Queguiner, S., Salvat, G., Vendevre, J. L., Carlier, V. & Ermel, G.** (1999). *Listeria monocytogenes* in pork slaughtering and cutting plants. Use of RAPD, PFGE and PCR-REA for tracing and molecular epidemiology. *Int J Food Microbiol.* 53, 127-140.

**Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J. C., Purcell, R., Rimmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. & Cossart, P.** (2001). Comparative genomics of *Listeria* species. *Science.* 294, 849-852.

**Gormon, T. & Phan-Thanh, L.** (1995). Identification and classification of *Listeria* by two-dimensional protein mapping. *Res Microbiol.* 146, 143-154.

**Graves, L.M., Swaminathan, B., Reeves, M.W., Hunter, S.B., Weaver, R.E., Plikaytis, B.D. & Schuchat, A.** (1994). Comparison of ribotyping and multilocus enzyme electrophoresis for subtyping of *Listeria monocytogenes* isolates. *J Clin Microbiol.* 32, 2936-2943.

**Gray, M. J., Zadoks, R. N., Fortes, E. D., Dogan, B., Cai, S., Chen, Y., Scott, V. N., Gombas, D. E., Boor, K. J. & Wiedmann, M.** (2004). *Listeria monocytogenes* isolates from foods and humans form distinct but overlapping populations. *Appl Environ Microbiol.* 70, 5833-5841.

**Grif, K., Hein, I., Wagner, M., Brandl, E., Mpamugo, O., McLauchlin, J., Dierich, M. P. & Allerberger, F.** (2001). Prevalence and characterization of *Listeria monocytogenes* in the feces of healthy Austrians. *Wien Klin Wochenschr.* 113, 737-742.

**Grimont, F. & Grimont, P. A.** (1986). Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Ann Inst Pasteur Microbiol.* 137B, 165-175.

**Grundling, A., Gonzalez, M. D. & Higgins, D. E.** (2003). Requirement of the *Listeria*

monocytogenes broad-range phospholipase PC-PLC during infection of human epithelial cells. *J Bacteriol.* 185, 6295-6307.

**Halsey, T. A., Vazquez-Torres, A., Gravdahl, D. J., Fang, F. C. & Libby, S. J.** (2004). The ferritin-like Dps protein is required for *Salmonella enterica* serovar *Typhimurium* oxidative stress resistance and virulence. *Infect Immun.* 72, 1155-1158.

**Hébraud, M. & Guzzo, J.** (2000). The main cold shock protein of *Listeria monocytogenes* belongs to the family of ferritin-like proteins. *FEMS Microbiology Letters.* 190, 29-34.

**Helloin, E., Jansch, L. & Phan-Thanh, L.** (2003). Carbon starvation survival of *Listeria monocytogenes* in planktonic state and in biofilm: a proteomic study. *Proteomics.* 3, 2052-2064.

**Herro, R., Poncet, S., Cossart, P., Buchrieser, C., Gouin, E., Glaser, P. & Deutscher, J.** (2005). How seryl-phosphorylated HPr inhibits PrfA, a transcription activator of *Listeria monocytogenes* virulence genes. *J Mol Microbiol Biotechnol.* 9, 224-234.

**Hess, J., Gentschev, I., Szalay, G., Ladel, C., Bubert, A., Goebel, W. & Kaufmann, S. H.** (1995). *Listeria monocytogenes* p60 supports host cell invasion by and in vivo survival of attenuated *Salmonella typhimurium*. *Infect Immun.* 63, 2047-2053.

**Ishikawa, T., Mizunoe, Y., Kawabata, S., Takade, A., Harada, M., Wai, S. N. & Yoshida, S.** (2003). The iron-binding protein Dps confers hydrogen peroxide stress resistance to *Campylobacter jejuni*. *J Bacteriol.* 185, 1010-1017.

**Jacquet, C., Rocourt, J. & Reynaud, A.** (1993). Study of *Listeria monocytogenes* contamination in a dairy plant and characterization of the strains isolated. *Int J Food Microbiol.* 20, 13-22.

**Jacquet, C., Doumith, M., Gordon, J. I., Martin, P. M., Cossart, P. & Lecuit, M.** (2004). A molecular marker for evaluating the pathogenic potential of foodborne *Listeria monocytogenes*. *J Infect Dis.* 189, 2094-2100.

**Jacquet, C., Gouin, E., Jeannel, D., Cossart, P. & Rocourt, J.** (2002). Expression of ActA, Ami, InlB, and listeriolysin O in *Listeria monocytogenes* of human and food origin. *Appl Environ Microbiol.* 68, 616-622.

**Jeffers, G.T., Bruce, J.L., McDonough, P.L., Scarlett, J., Boor, K.J. & Wiedmann, M.** (2001). Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases. *Microbiology*. 147, 1095-104 .

**Jeffery, C. J.** (1999). Moonlighting proteins. *Trends Biochem Sci*. 24, 8-11.

**Jonquieres, R., Bierne, H., Fiedler, F., Gounon, P. & Cossart, P.** (1999). Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of gram-positive bacteria. *Mol Microbiol*. 34, 902-914.

**Jonquieres, R., Bierne, H., Mengaud, J. & Cossart, P.** (1998). The inlA gene of *Listeria monocytogenes* LO28 harbors a nonsense mutation resulting in release of internalin. *Infect Immun*. 66, 3420-3422.

**Jonquieres, R., Pizarro-Cerda, J. & Cossart, P.** (2001). Synergy between the N- and C-terminal domains of InlB for efficient invasion of non-phagocytic cells by *Listeria monocytogenes*. *Mol Microbiol*. 42, 955-965.

**Kalmokoff, M. L., Austin, J. W., Wan, X. D., Sanders, G., Banerjee, S. & Farber, J. M.** (2001). Adsorption, attachment and biofilm formation among isolates of *Listeria monocytogenes* using model conditions. *J Appl Microbiol*. 91, 725-734.

**Kathariou, S., Metz, P., Hof, H. & Goebel, W.** (1987). Tn916-induced mutations in the hemolysin determinant affecting virulence of *Listeria monocytogenes* . *J Bacteriol*. 169, 1291-1297.

**Ko, R. & Smith, L. T.** (1999). Identification of an ATP-driven, osmoregulated glycine betaine transport system in *Listeria monocytogenes*. *Appl Environ Microbiol*. 65, 4040-4048.

**Kocks, C., Marchand, J. B., Gouin, E., d'Hauteville, H., Sansonetti, P. J., Carlier, M. F. & Cossart, P.** (1995). The unrelated surface proteins ActA of *Listeria monocytogenes* and IcsA of *Shigella flexneri* are sufficient to confer actin-based motility on *Listeria innocua* and *Escherichia coli* respectively. *Mol Microbiol*. 18, 413-423.

**Lasa, I., David, V., Gouin, E., Marchand, J. B. & Cossart, P.** (1995). The amino-terminal part of ActA is critical for the actin-based motility of *Listeria monocytogenes*; the central proline-rich region acts as a stimulator. *Mol Microbiol*. 18, 425-436.

**Lecuit, M., Dramsi, S., Gottardi, C., Fedor-Chaiken, M., Gumbiner, B. & Cossart, P.** (1999). A single amino acid in E-cadherin responsible for host specificity towards the human pathogen *Listeria monocytogenes*. *EMBO J.* 18, 3956-3963.

**Lecuit, M., Nelson, D. M., Smith, S. D., Khun, H., Huerre, M., Vacher-Lavenu, M. C., Gordon, J. I. & Cossart, P.** (2004). Targeting and crossing of the human maternofetal barrier by *Listeria monocytogenes*: role of internalin interaction with trophoblast E-cadherin. *Proc Natl Acad Sci U S A.* 101, 6152-6157.

**Lecuit, M., Vandormael-Pournin, S., Lefort, J., Huerre, M., Gounon, P., Dupuy, C., Babinet, C. & Cossart, P.** (2001). A transgenic model for listeriosis: role of internalin in crossing the intestinal barrier. *Science* . 292, 1722-1725.

**Lenz, L. L., Mohammadi, S., Geissler, A. & Portnoy, D. A.** (2003). SecA2-dependent secretion of autolytic enzymes promotes *Listeria monocytogenes* pathogenesis. *Proc Natl Acad Sci U S A.* 100, 12432-12437.

**Lessing, M. P., Curtis, G. D. & Bowler, I. C.** (1994). *Listeria ivanovii* infection. *J Infect.* 29, 230-231.

**Lety, M. A., Frehel, C., Beretti, J. L., Berche, P. & Charbit, A.** (2003). Modification of the signal sequence cleavage site of listeriolysin O does not affect protein secretion but impairs the virulence of *Listeria monocytogenes*. *Microbiology.* 149, 1249-1255.

**Liu, S., Graham, J. E., Bigelow, L., Morse, P. D. 2nd & Wilkinson, B. J.** (2002). Identification of *Listeria monocytogenes* genes expressed in response to growth at low temperature. *Appl Environ Microbiol.* 68, 1697-1705.

**Loessner, M. J.** (1991). Improved procedure for bacteriophage typing of *Listeria* strains and evaluation of new phages. *Appl Environ Microbiol.* 57, 882-884.

**Loessner, M. J. & Busse, M.** (1990). Bacteriophage typing of *Listeria* species. *Appl Environ Microbiol.* 56, 1912-1918.

**Loisel, T. P., Boujemaa, R., Pantaloni, D. & Carlier, M. F.** (1999). Reconstitution of actin-based motility of *Listeria* and *Shigella* using pure proteins. *Nature.* 401, 613-616.

**Loncarevic, S., Danielsson-Tham, M. L. & Tham, W.** (1995). Occurrence of *Listeria*

*monocytogenes* in soft and semi-soft cheeses in retail outlets in Sweden. *Int J Food Microbiol.* 26, 245-250.

**Lunden, J. M., Miettinen, M. K., Autio, T. J. & Korkeala, H. J.** (2000). Persistent *Listeria monocytogenes* strains show enhanced adherence to food contact surface after short contact times. *J Food Prot.* 63, 1204-1207.

**MacGowan, A. P., Bowker, K., McLauchlin, J., Bennett, P. M. & Reeves, D. S.** (1994). The occurrence and seasonal changes in the isolation of *Listeria* spp. in shop bought food stuffs, human faeces, sewage and soil from urban sources. *Int J Food Microbiol.* 21, 325-334 .

**Marquis, H., Doshi, V. & Portnoy, D. A.** (1995). The broad-range phospholipase C and a metalloprotease mediate listeriolysin O-independent escape of *Listeria monocytogenes* from a primary vacuole in human epithelial cells. *Infect Immun.* 63, 4531-4534.

**Marquis, H., Goldfine, H. & Portnoy, D. A.** (1997). Proteolytic pathways of activation and degradation of a bacterial phospholipase C during intracellular infection by *Listeria monocytogenes* . *J Cell Biol.* 137, 1381-1392.

**Mead, P. S., Slutsker, L., Dietz, V. , McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. & Tauxe, R. V.** (1999). Food-related illness and death in the United States. *Emerg Infect Dis.* 5, 607-625.

**Meissner, D., Vollstedt, A., van Dijl, J. M. & Freudl, R.** (2007). Comparative analysis of twin-arginine (Tat)-dependent protein secretion of a heterologous model protein (GFP) in three different Gram-positive bacteria. *Appl Microbiol Biotechnol.*

**Milenbachs, A. A., Brown, D. P., Moors, M. & Youngman, P.** (1997). Carbon-source regulation of virulence gene expression in *Listeria monocytogenes*. *Mol Microbiol.* 23, 1075-1085.

**Mohamed, W., Darji, A., Domann, E., Chiancone, E. & Chakraborty, T.** (2006). The ferritin-like protein Frm is a target for the humoral immune response to *Listeria monocytogenes* and is required for efficient bacterial survival. *Mol Genet Genomics.* 275, 344-353.

**Munk, M. E. & Kaufmann, S. H.** (1988). *Listeria monocytogenes* reactive T lymphocytes

in healthy individuals. *Microb Pathog.* 5, 49-54.

**Murray, E. G. D.** (1926). A disease of rabbits characterised by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacterium monocytogenes* (n.sp.). *J Pathol Bacteriol.* 29, 407-439.

**Nelson, K. E., Fouts, D. E., Mongodin, E. F., Ravel, J., DeBoy, R. T., Kolonay, J. F., Rasko, D. A., Angiuoli, S. V., Gill, S. R., Paulsen, I. T., Peterson, J., White, O., Nelson, W. C., Nierman, W. , Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Haft, D. H., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L. D., Uhlich, G. A., Bayles, D. O., Luchansky, J. B. & Fraser, C. M.** (2004). Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res.* 32, 2386-2395.

**Nicolas, J. A., Espaze, E. P., Catimel, B., Vidaud, N., Rocourt, J. & Courtieu, A. L.** (1989). Isolation of *Listeria* from French meat products. *Zentralbl Bakteriol.* 272, 242-247.

**Nightingale, K. K., Windham, K. & Wiedmann, M.** (2005). Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. *J Bacteriol.* 187, 5537-5551.

**Nolan, D. A., Chamblin, D. C. & Troller, J. A.** (1992). Minimal water activity levels for growth and survival of *Listeria monocytogenes* and *Listeria innocua*. *Int J Food Microbiol.* 16, 323-335.

**Nolling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y. I., Tatusov, R. L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M. J., Bennett, G. N., Koonin, E. V. & Smith, D. R.** (2001a). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol.* 183, 4823-4838.

**Nolling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y. I., Tatusov, R. L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M. J., Bennett, G. N., Koonin, E. V. & Smith, D. R.** (2001b). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol.* 183, 4823-4838.



**Norwood, D. E. & Gilmour, A.** (1999). Adherence of *Listeria monocytogenes* strains to stainless steel coupons. *J Appl Microbiol.* 86, 576-582.

**Notermans, S., Dufrenne, J., Teunis, P. & Chackraborty, T.** (1998). Studies on the risk assessment of *Listeria monocytogenes*. *J Food Prot.* 61, 244-248.

**O'Driscoll, B., Gahan, C. G. & Hill, C.** (1996). Adaptive acid tolerance response in *Listeria monocytogenes*: isolation of an acid-tolerant mutant which demonstrates increased virulence. *Applied and Environmental Microbiology.* 62, 1693-1698.

**Okamoto, M., Nakane, A. & Minagawa, T.** (1994). Host resistance to an intragastric infection with *Listeria monocytogenes* in mice depends on cellular immunity and intestinal bacterial flora. *Infect Immun.* 62, 3080-3085.

**Olier, M., Pierre, F., Lemaitre, J. P., Divies, C., Rousset, A. & Guzzo, J.** (2002). Assessment of the pathogenic potential of two *Listeria monocytogenes* human faecal carriage isolates. *Microbiology.* 148, 1855-1862.

**Olier, M., Pierre, F., Rousseaux, S., Lemaitre, J. P., Rousset, A., Piveteau, P. & Guzzo, J.** (2003). Expression of truncated Internalin A is involved in impaired internalization of some *Listeria monocytogenes* isolates carried asymptotically by humans. *Infect Immun.* 71, 1217-1224.

**Olsen, K. N., Larsen, M. H., Gahan, C. G., Kallipolitis, B., Wolf, X. A., Rea, R., Hill, C. & Ingmer, H.** (2005). The Dps-like protein Fri of *Listeria monocytogenes* promotes stress tolerance and intracellular multiplication in macrophage-like cells. *Microbiology.* 151, 925-933.

**Pagan, R., Condon, S. & Sala, F. J.** (1997). Effects of several factors on the heat-shock-induced thermotolerance of *Listeria monocytogenes*. *Appl Environ Microbiol.* 63, 3225-3232.

**Park, J. H., Lee, Y. S., Lim, Y. K., Kwon, S. H., Lee, C. U. & Yoon, B. S.** (2000). Specific binding of recombinant *Listeria monocytogenes* p60 protein to Caco-2 cells. *FEMS Microbiol Lett.* 186, 35-40.

**Pearson, L. J. & Marth, E. H.** (1990). *Listeria monocytogenes*--threat to a safe food supply: a review. *J Dairy Sci.* 73, 912-928.

**Periago, P. M., van Schaik, W., Abee, T. & Wouters, J. A.** (2002). Identification of proteins involved in the heat stress response of *Bacillus cereus* ATCC 14579. *Appl Environ Microbiol.* 68, 3486-3495.

**Phan-Thanh, L. & Gormon, T.** (1995). Analysis of heat and cold shock proteins in *Listeria* by two-dimensional electrophoresis. *Electrophoresis.* 16, 444-450.

**Phan-Thanh, L. & Gormon, T.** (1997). Stress proteins in *Listeria monocytogenes*. *Electrophoresis.* 18, 1464-1471.

**Phan-Thanh, L. & Mahouin, F.** (1999). A proteomic approach to study the acid response in *Listeria monocytogenes*. *Electrophoresis.* 20, 2214-2224.

**Phan-Thanh, L., Mahouin, F. & Alige, S.** (2000). Acid responses of *Listeria monocytogenes*. *Int J Food Microbiol.* 55, 121-126.

**Piffaretti, J. C., Kressebuch, H., Aeschbacher, M., Bille, J., Bannerman, E., Musser, J. M., Selander, R. K. & Rocourt, J.** (1989). Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. *Proc Natl Acad Sci U S A.* 86, 3818-3822.

**Pine, L., Kathariou, S., Quinn, F., George, V., Wenger, J. D. & Weaver, R. E.** (1991). Cytopathogenic effects in enterocytelike Caco-2 cells differentiate virulent from avirulent *Listeria* strains. *J Clin Microbiol.* 29, 990-996.

**Pirie, J. H. H.** (1940). *Listeria*: change of name for a genus of bacteria. *Nature.* 154, 264.

**Polidoro, M., De Biase, D., Montagnini, B., Guarrera, L., Cavallo, S., Valenti, P., Stefanini, S. & Chiancone, E.** (2002). The expression of the dodecameric ferritin in *Listeria spp.* is induced by iron limitation and stationary growth phase. *Gene.* 296, 121.

**Portnoy, D. A., Chakraborty, T., Goebel, W. & Cossart, P.** (1992). Molecular determinants of *Listeria monocytogenes* pathogenesis. *Infect Immun.* 60, 1263-1267.

**Promadej, N., Fiedler, F., Cossart, P., Dramsi, S. & Kathariou, S.** (1999). Cell wall teichoic acid glycosylation in *Listeria monocytogenes* serotype 4b requires gtcA, a novel, serogroup-specific gene. *J Bacteriol.* 181, 418-425.

**Ramnath, M., Rechinger, K. B., Jansch, L., Hastings, J. W., Knochel, S. & Gravesen, A.** (2003). Development of a *Listeria monocytogenes* EGDe partial proteome reference map and comparison with the protein profiles of food isolates. *Appl Environ Microbiol.* 69, 3368-3376.

**Raveneau, J., Geoffroy, C., Beretti, J. L., Gaillard, J. L., Alouf, J. E. & Berche, P.** (1992). Reduced virulence of a *Listeria monocytogenes* phospholipase-deficient mutant obtained by transposon insertion into the zinc metalloprotease gene. *Infect Immun.* 60, 916-921.

**Reinders, J. & Sickmann, A.** (2005). State-of-the-art in phosphoproteomics. *Proteomics.* 5, 4052-4061.

**Ripabelli, G., McLauchin, J. & Threlfall, E.J.** (2000 ). Amplified fragment length polymorphism (AFLP) analysis of *Listeria monocytogenes*. *Syst Appl Microbiol.* 23, 132-6.

**Roche, S. M., Gracieux, P., Milohanic, E., Albert, I., Virlogeux-Payant, I., Temoin, S., Grepinet, O., Kerouanton, A., Jacquet, C., Cossart, P. & Velge, P.** (2005). Investigation of specific substitutions in virulence genes characterizing phenotypic groups of low-virulence field strains of *Listeria monocytogenes*. *Appl Environ Microbiol.* 71, 6039-6048.

**Rocourt, J., Audurier, A., Courtieu, A. L., Durst, J., Ortel, S., Schrettenbrunner, A. & Taylor, A. G.** (1985). A multi-centre study on the phage typing of *Listeria monocytogenes*. *Zentralbl Bakteriol Mikrobiol Hyg [A].* 259, 489-497.

**Rocourt, J., Jacquet, C. & Reilly, A.** (2000). Epidemiology of human listeriosis and seafoods. *Int J Food Microbiol.* 62, 197-209.

**Rouquette, C., Ripio, M. T., Pellegrini, E., Bolla, J. M., Tascon, R. I., Vazquez-Boland, J. A. & Berche, P.** (1996). Identification of a ClpC ATPase required for stress tolerance and in vivo survival of *Listeria monocytogenes*. *Mol Microbiol.* 21, 977-987.

**Salvat, G., Coppen, P., Allo, J. C., Fenner, S., Laisney, M. J., Toquin, M. T., Humbert, F. & Colin, P.** (1997). Effects of AvGard treatment on the microbiological flora of poultry carcasses. *Br Poult Sci.* 38, 489-498.

**Sauders, B. D., Pettit, D., Currie, B., Suits, P., Evans, A., Stellrecht, K., Dryja, D. M.,**

**Slate, D. & Wiedmann, M.** (2005). Low prevalence of *Listeria monocytogenes* in human stool. *J Food Prot.* 68, 178-181.

**Schlech, W. F., Lavigne, P. M., Bortolussi, R. A., Allen, A. C., Haldane, E. V., Wort, A. J., Hightower, A. W., Johnson, S. E., King, S. H., Nicholls, E. S. & Broome, C. V.** (1983). Epidemic listeriosis-evidence for transmission by food. *New England Journal of Medecine.* 308, 203-206.

**Schluter, D., Domann, E., Buck, C., Hain, T., Hof, H., Chakraborty, T. & Deckert-Schluter, M.** (1998). Phosphatidylcholine-specific phospholipase C from *Listeria monocytogenes* is an important virulence factor in murine cerebral listeriosis. *Infect Immun.* 66, 5930-5938.

**Schmid, M. W., Ng, E. Y., Lampidis, R., Emmerth, M., Walcher, M., Kreft, J., Goebel, W., Wagner, M. & Schleifer, K. H.** (2005). Evolutionary history of the genus *Listeria* and its virulence genes. *Syst Appl Microbiol.* 28, 1-18.

**Schnupf, P., Portnoy, D. A. & Decatur, A. L.** (2006). Phosphorylation, ubiquitination and degradation of listeriolysin O in mammalian cells: role of the PEST-like sequence. *Cell Microbiol.* 8, 353-364.

**Schonberg, A., Teufel, P. & Weise, E.** (1989). Serovars of *Listeria monocytogenes* and *Listeria innocua* from food. *Acta Microbiol Hung.* 36, 249-253.

**Seeliger, H. P. R. & Höhne, K.** (1979). Serotyping of *Listeria monocytogenes* and related species. In *Methods in microbiology*, Academic press. pp.31-49.

**Seeliger, H. P. R. & Jones, D.** (1987 ). *Listeria*. In *Bergey's Manual of Systematic Bacteriology*, Eds 9th, Williams and Wilkins, Baltimore.

**Shen, Y., Naujokas, M., Park, M. & Ireton, K.** (2000). InIB-dependent internalization of *Listeria* is mediated by the Met receptor tyrosine kinase. *Cell.* 103, 501-10.

**Sim, J., Hood, D., Finnie, L., Wilson, M., Graham, C., Brett, M. & Hudson, J. A.** (2002). Series of incidents of *Listeria monocytogenes* non-invasive febrile gastroenteritis involving ready-to-eat meats. *Lett Appl Microbiol.* 35, 409-413.

**Sleator, R. D., Wouters, J., Gahan, C. G., Abee, T. & Hill, C.** (2001). Analysis of the

role of OpuC, an osmolyte transport system, in salt tolerance and virulence potential of *Listeria monocytogenes*. *Appl Environ Microbiol.* 67, 2692-2698.

**Smith, G. A., Marquis, H., Jones, S. , Johnston, N. C., Portnoy, D. A. & Goldfine, H.** (1995). The two distinct phospholipases C of *Listeria monocytogenes* have overlapping roles in escape from a vacuole and cell-to-cell spread. *Infect Immun.* 63, 4231-4237.

**Snapir, Y. M., Vaisbein, E. & Nassar, F.** (2006). Low virulence but potentially fatal outcome-*Listeria ivanovii*. *Eur J Intern Med.* 17, 286-287.

**Snyder, A. & Marquis, H.** (2003). Restricted translocation across the cell wall regulates secretion of the broad-range phospholipase C of *Listeria monocytogenes*. *J Bacteriol.* 185, 5953-5958.

**Sokolovic, Z., Fuchs, A. & Goebel, W.** (1990). Synthesis of species-specific stress proteins by virulent strains of *Listeria monocytogenes*. *Infect Immun.* 58, 3582-3587.

**Sokolovic, Z., Riedel, J., Wuenscher, M. & Goebel, W.** (1993). Surface-associated, PrfA-regulated proteins of *Listeria monocytogenes* synthesized under stress conditions. *Mol Microbiol.* 8, 219-227.

**Stack, H. M., Sleator, R. D., Bowers, M., Hill, C. & Gahan, C. G.** (2005). Role for HtrA in stress induction and virulence potential in *Listeria monocytogenes*. *Appl Environ Microbiol.* 71, 4241-4247.

**Stefanini, S., Cavallo, S., Montagnini, B. & Chiancone, E.** (1999). Incorporation of iron by the unusual dodecameric ferritin from *Listeria innocua*. *Biochem J.* 338 (Pt 1), 71-75.

**Su, M., Cavallo, S., Stefanini, S., Chiancone, E. & Chasteen, N. D.** (2005). The so-called *Listeria innocua* ferritin is a Dps protein. Iron incorporation, detoxification, and DNA protection properties. *Biochemistry.* 44, 5572-5578.

**Sword, C. P.** (1966). Mechanisms of pathogenesis in *Listeria monocytogenes* infection. I. Influence of iron. *J Bacteriol.* 92, 536-542.

**Tabouret, M., de Rycke, J. & Dubray, G.** (1992). Analysis of surface proteins of *Listeria* in relation to species, serovar and pathogenicity. *J Gen Microbiol.* 138, 743-753.

**Terplan, G. & Steinmeyer, S.** (1989). Investigations on the pathogenicity of *Listeria* spp. by experimental infection of the chick embryo. *Int J Food Microbiol.* 8, 277-280.

**Thevenot, D., Delignette-Muller, M. L., Christieans, S. & Vernozy-Rozand, C.** (2005). Prevalence of *Listeria monocytogenes* in 13 dried sausage processing plants and their products. *Int J Food Microbiol.* 102, 85-94.

**Tremoulet, F., Duche, O., Namane, A., Martinie, B. & Labadie, J. C.** (2002). Comparison of protein patterns of *Listeria monocytogenes* grown in biofilm or in planktonic mode by proteomic analysis. *FEMS Microbiol Lett.* 210, 25-31.

**Trost, M., Wehmhoner, D., Karst, U., Dieterich, G., Wehland, J. & Jansch, L.** (2005). Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species. *Proteomics.* 5, 1544-1557.

**Ueshima, J., Shoji, M., Ratnayake, D. B., Abe, K., Yoshida, S., Yamamoto, K. & Nakayama, K.** (2003). Purification, gene cloning, gene expression, and mutants of Dps from the obligate anaerobe *Porphyromonas gingivalis*. *Infect Immun.* 71, 1170-1178.

**Vaillant, V., de Valk, H., Baron, E., Ancelle, T., Colin, P., Delmas, M. C., Dufour, B., Pouillot, R., Le Strat, Y., Weinbreck, P., Jouglu, E. & Desenclos, J. C.** (2005). Foodborne infections in France. *Foodborne Pathog Dis.* 2, 221-232.

**Van Kessel, J. S., Karns, J. S., Gorski, L., McCluskey, B. J. & Perdue, M. L.** (2004). Prevalence of *Salmonellae*, *Listeria monocytogenes*, and fecal coliforms in bulk tank milk on US dairies. *J Dairy Sci.* 87, 2822-2830 .

**Van Langendonck, N., Bottreau, E., Bailly, S., Tabouret, M., Marly, J., Pardon, P. & Velge, P.** (1998). Tissue culture assays using Caco-2 cell line differentiate virulent from non-virulent *Listeria monocytogenes* strains. *J Appl Microbiol.* 85, 337-346.

**van Schaik, W. & Abee, T.** (2005). The role of sigmaB in the stress response of Gram-positive bacteria -- targets for food preservation and safety. *Curr Opin Biotechnol.* 16, 218-224.

**Vazquez-Boland, J. A., Kocks, C., Dramsi, S., Ohayon, H., Geoffroy, C., Mengaud, J. & Cossart, P.** (1992). Nucleotide sequence of the lecithinase operon of *Listeria*

*monocytogenes* and possible role of lecithinase in cell-to-cell spread. *Infect Immun.* 60, 219-230.

**Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J. & Kreft, J.** (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clin Microbiol Rev.* 14, 584-640.

**Vega, Y., Dickneite, C., Ripio, M. T., Bockmann, R., Gonzalez-Zorn, B., Novella, S., Dominguez-Bernal, G., Goebel, W. & Vazquez-Boland, J. A.** (1998). Functional similarities between the *Listeria monocytogenes* virulence regulator PrfA and cyclic AMP receptor protein: the PrfA\* (Gly145Ser) mutation increases binding affinity for target DNA. *J Bacteriol.* 180, 6655-6660.

**Walker, J. K., Morgan, J. H., McLauchlin, J., Grant, K. A. & Shallcross, J. A.** (1994). *Listeria innocua* isolated from a case of ovine meningoencephalitis. *Vet Microbiol.* 42, 245-253.

**Ward, R. L., Bernstein, D. I., Young, E. C., Sherwood, J. R., Knowlton, D. R. & Schiff, G. M.** (1986). Human rotavirus studies in volunteers: determination of infectious dose and serological response to infection. *J Infect Dis.* 154, 871-880.

**Ward, T. J., Gorski, L., Borucki, M. K., Mandrell, R. E., Hutchins, J. & Pupedis, K.** (2004). Intraspecific phylogeny and lineage group identification based on the prfA virulence gene cluster of *Listeria monocytogenes*. *J Bacteriol.* 186, 4994-5002.

**Way, S. S. & Wilson, C. B.** (2005). The *Mycobacterium tuberculosis* ESAT-6 homologue in *Listeria monocytogenes* is dispensable for growth *in vitro* and *in vivo*. *Infect Immun.* 73, 6151-6153.

**Wemekamp-Kamphuis, H. H., Wouters, J. A., Sleator, R. D., Gahan, C. G., Hill, C. & Abee, T.** (2002). Multiple deletions of the osmolyte transporters BetL, Gbu, and OpuC of *Listeria monocytogenes* affect virulence and growth at high osmolarity. *Appl Environ Microbiol.* 68, 4710-4716.

**Werbrouck, H., Grijspeerdt, K., Botteldoorn, N., Van Pamel, E., Rijpens, N., Van Damme, J., Uyttendaele, M., Herman, L. & Van Coillie, E.** (2006). Differential inlA and inlB expression and interaction with human intestinal and liver cells by *Listeria*

*monocytogenes* strains of different origins. *Appl Environ Microbiol.* 72, 3862-3871 .

**Wiedmann, M., Arvik, T. J., Hurley, R. J. & Boor, K. J.** (1998). General stress transcription factor sigmaB and its role in acid tolerance and virulence of *Listeria monocytogenes*. *J Bacteriol.* 180, 3650-3656.

**Wiedmann, M., Bruce, J. L., Keating, C., Johnson, A. E., McDonough, P. L. & Batt, C. A.** (1997). Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. *Infect Immun.* 65, 2707-2716.

**Wilesmith, J. W. & Gitter, M.** (1986). Epidemiology of ovine listeriosis in Great Britain. *Vet Rec.* 119, 467-470.

**Wilson, R. L., Brown, L. L., Kirkwood-Watts, D., Warren, T. K., Lund, S. A., King, D. S., Jones, K. F. & Hruby, D. E.** (2006). *Listeria monocytogenes* 10403S HtrA is necessary for resistance to cellular stress and virulence. *Infect Immun.* 74, 765-8.

**Wonderling, L. D., Wilkinson, B. J. & Bayles, D. O.** (2004). The htrA (degP) gene of *Listeria monocytogenes* 10403S is essential for optimal growth under stress conditions. *Appl Environ Microbiol.* 70, 1935-1943.

**Wuenscher, M. D., Kohler, S., Bubert, A., Gerike, U. & Goebel, W.** (1993). The iap gene of *Listeria monocytogenes* is essential for cell viability, and its gene product, p60, has bacteriolytic activity. *J Bacteriol.* 175, 3491-3501.

**Yamamoto, Y., Poole, L. B., Hantgan, R. R. & Kamio, Y.** (2002). An iron-binding protein, Dpr, from *Streptococcus mutans* prevents iron-dependent hydroxyl radical formation *in vitro*. *J Bacteriol.* 184, 2931-2939.

**Yang, X., Chiancone, E., Stefanini, S., Ilari, A. & Chasteen, N. D.** (2000). Iron oxidation and hydrolysis reactions of a novel ferritin from *Listeria innocua*. *Biochem J.* 349 (Pt 3), 783-786.

**Yeung, P. S., Zagorski, N. & Marquis, H.** (2005). The metalloprotease of *Listeria monocytogenes* controls cell wall translocation of the broad-range phospholipase C. *J Bacteriol.* 187, 2601-2608.



**Zhang, C., Zhang, M., Ju, J., Nietfeldt, J., Wise, J., Terry, P. M., Olson, M., Kachman, S. D., Wiedmann, M., Samadpour, M. & Benson, A. K.** (2003). Genome diversification in phylogenetic lineages I and II of *Listeria monocytogenes*: identification of segments unique to lineage II populations. *J Bacteriol.* 185, 5573-5584.

**Zhao, G., Ceci, P., Ilari, A., Giangiacomo, L., Laue, T. M., Chiancone, E. & Chasteen, N. D.** (2002). Iron and hydrogen peroxide detoxification properties of DNA-binding protein from starved cells. A ferritin-like DNA-binding protein of *Escherichia coli*. *J Biol Chem.* 277, 27689-27696.

# **PUBLICATIONS ET COMMUNICATIONS**

▪ **Articles scientifiques en deuxième auteur**

- Desvaux M, **Dumas E**, Chafsey I, Hebraud M :Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiol Lett.* 2006, 256: 1-15

- Meunier B, **Dumas E**, Picc I, Béchet D, Hébraud M\*<sup>1</sup>, Hocquette JF : Assessment of hierarchical clustering methodologies for proteomic data mining. *J. Proteome Res.* 2007, 6 (1), 358-366

▪ **Communications orales**

- **Dumas E**, Chafsey I, Chambon C Hebraud M Approche protéomique pour la caractérisation des souches de *Listeria monocytogenes* : recherche de marqueurs de risques – Journées de l'école doctorale 2005

- **Dumas E**, Meunier B, Berdague J.L., Chambon C, Chafsey I, Hebraud M : Approche protéomique pour la caractérisation des souches de *Listeria monocytogenes* : recherche de marqueurs de risques – 3<sup>e</sup> rencontre des microbiologistes du pôle clermontois (avril 2006).

- **Dumas E**, Meunier B, Berdague J.L., Chambon C, Chafsey I, Hebraud M : Comparison of two subproteomes from 12 strains of the foodborne pathogen *Listeria monocytogenes*: research of risk markers – Congrès de la Société Française d'Electrophorèse et d'Analyse protéique (SFEAP) 2006 . Saint-Malo, 16-18 octobre 2006.

▪ **Communications affichées**

- **Dumas E**, Dussurget O, Archambaud C, Chafsey I, Chambon C, Hébraud M, Cossart P. :Caracterisation du mutant ferritine (*Afri*) de *Listeria monocytogenes* EGDe - VI<sup>ème</sup> Congrès de la Société Française de Microbiologie. Bordeaux, 10-12 mai 2004.

- **Dumas E**, Dussurget O, Archambaud C, Chafsey I, Chambon C, Hébraud M, Cossart P. : Characterization of the ferritin mutant (*Afri*) of *Listeria monocytogenes* EGDe - 10<sup>èmes</sup> Journées de Gerland "Proteomics: New insight and prospects". Lyon, 8-9 novembre 2004

- **Dumas E**, Chambon C, Hebraud M : Approche protéomique pour la caractérisation de la biodiversité de *Listeria monocytogenes* : recherche de marqueurs de risques - 1<sup>er</sup> Symposium de Chimie et Biologie Analytiques. Montpellier, 26-29 septembre 2005.

- **Dumas E**, Chafsey I, Planchon S, Chambon C, Hebraud M : Base de données de gels d'électrophorèse bidimensionnelle du protéome membranaire et du sécrétome de *Listeria monocytogenes* EGDe - 1<sup>er</sup> Symposium de Chimie et Biologie Analytiques. Montpellier, 26-29 septembre 2005.

- **Dumas E**, Meunier B, Berdague J.L., Chambon C, Chafsey I, Hebraud M : Approche protéomique pour la caractérisation des souches de *Listeria monocytogenes* : recherche de marqueurs de risques – Rencontre des microbiologistes de l'INRA – Dourdan, juin 2006.

## Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure

Mickaël Desvaux, Emilie Dumas, Ingrid Chafsey & Michel Hébraud

Institut National de la Recherche Agronomique, Centre de Recherche Clermont-Ferrand - Theix - Lyon, Unité de Microbiologie, Equipe Qualité et Sécurité des Aliments, Site de Theix, Saint-Genès Champanelle, France

**Correspondence:** Mickaël Desvaux, Institut National de la Recherche Agronomique (INRA), Centre de Recherche Clermont - Ferrand - Theix - Lyon, UR 454-Unité de Microbiologie, Equipe Qualité et Sécurité des Aliments (QuaSA), Site de Theix, F-63122 Saint-Genès Champanelle, France. Tel.: +33 04 73 62 47 23; fax: +33 04 73 62 42 68; e-mail: mdesvaux@sancy.clermont.inra.fr

Received 21 October 2005; revised 8 December 2005; accepted 21 December 2005.  
First published online February 2006.

doi:10.1111/j.1574-6968.2006.00122.x

Editor: Ian Henderson

### Keywords

Gram-positive bacteria; cell surface display; membrane protein; cell wall anchoring; S-layer protein; bacterial surface organelle; bacterial protein secretion.

### Introduction

It was in 1884 that the Danish pharmacologist and physician Hans Christian Joachim Gram published the most-famous staining method in the world of bacteriology (Gram, 1884). With this method, slightly perfected over the years, bacteria are separated into two groups, i.e. retaining or not the stain and thus called Gram-positive and Gram-negative bacteria, respectively. The difference in staining was further related to profound divergence in structural organization of the bacterial cell envelope. While the cytoplasmic membrane of Gram-negative bacteria is surrounded by a thin cell wall beneath the outer membrane, i.e. two distinct membranes (diderm prokaryotes), Gram-positive have a thick cell wall but lack an outer membrane, i.e. possess a single cell membrane (monoderm prokaryotes) (Gupta, 1998).

Molecular analyses further revealed that contrary to Gram-negative bacteria, Gram-positive bacteria corresponds to a phylogenetically coherent grouping of prokaryotes within the domain *Bacteria* (Woese, 1987). In the

### Abstract

In the course of evolution, Gram-positive bacteria, defined here as prokaryotes from the domain *Bacteria* with a cell envelope composed of one biological membrane (monoderm) and a cell wall composed at least of peptidoglycan and covalently linked teichoic acids, have developed several mechanisms permitting to a cytoplasmic synthesized protein to be present on the bacterial cell surface. Four major types of cell surface displayed proteins are currently recognized: (i) transmembrane proteins, (ii) lipoproteins, (iii) LPXTG-like proteins and (iv) cell wall binding proteins. The subset of proteins exposed on the bacterial cell surface, and thus interacting with extracellular milieu, constitutes the surfaceome. Here, we review exhaustively the current molecular mechanisms involved in protein attachment within the cell envelope of Gram-positive bacteria, from single protein to macromolecular protein structure.

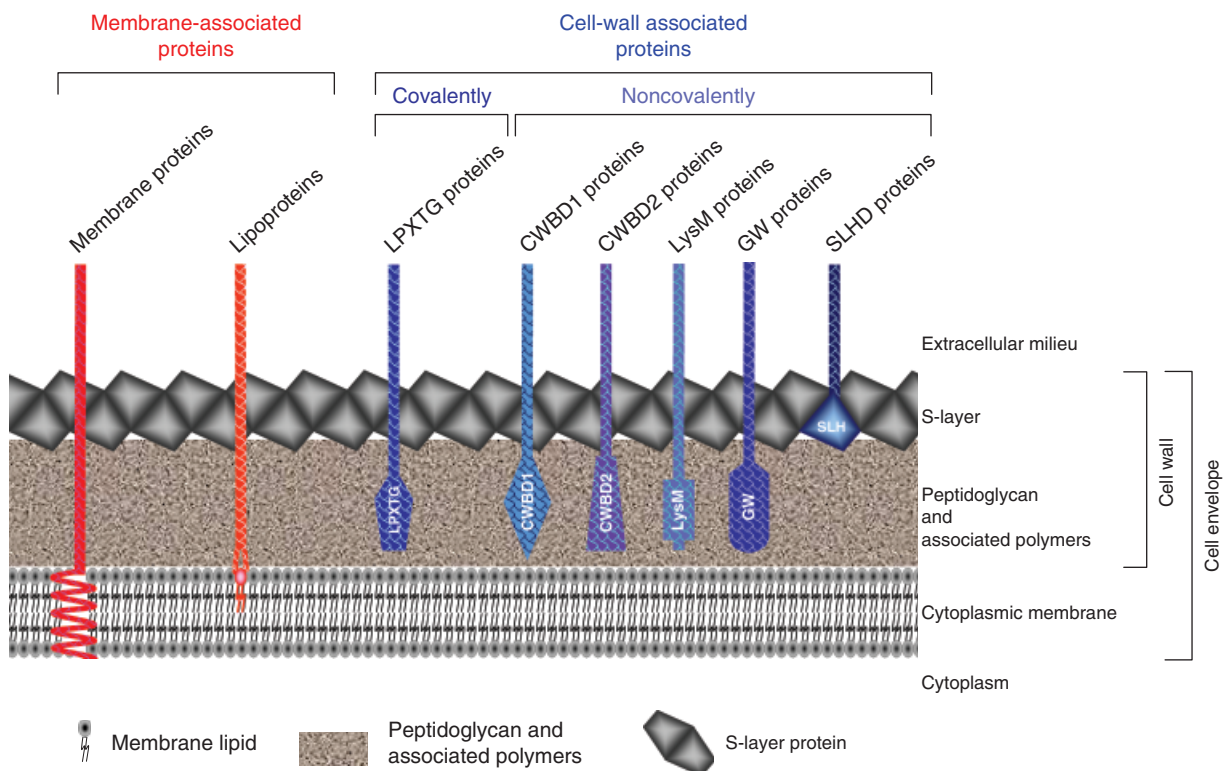
Bergey's Manual (Garrity, 2001), Gram-positive bacteria are divided into those that have low G+C moles percent, i.e. phylum BXIII *Firmicutes* and those that have high G+C moles percent, i.e. phylum BXIV *Actinobacteria*. Importantly, some of these Gram-positive bacteria do not retain the Gram stain, either because of (i) a too thin peptidoglycan layer, e.g. in the genus *Clostridium*, (ii) the absence of a cell wall, e.g. in the genus *Mycoplasma* or (iii) the presence of a waxy outer sheath preventing penetration of the stain, e.g. in the genus *Mycobacterium*. More confusingly, some bacteria not taxonomically-related to Gram-positive bacteria retain the Gram staining, e.g. some members of the phylum BIV *Deinococcus-Thermus*. Some bacteria clearly possessing a Gram-negative-like cell envelope architecture are in fact phylogenetically related to the taxonomic group of Gram-positive bacteria, e.g. *Thermotoga maritima* in phylum BII *Thermotogae* (Snel *et al.*, 2005) or *Fusobacterium nucleatum* in phylum BXXI *Fusobacteria* (Mira *et al.*, 2004). Some other phyla, like BVI *Chloroflexi* or BVII *Thermomicrobia*, regroup both monoderm and diderm bacteria (Botero *et al.*, 2004).

Even in some deep branches of the phylum *Firmicutes*, some bacteria clearly exhibit Gram-negative cell envelope (Shatkin, 2004). Therefore, it appears clearly that when employing the term ‘Gram-positive bacteria’ it is extremely important to specify if we refer to a Gram staining result, a cell envelope organization and/or a taxonomic group.

For the purpose of the present review, the term Gram-positive will be used to describe bacteria with a cell envelope composed of (i) a single biological membrane (monodermite), i.e. the cytoplasmic membrane and (ii) a cell wall composed at least of peptidoglycan, consisting of linear polysaccharide chains cross-linked by short peptides, covalently linked to teichoic acids, i.e. polyol phosphate polymers, not found in Gram-negative bacteria (Shockman & Barrett, 1983). Besides peptidoglycan, the rigid cell wall of Gram-positive bacteria contains large amounts of wall-associated polymers, also called ‘secondary’ cell wall polymers (SCWPs), which can be classified into three distinct groups: (i) teichuronic acids, (ii) teichoic acids, including lipoteichoic acids and (iii) other neutral or acidic polysaccharides, which cannot be assigned to the two former

groups, e.g. lipoglycans. The SCWPs present in various proportions are either covalently linked to the peptidoglycan backbone, e.g. teichuronic acids or tethered to a lipid anchor moiety, e.g. lipoglycans and lipoteichoic acids (Fischer, 1994). Except for teichoic and teichuronic acids, the structure and biosynthesis of other SCWPs are largely unknown.

Proteins displayed on the cell surface of Gram-positive bacteria must be attached to the cell envelope, i.e. either to the cytoplasmic membrane or cell wall components. Conversely, protein localization within the cell envelope is no guarantee that it is cell surface exposed. Four major types of cell surface displayed proteins are currently recognized: (i) proteins anchored to the cytoplasmic membrane by hydrophobic transmembrane domain(s), (ii) lipoproteins which are covalently attached to membrane lipids after cleavage by signal peptidase II, (iii) proteins containing C-terminal LPXTG-like motif and covalently attached to peptidoglycan by sortase and (iv) proteins recognizing some cell wall components by specific domains, i.e. some cell wall binding domains (Fig. 1). These basic anchoring mechanisms can further serve as foundation for the assembly of much more



**Fig. 1.** Schematic representation of the different types of surface proteins found in Gram-positive bacteria. Membrane-associated proteins can either (i) possess transmembrane domain(s) or (ii) be lipoproteins, and thus be covalently attached to long chain fatty acids of the cytoplasmic membrane. Cell wall-associated proteins can either (i) possess C-terminal LPXTG motif (or related motifs) and then be covalently anchored to peptidoglycan via sortase activity, or (ii) possess cell wall binding domain(s) (CWBDs) and thus remain stuck into the cell wall; four CWBDs are currently characterized, i.e. CWBD1, CWBD2, LysM, GW modules and SLHD. Importantly, since cell wall of Gram-positive bacteria is permeable, proteins can interact with their environment without ever having a domain poking out confines of the cell wall.

complex protein macromolecular structures, namely S-layer, pilus, flagellum or cellulosome. The present review will focus on the surfaceome in Gram-positive bacteria, i.e. the subset of proteins exposed on the surface of a bacterial cell (Cullen *et al.*, 2005), by describing how proteins can remain attached to the cell envelope.

## Cytoplasmic membrane anchoring of cell surface displayed proteins

### Mechanisms involving YidC homologues

In *Escherichia coli*, all inner membrane proteins are considered YidC-dependent (Froderberg *et al.*, 2003; Luirink *et al.*, 2005). In *Bacillus subtilis*, two homologues of YidC are found, i.e. SpoIIIJ and YqjG, and have been demonstrated to work independently for the insertion of membrane proteins (Tjalsma *et al.*, 2003, 2004). In *E. coli*, membrane protein integration into inner membrane can occur in two ways, involving either YidC alone or in association with the Sec (secretion) translocon, i.e. the SecYEGDFYaiC complex (Dalbey & Chen, 2004). The Sec-independent function of YidC is conserved and essential for bacterial growth.

YidC plays a major role in the folding step of transmembrane-spanning domains but the exact mechanism is still unclear (Dalbey & Chen, 2004; Van der Laan *et al.*, 2005). YidC would facilitate the membrane insertion by (i) providing a special amphiphilic surface overcoming the repulsion of the hydrophobic proteins and (ii) protecting polar residues against the hydrocarbon core of the membrane. The translocated protein would be stabilized and then folded by contact with YidC; transmembrane segments could also fold and interact with each other within the machinery. YidC would also function as an assembly site for hydrophobic domains and thus would attach individual subunits of a multi-component membrane complex.

Several regions of protein substrate play a role in the insertion into the bilayer membrane. The N-terminal signal peptide is either cleaved and removed from the precursor protein by SPase I (signal peptidase I) or uncleaved to act as membrane anchor domain (Tjalsma *et al.*, 2000, 2004). In the latter case, stop-transfer sequence halts the translocation of the protein chain. In addition, it was shown that the orientation of single spanning domain depends on topogenic elements carried by the protein. Type I and Type II membrane proteins initiate respectively the Nout-Cin or Nin-Cout topologies (Dalbey & Chen, 2004). The orientation would also depend on the presence of positively charged residues. Membrane protein can be integrated into the lipid bilayer in two ways, either hydrophobic segments (i) are inserted one at a time or (ii) are inserted by pairs, which corresponds to the helical hairpins model.

Cell surface displayed proteins anchored in the cytoplasmic membrane of Gram-positive bacteria have large variety of function ranging from metabolism of carbohydrates, proteins, lipids, nucleotides to signal transductions or virulence factors (Tjalsma *et al.*, 2000, 2004; Bunai *et al.*, 2005).

### Lipoproteins

Lipoproteins are characterized by a specific signal peptide of class 2 exhibiting a conserved lipobox sequence. Lipobox includes invariably a cysteine residue located just after the cleavage site (Tjalsma *et al.*, 2004). After translocation of the prolipoprotein precursor, prolipoprotein diacylglyceryl transferase (Lgt) adds the diacyl glyceryl group from a glycerophospholipid to the SH-group of cysteine residue of the lipobox (Tjalsma *et al.*, 1999). The diacylglyceryl group inserts into the lipid bilayer of the cytoplasmic membrane, preventing release of the protein. Subsequently, signal peptidase II (SPase II) cleaves off the signal peptide releasing the mature lipoprotein (Tjalsma *et al.*, 2000). Lipoprotein is further amino-acylated at the amino-terminal Cys residue by phospholipid/apolipoprotein transacylase (Lnt) resulting in the protein anchoring to membrane long chain fatty acid (Sankaran & Wu, 1994). In Gram-positive bacteria, Lnt homologues are not identifiable in many sequenced genomes and thus the extent of N-terminal amino acylation of lipoproteins is unclear.

As for the cell surface displayed transmembrane proteins, lipoproteins can be involved in a large variety of physiological functions, i.e. adhesins, transporters, receptors, enzymes or virulence factors (Sutcliffe & Russell, 1995).

## Cell wall anchoring of cell surface displayed proteins

### Covalent binding via sortase

In Gram-positive bacteria, a range of proteins is covalently link to the cell wall by sortases. In addition, to a Sec-dependent N-terminal signal peptide, protein substrates of sortases contain a C-terminal cell wall sorting signal, consisting of a LPXTG-like motif, a hydrophobic domain and a positively charged tail (Navarre & Schneewind, 1999). After cleavage of the Sec-signal peptide, the translocated protein is retained within the membrane by the hydrophobic domain. Then, a membrane-anchored sortase attacks the Thr-Gly bond of the LPXTG motif, capturing cleaved polypeptide as a thioester-linked acyl enzyme at its active site Cys residue. This complex is subsequently resolved by the nucleophilic attack of the amino group of the cross bridge within lipid II precursor. Surface proteins linked to lipid II are incorporated into the cell wall by the transglycosylation and transpeptidation reactions that generate peptidoglycan.

Sortase-like proteins have been identified in almost all Gram-positive bacteria (Pallen *et al.*, 2001). Based on phylogenetic criteria, sortases have been recently classified within 4 classes (Comfort & Clubb, 2004; Dramsi *et al.*, 2005). Class A includes prototypical StrA (sortase A) from *Staphylococcus aureus*. In general, SrtA-type sortases are necessary for the anchoring of the majority or all the LPXTG-containing proteins of a given bacterium. Bacteria always encode only a single SrtA-type homologue. Class B sortases include the prototypical StrB from *S. aureus*. Sortases of the StrB subfamily are found in few Gram-positive bacteria and recognize a novel type of sorting signal, i.e. a motif NP  $\begin{bmatrix} Q \\ K \end{bmatrix}$  TN. The sortase C class or subfamily-3 sortases, is the largest group with sortases found in both Low- and High-G+C Gram-positive bacteria. These sortases are often present in several copies per genome but they anchor far fewer proteins than the SrtA-like proteins. As SrtA, sortases of class C recognize LPXTG motifs but followed by a glycine residue instead of an acidic residue. At the difference of class A sortases, subfamily-3 sortase genes are adjacent to the genes of their potential substrate, suggesting a coordinate regulation. Moreover, SrtAs seem to be Type II membrane proteins, whereas subfamily 3 enzymes would be of Type I. The sortase D class, which gathers the subfamily 4 and 5, is the last class of sortases currently defined. While subfamily 4 processes proteins bearing the motif LPXTA, subfamily 5 anchors proteins exhibiting the motif LAXTG.

It is generally admitted that lots of surface proteins anchored by sortases are virulence factors but it must be stressed that LPXTG-like proteins can have alternative function particularly in saprophytic bacteria. Functions of LPXTG-like proteins are very diverse, ranging from adhesins (involved in eukaryotic host cells invasion and/or bio-film formation), antigens, receptors or enzymes (amidases, hydrolases, lyases, proteases, transferases).

## Noncovalent binding via cell wall binding domains (CWBDs)

### Choline-binding domain (PF01473)

In *Streptococcus pneumoniae*, investigations on LytA (autolytic amidase) revealed that the C-terminally localized choline binding domain is responsible for the specific recognition of choline residues of teichoic and lipoteichoic acids (Fernandez-Tornero *et al.*, 2001). The choline-binding domain, also called cell wall binding domain of Type 1 (CWBD1), is constituted of several imperfect repeats of ~20 amino acids each and characterized by the presence of conserved aromatic residues (Fig. 2). Such repeats are classified under the accession number PF01473 (<http://www.sanger.ac.uk/Software/Pfam/>) (Bateman *et al.*, 2004).

The number of copies ranges from 1 to more than 50 in the case of a putative enterotoxin from *Clostridium acetobutylicum* (Desvaux *et al.*, 2005) but it is generally around 5 or 10 multiple tandem repeats. Besides autolysins, CWBD1 can be found in sucrases, lactamases, proteases, bacteriocins, toxins, antigens and even in proteins involved in formation of fibril-like structure in *Peptostreptococcus micros* (Kremer *et al.*, 1999) (Fig. 2). In Gram-positive bacteria, CWBD1 are essentially found within the *Clostridiales* and *Lactobacillales* orders.

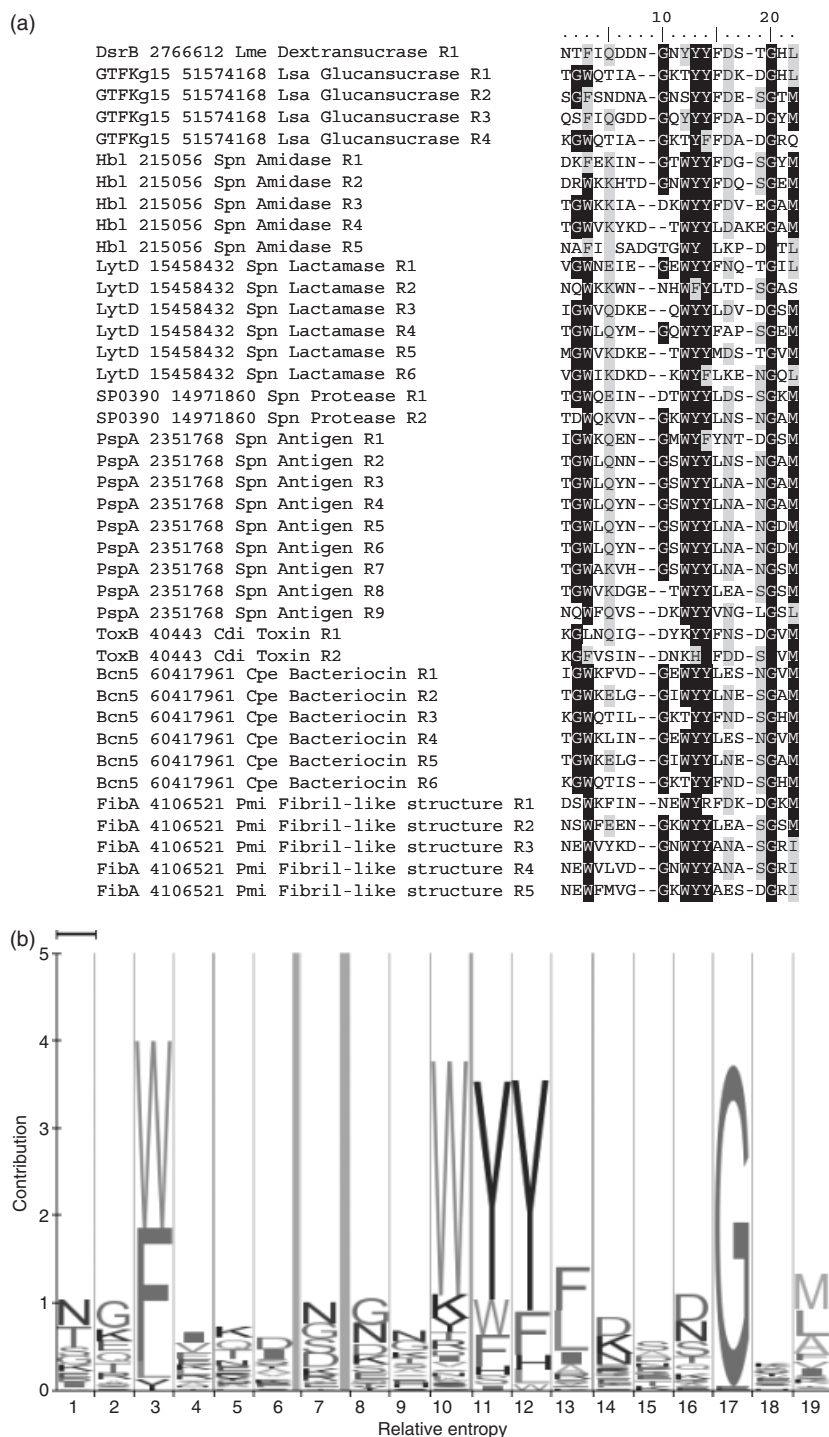
### Cell wall binding domain of Type 2 (PF04122)

From the studies of *B. subtilis* autolysin CwlB (Kuroda & Sekiguchi, 1991) and *Clostridium difficile* adhesin Cwp66 (Waligora *et al.*, 2001), it appeared that their N-terminal regions, which are composed of three repeated domains, were responsible for protein binding to cell wall. The cell wall binding domain of Type 2 (CWBD2; PF04122) is ~100 amino acids long (Fig. 3). CWBD2 are found in multiple tandem copies, between 1 and 3, and are generally localized in the N-terminal or C-terminal region. To date, the cell wall components involved in the interactions with CWBD2 have not been determined. In Gram-positive bacteria, such domain is found only among members of the *Bacillaceae* and *Clostridiaceae* families. Apart from autolysins and adhesins, CWBD2 are found in some S-layer proteins (Fig. 3) (see section about S-layer).

### Lysin motif domain (PF01476)

The lysin motif domain (LysM; PF01476), previously called peptidoglycan-binding domain of Type 2, got its name from its presence in various bacterial lysins (Joris *et al.*, 1992). LysM domains bind directly to peptidoglycan and binding is hindered by SCWPs, possibly lipoteichoic acids (Steen *et al.*, 2003). The LysM domain is ~40 residues long (Fig. 4) and multiple tandem copies, between 1 and 6, are generally localized in the N-terminal or C-terminal protein region. LysM is found in more than 1500 proteins, originating from both eukaryotic and prokaryotic cells, including *Archaea*, Gram-negative and Low- and High-G+C Gram-positive bacteria. Besides bacterial cell wall degradation, proteins bearing LysM domains can have other enzyme activities, such as peptidase, chitinase, esterase, reductase or nucleotidase, or act as antigens or binding proteins, e.g. to albumin, elastin or immunoglobulin (Fig. 4). Some autolysins can have a dual function, i.e. bacteriolytic activity and adhesive properties (Heilmann *et al.*, 2005). p60 (protein of 60 kDa) from *Listeria monocytogenes*, also called Iap (invasion-associated protein) or CwhA (cell wall hydrolase A), is even considered as a virulence factor (Dussurget *et al.*, 2004).



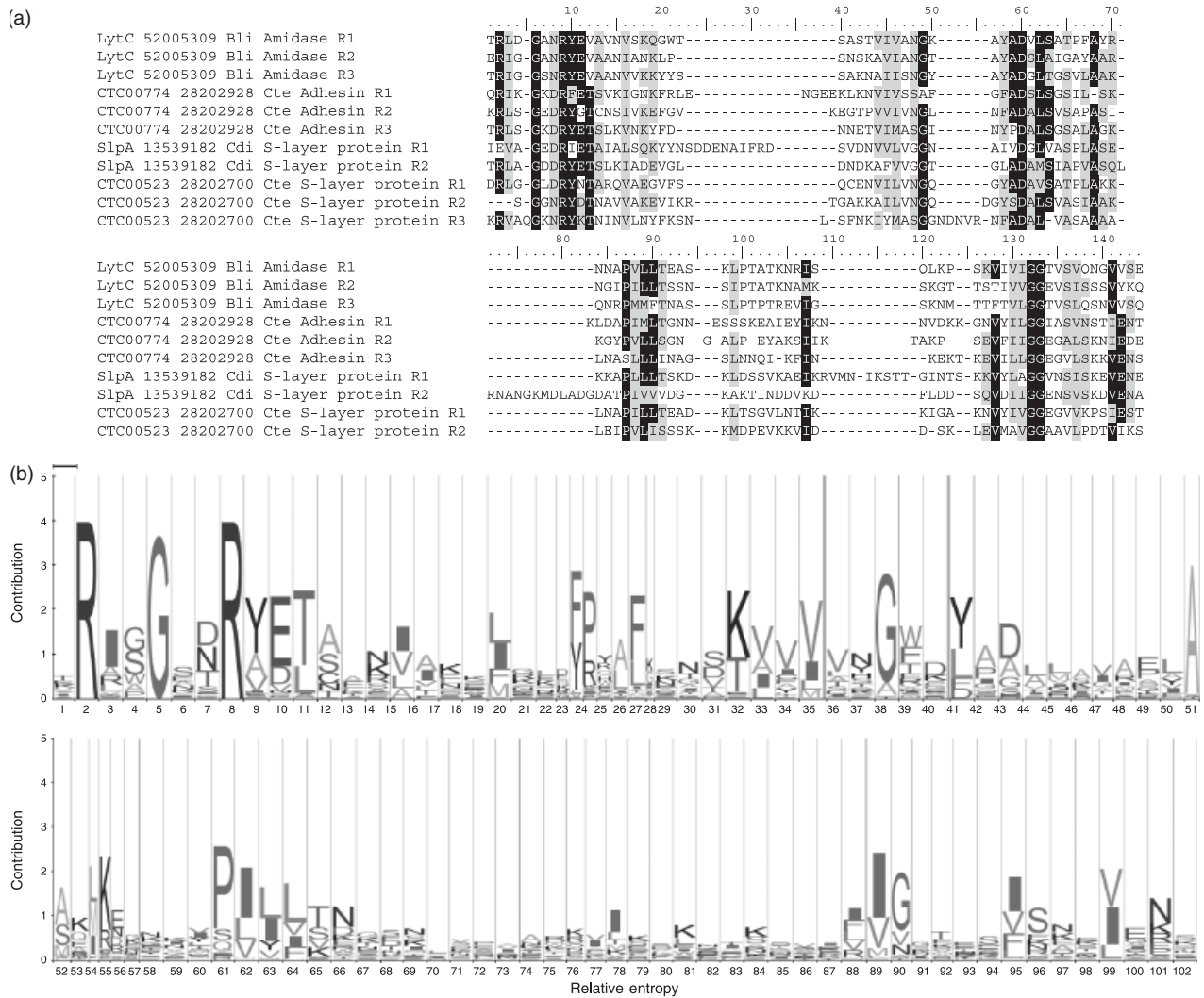


**Fig. 2.** The CWBD1 motif (PF01473). (a) Amino acid sequence alignment of CWBD1 repeats from some selected proteins found in Gram-positive bacteria. Alignment was performed using ClustalW (Higgins *et al.*, 1994) and visualized with BioEdit v7.0.4.1 (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) (Hall, 1999). Threshold was set up at 60% of the consensus; black and grey shading highlight identical and similar residues, respectively, at a given position. R, repeat; Cdi, *Clostridium difficile*; Cpe, *Clostridium perfringens*; Lme, *Leuconostoc mesenteroides*; Lsa, *Lactobacillus sakei*; Pmi, *Peptostreptococcus micros*; Spn, *Streptococcus pneumoniae*. (b) Alignment of CWBD1 motifs shown in the HMM-Logos format (Schuster-Bockler *et al.*, 2004). This Logo was generated from the 1082 proteins listed in Pfam as containing CWBD1 motifs (<http://www.sanger.ac.uk/Software/Pfam/>) (Bateman *et al.*, 2004). Briefly, in HMM-Logo graphical representation, the overall height of the stack indicates the sequence conservation at that position while the height of symbols within the stack indicates the relative frequency of each amino acid at that position; width of the column represents the expected contribution, i.e. the relative importance of this position in defining a protein family; pink column indicates position with no conserved residue (Schneider & Stephens, 1990; Schuster-Bockler *et al.*, 2004).

## GW modules

From the study of *L. monocytogenes* InlB (internalin B), a novel motif allowing noncovalent cell surface anchoring was uncovered (Jonquieres *et al.*, 1999). This motif of ~80 amino acids long and containing dipeptide Gly-Trp was

called GW module. Such a module is found in multicopy; three GW modules are found in the C-terminal region in InlB, whereas eight GW modules are found in Ami (amidase). The higher is the number of GW modules, the stronger is the attachment to bacterial cell wall. GW modules are necessary and sufficient for cell wall anchoring and



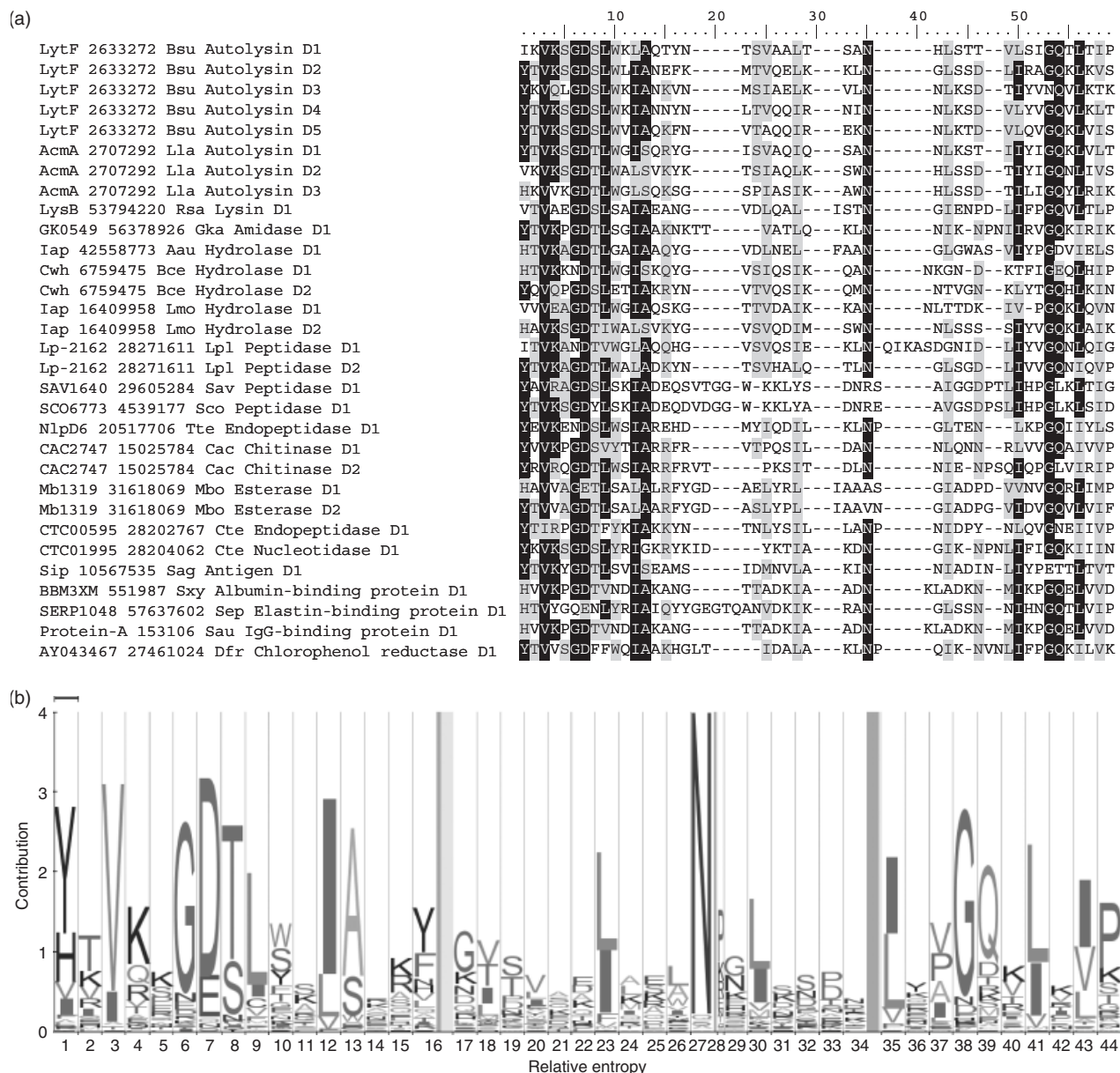
**Fig. 3.** The CWBD2 motif (PF04122). (a) Amino acid sequence alignment of CWBD2 repeats from some selected proteins found in Gram-positive bacteria. Threshold was set up at 60% of the consensus. R, repeat; Bli, *Bacillus licheniformis*; Cdi, *Clostridium difficile*; Cte, *Clostridium tetani*. (b) Alignment of CWBD2 motifs shown in the HMM-Logos format (Schuster-Bockler *et al.*, 2004). This Logo was generated from the 109 proteins listed in Pfam as containing CWBD2 motifs.

interact with lipoteichoic acids but proteins exhibiting only one GW module would not be surface attached. Besides the genus *Listeria*, GW proteins have been identified in *Staphylococcus* species (Cabanes *et al.*, 2002). GW proteins are mostly autolysins, which can also be implicated in adhesion to eukaryotic cells such as Aas (autolysin/adhesin of *Staphylococcus saprophyticus*) (Hell *et al.*, 1998).

### S-layer homology motif (PF00395)

S-layer homology domains (SLHDs) were originally discovered in (glyco)proteins forming the regular crystalline surface layer (S-layer) observed in some prokaryotic organisms, i.e. *Archaea* and *Bacteria* (Sára & Sleytr, 2000) (see section

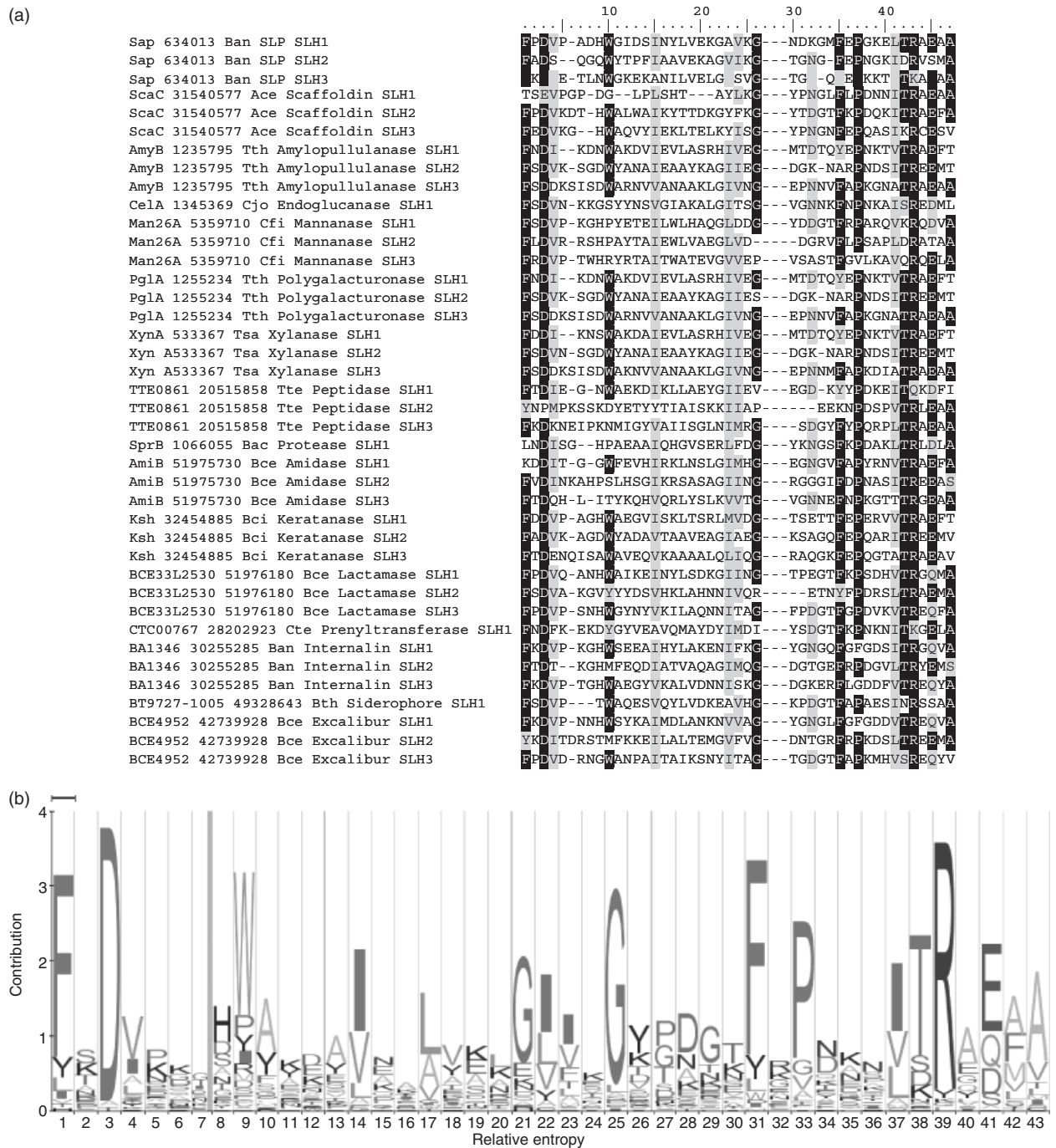
about S-layer). S-layer proteins bind very strongly to the cell wall but via noncovalent interactions. In Gram-positive bacteria, SLHDs can interact with pyruvylated SCWPs and/or peptidoglycan itself (Sára, 2001). Some SLHDs would also bind to SCWPs noncovalently linked to peptidoglycan network (Chauvaux *et al.*, 1999). Investigation on structure-function relationships revealed that SLH motifs are not sufficient for specific binding to SCWPs (Huber *et al.*, 2005). Thus, the molecular basis explaining SLHDs affinity and specificity of interaction with cell wall polymers are not completely elucidated. In S-layer proteins, SLHDs are (i) usually composed of 1-3 SLH modules/motifs (PF00395) of ~50 amino acids residues (Fig. 5) and (ii) systematically found in the N- or C-terminal region of the protein (Lupas



**Fig. 4.** The LysM motif (PF01476). (a) Amino acid sequence alignment of LysM domains from some selected proteins found in Gram-positive bacteria. Threshold was set up at 60% of the consensus. D, domain; Aau, *Arthrobacter aureus*; Bce, *Bacillus cereus*; Bsu, *Bacillus subtilis*; Cac, *Clostridium acetobutylicum*; Cte, *Clostridium tetani*; Dfr, *Desulfotobacterium frappierii*; Gka, *Geobacillus kaustophilus*; Lla, *Lactococcus lactis*; Lmo, *Listeria monocytogenes*; Lpl, *Lactobacillus plantarum*; Mbo, *Mycobacterium bovis*; Rsa, *Renibacterium salmoninarum*; Sau, *Staphylococcus aureus*; Sag, *Streptococcus agalactiae*; Sav, *Streptomyces avermitilis*; Sco, *Streptomyces coelicolor*; Sep, *Staphylococcus epidermidis*; Sxy, *Staphylococcus xylosus*; Tte, *Thermoanaerobacter tengcongensis*. (b) Alignment of LysM domains shown in the HMM-Logos format (Schuster-Bockler *et al.*, 2004). This Logo was generated from the 1668 proteins listed in Pfam as containing LysM domains.

*et al.*, 1994). Besides S-layer proteins, SLHDs are found in other cell surface displayed proteins, i.e. (i) enzymes, namely involved in polysaccharide degradation, e.g. cellulases, amylases, hemicellulases, galacturonidases, keratanases or pullanases and (ii) scaffoldins permitting the assembly of more complex macromolecular protein structure, called

cellulosomes (Engelhardt & Peters, 1998)(see section about Cellulosome). Bioinformatic analyses further suggest that some proteins possessing SLH motifs could be amidases, peptidases, proteases, siderophores, excaliburs (extracellular calcium-binding domain), or even internalins (Fig. 5).



**Fig. 5.** The SLH motif (PF00395). (a) Amino acid sequence alignment of SLH motifs from some selected S-layer and other cell-surface displayed proteins found in Gram-positive bacteria. Threshold was set up at 60% of the consensus. SLH, S-layer homology motif; SLP, S-layer protein; Ace, *Acetivibrio cellulolyticus*; Bac, *Bacillus* sp.; Ban, *Bacillus anthracis*; Bce, *Bacillus cereus*; Bci, *Bacillus circulans*; Cfi, *Cellulomonas fimi*; Cjo, *Clostridium josui*; Cte, *Clostridium tetani*; Tsa, *Thermoanaerobacterium saccharolyticum*; Tth, *Thermoanaerobacterium thermosulfurigenes*. (b) Alignment of SLH motifs shown in the HMM-Logos format (Schuster-Bockler et al., 2004). This Logo was generated from the 557 proteins listed in Pfam as containing SLH motifs.

### Other CWBDs

Another domain with a putative peptidoglycan-binding function, called peptidoglycan-binding domain of Type 1

(PBD1; PF01471), has been proposed in CwlA (cell wall hydrolase autolysin) (Foster, 1991). This domain is ~70 amino acids long and can be present in the N- or C-terminal regions. PBD1 is usually present in a single copy but up to

nine repeats can be found in some proteins. In Gram-positive bacteria, i.e. within *Clostridia*, *Bacilli* and *Actinobacteria* classes, proteins exhibiting such a domain are involved in bacterial cell wall degradation. Conclusive evidences of the involvement of PBD1 in protein attachment to bacterial cell wall components are still awaited.

From genomic analyses of *C. acetobutylicum*, a novel motif was hypothesized as involved in protein cell wall anchoring (Nölling *et al.*, 2001). The motif ChW (Clostridial hydrophobic domain with a conserved W residue), which is ~50 amino acids long and contains highly conserved Gly-Trp dipeptide motif as observed in GW modules, can be found repetitively several times along the protein sequence. Proteins bearing such repeats could be involved in adhesion or biopolymer degradation. Similar repeats have been only observed in *Streptomyces coelicolor*.

It is interesting to note that GW modules have been shown to be structurally and evolutionary related to SH3 but are unlikely to be functionally similar (Marino *et al.*, 2002). The GW modules are also related to prokaryotic SH3b domains. These domains of ~65 amino acids are homologous to eukaryotic SH3 domains and were first characterized in p60 proteins. The function of SH3b is unknown, but in lysostaphin from *Staphylococcus simulans* such a domain is present in the C-terminal region, which mediates protein binding to bacterial cell wall (Baba & Schneewind, 1996).

Some S-layer proteins do not exhibit SLHDs and thus possess alternative binding mechanisms to the cell wall, such as CWBD2 (see section about Cell wall binding domain of Type 2). In *Geobacillus stearothermophilus*, binding of SbsA (S-layer *Bacillus stearothermophilus* protein A) and SbsC occurs via a highly conserved N-terminal region, rich in R, Y and K, which recognizes a distinct SCWP (Schäffer *et al.*, 1999). In *Aneurinibacillus thermoaerophilus*, a S-layer glycoprotein is anchored very effectively to the cell wall, probably due to the doubling of a proposed N-terminal lectin-like binding motif (Steindl *et al.*, 2002). In *Corynebacterium glutamicum*, the S-layer protein PS2 (protein secreted 2) is most likely anchored via its C-terminal sequence to a hydrophobic layer within the cell wall (Chami *et al.*, 1997). In *Lactobacillus acidophilus*, SlpA (S-layer protein A) possess a highly conserved C-terminal sequence composed of putative carbohydrate-binding repeats most likely responsible for the anchoring to teichoic acids (Smit & Pouwels, 2002). Finally, in some members of the *Lactobacillales* order, while attachment of S-layer proteins seems to involve SCWPs, the location or nature of the cell wall binding domains involved are currently unknown (Avall-Jaaskelainen & Palva, 2005).

Some proteins normally considered as cytoplasmic are not translocated via a specific secretion pathway, such as Sec, but are most likely present in the extracellular milieu following autolysis of a subset of the bacterial population in a process called allolysis and would then be scavenged by

other cells (Guiral *et al.*, 2005); e.g. the enolase from *Streptococcus pneumoniae* can reassociate on bacterial cell surface by interacting probably with proteinaceous but uncharacterized receptor(s) (Bergmann *et al.*, 2001). Other streptococcal proteins such as glyceraldehydes 3-phosphate dehydrogenase, PavA (pneumococcal adhesion and virulence protein A) or FBP54 (fibronectin/fibrinogen-binding protein of 54 kDa) from *S. pneumoniae* and *S. pyogenes* respectively, have been proposed to be associated to the bacterial cell surface in a similar way (Chhatwal, 2002).

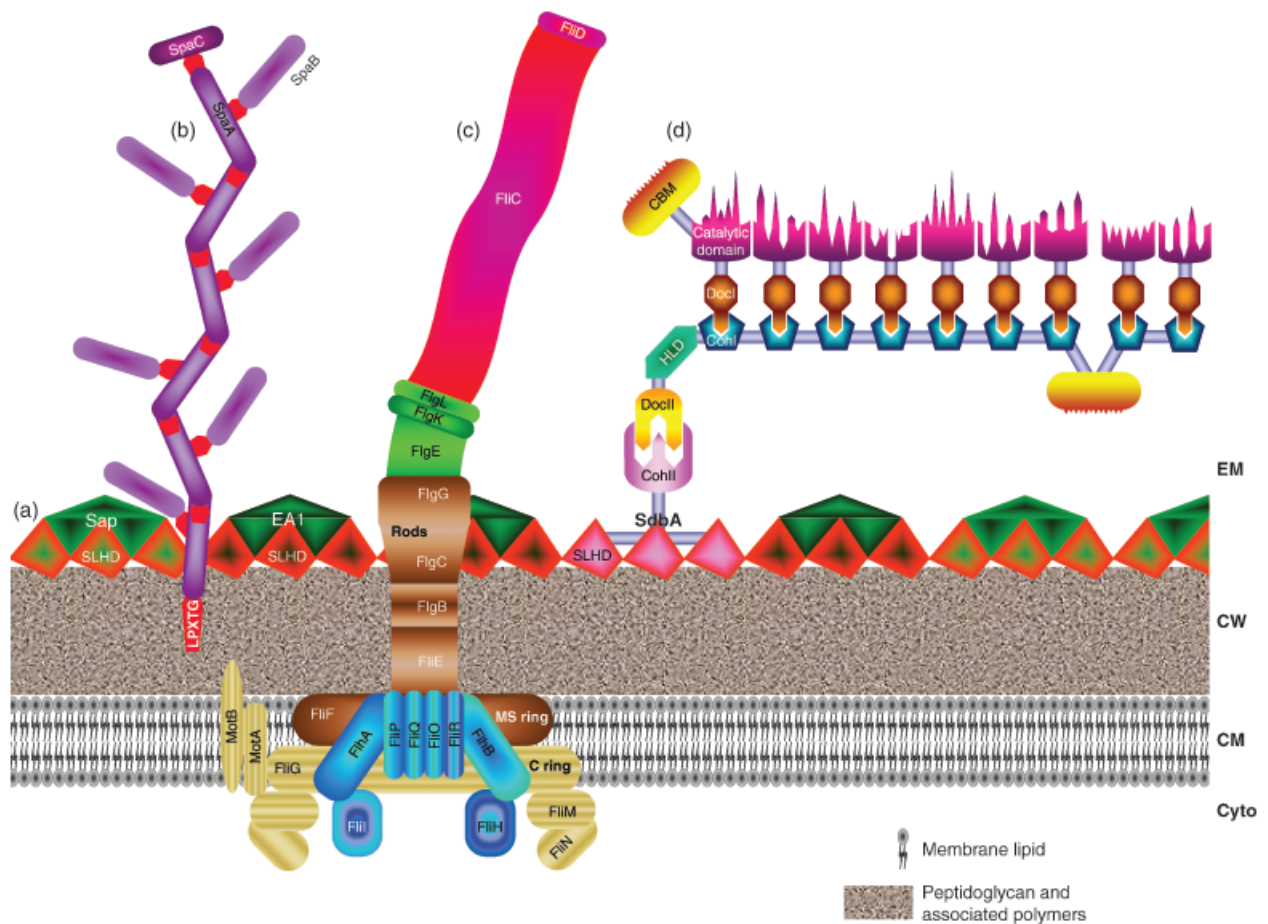
## Macromolecular protein surface structures

### S-layer

While S-layers are considered as widely distributed in almost all phylogenetic branches of *Archaea* and *Bacteria* (Engelhardt & Peters, 1998; Sára & Sleytr, 2000), they are not present in all Gram-positive bacteria, e.g. an S-layer is absent from members of *Listeria* and *Staphylococcus* genus. The S-layer, which entirely coats the bacterial cell surface, is formed of (glyco)proteins which bind by noncovalent interactions to cell wall components (see section about S-layer homology motif and Others CWBDs). S-layer proteins are arrayed in a two-dimensional lattice, which forms the outermost cell wall layer (Fig. 6a). As a result of monomolecular assembly of identical units, S-layers exhibit identical and regularly displayed pores. Most S-layers consist of single protein species. In *Clostridium difficile*, however, the S-layer consists of two different proteins derived from the same precursor (Calabi *et al.*, 2001). In *Bacillus anthracis*, two S-layer proteins are sequentially synthesized and lead to the formation of a patchwork pattern (Couture-Tosi *et al.*, 2002). In Gram-positive bacteria, S-layer proteins are synthesized as preproteins, i.e. with a cleavable N-terminal signal peptide, and thus translocated through the cytoplasmic membrane. Then, S-layer proteins self-assemble on the bacterial cell surface in a crystalline lattice, where proper orientation and incorporation depend on surface net charge and hydrophobicity. S-layers have been proposed to act as (i) precise molecular sieves with sharp cutoff levels, because of the presence of pores identical in size and morphology (ii) a protective shield, from attack by bacterial parasite e.g. or (iii) virulence factors by protecting pathogenic bacteria from immune system responses of the infected host.

### Pilus

In contrast to Gram-negative bacteria (Koebnik, 2001), only one pathway is currently recognized for pilus biogenesis in Gram-positive bacteria (Ton-That *et al.*, 2004; Lauer *et al.*, 2005). This pathway involves ordered cross-linking polymerization by transpeptidase sortase, where pilin subunits



**Fig. 6.** Schematic representation of the cell surface macromolecular structures currently recognized in Gram-positive bacteria. (a) Schematic representation of the S-layer is based on data from *Bacillus anthracis* (Couture-Tosi et al., 2002). After cytoplasmic membrane translocation of the S-layer proteins EA1 and Sap through Sec, the proteins, which are sequentially expressed, self-assemble and form a patchwork pattern. Both S-layer proteins possess three typical SLH motifs and bind to pyruvylated SCWPs. S-layer proteins are depicted in green and their SLH motifs are highlighted in red. (b) Pilus representation is based on the model proposed in *Corynebacterium diphtheriae* (Ton-That & Schneewind, 2004). After translocation via the Sec apparatus and cleavage of the signal peptides, sortase catalyses polymerization of pilins SpaA, SpaB and SpaC, which possess C-terminal LPXTG motif, using the side chain amino groups of pilin motif sequences to generate links between pilin subunits. SpaA, the major components of the pilus, is anchored into the cell wall and forms the pilus backbone. SpaB is spaced at regular intervals along the shaft and SpaC is located at the pilus tip. Pilins are violet and their LPXTG domain is red. (c) Flagellum representation is based on *Bacillus subtilis* model from KEGG (Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.jp/kegg/>). FEA permits translocation and assembly of flagellar components of the basal body, hook/junctions and filament. Components of flagellum motor/switch are depicted in yellow, FEA in blue, basal body in brown, hook/junctions in green and filament in red. (d) Cellulosome representation is based on the model proposed in *Clostridium thermocellum* (Bayer et al., 1998). After translocation of the different cellulosomal components presumably via Sec, cellulosome is assembled on the bacterial cell surface where it is attached via recognition of cohesin/dockerin domain of Type II to S-layer anchored proteins, such as SdbA. SLH motifs and CohI present in SdbA are depicted in red and pink, respectively. Cellulosome components DocI is coloured in yellow, HLD in green, DocI in blue, CohI in brown, CBM in orange and catalytic domains in violet. Cyto, cytoplasm; CW, cell wall; EM, extracellular milieu.

are covalently linked to one another (Fig. 6b) (see section about Covalent binding via sortase). Sortases involved in pili biogenesis belongs to class C but it is not excluded that other sortases could participate in the process. In *Corynebacterium diphtheriae*, such pili appear as thin fibre-like structure of 0.2–0.3  $\mu\text{m}$  in length and 2–6 nm in diameter. Such a pathway would be present in both Low- and High-G+C Gram-positive bacteria, namely in the genus *Actinomyces*, *Clostri-*

*dium*, *Corynebacterium*, *Enterococcus* and *Streptococcus*. These pili are involved in bacterial adhesion, colonization and biofilm formation as well as in bacterial pathogenicity.

In *Ruminococcus albus*, CbpC (cellulose-binding protein C) and GP25 (glycoprotein of 25 kDa) have been identified as exhibiting similarity with Type 4 pilin (Rakotoarivonina et al., 2002, 2005). GP25 was characterized as the major protein subunit of Type 4 pili (Tfp) present at the surface of

*R. albus*. Such pili would be involved in cellulose colonization, i.e. biofilm formation, and cellulose degradation. *Ruminococcus albus* is the only Gram-positive bacterium where Tfp have been reported. However, the secretion apparatus permitting their assembly remains unknown. From phylogenetic analyses, it is well known that Type II secretion system, the Tfp assembly apparatus, and Type IV secretion system of Gram-negative bacteria resemble to proteins found in the fimbriin-protein exporter (FPE) of Gram-positive bacteria and thus have been collectively called pilus/secretion/twitching motility/competence (PSTC) (Peabody *et al.*, 2003). The FPE, which has been essentially investigated in *B. subtilis*, permits the export and assembly of Type 4 prepilins participating to bacterial competence (Dubnau, 1999). Importantly, from investigations in *B. subtilis*, these Type 4 pilins do not form a proper Tfp but a transcell wall structure permitting presentation of exogenous DNA to translocation machinery.

Fibril-like surface structure found in *P. micros* are formed at least from a CWBD1 protein called FibA (fibril-like structure subunit A) (Kremer *et al.*, 1999). Molecular mechanisms permitting assembly of such structure, which is described as auto-aggregative and often exceeding 4 µm in length, are currently unknown. In *Arthrobacter photogonimos*, LipA (light-inducible pilin A) was proposed as the major pilin protein of surface pili (Yang & Hooper, 1998). While definitive evidence are missing to support LipA as a constituent of the observed pili on the cell surface of *A. photogonimos*, this system could constitute an alternative pilus formation pathway in Gram-positive bacteria. Finally, from genomic analyses it appeared that the recently unravelled Tad system involved in the formation of Flp (fimbrial low-molecular-weight protein) pili in Gram-negative bacteria, is also encoded on the genome of some Gram-positive bacteria, including *Actinobacteria* and *Clostridia* (Planet *et al.*, 2003; Desvaux *et al.*, 2005).

## Flagellum

Our understanding of the molecular mechanisms responsible for flagella assembly relies essentially on investigations in Gram-negative bacteria. The bacterial flagellum can be subdivided into 5 major parts: (i) the flagella motor/switch, which operates and controls a rotary mechanism; (ii) the basal body, a passive structure which receives torque from the motor and transmits it to the hook and then to the filament; (iii) the hook and junction proteins, which function as a universal joint; (iv) the flagellar filament and its distal cap, which function like a propeller and (v) the flagella export apparatus (FEA) (Macnab, 2004). Based on *B. subtilis* flagella model from Kyoto encyclopaedia of genes and genomes (<http://www.genome.jp/kegg/>) (Kanehisa *et al.*, 2004) (Fig. 6c), FEA is composed of the transmembrane

components FlhA, FlhB, FliO, FliP, FliQ and FliR, the chaperones FliJ, FliS and FliT, the ATPase FliI, and its regulator FliH. FliI would permit the coupling of ATP hydrolysis to protein export and seems negatively regulated by FliH (Bardy *et al.*, 2003). FEA permits the translocation and assembly of components of basal body, hook/junctions and filament. Proteins secreted through the FEA possess no cleavable signal peptide, their secretion is strictly Sec-independent (Aldridge & Hughes, 2001). It is worth mentioning that the FEA is related to the Type III secretion system of Gram-negative bacteria (Nguyen *et al.*, 2000). Besides a role in bacterial motility, flagella have been involved in bacterial adhesion, biofilm formation and bacterial pathogenicity (Jonson *et al.*, 2005). Flagella occurrence varies within a same genus or species.

## Cellulosome

In bacteria, cellulosomes are only found from within the phylum *Firmicutes*, families *Lachnospiraceae* and *Clostridiaceae* (Desvaux, 2005a). A cellulosome is an extracellular multienzymatic complex dedicated to degradation of plant cell wall polysaccharides, i.e. in bacterial colonization, biofilm formation and polysaccharide digestion (Bayer *et al.*, 2004). It is organized around a modular integrating protein, called scaffoldin, which coordinates the binding of different catalytic components. In the noncatalytic scaffolding protein CipA (cellulosome-integrating protein A) from *Clostridium thermocellum*, 9 cohesin domains of Type I (CohI) act as receptors for dockerin domains of Type I (DocI) borne by cellulosomal enzymes (Bayer *et al.*, 1998) (Fig. 6d). In *C. thermocellum*, cellulosomal enzymes can be cellulases, hemicellulases or chitinases. CipA also exhibits a dockerin domain of Type II (DocII) interacting with Type II cohesin domain (CohII) presents in cell surface proteins, namely SdbA (scaffoldin dockerin binding A), ORF2p (ORF 2 polypeptide) or OlpB (outer layer protein B) (Fig. 6d). Such proteins are anchored in the cell wall via SLHDs and thus permit cell surface display of the cellulosome. Contrary to SdbA, which bears only one CohII, OlpB and ORF2p harbour 2 and 4 CohII respectively, and as such can be considered as scaffoldins attaching several CipA proteins. A third S-layer anchored protein, OlpA, possesses a single CohI and was proposed to bind DocI-containing enzymes and/or to permit the temporary docking of cellulosomal enzyme before their final incorporation into the cellulosome. *C. thermocellum* cellulosome has a size ranging from 2.0 to 6.5 MDa depending on the bacterial strain and forms polycellulosomal protuberance of up to 100 MDa. From the crystallization of cohesin/dockerin complex in *C. thermocellum* (Carvalho *et al.*, 2003), it was proposed that a single DocI could simultaneously bind two CohI. As a consequence, such an interaction could permit the link

of two CipA proteins and thus explain the formation of polycellulosomes.

In *Bacteroides cellulosolvens* and *Acetivibrio cellulolyticus*, a higher level of complexity has been reached since cohesin/dockerin recognition involves two and three different scaffoldins respectively (Xu *et al.*, 2004). In *C. cellulovorans*, it has been demonstrated that the cellulosomal cellulase EngE (endoglucanase E) is cell wall anchored *via* SLHDs and its C-terminal DocI interacts with one of the CohI of the scaffoldin CbpA (cellulose binding protein A) (Kosugi *et al.*, 2002). It has also been speculated that HLDs present in CbpA could be involved in cellulosome adsorption on the bacterial cell surface since they would share homology with SLHDs (Doi & Tamaru, 2001). In *C. cellulolyticum* (Desvaux, 2005a, b), the way the cellulosome is attached to the bacterial cell surface remains unknown.

## Conclusion

Study of cell surface display mechanisms is essential for understanding bacterial virulence. Cell surface displayed proteins are frequently considered as potential virulence factors. However, a facet that should not be ignored by those studying bacterial pathogenesis is that many nonpathogenic bacteria also display cell surface proteins relevant to their life-style, e.g. plant-degradative enzymes in cellulolytic bacteria. Cell envelope-associated proteins can also mediate interactions and thus anchoring of capsule polysaccharides (Taylor & Roberts, 2005); such capsule can further mask cell surface-exposed proteins. Importantly, these mechanisms involved in protein anchoring within the cell envelope do not prevent such proteins to be present in the extracellular medium, e.g. because of bacterial cell lysis, natural loss of cell wall components and their associated proteins during bacterial growth, or proteolytic shaving of certain lipoproteins from the membrane (Venema *et al.*, 2003; Tjalsma *et al.*, 2004; Tjalsma & van Dijl, 2005). Another important process related to cell surface displayed proteins is biofilm formation, since besides individual adhesins, macromolecular structures such as flagella, pili and cellulosomes can also participate. It is worth stressing that in the attempt to identify the entire set of proteins sharing similar molecular mechanisms permitting protein attachment within the cell envelope of Gram-positive bacteria, proteomics appears as a very powerful approach (Calvo *et al.*, 2005; Pucciarelli *et al.*, 2005; Tjalsma & van Dijl, 2005).

From a biotechnological point of view and as proposed for autodisplay systems in Gram-negative bacteria (Lee *et al.*, 2003; Desvaux *et al.*, 2004a, b; Henderson *et al.*, 2004), systems involved in protein cell surface display in Gram-positive bacteria can potentially be used for (i) exposure of antigenic determinants for vaccine development, (ii) the display of receptor or ligand for purification

or binding assays, (iii) expression of peptide libraries for epitope mapping or antibody specificity test, (iv) functional domain analyses of heterologous proteins and (v) bioconversion by expressing enzymatic activity on the bacterial surface. Moreover, entering in the era of nanotechnology, nanomachines/nanostructures such as flagella, S-layers, pili and cellulosomes will undoubtedly be the focus point of in-depth investigations in order to develop future environmental, food or biomedical applications.

## Acknowledgements

This work was supported by the Institut National de la Recherche Agronomique (INRA). Emilie DUMAS is a PhD Research Fellow granted by the French Minister of National Education and Research (MENESR).

## References

- Aldridge P & Hughes KT (2001) How and when are substrates selected for Type III secretion? *Trends Microbiol* **9**: 209–214.
- Avall-Jaaskelainen S & Palva A (2005) *Lactobacillus* surface layers and their applications. *FEMS Microbiol Rev* **29**: 511–529.
- Baba T & Schneewind O (1996) Target cell specificity of a bacteriocin molecule: a C-terminal signal directs lysostaphin to the cell wall of *Staphylococcus aureus*. *EMBO J* **15**: 4789–4797.
- Bardy SL, Sandy YM & Jarrell KF (2003) Prokaryotic motility structures. *Microbiology* **149**: 295–304.
- Bateman A, Coin L, Durbin R, *et al.* (2004) The Pfam protein families database. *Nucleic Acids Res* **32**: D138–141.
- Bayer EA, Belaich JP, Shoham Y & Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol* **58**: 521–554.
- Bayer EA, Shimon LJ, Shoham Y & Lamed R (1998) Cellulosomes – structure and ultrastructure. *J Struct Biol* **124**: 221–234.
- Bergmann S, Rohde M, Chhatwal GS & Hammerschmidt S (2001)  $\alpha$ -Enolase of *Streptococcus pneumoniae* is a plasmin(ogen)-binding protein displayed on the bacterial cell surface. *Mol Microbiol* **40**: 1273–1287.
- Botero LM, Brown KB, Brumfield S, Burr M, Castenholz RW, Young M & McDermott TR (2004) *Thermobaculum terrenum* gen. nov., sp. nov.: a non-phototrophic Gram-positive thermophile representing an environmental clone group related to the *Chloroflexi* (green non-sulfur bacteria) and *Thermomicrobia*. *Arch Microbiol* **181**: 269–277.
- Bunai K, Nozaki M, Kakeshita H, Nemoto T & Yamane K (2005) Quantitation of de novo localized <sup>15</sup>N-labeled lipoproteins and membrane proteins having one and two transmembrane segments in a *Bacillus subtilis* *secA* temperature-sensitive mutant using 2D-PAGE and MALDI-TOF MS. *J Proteome Res* **4**: 826–836.



- Cabanes D, Dehoux P, Dussurget O, Frangeul L & Cossart P (2002) Surface proteins and the pathogenic potential of *Listeria monocytogenes*. *Trends Microbiol* **10**: 238–245.
- Calabi E, Ward S, Wren B, Paxton T, Panico M, Morris H, Dell A, Dougan G & Fairweather N (2001) Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol Microbiol* **40**: 1187–1199.
- Calvo E, Pucciarelli MG, Bierne H, Cossart P, Albar JP & Garcia-Del Portillo F (2005) Analysis of the *Listeria* cell wall proteome by two-dimensional nanoliquid chromatography coupled to mass spectrometry. *Proteomics* **5**: 433–443.
- Carvalho AL, Dias FMV, Prates JAM, Nagy T, Gilbert HJ, Davies GJ, Ferreira LMA, Romao MJ & Fontes CMGA (2003) Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. *Proc Natl Acad Sci USA* **100**: 13809–13814.
- Chami M, Bayan N, Peyret JL, Gulik-Krzywicki T, Leblon G & Shechter E (1997) The S-layer protein of *Corynebacterium glutamicum* is anchored to the cell wall by its C-terminal hydrophobic domain. *Mol Microbiol* **23**: 483–492.
- Chauvaux S, Matuschek M & Beguin P (1999) Distinct affinity of binding sites for S-layer homologous domains in *Clostridium thermocellum* and *Bacillus anthracis* cell envelopes. *J Bacteriol* **181**: 2455–2458.
- Chhatwal GS (2002) Anchorless adhesins and invasins of Gram-positive bacteria: a new class of virulence factors. *Trends Microbiol* **10**: 205–208.
- Comfort D & Clubb RT (2004) A comparative genome analysis identifies distinct sorting pathways in Gram-positive bacteria. *Infect Immun* **72**: 2710–2722.
- Couture-Tosi E, Delacroix H, Mignot T, Mesnage S, Chami M, Fouet A & Mosser G (2002) Structural analysis and evidence for dynamic emergence of *Bacillus anthracis* S-layer networks. *J Bacteriol* **184**: 6448–6456.
- Cullen PA, Xu X, Matsunaga J, Sanchez Y, Ko AI, Haake DA & Adler B (2005) Surfaceome of *Leptospira* spp. *Infect Immun* **73**: 4853–4863.
- Dalbey RE & Chen M (2004) Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta—Mol Cell Res* **1694**: 37–53.
- Desvaux M (2005a) The cellulosome of *Clostridium cellulolyticum*. *Enzymes Microbiol Technol* **37**: 373–385.
- Desvaux M (2005b) *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia. *FEMS Microbiol Rev* **29**: 741–764.
- Desvaux M, Khan A, Scott-Tucker A, Chaudhuri RR, Pallen MJ & Henderson IR (2005) Genomic analysis of the protein secretion systems in *Clostridium acetobutylicum* ATCC824. *Biochim Biophys Acta—Mol Cell Res* **1745**: 223–253.
- Desvaux M, Parham NJ & Henderson IR (2004a) The autotransporter secretion system. *Res Microbiol* **155**: 53–60.
- Desvaux M, Parham NJ & Henderson IR (2004b) Type V protein secretion: simplicity gone awry? *Curr Issues Mol Biol* **6**: 111–124.
- Doi RH & Tamaru Y (2001) The *Clostridium cellulovorans* cellulosome: an enzyme complex with plant cell wall degrading activity. *Chem Rec* **1**: 24–32.
- Drams S, Trieu-Cuot P & Bierne H (2005) Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res Microbiol* **156**: 289–297.
- Dubnau D (1999) DNA uptake in bacteria. *Annu Rev Microbiol* **53**: 217–244.
- Dussurget O, Pizarro-Cerda J & Cossart P (2004) Molecular determinants of *Listeria monocytogenes* virulence. *Annu Rev Microbiol* **58**: 587–610.
- Engelhardt H & Peters J (1998) Structural research on surface layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *J Struct Biol* **124**: 276–302.
- Fernandez-Tornero C, Lopez R, Garcia E, Gimenez-Gallego G & Romero A (2001) A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. *Nat Struct Biol* **8**: 1020–1024.
- Fischer W (1994) Lipoteichoic acid and lipids in the membrane of *Staphylococcus aureus*. *Med Microbiol Immunol* **183**: 61–76.
- Foster SJ (1991) Cloning, expression, sequence analysis and biochemical characterization of an autolytic amidase of *Bacillus subtilis* 168 trpC2. *J Gen Microbiol* **137**: 1987–1998.
- Froderberg L, Houben E, Samuelson JC, Chen M, Park SK, Phillips GJ, Dalbey R, Luirink J & De Gier JW (2003) Versatility of inner membrane protein biogenesis in *Escherichia coli*. *Mol Microbiol* **47**: 1015–1027.
- Garrity GM (2001) *Bergey's Manual of Systematic Bacteriology*, 2nd edn. Springer, Berlin Heidelberg.
- Gram H CJ (1884) Über die isolirte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschritte der Medizin* **2**: 185–189.
- Guiral S, Mitchell TJ, Martin B & Claverys JP (2005) Competence-programmed predation of noncompetent cells in the human pathogen *Streptococcus pneumoniae*: genetic requirements. *Proc Natl Acad Sci USA* **102**: 8710–8715.
- Gupta RS (1998) Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol Mol Biol Rev* **62**: 1435–1491.
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.
- Heilmann C, Hartleib J, Hussain MS & Peters G (2005) The multifunctional *Staphylococcus aureus* autolysin aaa mediates adherence to immobilized fibrinogen and fibronectin. *Infect Immun* **73**: 4793–4802.
- Hell W, Meyer HG & Gatermann SG (1998) Cloning of aas, a gene encoding a *Staphylococcus saprophyticus* surface protein with adhesive and autolytic properties. *Mol Microbiol* **29**: 871–881.
- Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC & Ala'Aldeen D (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **68**: 692–744.

- Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Huber C, Ilk N, Runzler D, Egelseer EM, Weigert S, Sleytr UB & Sara M (2005) The three S-layer-like homology motifs of the S-layer protein SbpA of *Bacillus sphaericus* CCM 2177 are not sufficient for binding to the pyruvylated secondary cell wall polymer. *Mol Microbiol* **55**: 197–205.
- Jonquieres R, Bierne H, Fiedler F, Gounon P & Cossart P (1999) Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of Gram-positive bacteria. *Mol Microbiol* **34**: 902–914.
- Jonson AB, Normark S & Rhen M (2005) Fimbriae, pili, flagella and bacterial virulence. *Contrib Microbiol* **12**: 67–89.
- Joris B, Englebort S, Chu CP, Kariyama R, Daneo-Moore L, Shockman GD & Ghuysen JM (1992) Modular design of the *Enterococcus hirae* muramidase-2 and *Streptococcus faecalis* autolysin. *FEMS Microbiol Lett* **91**: 257–264.
- Kanehisa M, Goto S, Kawashima S, Okuno Y & Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res* **32**: D277–280.
- Koebnik R (2001) The role of bacterial pili in protein and DNA translocation. *Trends Microbiol* **9**: 586–590.
- Kosugi A, Murashima K, Tamaru Y & Doi RH (2002) Cell-surface-anchoring role of N-terminal surface layer homology domains of *Clostridium cellulovorans* EngE. *J Bacteriol* **184**: 884–888.
- Kremer BH, Bijlsma JJ, Kusters JG, de Graaff J & van Steenberg TJ (1999) Cloning of *fibA*, encoding an immunogenic subunit of the fibril-like surface structure of *Peptostreptococcus micros*. *J Bacteriol* **181**: 2485–2491.
- Kuroda A & Sekiguchi J (1991) Molecular cloning and sequencing of a major *Bacillus subtilis* autolysin gene. *J Bacteriol* **173**: 7304–7312.
- Lauer P, Rinaudo CD, Soriani M, et al. (2005) Genome analysis reveals pili in Group B *Streptococcus*. *Science* **309**: 105.
- Lee SY, Choi JH & Xu Z (2003) Microbial cell-surface display. *Trends Biotechnol* **21**: 45–52.
- Luirink J, Von Heijne G, Houben E & De Gier JW (2005) Biogenesis of inner membrane proteins in *Escherichia coli*. *Annu Rev Microbiol* **59**: 329–355.
- Lupas A, Engelhardt H, Peters J, Santarius U, Volker S & Baumeister W (1994) Domain structure of the *Acetogenium kivui* surface layer revealed by electron crystallography and sequence analysis. *J Bacteriol* **176**: 1224–1233.
- Macnab RM (2004) Type III flagellar protein export and flagellar assembly. *Biochim Biophys Acta—Mol Cell Res* **1694**: 207–217.
- Marino M, Banerjee M, Jonquieres R, Cossart P & Ghosh P (2002) GW domains of the *Listeria monocytogenes* invasion protein InlB are SH3-like and mediate binding to host ligands. *EMBO J* **21**: 5623–5634.
- Mira A, Pushker R, Legault BA, Moreira D & Rodriguez-Valera F (2004) Evolutionary relationships of *Fusobacterium nucleatum* based on phylogenetic analysis and comparative genomics. *BMC Evol Biol* **4**: 50.
- Navarre WW & Schneewind O (1999) Surface proteins of Gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* **63**: 174–229.
- Nguyen L, Paulsen IT, Tchieu J, Hueck CJ & Saier MH Jr (2000) Phylogenetic analyses of the constituents of Type III protein secretion systems. *J Mol Microbiol Biotechnol* **2**: 125–144.
- Nölling J, Breton G, Omelchenko MV, et al. (2001) Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol* **183**: 4823–4838.
- Pallen MJ, Lam AC, Antonio M & Dunbar K (2001) An embarrassment of sortase – a richness of substrates? *Trends Microbiol* **9**: 97–101.
- Peabody CR, Chung YJ, Yen MR, Vidal-Ingigliardi D, Pugsley AP & Saier MH Jr (2003) Type II protein secretion and its relationship to bacterial Type 4 pili and archaeal flagella. *Microbiology* **149**: 3051–3072.
- Planet PJ, Kachlany SC, Fine DH, DeSalle R & Figurski DH (2003) The widespread colonization island of *Actinobacillus actinomycetemcomitans*. *Nat Genet* **34**: 193–198.
- Pucciarelli MG, Calvo E, Sabet C, Bierne H, Cossart P & Garcia-Del Portillo F (2005) Identification of substrates of the *Listeria monocytogenes* sortases A and B by a non-gel proteomic analysis. *Proteomics* **5**: 4808–4817.
- Rakotoarivonina H, Jubelin G, Hebraud M, Gaillard-Martinie B, Forano E & Mosoni P (2002) Adhesion to cellulose of the Gram-positive bacterium *Ruminococcus albus* involves Type 4 pili. *Microbiology* **148**: 1871–1880.
- Rakotoarivonina H, Larson MA, Morrison M, Girardeau JP, Gaillard-Martinie B, Forano E & Mosoni P (2005) The *Ruminococcus albus pilA1–pilA2* locus: expression and putative role of two adjacent *pil* genes in pilus formation and bacterial adhesion to cellulose. *Microbiology* **151**: 1291–1299.
- Sankaran K & Wu HC (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglycerol moiety from phosphatidylglycerol. *J Biol Chem* **269**: 19701–19706.
- Schneider TD & Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**: 6097–6100.
- Schuster-Bockler B, Schultz J & Rahmann S (2004) HMM Logos for visualization of protein families. *BMC Bioinform* **5**: 7.
- Schäffer C, Kahlig H, Christian R, Schulz G, Zayni S & Messner P (1999) The diacetamidodideoxyuronic-acid-containing glycan chain of *Bacillus stearothermophilus* NRS 2004/3a represents the secondary cell-wall polymer of wild-type *B. stearothermophilus* strains. *Microbiology* **145**: 1575–1583.
- Shatalkin AI (2004) Highest level of division in classification of organisms. 3. Monodermata and Didermata. *Zh Obshch Biol* **65**: 195–210.

- Shockman GD & Barrett JF (1983) Structure, function, and assembly of cell walls of Gram-positive bacteria. *Annu Rev Microbiol* **37**: 501–527.
- Smit E & Pouwels PH (2002) One repeat of the cell wall binding domain is sufficient for anchoring the *Lactobacillus acidophilus* surface layer protein. *J Bacteriol* **184**: 4617–4619.
- Snel B, Huynen MA & Dutilh BE (2005) Genome trees and the nature of genome evolution. *Annu Rev Microbiol* **59**: 191–209.
- Steen A, Buist G, Leenhouts KJ, El Khattabi M, Grijpstra F, Zomer AL, Venema G, Kuipers OP & Kok J (2003) Cell wall attachment of a widely distributed peptidoglycan binding domain is hindered by cell wall constituents. *J Biol Chem* **278**: 23874–23881.
- Steindl C, Schaffer C, Wugeditsch T, Graninger M, Matecko I, Muller N & Messner P (2002) The first biantennary bacterial secondary cell wall polymer and its influence on S-layer glycoprotein assembly. *Biochem J* **368**: 483–494.
- Sutcliffe IC & Russell RR (1995) Lipoproteins of Gram-positive bacteria. *J Bacteriol* **177**: 1123–1128.
- Sára M (2001) Conserved anchoring mechanisms between crystalline cell surface S-layer proteins and secondary cell wall polymers in Gram-positive bacteria? *Trends Microbiol* **9**: 47–49.
- Sára M & Sleytr UB (2000) S-Layer proteins. *J Bacteriol* **182**: 859–868.
- Taylor CM & Roberts IS (2005) Capsular polysaccharides and their role in virulence. *Contrib Microbiol* **12**: 55–66.
- Tjalsma H, Antelmann H, Jongbloed JD, *et al.* (2004) Proteomics of protein secretion by *Bacillus subtilis*: separating the “secrets” of the secretome. *Microbiol Mol Biol Rev* **68**: 207–233.
- Tjalsma H & van Dijl JM (2005) Proteomics-based consensus prediction of protein retention in a bacterial membrane. *Proteomics* **5**: 4472–4482.
- Tjalsma H, Kontinen VP, Pragai Z, Wu H, Meima R, Venema G, Bron S, Sarvas M & van Dijl JM (1999) The role of lipoprotein processing by signal peptidase II in the Gram-positive eubacterium *Bacillus subtilis*. Signal peptidase II is required for the efficient secretion of alpha-amylase, a non-lipoprotein. *J Biol Chem* **274**: 1698–1707.
- Tjalsma H, Bolhuis A, Jongbloed JD, Bron S & van Dijl JM (2000) Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* **64**: 515–547.
- Tjalsma H, Bron S & van Dijl JM (2003) Complementary impact of paralogous Oxa1-like proteins of *Bacillus subtilis* on post-translocational stages in protein secretion. *J Biol Chem* **278**: 15622–15632.
- Ton-That H, Marraffini LA & Schneewind O (2004) Protein sorting to the cell wall envelope of Gram-positive bacteria. *Biochim Biophys Acta—Mol Cell Res* **1694**: 269–278.
- Ton-That H & Schneewind O (2004) Assembly of pili in Gram-positive bacteria. *Trends Microbiol* **12**: 228–234.
- Van der Laan M, Nouwen NP & Driessen AJ (2005) YidC—an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. *Curr Opin Microbiol* **8**: 182–187.
- Venema R, Tjalsma H, van Dijl JM, de Jong A, Leenhouts K, Buist G & Venema G (2003) Active lipoprotein precursors in the Gram-positive eubacterium *Lactococcus lactis*. *J Biol Chem* **278**: 14739–14746.
- Waligora AJ, Hennequin C, Mullany P, Bourlioux P, Collignon A & Karjalainen T (2001) Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. *Infect Immun* **69**: 2144–2153.
- Woese CR (1987) Bacterial evolution. *Microbiol Rev* **51**: 221–271.
- Xu Q, Bayer EA, Goldman M, Kenig R, Shoham Y & Lamed R (2004) Architecture of the *Bacteroides cellulosolvens* cellulosome: description of a cell surface-anchoring scaffoldin and a family 48 cellulase. *J Bacteriol* **186**: 968–977.
- Yang HS & Hooper JK (1998) Molecular cloning of a light-inducible gene (*lipA*) encoding a novel pilin from *Arthrobacter photogonimos*. *FEMS Microbiol Lett* **163**: 85–89.

## Assessment of Hierarchical Clustering Methodologies for Proteomic Data Mining

Bruno Meunier,<sup>\*,†</sup> Emilie Dumas,<sup>‡</sup> Isabelle Picc,<sup>§</sup> Daniel Béchet,<sup>||</sup> Michel Hébraud,<sup>‡,⊥</sup> and Jean-François Hocquette<sup>†</sup>

UR 1213, Unité de Recherches sur les Herbivores, Equipe Croissance et Métabolisme du Muscle, INRA de Clermont-Ferrand/Theix, F-63122 Saint-Genès Champanelle, France, UR454 Microbiologie, Equipe Qualité et Sécurité des Aliments (QuaSA), INRA de Clermont-Ferrand/Theix, F-63122 Saint-Genès Champanelle, France, Laboratoire d'Immunologie EMI 0351, INSERM, 3 rue des Louvels, F-80036 Amiens, France, UMR 1019, Unité de Nutrition Humaine, INRA de Clermont-Ferrand/Theix, F-63122 Saint-Genès Champanelle, France, and Plate-forme protéomique, INRA de Clermont-Ferrand/Theix, F-63122 Saint-Genès Champanelle, France

Received July 12, 2006

**Abstract:** Hierarchical clustering methodology is a powerful data mining approach for a first exploration of proteomic data. It enables samples or proteins to be grouped blindly according to their expression profiles. Nevertheless, the clustering results depend on parameters such as data preprocessing, between-profile similarity measurement, and the dendrogram construction procedure. We assessed several clustering strategies by calculating the *F*-measure, a widely used quality metric. The combination, on logged matrix, of Pearson correlation and Ward's methods for data aggregation is among the best clustering strategies, at least with the data sets we studied. This study was carried out using PermutMatrix, a freely available software derived from transcriptomics.

**Keywords:** proteomics • bioinformatics • data mining • hierarchical clustering • 2-D PAGE

### Introduction

Quantitative proteomics using two-dimensional polyacrylamide gel electrophoresis techniques (2-D PAGE) generates ever more data, building a growing haystack in which the biochemist hopes to find the miraculous needle. The main difficulty in exploring these data lies in the high number of variables, in this particular case, the relative abundance of hundreds or thousands of protein spots ( $P \gg 100$ ), compared to the small number ( $n_{\text{bio}} < 10$  or more) of individuals or biological samples of interest analyzed by 2-D PAGE with often few replications ( $n_{\text{exp}} < 3-5$ ). Thus, data mining is an important tool for a first exploration of proteomic results, since the amount of data rises with the increasing sensitivity and availability of the technique,

although the data quality remains insufficient from a statistical point of view. In fact, abundance measurement distribution on a single gel is often asymmetrical and extended over a wide dynamic range. In addition, the data matrix may have many missing values (MV) due to a real absence of protein spots or miss-detection by the image analysis software due to any technical difficulty, from protein extraction to electrophoresis. In conclusion, proteomic experiments may prove to be a pitfall if special care is not taken and notably during bioinformatic analysis.<sup>1</sup>

The first objective of proteomic data mining is differential expression analysis. This consists of comparing two or more predefined biological conditions to accurately highlight the few protein spots of interest among noise. Current statistical methods give good results when data are correctly preprocessed.<sup>2,3</sup>

In this paper, we were interested in the second objective of proteomic data mining, that is, the need to group or to classify individuals according to their global expression profiles, without any *a priori* knowledge of the biological reasons for the existence of these groups. This may be of great interest for understanding complex biological systems, such as characterizing tumor samples for clinical diagnostics,<sup>4,5</sup> or investigating the genetic variability of different ecotypes, for example.<sup>6</sup> Thus, Appel et al.<sup>7</sup> introduced early (1988) the automatic classification of 2-D PAGE in the MELANIE program. In the same way, it may also be interesting to group protein spots with similar expression profiles among the samples,<sup>8</sup> for example, in kinetic studies<sup>9,10</sup> or for discovering biomarkers of a specific biological function. To this end, several multivariate analysis methods have been reviewed.<sup>11,12</sup> Here, we focus on the hierarchical clustering analysis (HCA), mainly used in proteomics.

HCA is a technique of choice for exploring and visualizing large data sets, such as in transcriptomic studies, with the analysis of microarrays.<sup>13</sup> Briefly, HCA consists in calculating the dissimilarity, usually called the distance, between the individuals (each one being analyzed with one chip or in one gel) with one individual corresponding generally to one column of the data matrix. In the case of two-way HCA, the same procedure is carried out on the variables (the genes or the proteins) which correspond to the rows of the data matrix.

\* To whom correspondence should be addressed. Phone: +33 473624097. Fax: +33 473624639. E-mail: bruno.meunier@clermont.inra.fr.

† UR1213, Unité de Recherches sur les Herbivores, Equipe Croissance et Métabolisme du Muscle, INRA de Clermont-Ferrand/Theix.

‡ UR454 Microbiologie, Equipe Qualité et Sécurité des Aliments (QuaSA), INRA de Clermont-Ferrand/Theix.

§ Laboratoire d'Immunologie EMI 0351, INSERM.

|| UMR1019, Unité de Nutrition Humaine, INRA de Clermont-Ferrand/Theix.

⊥ Plate-forme protéomique, INRA de Clermont-Ferrand/Theix.

About 87% of the 28 recent proteomic publications analyzed in this study preferred the Pearson correlation coefficient or the Euclidean distance for HCA (supplementary Table 1, Supporting Information). Once the distance matrix is calculated, an agglomerative clustering algorithm is generally performed to construct the final dendrogram, a tree-like representation in which the closest objects or clusters (connected with smallest branches) are mathematically the most similar. There too, several algorithms can be employed. In proteomics, the unweighted paired group average linkage (UPGMA), complete linkage, and Ward's methods are the most commonly used aggregation procedures. A synthetic description and comparison of these clustering methodologies applied to gene expression clustering is reviewed in D'haeseleer et al.<sup>14</sup> Therefore, HCA needs to be tuned, first for selection of the distance measurement and second for the aggregation procedure. At present, it appears difficult for the biochemist to make this choice on his own, since these parameters are not always clearly described in the proteomic literature. Moreover, the clustering result is sometimes announced as biologically doubtful in some of the publications, which leads the authors to declare that HCA is not relevant when applied to their data. Preprocessing of data is almost never mentioned, although this step is known to be essential. Finally, we identified as many statistical tools needed for HCA as publications, which also complicates the choice. Unlike proteomics, transcriptomics helps to cope with these difficulties, proposing a lot of free (or costly) bioinformatic tools and keys for their utilization. The effect of missing values, need for data transformation, and effect of clustering methods have also been largely evaluated.<sup>15</sup>

This article aims, therefore, to determine whether one of the tools and the methods developed for microarray technologies could be employed to improve and to facilitate HCA of complex proteomic data sets. PermutMatrix software<sup>16</sup> has been identified as one of the few freely available programs permitting evaluation of the overall major HCA methods described previously. Two real and complementary proteomic data sets have been used to evaluate the effects of distance and the aggregation procedure on the clustering results. Furthermore, several innovative data preprocessing operations are proposed to potentially enhance the quality of the results arising from HCA methods. All these methodologies are evaluated and then discussed in this paper.

## Materials and Methods

**Data Set Description.** The first data set (DS1) results from a proteomic analysis which was performed on 12 *Listeria monocytogenes* strains, a food-borne pathogen bacterial species. The strains belonged to three different serovars (1/2a, 1/2b, and 4b) and had three different origins and levels of virulence. For each strain, two protein extractions were made from two independent cultures, and at least three 2-D PAGE gels with 50  $\mu$ g of protein were performed per protein sample. Finally, more than 72 gels were performed. Each run for the 2-D PAGE separation was carried out with 12 (first dimension) or 6 (second dimension) protein samples obtained from 6 different bacterial strains which allowed the technical variability to be taken into account. Silver-stained 2-D PAGE gels were scanned using a GS-800 imaging densitometer (Bio-Rad Laboratories), and image analysis was performed using Image Master 2D Platinum (GE Healthcare). Saturated spots were excluded in order to work in the linear range of silver nitrate staining. The five most reproducible gels were selected for each strain by

using the scatter plot analysis tool (experimental replications *per* condition  $n_{\text{epx}} = 5$ ). The 60 selected gels for the 12 bacterial strains were then matched with a reference gel. The reference gel was produced with a mixture of equal quantities of protein extracts from the 12 strains to visualize a compilation of all the protein spots. Each gel (ex: S01\_1 for the gel 1 of the strain 1) was then characterized by its expression profile, a vector of 599 relative volume (% vol) values (corresponding to the  $P = 599$  protein spots matched across the  $N = 60$  gels). The final DS1 consists of a  $[599 \times 60]$  volume matrix (VM) with about 36% of missing values.

The second data set (DS2) concerns a differential proteomic analysis which investigated rat age-related sarcopenia.<sup>17</sup> This study was undertaken on 3 groups of 5 rats (samples A–F) slaughtered at the age of 7, 18, or 30 months. 2-D PAGE gels were made in triplicate (1, 2, 3) on each individual protein extract (biological replications *per* condition  $n_{\text{bio}} = 5$ , experimental replications *per* condition  $n_{\text{epx}} = 3$ ). The batch composition was carefully randomized to take the experimental variability into account. The 45 colloidal blue-stained gel images were acquired using a GS-800 imaging densitometer, and image analysis was performed using PDQuest 2-D analysis software (Bio-Rad Laboratories). Each gel (ex: 7mA1 for gel 1 of rat A sacrificed at 7 months) was then characterized by its expression profile, a vector of 341 relative volume (% vol) values (corresponding to the  $P = 341$  protein spots matched across the  $N = 45$  gels). The final DS2 consists of a  $[341 \times 45]$  volume matrix (VM) with about 4% of missing values.

**Data Preprocessing Step 1: Missing Value Imputation.** Experimental conditions largely affect protein expression, revealing strong disparities in the data files. The first is missing values, which are known to be very disturbing for clustering algorithms.<sup>18,19</sup> To deal with this problem and complete the data matrix, one solution is to convert the original matrix to a binary matrix (VM  $\rightarrow$  VM-Bin). Existing values are thus replaced by "1" and missing values are replaced by "0". This first approach is well-accepted, but it reduces the richness of information drastically, since the quantitative measures are permanently lost.

A second solution consists in replacing all missing values by zero (VM  $\rightarrow$  VM-Zero). This solution makes the hypothesis that a missing value corresponds, in most cases, to a protein undetectable by staining due to its very low expression level. This is not always true, which is why this approach is highly criticized in transcriptomics where the  $k$ -nearest neighbor (KNN) method is generally advised for imputing missing values.<sup>19</sup> But, in proteomics, the origin of missing values is fundamentally different, and the number of missing values may be too important ( $>10\%$ ) for an efficient imputation with a KNN algorithm. These last considerations are, in fact, valid for both procedures (VM-Bin, VM-Zero). Finally, the two processed data matrices (VM-Bin, VM-Zero) remain the same size [ $P \times N$ ] as the original one (VM), so that no protein spot is discarded before further analysis. This is the only advantage of these two procedures. Now we consider that these two MV imputation procedures are correct, since extreme care was taken over gel-making for both data sets. Consequently, we may suppose that missing values did not occur by random chance but were biologically meaningful.

A third solution is nevertheless considered. This consists in working only on the reliable protein spots. Thus, a protein spot is kept if at least  $(n_{\text{epx}} - 1)$  volume values are available for all samples. Then, any missing value in a sample is replaced by

the mean of the existing values for this protein spot in that sample. In the other cases, the protein spot is definitely excluded. This procedure results in a reduction of the original data matrix (VM) to a reliable one (VMR) of smaller size [ $P' \times N$ ],  $P' \leq P$ . Even if this procedure is not very stringent, it is nevertheless useful when a reasonable number of replications is available ( $n_{\text{rep}} > 3$ ).

Finally, it is important to note that, before HCA is applied, protein spots of interest (statistically significant ones, for example, or proteins with sufficient variance) do not have to be selected using an ANOVA filter as is usually done in proteomics.<sup>20</sup> We thus preferred here to deal with all the protein spots, without any *a-priori* knowledge or without any prior analysis (such as ANOVA), even if proteins whose variation is due only to experimental conditions may be disturbing for any clustering algorithm.

**Data Preprocessing Step 2: Volume Normalization.** A second problem inherent to proteomic data sets is the result of the wide dynamic range characteristic of the volume distribution. Thus, certain very abundant protein spots may weigh much more on the clustering results than the majority of weakly abundant ones. To deal with this problem, Vohradsky<sup>8</sup> proposed dividing the expression profiles of the samples by the maximum value. In the same way, Gion et al. proposed a log transformation.<sup>21</sup> We chose a logged ratio-based approach that also acts as a data normalization step. This type of procedure is classically used for transcriptomics because fold change is indeed an interesting way to investigate differential expression.<sup>3</sup> To this end, each protein spot volume was divided by the mean of all the existing values for this protein spot, present in all the  $N$  gels. Unfortunately, this procedure can only be performed on the reliable data matrix (VMR) because the ratio cannot be calculated on zero or missing values. This ratio-based data matrix was then submitted to the usual base 2 logarithmic transformation (VMR-Ratio). This approach allowed each sample (a single gel) to be compared to a common mean sample of reference (a synthetic average gel), so that all the samples could then be compared one-to-one (0, no difference with the mean; +1, 2-fold greater than the mean; -1, 2-fold less than the mean). In this respect, this approach is similar to a recent and powerful proteomic approach, fluorescent two-dimensional differential gel electrophoresis (2-D DIGE), for which dedicated software such as the Extended Data Analysis module of DeCyder (GE Healthcare) is available.

An alternative method consists in standardizing data row by row using the classic zero-mean and unit-standard deviation technique.<sup>22</sup> This normalization has the advantage of being applicable either to the raw data matrix (VM-Center) or to the reliable one (VMR-Center). Moreover, this procedure is almost always implemented in the statistical package or the bioinformatic software available in proteomics.

Finally, the data matrices were submitted to a two-way normalization, the first being vertical (% vol), which makes gels comparable, the second horizontal, and this time making the proteins comparable.

**Hierarchical Clustering Analysis (HCA): Methods and Tool.** Once correctly preprocessed, the different data matrices may be submitted to two-way HCA; that is, HCA is applied independently to the columns and to the rows. The distance measurement selected for this evaluation is based either on the popular Pearson correlation (PE) coefficient ( $1 - r$ ) or on the Euclidean metric (EU). The definition of these metrics is available in the Supporting Information. According to our own

assessment, PE is quoted in about 33% of the proteomic literature (supplementary Table 1, Supporting Information). It is also advised in transcriptomics<sup>14</sup> when absolute spot hybridization levels are used. In fact, PE is less sensitive to the scale factor than EU so it may also be more appropriate to protein spot volumes. EU is used in about 54% of the proteomic literature (supplementary Table 1, Supporting Information). It is generally advised for normal data, and consequently, it seems more suited to logged ratio data<sup>14</sup> than raw volumes whose distribution is non-normal. EU is also more sensitive to missing values because this metric is differently weighted according to the number of MV contained in the profiles. Finally, both metrics (EU and PE) are affected more strongly by very abundant protein spot variation (minority) than low-abundance ones (majority). Consequently, profile normalization and MV imputation are essential. A third distance based on the popular Jaccard index (JA) is used exclusively for binary matrix computing.<sup>23</sup> The definition of this metric is available in the Supporting Information.

Concerning the aggregation procedures, the most commonly used ones: UPGMA (UP), complete linkage (CO), and Ward's methods (WA) have been tested in this study. A brief description of these terms is available in the Supporting Information. In transcriptomics, use of the complete linkage method gives generally better results<sup>15</sup> than UPGMA, but the latter is widely used in proteomics (52%) based on listed references (supplementary Table 1, Supporting Information). In contrast, Ward's method is almost never used with microarray data sets, whereas 20% of proteomic authors choose this aggregation procedure (supplementary Table 1, Supporting Information). The reasons for such differences between transcriptomics and proteomics are unknown and probably result simply from different practices without any precise knowledge of the advantages and limitations of each procedure. In proteomics, raw volumes are assumed to be not-normally distributed and "variance-vs-mean"-dependent. We can notably observe an increase of variance among the highest spot volumes. Data normalization is then generally beneficial<sup>24</sup> to any parametric approach such as the Ward's agglomeration method based on intragroup variance minimization or the UPGMA based on average calculation.

To evaluate each clustering methodology [a combination of data preprocessing steps 1 and 2  $\times$  distance  $\times$  aggregation procedure], powerful and user-friendly software are needed. At present, transcriptomics offers a greater abundance of freely available data mining tools. The first and most popular one is certainly the Cluster+TreeView program by Eisen et al.<sup>13</sup> In addition to offering several clustering methodologies and the classic dendrogram display, this tool allows heat map visualization of the complete data matrix, a useful color representation in which each data point color is proportional to its value. Unfortunately, the clustering of binary data is not clearly designed, and Ward's method is not implemented in this tool, similarly to most other programs. Consequently, it is proposed here to use a new software program, PermutMatrix, which was also developed for microarray data but offers the entire panel of methods and visualization possibilities needed for this study. PermutMatrix version 1.8 is available free at <http://www.lirmm.fr/~caraux/PermutMatrix/>. Its convivial interface allows the five following steps for HCA:

(i) Uploading of a standard tab-delimited text file containing the data matrix

1. Columns represent samples (e.g., gels)

2. Rows represent protein spots
  - (ii) Setting of both clustering parameters, the distance and the aggregation procedure
    1. Distance = Pearson (PE), Euclidean (EU), or Jaccard (JA)
    2. Aggregation = UPGMA (UP), Complete (CO), or Ward (WA)
  - (iii) HCA application on the columns
  - (iv) HCA application on the rows
  - (v) Optimal reorganization of the trees using seriation methods (optional)
  - (vi) Clustering result visualization
    1. Dendrogram of the samples
    2. Dendrogram of the protein spots
    3. Heat map of the clustered data matrix

**Clustering Result Validation.** In most proteomic studies, HCA is first used to construct the dendrogram of the samples. Then the biochemist generally accepts at least one of the grouping results and interprets it based on his prior biological knowledge. Unfortunately, the choice of the optimal clustering solution is sometimes subjective and, consequently, debatable and sometimes inadequate. To validate a clustering result, its quality may also be evaluated statistically using various “internal” criteria, but these are known to be biased because they are highly correlated with the clustering algorithm.<sup>14</sup> Thus, “external” criteria are generally preferred when methodologies have to be compared.<sup>14,25</sup> Therefore, it is proposed to use an external quality criterion based on the most reasonable hypothesis that can be intuitively raised here: since experimental variability is generally expected to be lower than biological variability, natural partition of the data results in grouping the experimental replications in classes of  $n_{\text{epx}}$  gels. Then, a useful quality metric, the *F*-measure,<sup>25</sup> may be computed to compare each entire clustering solution to the natural partition. The *F*-measure is a combination of the “precision” and the “recall” assessment, where precision expresses the proportion of well-classed objects (e.g., gels) for each cluster and recall expresses the proportion of well-classed objects for each class. The *F*-measure is computed as described in the Supporting Information,<sup>25</sup> with this positive number taking 1 as maximum value when the clustering solution is equal to the natural partition. Finally *F*-measure is actually relevant because all replicate gels come from different batches. This condition is absolutely essential. Additionally, all the dendrograms constructed under PermutMatrix were positively validated by STATISTICA version 6.1 (StatSoft, France) which is a widely recognized and easy-to-use statistical analysis software. The topology of these dendrograms may also be a good indicator of the clustering quality. In fact, a well-balanced tree generally reflects biological reality better than an unbalanced one, plots in stairs being a good example.

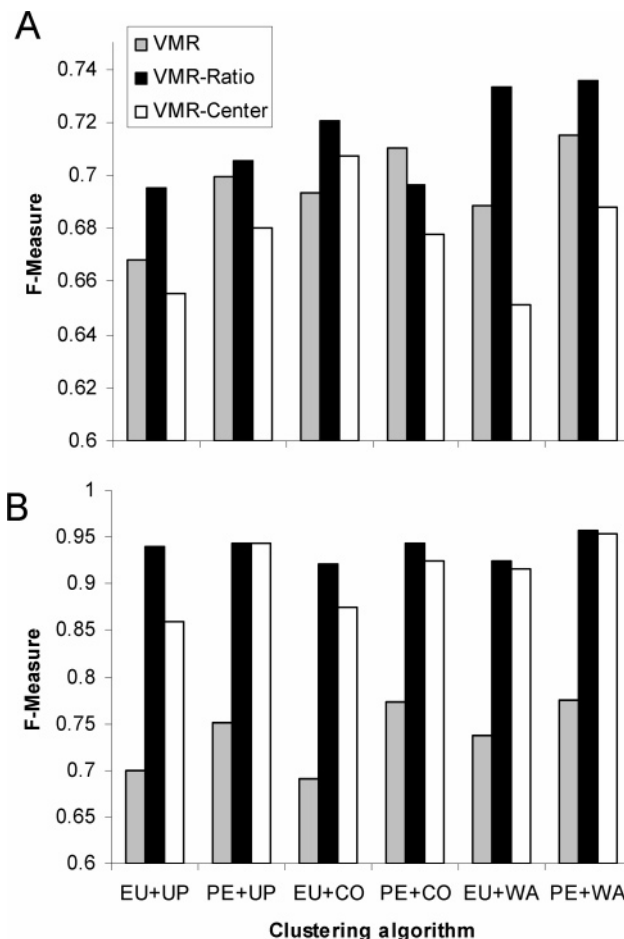
**Clustering Method Evaluation.** First, to evaluate the global effect of the clustering methodology on the clustering results, the *F*-measure was calculated for each clustering solution. All results were presented as means between both data sets.

Second, the effect of the missing value imputation method was compared using suitable clustering methodologies, that is, those giving the highest *F*-measure.

Finally, the global capability of the clustering methodology was evaluated from a practical point of view, including visual diagnostics, cluster extraction possibility, and conviviality.

## Results and Discussion

**Effect of the Algorithm and the Data Normalization.** We evaluated the effect of the global methodology on the clustering



**Figure 1.** Effects of the algorithm and the data normalization methods on the clustering results validated by *F*-measure. Bar graphs depict the *F*-measure levels obtained for each clustering solution. (A) Hierarchical clustering analysis (HCA) applied to the reliable matrix (VMR) of the first data set (DS1) without data normalization (gray bars), after logged ratio transformation (black bars), or 0-mean and 1-standard deviation normalization (white bars). The clustering algorithms result from a combination of a distance metric (EU = Euclidean or PE = Pearson correlation) and an aggregation procedure (UP = UPGMA, CO = Complete, or WA = Ward). (B) Similar analysis made with the second data set (DS2).

result of the samples. The selected algorithms resulted from a combination of a distance metric (EU or PE) and an aggregation procedure (UP, CO, or WA). They were executed on the reliable matrix of each data set (DS1 and DS2) after one of the proposed volume normalization procedures: (i) VMR (raw volume without normalization), (ii) VMR-Ratio, and (iii) VMR-Center. Bar graphs in Figure 1 depict the *F*-measure levels obtained for all possible clustering solutions.

**1. Data Normalization Effect.** The effect of the data normalization method was important, but a significant interaction was observed between data normalization and data set (Table 1A). The Ratio-based approach (VMR-Ratio) showed a higher performance for both data sets than the center-based (VMR-Center) and than the no-normalization methods (VMR), and this effect was higher for DS2 than for DS1. The worst method was VMR-Center for DS1 and VMR for DS2.

**2. Clustering Algorithm Effect.** The effect of the clustering algorithm was also important, but a significant interaction was observed between clustering algorithm and data set (Table 1B).

**Table 1.** Effects of the Data Normalization Method (A) and the Clustering Algorithm (B) on the Clustering Results for Both Data Sets (DS1 and DS2) Validated by the *F*-Measure<sup>a</sup>

A						
data set (DS)	data normalization method (DNM)					
	VMR	VMR-Ratio	VMR-Center			
DS1	0.696	<b>0.714</b>	0.677			
DS2	0.738	<b>0.938</b>	0.912			

B						
data set (DS)	clustering algorithm (CA)					
	EU+UP	PE+UP	EU+CO	PE+CO	EU+WA	PE+WA
DS1	0.673	0.695	0.707	0.695	0.691	<b>0.713</b>
DS2	0.833	0.879	0.829	0.880	0.859	<b>0.896</b>

<sup>a</sup> (A) Three data normalization methods (DNM) are evaluated: (i) VMR (raw volume without normalization), (ii) VMR-Ratio (logged ratio transformation), and (iii) VMR-Center (0-mean and 1-standard deviation normalization). (B) The six clustering algorithms (CA) tested are a combination of one distance [Pearson (PE) or Euclidean (EU)] plus one aggregation procedure [UPGMA (UP), Complete (CO), or Ward (WA)]. For each data set, highest *F*-measures (corresponding to best clustering strategies) are indicated in bold.

Nevertheless, the PE+WA algorithm was the best solution whatever the data set used, even if the difference between this algorithm and the others was sometimes low. Consequently, this result needs to be confirmed in further studies. Taken individually, PE was a better distance metric than EU in most of the cases, but this difference was also weak. Similarly, Ward's aggregation procedure appeared noticeably more powerful than the others.

**3. Validation of the *F*-Measure by Biological Considerations.** Additionally, it was interesting to note the high correlation between this *F*-measure, reflecting the quality of the replication grouping, and a biologically relevant grouping of the samples. Thus, in DS1 analysis, only the clustering methodologies giving an *F*-measure above about 0.7 led to clear separation of strains 1–6 from strains 7–12 in a balanced tree (Figure 2). This result was consistent with the phylogenetic analyses based on genomic data which showed that the 13 recognized serovars identified in the *L. monocytogenes* species were distributed between two different lineages, with serovars 1/2b and 4b strains in particular found in lineage I and serovar 1/2a strains found in lineage II.<sup>26,27</sup> Identification and validation of the protein markers splitting the two lineages is under way. Concerning DS2, a similar phenomenon appeared for *F*-measure above about 0.9, leading to a clear distinction of 30-month-old rats from others and a weaker distinction between 7- and 18-month-old rats (Figure 3). ANOVA analysis previously performed on this data set showed a similar result.<sup>17</sup> Indeed, mean differences observed for the significant differentially expressed proteins between 7- vs 18-, 18- vs 30-, and 7- vs 30-month-old rats were respectively 77%, 92%, and 107%, whereas the mean coefficients of experimental and biological variations were 25.3% ( $n_{\text{exp}} = 3$ ) and 29.3% ( $n_{\text{bio}} = 5$ ), respectively. Consequently, the biological differences detected by both ANOVA and clustering were higher than technical or biological variabilities, which validates our analysis.

In conclusion, even if our results were shown with only two data sets, we have demonstrated here the weakness of popular methodologies, (e.g., EU+UP) which surprisingly gave the lowest *F*-measure (Figure 1), the worst clustering results, and the most unbalanced trees (data not shown). Data normaliza-

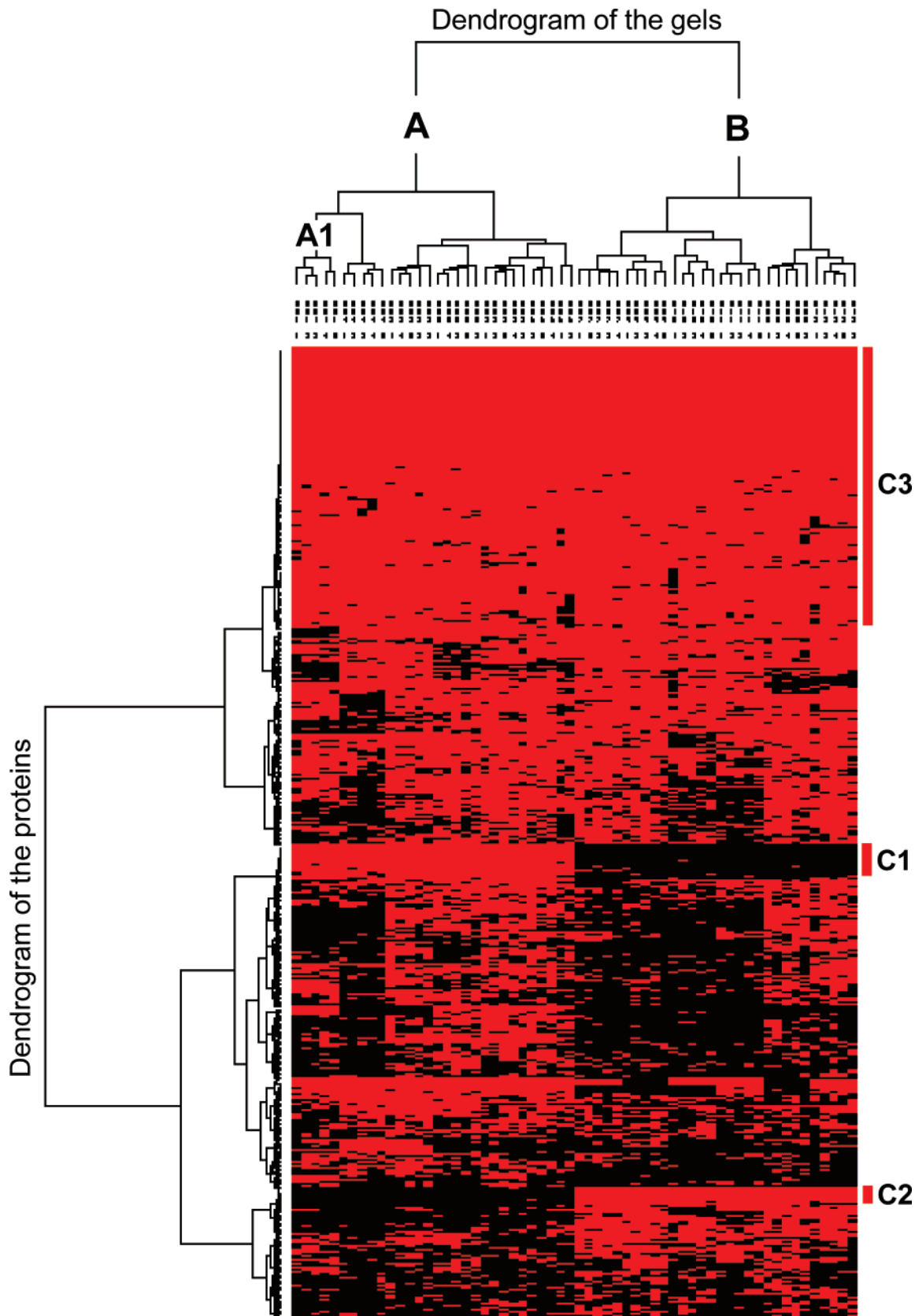
tion was also needed, although it is almost never mentioned in the publications. Thus, the application of the PE+WA algorithm to the VMR-Ratio gave the best *F*-measure in almost all the situations, leading to well-balanced and easily legible dendrograms (Figures 2 and 3). Unfortunately, this data normalization was only applicable on a reliable (and consequently generally reduced) data matrix containing no missing value.

**Effect of the Missing Value Imputation.** In this second comparative study, the overall missing value imputation techniques were evaluated based on the algorithm that performed best, adapted to each data set: JA+WA was selected for HCA of the binary matrix (VM-Bin) and PE+WA was used for the other matrices. Table 2 summarizes these methods and the maximum *F*-measure levels obtained for each best clustering solution. Interestingly, HCA applied to the binary matrix of the first data set led to an almost perfect grouping of replicate gels (*F*-measure = 0.97). Here, the effect of absent/present biomarkers seemed predominant compared to the quantitative values. Nevertheless, clustering of the VM-Zero-Center was equally efficient (*F*-measure = 0.91), with this matrix integrating the benefits of both quantitative and binary information. Because of 36% missing values, the clustering software was unable to process the raw matrix VM-Center. Finally, HCA of the VMR-Ratio gave the worst result, notably because only 140 out of a total of 599 proteins were taken into account. Thus, we may suppose these 140 proteins are not only common to all strains but most of them are housekeeping proteins characterized by similar expression profiles. Concerning the second data set, HCA applied to the binary matrix was this time clearly uninformative; 4% missing values seemed inadequate to contain pertinent or even discriminating information. The process applied to the three other normalized matrices (VM-Center, VM-Zero-Center, and VMR-Ratio) gave quite similar and satisfying results (0.93, 0.94, and 0.95, respectively), with the maximum *F*-measure value being obtained with PE+WA algorithm.

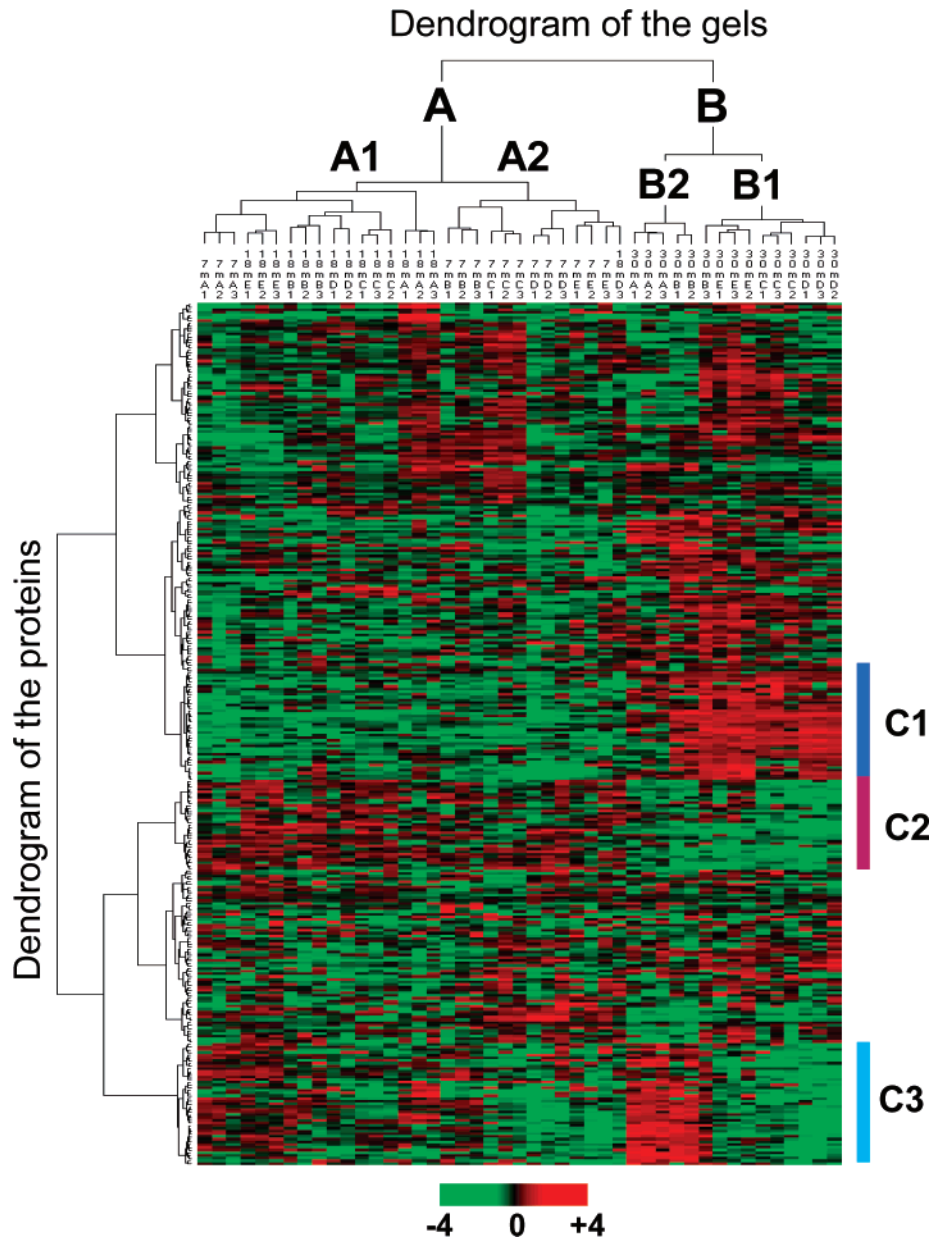
In conclusion, each data set is specific and, consequently, may need one or several suitable treatments. Here, both data sets were of interest and complementary for our study. DS1 was mostly characterized by genetic differences between samples involving notably numerous absent/present markers. A binary approach was informative and powerful in this case. On the contrary, DS2 was mainly constructed for differential expression analysis where small quantitative differences were expected. Here, a ratio-based approach seemed more sensitive for revealing proteins of interest.

**Major HCA Assets.** Once the samples are correctly grouped, either automatically using the well-adapted clustering methodology or manually leaving samples naturally arranged in order (e.g., in a kinetic study), it may be interesting to group proteins showing similar expression profiles. To this end, two-way HCA is a suitable preliminary data mining approach. Heat map visualization of the binary (Figure 2) or the quantitative (Figure 3) clustered data matrices gives powerful diagnostic information. Groups of coregulated proteins can be rapidly identified, either by their absence/presence pattern as shown in clusters C1 and C2 of Figure 2, or their quantitative difference, in clusters C1 and C2 of Figure 3. In this latter figure, only the data normalization techniques recommended previously (-Ratio or -Center) allowed this powerful visualization and the possibility for extraction of complex patterns. In a similar manner, it may be interesting to identify natural or





**Figure 2.** Two-way hierarchical clustering analysis (processed with PermutMatrix according to the Jaccard index and Ward's aggregation method) of the first data set (DS1, 599 proteins  $\times$  60 gels) after a binary transformation (VM-Bin). Heat map representation of the clustered data matrix in which each red cell represents an existing protein value (1) and each black cell represents a missing protein value (0). The dendrogram of the gels shows a balanced tree in which strains 1–6 are clearly separated (parent cluster [A]) from strains 7–12 (parent cluster [B]). Moreover, replicated gels are mostly grouped in separate subclusters (child cluster [A1], for example) leading to the higher *F*-measure obtained for this data set (0.97). The dendrogram of the proteins allows visual selection of coregulated proteins, for example, biomarkers present only in a predefined selection of strains: cluster [C1] of 22 proteins present only in the strains 1–6 and cluster [C2] of 12 proteins present only in the strains 7–12. In contrast, cluster [C3] is composed of 174 “reliable” proteins with relatively few missing values; thus, this cluster needs to be reanalyzed in a quantitative manner.



**Figure 3.** Two-way hierarchical clustering analysis (processed with PermutMatrix according to the Pearson distance and Ward's aggregation method) of the second data set (DS2, 318 proteins  $\times$  45 gels) after a logged-ratio transformation (VMR-Ratio). Heat map representation of the clustered data matrix in which each colored cell represents a protein value according to the color scale at the bottom of the figure. The dendrogram of the gels depicts a balanced tree in which all 30-month-old rats are clearly separated [B] from 7- and 18-month-old rats grouped in a common cluster [A] with a weak distinction between 7- (cluster [A2]) and 18-month-old rats (cluster [A1]). Clusters of coregulated proteins [C1] or [C2] illustrate this phenomenon, respectively, in terms of up-expression or down-expression in 30-month-old rats. Among the five rats aged 30 months, two (cluster [B2], characterized by [C3]) appear to act like outliers. Additionally, replicated gels are mostly grouped in separate subclusters leading to the higher  $F$ -measure obtained for this data set (0.95).

artificial (mainly due to experimental errors) groups of outliers, among the samples or the proteins. Thus, HCA of DS2 highlighted clearly two rats inside the "30-months" group (Figure 3, clusters B2/C3), whereas, on the contrary, few differences were observed inside the 7- and 18-months one. This first understanding of the data set, without any *a priori* knowledge, may orient future data analysis strategy such as new model construction for ANOVA or data filtering before further processing.

**General Conclusions.** As already shown in a previous work,<sup>3</sup> there is much similarity between transcriptomic and proteomic

data analysis strategy, and it needs to be exploited. Consequently, it is clear that bioinformatic tools, mainly developed for microarrays, could be adapted to 2-D PAGE-based studies. Unfortunately, this is not yet generally true and, in any case, needs careful attention. The most popular methodologies are sometimes weak. Here, we recommend PermutMatrix, a new free software implementing HCA methodology. We have shown that it is well-adapted to the present proteomic data sets. In addition, we have highlighted specific features of proteomic data: unlike in transcriptomics, where Ward's aggregation method is almost never employed, this procedure gives the best

**Table 2.** Effect of the Missing Value (MV) Imputation Method on the Clustering Results Validated by the *F*-Measure<sup>a</sup>

data set	matrix	size	MV (%)	best algorithm		<i>F</i> -measure MAX
				distance	aggregation	
DS1	VM-Center	599 × 60	36	-	-	-
	VM-Bin	599 × 60	0	Jaccard	Ward	<b>0.97</b>
	VM-Zero-Center	599 × 60	0	Pearson	Complete	0.91
	VMR-Ratio	140 × 60	0	Pearson	Ward	0.73
DS2	VM-Center	341 × 45	4	Pearson	Ward	0.93
	VM-Bin	341 × 45	0	Jaccard	Ward	0.39
	VM-Zero-Center	341 × 45	0	Pearson	Ward	0.94
	VMR-Ratio	318 × 45	0	Pearson	Ward	<b>0.95</b>

<sup>a</sup> Four methods are proposed: (i) no imputation (VM-Center), (ii) binary transformation (VM-Bin), (iii) replacement by zero (VM-Zero-Center), and (iv) replacement by mean (if available) of replicate values (VMR-Ratio). For each imputation technique, the corresponding matrix, reduced or not, is submitted to the HCA method (best algorithm and normalization procedure) giving the maximum *F*-measure.

clustering results in both these proteomic studies, notably, when associated with the Pearson-based distance metric. Similarly, the clustering of binary data appears particularly informative either to identify visually a protein spot strictly absent in one condition or to highlight gels with too many MVs. It is also important to note that the presence of replicate gels is an absolute prerequisite to select objectively the best clustering solutions among several, all of which give results that are sometimes quite different. This is especially valuable if technical replicate gels result from different batches, from extraction to staining. Interestingly, this well-tuned HCA approach gives the most biologically relevant grouping of the samples. Biological replicates are, thus, essential to strengthen the concluding remarks about the population of concern.<sup>28</sup> In conclusion, the two-way clustering approach is easy to implement, but special care must be taken with respect to data normalization, missing value imputation (which depends on data set specific features), and cluster validation (by “external” criteria). HCA may also be advantageously combined with other multivariate exploratory techniques such as principal components analysis (PCA) or *K*-means and Self-Organizing Maps (SOM)<sup>29</sup> clustering methods.

**Acknowledgment.** We thank Drs. I. Cassar-Malek, J. F. Martin, C. Jurie, and C. Bernard for constructive and helpful discussions. This work was supported by the AGENAE program. AGENAE is a national French program related to structural and functional genomics applied to animal science. The authors also thank the « Commissariat à l’aménagement et au Développement Economique du Massif Central » (France) for its financial support.

**Note Added after Print Publication.** On p 358, the affiliation addresses of the authors appear with incorrect zip codes in the version posted on the web December 14, 2006 (ASAP), and published in the January 2007 issue (Vol. 6, No. 1, pp 358–366; the correct electronic version of the paper was published February 7, 2007, and an Addition and Correction appears in the March issue (Vol. 6, No. 3).

**Supporting Information Available:** Supplementary Table 1 depicts the 28 recent proteomic publications analyzed in this study. In addition, a brief description of clustering terms is proposed. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Biron, D. G.; Brun, C.; Lefevre, T.; Lebarbenchon, C.; Loxdale, H. D.; Chevenet, F.; Brizard, J. P.; Thomas, F. The pitfalls of proteomics experiments without the correct use of bioinformatics tools. *Proteomics* **2006**, *6*, 5577–5596.
- (2) Chang, J.; Van Remmen, H.; Ward, W. F.; Regnier, F. E.; Richardson, A.; Cornell, J. Processing of data generated by 2-dimensional gel electrophoresis for statistical analysis: missing data, normalization, and statistics. *J. Proteome Res.* **2004**, *3*, 1210–1218.
- (3) Meunier, B.; Bouley, J.; Piec, I.; Bernard, C.; Picard, B.; Hocquette, J. F. Data analysis methods for detection of differential protein expression in two-dimensional gel electrophoresis. *Anal. Biochem.* **2005**, *340*, 226–230.
- (4) Iwadata, Y.; Sakaida, T.; Hiwasa, T.; Nagai, Y.; Ishikura, H.; Takiguchi, M.; Yamaura, A. Molecular classification and survival prediction in human gliomas based on proteome analysis. *Cancer Res.* **2004**, *64*, 2496–2501.
- (5) Harris, R. A.; Yang, A.; Stein, R. C.; Lucy, K.; Brusten, L.; Herath, A.; Parekh, R.; Waterfield, M. D.; O’Hare, M. J.; Neville, M. A.; Page, M. J.; Zvelebil, M. J. Cluster analysis of an extensive human breast cancer cell line protein expression map database. *Proteomics* **2002**, *2*, 212–223.
- (6) Chevalier, F.; Martin, O.; Rofidal, V.; Devauchelle, A. D.; Barreau, S.; Sommerer, N.; Rossignol, M. Proteomic investigation of natural variation between Arabidopsis ecotypes. *Proteomics* **2004**, *4*, 1372–1381.
- (7) Appel, R.; Hochstrasser, D.; Roch, C.; Funk, M.; Muller, A. F.; Pellegrini, C. Automatic classification of two-dimensional gel electrophoresis pictures by heuristic clustering analysis: a step toward machine learning. *Electrophoresis* **1988**, *9*, 136–142.
- (8) Vohradsky, J. Adaptive classification of two-dimensional gel electrophoretic spot patterns by neural networks and cluster analysis. *Electrophoresis* **1997**, *18*, 2749–2754.
- (9) Jimenez, C. R.; Stam, F. J.; Li, K. W.; Gouwenberg, Y.; Hornshaw, M. P.; De Winter, F.; Verhaagen, J.; Smit, A. B. Proteomics of the injured rat sciatic nerve reveals protein expression dynamics during regeneration. *Mol. Cell. Proteomics* **2005**, *4*, 120–132.
- (10) Culp, W. D.; Neal, R.; Massey, R.; Egevad, L.; Pisa, P.; Garland, D. Proteomic analysis of tumor establishment and growth in the B16–F10. mouse melanoma model. *J. Proteome Res.* **2006**, *5*, 1332–1343.
- (11) Dowsey, A. W.; Dunn, M. J.; Yang, G. Z. The role of bioinformatics in two-dimensional gel electrophoresis. *Proteomics* **2003**, *3*, 1567–1596.
- (12) Bensmail, H.; Haoudi, A. Postgenomics: proteomics and bioinformatics in Cancer Research. *J. Biomed. Biotechnol.* **2003**, *2003*, 217–230.
- (13) Eisen, M. B.; Spellman, P. T.; Brown, P. O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14863–14868.
- (14) D’haeseleer, P. How does gene expression clustering work? *Nat. Biotechnol.* **2005**, *23*, 1499–1501.
- (15) Gibbons, F. D.; Roth, F. P. Judging the quality of gene expression-based clustering methods using gene annotation. *Genome Res.* **2002**, *12*, 1574–1581.
- (16) Caraux, G.; Pinloche, S. PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* **2005**, *21*, 1280–1281.
- (17) Piec, I.; Listrat, A.; Alliot, J.; Chambon, C.; Taylor, R. G.; Bechet, D. Differential proteome analysis of aging in rat skeletal muscle. *FASEB J.* **2005**, *19*, 1143–1155.
- (18) Cox, B.; Kislinger, T.; Emili, A. Integrating gene and protein expression data: pattern analysis and profile mining. *Methods* **2005**, *35*, 303–314.
- (19) de Brevern, A. G.; Hazout, S.; Malpertuy, A. Influence of microarrays experiments missing values on the stability of gene groups by hierarchical clustering. *BMC Bioinf.* **2004**, *5*, 114.
- (20) Vohradsky, J.; Janda, I.; Grunfelder, B.; Berndt, P.; Roder, D.; Langen, H.; Weiser, J.; Jenal, U. Proteome of Caulobacter crescentus cell cycle publicly accessible on SWICZ server. *Proteomics* **2003**, *3*, 1874–1882.
- (21) Gion, J. M.; Lalanne, C.; Le, Provost, G.; Ferry-Dumazet, H.; Paiva, J.; Chaumeil, P.; Frigerio, J. M.; Brach, J.; Barre, A.; de Daruvar, A.; Claverol, S.; Bonneu, M.; Sommerer, N.; Negroni, L.; Plomion, C. The proteome of maritime pine wood forming tissue. *Proteomics* **2005**, *5*, 3731–3751.
- (22) Satoh, M.; Haruta-Satoh, E.; Omori, A.; Oh-Ishi, M.; Kodera, Y.; Furudate, S.; Maeda, T. Effect of thyroxine on abnormal pancreatic proteomes of the hypothyroid rdw rat. *Proteomics* **2005**, *5*, 1113–1124.

- (23) Zivy, M.; el Madidi, S.; Thiellement, H. Distance indices in a comparison between the A, D, I and R genomes of the Triticeae tribe. *Electrophoresis* **1995**, *16*, 1295–1300.
- (24) Gustafsson, J. S.; Ceasar, R.; Glasbey, C. A.; Blomberg, A.; Rudemo, M. Statistical exploration of variation in quantitative two-dimensional gel electrophoresis data. *Proteomics* **2004**, *4*, 3791–3799.
- (25) Handl, J.; Knowles, J.; Kell, D. B. Computational cluster validation in post-genomic data analysis. *Bioinformatics* **2005**, *21*, 3201–3212.
- (26) Zhang, C.; Zhang, M.; Ju, J.; Nietfeldt, J.; Wise, J.; Terry, P. M.; Olson, M.; Kachman, S. D.; Wiedmann, M.; Samadpour, M.; Benson, A. K. Genome diversification in phylogenetic lineages I and II of *Listeria monocytogenes*: identification of segments unique to lineage II populations. *J. Bacteriol.* **2003**, *185*, 5573–5584.
- (27) Doumith, M.; Cazalet, C.; Simoes, N.; Frangeul, L.; Jacquet, C.; Kunst, F.; Martin, P.; Cossart, P.; Glaser, P.; Buchrieser, C. New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. *Infect. Immun.* **2004**, *72*, 1072–1083.
- (28) Karp, N. A.; Spencer, M.; Lindsay, H.; O'Dell, K.; Lilley, K. S. Impact of replicate types on proteomic expression analysis. *J. Proteome Res.* **2005**, *4*, 1867–1871.
- (29) Herrero, J.; Dopazo, J. Combining hierarchical clustering and self-organizing maps for exploratory analysis of gene expression patterns. *J. Proteome Res.* **2002**, *1*, 467–470.

PR060343H

*APPROCHE PROTEOMIQUE POUR LA CARACTERISATION DES SOUCHES DE LISTERIA MONOCYTOGENES : RECHERCHE DE MARQUEURS DE LA PATHOGENICITE.*

DUMAS E., CHAFSEY I, CHAMBON C., HEBRAUD M.

*Listeria monocytogenes* est l'agent causal de la listériose, une infection alimentaire grave affectant essentiellement les personnes immunodéprimées, les personnes âgées et les femmes enceintes. Bien que la fréquence des cas de listériose ait diminué grâce aux efforts des industriels de la filière agroalimentaire et la mise en place d'un système de surveillance efficace, le risque sanitaire reste important avec un taux de mortalité d'environ 30%. Les gènes de virulence connus à ce jour sont présents dans les 13 sérovars de l'espèce. Cependant, plus de 90% des cas humains de listériose ne sont dus qu'aux sérovars 1/2a, 1/2b et 4b, et ce dernier semble être le plus virulent. Curieusement, alors qu'il est responsable de presque tous les cas épidémiques de listériose, il est beaucoup moins fréquemment isolé que les autres sérovars dans les aliments ou dans l'environnement industriel. Un projet en cours, réunissant plusieurs laboratoires et des approches complémentaires de génomique, post-génomique et divers tests de virulence, vise à mieux comprendre ce qui conditionne l'expression du pouvoir pathogène. L'objectif de l'étude est de développer une puce de typage afin de mieux évoluer et hiérarchiser les risques sanitaires auxquels s'exposent les consommateurs, en relation avec le sérovar et l'origine des souches. Dans le cadre de ce projet de recherche, nous sommes en charge de la caractérisation des protéomes de 12 souches de différents sérovars et de différentes origines (souches épidémiques, de portage asymptomatique et d'environnement). Les protéomes des protéines intracellulaires d'une part et des protéines sécrétées d'autre part, sont ainsi séparés par électrophorèse bidimensionnelle et comparés avec l'aide d'un logiciel d'analyse d'image. Des différences de profil protéique entre les 6 souches étudiées à ce jour, ont été mis en évidence. Les spots protéiques d'intérêt sont en cours d'identification par spectrométrie de masse MALDI TOF ou par comparaison avec des cartes protéiques de référence établies au laboratoire. Il apparaît d'ores et déjà qu'il est possible de séparer les souches de sérovar 4b et 1/2b de celles de sérovar 1/2a .

L'analyse des 6 autres souches et la compilation de l'ensemble des résultats devrait permettre de trouver des protéines marqueurs discriminantes des sérovars, de l'origine des souches, en relation avec leur pouvoir pathogène.

APPROCHE PROTEOMIQUE POUR LA CARACTERISATION DE SOUCHES DE  
*LISTERIA MONOCYTOGENES* : RECHERCHE DE MARQUEURS DE RISQUE

DUMAS E.<sup>1</sup>, MEUNIER B.<sup>2</sup>, BERDAGUE J.L.<sup>3</sup>, CHAMBON C.<sup>4</sup> et HEBRAUD M.<sup>1,4</sup>

UR Microbiologie - Equipe QuaSA<sup>1</sup>, URH - Equipe C2M<sup>2</sup>, UR QuaPA – Equipe T2A<sup>3</sup> et Plate-Forme Protéomique<sup>4</sup>, INRA site de Theix, 63122 Saint-Genès Champanelle

Mots-clés : *Listeria monocytogenes*, biodiversité, marqueurs de risque, protéomique.

*Listeria monocytogenes* est l'agent causal de la listériose, une infection grave d'origine alimentaire affectant essentiellement les personnes immunodéprimées, les personnes âgées et les femmes enceintes. Bien que la fréquence des cas de listériose ait diminué grâce aux efforts des industriels de la filière agroalimentaire et la mise en place d'un système de surveillance efficace, le risque sanitaire reste important avec un taux de mortalité d'environ 30%. Les gènes de virulence connus à ce jour sont présents dans les 13 sérovars de l'espèce. Cependant, 95% des cas de listériose humaine ne sont dus qu'aux sérovars 1/2a, 1/2b et 4b. Curieusement, alors que le sérovar 4b est prépondérant dans les cas épidémiques de listériose, il est beaucoup moins fréquemment isolé que les autres sérovars dans les aliments ou dans l'environnement industriel. Un projet en cours, réunissant plusieurs laboratoires et des approches complémentaires de génomique, post-génomique et divers tests de virulence, vise à mieux comprendre ce qui conditionne l'expression du pouvoir pathogène. Dans le cadre de ce projet de recherche, nous caractérisons les protéomes de 12 souches de différents sérovars et de différentes origines (épidémiques, portage asymptomatique et environnement). Les protéomes des protéines intracellulaires et des protéines sécrétées, sont séparés par électrophorèses bidimensionnelles puis comparés avec l'aide d'un logiciel d'analyse d'image. Les spots protéiques d'intérêt sont identifiés par spectrométrie de masse MALDI-TOF. Des différences de profil protéique entre les souches ont été mises en évidence. Ces différences sont corrélées avec l'évolution des sous-groupes au sein de l'espèce *L. monocytogenes* puisque les souches de sérovar 4b et 1/2b clustérisent indépendamment des souches 1/2a. Ainsi, des marqueurs discriminant des sérovars ont pu être sélectionnés. Par contre, le profil protéique des souches est relativement peu influencé par leur niveau de virulence et très peu par leur origine. A terme, ces connaissances seront utilisées pour une meilleure hiérarchisation et gestion du risque sanitaire auquel peuvent être exposés les consommateurs.

## *Comparison of two subproteomes from 12 strains of the foodborne pathogen Listeria monocytogenes: research of risk markers*

Dumas E.<sup>1</sup>, Meunier B.<sup>2</sup>, Berdagué J.L.<sup>3</sup>, Chafsey I.<sup>1</sup>, Chambon C.<sup>4</sup> and Hébraud M.<sup>1,4</sup>

UR Microbiologie - Equipe QuaSA<sup>1</sup>, URH - Equipe C2M<sup>2</sup>, UR QuaPA – Equipe T2A<sup>3</sup> et Plate-Forme Protéomique<sup>4</sup>, INRA site de Theix, 63122 Saint-Genès Champanelle

*Listeria monocytogenes* is a foodborne pathogen widespread in the environment and responsible for serious invasive and potentially fatal human and animal infections. This bacterium is the etiologic agent of listeriosis, which mainly affects pregnant women, neonates, the elderly and individuals with weakened immune systems due to medications or a disease. Although the incidence of *L. monocytogenes* decreased this last decade thanks to the installation of systematic and rigorous food control systems (HACCP) by the food manufacturers, it remains an important problem of public health with a mortality rate of 20 to 30%. The known virulence genes are present in the 13 serovars of the species. However, 96% of the human cases of listeriosis are only due to serovars 1/2a, 1/2b and 4b. While the serovar 4b is predominant in the epidemic cases of listeriosis, most of the strains isolated from food products belong to serovar 1/2a and 1/2b.

A research program joining together several laboratories and complementary approaches of genomic, post-genomic and virulence tests, aims to identify molecular markers of risks and to help to evaluate the sanitary risk to which the consumers are exposed with respect to contaminated foods.

In this project, we were in charge of the comparison and characterization of two subproteomes of 12 different *L. monocytogenes* strains. These strains came from various origins (epidemic, asymptomatic carriage and food plant environments). The three serovars mainly involved in human cases of listeriosis and three levels of virulence were represented among the 12 strains. The intracellular and secreted proteins were extracted and separated by two-dimensional gel electrophoresis. The comparison of the protein patterns (5 gels per strain and per subproteome) with the image analysis software Image Master 2D Platinum allowed to highlight significant qualitative and quantitative differences. The hierarchical clustering, based on data set of 219 and 159 selected protein spots for intracellular proteome and secretome, respectively, gave a very clear repartition of the 12 strains into two clusters in correlation with the evolution of the species. The estimated relative variance (ERV) of each spot was calculated for the three factors: serovar, virulence and origin. Protein spots with the highest ERV for each factor were used for factorial discriminant analyses and allowed to obtain very good discriminations for the three serovars and the three levels of virulence. These protein spots, considered as potential markers of interest, were firstly identified through their peptide mass fingerprint by using MALDI-TOF mass spectrometry. The identified proteins are currently analysed for their biological significance and for their potential validation (or not) as markers with respect to the factor considered.

The genomic, transcriptomic, proteomic and virulence results obtained by the different partners of this research program will be gathered in a data base and should contribute to a better risk assessment of *L. monocytogenes* in food products and food industries.

## Caractérisation du mutant ferritine ( $\Delta fri$ ) de *Listeria monocytogenes* EGDe

Dumas Emilie<sup>1</sup>, Dussurget Olivier<sup>3</sup>, Archambaud Cristel<sup>3</sup>, Chafsey Ingrid<sup>1</sup>, Chambon Christophe<sup>2</sup>, Cossart Pascale<sup>3</sup> et Hébraud Michel<sup>1,2</sup>

SRV-Microbiologie<sup>1</sup>, Plate-forme protéomique<sup>2</sup>, INRA Centre de Theix, 63122 Saint-Genès Champanelle, Unité des Interactions Bactéries-Cellules<sup>3</sup>, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15

### INTRODUCTION

*Listeria monocytogenes* est l'agent causal de la listériose, une infection grave d'origine alimentaire, mortelle dans 30% des cas, affectant essentiellement les personnes immunodéprimées, les personnes âgées et les femmes enceintes. Le fer est un élément essentiel pour son développement, mais il peut catalyser la formation de radicaux hydroxyles toxiques. Pour prévenir ce problème, les bactéries synthétisent des protéines de stockage du fer, les ferritines. L'expression du gène de la ferritine *fri* chez *Listeria* augmente lors de carences en fer, ce qui est inhabituel car chez d'autres microorganismes, la ferritine augmente dans des conditions riches en fer. De plus cette protéine est surexprimée après des chocs thermiques ou chimiques (SDS, désoxycholate, éthanol). La ferritine semble jouer plusieurs rôles dans la physiologie de la bactérie qui ne se limitent pas au stockage du fer. Afin de mieux comprendre les différentes fonctions de cette protéine, la souche EGDe de *L. monocytogenes* a été déléetée du gène codant la ferritine ( $\Delta fri$ ) et son protéome, son phénotype et sa virulence ont été étudiés.

### MATERIEL ET METHODES

**Souches** : *L. monocytogenes* EGDe (souche clinique de sérotype 1/2a) et son mutant  $\Delta fri$  déléeté du gène codant la ferritine

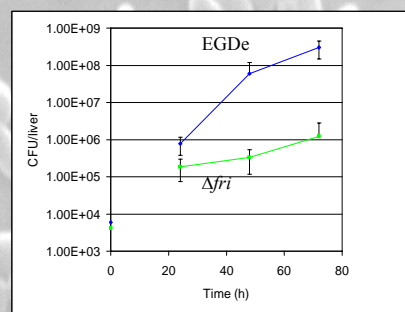
**Protéome** : Les bactéries ont été cultivées en milieu BHI à 37 °C, jusqu'en phase stationnaire. Les protéines ont été extraites et séparées par électrophorèses bidimensionnelles (IPG 3-10NL et 3-6). Les protéines d'intérêt ont été identifiées par spectrométrie de masse MALDI-TOF.

**Croissance** : Une pré-culture a été réalisée en milieu minimum (IMM) contenant 20 µg/ml de fer, puis les bactéries ont été cultivées en milieu IMM avec différentes concentration en fer.

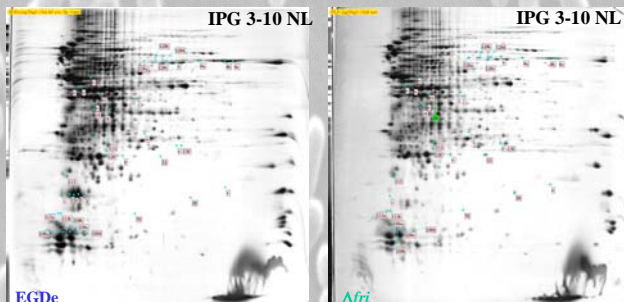
**Stress** : Les bactéries ont été cultivées à 37°C en milieu BHI. Les stress thermiques ont été appliqués en phase exponentielle en milieu liquide. La sensibilité au stress oxydatif a été évaluée par mesure des diamètres d'inhibition de croissance en utilisant la méthode des disques contenant l'agent oxydant sur milieu gélosé BHI.

**Virulence** : Des dilutions en série des bactéries (DL50) ou des doses sous-létales de bactéries (survie dans les organes) ont été injectées par voie intraveineuse à des souris BALB/c femelles de 6 semaines. La DL50 a été déterminée par la méthode probit. La survie dans les organes a été déterminée en étalant des dilutions en série des organes préalablement homogénéisés sur milieu gélosé BHI.

### Survie de *L. monocytogenes* EGDe et de $\Delta fri$ dans le foie chez la souris BALB/c

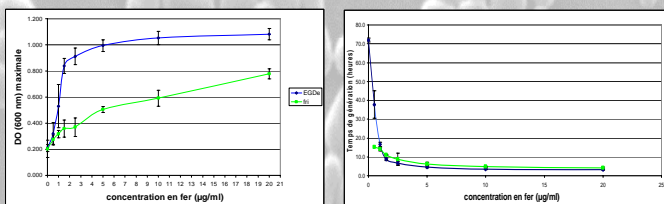


### Comparaison des protéomes de *L. monocytogenes* EGDe et de $\Delta fri$



numéro	catégorie fonctionnelle	gène	description
3	Adaptation to atypical conditions	frj	non-heme iron-binding ferritin
4a	Miscellaneous	frj	esterolysin O precursor
35	Regulation	ospA	catabolite control protein A
109	Similar to unknown proteins from other organisms	lmo2199	unknown protein
14	Metabolism of cofactors and prosthetic groups	lmo2101	similar to a protein required for pyridoxine synthesis
38	Transport / binding proteins and lipoproteins	lmo0783	similar to mannose-specific phosphotransferase system (PTS) component IIB
53	Similar to unknown proteins from other organisms	lmo1830	similar to conserved hypothetical proteins
70	Similar to unknown proteins from other organisms	lmo1468	similar to unknown proteins
112	Similar to unknown proteins from other organisms	lmo0794	similar to Bacillus subtilis YwB protein
113	Similar to unknown proteins from other organisms	lmo0798	conserved hypothetical protein
114	Adaptation to atypical conditions	RsbW	sigma B activity regulator RsbW
115	Similar to unknown proteins from other organisms	lmo2748	similar to Bacillus subtilis stress protein YdaG
128	Metabolism of nucleotides and nucleic acids	dra	similar to deoxyribose-phosphate aldolase
129	Metabolism of carbohydrates and related molecules (specific pathways)	lmo0913	similar to succinate semialdehyde dehydrogenase
130	Similar to unknown proteins from other organisms	lmo2391	conserved hypothetical protein similar to B.subtilis YHK protein

### DO<sub>600</sub> maximale et temps de génération de EGDe et de $\Delta fri$ en fonction de la concentration en fer



### Effet de différents stress sur $\Delta fri$ par rapport à EGDe

**Stress thermiques** : Après un choc thermique provoqué par le passage de 37°C à 4°C ou de 37°C à 45°C, le mutant présente une croissance similaire à la souche sauvage.

**Stress oxydatif** : Le mutant est plus sensible que la souche sauvage à la plumbagine, un réactif générant des anions radicalaires superoxydes.

### CONCLUSION

Notre étude montre que le mutant  $\Delta fri$  est atténué par rapport à la virulence de la souche sauvage EGDe chez la souris infectée par voie intraveineuse. Les résultats obtenus dans le foie sont similaires pour la rate et le cerveau. La DL50 du mutant  $\Delta fri$ , augmentée d'un facteur 2, confirme le rôle de Fri dans la virulence. La comparaison du protéome du mutant avec celui de la souche sauvage a permis d'observer 42 spots protéiques différenciellement exprimés. Parmi ceux-ci, un spot de la listériolysine O, facteur de virulence majeur, est sous-exprimé par le mutant, et la protéine RsbW (régulateur négatif de sigma B) est sur-exprimée. Des études phénotypiques ont montré que le mutant avait une croissance plus faible que la souche sauvage pour des concentrations en fer comprises entre 1 et 20 µg/ml. Pour des concentrations inférieures, le temps de génération augmentait fortement chez EGDe et de façon beaucoup moins prononcée chez le mutant. Bien que la ferritine soit surexprimée lors de chocs thermiques, ceux-ci n'ont pas d'effet significatif sur la croissance du mutant. Par contre, la ferritine semble impliquée dans la résistance au stress oxydatif. L'ensemble de ces résultats montre qu'en plus du stockage du fer, la ferritine de *L. monocytogenes* joue un rôle important dans l'adaptation à certains stress et dans la survie chez l'hôte.



## Characterization of the ferritin mutant of *Listeria monocytogenes* EGDe

Dumas Emilie<sup>1</sup>, Dussurget Olivier<sup>2</sup>, Archambaud Cristel<sup>3</sup>, Chafsey Ingrid<sup>1</sup>, Chambon Christophe<sup>2</sup>, Cossart Pascale<sup>3</sup> et Hébraud Michel<sup>1,2</sup>

SRV-Microbiologie<sup>1</sup>, Plate-forme protéomique<sup>2</sup>, INRA Centre de Theix, 63122 Saint-Genès Champanelle, Unité des Interactions Bactéries-Cellules<sup>3</sup>, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15

### INTRODUCTION

*Listeria monocytogenes* is a Gram-positive pathogen widespread in the environment and in foods. It is the causative agent of listeriosis, a severe foodborne disease, which mainly affects elderly, pregnant women, neonates and immuno-compromised persons. This bacterium is able to survive hostile environments and stress conditions. When it is submitted to heat or cold thermal shocks or to some chemical stresses (SDS, desoxycholate, éthanol), a protein involved in the iron storage, named ferritin, is strongly overexpressed. So, the ferritin seems to have several functions which are not limited to iron storage. In order to better understand the role of this protein, we characterised a deletion mutant ( $\Delta fri$ ) of *Listeria monocytogenes* EGDe by a proteomic approach and some growth and virulence tests.

### MATERIAL AND METHODS

**Strains** : *L. monocytogenes* EGDe (serotype 1/2a) and its mutant obtained by deletion of the ferritin gene ( $\Delta fri$ ).

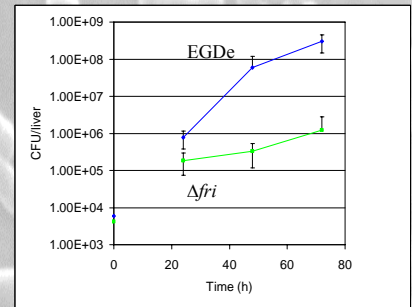
**Proteome** : Bacterial strains were grown in BHI at 37 °C, until the stationary phase. The soluble proteins were extracted and separated by two-dimensional electrophoresis (IPG 3-10NL and 3-6). The proteins whose expression varied between the two strains were identified by MALDI-TOF mass spectrometry.

**Growth** : Bacterial strains were precultured in a minimal medium (IMM) with 20 µg/ml of iron, then cultured in IMM supplemented with different iron concentrations.

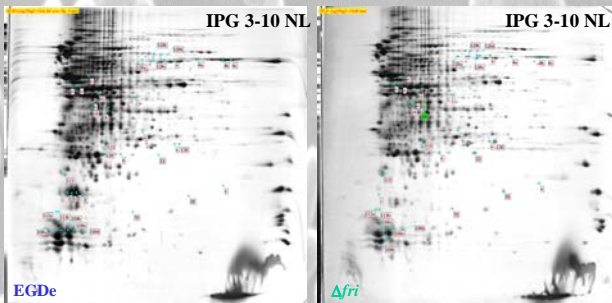
**Stress** : Bacterial strains were grown in BHI at 37 °C. Thermal shocks were applied in mid-log phase. Sensitivity to oxidative stress was measured by disk diffusion susceptibility assay using plumbagin as superoxide generator.

**Virulence** : Serial dilutions of bacteria (LD50) or 10<sup>4</sup> bacteria (survival in the liver) were injected intravenously in 6-week old female BALB/c mice. The LD50 was determined by the probit method. Survival in the liver was determined by spreading serial dilutions of the homogenized liver on BHI agar.

### Survival of *L. monocytogenes* EGDe and $\Delta fri$ in liver of BALB/c mice

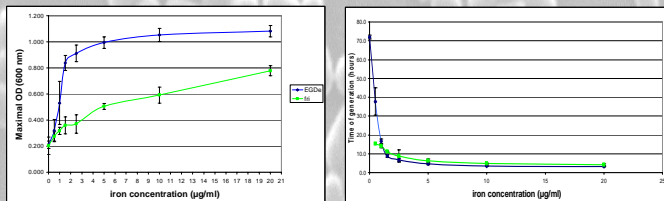


### Proteome comparison of *L. monocytogenes* $\Delta fri$ and EGDe



number	functional category	gene	description
3	Adaptation to atypical conditions	<i>fri</i>	non-heme iron-binding ferritin
4a	Miscellaneous	<i>chw</i>	listeriolysin O precursor
85	Regulation	<i>ccpA</i>	catabolite control protein A
109	Similar to unknown proteins from others organisms	<i>lmo2199</i>	unknown protein
14	Metabolism of coenzymes and prosthetic groups	<i>lmo2101</i>	similar to a protein required for pyridoxine synthesis
38	Transport / binding proteins and lipoproteins	<i>lmo0783</i>	similar to mannose-specific phosphotransferase system (PTS) component IIB
53	Similar to unknown proteins from others organisms	<i>lmo1930</i>	similar to conserved hypothetical proteins
70	Similar to unknown proteins from others organisms	<i>lmo1468</i>	similar to unknown proteins
112	Similar to unknown proteins from others organisms	<i>lmo0794</i>	similar to Bacillus subtilis YwB protein
113	Similar to unknown proteins from others organisms	<i>lmo0796</i>	conserved hypothetical protein
114	Adaptation to atypical conditions	<i>RsbW</i>	sigma B activity negative regulator RsbW
115	Similar to unknown proteins from others organisms	<i>lmo2748</i>	similar to Bacillus subtilis stress protein YdaG
128	Metabolism of nucleotides and nucleic acids	<i>dra</i>	similar to deoxyribose-phosphate aldolase
129	Metabolism of carbohydrates and related molecules (specific pathways)	<i>lmo0913</i>	similar to succinate semialdehyde dehydrogenase
130	Similar to unknown proteins from others organisms	<i>lmo2391</i>	conserved hypothetical protein similar to B. subtilis YHK protein

### Maximal OD<sub>600</sub> and time of generation of EGDe and $\Delta fri$ according to iron concentrations



### Effect of different stresses on $\Delta fri$ and EGDe

**Thermal stress** : Growth of wild-type and  $\Delta fri$  strains is similar after thermal shocks from 37°C to 4°C or 37°C to 45°C.

**Oxidative stress** : The  $\Delta fri$  mutant is more sensitive to the superoxide generator plumbagin.

### CONCLUSION

This study shows that the virulence of the  $\Delta fri$  mutant is attenuated compared to that of the wild strain in mice infected intravenously. Indeed, the LD50 of  $\Delta fri$  was twice higher, and survival of the  $\Delta fri$  strain in the liver was decreased by more than 2 logs at 72 h postinfection. The soluble proteins of the mutant and the wild strains were compared by a proteomic approach. This allowed the identification of several underexpressed proteins in the mutant, such as CcpA (Catabolite Control Protein A) and one spot of listeriolysin O, the major factor of virulence, and other overexpressed proteins involved in metabolism or transport (PTS). The growth curves in the presence of different iron concentrations showed that  $\Delta fri$  was characterized by lower OD<sub>max</sub>, prolonged lag phase but similar generation times than the wild strain. It was surprising to note that thermal shocks had no significant effect on the growth of the  $\Delta fri$  mutant while the ferritin was identified as a thermal shock inducible protein. On the other hand, ferritin seems to be involved in the resistance to oxidative stress. These results confirm that the ferritin is directly or indirectly involved in several functions, other than iron storage, such as adaptation to some stresses and survival in the host.

## BASE DE DONNEES DE GELS D'ELECTROPHORESE BIDIMENSIONNELLE DU PROTEOME MEMBRANAIRE ET DU SECRETOME DE *Listeria monocytogenes* EGDe

Emilie DUMAS<sup>1</sup>, Ingrid CHAFSEY<sup>1</sup>, Stella PLANCHON<sup>1</sup>, Christophe CHAMBON<sup>2</sup> et Michel HEBRAUD<sup>1,2</sup>  
 QuaPA-Microbiologie<sup>1</sup>, Plate forme protéomique<sup>2</sup>, INRA site de Theix, 63122 Saint-Genès Champanelle,  
 Tel : +33 (0)473.62.46.70 ; mél : hebraud@clermont.inra.fr

### INTRODUCTION

*Listeria monocytogenes* est une bactérie pathogène à Gram positif largement répandue dans l'environnement. Elle représente l'un des principaux risques sanitaires pour la filière agroalimentaire par sa capacité à coloniser les matières premières, les matériels et les surfaces des ateliers et son aptitude à survivre dans des environnements hostiles. Elle est à l'origine de la listériose, une infection grave (30 % de mortalité) qui affecte particulièrement les femmes enceintes, les nourrissons, les personnes âgées et/ou immunodéprimées.

L'objectif de notre travail est de construire une base de données avec des cartes de référence des différents sous-protéomes comprenant de nombreux spots protéiques identifiés. Cette base de données constitue une précieuse source d'informations pour toutes les analyses protéomiques effectuées sur *Listeria*.

### MATERIEL ET METHODES

**Souche** : *Listeria monocytogenes* EGDe, sérovar 1/2a.

**Conditions de culture** : milieu MCDB202 supplémenté, 37°C, 150 rpm. Les cellules sont récupérées à DO<sub>600</sub> = 0,9.

**Fractionnement** : • séparation des fractions extracellulaires et cellulaires par centrifugation.  
 • les cellules sont cassées aux ultrasons et centrifugées. Le surnageant est ultracentrifugé afin de séparer les fractions membranaires et cytosolubles.

**Extractions** : • les protéines de surface sont précipitées lors d'une délipidation de la fraction membranaire. Les protéines sont récupérées par centrifugation puis mises en solution dans un tampon d'IEF.  
 • les protéines sécrétées sont précipitées au TCA, le culot est lavé à l'acétone avant de remettre les protéines en solution dans un tampon d'IEF.

**Identification** : les protéines sont identifiées grâce à leur empreinte peptidique massique obtenue par spectrométrie de masse MALDI-TOF

### PROTEINES SECRETEES

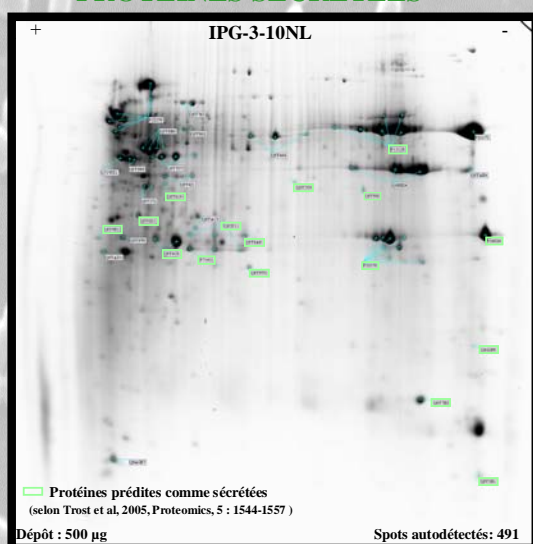


FIGURE 1: EBD des protéines sécrétées colorées au BC colloïdal.

Functional classification	Number of genes	Number of identified proteins	Number of spots identified	Predicted secretory proteins (Trust and al.)
Cell envelope and cellular processes	620	14	27	6
Intermediary metabolism	611	9	22	4
Information pathways	452	2	2	1
Others functions	149	3	10	1
Similar to unknown proteins	746	6	9	3
No similarity	260	2	2	1
TOTAL	2838	36	72	16

TABLEAU 1 : répartition des protéines identifiées dans le sécrétome selon leur catégorie fonctionnelle.

### PROTEINES MEMBRANAIRES

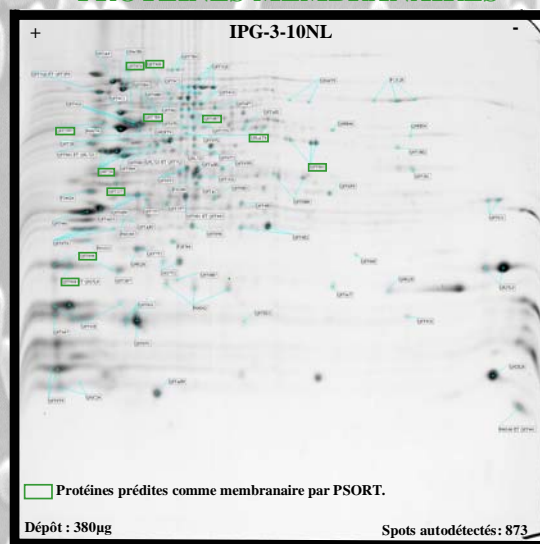


FIGURE 2 : EBD des protéines membranaires colorées au BC colloïdal.

Functional classification	Number of genes	Number of identified proteins	Number of spots identified	Predicted membrane proteins (PSORT)
Cell envelope and cellular processes	620	17	29	3
Intermediary metabolism	611	34	50	5
Information pathways	452	29	47	2
Others functions	149	10	14	0
Similar to unknown proteins	746	2	3	0
No similarity	260	0	0	0
TOTAL	2838	92	143	10

TABLEAU 2 : répartition des protéines identifiées dans la fraction membranaire selon leur catégorie fonctionnelle.

### CONCLUSION

Actuellement, 72 et 143 spots protéiques correspondant à 36 et 92 protéines ont été identifiées sur les cartes de référence des protéines sécrétées et des protéines membranaires, respectivement. D'autres identifications en cours viendront compléter cette base de données. La localisation de nombreuses protéines ne correspond pas aux prédictions telles qu'elles peuvent être faites grâce aux différents outils bioinformatiques. Plusieurs facteurs peuvent expliquer la présence non prédite de ces protéines : d'une part la prédiction par les logiciels n'est pas entièrement fiable ; d'autre part les techniques de fractionnement cellulaire ne permettent pas d'éviter les contaminations entre les sous-protéomes ; enfin, la localisation d'une protéine peut varier au cours de sa vie et selon sa (ses) fonction(s). La constitution de profils protéiques des sous-protéomes cellulaires fournissent de nouvelles informations qui apportent un éclairage complémentaire permettant de mieux appréhender la physiologie bactérienne. Les cartes de référence de ces deux sous-protéomes viendront compléter celles déjà réalisées en différents gradients de pH sur les protéines cytosoliques et qui sont disponibles sur le site web : [www.clermont.inra.fr/teproteome](http://www.clermont.inra.fr/teproteome).



## APPROCHE PROTEOMIQUE POUR LA CARACTERISATION DE LA BIODIVERSITE DE *Listeria monocytogenes* : RECHERCHE DE MARQUEURS DE RISQUE

Emilie DUMAS<sup>1</sup>, Christophe CHAMBON<sup>2</sup> et Michel HEBRAUD<sup>1,2</sup>

QUAPA -Microbiologie<sup>1</sup>, Plate forme protéomique<sup>2</sup>, INRA site de Theix, 63122 Saint-Genès Champanelle, Tel : +33 (0)473.62.46.70 ; mél : hebraud@clermont.inra.fr

### INTRODUCTION

*Listeria monocytogenes* est l'agent causal de la listériose, une infection grave d'origine alimentaire affectant essentiellement les personnes immunodéprimées, les personnes âgées et les femmes enceintes. Bien que la fréquence des cas de listériose ait diminué ces dernières années, le risque sanitaire reste important avec un taux de mortalité d'environ 30%. Les gènes de virulence connus à ce jour sont présents dans les 13 sérovars de l'espèce. Cependant, 95% des cas de listériose humaine ne sont dus qu'aux sérovars 1/2a, 1/2b et 4b. Curieusement, alors que le sérovar 4b est prépondérant dans les cas épidémiques de listériose, il est beaucoup moins fréquemment isolé que les autres sérovars dans les aliments ou dans l'environnement industriel. Un projet de recherche en cours, réunissant plusieurs laboratoires et des approches complémentaires de génomique, post-génomique et divers tests de virulence, vise à mieux comprendre ce qui conditionne l'expression du pouvoir pathogène. Dans le cadre de ce projet, nous caractérisons et comparons les sous-protéomes (cytosol et sécrétome) de souches de différents sérovars et différentes origines (épidémiques, portage asymptomatique et environnement) par électrophorèse bidimensionnelle (EBD). A terme, cette étude devrait permettre d'identifier de nouveaux marqueurs discriminants en fonction de la virulence, de l'origine ou du sérovar des souches.

### SOUCHES ETUDIEES

Origine	Nom de la souche	sérovar
Cas clinique de listériose humaine	CLIP80459	4b
	CLIP90602	1/2b
Listériose animale	EGDe	1/2a
Isolées de l'environnement agroalimentaire	CLIP93665	4b
	CLIP93649	1/2a
Portage asymptomatique	CLIP93667	4b

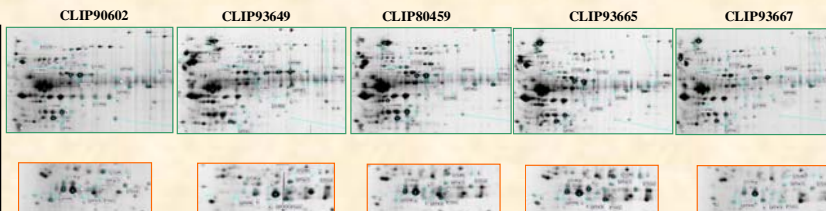
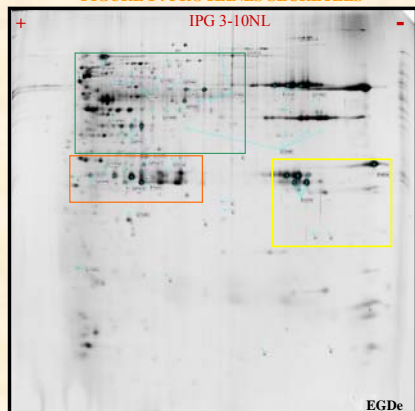
### MATERIEL ET METHODES

**Extraction des protéines :** Les souches de *L. monocytogenes* sont cultivées en milieu MCDB202, à 37°C (150 rpm) jusqu'à DO<sub>600</sub> = 0.9 (fin de phase exponentielle). Le culot cellulaire est séparé du surnageant par centrifugation. Les protéines sécrétées sont précipitées au TCA à partir du surnageant. Les cellules sont cassées aux ultrasons afin d'extraire les protéines cytosolubles. Les protéines cytosolubles et sécrétées sont ensuite remises en solution dans un tampon d'IEF.

**Électrophorèse bidimensionnelle :** les protéines sont séparées par EBD en utilisant des IPG strips pH 3-10NL (sécrétome) ou pH 4-7 (protéines cytosolubles) en première dimension et un gel d'acrylamide à 12.5% pour la deuxième dimension. Les protéines sont colorées au nitrate d'argent et les gels sont comparés à l'aide du logiciel Image Master 2D Platinum.

**Identification des protéines :** Les protéines d'intérêt sont identifiées grâce à leur empreinte peptidique massique obtenue par spectrométrie de masse MALDI-TOF.

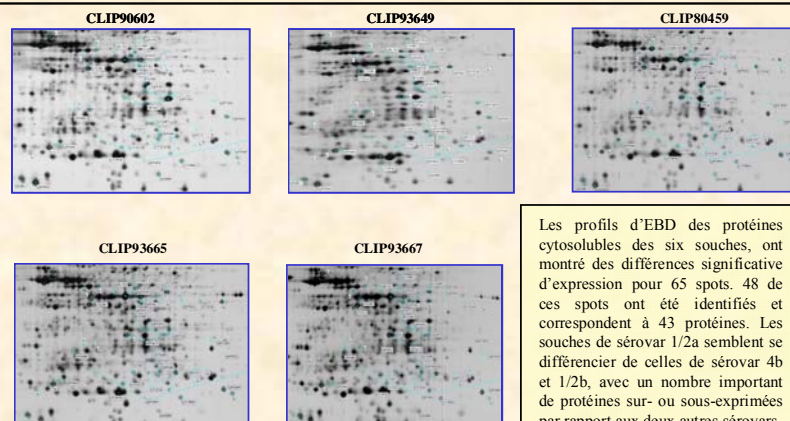
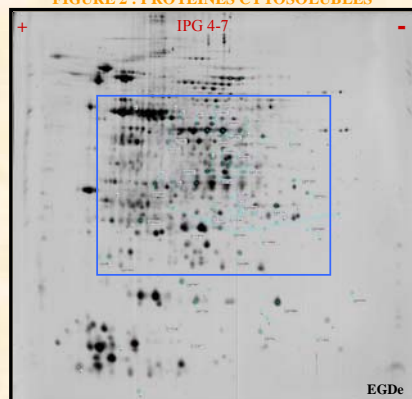
FIGURE 1 : PROTEINES SECRETEES



L'étude comparative du sécrétome de ces 6 premières souches montre qu'il existe une diversité importante de profil protéique. En effet, 64 spots protéiques présentent une différence significative d'expression entre les souches. 50 de ces spots ont été identifiés et correspondent à 30 protéines différentes. Au vu des premiers résultats, il apparaît que certaines protéines sont spécifiques ou ont d'importantes variations d'expression en fonction du sérovar. Par exemple, la protéine Q8Y4C8 (hypothetical protein lmo2522) est plus abondante dans les souches de sérovar 1/2a, tandis que 2 spots de la protéine Q71YE5 (sérine protéase) ne sont présents que chez les souches de sérovar 4b. Par contre, aucun marqueur spécifique de l'origine n'a encore pu être mis en évidence.

Figures 1 et 2 : EBD des sous-protéomes de EGDe et zooms de certaines zones chez les 5 autres souches. Le numéro d'accèsion Swiss-Prot des spots protéiques présentant une différence d'expression est indiqué (protéine identifiée chez l'une des 5 souches). x = protéine d'intérêt non identifiée.

FIGURE 2 : PROTEINES CYTOSOLUBLES



Les profils d'EBD des protéines cytosolubles des six souches, ont montré des différences significatives d'expression pour 65 spots. 48 de ces spots ont été identifiés et correspondent à 43 protéines. Les souches de sérovar 1/2a semblent se différencier de celles de sérovar 4b et 1/2b, avec un nombre important de protéines sur- ou sous-exprimées par rapport aux deux autres sérovars.

### CONCLUSION

A ce jour, l'analyse des 6 premières souches a permis de mettre en évidence des différences significatives dans les profils protéiques. La majorité des protéines différenciellement exprimées ont été identifiées. Les premières analyses montrent qu'il semble possible de discriminer les souches de sérovar 4b et 1/2b de celles de sérovar 1/2a. L'étude de 6 autres souches est en cours et viendra compléter ces premiers résultats. Des outils statistiques seront alors mis en œuvre sur l'ensemble des données issues de l'analyse des 12 souches afin d'aider à identifier et valider des marqueurs discriminants en fonction de la virulence, de l'origine et du sérovar. Les résultats de l'analyse protéomique seront confrontés à ceux des autres approches (génomique, transcriptomique) pour évaluer la pertinence des marqueurs potentiels. Une base de données, constituée à partir de l'ensemble des résultats issus de ces différentes approches, pourra contribuer à mieux comprendre l'écologie de cette bactérie et son processus infectieux chez l'homme. Ces connaissances seront également utilisées pour une meilleure gestion du risque et une amélioration de la sécurité microbiologique des aliments.



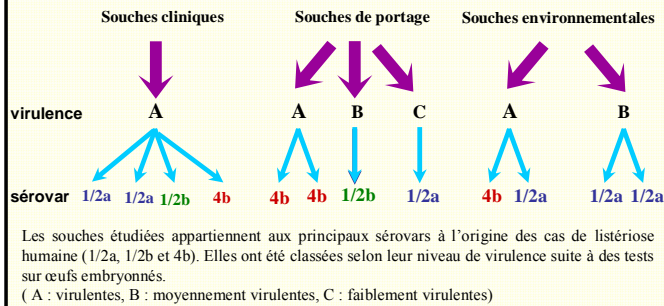
## APPROCHE PROTEOMIQUE POUR LA CARACTERISATION DES SOUCHES DE *LISTERIA MONOCYTOGENES* : RECHERCHE DE MARQUEURS DE RISQUE

Dumas E.<sup>1</sup>, Meunier B.<sup>2</sup>, Berdagué J.L.<sup>3</sup>, Chambon C.<sup>4</sup>, Chafsey I.<sup>1</sup> et Hébraud M.<sup>1,4</sup>  
 UR Microbiologie - Equipe QuaSA<sup>1</sup>, URH - Equipe C2M<sup>2</sup>, UR QuaPA – Equipe T2A<sup>3</sup> et Plate-Forme Protéomique<sup>4</sup>,  
 INRA site de Theix, 63122 Saint-Genès Champanelle

### INTRODUCTION

*Listeria monocytogenes* est l'agent causal de la listériose, une infection grave d'origine alimentaire affectant essentiellement les personnes immunodéprimées, les personnes âgées et les femmes enceintes. Bien que la fréquence des cas de listériose ait diminué ces dernières années, le risque sanitaire reste important avec un taux de mortalité d'environ 30%. Les gènes de virulence connus à ce jour sont présents dans les 13 sérovars de l'espèce. Cependant, 96% des cas de listériose humaine ne sont dus qu'aux sérovars 1/2a, 1/2b et 4b. Curieusement, alors que le sérovar 4b est prépondérant dans les cas épidémiques de listériose, il est beaucoup moins fréquemment isolé que les autres sérovars dans les aliments ou dans l'environnement industriel. Un projet de recherche en cours, réunissant plusieurs laboratoires et des approches complémentaires de génomique, post-génomique et divers tests de virulence, vise à mieux comprendre ce qui conditionne l'expression du pouvoir pathogène. Dans le cadre de ce projet, nous caractérisons et comparons les sous-protéomes (cytosol et sécrétome) de souches de différents sérovars et différentes origines (épidémiques, portage asymptomatique et environnement) par électrophorèse bidimensionnelle (EBD). L'objectif de cette étude est d'identifier de nouveaux marqueurs discriminants en fonction de la virulence, du sérovar ou de l'origine des souches.

### SOUCHES ETUDIÉES



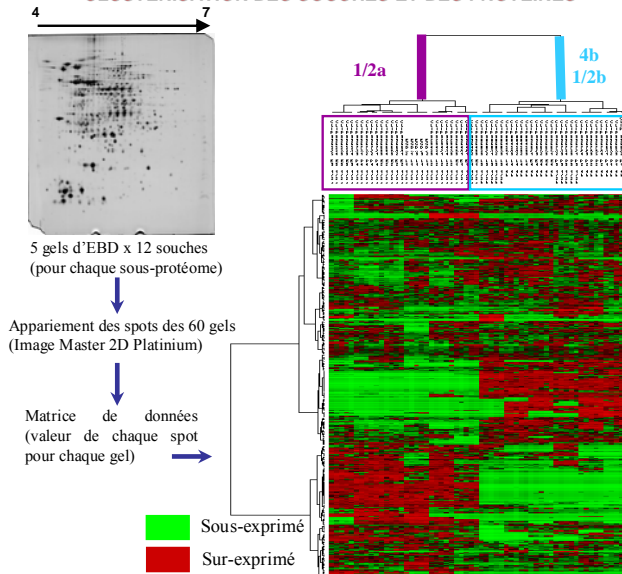
### METHODES

**Extraction des protéines** : Les souches de *L. monocytogenes* sont cultivées en milieu MCDDB202, à 37°C (150 rpm) jusqu'à DO<sub>600</sub> = 0.9 (fin de phase exponentielle). Le culot cellulaire est séparé du surnageant par centrifugation. Les protéines sécrétées sont précipitées au TCA à partir du surnageant. Les cellules sont cassées aux ultrasons afin d'en extraire les protéines cytosolubles. Les protéines cytosolubles et sécrétées sont ensuite remises en solution dans un tampon d'IEF.

**Électrophorèse bidimensionnelle** : les protéines sont séparées par EBD en première dimension et un gel d'acrylamide à 12.5% pour la deuxième dimension. Les protéines sont colorées au nitrate d'argent et les gels sont comparés à l'aide du logiciel Image Master 2D Platinum.

**Identification des protéines** : Les protéines d'intérêt sont identifiées grâce à leur empreinte peptidique massique obtenue par spectrométrie de masse MALDI-TOF.

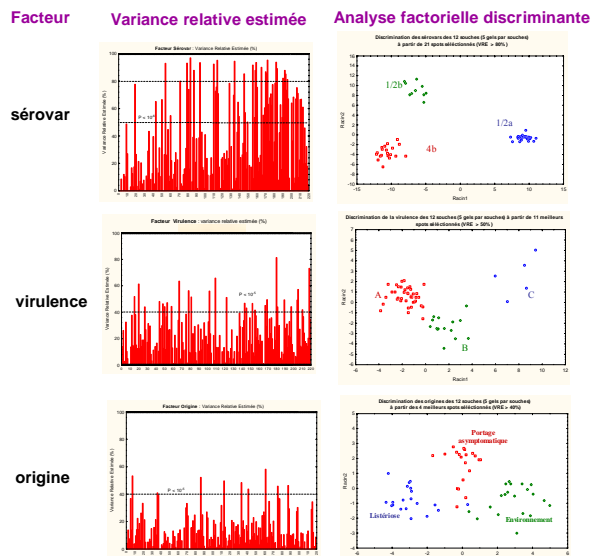
### CLUSTERISATION DES SOUCHES ET DES PROTEINES



Pour chaque souche, 5 gels d'électrophorèse bidimensionnelle par sous-protéome ont été sélectionnés. Les spots de chacun des gels ont été appariés avec ceux d'un gel de référence afin d'obtenir une matrice de données comprenant leur valeur pour chaque gel. La clusterisation hiérarchique des gels et des protéines a été réalisée grâce au logiciel Permut Matrix.

La clusterisation à partir des spots des 2 sous-protéomes montre que les différences observées entre les souches sont corrélées avec l'évolution des sous groupes au sein de l'espèce *L. monocytogenes* puisque les souches de serovar 4b et 1/2b (linéage I) clustrent indépendamment des souches 1/2a (linéage II). On peut observer qu'il existe des clusters de protéines spécifiques à chaque linéage.

### EFFET DU FACTEUR SEROVAR, VIRULENCE OU ORIGINE SUR LA VARIATION DES SPOTS PROTEIQUES



Afin de trouver des marqueurs discriminants du sérovar, de la virulence ou de l'origine, la variation relative estimée de chaque spot a été calculée pour ces différents facteurs dans les deux sous-protéomes. Une analyse factorielle discriminante a été réalisée à partir des spots protéomiques possédant les VRE les plus importantes. Seuls les résultats sur les protéines cytosolubles sont présentés ici.

Les 21 spots protéomiques ayant des VRE > 80% permettent de discriminer clairement les 3 sérovars étudiés. La discrimination des niveaux de virulence est moins pertinente mais quelques marqueurs potentiels apparaissent (11 spots avec VRE > 50%). Par contre, le facteur origine ne peut pas être discriminé par notre approche expérimentale.

### CONCLUSION

Les différences de profils protéomiques entre les souches permettent de discriminer clairement les 2 linéages et les 3 sérovars étudiés. L'AFD montre que cette discrimination reste excellente en ne conservant que les 5 meilleurs spots. Les protéines marqueurs du facteur sérovar ont été prélevées et identifiées par spectrométrie de masse MALDI-TOF. Le facteur virulence est moins bien discriminé et nécessiterait de tester plus de souches. Les marqueurs potentiels (VRE > 50%) sont en cours d'identification. Enfin, dans nos conditions expérimentales, il n'est pas possible d'identifier des marqueurs de l'origine des souches. L'analyse de l'ensemble des protéines marqueurs potentiels est en cours afin d'évaluer leur pertinence au niveau des gels d'EBD (isoformes, modifications post-traductionnelles), et des génomes (présence ou absence des gènes) et d'étudier leur signification biologique. Les résultats obtenus par les partenaires de ce programme de recherche seront réunis dans une base de données et devraient contribuer à une meilleure hiérarchisation et gestion du risque sanitaire auquel peut être exposé le consommateur.



---

## RÉSUMÉ

*Listeria monocytogenes* est l'agent étiologique de la listériose, une infection d'origine alimentaire peu fréquente mais avec un taux de mortalité de 25% chez l'homme. Nous avons d'abord caractérisé un mutant ferritine et montré l'importance du produit de ce gène dans la virulence, la résistance à des conditions de carence en fer et de stress thermique ou oxydatif. Nous nous sommes ensuite intéressés à la biodiversité de *L. monocytogenes*, par une approche protéomique, en étudiant 12 souches d'origines, de sérovars et de niveaux de virulence variés. Sur la base des profils des protéines intracellulaires et sécrétées, une classification hiérarchique a montré deux groupes distincts, l'un comprenant les souches de sérovar 1/2a, l'autre les souches de sérovar 4b et 1/2b. Des spots protéiques spécifiques de chaque souche et des différents sérovars ont été identifiés par spectrométrie de masse MALDI-TOF avec pour objectif de mieux hiérarchiser et gérer le risque dû à *L. monocytogenes*.

---

### ***Listeria monocytogenes*: characterization of a ferritin mutant and study of the biodiversity by a proteomic approach**

---

*Listeria monocytogenes* is the etiologic agent of listeriosis a foodborne infection rare but with a case fatality rate of 25% for humans. We have first characterized a ferritin mutant and shown the importance of this gene product in virulence, resistance to conditions of iron starvation and thermal or oxydative stresses. Then, we studied the biodiversity of *L. monocytogenes*, with a proteomic approach, by studying 12 strains of different origins, serovars and levels of virulence. Based on the intracellular and secreted protein patterns, a hierarchical clustering showed two distinct clusters, one including strains of serovar 1/2a, the other one strains of serovar 4b and 1/2b. Some protein spots either strain- or serovar-specifics were identified by MALDI-TOF mass spectrometry with the objective to better evaluate and manage the risk due to *L. monocytogenes*.

---

**Discipline** : Sciences des aliments - **Spécialisation doctorale** : Microbiologie.

---

**Mots-clés** : *Listeria monocytogenes* ; ferritine ; biodiversité ; sérovars 1/2a, 1/2b, 4b ; électrophorèse bidimensionnelle ; protéines intracellulaires, protéines extracellulaires, sécrétion.

---

Unité de Recherche 454 Microbiologie - Equipe Qualité et Sécurité des Aliments  
Institut National de la Recherche Agronomique - Clermont-Ferrand / Theix

---