



**HAL**  
open science

# Analyse protéomique et caractérisation de nouvelles protéines de paroi chez *Encephalitozoon cuniculi*, une microsporidie pathogène de l'homme

Damien Brosson

► **To cite this version:**

Damien Brosson. Analyse protéomique et caractérisation de nouvelles protéines de paroi chez *Encephalitozoon cuniculi*, une microsporidie pathogène de l'homme. Biologie cellulaire. Université Blaise Pascal - Clermont-Ferrand II, 2006. Français. NNT : 2006CLF21636 . tel-00688692

**HAL Id: tel-00688692**

**<https://theses.hal.science/tel-00688692v1>**

Submitted on 18 Apr 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 BLAISE PASCAL  
D.U : 1636

UNIVERSITE D'AUVERGNE

Ecole Doctorale  
des Sciences de la Vie et de la Santé  
N° d'ordre : 431

## **THESE**

pour l'obtention du grade de

### **DOCTEUR D'UNIVERSITE**

spécialité : **Microbiologie**

présentée et soutenue publiquement

par

**Damien BROSSON**

le 16 janvier 2006

ANALYSE PROTEOMIQUE ET CARACTERISATION DE NOUVELLES  
PROTEINES DE PAROI CHEZ *ENCEPHALITOZOOM CUNICULI* , UNE  
MICROSPORIDIE PATHOGENE DE L'HOMME

### **JURY**

Président : Pr. Gérard PRENSIER, Université Blaise Pascal, Clermont-Ferrand.  
Rapporteurs : Pr. Catherine BRAUN-BRETON, Université de Montpellier II.  
Dr. Gabriel PELTRE, Ecole Supérieure de Physique et Chimie Industrielles, Paris.  
Examineur : Dr. Catherine TEXIER, Université Blaise Pascal, Clermont-Ferrand.  
Directeur de thèse : Pr. Christian VIVARÈS, Université Blaise Pascal, Clermont-Ferrand.

*Laboratoire de biologie des protistes UMR CNRS 6023, équipe Parasitologie moléculaire et cellulaire,  
Université Blaise Pascal, Clermont-Ferrand.*

## 🔗 Résumé :

La microsporidie *Encephalitozoon cuniculi* est un parasite intracellulaire obligatoire capable d'infester un large spectre d'hôtes dont l'homme. Ce pathogène opportuniste est responsable de nombreuses infections principalement chez les individus immunodéprimés. La spore, forme de résistance et de dissémination, est protégée des stress environnementaux par une épaisse paroi composée d'une endospore chitineuse et d'une exospore glycoprotéique. Bien que constituant l'interface des relations hôte-pathogène, peu de données décrivent les protéines de cette structure. Le génome nucléaire de ce parasite a été séquencé et ces travaux se placent donc dans un contexte d'analyse post-génomique. Le premier volet de ces travaux a été de décrire le protéome exprimé dans les stades tardifs de développement d'*E. cuniculi*. Les protéines majeures de ces stades ont été extraites grâce à des traitements séquentiels et identifiées par spectrométrie de masse (MALDI-TOF-MS et LC-MS/MS) selon deux stratégies complémentaires : analyse des spots protéiques résolus par électrophorèse bidimensionnelle (2D-PAGE), ou analyse directe des extraits protéiques (approche type « Shotgun »). 177 protéines différentes ont ainsi pu être identifiées, permettant d'obtenir, pour la première fois chez une microsporidie, une vision globale de la physiologie de la spore. Le second objectif concerne la recherche de nouvelles protéines pariétales chez *E. cuniculi* et leur caractérisation au cours du cycle de développement. Une première exploitation des données du protéome a permis l'identification d'une chitine déacétylase putative localisée à l'interface membrane plasmique – endospore et probablement impliquée dans la formation de la paroi, de deux protéines localisées à l'endospore, ainsi que d'une protéine d'exospore. Nous avons également établi un crible bioinformatique pour sélectionner des protéines candidates à une localisation pariétale parmi les protéines à fonction inconnue du protéome de *E. cuniculi*, ces protéines devant être en partie spécifiques du parasite. Onze protéines candidates ont ainsi été déterminées et la validation de leur localisation est actuellement en cours. Le suivi de la localisation, au cours du cycle de développement du parasite, des protéines pariétales identifiées dans ce travail nous permet de proposer le premier modèle de morphogenèse de la paroi microsporidienne et d'en donner une vision dynamique qui n'était pas clairement envisagée jusqu'à présent.

## 🔗 Mots-clés :

Microsporidies, *Encephalitozoon cuniculi*

Protéomique, 2D-PAGE, Shotgun, spectrométrie de masse, physiologie cellulaire

Paroi, morphogenèse

UNIVERSITE BLAISE PASCAL

UNIVERSITE D'Auvergne

Ecole Doctorale  
des Sciences de la Vie et de la Santé

## THESE

pour l'obtention du grade de

### DOCTEUR D'UNIVERSITE

spécialité : **Microbiologie**

présentée et soutenue publiquement

par

**Damien BROSSON**

le 16 janvier 2006

ANALYSE PROTEOMIQUE ET CARACTERISATION DE NOUVELLES  
PROTEINES DE PAROI CHEZ *ENCEPHALITOZOON CUNICULI*,  
UNE MICROSPORIDIE PATHOGENE DE L'HOMME

### JURY

Président : Pr. Gérard PRENSIER, Université Blaise Pascal, Clermont-Ferrand.  
Rapporteurs : Pr. Catherine BRAUN-BRETON, Université de Montpellier II.  
Dr. Gabriel PELTRE, Ecole Supérieure de Physique et Chimie Industrielles, Paris.  
Examineur : Dr. Catherine TEXIER, Université Blaise Pascal, Clermont-Ferrand.  
Directeur de thèse : Pr. Christian VIVARÈS, Université Blaise Pascal, Clermont-Ferrand.



# REMERCIEMENTS

---

Ce travail a été réalisé au Laboratoire de Biologie des Protistes, UMR CNRS 6023, dirigé par Monsieur Christian Amblard. Je le remercie de m'avoir accueilli au sein de son unité.

Je remercie également Monsieur le Professeur Christian Vivarès, responsable de l'équipe Parasitologie Moléculaire et Cellulaire, de m'avoir proposé ce sujet de thèse et de l'avoir dirigé.

Je tiens particulièrement à adresser mes plus chaleureux remerciements à Catherine Texier pour avoir encadré cette thèse. Sa constante disponibilité, son soutien permanent et ses conseils m'ont guidé tout au long de ce travail et m'ont permis de le mener à bien.

Je remercie également Madame Catherine Braun-Breton et Monsieur Gabriel Peltre d'avoir accepté de juger ce travail.

Mes remerciements vont également à :

Gérard Prensier pour m'avoir fait profiter de son expertise en microscopie électronique, pour sa disponibilité et toute l'aide qu'il m'a apporté. Je le remercie également d'avoir accepté de présider ce jury.

Guy Méténier pour l'intérêt qu'il a témoigné à ce travail, pour ses discussions, son aide et ses conseils.

Frédéric Delbac et Emmanuel Cornillot qui ont toujours trouvé du temps pour discuter et pour me conseiller.

Enfin, je remercie tous les membres du Laboratoire de Biologie des Protistes, permanents ou de passage, qui ont contribué à ce travail et permis qu'il se déroule dans une ambiance chaleureuse et amicale.



# TABLE DES MATIERES

---

<b>REMERCIEMENTS</b>	<b>1</b>
<b>TABLE DES MATIERES</b>	<b>2</b>
<b>INTRODUCTION</b>	<b>5</b>
<b>ETUDE BIBLIOGRAPHIQUE</b>	<b>8</b>
<b>1. PRESENTATION GENERALE DES MICROSPORIDIES</b>	<b>9</b>
1.1. Un phylum fortement diversifié	9
1.2. Une phylogénie controversée : d'eucaryotes primitifs à champignons atypiques	10
1.3. Un large spectre d'hôtes et une pathogénicité variée	11
1.4. <i>Encephalitozoon cuniculi</i> , une microsporidie modèle	12
1.4.1. Structure de la spore	13
1.4.1.1. L'élément infectieux : le sporoplasme	13
1.4.1.2. L'appareil invasif	13
1.4.1.3. La paroi : protection de la spore et interface hôte-parasite	14
1.4.2. Cycle et stades de développement	14
1.4.3. L'invasion cellulaire : hypothèses pour un mécanisme original	15
1.4.4. Un génome eucaryote minimal	16
<b>2. LA PROTEOMIQUE : UN OUTIL D'ANALYSE POST-GENOMIQUE APPLICABLE AUX PARASITES UNICELLULAIRES</b>	<b>18</b>
2.1. La protéomique : à quelles fins ?	19
2.1.1. La protéomique descriptive	19
2.1.2. La protéomique différentielle	20
2.1.3. La protéomique fonctionnelle	20
2.1.4. La protéomique structurale	21
2.1.5. L'analyse des modifications post-traductionnelles	21
2.2. Principes de l'analyse protéomique basée sur l'électrophorèse bidimensionnelle	22
2.2.1. L'extraction protéique	22
2.2.2. De l'extrait protéique aux spots protéiques	22
2.2.2.1. Séparation sur gel 2D-PAGE des protéines extraites	22
2.2.2.2. Visualisation des spots protéiques	23
2.2.2.3. Traitement informatique des images 2D-PAGE	23





2.2.3.	Identification des protéines	24
<b>2.3.</b>	<b>La spectrométrie de masse : une technique d'identification de choix</b>	<b>24</b>
2.3.1.	Définition de la technique	24
2.3.2.	Principe de la spectrométrie de masse de type MS	25
2.3.3.	Principe de la spectrométrie de masse en tandem ou MS/MS	25
2.3.4.	Les sources d'ionisation	26
2.3.5.	Les analyseurs et spectromètres de masse	27
<b>2.4.</b>	<b>Améliorations techniques et haut débit</b>	<b>28</b>
2.4.1.	Amélioration de la protéomique basée sur la 2D-PAGE	28
2.4.2.	Vers une protéomique sans électrophorèse	30
2.4.3.	Vers de nouvelles technologies et de nouveaux concepts	31
<b>2.5.</b>	<b>Les levures, organismes modèles pour la protéomique des eucaryotes</b>	<b>32</b>
2.5.1.	Description et localisation des protéines	32
2.5.2.	Réseaux d'interactions et structures des protéines	33
2.5.3.	Analyses de pathogénicité	34
<b>2.6.</b>	<b>Analyse protéomique des parasites unicellulaires</b>	<b>35</b>
2.6.1.	Spécificités et difficultés de l'analyse protéomique chez les parasites	35
2.6.2.	Analyses descriptives du protéome total de parasites	36
2.6.3.	Analyses comparatives : dynamique des protéines au cours du cycle de vie du parasite	37
2.6.4.	Analyses descriptives de sous-protéomes : de la localisation à la fonction	38
2.6.5.	Analyses protéomiques et thérapeutiques	39
2.6.6.	L'analyse protéomique, un outil pour la recherche de cibles vaccinales	40
2.6.7.	Données de protéomique chez les microsporidies	40
<b>3.</b>	<b>LA PAROI DES CHAMPIGNONS UNICELLULAIRES</b>	<b>42</b>
<b>3.1.</b>	<b>Structure et mise en place</b>	<b>42</b>
3.1.1.	<i>Saccharomyces cerevisiae</i> , un organisme modèle pour la paroi des levures	42
3.1.1.1.	Composition, structure et mise en place de la paroi de la cellule végétative	42
3.1.1.1.1.	Le réseau de $\beta$ -glucane	42
3.1.1.1.2.	La chitine	43
3.1.1.1.3.	Protéines de paroi	44
3.1.1.1.4.	Dynamisme de la paroi	45
3.1.1.2.	Composition, structure et mise en place de la paroi de l'ascospore	46
3.1.2.	La paroi de <i>Candida albicans</i> , une structure impliquée dans la pathogénicité	48
3.1.3.	Les microsporidies	49
3.1.3.1.	La structure en deux feuillets de la paroi	49
3.1.3.1.1.	L'endospore	49
3.1.3.1.2.	L'exospore	50
3.1.3.1.3.	Les protéines pariétales	50
3.1.3.2.	Morphogénèse et mise en place	51
<b>3.2.</b>	<b>Rôle et fonction de la paroi microsporidienne</b>	<b>51</b>
3.2.1.	Maintien de la forme et adaptation au milieu extérieur	52
3.2.2.	Relation hôte-pathogène	52
3.2.2.1.	Adhésion / reconnaissance	52
3.2.2.2.	La paroi et la réponse immunitaire : intérêt pour le diagnostic	53



<b>RESULTATS</b>	<b>56</b>
<hr/>	
<b>4. ANALYSE DU PROTEOME DU PARASITE EUCARYOTE <i>ENCEPHALITOOZON CUNICULI</i>: UNE CARTE DE REFERENCE DES PROTEINES MAJEURES EXPRIMEES EN FIN DE SPOROGENIE</b>	<b>56</b>
<b>Résumé des travaux</b>	<b>57</b>
<b>Article :</b> "Proteomic analysis of the eukaryotic parasite <i>Encephalitozoon cuniculi</i> (Microsporidia): a reference map for proteins expressed in late sporogonial stages"	<b>59</b>
<b>5. RECHERCHE DE NOUVELLES PROTEINES DE PAROI CHEZ <i>E. CUNICULI</i></b>	<b>72</b>
<b>5.1. La chitine déacétylase putative de <i>E. cuniculi</i>: une protéine de surface probablement impliquée dans la formation de la paroi microsporidienne</b>	<b>72</b>
<b>Résumé des travaux</b>	<b>72</b>
<b>Article :</b> "The putative chitin deacetylase of <i>Encephalitozoon cuniculi</i> : a surface protein implicated in the microsporidian spore wall formation"	<b>75</b>
<b>5.2. Recherche de protéines candidates à une localisation pariétale par criblage bioinformatique du génome de <i>E. cuniculi</i></b>	<b>86</b>
5.2.1. Matériels et Méthodes	87
5.2.1.1. Analyses bioinformatiques	87
5.2.1.2. Biologie moléculaire	88
5.2.2. Résultats	89
5.2.2.1. Validation du crible bioinformatique sur des protéines expérimentalement identifiées comme pariétales	89
5.2.2.2. Recherche de candidates pariétales dans le protéome expérimental de <i>E. cuniculi</i>	90
5.2.2.3. Recherche de protéines candidates à une localisation pariétale dans l'ensemble des protéines d' <i>E. cuniculi</i>	91
<b>5.3. Proposition d'un modèle de mise en place de la paroi microsporidienne</b>	<b>92</b>
<b>DISCUSSION GENERALE ET PERSPECTIVES</b>	<b>93</b>
<hr/>	
<b>REFERENCES BIBLIOGRAPHIQUES</b>	<b>100</b>
<b>ANNEXES</b>	<b>123</b>



# *Introduction*



# INTRODUCTION

---

Les microsporidies sont des parasites intracellulaires obligatoires infestant un très large spectre d'hôtes qui inclut tous les vertébrés dont l'homme. Leur pathogénicité a des impacts économiques majeurs, et chez l'homme leur importance clinique a été révélée avec la pandémie du SIDA. Ces pathogènes opportunistes sont en effet responsables de nombreuses infections chez les individus immunodéprimés.

Les microsporidies ont longtemps été vues comme des eucaryotes primitifs, mais de récentes analyses phylogéniques nous font désormais considérer ces organismes soit comme de proches voisins des champignons, soit comme des champignons évolués. La forme infectieuse est une spore qui est libérée dans l'environnement après éclatement de la cellule hôte. Cette spore présente deux caractéristiques majeures : un tube polaire, enroulé dans le cytoplasme, qui a la capacité de se dévagner très rapidement pour permettre l'invasion cellulaire, et une paroi protéochitineuse très épaisse. Cette paroi est une structure rigide qui permet la survie de la spore dans l'environnement, le maintien de la forme cellulaire et permet l'accroissement de la pression interne conduisant à la dévagination du tube polaire. Cette structure est également à l'interface des relations hôte-pathogène : en particulier reconnaissance de la cellule hôte et des stimuli conduisant à l'invasion, ainsi que l'interface avec le système immunitaire.

Dans le cadre de l'analyse post-génomique de la microsporidie *Encephalitozoon cuniculi* (Texier *et al.*, 2005), parasite de mammifères et pathogène de l'homme, ces travaux de thèse ont eu pour premier objectif la description de son protéome exprimé dans les stades tardifs du cycle de développement (Brosson *et al.* en préparation). Les protéines majeures de ces stades ont été analysées par deux approches complémentaires, soit par séparation d'extraits protéiques sur gels d'électrophorèse bidimensionnelle (2D-PAGE) et analyse des spots par spectrométrie de masse, soit par analyse globale des extraits par spectrométrie de masse sans séparation préalable par 2D-PAGE (approche de type « shotgun »). Nous avons ainsi pu avoir, pour la première fois chez une microsporidie, une vision globale de la physiologie de la spore. D'autre part, parmi les 177 protéines majeures décrites, 25% correspondent à des protéines sans fonction prédite.

Le second volet de cette thèse, concerne la recherche de nouvelles protéines pariétales chez *E. cuniculi* et leur caractérisation au cours du cycle de développement. En effet, seule une protéine de surface, SWP1 avait été décrite auparavant chez cette microsporidie (Bohne *et al.*, 2000). Une





première exploitation des données du protéome a permis l'identification d'une chitine déacétylase putative localisée à l'interface membrane plasmique – endospore et probablement impliquée dans la formation de la paroi (Brosson *et al.*, 2005), ainsi que deux protéines localisées à l'endospore (Peuvel-Fanget *et al.*, sous presse). Nous avons également établi un crible bioinformatique pour sélectionner des protéines candidates à une localisation pariétale parmi les protéines à fonction inconnue du protéome de *E. cuniculi*, ces protéines devant être en partie spécifique du parasite. La validation de la localisation de ces candidates est en cours.

Ces travaux ont donné lieu aux publications suivantes :

### 🔗 Articles

Brosson D, Kuhn L, Delbac F, Garin J, Vivarès CP, Texier C.

Proteomic analysis of the eukaryotic parasite *Encephalitozoon cuniculi* (Microsporidia): a reference map for proteins expressed in late sporogonial stages. *Proteomics. Manuscrit soumis.*

Brosson D, Kuhn L, Prensier G, Vivarès CP, Texier C.

The putative chitin deacetylase of *Encephalitozoon cuniculi*: a surface protein implicated in the microsporidian spore wall formation. *FEMS Microbiology Letter* 2005 247:81-90.

Texier C, Brosson D, El Alaoui H, Méténier G, Vivarès CP.

Post-genomics of microsporidia, with emphasis on a model of minimal eukaryotic proteome: a review. *Folia Parasitologica* 2005 52 :15-22.

Peuvel-Fanget I, Polonais V, Brosson D, Texier C, Kuhn L, Vivarès C, Delbac F.

EnP1 and EnP2, two proteins associated with the *Encephalitozoon cuniculi* endospore, the chitin-rich inner layer of the microsporidian spore wall. *International Journal for Parasitology. Sous presse.*

### 🔗 Résumés de congrès

Brosson D, Texier C, Prensier G, Vivarès CP.

Proteomics of the Microsporidia *Encephalitozoon cuniculi* : from Proteins of Unknown Function to New Cell Wall Proteins. *Journal of Eukaryotic Microbiology* 2005; 52(2):3S (résumé 94A).

42ème congrès du Groupement des Protistologues de Langue Française, Montpellier, 2004. Communication orale.



Texier C, Brosson D, Kuhn L, Garin J, Vivarès CP.

Proteomics of the Microsporidian *Encephalitozoon cuniculi* spore : Application to the Identification of New Spore Wall Proteins. Journal of Eukaryotic Microbiology 2004. 51(2):28A (résumé 104).

41ème congrès du Groupement des Protistologues de Langue Française et 4ème biennale de Parasitologie, Clermont-Ferrand, 2003, communication orale.

Delbac F, Peuvel I, Brosson D, Méténier G, Vivarès CP.

Cell Invasion and Polar Tube Proteins in *Encephalitozoon spp.* Journal of Eukaryotic Microbiology 2004. 51(2).

41ème congrès du Groupement des Protistologues de Langue Française et 4ème biennale de Parasitologie, Clermont-Ferrand, 2003, communication orale.

Texier C, Roux F, Brosson D, Delbac F, Garin J, Vivarès CP.

Proteomics Analysis of *Encephalitozoon cuniculi*. Journal of Eukaryotic Microbiology 2003. 50(2):32A (résumé 114).

40ème congrès du Groupement des Protistologues de Langue Française, La Rochelle, 2002, communication orale.

## **🔗 Autres communications**

Texier C, Kuhn L, Brosson D, Garin J, Vivarès CP.

Protéomique de la microsporidie *Encephalitozoon cuniculi* : application à la recherche de nouvelles protéines pariétales. Congrès annuel de la Société Française de Biochimie et de Biologie Moléculaire, Lyon, 2003, communication par affiche.



*E t u d e*

*B i b l i o g r a p h i q u e*



# 1. PRESENTATION GENERALE DES MICROSPORIDIES

---

C'est le caractère pathogène des microsporidies qui permit leur découverte au cours du 19<sup>ème</sup> siècle. En quelques années, les élevages de vers à soie européens furent ravagés par une épizootie à très fort impact économique : la pébrine du ver à soie. Elle se caractérise par des vers dont les téguments sont recouverts de taches sombres, par une très forte mortalité et par une transmission à la descendance. En 1865, Pasteur s'intéressa à la maladie et mit au point une technique capable de la contrôler et de la prévenir. Cependant, c'est Nägeli qui identifia l'agent pathogène responsable en 1857, et le nomma *Nosema bombycis* mais l'assimila à une levure. En 1882, Balbiani, au vu des divergences morphologiques, suggéra de le classer dans un nouvel ordre : *Microsporidia*, dont les membres sont : « des organismes caractérisés par la petitesse de leurs spores comparée à celle des autres sporozoaires ». Pour tenir compte de leur très grande diversité, un phylum leur fut attribué par Sprague en 1971. Ces microsporidies sont des microorganismes pathogènes d'une très grande diversité, regroupant plus de 1200 espèces qui divergent aussi bien par leur morphologie et leur cycle de développement, que par le large spectre d'hôtes susceptibles d'être infestés. Ces parasites intracellulaires obligatoires sont des eucaryotes unicellulaires producteurs de spores (stade de résistance et de dissémination) qui partagent une même stratégie d'invasion faisant intervenir une structure spécifique, le tube polaire. Des microsporidies ont été mises en évidence dans pratiquement tous les groupes du règne animal : protozoaires, helminthes, arthropodes, oiseaux, poissons et mammifères. Leur impact sanitaire sur les animaux d'élevage peut être majeur. Chez les mammifères, dont l'homme, les microsporidioses, bien que souvent asymptomatiques, peuvent s'avérer dramatiques en particulier dans des contextes de co-infection ou d'individus immunodéprimés. Ainsi leur impact économique sur les élevages d'animaux, la prévalence importante dans la population humaine (de 2 à 22 % selon les études (Didier, 2005)) et l'issue fatale de plusieurs cas chez des patients atteints de SIDA ont fait évoluer le statut des microsporidies de curiosités biologiques à celui de pathogènes en émergence.

## 1.1. Un phylum fortement diversifié

Le phylum *Microsporidia* comprend des organismes unicellulaires eucaryotes, tous parasites intracellulaires obligatoires, répartis en 143 genres et plus de 1200 espèces (Vavra et Larsson,





1999). Ces microorganismes se caractérisent par la présence dans la spore d'un dispositif infectieux original, constitué par le tube polaire et ses annexes (capuchon polaire, polaroplaste, vacuole postérieure). La dévagination extrêmement rapide de ce tube, de l'ordre de quelques millisecondes (Frixione *et al.*, 1992), assure l'expulsion de l'élément infectieux ou sporoplasme (constitué du noyau et des composants cytoplasmiques de la spore) et permet son inoculation dans la cellule-hôte. Bien que partageant un mode de vie obligatoirement parasite intracellulaire et une stratégie commune d'invasion, les microsporidies forment un groupe hautement diversifié. En effet, elles présentent des variations au niveau de leurs morphologies (taille et forme de la spore, épaisseur de la paroi, nombre de noyaux, nombre d'enroulements du tube polaire) aussi bien entre espèces qu'au sein d'une même espèce chez des hôtes différents (Dunn et Smith, 2001). Une très grande diversité est également présente au niveau de la taille des génomes, de 2,3 Mpb pour *Encephalitozoon intestinalis* à 19,5 Mpb pour *Glugea atherinae*, ainsi qu'au niveau du nombre de chromosomes allant de 8 à 18 (Biderre *et al.*, 1995). Le cycle de développement présente aussi de grandes disparités à la fois dans son contexte (développement directement au contact du cytoplasme de la cellule-hôte ou bien à l'intérieur d'une vacuole parasitophore) que dans son déroulement (mérogonie par scission binaire ou bien formation d'un plasmode plurinucléé, nombre de division au cours de la sporogonie). De plus, la mise en évidence de formes monocaryotiques et diplocaryotiques chez les microsporidies du genre *Amblyospora* (Hazard et Brookbank, 1984), ainsi que la présence de complexes synaptonémaux suggèrent l'existence d'un cycle sexuel. Cette grande variabilité est également retrouvée dans le spectre d'hôtes, avec des espèces spécifiques d'un seul et unique hôte et d'autres véritablement ubiquistes. Ainsi, la classification historique des microsporidies qui a reposé bien souvent sur des variations phénotypiques, souvent observées sur un seul type de spores et chez un hôte unique, a pu conduire à une redondance au sein du phylum en référencant plusieurs fois une même espèce dont le polymorphisme n'avait pas été appréhendé. De plus, la description incomplète du cycle de développement de certaines espèces, a nécessité la création d'un genre spécifique, *Microsporidium*, pour regrouper ces espèces en attente de classification.

## **1.2. Une phylogénie controversée : d'eucaryotes primitifs à champignons atypiques**

Le positionnement des microsporidies dans l'arbre du vivant a longtemps été sujet à débat. Tout d'abord considérées comme des levures par Nägeli, les microsporidies furent ensuite classées parmi les plus primitifs des eucaryotes (Cavalier-Smith, 1998) au vu de l'absence d'organites considérés



comme quasi-universels chez les autres eucaryotes (mitochondrie et peroxyosome) et de la présence d'un appareil de Golgi très atypique (réseau tubulaire sans présence d'empilement de citernes). D'autres données appuyèrent cette origine primitive comme les coefficients de sédimentation des ribosomes similaires à ceux des procaryotes (Ishihara et Hayashin, 1968 ; Curgy *et al.*, 1980) et les résultats des premières études de phylogénie moléculaire basées sur les gènes de l'ARNr, et des facteurs d'élongation EF-1 $\alpha$  et EF-2 (Vossbrinck *et al.*, 1987 ; Kamaishi *et al.*, 1996). Cependant, à partir de 1997, l'identification de séquences codantes pour des protéines de stress adressées aux mitochondries (HSP70) permet d'envisager une autre hypothèse d'évolution des microsporidies : une évolution régressive avec perte secondaire des mitochondries (Hirt *et al.*, 1997 ; Germot *et al.*, 1997 ; Peyretailade *et al.*, 1998). Des études phylogéniques récentes, portant sur la grande sous-unité de l'ARN ribosomique, sur l' $\alpha$ - et la  $\beta$ - tubuline et sur plusieurs dizaines de protéines (van de Peer *et al.*, 2000 ; Keeling et Fast, 2002 ; Thomarat *et al.*, 2004), dressent le portrait d'un eucaryote dont la vitesse d'évolution est rapide, et dont les gènes sont sujets à une forte réduction de taille. De plus, ces études conduisent à considérer désormais les microsporidies soit comme une lignée qui aurait émergé des champignons, soit qui leur serait fortement apparentée. Ces analyses phylogéniques sont concordantes avec certaines particularités physiologiques qui sont partagées par les microsporidies et les champignons (Weiss et Vossbrinck, 1999) : appariement des noyaux en dicaryons chez certains genres, mitoses fermées avec présence de plaques polaires semblables à celles des levures (Vavra, 1976b ; Vivarès *et al.*, 1977), présence de tréhalose et de chitine dans la spore (Vavra, 1976a). Les premières analyses contradictoires de phylogénie moléculaire, qui conclurent toutes à une origine primitive des microsporidies, peuvent s'expliquer par la rapide évolution des gènes de l'ARNr et des facteurs d'élongation, conduisant à des séquences très divergentes, et donc propices au phénomène d'attraction des longues branches, responsable de constructions phylogéniques aberrantes (van de Peer *et al.*, 2000).

### **1.3. Un large spectre d'hôtes et une pathogénicité variée**

Le spectre d'hôtes des microsporidies englobe la quasi-totalité des groupes animaux (Canning et Lom, 1986 ; Canning, 1993). En dépit de l'impact sanitaire majeur que constituent les microsporidioses sur les élevages de poisson (Lom et Nilsen, 2003) et d'arthropodes (pébrine pour la sériciculture, « cotton shrimp disease » pour la crevetticulture), ce sont les microsporidioses humaines qui sont les mieux décrites dans la littérature. En effet, alors que les microsporidies peuvent infecter des sujets immunocompétents (van Gool *et al.*, 1994 ; Bryan *et al.*, 1997 ; Raynaud

Espèces microsporidiennes	Symptômes																	
	diarrhées	cholangites	cholécystites	hépatites	péritonites	encéphalites	kératoconjonctivites	rhinites	sinusites	bronchites	pneumonies	infections urinaires	urétrites	néphrites	prostatites	infections musculaires	infections disséminées	
<i>Enterocytozoon bienewsi</i>	x	x	x					x	x	x	x							
<i>Encephalitozoon intestinalis</i>	x	x	x				x	x	x		x	x		x			x	
<i>Encephalitozoon hellem</i> <sup>(1)</sup>							x	x	x	x	x	x	x	x	x		x	
<i>Encephalitozoon cuniculi</i>	x			x	x	x	x	x	x			x					x	
<i>Trachipleistophora hominis</i> <sup>(1)</sup>							x	x	x								x	
<i>Trachipleistophora anthropophthera</i> <sup>(1)</sup>						x											x	x
<i>Pleistophora</i> spp.																		x
<i>Vitaforma corneae</i> <sup>(2)</sup>							x					x						
<i>Nosema oculorum</i> <sup>(2)</sup>							x											
<i>Brachiola connor</i> <sup>(2)</sup>																		x
<i>Brachiola vesicularum</i> <sup>(1)</sup>																		x
<i>Brachiola algerae</i>							x											
<i>Microsporidium africanum</i> <sup>(2)</sup>							x											
<i>Microsporidium ceylonensis</i> <sup>(2)</sup>							x											

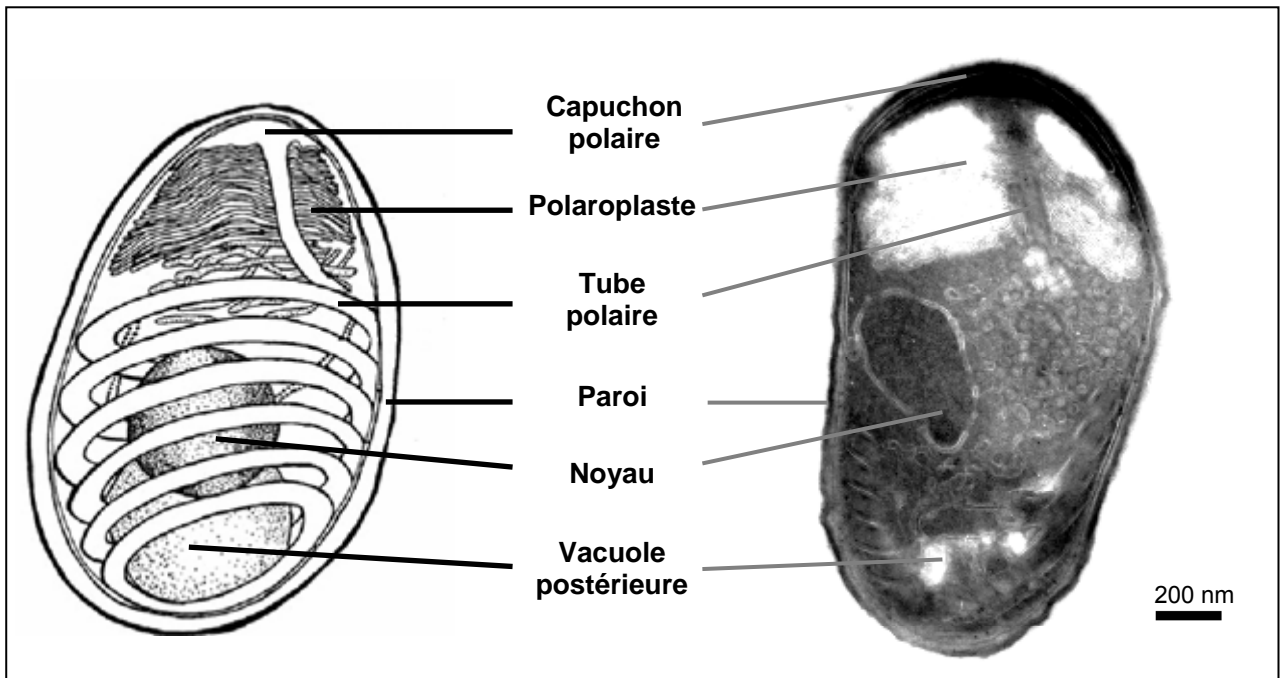
**Tableau 1.** Microsporidioses humaines et manifestations cliniques d'après Franzen et Muller (2001). Les espèces notées (1) n'ont été mises en évidence que chez des patients atteints de SIDA alors que celles notées (2) n'ont été identifiées que chez des individus immunocompétents.

*et al.*, 1998), la majorité des cas concerne des patients immunodéprimés (patients atteints de SIDA, ou soumis à des drogues immunosuppressives) chez qui les microsporidies provoquent des pathologies diverses (tableau 1) pouvant conduire au décès (Desportes-Livage, 1996 ; Muller *et al.*, 2001).

La prévalence des microsporidies dans la population générale est difficile à appréhender car les résultats des études portant sur des sujets non infectés par le HIV donnent des résultats variables selon l'espèce recherchée et la technique employée. Certaines études se basent sur la recherche d'anticorps anti-microsporidiens comme reflet d'une infection passée ou en cours et établissent une prévalence allant de 1,3 à 22 %. Les cohortes analysées ont été sélectionnées parmi des donneurs de sang, des femmes enceintes, des employés d'abattoir et des patients atteints de diarrhées chroniques (van Gool *et al.*, 1997 ; Kucerova-Pospisilova et Ditrich, 1998 ; Enriquez *et al.*, 1998 ; Del Aguila *et al.*, 2001 ; Kucerova-Pospisilova *et al.*, 2001). D'autres travaux ont recherché directement le parasite (par PCR ou microscopie) et donnent des valeurs voisines, entre 3,3 et 17,4%, pour une population d'enfants, de personnes âgées et de voyageurs tropicaux (Valperga *et al.*, 1999 ; Muller *et al.*, 2001 ; Lores *et al.*, 2002). Le caractère opportuniste de ce parasite, l'augmentation du nombre de patients immunodéprimés, le développement de meilleurs moyens de détection et une vigilance accrue à leur rencontre nous fait considérer, aujourd'hui, les microsporidies comme de véritables pathogènes en émergence.

#### **1.4. *Encephalitozoon cuniculi*, une microsporidie modèle**

En 1922, *Encephalitozoon cuniculi* a été la première microsporidie décrite chez un mammifère. Elle fut isolée dans le système nerveux et les reins d'un lapin atteint de paralysie motrice (Wright et Craighead, 1922). Par la suite, elle fut fréquemment rencontrée chez un grand nombre de mammifères dont les rongeurs, les lagomorphes, les carnivores, les ruminants et les primates. En 1959, la présence d'*E. cuniculi* a été clairement décrite dans le premier cas de microsporidiose humaine (Matsubayashi *et al.*, 1959). Son cycle de développement simple, à un seul hôte et sans forme sexuée a permis la mise au point d'une culture stable sur cellules de mammifère en 1969 par Shaddock. En 2001, la publication du génome d'*E. cuniculi* a montré sa forte compaction et sa réduction de taille (Katinka *et al.*, 2001). L'étude de ce génome « minimal » permet d'apporter des réponses à l'adaptation des microsporidies au parasitisme, et peut constituer un modèle simple pour l'étude comparée de parasites plus complexes (en particulier les membres du groupe des *apicomplexa*).



*Figure 1. Schéma et photographie en microscopie électronique à transmission (MET) d'une spore d'E. cuniculi. Sur le schéma, les deux feuilletts de la paroi sont représentés par un trait noir pour l'exospore et une région blanche pour l'endospore.*

### 1.4.1. Structure de la spore

La spore mature d'*E. cuniculi* est de taille réduite, de 2,0 à 2,5 µm sur 1,0 à 1,5 µm. Elle est protégée par une paroi épaisse qui enveloppe un sporoplasme contenant un noyau et un cytoplasme au sein duquel se structure l'ensemble de l'appareil invasif (figure 1).

#### 1.4.1.1. L'élément infectieux : le sporoplasme

Le sporoplasme est l'élément infectieux et pénètre dans la cellule-hôte lors de l'infestation. Il est constitué du noyau, du polaroplasme (cf chapitre 1.4.1.2. L'appareil invasif) et des composants cytoplasmiques. Comme toutes les microsporidies, le cytoplasme d'*E. cuniculi* se caractérise par une très grande richesse en ribosomes et d'une absence de mitochondrie ou de peroxyosome. Si certains éléments membranaires peuvent être qualifiés sans aucun doute de réticulum endoplasmique granuleux, d'autres pourraient simplement n'être que des vacuoles d'endo- ou d'exocytose (Vavra et Larsson, 1999). Quant à l'appareil de Golgi, son existence a longtemps été débattue car sa nature est très atypique. En effet, il ne se présente pas sous la forme classique d'un empilement de citernes plates. Cependant, il est désormais admis qu'il est constitué d'un réseau tubulaire présent chez les stades jeunes et porteur d'activités enzymatiques classiquement associées au Golgi (thiamine pyrophosphatase et nucléoside diphosphatase) (Vavra, 1968 ; Takvorian et Cali, 1994 ; Sokolova *et al.*, 2001a ; Sokolova *et al.*, 2001b).

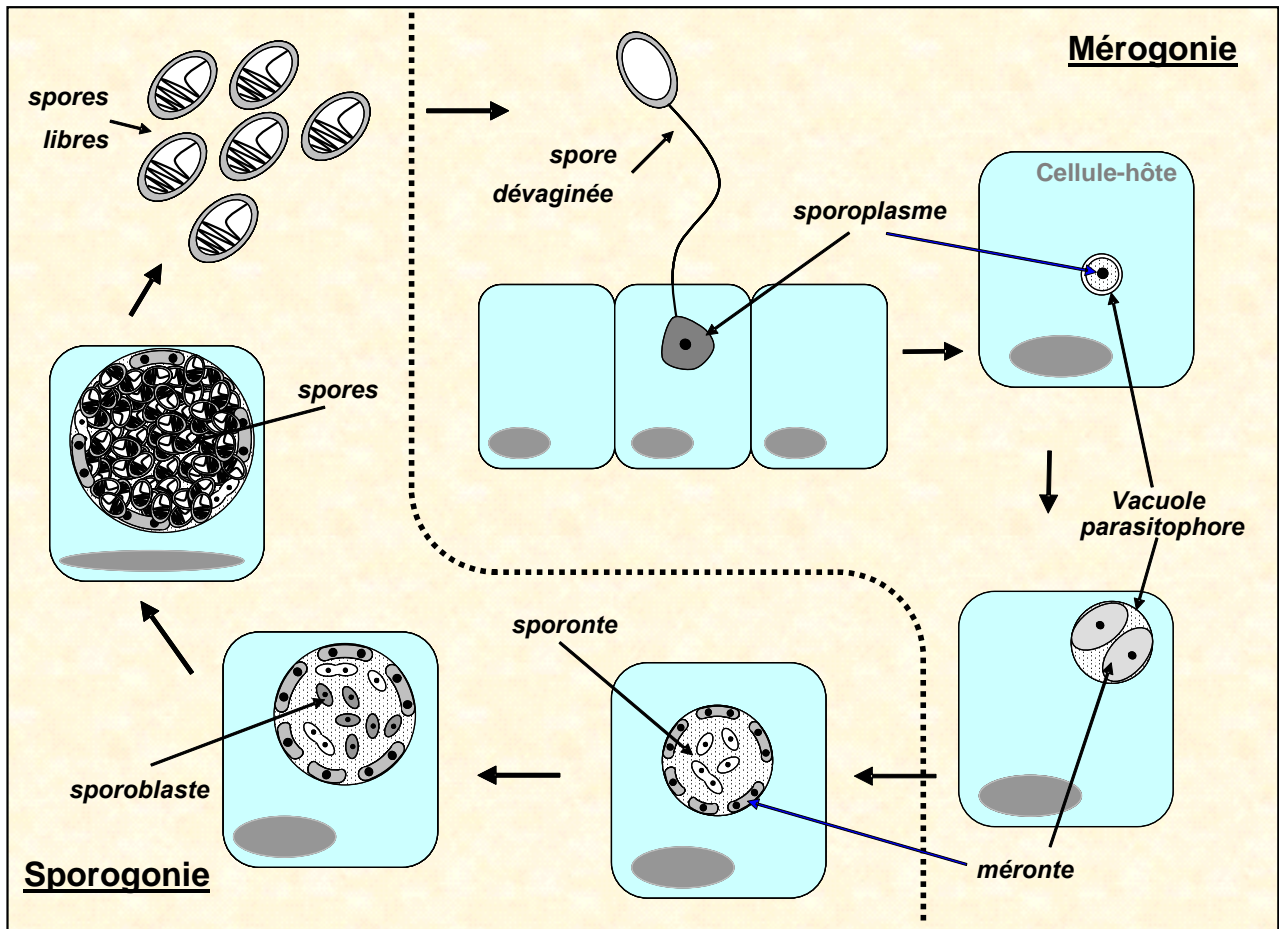
#### 1.4.1.2. L'appareil invasif

Les microsporidies se caractérisent par un appareil d'infestation original qui est constitué de quatre éléments : le tube polaire, le capuchon polaire, le polaroplasme, et la vacuole postérieure.

Le tube polaire est le composant principal de l'appareil d'invasion. Il s'agit d'une structure tubulaire enroulée dans le cytoplasme de la spore (entre cinq et six enroulements dans le cas d'*E. cuniculi*), dont la formation débute chez le sporoblaste<sup>1</sup> à partir du réticulum endoplasmique et des vésicules golgiennes (Vavra, 1976a). Trois protéines structurales (PTP1, PTP2 et PTP3) ont été mise en évidence chez *E. cuniculi* (Delbac *et al.*, 1998b ; Delbac *et al.*, 2001 ; Peuvél *et al.*, 2002). Le tube polaire vient se connecter au capuchon polaire qui est situé au pôle apical de la spore, à proximité d'une région pariétale moins épaisse. C'est à son niveau que la paroi se rompt lors de la dévagination du tube polaire. Le polaroplasme est un système membranaire volumineux dans la spore (pouvant constituer jusqu'à un tiers du volume sporal). Il est constitué d'une région lamellaire formée par l'empilement régulier de saccules très aplatis et limités par une membrane peu épaisse, et d'une région vésiculaire dense aux électrons. Lors de la germination de la spore et du passage du sporoplasme dans le tube polaire, le polaroplasme serait transféré simultanément de manière passive

<sup>1</sup> Sporoblaste : dernier stade de développement du parasite avant formation de la spore





**Figure 2.** Schéma du cycle de développement d'*E. cuniculi*. Le développement du parasite débute par la division du sporoplasme juste après son internalisation, et à lieu au sein d'une vacuole parasitophore dont l'origine membranaire n'est pas établie (Fasshauer et al., 2005). On peut diviser ce cycle en une phase de multiplication rapide (mérogonie) et une phase de différenciation (sporogonie) conduisant à la formation de spores matures. Ces spores sont finalement libérées, ainsi que tous les autres stades, lors de l'éclatement de la cellule-hôte.

pour former la membrane limitante du sporoplasme libéré à l'extérieur de la spore (Weidner *et al.*, 1984). La vacuole postérieure est une structure claire aux électrons, située au pôle basal de la spore, sa formation à partir de vésicules golgiennes intervient tardivement au cours de la sporogonie.

#### 1.4.1.3. La paroi : protection de la spore et interface hôte-parasite

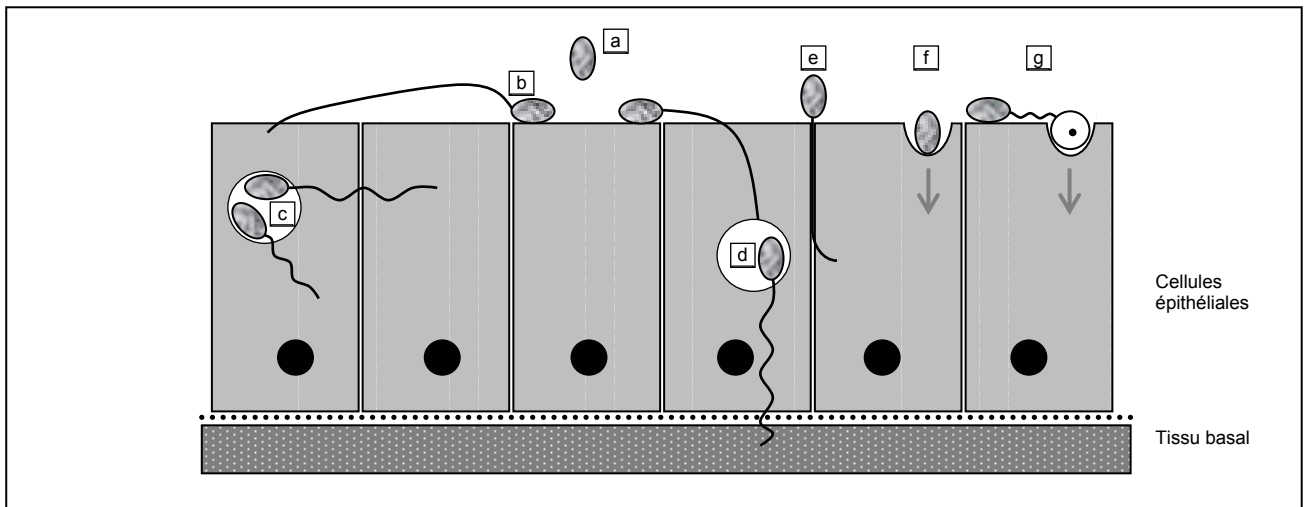
Mise en place très tôt au cours du cycle de développement du parasite (dès le stade sporonte<sup>1</sup>), la paroi joue un rôle majeur dans la protection de la spore contre les stress environnementaux, permettant une survie à long terme de la spore après sa libération par éclatement des cellules-hôtes (Shadduck et Polley, 1978 ; Koudela *et al.*, 1999). Il s'agit d'une épaisse structure rigide, de nature protéo-chitineuse, constituée de deux couches, l'endospore et l'exospore. La partie la plus interne de l'endospore semble être en contact très étroit avec la membrane cytoplasmique car ce contact demeure même après la libération du sporoplasme, celui-ci se dotant d'une nouvelle membrane plasmique en provenance du polaroplaste (Undeen et Frixione, 1991). Des données ultrastructurales de la paroi microsporidienne sont commentées dans la troisième partie de cette analyse bibliographique.

#### 1.4.2. Cycle et stades de développement

Le cycle de développement d'*E. cuniculi* (figure 2) comme toutes les espèces du genre *Encephalitozoon*, est un cycle direct asexuel qui se déroule dans une vacuole parasitophore à l'intérieur de la cellule-hôte. Il est constitué de deux étapes : une intense phase de prolifération (mérergie) et une étape de différenciation (sporogonie). Le développement du parasite au sein de la vacuole parasitophore se fait de façon centripète, les stades les plus jeunes étant dans la région périphérique de la vacuole, alors que les spores matures sont retrouvées dans la région centrale de celle-ci.

Le méronite est le premier stade issu de la division du sporoplasme, il s'agit d'une cellule de forme allongée, à large noyau qui se multiplie par simple fission binaire, et dont le cytoplasme, très riche en ribosomes, est très opaque aux électrons. Les méronites restent en contact étroit avec la membrane de la vacuole parasitophore durant tout le cycle de développement. La transition avec le sporonte est caractérisée par le dépôt à sa périphérie de matériel dense aux électrons, ainsi que par le détachement du parasite de la membrane de la vacuole (Pakes *et al.*, 1975). Les jeunes sporontes sont donc des stades libres dans la vacuole. Ils sont de forme allongée et vont s'arrondir durant leur développement. Parallèlement, le contenu cellulaire va commencer à s'organiser (début de formation du polaroplaste, accroissement de la densité du cytoplasme) et la paroi va poursuivre sa formation en s'épaississant. Finalement, les sporontes subissent une division cellulaire qui va

<sup>1</sup> Sporonte : stade de développement du parasite marquant la fin de la phase de multiplication et son engagement dans celle de différenciation conduisant à la spore mature



**Figure 3.** Les différentes voies d'entrée possibles des spores d' *E. cuniculi* (a) dans des cellules épithéliales. L'adhésion de la spore est suivie de la dévagination du tube polaire et la pénétration du parasite par la membrane apicale de la cellule (b). Des spores encore dans la vacuole parasitophore, peuvent infester secondairement leur cellule hôte ainsi que les cellules adjacentes (c) ou bien un autre tissu (d). La pénétration basolatérale se ferait par insinuation du tube polaire au niveau des jonctions intercellulaires (e). Enfin, une internalisation par endocytose directe de la spore (f) ou du sporoplasme (g) est aussi envisagée (d'après Vivarès et Méténier, 2001).

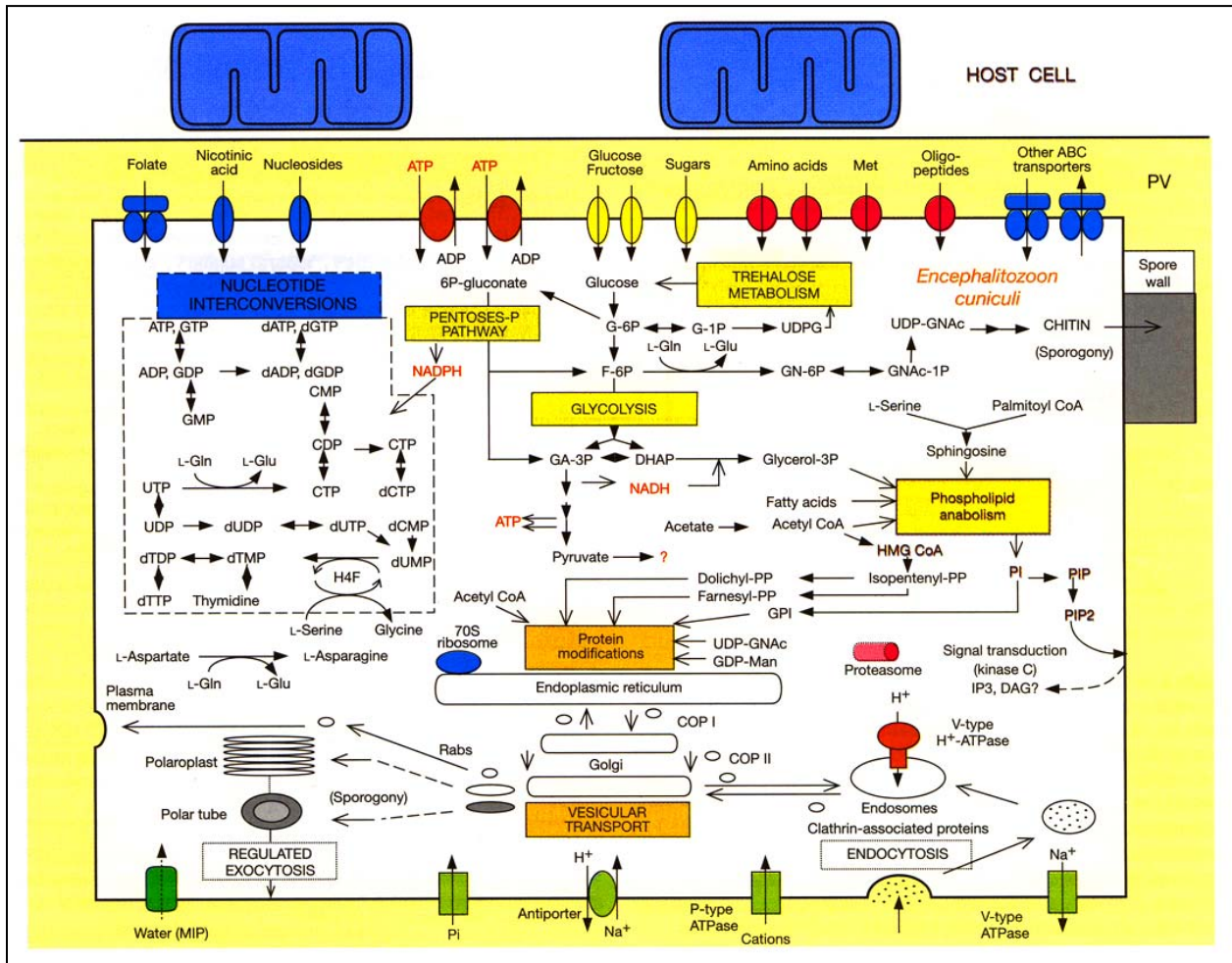
conduire à la formation de deux sporoblastes. Ceux-ci sont caractérisés par le début de formation du tube polaire, jusqu'alors indifférencié sous forme de vésicules et par l'accroissement de la complexité cytoplasmique (organisation du polaroplaste en région lamellaire et vésiculaire, apparition de la vacuole postérieure). Il semble clair que ce stade ne se divise plus. Il poursuit sa maturation par la formation complète de l'endospore ainsi que par une réduction de taille qui conduira finalement à la production d'une spore mature.

### **1.4.3. L'invasion cellulaire : hypothèses pour un mécanisme original**

Le mécanisme par lequel le sporoplasme envahit la cellule-hôte demeure sujet à débat et deux hypothèses sont proposées pour expliquer l'invasion cellulaire.

L'hypothèse classique pour la pénétration du parasite fait intervenir le tube polaire qui joue le rôle de l'aiguille d'une seringue. Des constituants présents au niveau de la paroi de la spore pourraient être impliqués dans une éventuelle reconnaissance tissulaire (Bohne *et al.*, 2000) et dans l'initiation du processus d'invasion. La nature de ces interactions n'est pas connue mais elles pourraient être le stimulus qui, au contact du tissu épithélial, déclencherait un influx d'eau dans la spore et induirait l'augmentation de la taille de la vacuole postérieure, ce qui conduirait à l'augmentation de la pression osmotique intrasporale. Cet accroissement de la pression provoquerait l'extrusion rapide du tube polaire au niveau du capuchon polaire (Frixione *et al.*, 1992). Deux origines possibles à cette augmentation de pression osmotique sont proposées : d'une part la création d'un gradient de protons à partir de l'environnement alcalin autour de la spore (Dall, 1983), ou bien la conversion des réserves de tréhalose en glucose (Undeen et Vander Meer, 1999). Il a aussi été montré que des flux d'ions  $\text{Ca}^{2+}$  joueraient un rôle dans la dévagination du tube polaire (Pleshinger et Weidner, 1985). Grâce à l'inertie de sa dévagination, le tube polaire traverse la membrane plasmique de la cellule hôte et le sporoplasme transite au travers du tube jusqu'au cytoplasme dans lequel va se former la future vacuole parasitophore. Les membranes nécessaires au transfert du sporoplasme ainsi que la membrane plasmique sont fournies par le polaroplaste (figure 3a à 3e).

D'autres modèles ont été proposés pour expliquer le processus invasif car des données expérimentales ne concordent pas avec le modèle classique d'invasion cellulaire. Ces hypothèses alternatives font intervenir directement la paroi. En effet, la vitesse du tube polaire lors de sa dévagination n'est pas suffisante pour lui permettre de traverser la membrane plasmique de la cellule hôte. Des auteurs font ainsi intervenir une internalisation directe par endocytose de la spore (Foucault et Drancourt, 2000) ou du sporoplasme (figure 3f et 3g). D'autres suggèrent que le tube



**Figure 4.** Reconstruction du métabolisme et des transports dans *E. cuniculi*, à partir de l'analyse de la séquence du génome. Le parasite est représenté dans une vacuole parasitophore (PV) au sein du cytoplasme de la cellule hôte. Les points d'interrogation correspondent aux incertitudes concernant le devenir du pyruvate et la nature des messagers secondaires des voies de transduction de signaux (Katinka et al., 2001).

polaire est une adaptation pour une cellule non mobile pour interagir avec des cellules plus ou moins distantes (Vivarès et Méténier, 2001).

#### 1.4.4. Un génome eucaryote minimal

Les premiers travaux portant sur le génome nucléaire d'*E. cuniculi* ont mis en évidence sa remarquable petitesse, la taille de son génome haploïde étant d'environ 2,9 Mpb et réparti sur 11 chromosomes linéaires de 217 à 315 kb (Biderre *et al.*, 1995). Cette taille est bien inférieure à celle du génome de nombreuses bactéries (*E. coli* : 4,6 Mpb) et bien plus petit que celui du champignon unicellulaire *Saccharomyces cerevisiae* (12,5 Mpb). Avec l'espèce voisine *Encephalitozoon intestinalis* (2,3 Mpb), ils font partie des plus petits génomes eucaryotes connus. Les caractéristiques majeures du génome d'*E. cuniculi* sont sa compaction et une réduction de taille qui s'opère à tous les niveaux (Katinka *et al.*, 2001). Au niveau de sa structure générale le génome est simplifié : peu de familles multigéniques et de séquences répétées, absence de minisatellites et d'éléments transposables, ainsi qu'une distance intergénique extrêmement raccourcie (129 bp en moyenne). Au niveau du nombre de gènes, l'annotation a mis en évidence un nombre réduit de séquences codantes (environ 2000 CDS dont 1909 non redondants) avec une perte de gènes conduisant à la simplification ou bien à la disparition de voies métaboliques entières (par exemple : absence totale de synthèse d'acides gras, absence de chaîne respiratoire, biosynthèse des acides aminés limitée à deux gènes). Les gènes présentent une quasi-absence d'introns (seulement 13). Enfin, cette réduction se retrouve au niveau de la séquence des protéines, dont la longueur moyenne est de 360 acides aminés, avec seulement six protéines dépassant les 2000 résidus. La comparaison de 350 protéines avec leurs homologues chez *S. cerevisiae* a montré une réduction de taille d'environ 15%, située principalement dans les domaines régulateurs (communication personnel E. Cornillot). L'annotation de ce premier génome de parasite a permis d'attribuer une fonction à 44% des séquences codantes, permettant de reconstruire son métabolisme (figure 4).

La disponibilité dans les banques de données publiques de ce génome et l'exploitation de ces données a ouvert le champ à la post-génomique (revue Texier *et al.*, 2005). Cela a permis, en particulier, de mettre en évidence la simplification des voies métaboliques en liaison avec le mode de vie parasitique ainsi que l'importance des échanges avec la cellule-hôte (Cornillot *et al.*, 2002). La disponibilité du génome a également permis d'effectuer une analyse phylogénique sur l'ensemble des protéines du génome, s'affranchissant de la subjectivité du choix des protéines soumises à l'analyse (Thomarat *et al.*, 2004). De manière plus générale, l'utilisation des orthologues réduits d'*E. cuniculi* peut, dans un contexte de génomique comparative, aider à discriminer entre les régions essentielles et non essentielles des protéines. Cette stratégie a été appliquée avec succès



pour déterminer les domaines catalytiques d'enzymes (application aux mécanismes de mRNA capping par Hausmann *et al.*, 2002) et, de même, pour résoudre la structure tridimensionnelle d'enzymes conservées (Fabrega *et al.*, 2004).

Cependant, aussi puissantes que soient les données issues de la recherche en génomique, « la réalité cellulaire est plus complexe que les potentialités du noyau seul » (Anderson et Anderson, 1998) et de nombreuses questions restent en suspens : quand, où et comment s'exprime une protéine ? Avec quel(s) autre(s) acteur(s) cellulaire(s) interagit-elle ? De plus, la génomique ne nous apporte aucune réponse sur les 56% de gènes pour lesquels aucune fonction n'a pu être déterminée par homologie de séquence. Seule l'analyse des éventuels produits de ces gènes pourrait nous guider vers la détermination d'une fonction. L'étude du génome comporte donc des limites, dues notamment à l'écroulement du dogme « un gène = une protéine = une activité ». En effet, de nombreux niveaux de contrôle rendent l'expression et l'activité d'une protéine plus subtiles : contexte spatio-temporel, modifications co- et post-traductionnelles, conditions physico-chimiques, interactions (autres protéines, acides nucléiques, substrats). Ceci ouvre donc la voie à la post-génomique, regroupant l'ensemble des analyses susceptibles, en s'appuyant sur les données de génomique, de décrire le résultat de ces régulations. Seule l'étude directe des protéines à l'échelle de la cellule, dans un contexte donné, constitue une description dynamique de la régulation des gènes et oriente vers la compréhension de la fonction protéique. Il apparaît donc clairement que la compréhension globale de systèmes biologiques, dans un contexte de post-génomique, nécessite deux approches qui sont indispensables et complémentaires : l'étude du génome et celle du protéome.





## 2. LA PROTEOMIQUE : un outil d'analyse post-génomique applicable aux parasites unicellulaires

---

Le terme « protéome » est né durant l'année 1994 dans plusieurs congrès pour désigner la totalité des protéines exprimées par un génome dans l'espace et dans le temps. Il s'agit donc, contrairement au génome, d'une image dynamique dépendant du tissu, du stade de développement et des conditions environnementales dans lesquelles l'organisme se situe. La protéomique qui constitue donc l'analyse du protéome, regroupe plusieurs champs d'investigation qui ont un même objectif : obtenir une vision globale et intégrée des processus biologiques par l'étude globale des protéines d'une cellule plutôt que de chacune individuellement. Ceci inclut non seulement une identification des protéines exprimées mais aussi l'élucidation de leur fonction et de leur localisation, les interactions protéines-protéines ainsi que l'étude des modifications post-traductionnelles.

La difficulté première de l'analyse d'un protéome est la nature même de l'échantillon étudié, qui est un mélange complexe de protéines. Ainsi, historiquement, l'analyse protéomique s'est appuyée sur la technique de séparation protéique la plus discriminante disponible, qui à l'époque était l'électrophorèse bidimensionnelle (2D-PAGE). Développée en 1970 (Kenrick et Margolis, 1970), cette technique permet de séparer efficacement les différentes protéines d'un extrait biologique. Dès ses premières applications, en 1975, elle mettait en évidence sa forte capacité résolutive (Scheele, 1975) s'illustrant par une carte protéique où s'individualisaient près de 1100 protéines majeures d'*Escherichia coli* (O'Farrell, 1975), ou bien plusieurs centaines à partir de tissus de souris (Klose, 1975).

Dans le but de pouvoir identifier ces protéines, la première technique utilisée fut un séquençage protéique par dégradation d'Edman, technique fiable et automatisable mais lente et assez peu sensible. Elle est désormais supplantée par une technologie plus rapide et beaucoup plus sensible, automatisable et adaptable à l'analyse haut-débit, la spectrométrie de masse qui est issue de l'analyse chimique. Elle permet d'obtenir la masse d'ions moléculaires, avec une précision inférieure au Dalton, et donc de discriminer des molécules de masses moléculaires même très proches. C'est grâce à un ensemble de progrès technologiques tels que la possibilité d'ioniser des protéines (Tanaka *et al.*, 1988 ; Fenn *et al.*, 1989) ou la disponibilité croissante des bases de données nucléiques, qu'à partir de 1993, différents auteurs ont fait évoluer cette technique (Henzel



*et al.*, 1993 ; Mann *et al.*, 1993 ; Pappin *et al.*, 1993 ; James *et al.*, 1993 ; Yates *et al.*, 1993). Ils ont pu corrélérer les informations provenant des spectres de masse peptidique avec les séquences protéiques déduites des génomes, conduisant à l'identification des protéines présentes dans les spots d'une électrophorèse bidimensionnelle.

Parallèlement à ces analyses d'identification par spectrométrie de masse, d'autres technologies qui ne seront pas décrites ici, ont été développées pour étudier les interactions protéiques au sein du protéome : phage display, puces à protéines, double hybride en levure.

## **2.1. La protéomique : à quelles fins ?**

La protéomique se décline généralement en plusieurs stratégies.

### **2.1.1. La protéomique descriptive**

Le but de la protéomique descriptive est de fournir une image de l'ensemble des protéines exprimées pour une cellule, un tissu ou un organisme donné (Payne et Garrels, 1997 ; Florens *et al.*, 2002). Elle permet d'obtenir une vision globale des protéines exprimées par un type cellulaire pour des conditions données. Mais elle peut aussi être appliquée, non pas à une cellule entière, mais à un compartiment cellulaire, ce qui constitue un sous-protéome c'est-à-dire l'ensemble des protéines contenues dans une fraction cellulaire (membrane, noyau, organelles, cytosol). Tout compartiment de la cellule pouvant être purifié, sans être altéré, peut être l'objet de ce type d'étude. En plus de fournir une image des protéines présentes dans un organite donné, ces études apportent une information sur la localisation des protéines identifiées, ce qui constitue une approche intéressante pour permettre l'attribution de fonction à des protéines hypothétiques. *A contrario*, l'identification de protéines à fonction connue, associées à certains organites permet de mieux comprendre le rôle de ceux-ci dans la biologie de la cellule. L'ensemble des données fournies par la protéomique descriptive représente donc un instantané des protéines exprimées dans des conditions et à un temps donnés. Ces résultats peuvent être représentés par une carte protéique, c'est-à-dire par la liste des protéines identifiées, associée aux spots protéiques correspondant sur une électrophorèse bidimensionnelle. Ils apportent les clés, en identifiant les acteurs protéiques, d'une meilleure compréhension de la physiologie de l'entité biologique étudiée.

D'autre part, les résultats fournis par la protéomique descriptive permettent d'avoir, dans un contexte de post-génomique, un retour critique sur l'annotation des génomes séquencés. Ils permettent de valider l'existence d'un gène prédit et d'identifier des séquences codantes non



prédites par l'annotation des génomes (Oshiro *et al.*, 2002). Ils permettent aussi de résoudre certains problèmes de prédiction bioinformatique tels que le choix des limites des séquences codantes (notamment le choix de la méthionine initiatrice), ainsi que la prédiction d'éventuels séquences introniques.

### **2.1.2. La protéomique différentielle**

Il s'agit d'étudier l'expression des protéines de tissus ou d'organismes entre deux conditions différentes et de déterminer les variations d'expression protéique de manière qualitative et quantitative. Ces résultats permettent d'établir une liste de protéines dont l'expression est variable : absence, sous- ou sur-expression entre les deux conditions. Par exemple, ceci permet de définir des variations d'expression de protéines en fonction du stade de développement (Florens *et al.*, 2002), en réponse à la présence de drogue (Cooper et Carucci, 2004) ou à un stress environnemental (Yin *et al.*, 2004). De façon générale cette approche permet d'éclairer sur les variations des mécanismes physiologiques de la cellule en réponse directe à une situation donnée.

### **2.1.3. La protéomique fonctionnelle**

Ici, le but est de pouvoir élucider les interactions entre protéines, pour établir des réseaux d'interaction, de façon à mieux appréhender leurs fonctions et leurs régulations. Des technologies comme le double hybride (Uetz *et al.*, 2000 ; Sprinzak *et al.*, 2003) ou les biopuces à protéines (Service, 2005) permettent de résoudre ces questions. Cependant, ces outils ne sont développés que chez la levure, ce qui oblige l'analyse des interactions protéine-protéine à être effectuée en système hétérologue, et avec un nombre limité de partenaires. Une technologie alternative, applicable aux organismes génétiquement transformables, repose sur une stratégie par spectrométrie de masse permettant une analyse *a posteriori* de ces interactions. Un gène d'intérêt est marqué par recombinaison homologue. La protéine exprimée interagit dans la cellule de manière native avec ses partenaires et dans les bonnes conditions physicochimiques (localisation cellulaire respectée). Les complexes protéiques associés à la protéine marquée sont extraits et analysés par spectrométrie de masse. Cette stratégie a permis de déterminer plus de 200 complexes chez la levure *Saccharomyces cerevisiae*, conduisant à l'annotation de 231 protéines à fonction inconnue (Gavin *et al.*, 2002).



#### **2.1.4. La protéomique structurale**

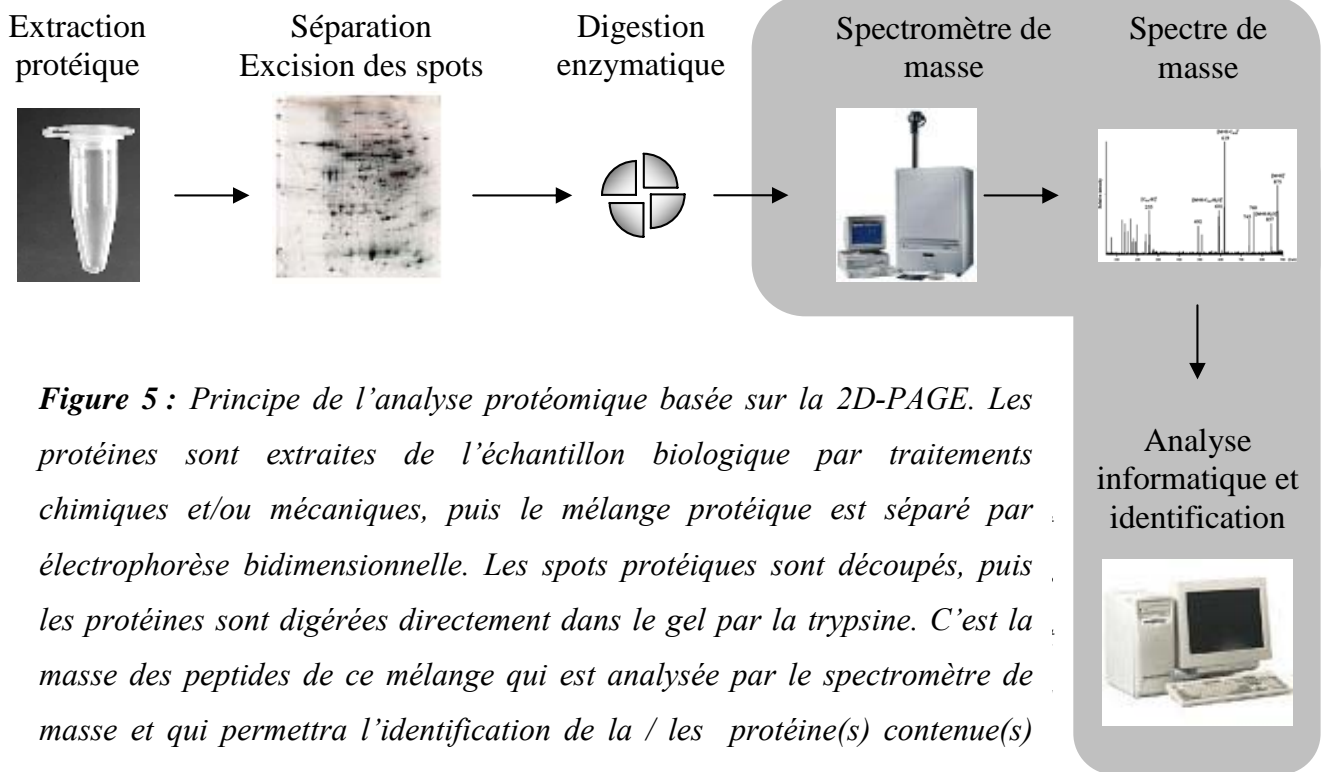
Le but de la protéomique structurale est de déterminer la relation entre la structure et la fonction d'une protéine, en particulier par la résolution de sa structure tridimensionnelle. Celle-ci est déterminée par des approches expérimentales, comme la radiocristallographie ou la spectrométrie à résonance magnétique nucléaire. Des travaux proposent des outils permettant de prédire une fonction protéique sur les bases de l'analyse de la comparaison de sa structure tridimensionnelle à celles connues (Jung et Lee, 2004). De plus, cette structure peut être utilisée pour interroger les réseaux d'interactions protéiques permettant de valider les interactions proposées. Cependant, cette analyse n'est pas compatible avec des études à grande échelle et constitue toujours, malgré le développement de plateformes de cristallisation, un goulot d'étranglement. Des solutions sont proposées pour lever ce problème. La stratégie retenue pour le projet de génomique structurale de la levure passe par l'automatisation des procédures d'obtention des cristaux protéiques (Leulliot *et al.*, 2005). D'autres travaux proposent des outils informatiques permettant aussi bien la prédiction de structures déjà connues que totalement nouvelles (Liu et Hsu, 2005).

#### **2.1.5. L'analyse des modifications post-traductionnelles**

Une protéine n'est mature et active qu'après plusieurs étapes complexes de modifications post-traductionnelles. Pouvoir obtenir des informations sur ces modifications est essentiel pour comprendre les régulations, la localisation et la fonction d'une protéine. Actuellement, le grand intérêt porté à l'étude de ces modifications a conduit à envisager leur analyse pour l'ensemble du protéome dans le but d'obtenir une vision globale des régulations cellulaires. On parle donc, par exemple, de glycoprotéomique pour désigner l'analyse de l'ensemble des glycoprotéines de la cellule, et de phosphoprotéomique pour l'étude des protéines phosphorylées.

Plusieurs stratégies ont été proposées pour mener ces études. D'une part, des protocoles d'extractions permettent d'enrichir un extrait protéique en protéines porteuses d'un type de modification (Tajiri *et al.*, 2005). D'autre part, ces modifications sont appréhendées directement par spectrométrie de masse dont la sensibilité (au dixième de Dalton) permet d'en déterminer avec précision la nature. Mais cela nécessite le développement de protocoles préparatoires et d'identification spécifiques (Sagi *et al.*, 2005).





**Figure 5 :** Principe de l'analyse protéomique basée sur la 2D-PAGE. Les protéines sont extraites de l'échantillon biologique par traitements chimiques et/ou mécaniques, puis le mélange protéique est séparé par électrophorèse bidimensionnelle. Les spots protéiques sont découpés, puis les protéines sont digérées directement dans le gel par la trypsine. C'est la masse des peptides de ce mélange qui est analysée par le spectromètre de masse et qui permettra l'identification de la / les protéine(s) contenue(s) dans le spot.

## **2.2. Principes de l'analyse protéomique basée sur l'électrophorèse bidimensionnelle**

Les stratégies d'analyse protéomique basée sur l'électrophorèse bidimensionnelle (ou approche classique) reposent sur trois grandes étapes critiques : l'extraction des protéines d'un échantillon biologique, leur séparation par électrophorèse bidimensionnelle et leur identification par spectrométrie de masse (figure 5). Une revue globale des protocoles et des avancées pour l'électrophorèse 2D-PAGE est disponible (Gorg *et al.*, 2004).

### **2.2.1. L'extraction protéique**

La qualité de l'extraction protéique est un point majeur de toute analyse protéomique, elle doit être effectuée dans des conditions de travail strictement contrôlées de manière à minimiser les contaminations par la kératine. L'extraction a pour but de solubiliser le plus grand nombre possible de protéines, en éliminant les liaisons entre elles (liaisons ioniques, hydrogène, hydrophobes et ponts disulfure) ou avec d'autres molécules biologiques (lipides, acides nucléiques). Elle doit permettre aussi de limiter leur dégradation et/ou modification par des enzymes et des réactions chimiques. Le tampon d'extraction joue également un rôle après le dépôt de l'extrait protéique sur le gel de première dimension, en limitant la précipitation des protéines durant l'isoélectrofocalisation (IEF), tout en aidant au maintien de la charge globale des protéines.

De nombreux protocoles ont été proposés pour s'adapter au matériel biologique étudié. Cependant, l'étape d'extraction se doit d'être la plus simple possible pour accroître la reproductibilité. Cette extraction se fait classiquement en présence de chaotropes (urée, thiourée), d'agent réducteur (DTT,  $\beta$ -mercaptoethanol) et de détergents (CHAPS, Triton X-100, SDS à faible concentration) (Rabilloud *et al.*, 1997), et peut être complétée par des traitements mécaniques (cycles de congélation / décongélation, sonication, broyage). Toutefois, l'emploi du SDS et du CHAPS comme détergent n'est pas suffisant pour permettre une extraction satisfaisante des protéines transmembranaires. Des détergents non ioniques ont donc été développés spécifiquement pour accroître la solubilisation des domaines hydrophobes, tout en conservant aux protéines leur charge ionique (Chevallet *et al.*, 1998). Les composés les plus efficaces sont des sulfobetaines de type zwitterion, les deux plus utilisés étant le SB 3-10 et l'ASB14 (Molloy, 2000).

### **2.2.2. De l'extrait protéique aux spots protéiques**

#### **2.2.2.1. Séparation sur gel 2D-PAGE des protéines extraites**



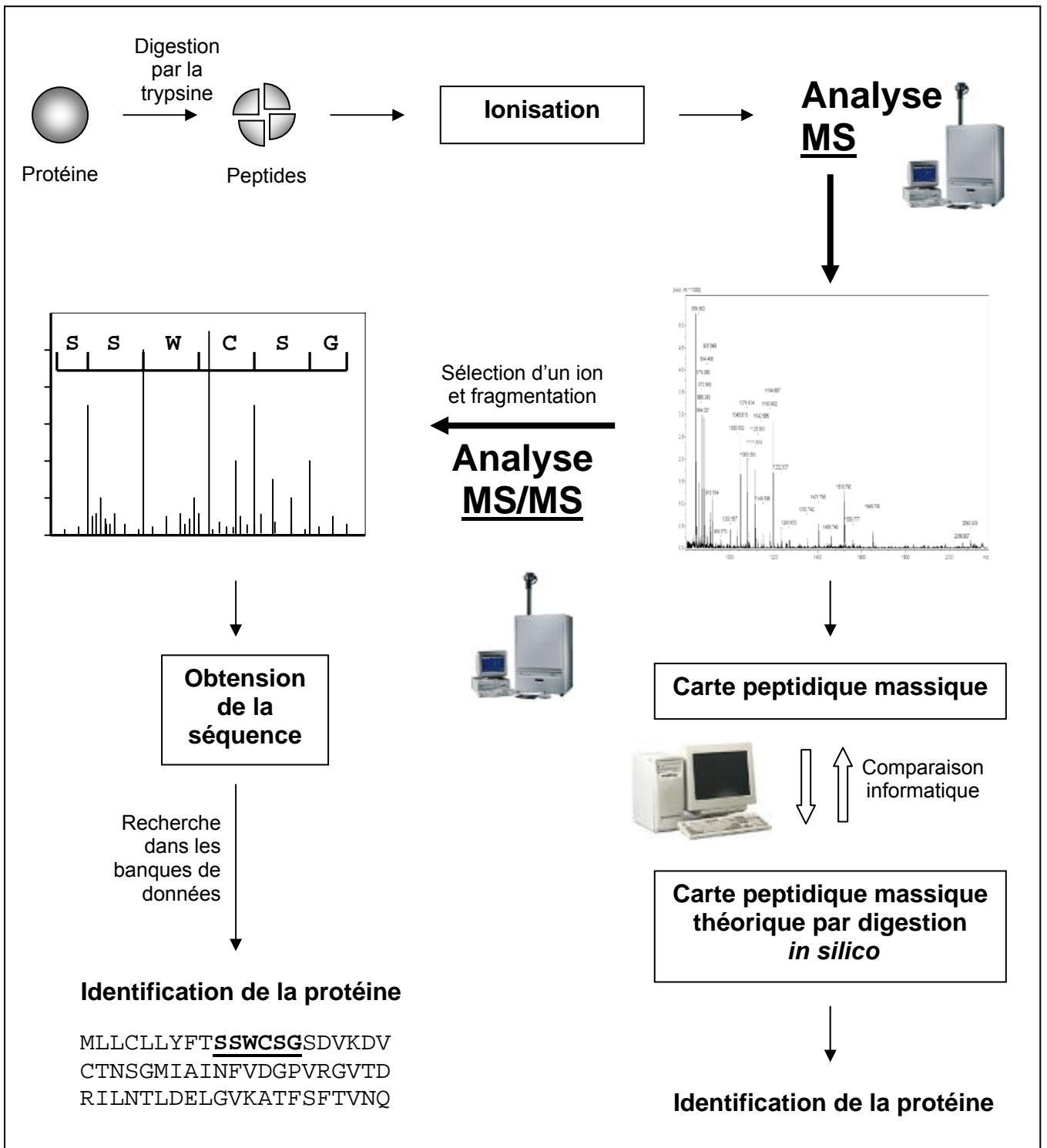
Bien que de nouvelles technologies de spectrométrie de masse permettent de s'affranchir de la séparation des protéines sur gel (voir le chapitre 2.4.2. Vers une protéomique sans électrophorèse), celle-ci reste encore pour la majorité des études une étape cruciale. La séparation la plus résolutive couramment utilisée est l'électrophorèse bidimensionnelle où les protéines sont séparées dans un premier temps selon leur point isoélectrique lors de l'isoélectrofocalisation (ou IEF), puis selon leur masse (SDS-PAGE). Cependant, depuis sa mise au point en 1975 par O'Farrell, cette technique a connu des apports non négligeables. L'un d'entre eux a été l'utilisation de gel d'isoélectrofocalisation à gradient de pH immobilisé (Bjellqvist *et al.*, 1982 ; Gorg *et al.*, 2000), permettant d'obtenir une bonne reproductibilité entre plusieurs gels.

#### **2.2.2.2. Visualisation des spots protéiques**

La dernière étape de l'électrophorèse 2D consiste à détecter les spots protéiques par coloration des gels. Plusieurs procédés de sensibilité différente existent et sont choisis en fonction de l'utilisation ultérieure des gels 2D. Le bleu de Coomassie permet de détecter un minimum de 100 ng de protéine par spot et présente l'avantage de donner une intensité de coloration proportionnelle à la quantité de protéines. La coloration au nitrate d'argent est, quant à elle, jusqu'à mille fois plus sensible et permet de détecter des spots contenant 0,1 ng de protéines. Elle présente néanmoins certains inconvénients : la stœchiométrie de la coloration n'est pas totalement linéaire, la reproductibilité est difficile à obtenir et certaines protéines sont peu ou pas colorées par cette méthode. Les protéines peuvent aussi être visualisées par des colorants fluorescents (CyDyes (Amersham), SYPRO Ruby (Rabilloud *et al.*, 2001)) dont la sensibilité, moins de 1 ng de protéine par spot, est bien supérieure au bleu de Coomassie, et dont la linéarité est maintenue sur plus de trois puissances de dix. Des avancées récentes, en particulier la simplification des protocoles (Lamanda *et al.*, 2004) rendent ces colorations très prometteuses. Des techniques alternatives existent pour visualiser les protéines, comme l'acquisition par phospho-imager de gel de protéines radiomarquées (Patterson et Latter, 1993). C'est actuellement la technique la plus sensible, permettant de visualiser moins de 12 copies d'une protéine par cellule et dont la linéarité ne présente pas de limite supérieure. Cependant, en plus des problèmes associés à la gestion de la radioactivité, l'acquisition d'image par cette technique est lente (plusieurs jours pour un gel) et nécessite un fort coût d'équipement.

#### **2.2.2.3. Traitement informatique des images 2D-PAGE**

Selon la technique de visualisation employée et de numérisation des gels, l'image obtenue nécessite une étape plus ou moins complexe de traitement graphique pour réduire le bruit de fond, éliminer les rayures et accroître le contraste. L'image obtenue est ensuite traitée par des logiciels d'analyse d'images 2D tels que Melanie (Amersham Biosciences) ou PD-Quest (BioRad). Ils permettent un



**Figure 6.** Identification protéique par spectrométrie de masse. Les protéines sont digérées par la trypsine (protéolyse après une arginine ou une lysine). Les peptides obtenus sont ionisés et analysés par le spectromètre de masse (MS). Chaque ion est représenté selon sa valeur  $m/z$  et son abondance. La liste des masses de ces ions est appelée carte peptidique massique et peut être comparée aux masses des peptides obtenues par digestion *in silico* du protéome théorique de l'organisme étudié ou des séquences protéiques présentes dans les bases de données internationales. Cette comparaison permet l'identification protéique. Dans les cas où cette analyse n'est pas possible ou pas suffisante une analyse par spectrométrie de masse en tandem (MS/MS) est nécessaire. Pour cela un peptide est sélectionné, fractionné et la masse de chacun des peptides issus de son fractionnement est analysée, ceci conduit à l'obtention d'une partie de la séquence de la protéine analysée qui est utilisée pour l'identification finale.

pré-traitement automatisé des spots et de leur quantification, mais cette étape doit être complétée manuellement avec pour référence le gel coloré. Ensuite, chaque spot d'un gel peut être associé à son équivalent sur un autre gel pour constituer un gel virtuel de référence. Cette étape, assistée par le logiciel, nécessite le positionnement manuel de repères invariants d'un gel à l'autre, à partir desquels le logiciel associe les autres spots. La comparaison entre gels de référence permet d'identifier des variations qualitatives et quantitatives entre les profils protéiques.

### **2.2.3. Identification des protéines**

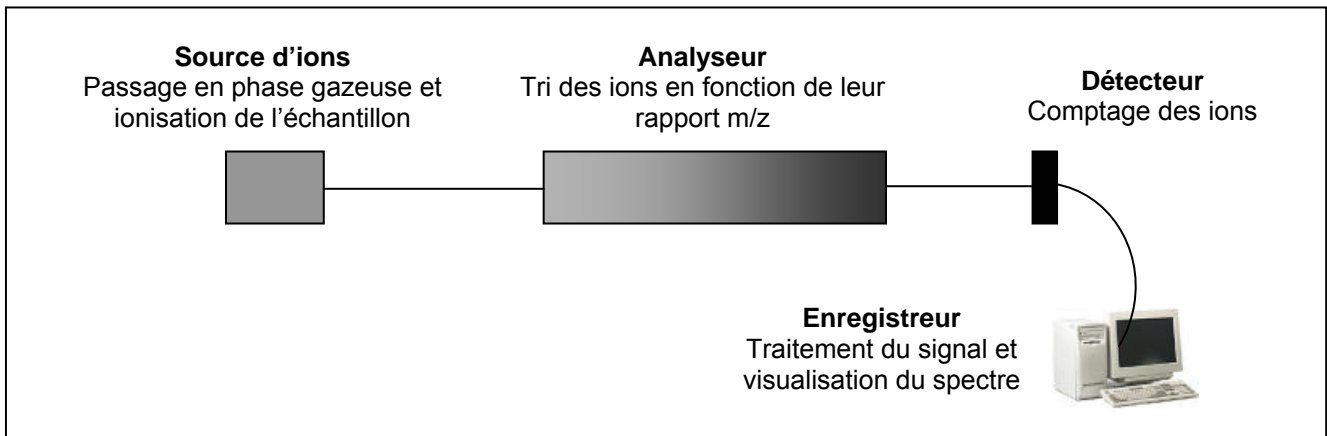
Après la coloration et l'acquisition informatique de l'image du gel, les spots individualisés sont découpés pour permettre l'identification des protéines. Historiquement, la première méthode utilisée pour identifier les protéines présentes dans les spots d'électrophorèse bidimensionnelle fut le microséquençage par dégradation d'Edman (Edman, 1949). Elle permet d'obtenir, par dégradation chimique, la séquence N-terminale de la protéine étudiée. Bien que cette technologie soit désormais automatisée, l'obtention d'une séquence est longue (4 à 8 heures) et nécessite une relative abondance en protéine (20 pmoles, c'est-à-dire 1 µg pour une protéine de 50 kDa). Cette technique est encore spécifiquement utilisée pour des cas précis comme l'analyse de la région N-terminale des protéines pour déterminer précisément le codon initiateur du gène ou le site de clivage d'un peptide signal. Cependant, pour l'analyse protéomique cette technique a été remplacée en routine par la spectrométrie de masse qui permet l'identification de protéines à une plus grande échelle, plus rapidement et avec une bien meilleure sensibilité.

## **2.3. La spectrométrie de masse : une technique d'identification de choix**

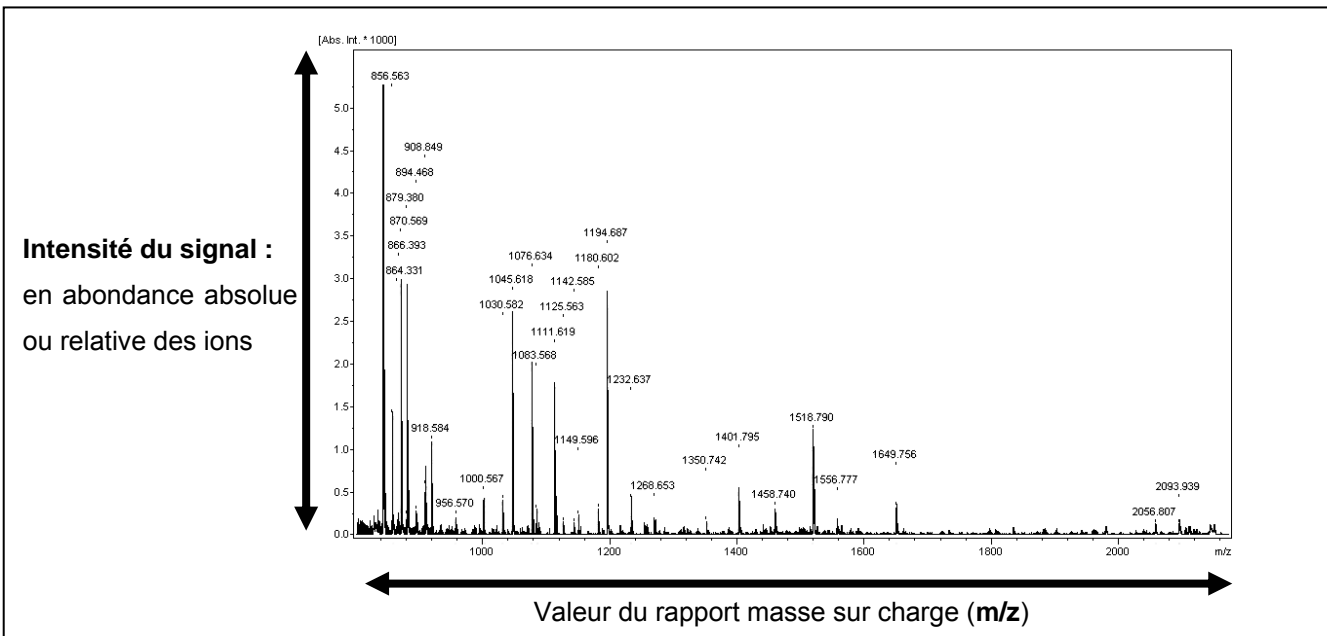
### **2.3.1. Définition de la technique**

La spectrométrie de masse permet de transformer des molécules en ions à l'état gazeux et d'obtenir leur masse moléculaire ( $m$ ) en analysant le rapport de leur masse sur leur charge ( $z$ ), noté  $m/z$ . Ce rapport  $m/z$  des ions formés à partir de l'échantillon est mesuré par un spectromètre de masse.

Le principe général de l'identification par spectrométrie de masse des protéines, préalablement digérées par la trypsine, est représenté sur la figure 6. L'analyse de la masse se divise en deux catégories suivant la technique utilisée conduisant soit à l'obtention de la masse de chaque peptide



**Figure 7.** Schéma de principe d'un spectromètre de masse de type MS.



**Figure 8.** Spectre de masse. Chaque ion est représenté en ordonnée par son abondance et en abscisse par sa valeur de rapport masse sur charge ( $m/z$ ) sous la forme d'un massif de pic (massif isotopique) correspondant aux isotopes naturels  $C^{13}$  de cet ion. Comme la nature des charges est connue car elle ne dépend que de la nature de l'ionisation, on peut, par déconvolution du spectre, obtenir de ce dernier la valeur de masse de l'ion analysé.

trypsique (spectrométrie de masse de type MS), soit de la masse de ces peptides associée à une séquence partielle (spectrométrie de masse en tandem ou MS/MS).

### **2.3.2. Principe de la spectrométrie de masse de type MS**

Le spectromètre de masse, dont le schéma de principe est illustré par la figure 7, est constitué des éléments suivants :

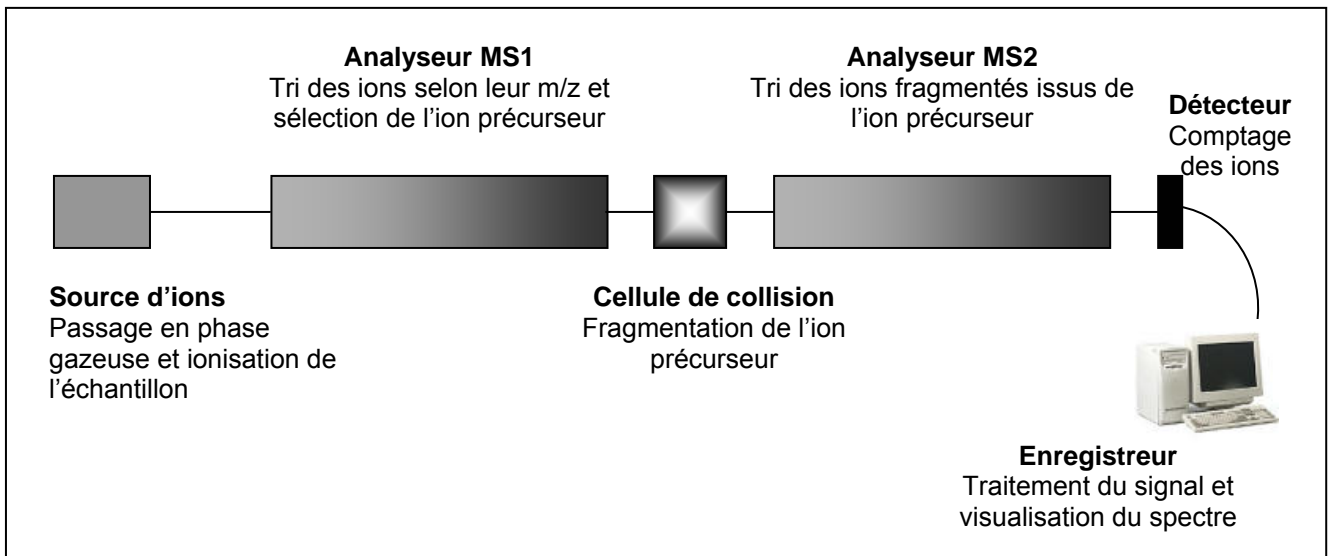
- Une source d'ions dans laquelle se produit le passage en phase gazeuse (vaporisation, sublimation, désorption) de l'échantillon à analyser et l'ionisation des molécules ;
- Un analyseur qui permet de trier les ions en fonction de leur rapport  $m/z$  ;
- Un détecteur qui compte les ions en leur associant leur rapport  $m/z$  ;
- Un enregistreur pour le traitement du signal et la visualisation des spectres.

L'analyse des peptides par spectrométrie de masse repose sur la détermination des rapports  $m/z$  d'un mélange de peptides et l'obtention d'un spectre de masse (figure 8). A partir d'un génome séquencé, on peut construire une base de données qui référence l'ensemble des peptides obtenus par digestion *in silico* de l'ensemble du protéome théorique d'un organisme. On peut donc comparer à cette base de données les masses mesurées des peptides analysés et ainsi identifier la protéine dont ils sont issus. Des algorithmes ont spécifiquement été développés afin d'automatiser cette identification, comme par exemple Mascot (Perkins *et al.*, 1999).

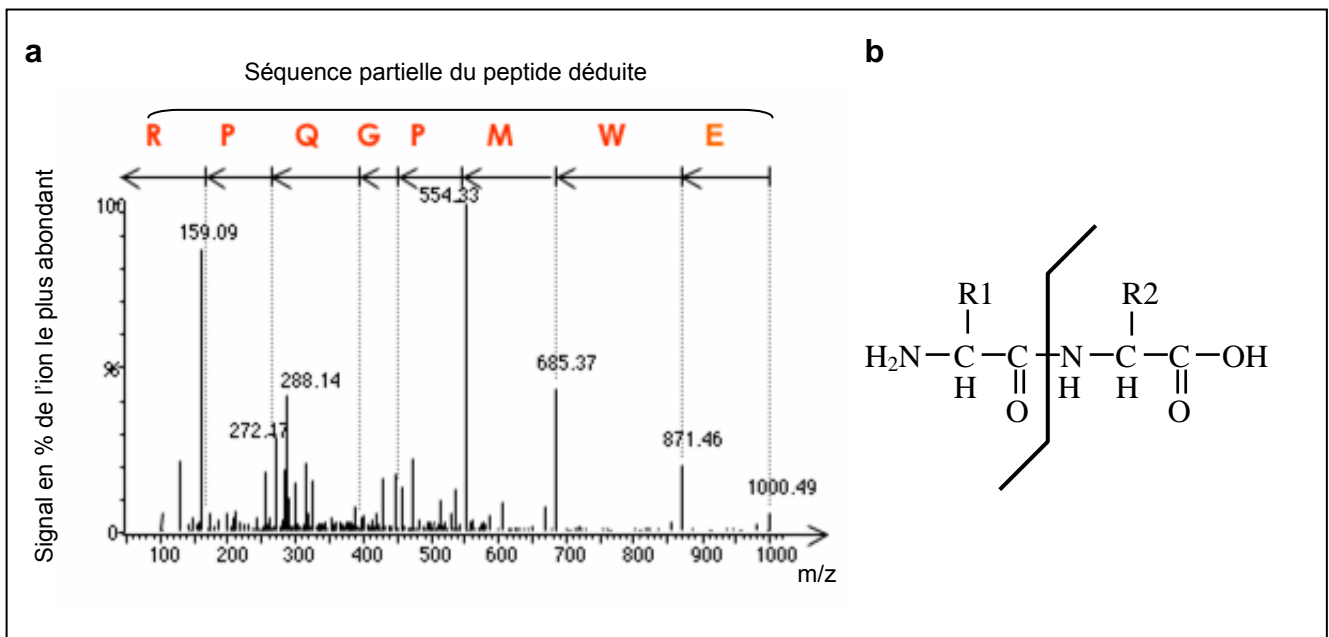
### **2.3.3. Principe de la spectrométrie de masse en tandem ou MS/MS**

La spectrométrie de masse en tandem qui combine deux analyseurs (figure 9) permet d'étudier les caractéristiques de fragmentation d'un peptide et ainsi d'accéder à des informations de séquence. Le principe général est simple : les ions formés dans l'ionisateur entrent dans le premier analyseur (MS1) qui permet de déterminer les masses moléculaires de tous les ions. Un ion de rapport  $m/z$  donné (ion précurseur) est sélectionné pour sortir de l'analyseur (MS1 est donc utilisé comme filtre d'ions). Il pénètre alors dans une cellule de collision où il est bombardé par des molécules de gaz (classiquement de l'azote ou de l'argon). Les collisions des peptides ionisés avec ces molécules vont conduire à leur fragmentation, majoritairement au niveau des liaisons peptidiques (figure 10b). Le rapport  $m/z$  de ces fragments est alors déterminé à l'aide du second analyseur (MS2). Le spectre de masse obtenu (figure 10a) permet de déduire la séquence partielle du peptide sur une dizaine d'acides aminés.





**Figure 9.** Schéma de principe d'un spectromètre de masse en tandem (MS/MS).



**Figure 10.** Spectre de masse MS/MS (a) après fractionnement d'un peptide ionisé dans la chambre de collision qui se fait préférentiellement au niveau de la liaison peptidique (b). Cela implique que la différence de masse entre deux pics adjacents correspond à la masse d'un acide aminé. On peut donc de proche en proche en déduire la séquence partielle du peptide.

### 2.3.4. Les sources d'ionisation

L'application à la biochimie du spectromètre de masse date du début des années 1980, jusqu'alors il s'agit d'un outil de référence des physiciens et des chimistes, mais le développement d'outils permettant l'ionisation de macromolécules (ionisation de type FABMS<sup>1</sup>, PDMS<sup>2</sup>, ou LSIMS<sup>3</sup> (Chapman, 1996)) ont permis son application à des problématiques biologiques. Cependant, c'est en 1988 que son champ d'application s'est considérablement élargi avec le développement de deux nouvelles techniques corrigeant les inconvénients des méthodes développées jusque là. Il s'agit de l'ionisation par électrospray ou ESI (Fenn *et al.*, 1989) et de la désorption / ionisation par laser de l'échantillon couplé à une matrice ou MALDI (Tanaka *et al.*, 1988). Le point crucial de ces techniques est l'ionisation simple (ou multiple) de tous les peptides sans qu'il y ait pour autant perte de leur intégrité. Ce sont aujourd'hui les techniques d'ionisation les plus utilisées en spectrométrie de masse appliquée aux protéines.

Le principe de l'ESI (pour ElectroSpray Ionisation) est montré figure 11. De nombreux solvants organiques (mais préférentiellement le méthanol, l'acétonitrile ou l'eau) sont compatibles avec cette technique alors que les tensioactifs, les substances non volatiles et les sels sont très mal tolérés car responsables de la formation de complexes ioniques rendant le spectre plus compliqué et diminuant la sensibilité. Cette analyse nécessite donc une étape préalable de purification des peptides avant leur ionisation (généralement sur colonne de dessalage). Les ions produits sont principalement des peptides multiprotonés avec des valeurs de  $m/z$  assez basses (inférieures à 2000) y compris pour des peptides de haute masse moléculaire. Cette source d'ions peut ainsi être couplée avec des analyseurs de gamme de masse limitée (analyseur quadripolaire ou ion trap).

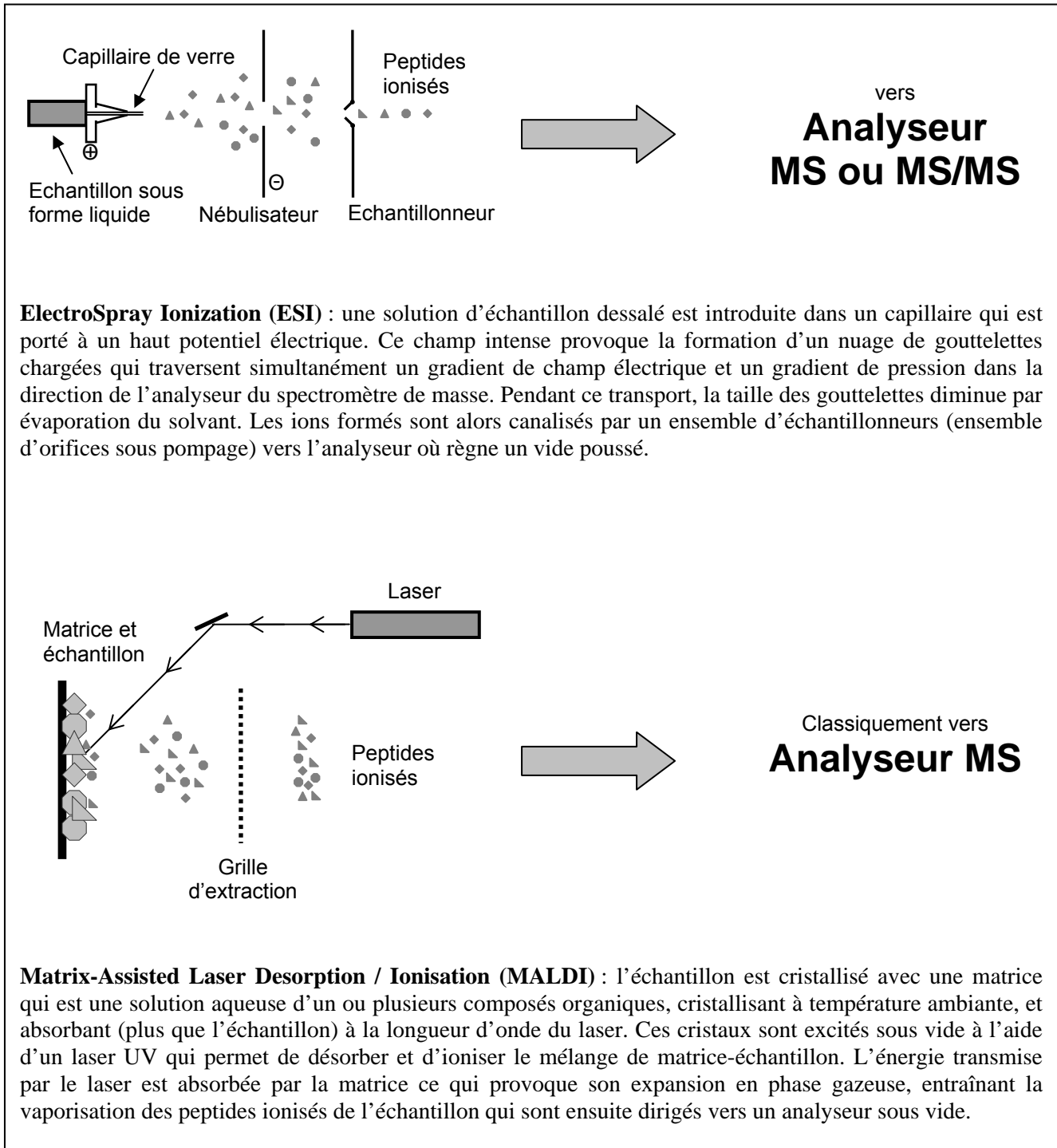
L'ionisation de type MALDI, pour Matrix-Assisted Laser Desorption / Ionisation (figure 11) concerne des peptides co-cristallisés avec une matrice organique. Il peut s'agir, par exemple, d'acide nicotinique (la plus ancienne) ou bien de dérivés de l'acide cinnamique, tel que l'acide  $\alpha$ -cyano-4-hydroxycinnamique, qui est l'une des matrices les plus utilisées. D'une part, la matrice permet la protection des peptides en absorbant l'énergie du laser pulsatif, et d'autre part, son expansion conduit à la vaporisation des peptides ionisés. Contrairement à l'ESI, cette technique est assez peu sensible à la présence de sels car les matrices présentent la caractéristique de co-cristalliser sélectivement avec les biopolymères les séparant ainsi des contaminants. Comme cette méthode d'ionisation va produire une grande variété d'ions peptidiques en même temps, avec un panel très large de masse et une charge ionique souvent unique, les valeurs de  $m/z$  mesurées vont être fortement étalées (de 600 à 300 000). Un ionisateur MALDI est donc nécessairement couplé à un analyseur n'ayant ni de limite maximum pour les masses, ni pour la quantité d'ions étudiés.

---

<sup>1</sup> FABMS pour Fast Atom Bombardment Mass Spectrometry

<sup>2</sup> PDMS pour Plasma Desorption Mass Spectrometry

<sup>3</sup> LSIMS pour Liquid Secondary Ion Mass Spectrometry



**ElectroSpray Ionization (ESI)** : une solution d'échantillon dessalé est introduite dans un capillaire qui est porté à un haut potentiel électrique. Ce champ intense provoque la formation d'un nuage de gouttelettes chargées qui traversent simultanément un gradient de champ électrique et un gradient de pression dans la direction de l'analyseur du spectromètre de masse. Pendant ce transport, la taille des gouttelettes diminue par évaporation du solvant. Les ions formés sont alors canalisés par un ensemble d'échantillonneurs (ensemble d'orifices sous pompage) vers l'analyseur où règne un vide poussé.

**Matrix-Assisted Laser Desorption / Ionisation (MALDI)** : l'échantillon est cristallisé avec une matrice qui est une solution aqueuse d'un ou plusieurs composés organiques, cristallisant à température ambiante, et absorbant (plus que l'échantillon) à la longueur d'onde du laser. Ces cristaux sont excités sous vide à l'aide d'un laser UV qui permet de désorber et d'ioniser le mélange de matrice-échantillon. L'énergie transmise par le laser est absorbée par la matrice ce qui provoque son expansion en phase gazeuse, entraînant la vaporisation des peptides ionisés de l'échantillon qui sont ensuite dirigés vers un analyseur sous vide.

**Figure 11.** Schéma de principe de deux technologies permettant l'ionisation et le passage en phase gazeuse des peptides : l'ionisation par électrospray (ElectroSpray Ionisation ou ESI), et la désorption / ionisation par laser (Matrix-Assited Laser Desorption/Ionisation ou MALDI).

### 2.3.5. Les analyseurs et spectromètres de masse

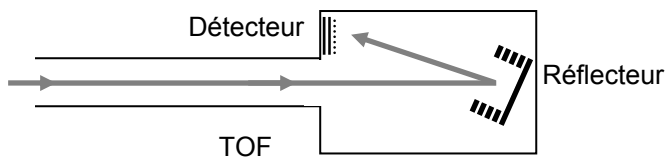
Le principe de fonctionnement des principaux analyseurs de masse est représenté figure 12.

Le spectromètre MALDI-TOF. Cet appareil associe une source d'ionisation de type MALDI à un analyseur de type Temps de Vol (Time Of Flight ou TOF). Il présente l'avantage d'avoir une gamme de masse pratiquement illimitée (supérieure à 1 MDa) car les ions ne sont déviés par aucun champ magnétique ou électrique. Ceci permet donc d'obtenir les masses de l'ensemble des ions provenant d'un échantillon complexe. De plus, toutes les étapes, de la préparation de l'échantillon au traitement du spectre de masse, peuvent être entièrement automatisées. Ce qui permet, dans un contexte d'analyses post-génomiques systématiques, de faire de cet analyseur un outil de choix pour les études de protéomique à grande échelle. La gamme de masses mesurées va usuellement de 600 Da à 300 kDa, avec une limite de sensibilité de détection de l'ordre d'une centaine de femtomoles et avec une précision de mesure de plus ou moins 50 ppm (Karas, 1996).

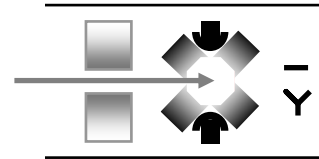
Le spectromètre ESI-MS/MS. Il s'agit du couplage d'une ionisation de type ESI avec un triple quadripôle. Grâce à la présence de deux analyseurs et d'une chambre de collision, il permet de fractionner un peptide de  $m/z$  précis, pour en déduire la séquence en acides aminés. Sa plage d'analyse se situe entre 10 et 4000 Da, avec une sensibilité de 80 femtomoles qui est cependant moindre pour les ions de masse supérieure à 2000 Da. La précision est de l'ordre de 100 ppm.

Le spectromètre ESI-IT (Ion Trap). L'analyseur à trappe ionique est classiquement couplé à la source électrospray pour produire des ions qui sont ensuite piégés dans la trappe par un champ électromagnétique et éjectés en fonction de leur  $m/z$  pour parvenir au détecteur. Il est possible de garder un ion dans la trappe. Cet ion parent peut alors y être fragmenté par collision avec des atomes d'hélium et les ions fils formés sont alors éjectés de la trappe. L'appareil fonctionne là en MS/MS. Bien que la précision de masse soit moins bonne qu'en MALDI-TOF, l'identification d'une protéine par ESI-IT est plus sûre grâce à l'analyse des ions de fragmentation. En contrepartie, elle prend plus de temps aussi bien pour obtenir les données que pour les traiter. La gamme de masse analysée se situe entre 15 et 4000 Da, avec une sensibilité de 80 femtomoles, pour une précision de l'ordre de 100 ppm.

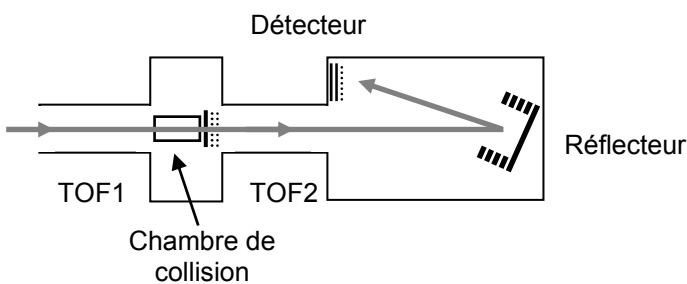
De nouveaux types d'analyseur de masse apparaissent depuis une dizaine d'années. Leur développement est une réponse pour remédier aux limitations des analyseurs existant, ainsi qu'une



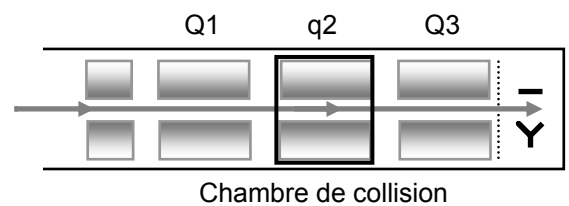
**Time-Of-Flight (TOF)** : les peptides ionisés entrent dans l'analyseur avec une même énergie cinétique. Le temps mis pour arriver au détecteur est donc fonction de leur masse et de leur charge.



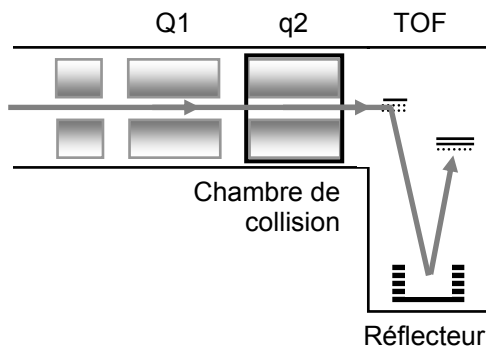
**Ion trap (IT)** : les fragments de peptides ionisés sont stockés dans un champ électromagnétique tridimensionnel et sont éjectés sélectivement selon leur rapport  $m/z$ .



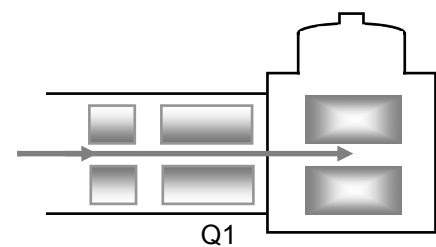
**TOF-TOF** : un ion peptidique est sélectionné dans le premier TOF, fractionné dans la chambre de collision, et les ions fragmentés sont analysés dans le second TOF.



**Triple quadripôle** : les ions d'un certain  $m/z$  sont sélectionnés dans le premier quadripôle (Q1), et fractionnés dans le second par collision avec des molécules de gaz inerte (q2). Le troisième quadripôle (Q3) permet de séparer les peptides fractionnés et d'obtenir leur masse.



**Q-TOF** : il s'agit d'un appareil hybride entre un triple quadripôle (sélection des ions d'un certain  $m/z$  puis fragmentation) couplé à un analyseur TOF.



**Résonance cyclotronique ionique et transformée de Fourier (FT-ICR)** : les peptides ionisés sont emprisonnés sous vide dans un champ magnétique où l'analyse de leur fréquence de rotation permet de déterminer leur  $m/z$ .

**Figure 12.** Schéma de principe des spectromètres de masse les plus courants. Tous à l'exception de l'analyseur TOF permettent des analyses MS/MS

réponse à des besoins techniques particuliers. La plupart sont des analyseurs hybrides, polyvalents, et compatibles avec une analyse MS/MS.

Le TOF-TOF. Développé pour remédier à la limite de l'analyseur TOF concernant l'obtention de spectres MS/MS. Ici un analyseur TOF permet la sélection de l'ion précurseur qui sera fragmenté dans une chambre de collision, puis la masse des ions obtenus est analysée par le second TOF.

Le Q-TOF. L'application principale du spectromètre de masse Q-TOF est le séquençage rapide des peptides, ainsi que l'analyse des modifications post-traductionnelles (Morris *et al.*, 1996). Cependant la source d'ionisation demeure un ESI, et nécessite une séparation préalable des peptides pour être compatible avec une analyse à haut débit.

Le FT-ICR. Le spectromètre de masse à résonance cyclotronique ionique et transformée de Fourier (Amster, 1996) est actuellement l'appareil le plus versatile, adaptable à une large variété d'analyses MS et MS/MS. Il s'agit d'une trappe ionique où les ions sont piégés dans un champ magnétique et la mesure de masse se fait par mesure de leur fréquence de rotation cyclotronique. Il s'agit actuellement de la méthode de spectrométrie ayant la plus large gamme de masse analysable (valeur de  $m/z$  de 30 à plus de 300 000), la meilleure résolution et la plus forte précision (2 ppm) pour des quantités de protéine injectées de 100 femtomoles. Cette précision en masse et la capacité à piéger des ions de haut poids moléculaire permettent également de s'intéresser à l'étude de complexes non covalents en conditions natives par ESI combinée à la FT-ICR.

Une réelle flexibilité dans l'analyse de masse est désormais atteinte avec ces appareils très polyvalents et très sensibles (particulièrement le FT-ICR). La diffusion de ce type d'instruments et leur disponibilité croissante permettent d'envisager des stratégies d'analyses à la fois sur une très large échelle (plusieurs milliers de peptides analysés par 24 heures) mais avec des précisions compatibles avec l'analyse des modifications post-traductionnelles ou des complexes protéiques.

## **2.4. Améliorations techniques et haut débit**

### **2.4.1. Amélioration de la protéomique basée sur la 2D-PAGE**

Les limitations de la protéomique basée sur l'électrophorèse bidimensionnelle se situaient, jusqu'à la fin des années 90, au niveau du nombre de protéines résolues sur les gels 2D-PAGE, ainsi que sur la reproductibilité d'un gel à l'autre. Des travaux ont été entrepris pour optimiser chaque étape du processus.



Concernant l'extraction protéique, l'un des biais qui apparaît dans des conditions compatibles avec l'électrophorèse bidimensionnelle est la sous-représentation des protéines membranaires, dans la mesure où elles sont difficiles à solubiliser sans interférer avec leur point isoélectrique. C'est pourquoi de nouveaux détergents, non ioniques ou zwitterioniques, ont été développés. Des études comparatives ont été menées d'où il ressort que les détergents de la famille des sulfobétaines (dont le SB 3-10 et l'ASB14), en combinaison avec de l'urée et de la thiourée, sont les plus efficaces pour solubiliser des protéines très hydrophobes comme les protéines transmembranaires (Molloy, 2000). Cependant, il est important de noter que l'efficacité d'un composé plutôt qu'un autre dépend à la fois de la nature des protéines et des constituants lipidiques de l'échantillon. Il ne s'agit donc pas d'utiliser un détergent universel pouvant extraire toutes les protéines membranaires, mais plutôt d'effectuer plusieurs extractions avec différents détergents, de manière à élargir le panel de protéines solubilisées (Luche *et al.*, 2003). Cette stratégie d'extraction protéique par traitements successifs n'est pas applicable uniquement aux protéines membranaires mais peut être utilisée pour solubiliser n'importe quelle protéine difficile à extraire, comme c'est le cas, par exemple, des protéines de paroi.

L'une des évolutions les plus significatives de la technologie 2D-PAGE fut le développement de gel de première dimension à gradient de pH immobilisé (IPG). Il s'agit d'un gel où sont copolymérisés des dérivés d'acrylamide comportant des groupements à effet tampon avec une matrice de polyacrylamide. Les groupements chargés, participant à la mise en place du gradient de pH, sont fixés au gel et permettent une focalisation précise des protéines tout en évitant les problèmes de dérive associés aux techniques plus anciennes (dus à la migration des ampholytes). La commercialisation de ce type de gel d'IEF par des sociétés telles que BioRad ou Amersham Biosciences a permis l'essor de la technique 2D-PAGE autorisant une haute reproductibilité d'un gel à l'autre (Molloy, 2000). De plus, ce type de gel autorise le dépôt de quantité importante de protéines (jusqu'à 5 mg pour des gels de 18 cm) permettant ainsi de visualiser des protéines peu abondantes (car peu exprimées ou mal extraites). Dans le but de séparer des spots protéiques avec une très haute résolution, des gels IPG ont été développés avec des gammes de pH réduites. Ces gels permettent de séparer les protéines sur quelques unités de pH (régions acide, neutre ou basique), ou bien désormais sur une seule unité de pH disponible sur l'ensemble de la gamme (de pH=3 à pH=11). L'analyse d'un même échantillon protéique sur plusieurs gels avec des unités de pH contiguës permet la reconstruction de gel « virtuel » avec un étalement de la gamme de pH sur plus d'un mètre (Gorg *et al.*, 2004). Ce procédé a été utilisé pour obtenir une carte 2D à haute résolution du protéome de *Leishmania major* où l'assemblage de cinq gels de pH 4 à 9 a permis la séparation de près de 3700 spots protéiques (Drummelsmith *et al.*, 2003).





Concernant la seconde dimension, la disponibilité dans le commerce de gel pré-coulé permet là aussi une standardisation du protocole et d'accroître la reproductibilité. De nouveaux types de coloration ont également accru la résolution des spots protéiques, il s'agit de fluorophore (Sypro Orange, Sypro Red, Sypro Ruby) avec une sensibilité de coloration presque équivalente à celle de l'argent (0,5 ng de protéine) et une linéarité d'intensité, une reproductibilité, une facilité et une rapidité proche de la coloration au bleu de Coomassie (Rabilloud *et al.*, 2001)

L'intégration de ces technologies a permis le développement de plateformes entièrement automatisées qui prennent en charge un échantillon protéique de l'extraction à l'analyse et l'identification des spots protéiques. Le gain de temps qui en découle, la très bonne reproductibilité de ces analyses permettent donc d'envisager une analyse protéomique à haut débit bien que basée sur des électrophorèses bidimensionnelles.

#### **2.4.2. Vers une protéomique sans électrophorèse**

Bien que les techniques ont fortement été améliorées pour une meilleure reproductibilité et un meilleur isolement des protéines, l'analyse protéomique par gel 2D-PAGE se heurte encore aux limitations associées à la séparation des protéines sur gel. D'une part les protéines faiblement ou moyennement exprimées sont en effet rarement détectées (Gygi *et al.*, 2000) en raison de leur concentration se situant aux limites de détection des colorations. D'autre part, elles sont souvent masquées par des protéines plus fortement exprimées. De plus, les techniques d'extraction biaisent les analyses en provoquant une forte sous-représentation des protéines les moins solubles (comme les protéines membranaires) et des protéines basiques, difficiles à séparer.

Pour palier ces défauts, une stratégie d'identification s'affranchissant totalement de la séparation 2D-PAGE a été développée. Il s'agit d'une colonne de chromatographie liquide (LC) en phase inverse directement couplée en ligne à un spectromètre de masse de type ESI-MS/MS. L'analyse porte sur un échantillon protéique qui est digéré par la trypsine, séparé par chromatographie liquide puis analysé par le spectromètre de masse. L'évolution récente de cette technologie repose sur l'adjonction d'une chromatographie en phase liquide bidimensionnelle, qui permet une séparation plus précise des peptides. Cette technologie, dite « MudPIT » pour Multidimensional Protein Identification Technology ou « Shotgun », autorise l'analyse directe d'extraits protéiques bien plus complexes que ne le permet la LC-MS/MS comme par exemple des lysats cellulaires complets (Link *et al.*, 1999). La première dimension de la colonne HPLC est une résine échangeuse d'ions qui capte les peptides digérés, puis un gradient de sel désorbe séquentiellement ces peptides vers une colonne en phase inverse. Un gradient de solvants élue enfin les peptides vers l'analyseur de masse. L'intégralité de l'échantillon peut ainsi être séparée et analysée sur une durée de 6 à 36



heures avec une gamme d'analyses permettant de discriminer entre des peptides dont la différence de concentration est de l'ordre de cinq log. Le peu de manipulation nécessaire en routine, ainsi que la grande sensibilité de ce type de stratégie permet l'utilisation de cette technologie pour des analyses à grande échelle ou haut débit, comme par exemple l'identification du protéome de *Saccharomyces cerevisiae*, où 1484 protéines ont pu être identifiées dans un seul échantillon de lysat cellulaire (Washburn *et al.*, 2001).

### **2.4.3. Vers de nouvelles technologies et de nouveaux concepts**

La protéomique est intimement liée à quatre domaines : la génomique, la biochimie des protéines, la spectrométrie de masse et la bioinformatique. Une seule technique n'est pas suffisante pour résoudre toutes les problématiques de la protéomique, mais l'ensemble des progrès effectués dans ces quatre domaines permet d'envisager de nouvelles approches pour résoudre ces problématiques. Ces progrès peuvent reposer à la fois sur le développement de nouvelles technologies et sur la proposition de nouveaux concepts. Récemment, de nouvelles méthodologies basées sur la spectrométrie de masse ont permis de caractériser des cellules ou des tissus. Ces échantillons biologiques sont directement soumis au processus d'ionisation, par exposition au laser d'un ionisateur MALDI, chaque impulsion du laser pouvant être dirigée contre une région définie du tissu. L'imagerie par spectrométrie de masse et l'obtention de profil directement sur un échantillon biologique permettent d'associer à une région donnée d'un tissu ou d'un ensemble de cellules un profil peptidique caractéristique (Chaurand *et al.*, 2004). Ces profils sont constitués de 500 à 1000 peptides dont la masse est comprise entre 2 et 70 kDa. Ces informations peuvent ensuite être corrélées avec des changements dans les niveaux d'expression de ces marqueurs en association avec des maladies ou avec des réponses à des traitements. Ces techniques permettent de caractériser de multiples tissus (particulièrement des cellules cancéreuses), ainsi que la réponse de tumeurs à des thérapies, elles peuvent aussi être employées pour l'identification et la caractérisation de pathogènes. Ces techniques ont été appliquées, en particulier, au suivi de cancers cérébraux ou pulmonaires, en permettant la classification des cellules cancéreuses et l'évaluation de leur progression (Yanagisawa *et al.*, 2003 ; Schwartz *et al.*, 2004).

De nouveaux concepts ont également été proposés pour prendre en compte le contexte dans lequel se situe une analyse protéomique. C'est par exemple le cas de la notion de « parasito-protéomique » où viennent s'intégrer l'analyse du protéome d'un parasite avec l'analyse du protéome de la cellule hôte (Biron *et al.*, 2005). Les éléments de réponse de la cellule hôte peuvent ainsi être caractérisés et définis comme spécifiques ou non à l'invasion par un parasite donné. Ce type d'analyse permet



de révéler fondamentalement les protéines du parasite dont la fonction est l'interaction directe avec la cellule-hôte et qui conduisent à des modifications de son métabolisme ou de sa physiologie (dérégulation du cycle cellulaire, modification des transports intracellulaires, réorganisation du cytosquelette). Ainsi, cette approche, en plus de caractériser plus précisément le réseau d'interactions entre l'hôte et le pathogène, permet de définir de nouvelles voies dans la recherche de traitements et de vaccins.

## **2.5. Les levures, organismes modèles pour la protéomique des eucaryotes**

En 1996, le génome de la levure *Saccharomyces cerevisiae* fut le premier génome eucaryote séquencé (Goffeau *et al.*, 1996), ouvrant l'étude de l'expression de ces 6234 ORF<sup>1</sup>. Les premières études du protéome de cette levure ont montré une différence significative entre le taux de transcription d'un gène et le niveau d'expression d'une protéine (Griffin *et al.*, 2002). Ceci met en évidence qu'une description solide de l'expression protéique nécessite une analyse protéomique et ne peut être déduite de la seule analyse des transcrits.

### **2.5.1. Description et localisation des protéines**

Le protéome de *Saccharomyces cerevisiae* est considéré comme un modèle pour l'analyse des protéomes eucaryotes. Son analyse descriptive a apporté une vision globale des protéines exprimées. Parmi les premières études descriptives, la plus complète a permis d'identifier plus de 600 spots protéiques, correspondant à près de 250 protéines (Boucherie *et al.*, 1996). En outre, des travaux ont permis de valider le transfert des techniques de protéomique « classique » jusqu'alors développées pour l'analyse des protéomes de procaryotes (Gygi *et al.*, 2000). C'est aussi sur ce protéome que les techniques d'analyse protéomique à haut débit (LC-MS/MS, MudPIT) ont été développées, validées et comparées (Washburn *et al.*, 2001 ; Peng *et al.*, 2003). Par exemple, l'étude de Peng et de ses collaborateurs a permis d'obtenir, en une seule analyse et sur un unique extrait protéique, plus de 162 000 spectres MS/MS, correspondant à 26 815 peptides, permettant l'identification de 1 504 protéines.

L'étude comparée des cartes bidimensionnelles ou des profils d'expression protéique a permis d'établir des différences d'expression protéique vues au niveau de l'ensemble du protéome, apportant un éclairage sur l'adaptation physiologique et métabolique entre deux conditions (réponse à la déshydratation (Blomberg, 1997), adaptation à des milieux de culture différents (Kolkman *et*

---

<sup>1</sup> ORF pour Open Reading Frame, ou cadre ouvert de lecture. C'est la région de l'ADN qui sépare deux codons STOP



*al.*, 2005), privation en acide aminé (Yin *et al.*, 2004), adaptation de souches d'intérêt agro-alimentaire (Trabalzini *et al.*, 2003)). De même, la protéomique comparative permet de mesurer l'impact toxicologique de composés chimiques au niveau de l'ensemble des protéines de la levure. Par exemple, l'analyse des variations des niveaux d'expression protéique en réponse à un xénobiotique (un herbicide : l'acide 2,4-dichlorophénoxyacétique) apporte des informations à la fois sur les voies de détoxification activées mais aussi sur les dérèglements cellulaires induits par le composé (Teixeira *et al.*, 2005).

Une autre donnée importante pour la compréhension du rôle des protéines est l'information sur leur localisation dans les microenvironnements cellulaires que constituent les organelles des cellules eucaryotes (Simpson et Pepperkok, 2003). Cette information de localisation peut être soit directement obtenue par des techniques de marquage des protéines et de localisation directe dans la cellule ou bien indirectement pour l'analyse des sous-protéomes cellulaires. Bien que des études à petite échelle ont été effectuées précédemment, c'est en 2003 que Huh et ses collaborateurs ont réalisé la première étude de localisation des protéines à l'échelle de tout le protéome, en fusionnant 75% des ORF de *S. cerevisiae* avec la séquence codante d'une GFP. Les résultats de localisation effectuée sur cellule vivante concordent à 80% dans le cas de protéines préalablement localisées. Un sous-protéome de la levure a particulièrement été étudié, il s'agit de celui de la mitochondrie, pour lequel 750 protéines différentes ont été identifiées grâce à plusieurs stratégies de séparation et d'identification (2D-PAGE et 1D-PAGE couplées à de la spectrométrie de masse MALDI-TOF, LC-MS/MS et MudPIT). Ces travaux ont mis en évidence l'ensemble des protéines mitochondriales impliquées dans la chaîne respiratoire et dans le cycle de Krebs. En comparant ces résultats avec les protéines mitochondriales décrites dans la littérature, les protéines identifiées représenteraient, d'après les auteurs, 90% des protéines mitochondriales totales (Sickmann *et al.*, 2003).

### **2.5.2. Réseaux d'interactions et structures des protéines**

Les réseaux d'interactions entre protéines peuvent être appréhendés par une analyse directe des complexes protéiques par spectrométrie de masse (Deshaies *et al.*, 2002) ou bien par des techniques de double hybride (Uetz *et al.*, 2000). L'analyse de ces réseaux permet de décrire les acteurs impliqués dans l'activité biologique de la cellule et dans sa régulation, mais aussi d'assigner un rôle possible à une protéine à fonction inconnue selon les partenaires avec lesquels elle interagit. (Schwikowski *et al.*, 2000). L'ensemble des interactions protéine-protéine mises en évidence par les différentes études à large échelle se chiffre à environ 15 000 chez *S. cerevisiae*. Cependant, ces techniques génèrent un très grand nombre de faux positifs que certains auteurs chiffrent jusqu'à





hauteur de 50% (von Mering *et al.*, 2002). Cela implique la nécessité d'établir plusieurs niveaux de contrôle pour confirmer ces réseaux d'interaction. En effet, au vu du trop grand nombre de partenaires impliqués, la validation expérimentale individuelle de chaque interaction n'est pas envisageable. Des techniques alternatives ont alors été proposées, parmi lesquelles la reconnaissance des domaines d'interaction par phage display ou bien par analyse bioinformatique (Sprinzak *et al.*, 2003 ; Causier, 2004). Le développement récent d'algorithmes permettant de proposer une fonction, selon ses partenaires, à une protéine, révèle le formidable potentiel de l'exploitation de ces cartes d'interactions mais aussi la nécessité de pouvoir s'assurer de leur fidélité au modèle vivant (Nabieva *et al.*, 2005).

### 2.5.3. Analyses de pathogénicité

Les études protéomiques de *S. cerevisiae* ont apporté, comme nous l'avons vu, des techniques et des descriptions de profils d'expression protéique d'un eucaryote simple. Ces données peuvent être exploitées, entre autres, pour l'analyse des facteurs de pathogénicité de la levure dimorphique *Candida albicans*. En effet, ce commensal, dont le génome a été séquencé en 2004 (Jones *et al.*, 2004), est un pathogène d'intérêt médical majeur. Il se présente sous deux formes de vie différentes, levure ou filamenteuse. Quelles peuvent être les protéines spécifiques de chaque état, et comment est contrôlé le passage d'une phase à l'autre ? L'analyse comparative des protéomes des deux formes a mis en évidence très peu de protéines différentes dans les fractions solubles des deux stades (Choi *et al.*, 2003). Seules 12 protéines ont été spécifiquement identifiées chez la forme filamenteuse (Hernandez *et al.*, 2004), il s'agit d'enzymes impliquées dans des voies métaboliques (biosynthèse des acides aminés, métabolisme à un carbone, alcool déshydrogénase), reflétant plutôt des différences dans les milieux de cultures utilisés, plutôt qu'entre les deux formes.

Il s'agit également d'un microorganisme d'intérêt pour appréhender les relations hôte-pathogène dans un cadre moins complexe que celui d'un parasite intracellulaire. Ceci a permis d'identifier les adhésines de *C. albicans* impliquées dans la reconnaissance et la liaison aux protéines de la cellule-hôte (principalement au plasminogène) (Jong *et al.*, 2003). De plus, des immunomarquages d'électrophorèses bidimensionnelles par des sérums de patients donnent des indications sur la réponse sérologique de l'hôte lors d'une infection à *Candida*. Ces travaux mettent en évidence que les protéines majoritairement immunogènes sont des protéines glycosylées ainsi que des HSP (Fernandez-Arenas *et al.*, 2004). La recherche de protéines induisant la production d'anticorps protecteurs est actuellement en cours dans le but d'obtenir un vaccin.

Les traitements sont actuellement à améliorer pour traiter efficacement les candidoses. L'analyse protéomique constitue un axe de recherche à la fois pour comprendre les modes d'actions des



produits actuellement utilisés (Bruneau *et al.*, 2003) mais aussi pour la recherche de nouvelles cibles comme par exemple la recherche de nouvelles protéines de paroi (Urban *et al.*, 2003 ; De Groot *et al.*, 2004) pour le développement de nouveaux agents thérapeutiques. De plus, les analyses de protéomique comparative entre *S. cerevisiae* et *C. albicans* font ressortir des différences physiologiques ou métaboliques qui pourraient être exploitées à des fins thérapeutiques (Yin *et al.*, 2004).

## **2.6. Analyse protéomique de parasites unicellulaires**

### **2.6.1. Spécificités et difficultés de l'analyse protéomique chez les parasites**

C'est avec la disponibilité croissante des données de séquences génomiques, qu'elles soient complètes ou bien partielles, qu'a véritablement débuté l'analyse globale des profils d'expressions protéiques conduisant à une meilleure description de la physiologie et du métabolisme des parasites unicellulaires. Cependant, ces projets se heurtent à de nombreuses difficultés, particulièrement dues au mode de vie de ces parasites. En effet, la majorité d'entre eux sont des organismes intracellulaires obligatoires, qui nécessitent d'être séparés des débris des cellules-hôtes avant d'être étudiés. Cette étape de purification doit être assez drastique pour éliminer les contaminations de la cellule-hôte mais suffisamment ménagée pour ne pas endommager les protéines du parasite, ce qui peut constituer un véritable défi pour les stades parasitaires les plus fragiles. Cela implique également que des niveaux supplémentaires de contrôle doivent être mis en place pour vérifier qu'une protéine identifiée soit bien d'origine parasitaire et non pas une contamination de la cellule-hôte (vérification des purifications en microscopie, recherche de marqueurs cellulaires dans les extraits protéiques). De plus, ces organismes présentent souvent un cycle de vie complexe, avec plusieurs stades de développement et parfois de multiples hôtes, multipliant ainsi le nombre d'analyses à effectuer pour avoir une vision générale de l'expression protéique au cours du cycle de vie. Cependant, certains stades de développement peuvent être difficiles à obtenir pour certains parasites, particulièrement si celui-ci n'est pas cultivable *in vitro*, ou bien si le stade d'intérêt ne peut être différencié en culture cellulaire, comme c'est le cas, par exemple, des sporozoïtes de *Plasmodium* uniquement présents dans les glandes salivaires de l'anophèle. L'obtention de ce matériel biologique nécessite donc la production de parasites chez l'organisme hôte. Cela implique que la récolte du parasite nécessitera, dans la majorité des cas, des dissections précises pouvant conduire à diminuer drastiquement la quantité de matériel biologique disponible. Chez d'autres parasites, la présence simultanée de tous les stades de développement dans la cellule-hôte, comme



c'est le cas pour les microsporidies, nécessite de résoudre la problématique de la séparation des stades avant d'espérer avoir une image dynamique de leur protéome.

Il faut, de plus, conserver à l'esprit que l'analyse protéomique présente deux biais majeurs. Le premier est la très faible représentation des protéines membranaires qui sont difficilement solubilisables. Ces protéines membranaires, parmi lesquelles se trouvent des protéines de transduction de signal ou des récepteurs, ne peuvent être appréhendées que par des techniques spécifiques d'extraction qui ne sont pas directement compatibles avec les analyses à haut débit. Le second biais provient du fait que la majorité des analyses sur parasites sont effectuées sur des stades cultivés *in vitro*, et ne sont donc pas soumis aux influences de l'organisme hôte, relativisant la représentativité d'un tel profil d'expression protéique. Ces doutes pourraient être levés en comparant les protéomes de stades cultivés *in vitro* avec des stades prélevés *in vivo* sur des organismes parasités, la faible quantité de matériel n'étant pas limitatif dans le cas d'analyse de type MudPIT.

### **2.6.2. Analyses descriptives du protéome total de parasites**

Les premières analyses des protéomes parasitaires se sont dans un premier temps orientées vers l'optimisation des techniques de protéomique classique (extraction et électrophorèse bidimensionnelle). Ainsi, chez *Leishmania major*, un kinétoplaste dont le séquençage du génome vient très récemment d'être publié (Ivens *et al.*, 2005), les analyses protéomiques ont été principalement orientées sur l'optimisation des techniques d'extraction (Acestor *et al.*, 2002) et d'électrophorèse bidimensionnelle. Des travaux ont montré qu'il était possible de séparer près de 3700 spots protéiques sur des gels d'électrophorèse à gammes de pH réduites (El Fakhry *et al.*, 2002). Cependant, lors de ces travaux, l'absence de données de séquence a limité l'identification protéique à une dizaine de spots uniquement. De même, les autres études descriptives, à l'exception de celle de *Plasmodium*, sont actuellement limitées à l'établissement de cartes protéiques préliminaires comme pour les stades tachyzoïtes de *Toxoplasma gondii*, pour lesquels 500-600 spots ont été séparés mais seulement 71 ont été identifiés (Cohen *et al.*, 2002).

Mais l'accroissement de la disponibilité des données de séquence de génomes parasitaires a rendu l'application des techniques à plus haut débit (MudPIT, LC-MS/MS) possible pour l'analyse des protéomes de parasite. Ces techniques présentent l'avantage, au delà du gain de temps de l'analyse en routine, d'éviter la perte de matériel associée aux techniques par 2D-PAGE, tout en étant beaucoup plus sensibles et, par conséquent, d'être moins consommatrices d'extrait protéique. Actuellement, le protéome le plus complètement décrit est celui de l'apicomplexe *Plasmodium*

	<i>Organisme</i>	<i>Génome</i>	<i>Stades ou organites analysés</i>	<i>Protéines totales identifiées (spécifiques du stade)</i>	<i>Technique</i>	<i>Références</i>
A P I C O M P L E X E S	<i>Plasmodium falciparum</i>	(Gardner et al., 2002)	sporozoïte	1049 (513)	MudPit	(Florens et al., 2002)
			mérozoïte	839 (204)		
			trophozoïte	1036 (286)		
			gamétocyte	1147 (376)		
			trophozoïte	714 (226)		
	<i>Eimeria tenella</i>	en cours	sporozoïte	28	2D-PAGE MALDI-TOF, LC-MS/MS	(de Venevelles et al., 2004)
			micronème	3	2D-PAGE MALDI-TOF, CAF-MALDI-MS	(Bromley et al., 2003)
			tachyzoïte	71	2D-PAGE MALDI-TOF	(Cohen et al., 2002)
			micronème granule dense rhoptrie	71	2D-PAGE MALDI-TOF, MudPIT	(Zhou et al., 2005)
			tachyzoïte	20	2D-PAGE MALDI-TOF	(Lee et al., 2003)
K I N E T O P L A S T E S	<i>Leishmania major</i>	(Ivens et al., 2005)	tous stades	49	2D-PAGE MALDI-TOF	(Bente et al., 2003)
			tous stades	47	Capillary HPLC-QTOF	(Nugent et al., 2004)
	<i>Trypanosoma cruzi</i>	(El Sayed et al., 2005)	amastigote	1871 (691)	LC-MS/MS	(Atwood et al., 2005)
			trypomastigote métacyclique trypomastigote	1486 (582) 2339 (969)		
			épimastigote	1861 (732)		
			épimastigote	45	2D-PAGE MALDI-TOF	(Parodi-Talice et al., 2004)
ENTAMOEBIDES	<i>Entamoeba histolytica</i>	(Loftus et al., 2005)	phagosome	85	LC-MS/MS	(Okada et al., 2005)
MICROSPORIDIIES	<i>Encephalitozoon cuniculi</i>	(Katinka et al., 2001)	spores	-	2D-PAGE	(Moura et Visvesvara, 2001)

*Autres programmes de séquençage de génomes de parasites unicellulaires*

**APICOMPLEXES**

*Theileria parva* (Gardner et al., 2005)  
*Theileria annulata* (Pain et al., 2005)  
*Plasmodium berghei* (Hall et al., 2005)  
*Plasmodium chabaudi* (Hall et al., 2005)  
*Cryptosporidium hominis* (Xu et al., 2004)  
*Cryptosporidium parvum* (Abrahamsen et al., 2004)  
*Babesia bovis* en cours

**KINETOPLASTES**

*Trypanosoma brucei* (Berriman et al., 2005)

**MICROSPORIDIA**

*Antonospora locustae* en cours  
*Glugea aterinae* en cours  
*Nosema bombycis* en cours

**AUTRES**

*Giardia lamblia* en cours  
*Trichomonas vaginalis* en cours

**Tableau 2.** Etat des lieux des programmes de séquençage des génomes et analyses protéomiques descriptives et comparatives de stades ou d'organites de parasites unicellulaires.

*falciparum*. L'établissement du profil d'expression des quatre stades majeurs du cycle de développement (Johnson *et al.*, 2004) (tableau 2) grâce aux techniques à haut débit (Mudpit et LC-MS/MS) permet une analyse à large échelle avec très peu de matériel (particulièrement intéressant pour l'étude des sporozoïtes qui proviennent de microdissections manuelles de glandes salivaires d'anophèle). L'analyse descriptive de l'ensemble des protéines exprimées d'un stade parasitaire permet d'avoir une idée de la physiologie et des métabolismes en cours chez ce stade. Par exemple, la majorité des protéines identifiées chez le stade sporozoïte sont associées aux organites apicaux, alors que chez les mérozoïtes, trophozoïtes et gamétocytes, prédominent les protéines impliquées dans les voies métaboliques, ainsi que dans la synthèse et le devenir des protéines. Cette analyse descriptive permet aussi d'effectuer une analyse post-génomique en validant l'annotation du génome. En effet, dans l'analyse totale du protéome de *P. falciparum* plus de 51% des protéines identifiées sont annotées comme hypothétiques, leur expression valide donc bien comme séquences codantes les ORF identifiés par les algorithmes de détection de gènes (Florens *et al.*, 2002) et nous renseigne sur leur stade d'expression.

### **2.6.3. Analyses comparatives : dynamique des protéines au cours du cycle de vie du parasite**

L'analyse comparative des profils d'expression obtenus pour les divers stades de développement d'un parasite est une véritable pierre de rosette pour éclaircir la biologie de ces pathogènes. En effet, cette analyse apporte des réponses quant aux régulations biologiques, à la physiologie et à la morphogenèse propre à chaque stade. En effet, si l'on reprend l'analyse du protéome des stades de *P. falciparum* (Florens *et al.*, 2002), la comparaison met en évidence que de très nombreuses protéines sont spécifiques d'un stade particulier (jusqu'à 50% de celles du sporozoïte). Cependant, les profils d'expression de certaines protéines ne sont pas ceux attendus comme par exemple pour les protéines de surfaces Var et Rif. Ces protéines sont connues pour permettre l'échappement des stades sanguins au système immunitaire, mais leurs profils d'expression montrent qu'elles sont bien plus largement exprimées, ce qui suggère qu'elles puissent avoir d'autres fonctions. La récente analyse du protéome complet de *T. cruzi* (Atwood *et al.*, 2005) a permis d'établir les profils généraux d'expression protéique des quatre stades et de définir les variations majeures d'un stade à l'autre (adaptation du métabolisme des acides aminés selon l'hôte, accroissement des défenses anti-oxydantes pour les stades présents chez l'insecte, accroissement de l'expression des protéines leurres de surface pour les stades présents chez l'homme). Cette analyse a aussi montré qu'une faible proportion des protéines trans-sialidases correspond à des formes enzymatiquement actives





(13 sur 223 identifiées), confirmant le rôle de leurre des formes non actives vis-à-vis du système immunitaire.

Pour *Leishmania*, les études des profils d'expression 2D-PAGE n'ont permis pour l'instant d'identifier que 49 (Bente *et al.*, 2003) et 47 (Nugent *et al.*, 2004) protéines différentes, car les identifications protéiques à grandes échelles étaient complexes chez ce parasite en raison de données de génomique réduites jusqu'à présent. Les protéines identifiées ont été choisies pour être des protéines spécifiques d'un seul stade de développement et pour cela ont été sélectionnées parmi les spots protéiques définis comme spécifiques par des études comparatives entre les stades promastigote et amastigote (Bente *et al.*, 2003 ; Nugent *et al.*, 2004). Ces premiers résultats montrent que les changements métaboliques ont lieu au cours des 24 premières heures de différenciation dans les grandes voies de régulations cellulaires. Un autre exemple d'analyse comparative est la recherche de marqueurs d'infection et de virulence chez *T. gondii*. Des travaux ont comparé les cartes protéomiques de tachyzoïtes d'une souche virulente et d'une souche atténuée (Nischik *et al.*, 2001). Parmi les 200 spots protéiques obtenus, 35 ont un niveau d'expression plus faible chez la souche atténuée, montrant une corrélation entre la virulence du parasite et les synthèses d'actine, de catalase et de plusieurs protéines excrétées.

#### **2.6.4. Analyses descriptives de sous-protéomes : de la localisation à la fonction**

L'analyse des sous-protéomes de la cellule parasitaire, comme par exemple les organelles impliqués dans les mécanismes d'invasion constitue un élément important de la problématique des parasites. L'information sur la localisation des protéines identifiées, est une approche intéressante pour permettre l'attribution de fonction aux nombreuses protéines hypothétiques des parasites. Au contraire, l'identification de protéines à fonction connue, associées à certains organites permet de mieux comprendre le rôle de ceux-ci dans la biologie de la cellule. Or, dans le cas des parasites où la compartimentation peut être très complexe, il peut être difficile d'obtenir des fractions pures pour un organelle donné. Ainsi, en l'absence de marqueur permettant d'identifier sans ambiguïté un type d'organites purifiés et d'évaluer son degré de pureté, ces résultats devront être validés par des approches de biologie plus traditionnelles : immunolocalisation, expression transgénique ou inhibition de gène.

Cette approche a été appliquée à l'identification moléculaire et la caractérisation de la biogenèse du phagosome d'*Entamoeba histolytica* (Okada *et al.*, 2005) montrant son implication dans les mécanismes d'adhésion, dans le réarrangement du cytosquelette, dans le trafic vésiculaire et dans la dégradation. Chez les apicomplexes, où la compartimentation cellulaire est très importante, de



nombreuses études ont été effectuées pour attribuer un rôle précis au cours de l'invasion cellulaire aux trois organites majeurs que sont les rhoptries, les micronèmes et les granules denses (Blackman et Bannister, 2001). Chez *T. gondii*, l'approche protéomique a permis d'identifier les protéines majeures localisées dans les rhoptries et dans les granules denses (Cohen *et al.*, 2002). Chez *Eimeria tenella*, 37 protéines de micronème ont été identifiées (Bromley *et al.*, 2003). Il ressort de ces études que tous ces organelles contiennent un mélange complexe de protéines qui leurs sont, à tous trois, souvent communs, et qui jouent un rôle actif dans les processus d'invasion. Les protéines de micronème sont importantes dans la phase de reconnaissance cellulaire et sont impliquées dans la mobilité du parasite alors que celles des rhoptries semblent être à l'origine de la formation de la vacuole parasitophore. Les protéines des granules denses sont, quant à elles, prédominantes dans les modifications de la cellule-hôte après l'invasion.

### **2.6.5. Analyses protéomiques et thérapeutiques**

L'étude des interactions entre les composés antiparasitaires et les parasites est devenue une problématique majeure en parasitologie. En effet, les résistances des parasites aux traitements disponibles constituent un réel problème sanitaire et le développement de nouvelles chimiothérapies ou vaccins un challenge pour les parasitologues. Actuellement, les premiers apports de la protéomique à cette problématique sont plutôt centrés sur la description générale, à l'échelle de la cellule, des mécanismes de résistance et des voies métaboliques mises en jeu. C'est le cas de la première analyse protéomique de la résistance d'un parasite à une chimiothérapie, le méthotrexate, qui a été effectuée sur *L. major* (Drummel-Smith *et al.*, 2003 ; Drummel-Smith *et al.*, 2004) . Elle montre les modifications physiologiques induites par le composé et les voies de détoxication activées lors de cette résistance comme la surexpression des protéines impliquées dans les voies de réponse au stress, ainsi que le rôle d'enzymes impliquées dans le métabolisme des acides aminés, qui, jusque là, n'avaient pas été associées à ce mécanisme de résistance. Ces résultats confirment que l'acquisition de la résistance à une molécule active est un mécanisme complexe qui fait intervenir une grande variété de réponses cellulaires. Le même type d'analyse a été effectué pour comprendre les mécanismes de résistance à la chloroquine chez *Plasmodium* (Gunasekera *et al.*, 2003).

L'étape suivante est le développement de nouveaux composés thérapeutiques, qui nécessite de découvrir de nouvelles cibles chez le parasite, cibles qui doivent être spécifiques. Cela nécessite d'acquérir des données descriptives fondamentales pour avoir une meilleure compréhension de la physiologie du parasite, ainsi qu'une meilleure connaissance du métabolisme spécifique de chaque



stade. Ainsi, les données acquises par l'intermédiaire des différents types d'analyses protéomiques doivent être intégrées dans l'optique de la recherche de nouvelles cibles thérapeutiques.

### **2.6.6. L'analyse protéomique, un outil pour la recherche de cibles vaccinales**

A cause des mécanismes d'échappement qu'ont développés les parasites pour contrer le système immunitaire (variabilité antigénique, leurres protéiques), le développement d'une cible vaccinale et d'une réponse immune protectrice passe par la détection à grande échelle des interactions entre le parasite et l'hôte, comme par exemple la définition des épitopes d'interaction entre *Plasmodium* et les lymphocytes T CD4+ et CD8+ (Good et Doolan, 1999). De même, pour comprendre la réponse de l'hôte à l'infection par *T. gondii*, la réactivité de sera (Geissler *et al.*, 1999) et de cellules T (Dlugonska *et al.*, 2001) de patients a été testée contre des protéines parasitaires séparées par 2D-PAGE.

L'identification dans le protéome de *T. cruzi* de protéines exprimées en abondance dans les stades présents chez l'homme, constitue une piste qui pourrait permettre le développement de vaccin sur ces nouveaux candidats (Atwood *et al.*, 2005). Chez *Plasmodium* spp, bien que complexe, le développement d'un vaccin semble être réalisable car une réponse protectrice a pu être obtenue par immunisation de sporozoïtes irradiés (chez l'homme (Doolan *et al.*, 2003), le singe et le rongeur). Mais les protéines de surface actuellement connues sont des protéines à forte variabilité antigénique et l'étude fondamentale des constituants externes de ces parasites pourrait donc être vue aussi dans l'optique de recherche de nouvelles protéines antigéniques, candidates possibles à des essais de vaccins. Par exemple, une étude de criblage par analyse protéomique (Florens *et al.*, 2004) a permis de mettre en évidence de nouvelles protéines de surface. Parmi celles-ci, une protéine à fonction inconnue, ne présentant pas de polymorphisme et pouvant être impliquée dans la formation d'un complexe multiprotéique, semble être un candidat intéressant pour servir de cible à un vaccin préventif.

### **2.6.7. Données de protéomique chez les microsporidies**

Les données sur le protéome des microsporidies sont limitées seulement à trois publications. Une première étude (Langley *et al.*, 1987) porte sur l'analyse comparative des profils électrophorétiques bidimensionnels des protéines de spores de trois espèces parasites d'insectes (*Nosema bombycis*, *Antonospora (Nosema) locustae* et *Vairimorpha necatrix*). Elle met en évidence, malgré le peu de spots visibles, des profils bidimensionnels peu similaires. Plus récemment, 200 spots protéiques ont



été séparés par 2D-PAGE à partir de spores d'*Encephalitozoon intestinalis*, une microsporidie humaine, et 95 d'entre eux ont été montrés comme immunogènes chez le lapin, cependant aucune identification n'a été réalisée (Moura et Visvesvara, 2001). De plus, la spectrométrie de masse MALDI-TOF a été proposée par Moura *et al.* comme technique rapide d'identification des microsporidies par détection des profils protéiques de petits peptides compris entre deux et huit kDa (Moura *et al.*, 2003). Le peu de publications traitant de protéomique microsporidienne est également à mettre en parallèle avec le peu de données publiées individuellement sur des protéines microsporidiennes. Une revue reprend les études s'intéressant aux produits des gènes d'*E. cuniculi* publiées entre 2002 et 2004 (Texier *et al.*, 2005).





## 3. LA PAROI DES CHAMPIGNONS UNICELLULAIRES

---

### 3.1. Structure et mise en place

#### 3.1.1. *Saccharomyces cerevisiae*, un organisme modèle pour la paroi des levures

Comme chez toutes les cellules fongiques unicellulaires, la morphologie de *Saccharomyces cerevisiae* est principalement déterminée par sa paroi. Cette structure externe assure la protection de la cellule face aux changements de l'environnement, définit sa forme et régule les échanges avec le milieu extérieur. Cependant, malgré sa rigidité, cette paroi est une structure extrêmement dynamique, pouvant s'adapter aux modifications morphologiques de la levure : croissance cellulaire, bourgeonnement cellulaire et sporulation. Elle peut subir des modifications qualitatives en réponse aux conditions environnementales. Le cycle de développement de *S. cerevisiae* présente trois formes cellulaires différentes : haploïde, diploïde et sporale. Du point de vue de la structure de la paroi les formes haploïdes et diploïdes sont très similaires. Par contre, le rôle protecteur de la paroi de la spore ou ascospore nécessite une structure pariétale très différente.

##### 3.1.1.1. Composition, structure et mise en place de la paroi de la cellule végétative

La paroi des formes végétatives de *S. cerevisiae* est une structure d'environ 200 nm d'épaisseur (Klis *et al.*, 2002). Sa composition globale est d'environ 60% de glucane, 40% de mannoprotéines, et 1% de chitine. L'analyse en microscopie électronique à transmission (MET), révèle en coloration négative une structure en doubles feuillets : le plus interne est transparent aux électrons, il est principalement constitué de  $\beta$ -1,3-glucane et de chitine, alors que le plus externe, dense aux électrons, est composé de mannoprotéines (Osumi, 1998).

##### 3.1.1.1.1. Le réseau de $\beta$ -glucane

La couche la plus interne de la paroi est constituée d'un réseau qui est distribué uniformément autour de la cellule et composé de polymères linéaires et ramifiés de molécules de  $\beta$ -1,3-glucane, sur lesquelles sont liées par liaisons covalentes des molécules de  $\beta$ -1,6-glucane et de chitine. *In*



*vitro*, ce polymère est structuré en simple filament, alors qu'il est ramifié *in vivo*, ce qui suppose des modifications des chaînes de glucanes néoformées durant leur synthèse ou peu après (Shematek et Cabib, 1980). Plusieurs travaux ont montré l'implication de nombreuses protéines dans ces modifications, mais leur rôle précis n'a pas encore été défini. Ces enzymes sont généralement non associées à la paroi, et peuvent être retrouvées dans le milieu de culture. C'est par exemple le cas de l'exoglucanase Exg1, responsable de 90% de l'activité glucanase dans la cellule végétative (Orlean, 1997), ainsi que l'endonucléase Bgl2, présentant aussi une activité transglucosylase qui permettrait la connexion des chaînes néoformées sur des chaînes existantes (Goldman *et al.*, 1995). En phase stationnaire de croissance, chaque filament est constitué d'environ 1500 monomères de glucose, mais certaines études montrent de plus hauts degrés de polymérisation (Muller *et al.*, 1997). Ces filaments sont maintenus solidaires entre eux par des liaisons hydrogène. Ce réseau apporte à la paroi à la fois de la résistance et de l'élasticité, grâce aux structures en hélice à divers degrés d'extension, que forment les chaînes de  $\beta$ -1,3-glucane (Ress *et al.*, 1982 ; Krainer *et al.*, 1994) et qui peuvent se déformer pour s'adapter aux variations morphologiques de la cellule. De plus, les deux complexes  $\beta$ -1,3-glucane synthases, Fks1 et Gsc2/Fks2, sont soumis à des régulations environnementales impliquant de fortes variations dans le degré de polymérisation des chaînes (Cabib *et al.*, 1998). Ces deux complexes sont notamment constitués de deux protéines transmembranaires, les protéines Fks1 et Gsc2, qui sembleraient en être les unités catalytiques. Cependant, elles sont aussi suspectées de former des pores transmembranaires permettant le transit des chaînes de glucane vers la paroi (Moor et Muhlethaler, 1963).

En phase stationnaire de croissance, la quantité de résidus de  $\beta$ -1,6-glucane dans la couche la plus interne de la paroi est estimée à environ 3 à 4 % (Montijn *et al.*, 1999). Cependant, ce taux est dépendant des conditions de milieu, le  $\beta$ -1,6-glucane étant plus abondant lors d'un développement en milieu peu favorable. Ces résidus servent d'accepteurs de liaison au sein du réseau de glucane, permettant la liaison de certaines protéines pariétales ainsi que des molécules de chitine (Kollar *et al.*, 1997).

### **3.1.1.1.2. La chitine**

La chitine, polymère de  $\beta$ -1,4-N-acétylglucosamine, est présente en faible quantité dans la paroi des levures végétatives mais sa présence est essentielle à la viabilité de la cellule (Shaw *et al.*, 1991). Les chaînes linéaires formées s'associent en microfibrilles qui sont maintenues entre elles par des liaisons hydrogène. Elles sont principalement rencontrées sous forme de chaînes d'environ 190 résidus au niveau du site de bourgeonnement lors de la division cellulaire, puis après cytodièrese au niveau de la cicatrice de bourgeonnement chez la cellule mère (Holan *et al.*, 1981). De la chitine est également retrouvée répartie uniformément dans l'ensemble de la paroi de la cellule mère, liée à des

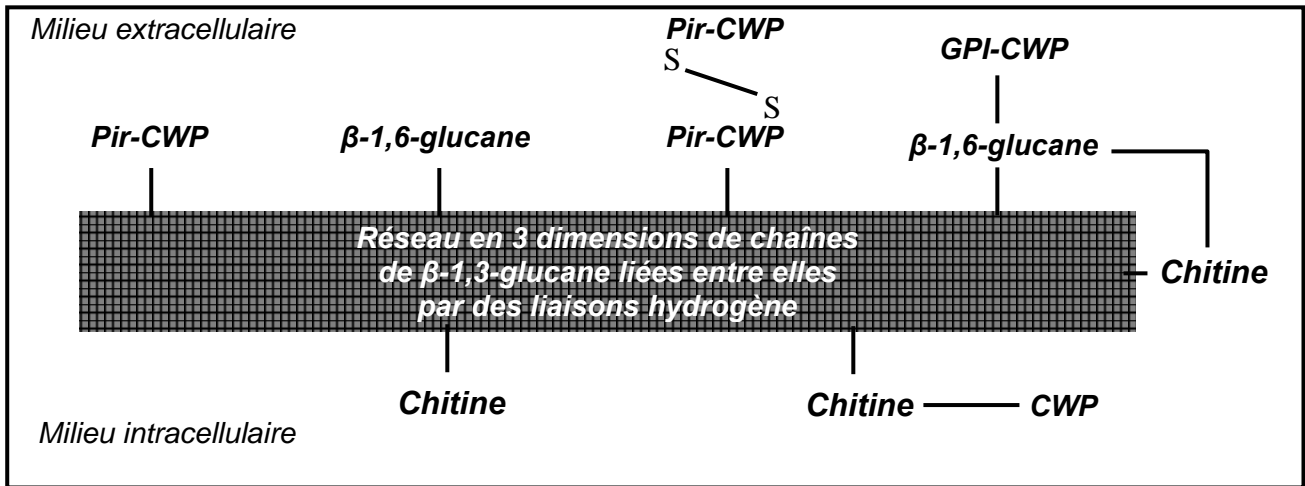


résidus de  $\beta$ -1,3- et  $\beta$ -1,6- glucane (Molano *et al.*, 1980) alors que sa présence n'est pas détectée chez la cellule fille (Shaw *et al.*, 1991). Le dépôt de chitine est vu, par certains auteurs (Klis *et al.*, 2002), comme un mécanisme de secours et de protection de la cellule. Dans le cas de cellules déficientes dans la formation de la paroi, celle-ci peut être constituée à plus de 20% de chitine (contre 1 à 2 % dans le cas d'une cellule normale) (Dallies *et al.*, 1998). Elle serait donc déposée pour renforcer les zones fragilisées comme c'est le cas au niveau du septum et des régions adjacentes lors de la division cellulaire.

La synthèse de chitine chez *S. cerevisiae* fait intervenir trois chitine synthases (CSI, CSII et CSIII) qui sont régulées indépendamment (Chuang et Schekman, 1996 ; Ziman *et al.*, 1996). L'activité de la protéine CSII est strictement régulée par le cycle cellulaire, avec un pic d'activité juste avant la séparation cellulaire. Sa localisation au niveau du septum de division permet d'envisager un rôle actif dans la mise en place de la chitine au niveau de l'anneau de division (Cid *et al.*, 1995). Les chitine synthases CSI et CSIII ont été localisées au niveau du cytoplasme mais aussi au niveau des chitosomes. Ces vésicules semblent être dérivées des voies d'endocytose, ce qui suggère que la régulation temporelle et spatiale de la synthèse de chitine au niveau de la paroi puisse être médiée par la mobilisation du stock endosomique de ces deux enzymes (Ziman *et al.*, 1996).

### 3.1.1.1.3. *Protéines de paroi*

Les protéines de paroi (ou CWP pour Cell Wall Proteins) sont classées en deux grandes familles selon leur caractéristique de solubilité : celles faiblement liées à la paroi et solubles à l'aide de détergents (SDS) et celles liées par liaisons covalentes qui nécessitent des traitements de solubilisation plus complexes, comme des traitements alcalins modérés, des agents réducteurs, des détergents ou des digestions enzymatiques. Ces traitements peuvent être effectués de manière séquentielle pour faciliter l'extraction protéique. Par exemple l'utilisation d'une protéase permet d'éliminer les protéines les plus externes et ainsi de faciliter l'action d'une  $\beta$ -glucanase qui pourra solubiliser les protéines liées au réseau glucidique. Une technique alternative est l'analyse de la régénération de protoplastes où des cellules végétatives sont maintenues en survie après avoir été déplétées de leur paroi. L'identification des protéines excrétées dans le milieu permet d'avoir des informations sur les acteurs impliqués dans la mise en place précoce de la paroi (Pardo *et al.*, 1999). Les protéines les plus solubles de la paroi sont principalement des enzymes impliquées dans la synthèse, la mise en place et les modifications du réseau de glucides de la paroi ainsi que dans la liaison des protéines à ce réseau (par exemple les familles des glucane synthases, exoglucanases, chitine synthase, chitine déacétylase, chitinase). On retrouve également parmi ces protéines solubles des familles de protéines de réponse au stress, des enzymes glycolytiques ainsi que des protéines à fonction inconnue. Les protéines liées de façon covalente au réseau glucidique sont des



**Figure 13.** Modèle de l'organisation de la paroi de la cellule végétative de *S. cerevisiae*. Représentation de l'accrochage des protéines pariétales (CWP, Pir-CWP et GPI-CWP) au réseau de  $\beta$ -1,3-glucane et aux composés glucidiques de la paroi.

mannoprotéines qui forment la couche la plus externe de la paroi. La majorité des asparagines de ces protéines porte de longues chaînes ramifiées de résidus de mannose (N-glycosylation). Les sérines ainsi que les thréonines peuvent aussi être glycosylées mais par de plus courtes chaînes (O-glycosylation). Ceci participe, avec les liaisons interprotéiques par ponts disulfure, à rendre cette couche extrêmement imperméable aux macromolécules. Ces mannoprotéines sont des protéines très hautement glycosylées, avec un taux de résidus glucidiques représentant parfois plus de 90% de la masse de la molécule (Orlean, 1997). La fonction de ces protéines est principalement structurale. La plus abondante est le mannane, une mannoprotéine riche en mannose, liée de manière covalente aux autres protéines (Pardo *et al.*, 1999).

Les protéines liées de manière covalente à la paroi présentent plusieurs types d'ancrage possible (figure 13). Les Pir-CWP (Protein with Internal Repeats) sont directement liées au  $\beta$ -1,3-glucane par une liaison osidique et sont labiles dans des conditions de lyse alcaline modérée (Kapteyn *et al.*, 2000). Elles présentent toute une structure similaire : un peptide signal, le site *kex2* qui est une région constituée de répétitions variable selon les protéines et une région C-terminale hautement conservée présentant 4 cystéines à des positions conservées. Cette région conservée peut comporter jusqu'à 11 répétitions qui sont les sites servant d'ancrage aux chaînes de  $\beta$ -1,3-glucane (Castillo *et al.*, 2003). Cela permet à une protéine Pir-CWP d'interconnecter plusieurs chaînes de glucane par l'intermédiaire de plusieurs domaines de liaison et ainsi de rigidifier la paroi. Les GPI-CWP sont, quant à elles, liées à la paroi par l'intermédiaire d'une chaîne de  $\beta$ -1,6-glucane (Frieman et Cormack, 2004). Ces protéines portent une molécule de glycosylphosphatidylinositol (GPI) qui permet sa liaison au  $\beta$ -1,6-glucane par l'intermédiaire d'un pont phosphodiester. La synthèse des GPI-CWP s'effectue dans la cellule, puis elles sont exportées vers la membrane plasmique où la liaison de la protéine au résidu lipidique est hydrolysée, libérant ainsi la protéine de son ancre membranaire pour lui permettre d'être intégrée à la paroi (Lu *et al.*, 1995). Dans le génome de *S. cerevisiae* (Goffeau *et al.*, 1996), plus de 50 protéines ont été identifiées comme de putatives protéines de paroi avec un signal GPI potentiel (Caro *et al.*, 1997). Un troisième type d'ancrage a été proposé pour des protéines qui se sont avérées être liées à d'autre CWP par des ponts disulfure (Mrsa *et al.*, 1997). Enfin, quelques protéines sont liées directement à la chitine par de courtes chaînes de  $\beta$ -1,6-glucane (Kollar *et al.*, 1997)

#### **3.1.1.1.4. Dynamisme de la paroi**

Des levures privées de leur paroi (ou protoplastes) et maintenues dans des conditions de protection de pression osmique et d'apport nutritif sont capables de régénérer *de novo* une paroi complète. Cependant, dans des conditions de croissance standard, le développement de la structure pariétale est basé sur la structure existante de la cellule mère. La paroi sert donc de matrice à l'incorporation



<b>F O N C T I O N</b>	<b>P R O T E I N E S</b>
<b>Perméabilité de la paroi</b>	<i>Ccw12, Ydr13C</i>
<b>Adhésion</b>	
Reproduction	<i>Sag1, Aga1, Aga2, Fig2</i>
Floculation	<i>Flo1, Flo5, Flo9, Flo10</i>
Formation de biofilm	<i>Flo11</i>
Invasion / croissance hyphale	<i>Flo11</i>
<b>Mise en place de la paroi</b>	
Organisation des glucanes	<i>Crh1, Crh2, Crr1, Fig2, Bgl2</i>
Division cellulaire	<i>Egt2, Pry3, Cts1, Dse2, Dse4, Scw11</i>
Protéine de structure	<i>Cwp1, Pir1, Pir2/HSP150, Pir3, Pir4/Cis3</i>
<b>Croissance isotropique</b>	<i>Pir1, Pir2/HSP150, Pir3, Pir4/Cis3</i>
<b>Condition de croissance</b>	
Sporulation	<i>Spol</i>
Phase stationnaire	<i>Sed1, Spi1</i>
Croissance en anaérobie	<i>Dan1, Dan4, Tip1, Tir1, Tir2, Tir 3, Tir4</i>
Basse température	<i>Tip1, Tir1, Tir2, Tir3</i>
<b>Métabolisme</b>	
Estérase	<i>Tip1</i>
Chélation du fer	<i>Fit1, Fit2, Fit3</i>
<b>Protéines de réponse au stress</b>	<i>Famille des Hsp150 et des Hsp70</i>

**Tableau 3.** Fonction des principales protéines pariétales chez *S. cerevisiae* d'après Klis et al. (2002).

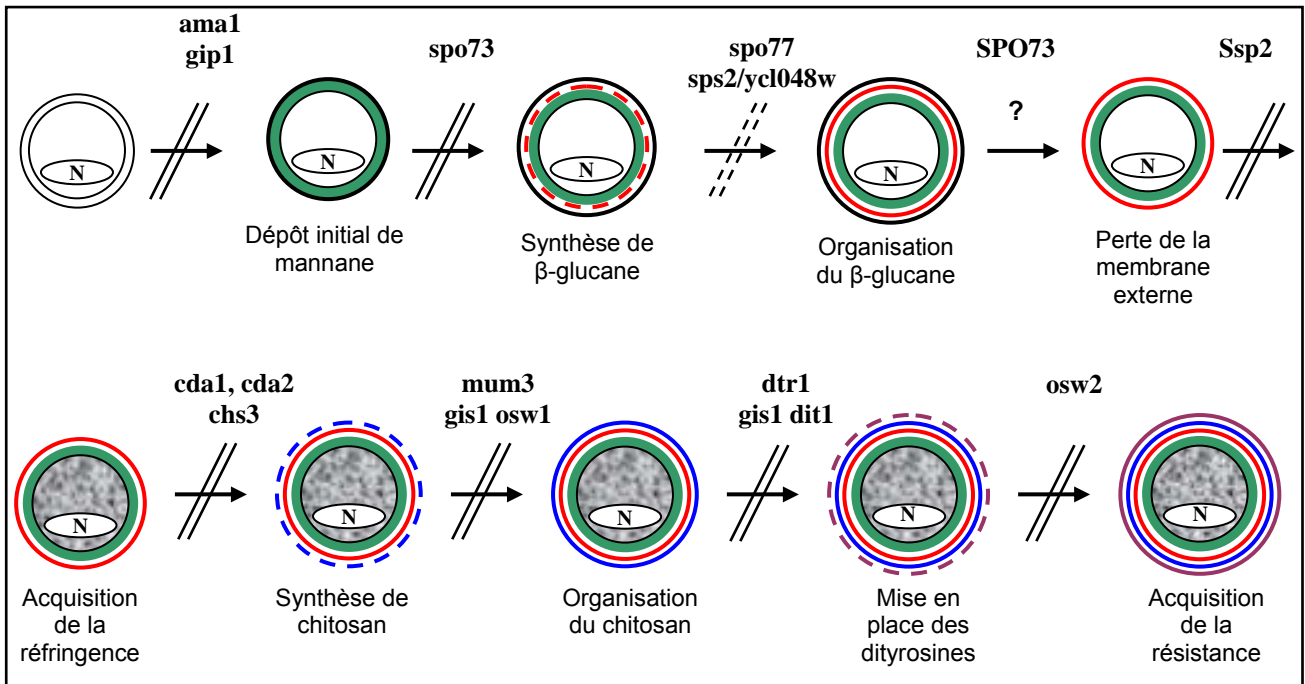
de nouveaux matériaux aux points de croissance, ce qui permet l'extension et le remodelage de la paroi des deux cellules. Il est estimé que 1200 gènes affectent directement la composition et l'organisation de la paroi (de Groot *et al.*, 2001).

La mise en place des protéines de paroi se fait de manière modulaire. En effet, des complexes covalents composés soit de protéines Pir-CWP, soit de protéines GPI-CWP, définissent deux types de blocs majeurs qui fonctionnent comme des unités structurales. Le complexe protéique le plus abondant lors d'une croissance en milieu riche, est formée de GPI-CWP liées à du  $\beta$ -1,6-glucane lui-même lié à du  $\beta$ -1,3-glucane. Les GPI-CWP de ce complexe sont principalement Ccw12, Tip1, Ssr1, Cwp1 et Sag1 (les fonctions de ces protéines sont indiquées dans le tableau 3). Cependant, la nature et la composition de ces complexes sont très variables en fonction des conditions environnementales ou cellulaires (quantité de nutriments présents dans l'environnement, inactivation de certaines enzymes, lyse enzymatique). Les proportions des différents constituants pariétaux, y compris les protéines, sont extrêmement variables, en réponse à de très nombreux stimuli, comme c'est par exemple le cas avec l'accroissement de la quantité de  $\beta$ -1,6-glucane et de chitine qui viennent compenser un appauvrissement de  $\beta$ -1,3-glucane dans la paroi.

La flexibilité de ce modèle structural pourrait expliquer les modifications spatiales et temporelles qui ont lieu dans la paroi au cours de la vie de la cellule. La fraction glucidique est synthétisée à la surface de la membrane plasmique. Les mannoprotéines sont synthétisées à l'intérieur de la cellule puis sont exportées à la surface, orientées vers les voies de sécrétion grâce à la présence d'un peptide signal. Au cours de ce transit, elles sont modifiées avec adjonction principalement de motifs glucidiques. Les deux types de macromolécules sont interconnectés à l'extérieur de la cellule par l'intermédiaire de liens glucidiques pour former des blocs qui sont ensuite eux-même assemblés pour réaliser la paroi mature. Ces étapes finales d'assemblage sont réalisées par des enzymes spécifiques situées soit dans la membrane plasmique soit dans la paroi elle-même.

### 3.1.1.2. Composition, structure et mise en place de la paroi de l'ascospore

En réponse à la privation de nutriments, la forme diploïde de *S. cerevisiae* sort du cycle de multiplication végétatif par une méiose qui conduit à la différenciation de quatre cellules filles en spores haploïdes (Kupiec *et al.*, 1997). Il s'agit d'une forme quiescente de survie, permettant le maintien dans l'environnement jusqu'au retour de nutriments. Sa résistance est assurée principalement par sa paroi. Il s'agit d'une structure extracellulaire stratifiée, plus complexe que la paroi de la cellule végétative, et qui est composée de quatre couches distinctes. Ses constituants principaux sont, de la couche la plus interne à la plus externe : des glycoprotéines, du  $\beta$ -1,3-glucane (Kreger-Van Rij, 1978), du chitosan (Briza *et al.*, 1988) qui est un polymère provenant de la déacétylation de la chitine et enfin une couche de dityrosine qui est un dimère de molécules de



**Figure 14.** Morphogénèse de la paroi de la spore de *S. cerevisiae*. Le schéma représente l'apparence de la spore à chaque étape. Les flèches barrées de lignes parallèles représentent les blocages induits par une mutation. Celles barrées de lignes en pointillé représentent les mutants déficients dans une étape mais qui poursuivent tout de même leur morphogénèse. Les lignes noires représentent les membranes, en vert les mannoprotéines, en rouge le  $\beta$ -glucane, en bleu le chitosan et en marron la dityrosine. N indique le noyau. D'après Coluccio et al. (2004).

tyrosine (Briza *et al.*, 1986). Les deux couches les plus internes sont similaires à celles de la cellule végétative mais en position inversée, les deux plus externes jouent un rôle majeur dans les mécanismes de résistance.

La morphogenèse de cette paroi est totalement indépendante de celle de la cellule végétative et se constitue *ex nihilo* au niveau de la membrane plasmique sans aucune matrice, de la couche la plus interne vers la plus externe (Tachikawa *et al.*, 2001). Le criblage de mutants déficients dans la capacité à sporuler a permis de proposer un modèle de cette morphogenèse (figure 14), faisant intervenir les acteurs protéiques impliqués dans la formation et dans la régulation de la paroi de l'ascospore (Rabitsch *et al.*, 2001 ; Coluccio *et al.*, 2004).

A l'issue de la méiose, les prospores sont entourées par deux membranes (Neiman, 1998). Entre celles-ci s'accumulent tout d'abord les mannoprotéines. Ensuite, le  $\beta$ -1,3-glucane est déposé, mais ce dépôt nécessite d'être organisé en un réseau cohérent pour permettre l'acquisition de la forme sphérique de l'ascospore. En effet, les mutants pour les protéines responsables de cette organisation (SPO77 et SPS2) présentent une forme allongée tout en étant cependant capable de poursuivre la formation de leur paroi. C'est aussi pendant la synthèse de cette couche de glucane que la membrane externe est résorbée et que le cytoplasme de la spore acquiert son caractère réfringent. La couche de chitosan est mise en place par un complexe d'au moins quatre enzymes : la chitine synthase III qui produit au niveau de la paroi les chaînes de chitine et deux chitine déacétylases CDA1 et CDA2 qui retirent le groupement acétyl aux résidus N-acétylglucosamine (Christodoulidou *et al.*, 1996 ; Christodoulidou *et al.*, 1999). On peut supposer que la déacétylation a lieu immédiatement après la synthèse des chaînes de chitine au niveau de la membrane pour donner le chitosan incorporé à la paroi. Enfin, la chitinase CTS1 est responsable de la dégradation des résidus de chitine n'ayant pas été déacétylés et permet l'obtention d'une structure pariétale exclusivement composée de chitosan. Comme pour la couche de glucane, le dépôt de chitosan n'est pas suffisant pour l'obtention d'une couche ordonnée. Son organisation est de plus indispensable pour que les composants les plus externes puissent être mis en place. Elle est réalisée par les protéines Mum3 et Osw1.

La couche la plus externe de la paroi est donc la dernière à être mise en place et est constituée de molécules de dityrosine liées entre elles (Briza *et al.*, 1986 ; Briza *et al.*, 1994). Les monomères de dityrosine sont produits dans le cytoplasme par les protéines Dit1 et Dit2, puis exportés vers la paroi par la protéine Dtr1. La protéine Osw2 est nécessaire à l'agencement final de cette couche et à l'obtention de la résistance aux facteurs extérieurs. La synthèse de ces protéines est sous la dépendance du facteur de transcription Gis1.

Les quatre ascospores produites à partir d'une cellule mère semblent être liées entre elles par des « extensions » de la paroi appelées *interspore bridges* (Coluccio et Neiman, 2004). Leur analyse en



microscopie électronique à transmission et à balayage révèle une structure qui est en continuité avec les couches externes de la paroi et qui en présente le même aspect. Cela suggère que ces « ponts » sont composés de chitosan ainsi que de dityrosine. Cependant, la nature du matériel composant la région centrale de ces extensions n'est pas encore connue. Ces prolongements pourraient favoriser la conjugaison sexuelle entre spores sœurs et ainsi favoriser le maintien de l'hétérozygotie de *S. cerevisiae*.

### **3.1.2. La paroi de *Candida albicans*, une structure impliquée dans la pathogénicité**

La paroi de *Candida albicans* joue un rôle majeur dans sa morphogénèse et sa pathogénicité, et représente une cible d'intérêt pour la recherche de nouvelles molécules antifongiques. De nombreuses études ont montré que la paroi de *Candida albicans* présente une organisation similaire à celle de *S. cerevisiae* tant au niveau de la structure que des constituants (Kapteyn *et al.*, 2000 ; Klis *et al.*, 2001). Bien que similaire entre les deux champignons, cette structure présente pourtant des fonctions variées d'une espèce à l'autre. En effet, *C. albicans* est un champignon commensal mais qui peut être pathogène grâce à des caractéristiques d'adhésion cellulaire, et qui possède la capacité de moduler la réponse immunitaire de l'hôte. L'analyse de ce sous-protéome a été effectuée grâce à des extraits protéiques enrichis en matériel pariétal, provenant d'extractions différentielles sur des parois isolées (Pitarch *et al.*, 2002). Cette étude montre qu'un grand nombre de protéines est impliqué dans la mise en place de la paroi, et que les remaniements morphologiques faisant suite à la formation du tube germinatif sont reliés à des changements dans la composition des protéines de paroi (protéines de réponse au stress, enzymes, facteurs de virulence putatifs). Ces données sont cohérentes avec les résultats des études comparatives des transcriptomes des deux états (Fradin *et al.*, 2003).

De nombreux constituants de la paroi sont reconnus par les anticorps de patients atteints de candidose et jouent un rôle dans la pathogénicité (Nimrichter *et al.*, 2005). C'est particulièrement le cas des protéines de réponse au stress, Hsp90 et la famille des protéines Hsp70, qui ont été caractérisées comme composants de la paroi. Des anticorps dirigés contre la protéine immunodominante Hsp90 présentent un caractère protecteur vis-à-vis de la létalité des candidoses chez des sujets immunodéprimés. Des glucosylcéramides (GSL) sont également présents dans la paroi : ce sont des adhésines qui peuvent se fixer aux composés hydrophiles de la membrane plasmique. Ils ont une fonction directe dans la pathogénicité, car des anticorps anti-GSL ou des inhibiteurs de la GSL synthase empêchent la formation du tube germinatif. Leur spécificité dans le groupe des champignons en font des cibles de choix pour le développement d'agents antifongiques.



Un grand nombre de mannoprotéines (dont l'antigène majeur Mp58, une mannoprotéine fixant le fibrinogène), joue également un rôle important dans l'adhésion à la cellule-hôte. Elles peuvent être divisées en deux groupes, les peptidomannanes de haut poids moléculaire et les mannoprotéines à faible poids moléculaire. Ces deux groupes présentent des chaînes glycosidiques similaires mais des peptides ou des protéines différentes. Il s'agit principalement de protéines de type lectine, de récepteurs et de ligands, d'intégrines ainsi que d'adhésines.

Des études de protéomique comparative entre *S. cerevisiae* et *C. albicans* ont également permis le recensement des protéines exposées à la surface et spécifiques à *Candida*. Ces protéines fortement exposées au système immunitaire sont des candidates potentielles pour être des antigènes majeurs ainsi que des facteurs de virulence et d'adhésion. Ces analyses ont permis de mettre en évidence qu'une GPI-CWP était directement impliquée dans l'adhésion de *C. glabrata* aux cellules épithéliales humaines (Cormack *et al.*, 1999). A plus grande échelle, elles ont également permis de recenser des protéines exposées à la surface et fortement labiles (Urban *et al.*, 2003) ou bien liées de manière covalente à la paroi (De Groot *et al.*, 2004).

### **3.1.3. Les microsporidies**

#### **3.1.3.1. La structure en deux feuillets de la paroi**

L'analyse de la paroi de la spore microsporidienne par microscopie électronique à transmission révèle une structure constituée de deux feuillets en contact intime. Le plus interne, l'endospore, est situé au contact direct de la membrane plasmique et supporte la couche la plus interne ou exospore.

##### **3.1.3.1.1. L'endospore**

L'endospore est une région qui apparaît transparente aux électrons et qui ne présente pas d'organisation apparente. Elle recouvre uniformément la spore avec une épaisseur constante d'environ quatre vingt nanomètres, à l'exception de la région du capuchon polaire où vient s'ancrer le tube polaire. Cette région où se fera l'extrusion du tube polaire possède une endospore bien plus fine, de l'ordre d'une trentaine de nanomètre (pour *Encephalitozoon cuniculi*). L'endospore est une structure riche en polysaccharides. La chitine est le composé majoritaire (Erickson et Blanquet, 1969), sa présence a été confirmée par des réactions de coloration (Calcofluor White M2R, Uvitex 2B) ainsi que par spectrométrie à infrarouge et diffraction aux rayons X (Vavra, 1976a). Des analyses par cryofractures et cryodécoupages, effectuées sur le genre *Encephalitozoon*, révèlent que l'endospore est en réalité un espace parcouru par un réseau de fibrilles formant des ponts entre l'exospore et la membrane plasmique (Bigliardi *et al.*, 1996). Les auteurs ont avancé l'hypothèse que ces microfibrilles sont des filaments d' $\alpha$ -chitine interconnectés entre eux par l'intermédiaire de





résidus de  $\beta$ -chitine ou par l'intermédiaire de protéines pour former un réseau cohérent. Celui-ci participe au maintien de la forme de la cellule et rend donc la membrane plasmique solidaire de la paroi, expliquant pourquoi la membrane plasmique reste associée à la paroi après expulsion du sporoplasme.

### **3.1.3.1.2. L'exospore**

L'exospore est une région qui apparaît dense aux électrons et principalement composée de protéines. Les premières analyses biochimiques ont montré une richesse en acides aminés impliquant une structure riche en protéines (Vivarès *et al.*, 1976 ; Vavra, 1976a). La variabilité de son aspect selon le genre et l'espèce constitue un critère important d'identification. Elle peut être homogène ou stratifiée, et son aspect externe peut être nu ou bien décoré d'ornementations allant jusqu'à des projections de matériel pariétal. D'une extrême à l'autre son épaisseur peut varier d'une dizaine de nanomètres à près de 200 (Larsson, 1986). Chez le genre *Encephalitozoon*, l'exospore a tout d'abord été décrite comme une structure homogène (Barker, 1975), mais les études en cryofracture et cryodécapage (Vavra *et al.*, 1986 ; Bigliardi *et al.*, 1996) montrent une structure bien plus complexe. Elle semble être constituée de trois couches de nature, structure et densité différentes, qui adhèrent cependant intimement les unes aux autres. La couche la plus interne, assez épaisse (une quinzaine de nanomètres) est composée d'un réseau filamenteux. Directement dessus, repose une couche intermédiaire peu dense aux électrons, d'une épaisseur de deux à trois nanomètres. Elle est constituée de fibrilles de quatre nanomètres agencées parallèlement et qui forment une monocouche régulière d'aspect strié. Enfin, la couche externe semble être composée de blocs d'aspect flocculeux, hérissés de spicules et séparés de manière régulière par des espaces peu denses aux électrons.

### **3.1.3.1.3. Les protéines pariétales**

Peu d'études viennent éclairer la structure moléculaire de la paroi microsporidienne. A ce jour, seules deux protéines ont été identifiées au niveau de la paroi. La protéine SWP1 (Spore Wall Protein) de 50 kDa a été localisée au niveau de l'exospore chez *E. cuniculi* ainsi que *E. intestinalis* et un gène homologue a été identifié chez *E. hellem* (Bohne *et al.*, 2000). Aucune homologie de domaine n'a été retrouvée pour cette protéine qui présente une région C-terminale originale, constituée d'une série de répétitions riches en sérine et en glycine. Chez *E. intestinalis*, une seconde protéine de paroi a été identifiée (Hayman *et al.*, 2001). Cette protéine, SWP2, de 150 kDa est très similaire à SWP1 au niveau de sa région N-terminale, mais elles diffèrent par leur région C-terminale qui est constituée d'une cinquantaine de répétitions de 12 à 15 acides aminés chez SWP2. Des expériences d'immunoprécipitation semblent montrer que ces deux protéines forment un

<i>Anticorps développé contre</i>	<i>Nombre et nature des anticorps</i>	<i>Western Blot chez la microsporidie d'origine</i>	<i>Localisation</i>	<i>Réactivité croisée sur E. cuniculi montrée par Western blot</i>	<i>Auteurs</i>
<i>E. intestinalis</i>	1 - Mab	130 kDa	paroi <sup>i</sup>	non testé	Beckers et al., 1996
<i>E. intestinalis</i>	6 - Mab	40 à 60 kDa	paroi <sup>i</sup>	40 à 60 kDa	Lujan et al., 1998
	1 - Mab	23 kDa	paroi <sup>i</sup>	23 kDa	
	1 - Mab	55 kDa	paroi <sup>i</sup>	55 kDa	
<i>E. hellem</i>	1 - Mab	40 à 60 kDa	paroi <sup>i</sup>	40 à 60 kDa	
<i>E. cuniculi</i>	1 - Pab	30 kDa	exospore <sup>e</sup>	30 kDa	Delbac et al., 1998
	1 - Pab	33 et 55 kDa	endospore <sup>e</sup>	33 et 55 kDa	
	1 - Pab	55 kDa	endospore <sup>e</sup>	55 kDa	
<i>E. intestinalis</i>	1 - Mab	bande à 125 kDa	exospore <sup>e</sup>	ne croise pas	Prigneau et al., 2000
	1 - Mab	doublet à 55/60 kDa	endospore <sup>e</sup>	doublet 55/60 kDa	
	1 - Mab	doublet à 55/60 kDa	paroi <sup>e</sup>	ne croise pas	
	1 - Mab	doublet à 55/60 kDa	endospore <sup>e</sup>	ne croise pas	
	1 - Mab	bande à 55/60 kDa	endospore <sup>e</sup>	ne croise pas	

**Tableau 4.** Bilan des tailles des antigènes reconnus par des anticorps dirigés contre la surface des microsporidies du genre *Encephalitozoon*. Mab pour anticorps monoclonal. Pab pour anticorps polyclonal. Les expériences d'immunolocalisation avec un i en exposant ont été effectuées par IFA, celles avec un e par microscopie électronique à transmission.

complexe protéique dès le stade méronte. L'analyse bioinformatique de leur séquence révèle la présence d'un peptide signal et aucun ancrage membranaire, suggérant leur sécrétion à partir du réticulum endoplasmique.

D'autres études, basée sur des expériences de cytochimie, se sont intéressées aux protéines pariétales des microsporidies (Beckers *et al.*, 1996 ; Lujan *et al.*, 1998 ; Delbac *et al.*, 1998a ; Prigneau *et al.*, 2000). Elles ont mis en évidence, grâce à des anticorps monoclonaux et polyclonaux, des antigènes localisés au niveau de la paroi. L'utilisation de ces anticorps en immunodétection sur des western blots révèle des protéines dont la masse apparente est majoritairement située entre 40 et 60 kDa (tableau 4). Cependant aucune de ces protéines n'a été identifiée.

### **3.1.3.2. Morphogenèse et mise en place**

La mise en place de la paroi débute très tôt au cours du cycle de développement par des dépôts irréguliers de matériel dense aux électrons à la périphérie de la cellule. Ils sont considérés comme correspondant à la future exospore et leur apparition marque la transition entre le stade méronte et le stade sporonte. L'endospore se forme plus tardivement entre la membrane plasmique et cette structure pariétale dense aux électrons. Elle apparaît chez le jeune sporoblaste et s'épaissit continuellement jusqu'à différenciation totale de la spore. Sa formation complète est concomitante avec l'aboutissement du cycle de développement et peut être interprétée comme la finalisation de la maturation de la spore, où l'endospore vient sceller la cellule, séparant complètement le sporoplasme du milieu extérieur (Vavra et Larsson, 1999).

La protéine SWP1 est absente chez le méronte, mais est localisée chez le jeune sporonte au niveau des dépôts de matériel dense. Elle demeure durant tout le développement de la spore à la surface de la paroi. SWP2 est mise en place plus tardivement, lorsque que la structure pariétale du sporonte est complètement formée (Hayman *et al.*, 2001). Les auteurs ont proposé que ces deux protéines puissent être regroupées en un complexe localisé au niveau le plus externe de l'exospore. Il s'agit à ce jour des seules données moléculaires connues sur la mise en place de la paroi microsporidienne.

## **3.2. Rôle et fonction de la paroi microsporidienne**

La paroi des champignons est une structure dont la nature et la fonction sont bien documentées. Nous allons nous appuyer sur ces études pour éclairer les fonctions de la paroi microsporidienne.



### **3.2.1. Maintien de la forme et adaptation au milieu extérieur**

L'une des fonctions principales de la paroi chez les microsporidies, comme chez de nombreux microorganismes (levures, bactéries), est le maintien de la forme cellulaire. Cette forme est principalement assurée par la rigidité de la paroi. De plus, chez les microsporidies cette nécessité d'une extrême rigidité est directement liée à la capacité d'invasion cellulaire. En effet, le processus d'extrusion du tube polaire est la conséquence d'un accroissement de la pression intracellulaire qui n'est possible que grâce à la résistance de la paroi (Frixione *et al.*, 1997). Cette paroi est également remarquable pour la protection qu'elle apporte au contenu cytoplasmique. Elle permet la résistance de la spore lors de l'exposition aux stress environnementaux après libération du parasite de la cellule-hôte. Sa capacité de résistance permet ainsi de conserver intact le pouvoir infectieux de la spore pendant plusieurs années à 4°C dans l'eau ou congelée jusqu'à de très faibles températures (-80°C). De même, elle permet la survie de la spore pendant plusieurs heures à des pH acides et basiques (Shadduck et Polley, 1978 ; Koudela *et al.*, 1999).

Nous avons vu que chez *Saccharomyces cerevisiae*, la paroi est une structure remarquable par son dynamisme et sa capacité d'adaptation en réponse aux stress environnementaux et aux perturbations génétiques. L'intégrité cellulaire est assurée par la variation relative et le degré de liaison des polymères constituant la paroi. La synthèse d'un composé est intimement contrôlée par la présence des autres, lui permettant ainsi d'être incorporé de manière plus importante pour en compenser un autre. Cette structure est donc capable de résister à des modifications (thermique, pH, oxydation) du milieu extérieur, tout en maintenant possible les transferts de nutriments ainsi que les sécrétions de la cellule. Par comparaison, le petit nombre d'études dédiées à ce sujet laisse apparaître une paroi microsporidienne bien moins dynamique. En effet, après avoir subi de nombreuses modifications au cours du cycle de développement, elle semble ne pas présenter de dynamique chez la spore mature et isoler complètement le milieu intrasporal du milieu extérieur (Vavra et Larsson, 1999).

### **3.2.2. Relation hôte-pathogène**

#### **3.2.2.1. Adhésion / reconnaissance**

Les protéines pariétales, par leur positionnement à la périphérie de la cellule, jouent un rôle majeur dans la liaison et la reconnaissance des protéines cibles de la cellule-hôte et sont intimement liées aux mécanismes d'activation des pathogènes. Ainsi, chez les champignons du genre *Candida*, plusieurs protéines localisées à la paroi sont impliquées dans les mécanismes de liaison à la surface cellulaire par reconnaissance de résidus glycosidiques (Cormack *et al.*, 1999 ; Santoni *et al.*, 2002 ; Alberti-Segui *et al.*, 2004).



Chez les microsporidies, les signaux conduisant à l'activation de l'invasion n'ont pas été établis. Mais on peut supposer que le mécanisme de reconnaissance de la cellule-hôte fasse intervenir des protéines pariétales quel que soit le modèle d'invasion cellulaire retenu par perforation de la membrane de la cellule-hôte par le tube polaire ou par endocytose induite (Weidner et Sibley, 1985 ; Couzinet *et al.*, 2000). Cependant, plusieurs études rapportent le rôle des sucres dans ces mécanismes d'adhésion et de reconnaissance. Des travaux ont montré que la présence de glycosaminoglycanes (ou GAGs) à la surface de la cellule-hôte était nécessaire à l'établissement de l'adhésion cellulaire chez *E. intestinalis* (Hayman *et al.*, 2005). C'est par cette famille de molécules que d'autres parasites se fixent à leur cible. C'est par exemple le cas des sporozoïtes de *P. falciparum* qui se lient à l'héparine et à l'héparan sulfate présent à la surface des hépatocytes et des cellules placentaires (Frevert *et al.*, 1993 ; Wadstrom et Ljungh, 1999). Les travaux de Hayman ont également montré que le taux d'infestation de la spore microsporidienne était fortement réduit lorsqu'on inhibait son adhésion par des anticorps dirigés contre des épitopes localisés à la paroi. Pour la première fois, est mise en évidence une relation directe entre adhésion et invasion chez les microsporidies. D'autre part, chez *E. cuniculi*, des expériences montrent une reconnaissance par la concanavaline A des protéines PTP1 et SWP1 (Bohne *et al.*, 2000 ; Xu *et al.*, 2003). Leur glycosylation pourrait être une O-mannosylation, seule voie de glycosylation déduite du génome d'*E. cuniculi*. Ces glycosylations joueraient un rôle significatif dans la capacité de la spore à envahir la cellule-hôte (Xu *et al.*, 2004 ; Xu et Weiss, 2005).

### 3.2.2.2. La paroi et la réponse immunitaire : intérêt pour le diagnostic.

Un autre aspect important de l'interaction avec l'hôte est la capacité du pathogène à moduler la réponse immunitaire de l'hôte. Chez *Candida* spp, la capacité des constituants de la paroi à activer ou déprimer la réponse immunitaire a été largement étudiée (Cassone, 1989). Le mannane et les mannoprotéines montrent une très forte activité élicitrice sur l'intégralité des agents du système immunitaire : cellules NK, phagocytes, immunités cellulaire et humorale (Domer, 1989 ; Palma *et al.*, 1992). Cependant, la paroi n'est pas seule responsable de l'activité pathogène de *Candida*, et le rôle des protéines sécrétées a été clairement démontré.

Chez les microsporidies, la réponse immunitaire de l'hôte durant l'infection a été étudiée, tant au niveau des mécanismes immunitaires mis en œuvre que de l'immunogénicité des protéines microsporidiennes. La réponse humorale anti-microsporidienne chez l'animal et chez l'homme est principalement dirigée contre les constituants les plus externes du parasite (protéines de tube polaire et de paroi) (Weiss *et al.*, 1992 et données personnelles). La forte immunogénicité de ces structures peut s'expliquer par leur exposition directe au système immunitaire, ainsi que par la présence de résidus glucidiques exposés connus pour leur activité élicitrice. Cependant, chez la souris, la





réponse humorale n'est pas suffisante pour prévenir la mortalité de l'hôte et éliminer le parasite (Khan *et al.*, 2001). La régulation de l'infection nécessitant en fait la mobilisation de l'immunité cellulaire (Braunfuchsova *et al.*, 2001).

La nature immunogène de la paroi autorise l'utilisation des anticorps dirigés contre les protéines pariétales comme marqueurs épidémiologiques utilisables pour le diagnostic (Franzen et Muller, 2001) et le suivi à long terme des microsporidioses (van Gool *et al.*, 2004). La production de protéines recombinantes reconnues par ces anticorps permettrait la mise au point de kits de détection (par technique ELISA) fiables et reproductibles. De plus, la fraction glycosidique de la paroi peut aussi être utilisée pour permettre la détection directe des spores microsporidiennes. C'est par exemple le cas de la chitine dont l'affinité pour les fluorochromes Uvitex 2B et Calcofluor White permet la coloration et la visualisation directe de la paroi (Vavra et Chalupsky, 1982 ; Ignatius *et al.*, 1997), permettant ainsi une détection et un dénombrement simples et directs.



# *R é s u l t a t s*



#### 4. ANALYSE DU PROTEOME du parasite eucaryote *Encephalitozoon cuniculi* : une carte de référence des protéines majeures exprimées en fin de sporogonie

---

Dans le contexte de l'exploitation du génome d'*E. cuniculi*, nous avons mené, avec l'équipe de J. Garin (Laboratoire de Chimie des Protéines, ERM 201 INSERM/CEA/UJF, CEA Grenoble), la première analyse descriptive du protéome d'une microsporidie. Ces travaux nous ont permis d'obtenir une première vision intégrée de la physiologie du parasite.

Pour des contraintes techniques liées à la séparation et à la purification des différents stades du parasite, nous avons centré notre étude sur les stades tardifs de la sporogénèse qui sont plus facilement accessibles à l'analyse moléculaire. En effet, des quantités suffisantes de cellules d'*E. cuniculi* sont difficiles à obtenir à partir de cultures de cellules adhérentes infestées où le parasite se multiplie de façon asynchrone dans des vacuoles parasitophores jusqu'à éclatement de ces cellules-hôtes. Après récolte des surnageants de culture, les stades parasitaires les plus jeunes ne résistent pas aux étapes de purification des débris cellulaires et seuls les stades les plus résistants peuvent être étudiés. Il s'agit des stades où la paroi est mise en place et assure une protection suffisante, c'est-à-dire ceux de fin de sporogonie ainsi que les spores matures.

Les échantillons protéiques ont été obtenus par extractions séquentielles d'un même échantillon de parasites. Les protéines ont tout d'abord été extraites en conditions dénaturantes à l'aide de chaotropes, de détergents et d'agents réducteurs (solution d'urée, thiourée, CHAPS, SDS et DTT) et une combinaison de traitements thermiques et de sonication. Une extraction par traitement alcalin a été ensuite effectuée sur le culot de matériel insoluble. Enfin, les protéines résistantes aux précédents traitements ont été solubilisées par solution de Laemmli. Les protéines du premier extrait ont été séparées par électrophorèses 2D-PAGE (gels de 18 cm, gammes de pI 3-10 et 5-8, coloration au bleu de Coomassie). 402 spots protéiques ont été sélectionnés et analysés par spectrométrie de masse de type MALDI-TOF et MS/MS. Ces spots présentent une masse apparente comprise entre 10 et 250 kDa, et des pI étalés sur l'intégralité de la gamme de pH, avec cependant une plus forte concentration dans la fenêtre de pI comprise entre 5 et 8. Les extraits protéiques obtenus par traitement alcalin et par solution de Laemmli sont incompatibles avec la séparation par 2D-PAGE. Ils ont ainsi chacun fait l'objet d'une analyse de type « Shotgun » : digestion trypsique de l'échantillon puis analyse des peptides par LC-MS/MS.



Nous avons ainsi pu identifier 139 protéines dans l'extrait urée, 17 dans l'extrait alcalin et 50 dans l'extraction au Laemmli. Certaines de ces protéines étant communes aux trois analyses, nous avons finalement identifié 177 protéines différentes correspondant aux stades tardifs de formation de la spore de *E. cuniculi*, ce qui représente 9% des 2000 séquences codantes du génome nucléaire. De plus, les extractions différentielles effectuées sont de plus en plus stringentes, en particulier dans la solubilisation des protéines comportant des domaines transmembranaires.

Les protéines identifiées ont été distribuées en 12 classes fonctionnelles. La plus représentée, avec près de 25% des protéines identifiées, correspond aux protéines sans fonction prédite, suggérant qu'une large proportion des gènes d'*E. cuniculi* codant pour des protéines à fonction inconnue est exprimée. Cela constitue également une validation de l'annotation automatique du génome. Pour les protéines dont la fonction est connue, la sur-représentation de certaines classes fonctionnelles par rapport au génome est cohérente avec les données connues sur la physiologie et la biochimie du parasite comme, par exemple, la protection contre les stress environnementaux.

Une part importante des protéines identifiées (42,9% des 120 protéines identifiées dans l'extrait urée) présente de multiples isoformes (variations de masse et de pI), comme par exemple l'actine (cinq isoformes). Ces variations sont probablement dues à des modifications post-traductionnelles importantes. De nombreuses enzymes impliquées dans ces mécanismes ont d'ailleurs été annotées dans le génome d'*E. cuniculi* : glycosylation, acétylation, myristoylation, isoprénylation, phosphorylation et protéolyse. Ces modifications post-traductionnelles sont probablement un moyen d'accroître la diversité moléculaire et fonctionnelle dans le contexte d'un parasite disposant d'un faible répertoire protéique, conséquence de son petit génome. De plus, ces modifications peuvent intervenir dans les mécanismes de régulation cellulaire, comme par exemple la phosphorylation dans la transduction des signaux de la membrane plasmique vers le noyau. Aux vues des données de ce protéome, nous proposons un modèle de contrôle du cycle cellulaire au niveau sporal. D'autre part, l'identification de plusieurs composants de la machinerie de traduction ainsi que du métabolisme énergétique semble être paradoxale dans des stades de fin de sporogénèse, où un ralentissement des mécanismes cellulaires dans la forme de dissémination du parasite (la spore) est plutôt attendu. Il peut s'agir du reflet de synthèses protéiques tardives chez les sporoblastes, ou bien alors de mise en réserve de matériel pour l'initiation immédiate d'un nouveau cycle de développement après l'invasion cellulaire.

En conclusion, ces résultats permettent de dresser l'image d'une cellule où de nombreux mécanismes de défense, de détoxification et de maintien de la structure des protéines permettent la survie de la spore dans l'environnement ainsi que la conservation de son pouvoir infectieux.



**Table S1:** Identified *E. cuniculi* spore proteins.

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>v</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vi</sup>		
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)
<b>1, Metabolism</b>											
<b><u>nucleotide metabolism: deoxyribonucleotide metabolism</u></b>											
ECU08_1380	glutaredoxin	U	a	a	12/5.4	-	-	-	60	1	24
<b><u>nucleotide metabolism: other</u></b>											
ECU06_1530	nucleoside diphosphate kinase A	U	+	21/5.8	17/6.3	72	8/28	43	117	3	22
<b><u>phosphate metabolism</u></b>											
ECU10_0340	inorganic pyrophosphatase	U	++	MF(3)	32/5.4	60	7/18	26	249	5	19
<b><u>c compound and carbohydrate metabolism</u></b>											
ECU01_0970	aldose reductase	1 U	+++	MF(2)	33/8.5	116	14/39	46	-	-	-
ECU07_1280	glucosamine fructose-6-phosphate aminotransferase	U-L	+	74/5.9	78/6.2	125	15/19	14	372	5	22
ECU01_0650	phosphoacetylglucosamine mutase	L	a	a	59/5.5	-	-	-	333	4	18
ECU11_1780	UDP-N-acetylglucosamine pyrophosphorylase	L	a	a	38/8.7	-	-	-	63	1	7
<b><u>lipid, fatty acid and isoprenoid metabolism: biosynthesis</u></b>											
ECU11_0870	ethanolamine phosphate cytidyltransferase	2 U	+++	MF(4)	37/6.0	179	19/35	63	104	2	27
<b><u>lipid, fatty acid and isoprenoid metabolism: degradation</u></b>											
ECU06_0940	acetyl-CoA-C-acyltransferase	U	+	40/6.1	42/6.4	82	14/40	33	-	-	-
<b>2, Energy</b>											
<b><u>glycolysis</u></b>											
ECU01_0240	fructose-bisphosphate aldolase	U-L	+	MF(3)	38/6.8	61	6/15	15	370	4	29
ECU05_0650	glucose-6-phosphate isomerase	L	a	a	58/6.9	-	-	-	352	5	21
ECU09_0640	pyruvate kinase	U	+	54/5.9	58/5.8	54	8/21	14	-	-	-
<b><u>pentose phosphate pathway</u></b>											
ECU08_1850	glucose-6-phosphate 1-dehydrogenase	U	++	46/5.3	50/5.4	189	20/20	40	283	5	17
<b><u>trehalose metabolism</u></b>											
ECU01_0800	alpha,alpha-trehalose-phosphate synthase, catalytic subunit	U	++	MF(2)	53/5.5	49	8/25	17	102	1	2
ECU01_0870	trehalose-6-phosphate phosphatase	N	a	a	82/6.3	-	-	-	36	1	1
ECU03_0280	UTP glucose-1-phosphate uridylyltransferase	U-L	+	MF(2)	56/6.5	69	10/24	23	105	2	4

(continues)

i Coding DNA Sequence number in the *E. Cuniculi* genome

ii Numbers in bold refer to the most abundant proteins (see Table 1 and Figure 2)

iii U = urea extract; N= NaOH extract; L= Laemmli extract

iv See Materials and methods for determination of the protein abundance

v Molecular mass (MM) is in kDa; MF(x) means "multiple forms (number of isoforms found)"

vi When several MS data were obtained for the identification of a given protein, only the one with the highest Mascot score is given.

<sup>a</sup>: not determined (Some parameters were only determined on 2D gels)

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>v</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vi</sup>		
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)
<b>3, Cell growth, cell division and DNA synthesis</b>											
<b><u>DNA replication factor A and C</u></b>											
ECU09_1330	DNA replication factor C subunit	3 U	+++	38/5.1	32/5.2	56	7/18	20	-	-	-
<b><u>recombination and DNA repair</u></b>											
ECU07_0610	DNA repair protein RAD50	U	+	200/4.1	146/5.7	48	13/28	8	-	-	-
ECU09_1390	RUVB-like DNA helicase	U	+	50/5.6	47/5.7	54	6/18	51	-	-	-
<b><u>cell cycle control and mitosis: protein phosphatases</u></b>											
ECU05_0440	ser/thr protein phosphatase 1-1, gamma catalytic subunit	U	++	22/4.6	52/6.2	-	-	-	239	2	13
<b><u>cytokinesis</u></b>											
ECU11_1950	septin	U	+	43/5.0	47/5.6	-	-	-	301	6	18
ECU09_0820	septin	U	+	32/5.3	34/5.8	91	9/12	23	175	4	15
ECU01_1370	septin-like protein	4 U	+++	MF(2)	43/5.3	-	-	-	175	2	6
<b>4, Transcription</b>											
<b><u>tRNA modification</u></b>											
ECU08_1450	N2,N2-dimethylguanosine-tRNA methyltransferase	U	+	48/5.5	48/8.6	95	7/12	20	-	-	-
<b><u>transcriptional initiation factors</u></b>											
ECU01_1330	transcription initiation factor (TFIID family)	U	a	a	15/8.7	-	-	-	70	1	10
<b><u>pre-mRNA splicing factors</u></b>											
ECU05_1440	arginine/serine-rich pre-mRNA splicing factor	U	++	MF(3)	35/6.1	76	5/11	28	-	-	-
<b><u>mRNA processing (5'-end, 3'-end processing)</u></b>											
ECU08_0380	cleavage stimulation factor	U	+	MF(2)	25/6.0	-	-	-	118	1	11
ECU10_1110	polyadenylate-binding protein 2	U	+	48/5.8	57/7.6	67	13/47	24	48	1	3
<b>5, Protein synthesis</b>											
<b><u>60S ribosomal proteins</u></b>											
ECU07_0820	acidic ribosomal protein P0 (A0)	U	+	31/8.5	32/8.8	71	9/37	25	-	-	-
ECU07_0110	acidic ribosomal protein P2 (A4)	U	a	a	12/4.2	-	-	-	61	1	11
ECU06_1120	60S ribosomal protein L36	L	a	a	12/10.9	-	-	-	90	1	36
ECU07_1820	60S ribosomal protein L35	L	a	a	14/11.3	-	-	-	68	1	16
= ECU11_2060 = ECU10_0070											
ECU03_0710	60S ribosomal protein L34	L	a	a	12/11.5	-	-	-	78	1	17
ECU04_0740	60S ribosomal protein L22	U-N	a	a	13/9.3	-	-	-	41	1	41
ECU08_2010	60S ribosomal protein L12	U	a	a	18/9.5	-	-	-	210	4	36
ECU02_0610	60S ribosomal protein L11	N	a	a	19/9.7	-	-	-	52	1	6
ECU08_1570	60S ribosomal protein L10	L	a	a	25/10.5	-	-	-	139	2	15
ECU05_0600	60S ribosomal protein L10A (L1)	U	+	32/8.0	25/8.2	59	7/23	29	179	3	21
ECU02_0750	60S ribosomal protein L7A (L8)	N	a	a	24/9.5	-	-	-	85	1	12
ECU03_1220	60S ribosomal protein L3	L	a	a	42/10.4	-	-	-	213	3	23
<b><u>40S ribosomal proteins</u></b>											
ECU06_1445	40S ribosomal protein S26	N	a	a	12/10.4	-	-	-	65	1	11
ECU10_0400	40S ribosomal protein S23	N	a	a	15/10.4	-	-	-	36	1	13

(continues)

---

**Proteomic analysis of the eukaryotic parasite *Encephalitozoon cuniculi* (Microsporidia): a reference map for proteins expressed in late sporogonial stages.**

---

Damien BROSSON <sup>a</sup>, Lauriane KUHN <sup>bc</sup>, Frédéric DELBAC <sup>a</sup>, Jérôme GARIN <sup>b</sup>, Christian P. VIVARES <sup>a</sup>, Catherine TEXIER <sup>a</sup>

<sup>a</sup> Equipe Parasitologie Moléculaire et Cellulaire, LBP, UMR CNRS 6023, Université Blaise Pascal, 63177 Aubière, FRANCE

<sup>b</sup> Laboratoire de Chimie des Protéines, ERM 201 INSERM/CEA/UJF, CEA/Grenoble Cedex 9, 38054 Grenoble, FRANCE.

<sup>c</sup> Fondation Rhône-Alpes Futur, 89 rue Bellecombe, 731 Z, 69003 Lyon Cedex 03, FRANCE

D. Brossson and L. Kuhn equally participated in this work and are both considered as first authors.

*Manuscrit soumis à la revue Proteomics*

**Abstract:**

The microsporidian *Encephalitozoon cuniculi* is a unicellular obligate intracellular parasite considered as an emerging opportunistic human pathogen. The differentiation phase of its life cycle leads to the formation of stress-resistant spores. The *E. cuniculi* genome (2.9 Mbp) having been sequenced, we undertook a descriptive proteomic study of a spore-rich cell population isolated from culture supernatants. A combination of two-dimensional electrophoresis (2-DE) and 2-DE-free techniques was applied to whole-cell protein extracts. Protein identification was performed using an automated MALDI-TOF-MS platform and a nanoLC-MS/MS instrument. A reference 2-DE map of about 350 major spots with multiple isoforms was obtained, and for the first time in Microsporidia, a large set of unique proteins

(177) including proteins with unknown function in a proportion of 25.6% was identified. The data are mainly discussed with reference to secretion and spore structural features, energy and carbohydrate metabolism, cell cycle control and parasite survival in the environment.

**1. Introduction**

Microsporidia are obligate intracellular parasites that infect a wide variety of animals ranging from insects to mammals. Phylogenetic analyses based on some conserved proteins have shown that microsporidia are related to fungi [1]. This phylum of amitochondriate unicellular eukaryotes consists in about 150 genera with more than a thousand different species. Seven genera (*Encephalitozoon*, *Enterocytozoon*, *Pleistophora*, *Trachipleistophora*, *Vittaforma*,

Table S1 (continues)

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>iv</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vi</sup>			
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)	
<b>40S ribosomal proteins (continues)</b>												
ECU06_1110	40S ribosomal protein S18	U	a	a	17/10.6	-	-	-	66	1	9	
ECU02_0770	40S ribosomal protein S17	U	a	a	14/9.0	-	-	-	58	1	23	
ECU03_0650	40S ribosomal protein S14	L	a	a	14/11.1	-	-	-	50	1	13	
ECU08_1060	40S ribosomal protein S13	L	a	a	17/10.4	-	-	-	35	1	7	
ECU05_0920	40S ribosomal protein S9	N	a	a	21/9.7	-	-	-	47	1	6	
ECU11_0780	40S ribosomal protein S7	U	++	21/8.3	20/9.1	71	10/31	50	137	3	22	
ECU04_0140	40S ribosomal protein S5	U-L	a	a	23/10.6	-	-	-	172	2	6	
ECU08_0870	40S ribosomal protein S4	N	a	a	30/9.3	-	-	-	34	1	7	
ECU09_1250	40S ribosomal protein S3	5	U	+++	MF(3)	26/9.2	62	9/39	34	58	1	
ECU04_0450	40S ribosomal protein SA ou P40	U	+	32/5.7	28/5.8	131	13/21	44	-	-	-	
ECU11_1620	40S ribosomal protein S19	U	+	20/6.0	20/7.0	65	7/17	34	148	3	25	
<b>translation: initiation</b>												
ECU09_1200	translation initiation factor eIF4A	U-N-L	+	50/5.4	49/5.6	138	24/30	49	94	2	12	
<b>translation: elongation</b>												
ECU04_1100	translation elongation factor eEF1 $\alpha$	6	U-N-L	+++	MF(4)	56/9.1	156	22/25	44	527	8	
ECU11_1460	translation elongation factor eEF2	U-L	+	84/6.7	94/7.0	75	11/24	14	375	7	18	
<b>translational control</b>												
ECU10_0620	nonsense-mediated mRNA decay protein	L	a	a	107/4.8	-	-	-	101	1	10	
<b>tRNA synthetases</b>												
ECU08_0550	arginyl-tRNA synthetase	L	a	a	64/5.9	-	-	-	55	1	8	
ECU08_0490	cysteinyl-tRNA synthetase	U	+	57/6.0	56/6.3	58	7/18	11	-	-	-	
ECU10_1790	glycyl-tRNA synthetase	U	++	68/6.0	67/6.2	131	13/19	21	-	-	-	
ECU02_1360	prolyl-tRNA synthetase	U	+	59/6.0	59/6.9	109	14/20	28	-	-	-	
ECU04_0750	seryl-tRNA synthetase	U	+	52/5.9	49/6.2	39	15/51	22	-	-	-	
ECU11_0530	tryptophanyl-tRNA synthetase	7	U-L	+++	MF(2)	44/6.4	175	22/35	48	83	2	
<b>6, Protein destination</b>												
<b>protein folding and stabilization : T Complex Protein 1 (TCP1)</b>												
ECU03_0220	TCP1 alpha subunit	U	++	MF(3)	58/6.3	67	9/24	16	85	1	9	
ECU09_0480	TCP1 beta subunit	U	++	MF(2)	56/6.3	232	24/29	45	-	-	-	
ECU10_0240	TCP1 gamma subunit	U	++	MF(2)	57/5.7	100	13/15	24	-	-	-	
ECU02_0520	TCP1 delta subunit	U-L	+	51/5.6	53/5.7	-	-	-	68	1	3	
ECU06_0990	TCP1 epsilon subunit	U	++	MF(3)	58/4.9	61	7/18	22	-	-	-	
ECU10_0630	TCP1 eta subunit	U	++	MF(2)	56/5.8	148	17/19	34	334	5	20	
ECU04_1020	TCP1 theta subunit	L	a	a	54/8.3	-	-	-	56	2	4	
<b>protein folding and stabilization : chaperones</b>												
ECU11_1830	heat-shock protein, HSP70 family	L	a	a	76/5.1	-	-	-	56	1	4	
ECU03_0520	heat-shock protein, HSP70 family	8	U-N-L	+++	MF(7)	75/5.4	185	20/23	34	694	12	
ECU02_1100	heat-shock protein HSP 90	U-L	+	75/4.9	79/5.1	209	26/37	36	727	1	14	
ECU02_0100	HS70 protein cognate 4 (HSC70)	9	U	+++	MF(6)	76/5.2	215	31/32	43	404	7	

(continues)

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>iv</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vi</sup>			
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)	
<b>protein folding and stabilization : isomerases</b>												
ECU08_0470	peptidylprolyl <i>cis-trans</i> isomerase (cyclophilin)	U	++	MF(3)	19/6.6	75	8/14	44	195	6	48	
ECU09_1560	protein disulfide isomerase	U	+	49/4.7	50/4.9	-	-	-	68	1	3	
<b>protein modification: acetylation</b>												
ECU07_0750	histone acetyltransferase type B, subunit	U	+	49/4.7	43/4.9	-	-	-	56	1	5	
<b>protein modification : glycosylation</b>												
ECU05_0260	phosphomannomutase	U	++	MF(2)	30/5.8	147	14/26	43	-	-	-	
<b>protein modification : prenylation</b>												
ECU09_1870	dehydrodolichyl diphosphate synthase ( <i>cis</i> -prenyltransferase)	U	+	34/5.1	31/9.5	62	6/17	20	-	-	-	
<b>protein modification: processing</b>												
ECU10_0750	methionine aminopeptidase 2	U	+	41/5.8	40/6	-	-	-	299	4	27	
<b>proteolyse: cytoplasmic degradation: 26S proteasome</b>												
ECU11_1670	proteasome alpha-type subunit	U	++	29/6.0	27/6.5	101	11/21	35	-	-	-	
ECU07_1420	proteasome alpha-type subunit	U	+	28/5.2	33/5.5	-	-	-	128	1	4	
ECU09_0330	proteasome alpha-type subunit	U	++	30/5.7	27/5.8	130	13/22	54	-	-	-	
ECU05_1340	proteasome alpha-type subunit	U	++	30/5.5	26/5.9	89	12/19	38	-	-	-	
ECU05_1400	proteasome alpha-type subunit	U	++	32/5.9	26/6.3	96	11/19	41	-	-	-	
ECU05_0290	proteasome beta-type subunit	U	+	29/5.7	26/6.0	66	9/22	30	249	4	10	
ECU10_1450	proteasome alpha-type subunit	U	a	22/7.6	-	-	-	-	95	1	10	
ECU09_0720	proteasome beta-type subunit	U	+	30/6.1	25/6.7	-	-	-	245	3	20	
ECU08_0280	proteasome beta-type subunit	U	+	26/5.4	22/5.7	73	10/37	31	-	-	-	
ECU02_0480	proteasome regulatory subunit 1	L	a	a	88/4.9	-	-	-	47	1	4	
ECU05_1540	proteasome regulatory subunit 3	L	a	a	43/8.2	-	-	-	53	1	10	
ECU04_0310	proteasome regulatory subunit 4	L	a	a	89/5.0	-	-	-	41	1	1	
ECU07_1640	proteasome regulatory subunit 4	U	+	49/6.3	47/6.4	129	12/12	35	-	-	-	
ECU10_1420	proteasome regulatory subunit 6A (tat-binding protein 1)	U	++	45/5.9	45/6.1	146	16/23	44	-	-	-	
ECU08_1970	proteasome regulatory subunit 6	U	+	44/8.4	43/8.6	101	14/32	28	-	-	-	
ECU10_0130	= ECU10_0130											
ECU09_1840	proteasome regulatory subunit 8	U	+	41/8.5	51/8.4	-	-	-	171	3	6	
ECU11_0570	proteasome regulatory subunit 11	U	+	33/6.2	33/6.7	-	-	-	109	2	9	
ECU01_1230	protein of the CDC48/PAS1/SEC28 family of ATPases (AAA)	10	U-L	+++	MF(3)	87/5.8	125	18/20	27	131	3	
<b>proteolyse: cytoplasmic degradation : ubiquitin conjugating enzyme E2</b>												
ECU01_1010	ubiquitin-conjugating E2 subunit	11	U	+++	MF(2)	25/5.3	51	7/41	21	-	-	
<b>proteolyse: other</b>												
ECU10_1770	cytosolic aminopeptidase (X-Leu/X-Pro)	U-L	++	MF(3)	60/7.6	181	29/38	36	332	4	8	

**7, Transport facilitation****ABC transporters**

ECU11\_1340 similarity to yeast ABC transporter-like fragment

L a a 66/6.0 - - - 41 1 6

(continues)

*Brachiola* and *Nosema*) are currently known to contain human-infecting species with a high seroprevalence in immunocompetent populations [2]. These emerging human pathogens are responsible for various digestive and nervous clinical syndromes in immunocompromised AIDS and organ transplant patients [3]. Microsporidia display an original invasion mechanism, involving the polar tube, a cylindrical and very long structure that is coiled within the spore. It can be quickly extruded at the anterior pole of the spore and is used to transfer the spore contents or sporoplasm into a target cell. The cycle of the internalized parasite comprises a proliferation phase or merogony (binary or multiple divisions), followed by a multi-step differentiation phase or sporogony: (i) meront-to-sporont transformation (onset of cell wall formation), (ii) sporont division leading to sporoblasts and (iii) conversion of sporoblasts into spores [4]. According to the considered genus, the intracellular development takes place either in direct contact with host cytoplasm, inside a secreted vesicle-like structure (sporophorous vesicle) or less frequently inside a membrane-bounded vacuole (parasitophorous vacuole). Finally, the spores equipped with a complete cell wall and invasion apparatus are released in the environment by host cell disruption.

The *Encephalitozoon* genus includes three species (*E. cuniculi*, *E. hellem*, *E. intestinalis*) that are uninucleate and multiply via binary division inside parasitophorous vacuoles [4].

*Encephalitozoon cuniculi* was the first eukaryotic parasite the nuclear genome of which was completely sequenced [5], providing an insight into its potential metabolism. Its minimal genome among eukaryotes (~2.9 Mbp) comprises about 2000 coding sequences, 55% of these corresponding to proteins of unknown or hypothetical function. The nuclear genome of *Antonospora (Nosema) locustae*, a microsporidia parasite of insects, is in progress (“*Nosema locustae* Genome Project” at <http://jbpc.mbl.edu/Nosema/>). Other genome sequences are now available for several protistan parasites including apicomplexans [6-11], kinetoplastids [12-14] and the amitochondriate *Entamoeba histolytica* [15].

With these increasing genome projects, parasitological problems can be addressed by proteomic analysis [16-18] using either the conventional strategy which involves two-dimensional polyacrylamide gel electrophoresis (2-DE) and protein identification by mass spectrometry, or the Multidimensional Protein Identification Technology (MudPIT) which is a LC-MS/MS particular implementation devoted to 2-DE-free shotgun proteomics [19, 20]. As regard to apicomplexan parasites, the malaria agent *Plasmodium falciparum* was subjected to various proteomic studies providing a wealth of information on protein expression during life cycle [21, 22], novel surface proteins [23], host-parasite interactions with delineation of

Table S1 (continues)

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>v</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vi</sup>			
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)	
<b>8, Intracellular transport</b>												
<u>nuclear transport</u>												
ECU04_1560	nuclear GTP binding protein	U	++	28/6.3	24/6.9	64	9/26	35	286	4	14	
ECU02_0140	ran-specific GTPase-activating protein	U	+	30/6.0	25/6.3	110	9/19	39	-	-	-	
<u>vesicular transport: ADP ribosylation factor (ARF)</u>												
ECU08_1930 = ECU10_0090	GTP-binding protein, ADP ribosylation factor family	U	++	26/5.8	24/6.3	78	8/15	28	110	4	24	
ECU05_0090	GTP-binding protein, SAR-type, ARF family	12 U	+++	MF(2)	25/5.4	81	9/32	49	253	5	32	
<u>vesicular transport: COP I coat</u>												
ECU08_0680	coatamer complex, zeta subunit	U	++	22/4.6	19/5.4	53	7/31	30	22	1	15	
<u>vesicular transport: COPII coat</u>												
ECU11_0790	COPII, protein transport SEC23	L	a	a	81/5.8	-	-	-	59	1	7	
<u>vacuolar transport: vacuolar H<sup>+</sup>-ATPase (V0V1)</u>												
ECU11_1280	H <sup>+</sup> -ATPase V <sub>1</sub> , A subunit	U	+	63/5.4	73/8.5	149	22/27	33	-	-	-	
ECU10_1040	H <sup>+</sup> -ATPase V <sub>1</sub> , B subunit	U	++	MF(3)	53/5.1	80	14/33	32	212	3	14	
<u>vacuolar transport: other</u>												
ECU03_0900	vacuolar protein sorting-associated protein	L	a	a	59/6.0	-	-	-	36	1	3	
<u>other intracellular transport activities</u>												
ECU07_0760	DNAJ homolog 2 involved in mitochondrial protein import	U	+	43/8.5	44/9.0	107	15/36	31	-	-	-	
ECU03_1010	similarity to 14.3.3 protein	U	++	MF(2)	30/5.1	157	16/31	45	-	-	-	
<b>9, Cellular organization and biogenesis</b>												
<u>cell surface</u>												
ECU09_1950	similarity to mammalian integral membrane protein	N-L	a	a	20/9.0	-	-	-	111	1	7	
ECU10_1660	spore wall protein 1 (SWP1)	13 U-N-L	+++	50/4.0	46/5.1	66	13/17	12	157	4	8	
ECU11_1320	similarity to Plasmodium membrane protein	U	a	a	18/4.5	-	-	-	46	1	9	
<u>Cytoskeleton</u>												
ECU01_0460	actin	14 U-N-L	+++	MF(5)	42/5.8	146	14/18	41	535	9	40	
ECU04_1000	myosin heavy chain, non muscle type A	U	++	MF(2)	199/6.0	74	19/34	9	-	-	-	
ECU03_0820	beta-tubulin	U	++	49/4.9	52/5.1	49	13/27	21	-	-	-	
<u>chromosome structure</u>												
ECU02_0720	histone H2A	U	a	a	14/9.7	-	-	-	90	2	17	
ECU08_0410	histone H2B	U-N	a	a	14/10.0	-	-	-	195	2	26	
<u>Other</u>												
ECU06_0250	polar tube protein 1 (PTP1)	15 U	+++	MF(3)	37/4.8	-	-	-	96	2	18	
ECU06_0240	polar tube protein 2 (PTP2)	16 U-L	+++	MF(4)	30/8.7	139	17/38	35	223	3	24	
ECU11_1440	polar tube protein 3 (PTP3)	U	+	MF(2)	141/6.0	70	9/9	11	-	-	-	
ECU01_0840	similarity to chloroplast 20kDa-RNP	U	a	a	18/9.1	-	-	-	34	1	6	

(continues)

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>v</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vi</sup>			
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)	
<b>10, Cellular communication / signal transduction</b>												
<u>guanine nucleotide binding proteins</u>												
ECU08_1110	guanine nucleotide binding protein, beta subunit	U	++	MF(3)	37/5.6	102	15/19	34	117	2	30	
<u>Other</u>												
ECU04_1480	ZRP1	17 U	+++	MF(5)	24/5.4	90	10/14	48	-	-	-	
<b>11, Cell rescue, defence, cell death and ageing</b>												
<u>DNA repair (direct repair, base excision...)</u>												
ECU05_1590 = ECU11_0140	3-methyl-adenine DNA glycosylase	18 U	+++	MF(2)	24/9.4	56	11/34	29	87	1	17	
ECU08_0880	DNA lyase, endonuclease 3	U	++	33/8.5	27/9.4	49	7/21	30	-	-	-	
<u>Detoxification</u>												
ECU02_0440	glutathione peroxidase	U	a	a	20/9.0	-	-	-	102	1	14	
ECU02_0580	hydroxyacyl glutathione hydrolase	19 U	+++	MF(4)	30/8.0	120	12/14	45	51	1	8	
ECU01_0930	thioredoxin H-type	20 U	+++	MF(2)	24/9.1	72	7/9	30	119	2	27	
ECU03_1190	thioredoxin peroxidase	U	++	MF(3)	20/5.4	45	6/26	34	101	2	18	
<b>12, Unassigned role</b>												
ECU01_0820	cysteine-rich protein (endospore protein EnP1, Peuvél-Fanget <i>et al.</i> , Int. J. Parasitol. in press)	21 U-N-L	+++	17/5.7	41/9.3	-	-	-	512	4	42	
ECU01_1270	serine-rich protein (endospore protein EnP2, Peuvél-Fanget <i>et al.</i> , Int. J. Parasitol. in press)	U-L	+	MF(2)	23/8.5	68	9/20	31	204	2	11	
ECU02_0390	possible nuclease for DNA repair	L	a	a	60/8.2	-	-	-	52	1	1	
ECU02_1470	similarity to integral membrane proteins (YIPC_ yeast)	L	a	a	18/7.7	-	-	-	47	2	11	
ECU03_1070	domain: transcription initiation factor TFIIA	U	a	a	18/5.0	-	-	-	40	1	8	
ECU03_1610	domains: late embryogenesis abundant protein (in higher plant seed embryos and under dehydration stress)	U	++	MF(2)	18/8.8	-	-	-	84	2	12	
ECU06_0650	possible adenylate kinase	U	+	20/5.3	24/9.1	-	-	-	220	3	26	
ECU06_0670	possible protein kinase	L	a	a	31/6.6	-	-	-	106	2	12	
ECU07_0040 = ECU09_2030	member of a multigenic subtelomeric family / domain: MASE1 (transduction)	L	a	a	31/8.6	-	-	-	49	1	4	
ECU07_0920	domain: exonuclease-phosphatase family	22 U	+++	MF(4)	35/5.7	122	19/29	54	102	2	6	
ECU08_0270	domain: DUF220 (unknown function)	L	a	a	45/5.2	-	-	-	36	1	7	
ECU08_0330	possible GTP-binding protein	U	+	MF(2)	45/9.6	45	11/33	27	61	1	3	
ECU08_0540	domain: C4-dicarboxylate/malic acid transport family	L	a	a	32/5.6	-	-	-	102	2	14	
ECU08_1610	protein of the UPF0023 family (hypothetical proteins)	23 U	+++	MF(2)	26/8.5	71	9/31	28	-	-	-	
ECU08_1730	similarity to ECU08_1700, ECU08_1710 and ECU08_1720 / domain: ricin B lectin (QxW lectin repeat)	U	++	23/6.1	23/6.1	54	7/24	38	145	3	18	
ECU08_2020	member of a multigenic subtelomeric family / domain: ricin B lectin (QxW lectin repeat)	24 U	+++	MF(4)	29/6.8	73	9/29	32	99	1	8	
ECU09_1370	possible translation initiation factor 5A (eIF-5A)	U	+	22/5.8	17/6.7	115	9/15	60	98	2	17	
ECU09_1460	WD repeat domain (G beta repeat)	U	+	33/5.6	34/5.8	29	3/10	15	-	-	-	
ECU09_1470	possible RNA binding protein	25 U	+++	27/8.1	18/9.6	52	5/10	33	135	2	23	
ECU10_0880	possible GTPase activator protein	L	a	a	35/6.0	-	-	-	202	3	18	

(continues)

T cell epitopes [24] and drug-induced proteome alterations [25]. In *Toxoplasma gondii*, a preliminary proteome map was assigned to the tachyzoite stage [26] and the content of the apical secretory organelle was described [27] while host serological and T-cell responses were investigated [28, 29]. 2-DE reference maps were elaborated for tachyzoites and antigens in *Neospora caninum* [30, 31] or sporozoites and microneme secretory organelles in *Eimeria tenella* [32, 33]. In the case of kinetoplastids, *Leishmania* proteomic studies focused on protein extraction methods [34], 2-DE map elaboration [35, 36], and stage differentiation [37, 38]. A preliminary 2-DE study was reported for *Trypanosoma brucei* [39] and protein profiles characterizing four different cell stages were described in *T. cruzi* [40, 41]. Post-genomic experimental investigations of *E. cuniculi* proteins [reviewed in 42] and proteomic data relevant to other microsporidia are rather limited. Three insect-infecting species (*Antonospora locustae*, *Nosema bombycis* and *Vairimorpha necatrix*) were distinguished by comparing their 2-DE spore protein patterns [43]. More recently, about 200 protein spots were separated by 2-DE from spores of the human pathogen *E. intestinalis* and 95 spots were shown to be immunogenic in the rabbit [44], but none was identified at the sequence level. MALDI-TOF mass spectrometry analysis was evaluated for a rapid discrimination of microsporidian

species on the basis of peptide biomarkers in the 2-8-kDa range [45].

Exploiting the genome database [5], the present proteomic study provides a reference map of the major proteins expressed during late sporogony of *E. cuniculi*. In addition, proteins of unknown function have been screened for new cell wall candidates.

## 2. Materials and methods

### 2.1. Microsporidian spore production and protein extraction

The *Encephalitozoon cuniculi* isolate GB-M1 used for genome sequencing, was cultured in Madin-Darby Canine kidney (MDCK) or Human Foreskin Fibroblast (HFF) cells as previously described [46]. After collection of culture supernatants, parasites were sedimented (5,000xg, 5 min) and heated at 65°C in a 1% SDS solution. A spore-rich fraction was recovered after several washes in water to remove host cell debris. Parasites were disrupted in a lysis buffer ( $10^9$ - $10^{10}$  cells/mL) containing 100 mM DTT, 4% CHAPS and 0.2% SDS, by repeated cycles of freezing-thawing and sonication (Deltasonic 1320, 300 W, 28 kHz). Proteins of broken cells were first extracted with a solution containing 7 M urea, 2 M thiourea, 100 mM DTT, 4% CHAPS and 0.2% SDS for 6 h at room temperature. After centrifugation (20,800xg, 5 min), the supernatant was collected ("urea extract"). The sedimented material was treated with 30 mM NaOH overnight at 4°C and centrifuged as before to

Table S1 (continues and ends)

CDS number <sup>i</sup>	Protein function <sup>ii</sup>	Extract <sup>iii</sup>	Abundance <sup>iv</sup>	MM/pl exp <sup>v</sup>	MM/pl calc <sup>v</sup>	MALDI-TOF-MS <sup>vi</sup>			LC-MS/MS <sup>vii</sup>			
						Mascot Score	peptides	coverage (%)	Mascot Score	peptides	coverage (%)	
<b>Unassigned role (continues)</b>												
ECU11_0510	possible chitin deacetylase ( <i>EcCDA</i> , Brosson <i>et al.</i> , FEMS Microbiol. Lett. 2005, 247, 81-90)	26	U	+++	MF(2)	28/4.7	120	9/10	39	61	1	5
ECU01_0250			U	+	MF(2)	19/5.5	69	5/10	30	59	2	16
ECU01_0420		27	U	+++	32/8.5	28/9.2	107	13/37	42	-	-	-
ECU01_0440			U	++	22/4.6	15/5.1	51	5/19	32	64	2	24
ECU01_0910			U	a	18/9.7	-	-	-	-	59	1	2
ECU02_0150	Located at the cell wall (this work)		U	++	MF(3)	25/5.0	57	7/11	28	-	-	-
ECU03_1535			U	++	18/5.7	11/5.8	-	-	-	135	4	39
ECU04_1080			L	a	13/6.3	-	-	-	-	41	1	11
ECU05_0110			U	+	45/6.0	36/6.5	-	-	-	239	3	10
ECU05_0140			U-L	a	51/9.1	-	-	-	-	386	5	34
ECU05_0510		28	U	+++	27/4.8	27/5.1	58	9/22	24	-	-	-
ECU05_0590		29	U-L	+++	MF(2)	28/6.0	94	7/17	32	56	1	4
ECU05_1040			N	a	29/8.6	-	-	-	-	85	2	12
ECU07_0400	Located at the Golgi network (this work)	30	U	+++	MF(2)	23/5.6	115	12/23	49	117	2	35
ECU07_0530			L	a	48/8.5	-	-	-	-	73	1	17
ECU07_1090			U	++	37/6.5	31/6.6	67	8/8	38	-	-	-
ECU07_1120			U	++	MF(2)	49/5.2	49	7/21	15	54	2	15
ECU08_1280			U	++	50/5.3	52/5.4	54	8/31	21	-	-	-
ECU08_1550			U	a	11/9.3	-	-	-	-	76	2	18
ECU09_1400			U	++	33/4.7	29/4.8	84	11/27	31	-	-	-
ECU09_1820		31	U	+++	MF(3)	26/6.5	115	12/15	50	-	-	-
ECU09_1880			U	+	34/5.7	47/6.0	72	6/10	11	-	-	-
ECU10_1070			U	+	22/4.6	11/4.4	-	-	-	75	1	31
ECU11_1210			U	++	MF(2)	33/5	97	14/27	42	-	-	-
ECU11_1270			U	+	46/5.2	46/5.4	-	-	-	69	1	3

collect the supernatant ("NaOH extract"). The NaOH-insoluble material was finally boiled for 10 min in Laemmli solution containing 2.5% SDS, 0.125 M Tris HCl pH 6.8, 20% glycerol, 2 mM EDTA and 100 mM DTT ("Laemmli extract"). After quantification using the Plus-One 2D Quant Kit (Amersham), the different protein samples were characterized by SDS-PAGE on 12% polyacrylamide gels and stored at -20°C.

## **2.2. Two-dimensional electrophoresis (2-DE)**

Isoelectrofocalisation (IEF) was performed along linear immobilized pH gradient strips of 17 cm (BioRad) in rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 2 mM tributyl phosphine (TBP) and 0.5% ampholytes) with the IPGPhor apparatus (Amersham). Migration was stopped at 50-55000 Vh and 85-90000 Vh for pH 3-10 and pH 5-8 IEF gradient, respectively. Before SDS-PAGE on 10-12% polyacrylamide gels (18 x 18 cm), strips were equilibrated with 50 mM Tris HCl pH 8.8, 6 M urea, 30% glycerol, 2% SDS and 100 mM DTT, and then 135 mM iodoacetamide. Gels were fixed in 7.5% acetic acid / 30% ethanol, stained with Coomassie Blue (BioRad), destained with 30% ethanol and stored at 4°C in water until image analysis and spot excision.

## **2.3. 2-DE gel analysis**

2-DE gel analysis was performed using the PDQuest 6.2.1 software (BioRad). The molecular mass and pI assigned to individual spots were evaluated by comparison with 2-

DE SDS-PAGE standards (BioRad). The relative abundance of each spot (RA) in a given gel was determined by dividing its intensity by the intensity of a reference spot that was shared by all 2-DE profiles. The relative abundance of each identified protein was calculated by summing the relative abundances of all the spots ( $\Sigma$ RA) containing this protein and was coded as either low (+:  $\Sigma$ RA <3), intermediate (++:  $3 \leq \Sigma$ RA < 10) or high (+++:  $\Sigma$ RA  $\geq$ 10).

## **2.4. Tryptic digestion for mass spectrometry analysis**

Protein spots were manually excised from Coomassie blue-stained 2-DE gels and transferred in 96-well-microtitration plates. Sample preparation was fully automated (Mass Prep Station Multiprobe II, Micromass) as previously described [47]. Briefly, after destaining and dehydration, excised proteins were alkylated with iodoacetamide and digested with 0.1  $\mu$ g/ $\mu$ L modified trypsin (Promega, sequencing grade) in 25 mM  $\text{NH}_4\text{HCO}_3$  for 5 h at 37°C.

## **2.5. MALDI-TOF-MS analysis and database search**

MALDI-TOF-MS analysis was carried out as previously described [47] using a MALDI-TOF mass spectrometer (Autoflex, Bruker Daltonik) in reflector mode over a 0-4200 Da mass range. Spectrum acquisition and annotation were done automatically. The *E. cuniculi*-specific database was explored using an intranet version of Mascot 1.7, with peptide masses assumed to be monoisotopic





and  $[M+H]^+$  (protonated molecular ions), and allowing some peptide modifications (cysteine carbamidomethylation, protein N-acetylation and methionine oxidation). The maximum number of missed cleavages was set to 1 and the mass tolerance to  $\pm 150$  ppm. Proteins showing a score higher than 70 and an average coverage higher than 20% were automatically validated. If not, they were analyzed by nanoLC-MS/MS.

## 2.6. NanoLC-MS/MS analysis and database search

Peptides were extracted automatically by diffusion in 50% acetonitrile for 15 min, then in 5% formic acid and finally in 100% acetonitrile (Mass Prep station). After drying, extracted peptides were solubilized in water containing 2% acetonitrile and 5% trifluoroacetic acid before being transferred to vials compatible with nanoLC-MS/MS analysis (CapLC and Q-TOF Ultima, Waters). Nano-liquid chromatography was run on a 300  $\mu\text{m}$  x 5 mm PepMap C18 precolumn for peptide concentration and a 75  $\mu\text{m}$  x 150 mm C18 column for peptide elution. Chromatography conditions consisted in a 75-min run, at a 200 nL/min flow rate with a gradient of acetonitrile/water/formic acid (A: 2% / 98% / 0.1% and B: 80% / 20% / 0.1%). Spectra were calibrated using the glu-fibrino peptide fragmentation in MS/MS mode. MS and MS/MS data were acquired and processed automatically with the MassLynx 3.5 software. Consecutive searches against the specific database from *E. cuniculi* were

performed with Mascot 1.7. Peptides showing a score higher than 40 were automatically validated. Peptides with scores between 20 and 40 were checked manually. The remaining unassigned peptide sequences were interpreted manually.

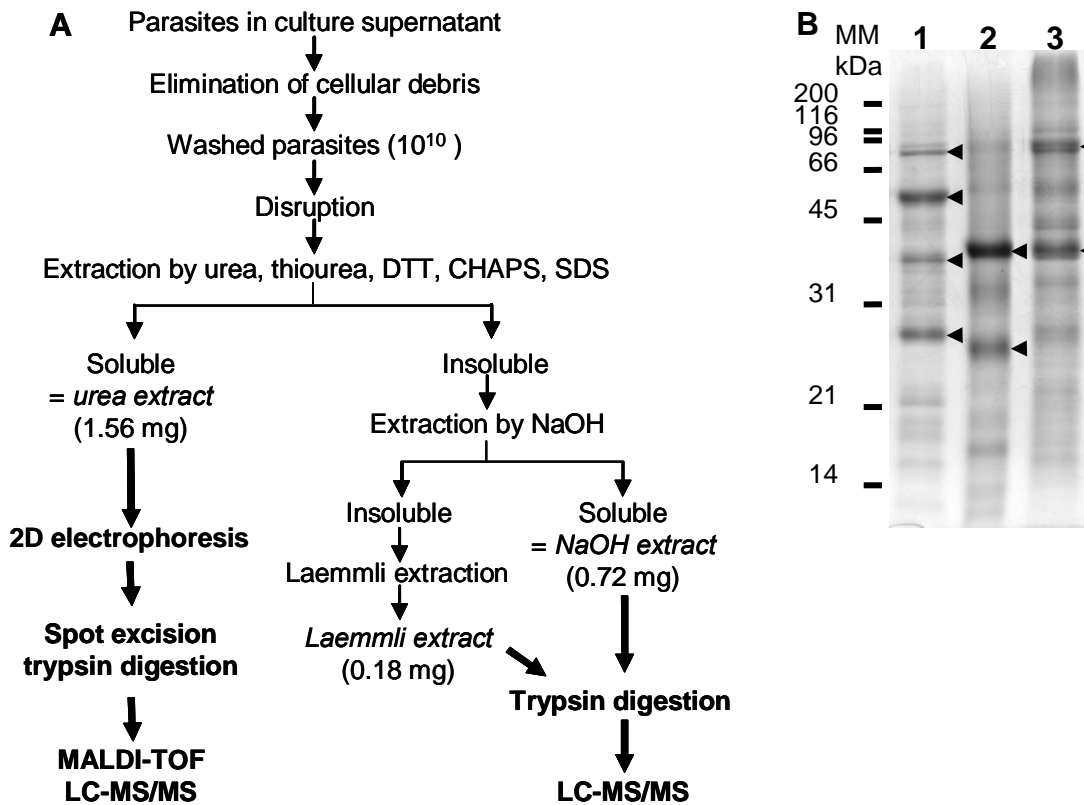
## 2.7. “Shotgun” identification

For identification by a “shotgun” strategy (without 2-DE), SDS-PAGE migration was stopped when proteins reached the boundary between concentration and separation gels [48]. Protein material was excised as a single band and subjected to trypsin digestion. Resulting peptide mixture was analysed by nanoLC-MS/MS before database searches as previously explained.

## 2.8. Sequence analysis

Physico-chemical parameters of the protein sequence were calculated with ProtParam ([www.expasy.org](http://www.expasy.org)). The SignalP program version 2 ([www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), the HMMTOP program version 2.0 ([www.enzim.hu/hmmtop](http://www.enzim.hu/hmmtop)) and the PSORT II algorithm ([www.psort.nibb.ac.jp/form2.html](http://www.psort.nibb.ac.jp/form2.html)) were used to predict signal peptide, transmembrane helices and protein localisation, respectively. Searches for protein homologies and protein domains were done using the BLAST program ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and the Pfam database ([www.sanger.ac.uk/Software/Pfam/search.shtml](http://www.sanger.ac.uk/Software/Pfam/search.shtml)), respectively.

## 2.9. Recombinant protein and polyclonal antibody production



**Figure 15.** A) Phase contrast microscopy of the purified *E. cuniculi* cells. Refrangent mature spores (*Sp*) are mixed with late sporoblasts (*Sb*). Bar:  $1\mu\text{m}$  B) Extraction procedure (for details, see Material and methods). C) SDS-PAGE profiles of protein extracts ( $7\mu\text{g}$  per lane). The urea extract (lane 1) is characterized by four major bands (arrowheads) at 70, 50-55, 35 and 28 kDa. Each of the two other extracts present two intense bands: 40 and 26 kDa for NaOH extract (lane 2), 70 and 40 kDa for Laemmli extract (lane 3).

The ECU07\_0400 (from residue 19 to residue 207) and ECU02\_0150 (the 26-221 sequence) proteins were expressed as truncated forms in fusion with glutathione-S-transferase (GST) at the N-terminal extremity and an 8xHis-tag at the C-terminal extremity. PCR reactions were performed according to standard conditions, and purified products were cloned in frame with GST into the prokaryotic expression vector pGEX-4T1 (Pharmacia) modified to provide a C-terminal 8xHis-tag after the *XhoI* site. Resulting recombinant plasmids were individually introduced in the *E. coli* BL21+ strain. After induction with IPTG, bacterial proteins were extracted with Laemmli solution, analysed by SDS-PAGE and used to produce antiserum in SWISS mice. Mice were injected intraperitoneally with samples homogenized with Freund's adjuvant (Sigma). Sera were collected 2 weeks after the last injection and stored at -20°C.

### **2.10. Transmission electron microscopy (TEM) immunocytochemistry**

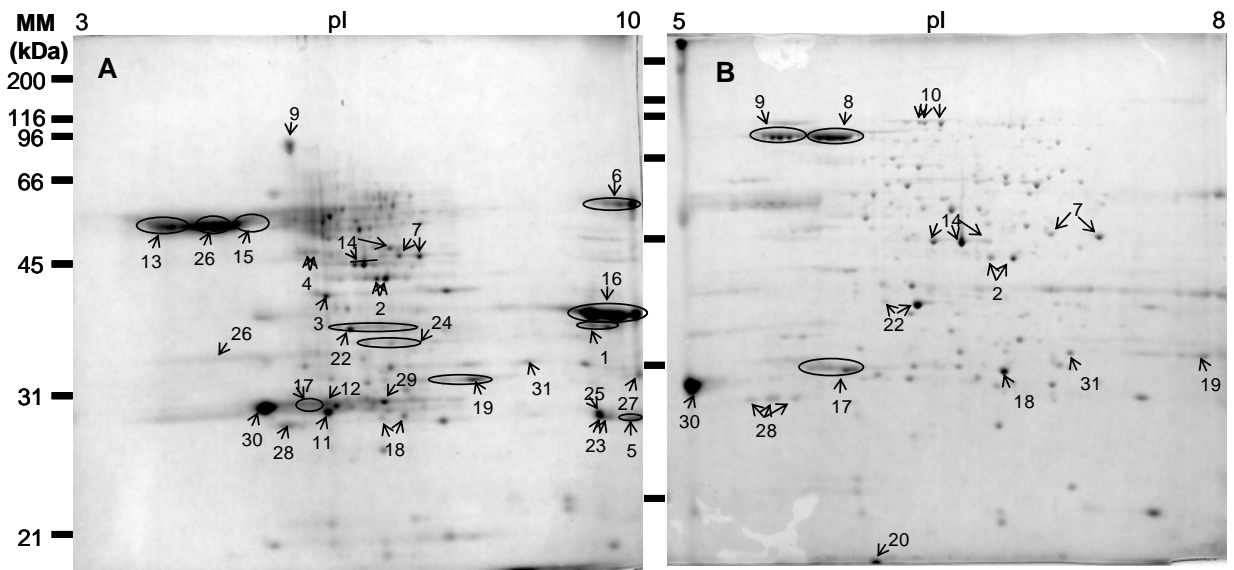
TEM immunolabelling experiments were done with *E. cuniculi*-infected MDCK cells as previously described [47]. Briefly, 85-nm ultrathin sections were performed with the FCS system on the Leica UltraCutS ultramicrotome before transfer on collodion-coated nickel grids. After blocking with BSA, sections were incubated with various dilutions of polyclonal antibodies, and then with protein A or goat anti-mouse antibody conjugated with 5-nm colloidal gold particles

(Sigma). After washing, the grids were stained and embedded with 0.5% aqueous uranyl acetate-1.6% methylcellulose prior to examination with a JEOL 1200EX transmission electron microscope.

## **3. Results and discussion**

The large-scale separation of different *Encephalitozoon cuniculi* cell stages from infected adherent mammalian cells is still a difficult task, especially because this small-sized parasite (mean spore size: 2.0-2.5 x 1.0-1.5 µm) develops asynchronously inside parasitophorous vacuoles and wall-lacking merogonial stages are in close contact with the vacuolar membrane [49]. This led us to restrict the present study to the sporadic stage using a procedure of purification from culture supernatants. Large amounts of spores are indeed released in the culture medium following natural disruption of heavily infected host cells. In fact, TEM observations on fractions derived from *E. cuniculi* culture supernatants have revealed the presence of other sporogonial stages (sporonts, early and late sporoblasts) [50]. Phase contrast microscopy examination of the parasite population here purified also showed a significant heterogeneity, the most frequent cells being spores and late sporoblasts (not shown). Thus, more appropriate terms to designate this population may be either "late sporogonial stages" or "sporoblast-to-spore stages".

### **3.1. Differential protein extractions**



**Figure 16.** Two-dimensional electrophoresis maps of proteins (urea extract) from a spore-rich *E. cuniculi* fraction. Isoelectric focusing was carried out using either a pH 3-10 (A) or a pH 5-8 (B) gradient. Numbers refer to abundant proteins (see also Table 15).

Protein extraction needed harsh breakage conditions combined with sequential extraction treatments (Figure 15A). First, we applied a highly denaturing mixture composed of detergents and high molarity chaotrope and reducing agents (“urea extract”). The alkaline treatment of the insoluble material (“NaOH extract”) was known to solubilize fungal spore wall proteins bounded to polysaccharides by their O-glycosylated moities or by other uncharacterized bonds [51]. Finally, alkali-resistant proteins were solubilized by boiling in standard Laemmli solution (“Laemmli extract”). The three extracts exhibited rather different SDS-PAGE patterns (Figure 15B).

### 3.2. Protein separation and mass spectrometry identification strategies

Two identification strategies were used depending on the protein extract (Figure 15A). The urea extract was analyzed by a 2-DE proteomic method. After 2-DE separation, Coomassie blue stained spots were manually excised, and transferred to a highly automated MALDI-TOF chain. The automation was indeed applied to the steps of spot preparation, trypsin digestion and spotting of the digested samples on the MALDI target, as well as to data acquisition and interpretation. For some spots, the identification was confirmed or obtained by LC-MS/MS analysis. Wide pH gradient gels (pH 3-10) reproducibly revealed about 140 different spots, 70% of these being comprised in the 25-100 kDa / pI 5-8 window (Figure 16A).

Further focusing on this window revealed about 250 spots (Figure 16B). Major spots distributed throughout the ~350-spot reference map were subjected to mass spectrometry. As a result, a total of 139 different proteins were identified in the urea extract. The NaOH and Laemmli extracts were roughly analyzed by a 2-DE-free shotgun proteomic strategy. Each extract was concentrated in one unique acrylamide-gel band and digested with trypsin and the peptide mixture was processed using LC-MS/MS. We identified 17 proteins in the NaOH extract and 50 proteins in the Laemmli one. These two extracts shared 8 and 21 proteins, respectively, with the urea one.

Thus, 177 different proteins associated with the *E. cuniculi* spore-rich fraction have been revealed in this study (Table S1 in supplementary material), 31 highly abundant proteins providing important landmarks on the 2-DE map (Table 5 and Figure 16). Gel-free MudPIT proteomic studies allowed the identification, per life stage, of 97 to 513 unique proteins in *P. falciparum* [21, 22] and 582 to 969 ones in *T. cruzi* [41]. Contrastingly, the number of unique proteins identified on 2-DE gels was only 20 to 30 for other apicomplexan parasites [26, 30, 32], 67 in *Leishmania* spp [37] and 45 in *T. brucei* [40]. An intermediate case is therefore offered by the present analysis in *E. cuniculi*. It should be noted that the 177 proteins represent ~9% of the coding sequences of the small microsporidian genome. Even if an

Class function	Spot Number	ORF Number	Protein function
1. Metabolism	1	ECU01_0970	aldose reductase
	2	ECU11_0870	ethanolamine phosphate cytidyltransferase
3. Cell growth, cell division and DNA synthesis	3	ECU09_1330	DNA replication factor C subunit
	4	ECU01_1370	septin-like protein
5. Protein synthesis	5	ECU09_1250	40S ribosomal protein S3
	6	ECU04_1100	translation elongation factor eEF1 $\alpha$
	7	ECU11_0530	tryptophanyl-tRNA synthetase
6. Protein destination	8	ECU03_0520	heat-shock protein, HSP70 family
	9	ECU02_0100	HS70 protein cognate 4 (HSC70)
	10	ECU01_1230	protein of the CDC48/PAS1/SEC28 family of ATPases (AAA)
	11	ECU01_1010	ubiquitin-conjugating E2 subunit
8. Intracellular transport	12	ECU05_0090	GTP binding protein, SAR-type, ARF family
9. Cellular organisation and biogenesis	13	ECU10_1660	spore wall protein 1 (SWP1)
	14	ECU01_0460	actin
	15	ECU06_0250	polar tube protein 1 (PTP1)
	16	ECU06_0240	polar tube protein 2 (PTP2)
10. Cellular communication / signal transduction	17	ECU04_1480	ZRP1
11. Cell rescue, defence, cell death and ageing	18	ECU05_1590 = ECU11_0140	3-methyl-adenine DNA glycosylase
	19	ECU02_0580	hydroxyacyl glutathione hydrolase
	20	ECU01_0930	thioredoxin H-type
	21	ECU01_0820	cysteine-rich protein (endospore protein EnP1, [76])
12. Unassigned role	22	ECU07_0920	exonuclease-phosphatase family domain
	23	ECU08_1610	UPF0023 family (hypothetical proteins)
	24	ECU08_2020	member of a multigenic subtelomeric family / domain: ricin B lectin (QxW lectin repeat)
	25	ECU09_1470	possible RNA binding protein
	26	ECU11_0510	Possible chitin deacetylase [47]
	27	ECU01_0420	
	28	ECU05_0510	
	29	ECU05_0590	
	30	ECU07_0400	
	31	ECU09_1820	

**Table 5.** Classification of highly abundant proteins identified in *E. cuciculi*. Spot numbers refer to Figure 2. For determination of protein abundance, see Material and methods.

A	Extract	Identified proteins	Number of proteins with			B	Number of urea-extracted proteins with	
			0 TM	1-3 TM	4-8 TM		1 isoform	69 (57.1%)
	Urea	139	108 (78%)	30 (22%)	1 (1%)	2 to 3 isoforms	41 (34.5%)	
	NaOH	17	12 (71%)	5 (29%)	0 (0%)	4 to 7 isoforms	10 (8.4%)	
	Laemmli	50	32 (64%)	15 (30%)	3 (6%)			

**Table 6.** General features of the proteins identified in *E. cuciculi*. A- Efficiency of the sequential extraction in terms of hydrophobic proteins. TM: transmembrane helix. B- Distribution of 120 urea-extracted proteins according to their isoforms.

underestimation seems very likely, it is clear that all gene products are not synthesized simultaneously during *E. cuniculi* spore differentiation.

### 3.3. Global analysis of the identified proteins

Each extraction procedure had targets with peculiar physicochemical properties as exemplified by basic histones and ribosomal proteins which were mainly recovered in the NaOH extract (Table S1). The three successive treatments were more and more stringent in solubilizing membrane proteins (Table 6A), the Laemmli extract containing proteins with up to 8 transmembrane segments. Numerous urea-extracted proteins exhibited isoforms (Figure 16, Table 6B and S1) that should be mainly due to post-translational modifications, given that most *E. cuniculi* proteins are encoded by single-copy genes [5]. While experimental molecular masses (MM) were globally centred on theoretical values  $\pm 6$  kDa (Figure 17A), up to 29% of the spots exhibited a pI lower than the calculated pI  $\pm 0.4$  pH unities (Figure 17B). A same protein was frequently identified in several of the latter spots, indicating the presence of pI variants. Different post-translational modifications affecting protein pI were expected to take place in *E. cuniculi* since genome annotation predicted various enzymes involved in acetylation, myristoylation, isoprenylation, phosphorylation and proteolysis [5]. Some of these enzymes are represented in the list of

expressed proteins (Table S1). For example, the five isoforms of the actin protein ECU01\_0460 (Figure 16, spots 14) we unambiguously identified, may be due to post-translational modifications predicted by the *E. cuniculi* genome [5] such as acetylation [52], phosphorylation which is associated with dormancy in slime molds [53, 54] or the glutaredoxin-dependant reversible S-glutathionylation [55] supposed to protect actin from oxidative stress [56].

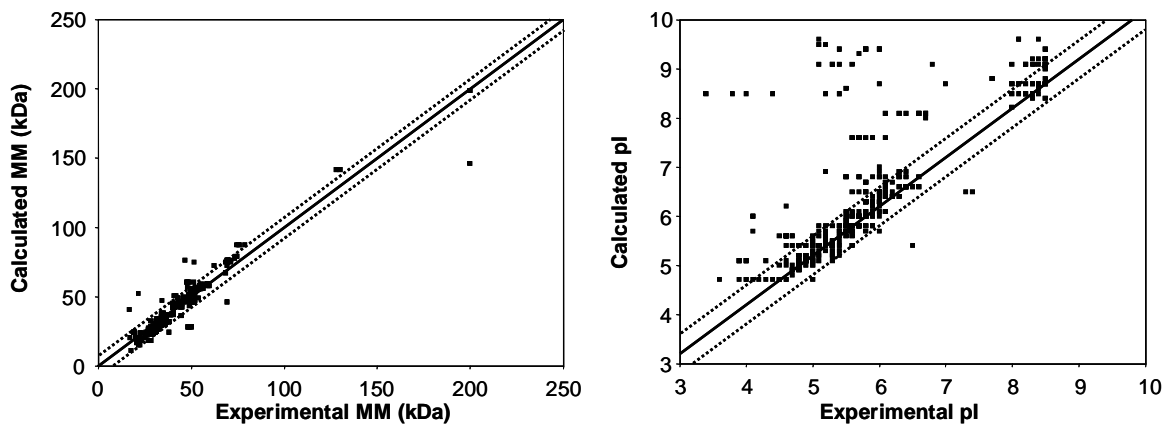
### 3.4. Biological functions of major proteins

The *E. cuniculi* proteins which are mainly eukaryotic housekeeping proteins (Table S1) have been distributed into 12 functional classes (Figure 18A). Two-third of these proteins account for only three classes: protein synthesis (19.9%), protein destination (21.0%) and unassigned function (25.6%). Four classes (Energy, Protein synthesis, Protein destination and Cell defence) are over-represented in the proteome when compared to the genome while two others (Transcription and Transport facilitation) are under-represented (Figure 18B). This may reflect differences in protein solubility and/or relative amount for the studied parasite stages. The following discussion of *E. cuniculi* spore physiology is based upon proteins with a predicted function.

#### 3.4.1. Secretion and specialized structures of the spore.

During *E. cuniculi* sporogony, sporoblast-to-spore morphogenesis involves an intense activity of the endomembrane system, marked





**Figure 17.** Comparison of experimental and calculated molecular mass (MMs, left panel) or isoelectric points (pIs, right panel) for a set of 402 spots. The distribution of experimental MMs is globally centred on the theoretical values (full line), 89% of the spots exhibiting a MM equal to the calculated one  $\pm 6$  kDa (dashed lines). In contrast, the distribution of experimental pIs is systematically shifted down by 0.2 pH unities (full line). Moreover, only 71% of the spots exhibit a corrected pI centered on the calculated pI  $\pm 0.4$  pH unities (dashed lines). This indicates that a given protein should be looked for on a 2-DE gel in the window (calculated MM  $\pm 6$  kDa; calculated pI  $\pm 0.4$ ).

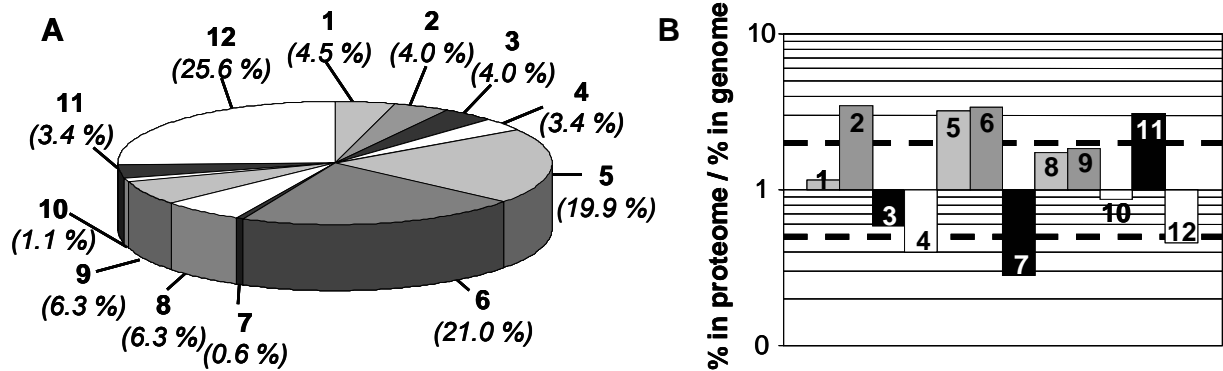
by the enlargement of an unusual Golgi complex at the origin of the polar tube-forming vesicles [4]. Moreover, the cells complete the secretion of cell wall components including chitin and proteins, and the reorganization of both layers of the cell wall (unpublished observation). Many proteins identified in this study are indeed involved in protein synthesis (ribosomal proteins, translation control and tRNA synthetases) and modification (glycosylation, acetylation, prenylation and processing). The functionality of the atypical microsporidian Golgi apparatus which is non-stacked and appears as a set of vesicles embedded in a high electron-density matrix [4], is suggested by the expression of proteins classically required for Golgi vesicular traffic (COPI, COPII, ADP ribosylation factor) and is consistent with the cytochemical detection of a Golgi-specific thiamine pyrophosphatase activity [57].

Proteins from two specialized structures of the *E. cuniculi* spore (polar tube and cell wall) provided major landmarks in the map. The SWP1 protein, produced early during sporogony and representing a major component of the outer spore wall layer or exospore [58] was assigned to spot 13 (Figure 16A). Three proteins have been shown to be located to the polar tube (PTPs) and to be synthesized during the second half of sporogony [59-61]. PTP3 (150 kDa) and the very abundant PTP2 (35 kDa, Figure 16A,

spot 16) were easily detected and exhibited isoforms (2 and 4, respectively. Table S1).

### 3.4.2. Energy and carbohydrate metabolism

Several enzymes involved in carbohydrate metabolism (glycolysis, pentose phosphate and hexosamine pathways, trehalose and chitin metabolisms) have been identified (Table S1). Spores of the cricket parasite *Nosema grylli* were shown to have more or less significant activities of glycolytic enzymes [62] and pyruvate kinase activity was detected at the same level in isolated intracellular stages and in spores [63], suggesting that spores may retain the capacity to degrade glucose. The identification of two key enzymes for trehalose synthesis suggests that trehalose, a non-reducing disaccharide frequently found among fungi, may accumulate during sporogony. In the microsporidian *Nosema algerae*, the level of trehalose decreased during *in vitro* spore germination and trehalose hydrolysis was considered to be at the origin of the rapid increase in intrasporal osmotic pressure required for polar tube extrusion [64]. Trehalose was also thought to be implicated in microsporidian spore protection against desiccation or freezing, or to be an energy source [65]. We identified two of the three enzymes involved in the synthesis of chitin which is secreted during sporoblast-to-spore differentiation to form the inner spore wall layer or endospore (phosphoacetylglucosamine mutase and UDP-



**Figure 18.** Functional distribution of 177 *E. cuniculi* proteins. A) Repartition of the proteins among 12 classes. B) Comparison of proteome- and genome-based class distribution. Over- and under-expressed classes in the proteome are suggested by ratios that are more than 2 and less than 0.5, respectively. The proteins identified by mass spectrometry were classified as those predicted after genome annotation [5]. Classes: 1: Metabolism; 2: Energy; 3: Cell growth, cell division and DNA synthesis; 4: Transcription; 5: Protein synthesis; 6: Protein destination; 7: Transport facilitation; 8: Intracellular transport; 9: Cellular organization and biogenesis; 10: Cellular communication, signal transduction; 11: Cell rescue, defence, cell death and ageing; 12: Unassigned function

GlcNAc pyrophosphorylase). Our failure in the detection of chitin synthase is likely due to a poor extraction of the protein that contains seven predicted transmembrane helices, not to a true lack of expression. This is corroborated by the presence of the putative chitin deacetylase [47] which transforms chitin into chitosan.

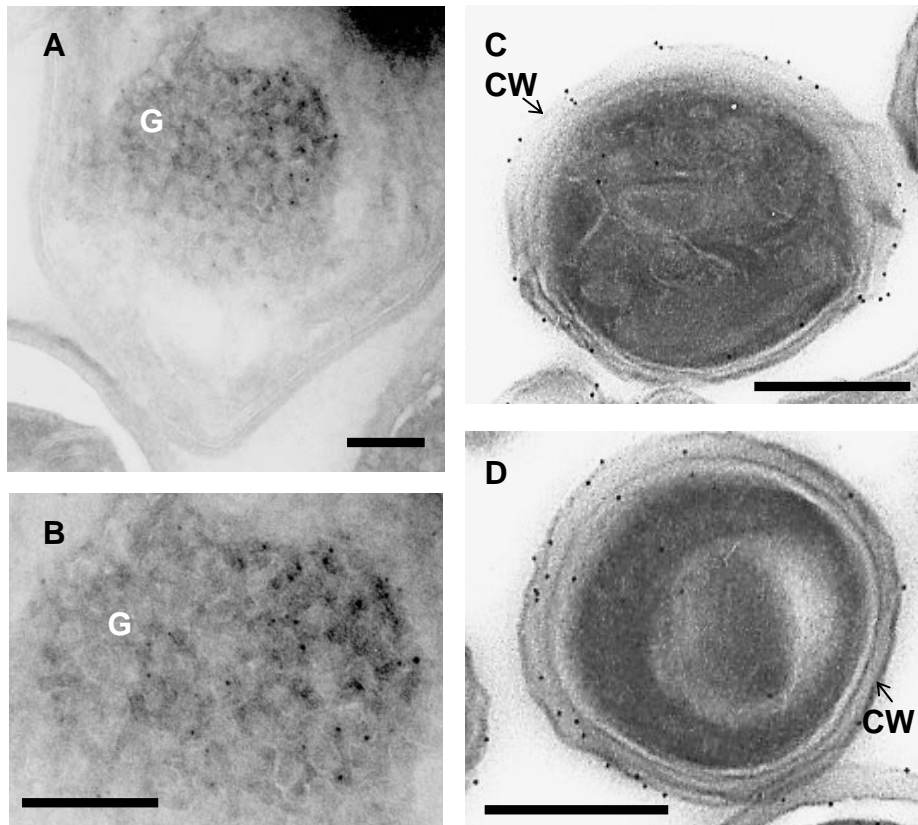
### 3.4.3. Cell cycle control

Five isoforms for ZPR1 protein (Figure 16, spots 17) also represented important landmarks in the studied proteome. ZPR1 accumulates in the nucleolus of proliferating mammalian cells and is involved in preribosomal RNA expression [66]. In *Saccharomyces cerevisiae*, ZPR1 interacts with eEF-1 $\alpha$  and the corresponding complex redistributes from cytoplasm to nucleus when glucose-starved cells respond to extracellular growth stimuli, in particular nutrient stimulation [67]. This interaction is essential for normal cellular proliferation, the yeast cells accumulating in the G2 phase of the cell cycle when altered. By promoting ZPR1 nuclear export, the peptidyl-prolyl isomerase activity of the cyclophilin A is involved in the cellular distribution of the ZPR1-eEF-1 $\alpha$  complex [68]. Both ZPR1 partners, the abundant eEF-1 $\alpha$  (ECU04\_1100) and cyclophilin (ECU08\_0470), were expressed in the spore-rich *E. cuniculi* population here analyzed. The nuclear DNA content of *E. cuniculi* spores is representative of a G2 level [69]. We hypothesize that cyclophilin may promote ZPR1 cytoplasmic accumulation and

may impair the ZPR1-eEF-1 $\alpha$  interaction during spore differentiation, thus leading to a blockage of the cell cycle in the G2 phase. Possibly, nutrient stimulation from the host cell may rescue the ZPR1-eEF-1 $\alpha$  interaction in the intracellular sporoplasm, allowing the translocation of the ZPR1-eEF-1 $\alpha$  complex in the nucleolus for *de novo* rRNA synthesis activation.

### 3.4.4. Parasite survival in the environment

Several abundant proteins may be involved in parasite integrity and its long survival after release in the environment (Tables 5 and S1). The presence of detoxification proteins, such as glutathione peroxidase or thioredoxin peroxidase, suggests glutathione defence mechanisms against oxidative, thermal or osmotic stresses acting in the *E. cuniculi* spore, similar to those in yeasts [70], filamentous fungi [71] or *P. falciparum* [72]. Stress-altered proteins may be helped in their folding by microsporidian major chaperone classes (HSP70, HSP90 and TCP1) or eliminated by *E. cuniculi* 26S proteasome. Interestingly, HSP70 and HSP90 proteins play a critical role in *T. gondii* developmental stage switch [73, 74], HSP90 acting as an extracellular chaperone for molecules implicated in host invasion [75]. Evidence for the expression of proteins involved in DNA protection or repair (RuvB-like protein, Rad50, DNA lyase and 3-methyl-adenine DNA glycosylase) was also obtained.



**Figure 19.** Immunogold electron microscopy of cryosections of *E. cuniculi* cells using mouse polyclonal antibodies raised against recombinant ECU07\_0400 (A, B) and ECU02\_0150 (C, D) proteins. A and B- In a late sporoblast, the ECU07\_0400 protein is associated with the Golgi apparatus appearing as a vesicular network embedded in an electron-dense matrix (magnified in B) Bar: 200 nm. C and D- The ECU02\_0150 protein is localized at the spore cell wall. Bar: 400 nm. CW: cell wall; G: Golgi apparatus.

### 3.5. Using proteins with unassigned functions to look for specific *E. cuniculi* proteins

The proteins of unassigned function correspond to 25.6% of the proteins identified here (Figure 18). This validates as true coding sequences some of the ORFs predicted by gene modelling algorithms (56% of the *E. cuniculi* genes, [5]). Moreover, 11 of the 31 most abundant proteins belong to this group, suggesting a significant contribution to major structural and/or physiological characteristics of the parasitic cell. How to decipher the function of proteins with very few or no sequence homologies, is a general problem in post-genomic research. A complete definition of a protein function implies the knowledge of biological activity, interaction partners, localization features and possible changes in expression during development. Immunolocalization with specific antibodies can guide further research as exemplified by the highly abundant 28-kDa ECU07\_0400 protein (Figure 16, spot 30). Indeed, while sequence analysis gives no clues about its function, an immunocytochemical study has revealed that this protein is associated with endoplasmic reticulum and Golgi vesicles (Figures 19A and 19B).

Unassigned proteins should include a set of various microsporidia-specific proteins such as those required for spore wall formation and parasite-host surface interactions. Very few *E. cuniculi* wall proteins have been characterized so far, including an exospore protein (SWP1)

[58], a putative chitin deacetylase (ECU11\_0510) located to the plasma membrane-endospore interface [47] and the EnP1 and EnP2 proteins that are targeted to the endospore [76]. To look for new wall proteins which must be either secreted or anchored to the plasma membrane, a sequence analysis screening based on the prediction of signal peptide, transmembrane helices, GPI anchor and cellular location was applied to the pool of expressed proteins with unknown function. The candidate ECU02\_0150 protein (220 amino acids residues) was predicted to be secreted then its location to the cell wall was demonstrated by TEM immunolabelling with mouse polyclonal antibodies (Figures 19C and 19D).

### 4. Concluding remarks

Sporoblast-to-spore stages were harvested from culture supernatants to get enough proteins for 2-DE and 2-DE-free analyses. The identification of 177 unique proteins in a map of ~350 spots gave the first global view of proteins expressed in late sporogony of the microsporidian parasite life cycle. Taking into consideration the very low quantities of parasite available, comparative developmental stage proteomics will require very sensitive identification methods. We are currently investigating some protein subgroups such as (i) phosphoproteins which might be implicated in a signal transduction pathway initiating polar tube extrusion, and (ii) glycoproteins for which the diversity of



carbohydrate moieties remain to be elucidated, only O-mannosylated chains and GPI anchors having been predicted following genome annotation [5].

### Acknowledgements

We would like to thank the Institut National de la Santé et de la Recherche Médicale (INSERM), the life science division of the Commissariat à l'Energie Atomique (CEA/DSV), the french Genopole network (RNG) and Joseph Fourier University for their support. D. Brosseau was supported by a grant from "Ministère de l'Education, de la Recherche et de la Technologie". We thank G. Méténier for helpful discussion and critical reading of this manuscript, G. Prensier for his transmission electron microscopy expertise, and F. Roux and M. Court for some mass spectrometry analysis. We also thank B. Chebance and I. Wawrzyniak for *E. cuniculi* cell culture.

### 5. References

[1] Keeling, P.J., Luker, M.A., Palmer, J.D., *Mol. Biol. Evol.* 2000, *17*, 23-31.  
 [2] van Gool, T., Vetter, J.C.M., Weinmayr, B., Van Dam, A., *et al.*, *J. Infect. Dis.* 1997, *175*, 1020-1024.  
 [3] Kotler, D.P., Orenstein J.M., in: Wittner, M. (Ed), *The Microsporidia and Microsporidiosis*, ASM Press, Washington DC 1999, pp. 258-292.  
 [4] Vavra, J., Larson, J., in: Wittner, M. (Ed), *The Microsporidia and Microsporidiosis*, ASM Press, Washington DC 1999, pp. 7-84.  
 [5] Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., *et al.*, *Nature* 2001, *414*, 450-453.

[6] Gardner, M.J., Hall, N., Fung, E., White, O., *et al.*, *Nature* 2002, *419*, 498-511.  
 [7] Hall, N., Karras, M., Raine, J.D., Carlton, J.M., *et al.*, *Science* 2005, *307*, 82-86.  
 [8] Xu, P., Widmer, G., Wang, Y., Ozaki, L.S., *et al.*, *Nature* 2004, *431*, 1107-1112.  
 [9] Abrahamsen, M.S., Templeton, T.J., Enomoto, S., Abrahante, J.E., *et al.*, *Science* 2004, *304*, 441-445.  
 [10] Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., *et al.*, *Science* 2005, *309*, 134-137.  
 [11] Pain, A., Renauld, H., Berriman, M., Merphy, L., *et al.*, *Science* 2005, *309*, 131-133.  
 [12] Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., *et al.*, *Science* 2005, *309*, 436-442.  
 [13] Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., *et al.*, *Science* 2005, *309*, 416-422.  
 [14] El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., *et al.*, *Science* 2005, *309*, 409-415.  
 [15] Loftus, B., Anderson, I., Davies, R., Alsmark, U.C., *et al.*, *Nature* 2005, *433*, 865-868.  
 [16] Barrett, J., Jefferies, J.R., Brophy, P.M., *Parasitol. Today* 2000, *16*, 400-403.  
 [17] Ashton, P.D., Curwen, R.S., Wilson, R.A., *Trends Parasitol.* 2001, *17*, 198-202.  
 [18] Biron, D.G., Moura, H., Marché, L., Hughes, A.L., Thomas, F., *Trends Parasitol.* 2005, *21*, 162-168.  
 [19] Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., *et al.*, *Nat. Biotechnol.* 1999, *17*, 676-682.  
 [20] Wolters, D.A., Washburn, M.P., Yates, J.R. 3rd, *Anal. Chem.* 2001, *73*, 5683-5690.  
 [21] Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., *et al.*, *Nature* 2002, *419*, 520-526.  
 [22] Lasonder, E., Ishihama, Y., Andersen, J.S., Vermunt, A.M.W., *et al.*, *Nature* 2002, *419*, 537-542.  
 [23] Florens, L., Liu, X., Wang, Y., Yang, S., *et al.*, *Mol. Biochem. Parasitol.* 2004, *135*, 1-11.  
 [24] Doolan, D.L., Southwood, S., Freilich, D.A., Sidney, J., *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, *100*, 9952-9957.  
 [25] Makanga, M., Bray, P.G., Horrocks, P., Ward, S.A., *Proteomics* 2005, *5*, 1849-1858.





- [26] Cohen, A.M., Rumpel, K., Coombs, G.H., Wastling, J.M., *Int. J. Parasitol.* 2002, *32*, 39-51.
- [27] Zhou, X.W., Kafsack, B.F.C., Cole, R.N., Beckett, P., *et al.*, *J. Biol. Chem.* 2005, *280*, 34233-34244.
- [28] Geissler, S., Sokolowska-Köhler, W., Bollmann, R., Jungblut, P.R., Presber, W., *FEMS Immunol. Med. Microbiol.* 1999, *25*, 299-311.
- [29] Reichmann, G., Dlugonska, H., Hiszczynska-Sawicka, E., Fischer, H.G., *Microbes Infect.* 2001, *3*, 779-787.
- [30] Lee, E.G., Kim, J.H., Shin, Y.S., Shin, G.W., *et al.*, *Proteomics* 2003, *3*, 2339-2350.
- [31] Lee, E.G., Kim, J.H., Shin, Y.S., Shin, G.W., *et al.*, *J. Chromatogr. B.* 2005, *815*, 305-314.
- [32] de Venevelles, P., Chich, J.F., Faigle, W., Loew, D., *et al.*, *Int. J. Parasitol.* 2004, *34*, 1321-1331.
- [33] Bromley, E., Leeds, N., Clark, J., McGregor, E., *et al.*, *Proteomics* 2003, *3*, 1553-1561.
- [34] Acestor, N., Masina, S., Walker, J., Saravia, N.G., *et al.*, *Proteomics* 2002, *2*, 877-879.
- [35] El Fakhry, Y., Ouellette, M., Papadopoulou, B., *Proteomics* 2002, *2*, 1007-1017.
- [36] Drummel-Smith, J., Brochu, V., Girard, I., Messier, N., Ouellette, M., *Mol. Cell. Proteomics* 2003, *2*, 146-155.
- [37] Bente, M., Harder, S., Wiesgigl, M., Heukeshoven, J., *et al.*, *Proteomics* 2003, *3*, 1811-1829.
- [38] Nugent, P.G., Karsani, S.A., Wait, R., Tempero, J., Smith, D.F., *Mol. Biochem. Parasitol.* 2004, *136*, 51-62.
- [39] van Deursen, F.J., Thornton, D.J., Matthews, K.R., *Mol. Biochem. Parasitol.* 2003, *128*, 107-110.
- [40] Parodi-Talice, A., Duran, R., Arrambide, N., Prieto, V., *et al.*, *Int. J. Parasitol.* 2004, *34*, 881-886.
- [41] Atwood, J.A., Weatherly, D.B., Minning, T.A., Bundy, B., *et al.*, *Science* 2005, *309*, 473-476.
- [42] Texier, C., Brosson, D., El Alaoui, H., Méténier, G., Vivarès, C.P., *Folia Parasitol. (Praha)* 2005, *52*, 15-22.
- [43] Langley, R.C. Jr, Cali, A., Somberg, E.W., *J. Parasitol.* 1987, *73*, 910-918.
- [44] Moura, H., Visvesvara, G.S., *J. Eukaryot. Microbiol.* 2001, *suppl.*, 56S-59S.
- [45] Moura, H., Ospina, M., Woolfitt, A.R., Barr, J.R., Visvesvara, G.S., *J. Eukaryot. Microbiol.* 2003, *50*, 156-163.
- [46] Delbac, F., Duffieux, F., David, D., Méténier, G., Vivarès, C.P., *J. Eukaryot. Microbiol.* 1998, *45*, 224-231.
- [47] Brosson, D., Kuhn, L., Prensier, G., Vivarès, C.P., Texier, C., *FEMS Microbiol. Lett.* 2005, *247*, 81-90.
- [48] Ferro, M., Salvi, D., Brugière, S., Miras, S. *et al.* *Mol. Cell. Proteomics* 2003, *2*, 325-345.
- [49] Sprague, V., Vernick, S.H., *J. Protozool.*, 1971, *18*, 560-569.
- [50] Green, L.C., Didier, P.J., Didier, E.S., *J. Eukaryot. Microbiol.*, 1999, *46*, 434-438.
- [51] Pitarch, A., Sanchez, M., Nombela, C., Gil, C., *Mol. Cell. Proteomics* 2002, *1*, 967-982.
- [52] Schmitz, S., Grainger, M., Howell, S., Calder, L.J., *et al.* *J. Mol. Biol.* 2005, *349*, 113-125.
- [53] Gauthier, M.L., Lydan, M.A., O'Day, D.H., Cotter, D.A., *Cell. Signal.* 1997, *9*, 79-83.
- [54] Furuhashi, K., *J. Eukaryot. Microbiol.* 2002, *49*, 129-133.
- [55] Starke, D.W., Chock, P.B., Mieyal, J.J., *J. Biol. Chem.* 2003, *278*, 14607-14613.
- [56] Dalle-Donne, I., Rossi, R., Milzani, A., Di Simplicio, P., Colombo, R., *Free Radic. Biol. Med.* 2001, *31*, 1624-1632.
- [57] Takvorian, P.M., Cali, A., *J. Eukaryot. Microbiol.* 1994, *41*, 63S-64S.
- [58] Bohne, W., Ferguson, D.J.P., Kohler, K., Gross, U., *Infect. Immun.* 2000, *68*, 2268-2275.
- [59] Delbac, F., Peyret, P., Méténier, G., David, D., *et al.*, *Mol. Microbiol.* 1998, *29*, 825-834.
- [60] Delbac, F., Peuvél, I., Méténier, G., Peyretailade, E., Vivarès, C.P., *Infect. Immun.* 2001, *69*, 1016-1024.
- [61] Peuvél, I., Peyret, P., Méténier, G., Vivarès, C.P., Delbac, F., *Mol. Biochem. Parasitol.* 2002, *122*, 69-80.
- [62] Dolgikh, V.V., Sokolova, J.J., Issi, I.V., *J. Eukaryot. Microbiol.* 1997, *44*, 246-249.
- [63] Dolgikh, V.V., *Protistology* 2000, *1*, 87-91.



- [64] Undeen, A.H., Vander Meer, R.K., J. Eukaryot. Microbiol. 1994, *41*, 129-132.
- [65] Méténier, G., Vivarès, C.P., Microbes Infect. 2001, *3*, 407-415.
- [66] Galcheva-Gargova, Z., Gangwani, L., Konstantinov, K.N., Mikrut, M., *et al.*, Mol. Biol. Cell 1998, *9*, 2963-2971.
- [67] Gangwani, L., Mikrut, M., Galcheva-Gargova, Z., Davis, R.J., J. Cell Biol. 1998, *143*, 1471-1484.
- [68] Ansari, H., Greco, G., Luban, J., Mol. Cell. Biol. 2002, *22*, 6993-7003.
- [69] Biderre, C., Pagès, M., Méténier, G., Canning, E.U., Vivarès, C.P., Mol. Biochem. Parasitol. 1995, *74*, 229-231.
- [70] Carmel-Harel, O., Storz, G., Annu. Rev. Microbiol. 2000, *54*, 439-461.
- [71] Pocsí, I., Prade, R.A., Penninckx, M.J., Adv. Microb. Physiol. 2004, *49*, 1-76.
- [72] Becker, K., Rahlfs, S., Nickel, C., Schirmer, R.H., Biol. Chem. 2003, *384*, 551-566.
- [73] Weiss, L.M., Ma, Y.F., Takvorian, P.M., Tanowitz, H.B., Wittner, M., Infect. Immun. 1998, *66*, 3295-3302.
- [74] Echeverria, P.C., Matrajt, M., Harb, O.S., Zappia, M.P. *et al.*, J. Mol. Biol. 2005, *350*, 723-734.
- [75] Ahn, H.J., Kim, S., Nam, H.W., Biochem. Biophys. Res. Commun. 2003, *311*, 654-659.
- [76] Peuvél-Fanget, I., Polonais, V., Brosson, D., Texier, C. *et al.* Int. J. Parasitol. *In press*



## 5. RECHERCHE DE NOUVELLES PROTEINES DE PAROI

---

La spore microsporidienne présente deux structures caractéristiques majeures : le tube polaire et la paroi. Cette paroi protéochitineuse très épaisse est décrite dans les analyses ultrastructurales comme constituée d'une endospore, épaisse et riche en chitine, ainsi que d'une exospore riche en glycoprotéines et subdivisée en trois feuillets structurés. Elle permet le maintien de l'intégrité de la spore dans l'environnement et participe, par sa rigidité, au processus d'invasion en permettant l'accroissement de la pression interne conduisant à la dévagination du tube polaire. Cette structure est également l'interface des relations hôte-pathogène, vecteur de la reconnaissance de la cellule hôte et des stimuli conduisant à l'invasion. C'est aussi la structure de la spore qui est en contact avec le système immunitaire, elliciteur et cible de la réponse humorale. L'aspect moléculaire de la paroi est très peu connu, et seules deux protéines ont été décrites dans le genre *Encephalitozoon* : SWP1 (Bohne *et al.*, 2000) et SWP2 (Hayman *et al.*, 2001). Nous nous sommes donc intéressés à mettre en évidence de nouveaux constituants de la région pariétale chez *E. cuniculi*.

### **5.1. La chitine déacétylase putative de *E. cuniculi* : une protéine de surface probablement impliquée dans la formation de la paroi microsporidienne**

Nous disposons chez *E. cuniculi* d'un anticorps reconnaissant une protéine de 33 kDa que des observations par microscopie par épifluorescence ont localisée à la périphérie de la cellule. Cet anticorps reconnaît deux régions à 33 kDa (pI de 6 à 7) et 55 kDa (pI de 4 à 7). La cible de cet anticorps a été identifiée par spectrométrie de masse MALDI-TOF au cours de l'analyse protéomique présentée précédemment. Pour les deux régions, une unique protéine de 254 acides aminés a été identifiée. Cette protéine est le produit du gène ECU11\_0510, présent en une seule copie dans le génome de *E. cuniculi* (Katinka *et al.*, 2001). Son analyse bioinformatique suggère une protéine membranaire appartenant à la famille des polysaccharide déacétylases. La chitine, présente en abondance dans l'endospore microsporidienne, est le seul polysaccharide que la reconstruction du métabolisme d'après les données du génome prédit pour être synthétisé. Nous



suggérons donc une activité chitine déacétylase (CDA) pour cette protéine, et nous proposons l'appellation « *EcCDA* ». Cette enzyme convertirait la chitine en chitosan, un polymère présent dans la paroi de nombreux champignons (*Mucor rouxii*, ascospore de *S. cerevisiae* (Briza et al., 1988)), et particulièrement résistant à l'action des chitinases. Un gène homologue est présent chez la microsporidie parasite d'insecte *Antonospora locustae*<sup>1</sup> (précédemment *Nosema locustae*). Les deux séquences sont très proches avec 40% d'identité.

L'analyse de la localisation de cette protéine par MET révèle la présence de cette protéine très tôt dans le cytoplasme des mérontes, puis une mise en place à l'interface de la membrane plasmique et de l'endospore dès le début de la sporogonie. Cette localisation est cohérente avec le rôle suggéré de cette enzyme dans la formation de la paroi. Les dernières étapes de la sporogonie sont caractérisées en particulier par un épaissement de l'endospore riche en chitine, dont le mécanisme moléculaire n'est pas connu. La protéine *EcCDA* est régulièrement distribuée au niveau de la membrane plasmique de toute la cellule, ce qui permet d'envisager une production de chitine tout autour de la cellule. Chez *Mucor rouxii*, les CDA forment en effet un complexe fonctionnel avec les chitine synthases, les chaînes de chitine néoformées étant immédiatement prises en charge par les CDA pour être déacétylées (Kafetzopoulos *et al.*, 1993).

Nous avons également mis en évidence une structure riche en membranes, observée à partir du stade sporoblaste. Cette structure mesure 300 nm sur 200 nm, semble s'ouvrir sur la paroi et réagit avec les anticorps anti-CDA. Les anticorps sont clairement associés aux membranes de cette structure, qui contient un matériel dense aux électrons, similaire à celui de l'endospore. Une structure similaire a déjà été évoquée chez les microsporidies sous le nom de « scindosome » ou « paramural body » (Vavra, 1976a ; Vivarès et Sprague, 1979). Nous émettons deux interprétations pour cette structure. Les dernières étapes de formation de la spore s'accompagnent d'une réduction de taille du parasite, ces structures pourraient donc correspondre à des régions d'endocytose et de récupération de l'excédent de membrane plasmique et de matériel pariétal. Cependant, ces structures pourraient être, chez les microsporidies, l'équivalent des chitosomes fongiques (Bracker *et al.*, 1976), c'est-à-dire des structures vésiculaires cytoplasmiques, où s'effectue la production de chitine et de chitosan, et qui, par exocytose, permettent l'incorporation à la paroi de leur contenu.

La présence d'une chitine déacétylase chez les microsporidies est un élément intéressant qui vient supporter leur proximité phylogénique avec les champignons. En effet, les chitine déacétylases sont présentes chez de nombreux champignons où elles sont responsables de la conversion de la chitine en chitosan, un polymère résistant aux chitinases. La présence de ce polymère dans la paroi microsporidienne participerait à sa remarquable résistance, particulièrement vis-à-vis des chitinases.

<sup>1</sup> *Nosema locustae* Genome Projet, Marine Biological Laboratory, <http://jbcpl.nbl.edu/Nosema/index.html>





Les CDA fongiques (Tsigos *et al.*, 2000) sont des glycoprotéines dont la taille varie de 12 à 150 kDa et dont le contenu glycosidique représente entre 30% et 67% de la masse chez *Colletotrichum lindemuthianum* (Tsigos et Bouriotis, 1995). La chitine déacétylase d'*E. cuniculi* présente deux isoformes de masse différente (33 kDa et 55kDa) ce qui laissent supposer également une forte glycosylation.

Il est intéressant également de faire un parallèle entre la présence de chitosan au sein de l'endospore et l'ultrastructure de l'endospore microsporidienne. Aucune couche de chitosan bien individualisée n'a été décrite au sein de l'endospore comme c'est le cas dans la paroi de l'ascospore de *S. cerevisiae* (Briza *et al.*, 1988). Il pourrait non pas être uniformément repartit dans l'ensemble de l'endospore microsporidienne mais être plus abondant dans des zones de tension, permettant à la paroi d'acquérir des caractéristiques particulières de résistance ou de structure, comme par exemple dans la région du capuchon polaire. L'analyse de l'ultrastructure de l'endospore a révélée la présence de structures filamenteuses qui semblent relier la membrane plasmique à l'exospore. Ces structures semblent être responsables du maintien intime de la membrane plasmique avec la paroi après externalisation du sporoplasme. Nous pouvons émettre l'hypothèse de la participation, au moins en partie du chitosan à ces structures.

De récents travaux portant sur les CDA de champignons ont suggéré que cette enzyme était impliquée dans les interactions hôte-pathogène. D'une part, une mannoprotéine sécrétée par le champignon *Cryptococcus neoformans* et identifié comme un homologue des chitine déacétylases, a été décrite comme fortement immunogène vis-à-vis de cellules T (Levitz *et al.*, 2001). D'autre part, chez *Colletotrichum lindemuthianum*, un champignon pathogène de plante, l'invasion est facilitée par l'activité d'une chitine déacétylase sécrétée sur les oligomères de chitine provenant de la dégradation de la paroi fongique par les chitinases de la plante. En fait, ces oligomères de chitine ont un effet éliciteur sur les mécanismes de défense de la plante (production de callose ou lignification) alors que ces mêmes oligomères après déacétylation ne présentent plus cet effet (Tsigos et Bouriotis, 1995). De façon intéressante, des chitinases circulantes ont été mise en évidence dans le sérum de plusieurs mammifères dont l'homme (Escott et Adams, 1995). Chez le cochon d'Inde, lors d'une infection intraveineuse par le champignon *Aspergillus fumigatus*, l'activité des chitinases est augmentée dans le sérum et les tissus (Overdijk *et al.*, 1999). Cela pourrait supporter le rôle possible des chitinases comme mécanismes de défense contre les pathogènes riches en chitine. Dans un tel contexte, la capacité de certains microorganismes à synthétiser du chitosan, comme les microsporidies, pourrait être vu comme un mécanisme de résistance à la dégradation de la chitine par l'hôte.



---

**The putative chitin deacetylase of *Encephalitozoon cuniculi*: a surface protein implicated in the microsporidian spore wall formation.**

---

Damien BROSSON<sup>a</sup>, Lauriane KUHN<sup>b</sup>, Gérard PRENSIER<sup>a</sup>, Christian P. VIVARES<sup>a</sup>, Catherine TEXIER<sup>a</sup>

<sup>a</sup> Equipe Parasitologie Moléculaire et Cellulaire, LBP, UMR CNRS 6023, Université Blaise Pascal, 63177 Aubière, France

<sup>b</sup> Laboratoire de Chimie des Protéines, DBMS, CEA, 38054 Grenoble, France.

*Publié dans FEMS Microbiology Letters en Juin 2005, volume 247, pages 81 à 90.*

### Abstract

Microsporidia are fungal-like unicellular eukaryotes which develop as obligate intracellular parasites. They differentiate into resistant spores that are protected by a thick cell wall composed of glycoproteins and chitin. Despite an extensive description of the fibrillar structure of this wall, very little is known about its protein components and deposit mechanisms. In this study on the human pathogen *Encephalitozoon cuniculi*, we identify by mass spectrometry the target of polyclonal antibodies previously raised against a 33-kDa protein located at the outer face of the parasite plasma membrane. This 254-amino acid protein is encoded by the ECU11\_0510 open reading frame and presents two isoforms of 33 and 55 kDa. Sequence analysis supports an assignment to the polysaccharide deacetylase family with a suspected chitin deacetylase activity (*EcCDA*). As demonstrated by TEM studies, *EcCDA* is present at the plasma membrane of the early stages of *E. cuniculi* life-cycle. At

the sporoblast stage, the enzyme accumulates especially in paramural bodies which are convolutions of the plasma membrane opened to the wall. The identification of an *EcCDA* homologue in the insect parasite *Antonospora locustae* (ex *Nosema locustae*) suggests a widespread distribution of this enzyme among Microsporidia. This characterization of a new microsporidian surface protein creates new perspectives to understand spore wall formation and spore resistance.

### 1. Introduction

The obligate intracellular Microsporidia are spore-forming amitochondriate unicellular eukaryotes that can parasitize a wide variety of animals ranging from protists to mammals. Some species, including *Encephalitozoon cuniculi*, are opportunist human pathogens [1] with a high seroprevalence in immunocompetent populations [2-4]. Phylogenetic analyses based on some conserved proteins suggest that Microsporidia are related to fungi [5,6]. Complete



sequencing of the *E. cuniculi* genome [7] has also provided data consistent with a fungal origin of Microsporidia [8].

Microsporidian spores display an original invasion mechanism, involving a coiled structure (the polar tube) that can be quickly extruded to transfer the spore content or sporoplasm into a target cell. The parasite enters a proliferation phase or merogony and the resulting meronts then undergo a differentiation step (sporogony). They transform into sporonts (beginning of cell wall formation), then into sporoblasts (beginning of polar tube elaboration) and finally into mature spores protected by a thick wall. The microsporidian spore wall plays a major role in the protection against environmental stresses, allowing long-term survival of the parasite after its release from the host cell [9,10]. While the classical concept of the parasite penetration mode involves the polar tube acting as a needle-syringe system, microsporidian spores may also be internalized via attachment to the host plasma membrane then endocytosis [11]. If this occurs, the induced endocytosis process should depend on interactions with the wall.

The microsporidian spore wall usually comprises a proteinaceous outer electron-dense layer (exospore) and a chitinous inner electron-transparent layer (endospore) [12-14], which is of fibrillar nature and connected to the plasma membrane [13,15]. The stippled deposit of an electron-dense material on the outer side of the parasite plasma membrane is

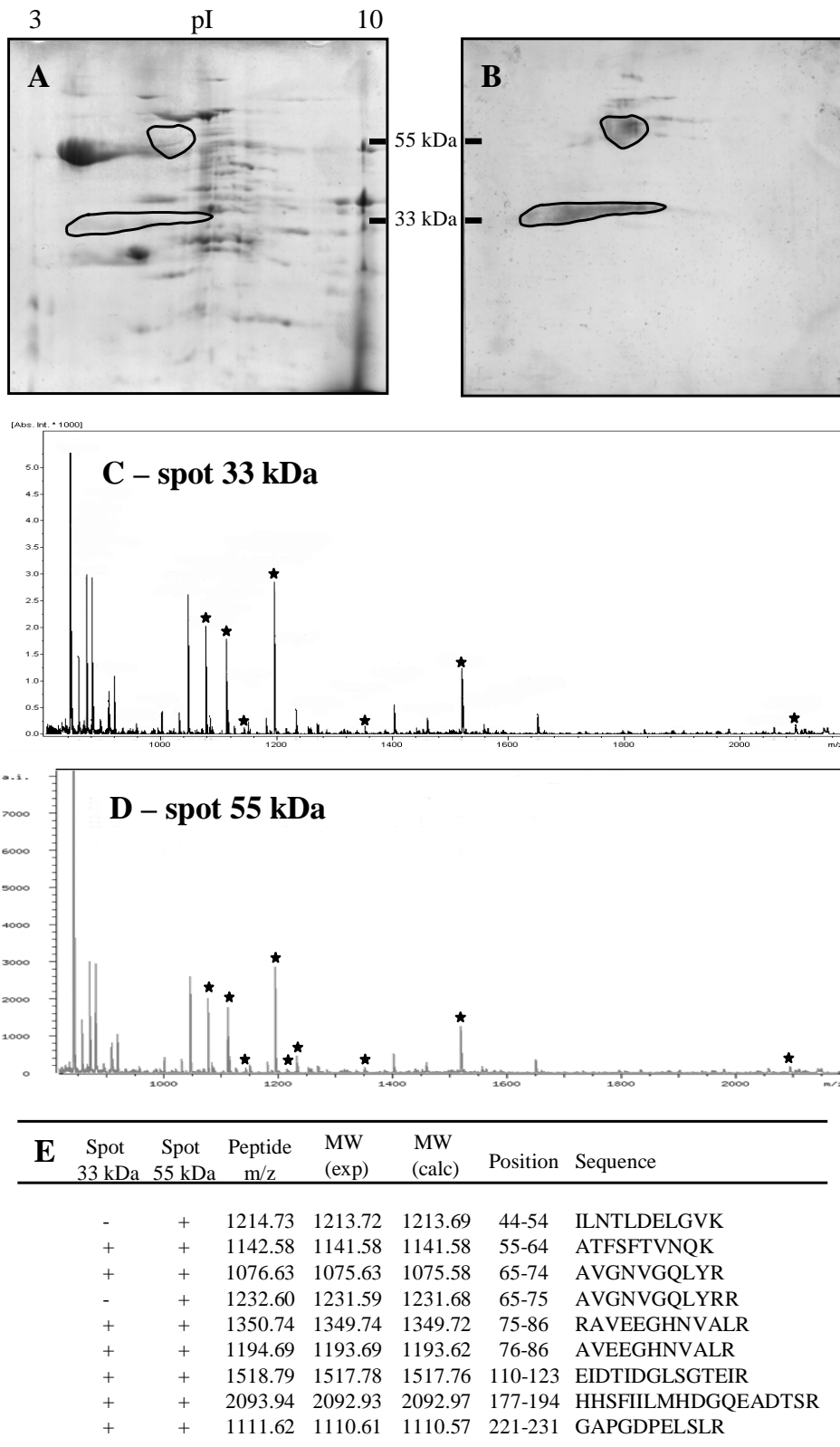
the main visible feature of the transition from merogony to sporogony [16,17]. This deposit turns into a continuous layer in young sporoblasts, and should be at the origin of the exospore. The electron-transparent endospore is built in later stages and thus appears as a cytological marker of the mature spore.

Few studies have been dedicated to the characterisation of spore wall proteins. Immunostaining with polyclonal antibodies raised against partially purified *E. cuniculi* proteins has revealed a 30-kDa antigen within the exospore and a 33-kDa one close to the spore plasma membrane [18]. Several monoclonal antibodies have also been reported to be specific to spore wall antigens [19-21]. To date, the knowledge of the primary structure of such antigens is limited to two exospore components with unknown function: SWP1 (50 kDa) in both *E. cuniculi* [22] and *E. intestinalis* [23], and SWP2 (150 kDa) in *E. intestinalis* only [23]. Here, we provide evidence for an *E. cuniculi* surface protein belonging to the polysaccharide deacetylase family that should be active in chitin deacetylation during all sporogonic stages.

## 2. Materials and Methods

### 2.1. Microsporidian spore production and protein extraction

The *E. cuniculi* spores were produced on Madin-Darby Canine kidney (MDCK) or Human Foreskin Fibroblast (HFF) cells as previously described [18].



**Figure 20.** Migration on 2-D gel and identification by mass spectrometry of the anti-33kDa PAb target. E. cuniculi spore proteins were separated by 2-D electrophoresis and either stained by Coomassie Blue (A) or subjected to Western blot immunodetection with a 1:4000 dilution of the anti-33kDa PAb (B). Trypsin digests of the 33-kDa spot (C) and the 55-kDa spot (D) was analyzed by MALDI-TOF-MS. Seven (nine) experimental peptides for the 33-kDa spot (55-kDa one), labelled with stars in C (in D), matched the ECU11\_0150 protein with a score of 60 (100) and sequence coverage of 29% (33.8%) (E). Only one miscleavage was observed. Differences between experimental and calculated molecular weights (MW) do not exceed 0.09 Da.

Parasites ( $10^9$ - $10^{10}$ ) were disrupted in lysis buffer containing 100 mM DTT, 4% CHAPS and 0.2% SDS, by repeated cycle of freezing-thawing and sonication (Deltasonic 1320, 300W, 28 kHz). Disrupted spores were then incubated in the extraction solution (7 M urea, 2 M thiourea, 100 mM DTT, 4% CHAPS, 0.2% SDS) for 6h at room temperature. Soluble proteins were collected and stored at  $-20^{\circ}\text{C}$ .

## 2.2. Recombinant ECU11\_0510 protein expression

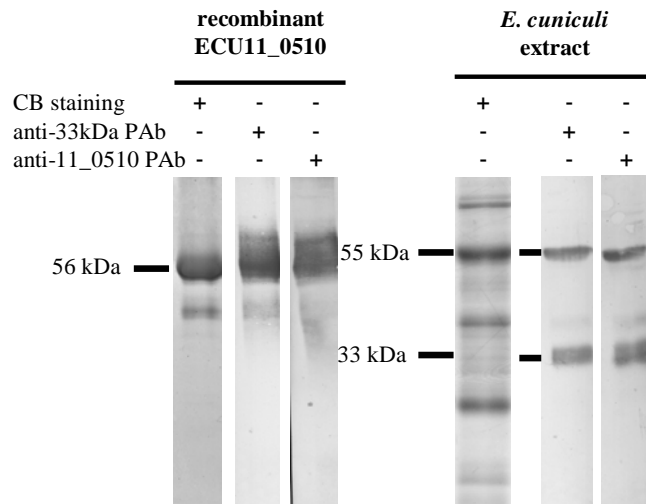
The ECU11\_0510 protein was expressed as a truncated form in fusion with glutathione-S-transferase (GST) at the N-terminal and an 8xHis-tag at the C-terminal extremity. PCR primers were designed (5'-CGGAATTCTTTGTGGACGGTCCTGTC-3' and 5'-CGCTCGAGTCCTATGCTCTCAACGCC-3' containing an *EcoRI* and a *XhoI* restriction site, respectively) to amplify a 624-bp genomic DNA fragment corresponding to the amino acid region 32-239 of the 254-amino acid ECU11\_0510 protein. PCR reactions were performed according to standard conditions with an annealing temperature of  $55^{\circ}\text{C}$ , and the purified product was cloned in frame with GST into the prokaryotic expression vector pGEX-4T1 modified to provide a C-terminal 8xHis-tag after the *XhoI* site. The resulting recombinant plasmid was introduced in the *E. coli* BL21+ strain. After induction with IPTG, bacterial proteins were

extracted with Laemmli solution and analysed by SDS-PAGE.

## 2.3. Gel electrophoresis and Western blot analysis

Protein samples were analysed by SDS-PAGE on 12% polyacrylamide gels. For two-dimensional (2-D) gel electrophoresis, protein isoelectrofocalisation was performed along linear immobilized pH gradient strips of 7 and 17 cm (BioRad) using a rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS) supplemented with 2 mM tributyl phosphine (TBP) and 0.5% ampholytes (BioRad) with the IPGPhor apparatus (Amersham). Before standard SDS-PAGE on 12% polyacrylamide gel, strips were equilibrated with 50 mM Tris HCl pH8.8, 6 M urea, 30% glycerol, 2% SDS containing 100 mM DTT, and then 135 mM iodoacetamide. For mass spectrometry analysis, gels were fixed in a 7.5% acetic acid / 30% ethanol solution, stained with Coomassie Blue (BioRad) and destained with 30% ethanol. For immunoblotting studies, proteins were transferred onto polyvinylidene difluoride (PVDF) membrane (Millipore). After blocking for 1h in PBS-5% skim milk, membranes were incubated for 3h with appropriate dilutions of polyclonal mouse antibodies in PBS 0.1% Triton X-100, treated with a 1:10000 dilution of alkaline phosphatase – conjugated goat anti mouse IgG (H+L) (Promega) for 1h at room temperature and revealed with NBT/BCIP (Promega).





**Figure 21.** Immunoblotting with anti-33kDa PAb and anti-11\_0510 Pab against both the recombinant ECU11\_0510 protein and *E. cuniculi* extract. The 56-kDa recombinant form of the ECU11\_0510 protein (fused in N-terminal with GST and in C-terminal with 8xHis tag) was injected into BALB/c mice to produce polyclonal antibodies (anti-11\_0510 PAb). Both antibodies reacted in a similar way. (CB: Coomassie blue).

#### 2.4. Tryptic digestion for mass spectrometry analysis

Protein spots were manually excised from Coomassie blue-stained 2-D gels and transferred in 96-well-microtitration plates. Sample preparation was fully automated (Mass Prep Station Multiprobe II, Micromass). Briefly, excised gels were washed with destaining solutions (25 mM  $\text{NH}_4\text{HCO}_3$  for 15 min and 50% (v/v) acetonitrile containing 25 mM  $\text{NH}_4\text{HCO}_3$  for 15 min). After dehydration with 100% acetonitrile and drying, proteins were reduced in 25 mM  $\text{NH}_4\text{HCO}_3$ , 10 mM DTT for 1h at 37°C, and alkylated in 25 mM  $\text{NH}_4\text{HCO}_3$ , 55 mM iodoacetamide for 30 min at 37°C. After washing with destaining solutions and pure water, gels were shrunk again with 100% acetonitrile. Depending on protein amount, 2 to 3  $\mu\text{L}$  of 0.1  $\mu\text{g}/\mu\text{L}$  modified trypsin (Promega, sequencing grade) in 25 mM  $\text{NH}_4\text{HCO}_3$  were added over the gel spots for 30 min. Seven  $\mu\text{L}$  of 25 mM  $\text{NH}_4\text{HCO}_3$  were then added to cover the gel spots and digestion allowed to continue for 5h at 37°C.

#### 2.5. MALDI-TOF-MS analysis and database search

For MALDI-TOF-MS analysis, 0.5  $\mu\text{L}$  of trypsin peptide mixture was mixed with 0.5  $\mu\text{L}$  of  $\alpha$ -cyano-4-hydroxycinnamic acid at half saturation in 60% acetonitrile / 0.1% (v/v) trifluoroacetic acid (TFA). The resulting solution was automatically spotted on the target plate and rinsed by 2  $\mu\text{L}$  of 0.1% TFA for 30s, liquid being then blown off by

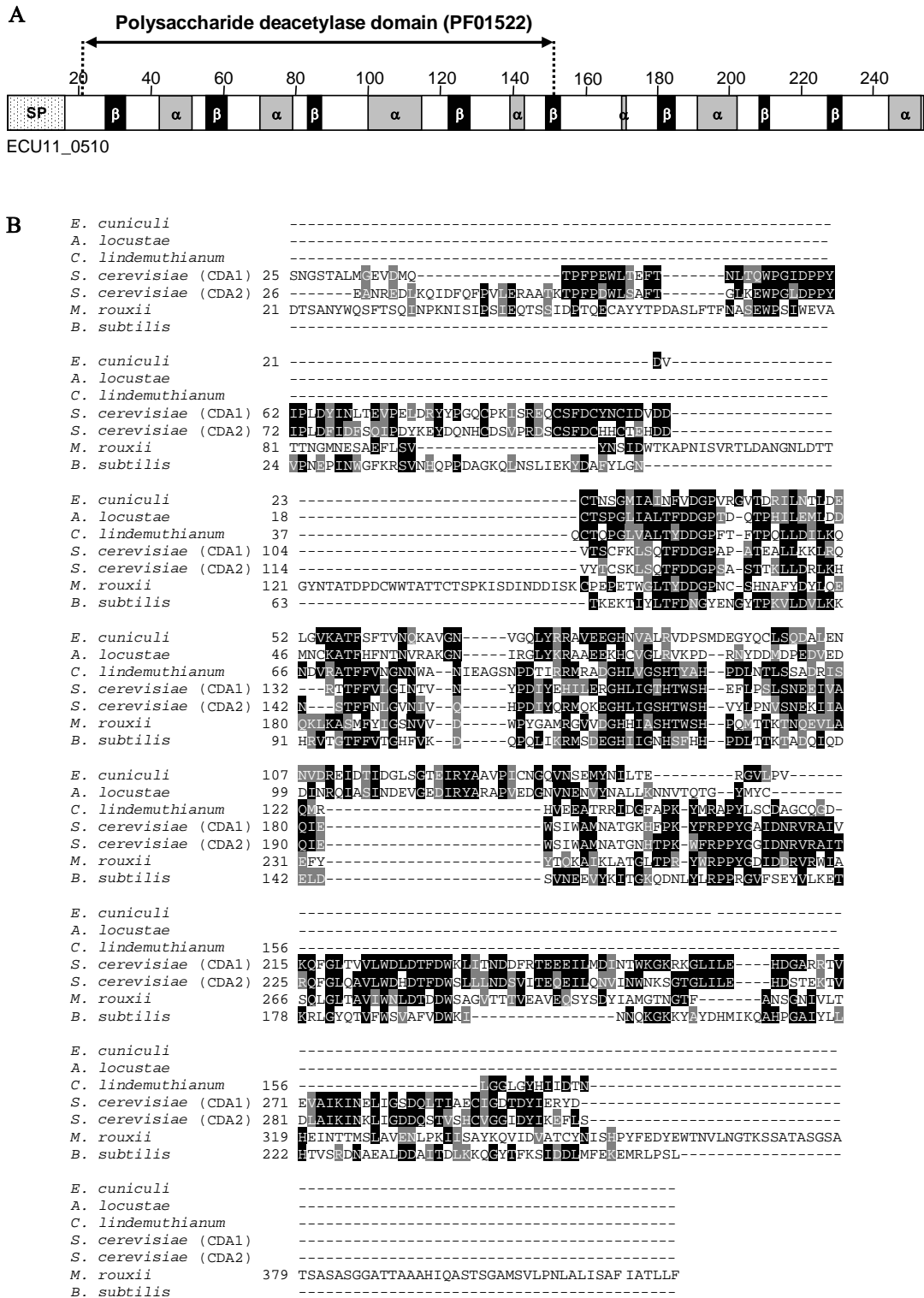
pressurized air. Peptide mixtures were analysed with a MALDI-TOF mass spectrometer (Autoflex, Bruker Daltonik) in reflector mode over a mass range of 0-4200 Da. For each sample, spectrum acquisition and annotation were done automatically. Spectra from 200 shots at several positions were combined to generate a peptide mass fingerprint for database searches. Spectra were calibrated using a mixture of angiotensin II (1046.54 Da), substance P (1347.74 Da), bombesin (1619.82 Da) and ACTH (18-39) (2465.20 Da) as an external standard. The *E. cuniculi* specific database was explored with an intranet version of Mascot 1.7. All peptide masses were assumed to be monoisotopic and  $[\text{M}+\text{H}]^+$  (protonated molecular ions). Peptide modifications allowed during the search were cysteine carbamidomethylation, protein N-acetylation and methionine oxidation. The maximum number of missed cleavages was set to 1 and the mass tolerance to  $\pm 150$  ppm.

#### 2.6. Polyclonal antibody production

Antiserum to *E. coli* expressed recombinant protein was produced in BALB/c mice from crushed SDS-PAGE protein bands. Every two weeks, mice were injected intraperitoneally with samples homogenized with Freund's adjuvant (Sigma). Serum was collected 2 weeks after the last injection and stored at -20°C.

#### 2.7. Transmission electron microscopy immunolabelling

*E. cuniculi*-infected MDCK cells were pelleted, then fixed with 4%



**Figure 22.** Sequence analysis of the ECU11\_0510 protein (A) and alignment with chitin deacetylase domains from other organisms (B). A) The full-length protein (254 amino acids) is represented with different predicted elements: signal peptide (SP), secondary structure consensus with beta strand in black and alpha helix in grey. The ECU11\_0510 polysaccharide deacetylase domain encompasses the position 21-150 and is included in the pfamA01522 domain (score= 11.1, Evalue= 3.8e-05). The predicted  $\alpha / \beta$  class structure fits with the crystal three-dimensional model of *Bacillus subtilis* polysaccharide deacetylase (PDB: 1NY1). B) Sequence alignment of chitin deacetylase domains from the mammalian *Microsporidia* *Encephalitozoon cuniculi* (ECU11\_0510), the insect *Microsporidia* *Antonospora* (formerly *Nosema*) *locustae* (sequence from contig\_159\_1742:2506, <http://jbpc.mbl.edu/Nosema/index.html>, pfam: score=56.4, Evalue=8.3e-14), the fungi *Colletotrichum lindemuthianum*, *Saccharomyces cerevisiae*, *Mucor rouxii* and the bacterium *Bacillus subtilis*. Identical and similar residues are highlighted in black and grey respectively. The microsporidian domain is markedly reduced but presents a strong homology with the other proteins on its 60-70 first amino acids.

paraformaldehyde and 0.1% glutaraldehyde in 0.1 M phosphate buffer pH 7.2. After washing, the pellet was impregnated with a cryoprotective agent (25% glycerol, 5% DMSO in phosphate buffer) and frozen in pasty nitrogen before storage in liquid nitrogen. The 85 nm ultrathin sections were performed with the FCS system on the Leica UltraCut S ultramicrotome. The sections were picked on collodion coated nickel grids and stored on 2.3 M sucrose before immunolabelling. After blocking for 30 min with 2% BSA in PBS, sections were incubated for 1h30 with various dilutions of polyclonal antibodies, and then for 1h with a 1:100 dilution of protein A or goat anti-mouse antibody conjugated with 5 nm colloidal gold particles (Sigma). After washing, the grids were stained and embedded with 0.5% aqueous uranyl acetate 1.6% methylcellulose for 10 min prior to examination with a JEOL 1200EX transmission electron microscope.

### 2.8. Sequence analysis

Physico-chemical parameters of the protein sequence were calculated with ProtParam ([www.expasy.org](http://www.expasy.org)). The SignalP program version 2 ([www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), the TMPRED program ([www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)) and the PSORT II algorithm ([www.psort.nibb.ac.jp/form2.html](http://www.psort.nibb.ac.jp/form2.html)) were used to predict signal peptide, transmembrane helices and predict protein localisation, respectively. Consensus secondary structure prediction were performed on NPS@

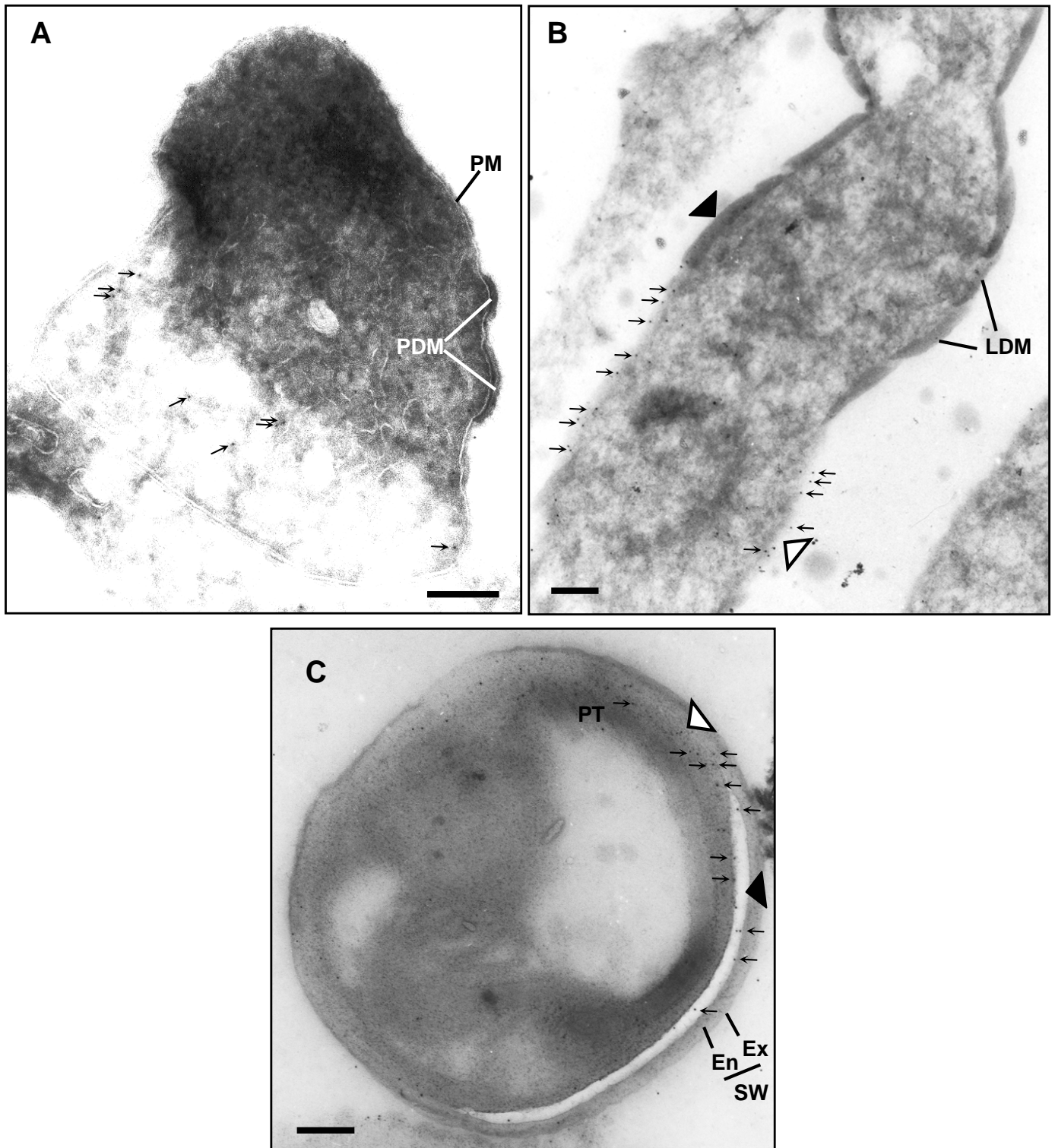
Network Protein Sequence Analysis (<http://npsa-pbil.ibcp.fr>) [24] using the DSC, GOR4, PHD and SOPMA algorithms. Search for protein homologies and protein domains were done using the BLAST program ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and the Pfam database ([www.sanger.ac.uk/Software/Pfam/search.shtml](http://www.sanger.ac.uk/Software/Pfam/search.shtml)), respectively. Multiple sequence alignments were performed with Dialign program ([bibiserv.techfak.uni-bielefeld.de/dialign/](http://bibiserv.techfak.uni-bielefeld.de/dialign/)) and analyzed for amino acid identity and similarity with Boxshade ([www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

## 3. Results and Discussion

Mouse polyclonal antibodies (PAb) were previously raised against several SDS-PAGE separated *E. cuniculi* spore proteins and analysed by Western blotting and indirect immunofluorescence [25]. One of these, directed against a 33-kDa protein band (anti-33kDa PAb), showed a labelling close to the outer face of the parasite plasma membrane. This prompted us to identify the protein target of this anti-33kDa PAb.

### 3.1. The anti-33kDa PAb recognizes the ECU11\_0510 protein

To get a better separation of the targeted protein, we performed 2-D electrophoresis of proteins extracted from *E. cuniculi* spores (Fig. 20A). After Western blotting, the anti-33kDa PAb highlighted two major regions at 33 kDa and 55 kDa (Fig. 20B). MALDI-TOF



**Figure 23.** Immunogold electron microscopy of cryosections of different *E. cuniculi* developmental stages using the anti-11\_0510 PAb and secondary antibody conjugated with 5 nm colloidal gold (bars = 200 nm). The labelling (small arrows) is observed as early as the meront-to-sporont transition in the cytoplasm (A). In dividing sporonts (B), gold particles are found at the periphery of the cell. Finally, the antibodies react with the innermost part of the endospore of the mature spore (C). Black arrowheads highlight transversal sections of the spore wall, while the white ones show grazing section planes. PM: plasma membrane. PDM: patches of dense material. LDM: lamellar dense material. SW: spore wall. Ex: exospore. En: endospore. PT: polar tube.

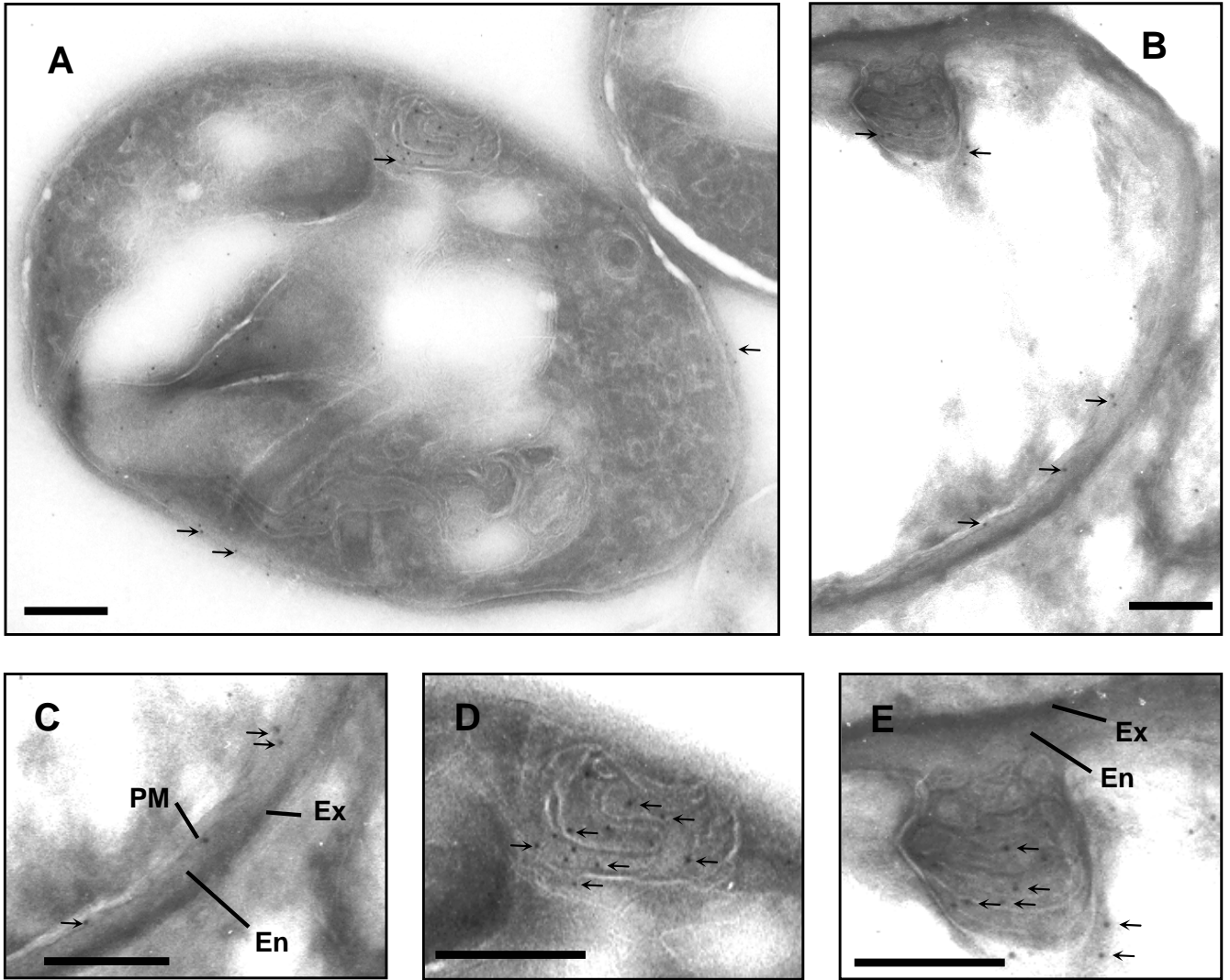
mass spectrometry data obtained with the 33-kDa region (Fig. 20C) matched the ECU11\_0510 protein (SWISS-PROT: Q8SU65) in the *E. cuniculi* proteome database deduced from genome sequencing [7] (Fig. 20E). Identical data were obtained with the 55-kDa spot (Fig. 20D and 20E). We produced a recombinant form of the *EcCDA* protein which was slightly truncated to exclude predicted transmembrane helices (see below). The purified recombinant protein reacted with the anti-33kDa PAb (Fig. 21) and was also used to produce specific PAb in BALB/c mice (anti-11\_0510 PAb). These new antibodies exhibited a similar reaction pattern as the anti-33kDa PAb either on the recombinant ECU11\_0510 protein or on *E. cuniculi* extract (Fig. 21).

### 3.2. The ECU11\_0510 protein is predicted to be a chitin deacetylase (*EcCDA*)

As supported by genome data [7], the ECU11\_0510 open reading frame is a 765-bp single-copy gene which is located on chromosome 11 and which codes for a 254-amino acid protein (28.1 kDa, pI 4.56). As summarized in Fig. 22A, the extreme N-terminal region displays the characteristics of a signal peptide with a predicted cleavage site between positions 15 (Gly) and 16 (Ser) corroborated by the TMPred prediction of a transmembrane helix (TMH) in the 1-18 region. An additional TMH should be located at the 234-253 C-terminal extremity in an outside-inside orientation, suggesting that the protein attached to the membrane via this

TMH has a large extracellular domain. This is in accordance with the efficient recovery of ECU11\_0510 (not shown) when using an improved membrane protein solubility mixture characterized by the amidosulfobetaine ASB14 detergent and high chaotrope molarity [26,27]. PSORT program also predicts an extracellular location including cell wall (66.7%, k=9/23).

The ECU11\_0510 protein contains the PfamA conserved domain 15022 representative of the polysaccharide deacetylase family (Fig. 22B). This family includes *Rhizobium* NodB (nodulation protein B), endoxylanases and fungal chitooligosaccharide deacetylases. The presence of chitin in the microsporidian spore wall has been unambiguously demonstrated [12,13]. Moreover, chitin is the sole polysaccharide that metabolism reconstruction predicts to be synthesized in *E. cuniculi* [7]. *N*-acetylglucosamine 6-P is converted into *N*-acetylglucosamine 1-P, then into UDP-*N*-acetylglucosamine via the activities of phosphoacetylglucosamine mutase and UDP-GlcNAc pyrophosphorylase. Chitin synthase carries out the polymerisation of *N*-acetylglucosamine into chitin chains. We thus assume that the ECU11\_0510 protein may be a chitin deacetylase enzyme (*EcCDA*) which deacetylates chitin into chitosan. Enzymatic assays with a purified recombinant protein that was expressed in *E. coli*, were unsuccessful. The recombinant protein may lack some post-translational modifications required for the biological activity as it is the



**Figure 24.** Immunogold electron microscopy of cryosections of *E. cuniculi* sporoblasts using the anti-11\_0510 PAb and secondary reagent labelled with 5 nm colloidal gold (bars = 200 nm). In sporoblasts (A-B), the labelling (small arrows) is associated with the constructing wall (C) and concentrates in peripheral structures (D-E). PM: plasma membrane. Ex: exospore. En: endospore

case for the yeast CDA2p which loses its activity after deglycosylation [28].

Interestingly, strong homologues of the chitosan pathway enzymes, including CDA, were identified in the partially sequenced genome of an insect Microsporidia, *Nosema locustae* renamed *Antonosporea locustae* [29] ("*Nosema locustae* Genome Project, Marine Biological Laboratory at Woods Hole, funded by NSF award number 0135272", <http://jbcpl.mbl.edu/Nosema/index.html>). The two microsporidian CDA sequences are of nearly identical size and exhibit 40% identity (Fig. 22B).

The presence of CDA in Microsporidia provides an additional support to the fungal origin of these intracellular parasites. Chitin deacetylases have been characterized in several fungi [30], including zygomycetes (*Mucor rouxii* [31,32]) and ascomycetes (*Colletotrichum lindemuthianum* [33], *Saccharomyces cerevisiae* [34,35]). These enzymes are monomeric glycoproteins ranging from 12 kDa to 150 kDa in size, with a carbohydrate content ranging from 30% to 67% by weight in *Colletotrichum lindemuthianum*, [33]. A well-defined role in cell-wall chitosan biosynthesis has been demonstrated for CDA secreted in the periplasm. In *M. rouxii*, chitin synthase first polymerizes chitin chains, then the *N*-acetamido bonds are hydrolysed by CDA [30]. In *S. cerevisiae*, two CDAs are required for the biosynthesis of a chitosan-containing ascospore wall [36], resistant to

stress conditions such as glucanase or heat treatment [35,37]. In the microsporidian spore wall, chitosan could provide interactions with the other wall components and could thus account for the extreme rigidity and resistance of the parasite wall to chitinase treatments and to different stresses such as freezing, heating and extreme pH [9,10]. In some pathogen fungi, CDA is extracellular and is thought to be implicated in the bypass of host-defence mechanisms during *C. lindemuthianum* plant invasion [33] or *Aspergillus fumigatus* mammalian infestation [38].

### 3.3. The *EcCDA* protein presents two isoforms

In the *E. cuniculi* extract, both PAb recognized the two ECU11\_0510 isoforms at 33 and 55 kDa (Fig. 21). Several fungi also exhibit multiple CDA glycoproteins: two, three, four and up to five isoforms were described in *Mucor rouxii* [32], *Metarizium anisophilae* [39], *Rhizopus nigricans* [40] and *Uromyces viciae* [41], respectively. The *E. cuniculi* genome sequence predicts only few genes involved in protein glycosylation, all of them being related to O-mannosylation [7]. Sixteen serine and eleven threonine residues are potential O-glycosylation sites in *EcCDA*. While O-mannoproteins were recently detected in the spore wall of *E. cuniculi* and in the SDS-PAGE 55-kDa region (Méténier, G., Mazet, M. and Taupin, V., personal communication), concavalin A labelling did not co-localize with the ECU11\_0510 protein on 2-D gels (not shown).





### 3.4. *EcCDA* is located at the plasma membrane - endospore interface of the parasite

Immunocytochemical experiments at the ultrastructural level with the anti-11\_0510 PAb, revealed a slight gold labelling of the cytoplasm of stages representative of the meront-to-sporont transition (Fig. 23A). Such cells are indeed still tightly attached to the parasitophorous vacuole membrane by one side, and their engagement in early sporogony is marked by patches of dense material abutting on the plasma membrane. As sporogony progresses, the labelling tends to accumulate at the periphery of the parasitic cells. In longitudinal sections of dividing sporonts (Fig. 23B), gold particles are scattered just beneath the lamellar dense material deposit on the plasma membrane. In mature spores (Fig. 23C), gold particles are mainly seen at the plasma membrane - endospore interface. In sporoblasts, the *EcCDA* is also present in the inner part of the cell wall tightly associated with the plasma membrane (Fig. 24A-B-C). This location is consistent with biochemical and sequence data, and suggests a role in cell-wall formation. Late sporogony is characterized by the thickening of the chitin-containing endospore [12], but the mechanism of chitin deposition is unknown. Since the *EcCDA* enzyme is regularly distributed at the plasma membrane during the whole sporogony, chitin may be produced all around the cell.

Late sporoblasts frequently harbour a membrane-rich structure (300 x 200-250 nm) which reacts with *EcCDA*-specific PAb (Fig. 24A-B-D-E), and which opens to the wall. Gold particles are clearly associated with membranes of this structure that is filled with an electron-dense material similar to that of the endospore. Figure 24E suggests that this material is flowing between the plasma membrane and the future exospore. In different Microsporidia such as *Pleistophora sp.*, *Stempellia sp.* and *Ameson pulvis*, a similar membranous whorl termed “scindosome” or “paramural body” has been observed at the furrowing sites during division stage of the parasite [13,42].

Since the sporoblast decreases in size as its differentiation proceeds, the *E. cuniculi* paramural body could be interpreted as an endocytic recuperation of membrane/endospore material. Alternatively, chitin might be synthesized and deacetylated within the paramural body before being incorporated to the wall by an exocytosis-like process. As a result, this structure would be the equivalent of the fungal chitosome, an organelle which is specialized for chitin biosynthesis [43,44]. The chitosome hypothesis should be tested through further studies on the expression and location of microsporidian chitin synthase (ECU01\_1390).



## Acknowledgments

D. Brosson was supported by a grant from “Ministère de l’Education, de la Recherche et de la Technologie”. We thank G. Méténier for helpful discussion and critical reading of this manuscript. We also thank B. Chebance and I. Wawrzyniak for *E. cuniculi* cell culture and R. Guerry and A. Voldoire for technical advice.

## References

- [1] Kotler, D.P. and Orenstein, J.M. (1999) Clinical syndromes associated with microsporidiosis. In: *The Microsporidia and Microsporidiosis* (Wittner, editor) pp. 258-292. ASM Press, Washington D.C.
- [2] van Gool, T., Vetter, J.C., Weinmayr, B., Van Dam, A., Derouin, F. and Dankert, J. (1997) High seroprevalence of *Encephalitozoon* species in immunocompetent subjects. *J. Infect. Dis.* 1997, 175, 1020-1024.
- [3] Raynaud, L., Delbac, F., Broussolle, V., Rabodonirina, M., Girault, V., Wallon, M., Cozon, G., Vivares, C.P. and Peyron, F. (1998) Identification of *Encephalitozoon intestinalis* in travelers with chronic diarrhea by specific PCR amplification. *J. Clin. Microbiol.* 36, 37-40.
- [4] van Gool, T., Biderre, C., Delbac, F., Wentink-Bonnema, E., Peek, R. and Vivares, C.P. (2004) Serodiagnostic studies in an immunocompetent individual infected with *Encephalitozoon cuniculi*. *J. Infect. Dis.* 189, 2243-2249.
- [5] Keeling, P.J., Luker, M.A. and Palmer, J.D. (2000) Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. *Mol. Biol. Evol.* 17, 23-31.
- [6] Keeling, P.J. (2003) Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. *Fungal. Genet. Biol.* 38, 298-309.
- [7] Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretailade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J. and Vivarès C.P. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414, 450-453.
- [8] Thomarat, F., Vivarès, C.P. and Gouy, M. (2004) Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. *J. Mol. Evol.* 59, 780-791.
- [9] Shadduck, J.A. and Polley, M.B. (1978) Some factors influencing the in vitro infectivity and replication of *Encephalitozoon cuniculi*. *J. Protozool.* 25, 491-496.
- [10] Koudela, B., Kucerova, S. and Hudcovic, T. (1999) Effect of low and high temperatures on infectivity of *Encephalitozoon cuniculi* spores suspended in water. *Folia Parasitol. (Praha)* 43, 171-174.
- [11] Vivarès, C.P. and Méténier, G. (2001) The microsporidian *Encephalitozoon*. *Bioessays* 23, 194-202.
- [12] Vavra, J. and Larson, J. (1999) Structure of the Microsporidia. In: *The Microsporidia and Microsporidiosis* (Wittner, editor), pp. 7-84. ASM Press, Washington D.C.
- [13] Vavra, J. (1976) Structure of the Microsporidia. In: *Comparative pathology. Volume 1: Biology of the microsporidia* (L.A. B, T.C. C, editors), pp. 1-86. Plenum Press, New York London.
- [14] Vivarès, C.P., Loubes, C. and Bouix, G. (1976) Recherches cytochimiques approfondies sur les Microsporidies parasites du Crabe Vert de la Méditerranée *Carcinus mediterraneus* Czerniavsky, 1884. *Ann. Parasitol. Hum. Comp.* 51, 1-14.
- [15] Bigliardi, E., Selmi, M.G., Lupetti, P., Corona, S., Gatti, S., Scaglia, M. and Sacchi, L. (1996) Microsporidian spore wall: ultrastructural findings on



- Encephalitozoon hellem* exospore. J. Eukaryot. Microbiol. 43, 181-186.
- [16] Pakes, S.P., Shadduck, J.A. and Cali, A. (1975) Fine structure of *Encephalitozoon cuniculi* from rabbits, mice and hamsters. J. Protozool. 22, 481-488.
- [17] Hamilton, R.C., Cox, J.C. and Pye, D. (1977) Wall structure of the sporonts of *Encephalitozoon cuniculi* grown in human fibroblasts. J. Gen. Microbiol. 98, 305-307.
- [18] Delbac, F., Duffieux, F., David, D., Méténier, G. and Vivarès, C.P. (1998) Immunocytochemical identification of spore proteins in two microsporidia, with emphasis on extrusion apparatus. J. Eukaryot. Microbiol. 45, 224-231.
- [19] Beckers, P.J., Derks, G.J., van Gool, T., Rietveld, F.J. and Sauerwein, R.W. (1996) *Encephalocytozoon intestinalis*-specific monoclonal antibodies for laboratory diagnosis of microsporidiosis. J. Clin. Microbiol. 34, 282-285.
- [20] Lujan, H.D., Conrad, J.T., Clark, C.G., Touz, M.C., Delbac, F., Vivarès, C.P. and Nash, T.E. (1998) Detection of microsporidia spore-specific antigens by monoclonal antibodies. Hybridoma 17, 237-243.
- [21] Visvesvara, G.S., Leitch, G.J., da Silva, A.J., Croppo, G.P., Moura, H., Wallace, S., Slemenda, S.B., Schwartz, D.A., Moss, D., Bryan, R.T., and Pieniazek, N.J. (1994) Polyclonal and monoclonal antibody and PCR-amplified small-subunit rRNA identification of a microsporidian, *Encephalitozoon hellem*, isolated from an AIDS patient with disseminated infection. J. Clin. Microbiol. 32, 2760-2768.
- [22] Bohne, W., Ferguson, D.J., Kohler, K. and Gross, U. (2000) Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cuniculi*. Infect. Immun. 68, 2268-2275.
- [23] Hayman, J.R., Hayes, S.F., Amon, J. and Nash, T.E. (2001) Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*. Infect. Immun. 69, 7057-7066.
- [24] Combet, C., Blanchet, C., Geourjon, C. and Deleage, G. (2000) NPS@: network protein sequence analysis. Trends Biochem. Sci. 25, 147-150.
- [25] Delbac, F., Peyret, P., Méténier, G., David, D., Danchin, A. and Vivarès C.P. (1998) On proteins of the microsporidian invasive apparatus: complete sequence of a polar tube protein of *Encephalitozoon cuniculi*. Mol. Microbiol. 29, 825-834.
- [26] Chevallet, M., Santoni, V., Poinas, A., Rouquie, D., Fuchs, A., Kieffer, S., Rossignol, M., Lunardi, J., Garin, J. and Rabilloud, T. (1998) New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. Electrophoresis 19, 1901-1909.
- [27] Molloy, M.P. (2000) Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. Anal. Biochem. 280, 1-10.
- [28] Martinou, A., Koutsoulis, D. and Bouriotis, V. (2002) Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from *Saccharomyces cerevisiae*. Protein Expr. Purif. 24, 111-116.
- [29] Slamovits, C.H., Williams, B.A. and Keeling, P.J. (2004) Transfer of *Nosema locustae* (Microsporidia) to *Antonospora locustae* n. comb. based on molecular and ultrastructural data. J. Eukaryot. Microbiol. 51, 207-213.
- [30] Tsigos, I., Martinou, A., Kafetzopoulos, D. and Bouriotis, V. (2000) Chitin deacetylases: new, versatile tools in biotechnology. Trends Biotechnol 18, 305-312.
- [31] Kafetzopoulos, D., Thireos, G., Vournakis, J.N. and Bouriotis, V. (1993) The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products. Proc. Natl. Acad. Sci. U S A 90, 8005-8008.
- [32] Kafetzopoulos, D., Martinou, A. and Bouriotis, V. (1993) Bioconversion of chitin to chitosan: purification and characterization of chitin deacetylase from *Mucor rouxii*. Proc. Natl. Acad. Sci. U S A 90, 2564-2568.



- [33] Tsigos, I. and Bouriotis, V. (1995) Purification and characterization of chitin deacetylase from *Colletotrichum lindemuthianum*. J. Biol. Chem. 270, 26286-26291.
- [34] Mishra, C., Semino, C.E., McCreath, K.J., de la Vega, H., Jones, B.J., Specht, C.A. and Robbins, P.W. (1997) Cloning and expression of two chitin deacetylase genes of *Saccharomyces cerevisiae*. Yeast 13, 327-336.
- [35] Christodoulidou, A., Bouriotis, V. and Thireos, G. (1996) Two sporulation-specific chitin deacetylase-encoding genes are required for the ascospore wall rigidity of *Saccharomyces cerevisiae*. J. Biol. Chem. 271, 31420-31425.
- [36] Briza, P., Ellinger, A., Winkler, G. and Breitenbach, M. (1988) Chemical composition of the yeast ascospore wall. The second outer layer consists of chitosan. J. Biol. Chem. 263, 11569-11574.
- [37] Christodoulidou, A., Briza, P., Ellinger, A. and Bouriotis, V. (1999) Yeast ascospore wall assembly requires two chitin deacetylase isozymes. FEBS Lett 460, 275-279.
- [38] Overdijk, B., Van Steijn, G.J. and Odds, F.C. (1999) Distribution of chitinase in guinea pig tissues and increases in levels of this enzyme after systemic infection with *Aspergillus fumigatus*. Microbiology 145, 259-269.
- [39] Nahar, P., Ghormade, V., Deshpande, M.V. (2004) The extracellular constitutive production of chitin deacetylase in *Metarhizium anisopliae*: possible edge to entomopathogenic fungi in the biological control of insect pests. J. Invertebr. Pathol. 85, 80-88.
- [40] Trudel, J. and Asselin, A. (1990) Detection of chitin deacetylase activity after polyacrylamide gel electrophoresis. Anal. Biochem. 189, 249-253.
- [41] Deising, H. and Siegrist, J. Chitin deacetylase activity of the rust *Uromyces viviae-fabiae* is controlled by fungal morphogenesis. FEMS Microbiol. Lett. 127, 207-212.
- [42] Vivarès, C.P. and Sprague, V. (1979) The Fine Structure of *Ameson pulvis* (Microspora, Microsporidia) and its Implication Regarding Classification and Chromosome Cycle. J. Invertebr. Pathol. 33, 40-52.
- [43] Bracker, C.E., Ruiz-Herrera, J. and Bartnicki-Garcia, S. (1976) Structure and transformation of chitin synthetase particles (chitosomes) during microfibril synthesis in vitro. Proc. Natl. Acad. Sci. U S A 73, 4570-4574.
- [44] Ruiz-Herrera, J., Sentandreu Ramon, R. and Martinez, J.P. (1992) Chitin biosynthesis in fungi. In: Handbook of Applied Mycology. Volume 4: Fungal Biotechnology (Arora, D.K., Elander, R.P., Mukerji, K.G., editors), Marcel Dekker, INC, New York, Basel, Hong Kong.



<i>Signaux et motifs</i>	<i>Critères</i>
<i>N-terminal signal peptide</i>	<i>Modified McGeoch's method and the cleavage-site consensus</i>
<i>Mitochondrial-targeting signal</i>	<i>Amino acid composition of the N-terminal 20 residues and some weak cleavage-site consensus</i>
<i>Nuclear-localization signals</i>	<i>Combined score for various empirical rules</i>
<i>ER-lumen-retention signal</i>	<i>The KDEL-like motif at the C-terminus</i>
<i>ER-membrane-retention signal</i>	<i>Motifs: XXRR-like (N-terminal) or KKXX-like (C-terminal)</i>
<i>Peroxisomal-targeting signal</i>	<i>PTS1 motif at the C-terminus and the PTS2 motif</i>
<i>Vacuolar-targeting signal</i>	<i>[TIK]LP[NKI] motif</i>
<i>Golgi-transport signal</i>	<i>The YQRL motif (preferentially at the cytoplasmic tail)</i>
<i>Tyrosine-containing motif</i>	<i>Number of tyrosine residues in the cytoplasmic tail</i>
<i>Dileucine motif</i>	<i>At the cytoplasmic tail</i>
<i>Membrane span(s)/topology</i>	<i>Maximum hydrophobicity and the number of predicted spans; charge difference across the most N-terminal transmembrane segment</i>
<i>RNA-binding motif</i>	<i>RNP-1 motif</i>
<i>Actinin-type actin-binding motifs</i>	<i>From PROSITE</i>
<i>Isoprenyl motif</i>	<i>CaaX motif at the C-terminus</i>
<i>GPI-anchor</i>	<i>Type-1a membrane protein with very short tail</i>
<i>N-myristoylation motif</i>	<i>At the N-terminus</i>
<i>DNA-binding motifs</i>	<i>63 motifs from PROSITE</i>
<i>Ribosomal-protein motifs</i>	<i>71 motifs from PROSITE</i>
<i>Prokaryotic DNA-binding motifs</i>	<i>33 motifs from PROSITE</i>
<i>Amino acid composition</i>	<i>Neural network score that discriminates between cytoplasmic and nuclear proteins</i>
<i>Coiled-coil structure</i>	<i>Number of residues in the predicted coiled-coil state</i>

**Tableau 7.** Motifs reconnus par PSORT II. D'après Nakai et Horton (1999).

## **5.2. Recherche de protéines candidates à une localisation pariétale par criblage bioinformatique du génome de *E. cuniculi***

Pour dresser un inventaire des protéines potentiellement localisées au niveau de la paroi et de la membrane plasmique chez *E. cuniculi*, nous avons défini un ensemble de caractéristiques de séquence compatible avec cette localisation. En fonction de ces critères nous avons établi un crible bioinformatique qui a été confronté aux protéines majeures pariétales connues puis appliqué aux protéines à fonction inconnue présentes dans le protéome sporal (cf chapitre 4. Analyse du protéome...). Les protéines retenues ont fait l'objet d'une investigation expérimentale de leur localisation, ce qui a conduit à la validation de nos choix bioinformatiques. Nous avons ensuite appliqué ces critères à l'ensemble des protéines à fonction inconnue déduites du génome d'*E. cuniculi*.

Chez les eucaryotes, les protéines qui sont destinées à une localisation extracellulaire ou de surface sont exportées par la voie d'exocytose générale. Les protéines qui empruntent cette voie, présentent une séquence signal à leur extrémité N-terminale, requise pour l'entrée dans le réticulum endoplasmique, première étape de la voie de sécrétion. Cependant, quelques protéines de levure identifiées expérimentalement comme extracellulaires, dont notamment HSP70, ne présentent pas cette séquence signal. Le mécanisme moléculaire de ce transport n'est pas encore connu et aucune séquence spécifique n'a pu être mise en évidence (Cleves *et al.*, 1996 ; Pardo *et al.*, 1999). Dans cette étude, nous avons postulé que la majorité des protéines extracellulaires ou de surface suivent la voie d'exocytose générale et qu'elles possèdent un peptide signal à leur extrémité N-terminale. Nous avons utilisé un algorithme bioinformatique (SignalP) permettant de prédire la présence de ce peptide signal. D'autres outils informatiques ont également été utilisés pour compléter notre prédiction de localisation tel qu'une analyse globale multicritère conduisant à une localisation putative (PsortII), une prédiction d'ancrage GPI (DGPI et Big-PI Predictor), ainsi qu'une recherche d'homologies de séquence ou de domaine qui serait une éventuelle indication pour une localisation membranaire ou pariétale.

A.

<i>CDS</i>	<i>Taille du CDS</i>	<i>Taille du produit d'amplification</i>	<i>Limites de la région amplifiée par rapport au début du CDS</i>
ECU01_0820	1074	762	72-834
ECU01_1270	666	612	51-663
ECU11_0510	762	624	96-717
ECU02_0150	663	597	75-660
ECU07_0920	933	701	68-769
ECU08_1730	645	552	54-606
ECU05_0590	756	690	60-750
ECU11_1210	888	831	48-879
ECU09_1460	906	750	54-804

B.

<i>Orientation</i>	<i>Séquence nucléique des oligonucléotides</i>	<i>Site de restriction</i>
<b>ECU01_0820</b>		
Sens	CG <b>GAA TTC</b> ACA TGC GAG CAG GAC CTG AGG	EcoRI
Revers	CG <b>CTC GAG</b> CTG TCC TGG GCA TCG GAG GC	Xho
<b>ECU01_1270</b>		
Sens	CG <b>GGA TCC</b> AGC AGT GGC CTC TAC ATA CC	BamHI
Revers	CG <b>GAA TTC</b> CAT CAC AAT TGA GAA CAG AGC	EcoRI
<b>ECU11_0510</b>		
Sens	CG <b>GAA TTC</b> AGG GAG TTG GAG GAG TTC G	EcoRI
Revers	CG <b>CTC GAG</b> ACT CTC TTC GTA CAG CGG	Xho
<b>ECU02_0150</b>		
Sens	CG <b>GGA TCC</b> ATT CCA GTA GAA GTA TAT G	BamHI
Revers	CG <b>GAA TTC</b> GTG GCA GAC TCC ACC GG	EcoRI
<b>ECU07_0920</b>		
Sens	CG <b>GAA TTC</b> GAT GTC TCG TTG AGC C	EcoRI
Revers	CG <b>CTC GAG</b> GAT GTC CGG AAG GTTC	Xho
<b>ECU08_1730</b>		
Sens	CG <b>GGA TCC</b> GCA AAG AAC AAC GAA GG	BamHI
Revers	GG <b>GAA TTC</b> CTC GTC CTC ATA GGG	EcoRI
<b>ECU05_0590</b>		
Sens	CG <b>GGA TCC</b> TAC GAC CCT GTT CTT CAG C	BamHI
Revers	CG <b>CTC GAG</b> CTT AGA TCC TAT TTG TTC	Xho
<b>ECU11_1210</b>		
Sens	CG <b>GGA TCC</b> CGG GAA GGC CTT GCC TTG	BamHI
Revers	CG <b>GAA TTC</b> GCA CTC GTC ATC GAC AGA	EcoRI
<b>ECU09_1460</b>		
Sens	CG <b>GGA TCC</b> GAC TCC ACG CTA GTT GTG	BamHI
Revers	CG <b>GAA TTC</b> CTC CAT GGA CGC TTT TCC	EcoRI

**Tableau 8.** Caractéristiques des CDS retenus par le crible bioinformatique de recherche de protéines pariétales (A), et des oligonucléotides permettant leur amplification génique (B). Les réactions de PCR ont été effectuées selon des conditions standard avec une température d'hybridation de 55°C.

## 5.2.1. Matériels et Méthodes

### 5.2.1.1. Analyses bioinformatiques

#### 5.2.1.1.1. Prédiction d'un peptide signal

Chez les eucaryotes, les protéines qui sont véhiculées vers l'extérieur par l'intermédiaire du système endomembranaire présentent habituellement une séquence spécifique (appelée peptide signal) qui est clivée après translocation au travers de la membrane du réticulum endoplasmique. Il s'agit communément d'une région chargée positivement à l'extrémité N-terminale, suivie d'une région hydrophobe organisée en hélice, puis d'une région polaire et non chargée. Les acides aminés en -3 et en -1 par rapport au site de clivage doivent être de petites tailles et non chargés pour que le clivage s'effectue correctement (von Heijne, 1983).

La présence d'un peptide signal a été déterminée par deux analyses complémentaires, dont en premier lieu une prédiction par l'algorithme SignalP 3.0 (Nielsen *et al.*, 1997), disponible à l'adresse <http://www.cbs.dtu.dk/services/SignalP>. Des travaux (Zhang et Henzel, 2004) ont expertisé la robustesse des algorithmes de prédiction de peptides signaux en comparant les résultats de prédiction aux données expérimentales pour 270 protéines humaines caractérisées comme protéines sécrétées. L'algorithme le plus efficace de cette étude était SignalP avec un taux de reconnaissance du site de clivage de 78%. Le second algorithme que nous avons utilisé est le logiciel TMPred ([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)) qui permet de prédire l'ensemble des hélices transmembranaires d'une protéine. Nous avons donc ainsi pu confirmer la présence d'une hélice transmembranaire en région N-terminale grâce à l'algorithme TMPred. En outre, dans le cas de protéines présentant un grand nombre d'hélices transmembranaires, on peut légitimement supposer qu'il s'agisse d'une protéine intrinsèque, c'est pourquoi nous avons exclu de notre crible les protéines présentant plus de quatre segments transmembranaires, indication de protéines intrinsèques, tels que les transporteurs et les pores cellulaires. Leur intégration au sein de la membrane plasmique les rendant difficilement accessibles aux anticorps, leur localisation par immunocytochimie est complexe.

#### 5.2.1.1.2. Recherche de signaux d'adressage et de rétention aux organites cellulaires

Parmi les protéines qui sont prises en charge par le système sécrétoire, certaines sont dirigées ou retenues dans le réticulum, l'appareil de Golgi, les vésicules de sécrétion et la membrane plasmique, d'autres suivent la voie « par défaut » conduisant à leur sécrétion. Plusieurs signaux ont été décrits, permettant aussi bien d'adresser les protéines vers un organelle cible, ou au contraire, de les maintenir dans un compartiment précis. Il est donc possible de voir les protéines extracellulaires

	Nombre d'AA	MM th	pl th	MM exp	pl exp	SignalP	TMPred	Psort II	DGPI et Big-PI predictor	Homologie
<b>ECU01_0820</b>	357	40,6	9,3	16,8	5,7	1-16	1-17	44,4% extracellulaire 22,2% cytoplasme 22,2% mitochondrie	non	domaine "keratin high sulfur B2" (riche en cystéine)
<b>ECU01_1270</b>	221	28,2	5,2	22,5	8,5	1-17 1-20	6-23 124-149 201-221	55,6% extracellulaire 44,4% mb plasmique	ancree GPI avec clivage en 193	protéine riche en sérine
<b>ECU11_0510</b>	254	28,1	4,6	32,6	4,5	1-15	1-18 234-253	66,7% extracellulaire 3x11% réticulum / vacuole / mitochondrie	ancree GPI (sans clivage)	domaine polysaccharide déacétylase
<b>ECU02_0150</b>	220	24,6	5,0	27,5	4,8 5,0	1-20	1-23	66,7% extracellulaire 22,2% vacuole	non	
<b>ECU07_0920</b>	311	35,2	5,7	35,2	5,4	1-13	1-17	55,6% extracellulaire 2x11% vacuole / noyau 11% cytoplasme 11% mitochondrie	non	domaine exonucléase phosphatase
<b>ECU08_1730</b>	214	23,4	6,1	23,0	6,1	1-14	1-17	21,7% extracellulaire 21,7% mb plasmique 17,4% cytoplasme 17,4% noyau	non	similaire à ECU 08_1700, ECU08_1710 et ECU 08_1720 domaine "ricin B lectine"
<b>ECU05_590</b>	251	28,4	6,0	29,6 28,2	5,8	1-15	1-20 183-200	55,6% extracellulaire 2x11% vacuole / noyau 11% cytoplasme 11% mitochondrie	non	
<b>ECU11_1210</b>	295	33,0	5,0	36,4	4,7	1-23	1-24	66,7% extracellulaire 22,2% réticulum 11,1% vacuole	non	
<b>ECU09_1460</b>	301	33,9	5,8	32,8	5,6	1-16	39-60 200-223	33,3% extracellulaire 22,2% mb plasmique	non	domaine répété "WD40"

**Tableau 9.** Bilan des protéines candidates à une localisation extracellulaire retenues parmi les protéines à fonction inconnue du protéome sporal de *E. cuniculi*. AA : acides aminés, MM : masse moléculaire, pI : point isoélectrique, exp : expérimental, th : théorique, mb : membrane.

comme ne présentant ni signaux de rétention, ni signaux d'adressage autres que le peptide signal (Nakai et Horton, 1999). Le programme PSORT II (<http://www.psort.nibb.ac.jp/form2.html>) a été développé pour prédire la localisation cellulaire de protéines, en recherchant dans leur séquence la présence de signaux caractéristiques d'une localisation cellulaire (tableau 7). L'absence de ces signaux, à l'exception de ceux responsables d'ancrage membranaire et du peptide signal est le second critère que nous avons retenu pour notre analyse.

#### 5.2.1.1.3. *Prédiction d'ancre GPI et recherche d'homologie de séquences*

Chez les champignons, et particulièrement *S. cerevisiae*, toute une famille de protéines pariétales est ancrée à la membrane plasmique par l'intermédiaire d'un résidu GPI (Frieman et Cormack, 2004). La région où s'effectue la liaison enzymatique de l'ancre GPI se situe au niveau de l'extrémité C-terminale. Elle est constituée d'une partie hydrophobe de 15 à 30 acides aminés, suivie d'une partie hydrophile (environ 8 acides aminés) puis du site d'ancrage qui est un petit acide aminé, souvent la sérine (48%). Nous avons utilisé deux algorithmes : DGPI et big-PI predictor pour prédire ce type de structure, disponibles respectivement aux adresses suivantes [http://129.194.185.165/dgpi/index\\_en.html](http://129.194.185.165/dgpi/index_en.html) et [http://mendel.imp.univie.ac.at/sat/gpi/gpi\\_server.html](http://mendel.imp.univie.ac.at/sat/gpi/gpi_server.html).

Nous avons également inclus dans notre analyse la recherche d'homologies de séquence par BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>) ainsi que de domaine par Prosite (<http://www.expasy.org/prosite>). L'homologie entre protéines ou bien la présence de domaines conservés peut être une indication permettant d'émettre des hypothèses sur la fonction d'une protéine, et donc sur sa localisation.

#### 5.2.1.2. **Biologie moléculaire**

##### 5.2.1.2.1. *Expression et clonage des protéines d'intérêt*

Les protéines ont été exprimées sous forme tronquée en fusion avec la glutathion-S-transférase (GST) à l'extrémité N-terminale et avec un motif de 8xHis à l'extrémité C-terminale. L'ensemble des données de clonage et d'amplification génique est résumé dans le tableau 8. Les produits d'amplification ont été purifiés puis clonés en phase avec la GST dans le vecteur d'expression procaryote pGEX-4T1 modifié, avant le site XhoI, pour présenter un motif de huit histidines à l'extrémité C-terminale de la protéine recombinante. Le plasmide recombinant a été introduit dans une souche d'*E. coli* BL21+ par choc thermique. Après induction avec 2 mM d'IPTG pendant 4 heures, les protéines bactériennes ont été extraites en solution de Laemmli (Tris 125 mM, Glycérol 10%, DTT 0,125 mM, SDS 2%, trace de bleu de Bromophénol) et analysées par SDS-PAGE.

<b>Protéine</b>	<b>Localisation par IFA</b>	<b>Localisation par MET</b>
ECU01_0820	Paroi	Localisation à la surface de la paroi chez les stades jeunes puis à l'endospore chez la spore (EnP1) <sup>a</sup>
ECU01_1270	Paroi	Localisation à la surface de la paroi chez les stades jeunes puis à l'endospore chez la spore (EnP2) <sup>a</sup>
ECU11_0510	Paroi	Localisation à l'interface membrane plasmique endospore dès la mise en place de la paroi (EcCDA) <sup>b</sup>
ECU02_0150	Paroi	Localisation au niveau de l'exospore dès la mise en place de la paroi (ExP1) <sup>c</sup>
ECU07_0920 ECU08_1730	Cytoplasme dans les stades jeunes	Marquage plutôt cytoplasmique difficile à interpréter
ECU05_0590		
ECU11_1210	Non effectué car problèmes à l'étape de clonage du gène d'intérêt	
ECU09_1460		

**Tableau 10.** Localisation cellulaire des protéines candidates à une localisation pariétale. <sup>a</sup> : Peuvél-Fanget et al., sous presse, <sup>b</sup> : Brosson et al., 2005, <sup>c</sup> : Brosson et al., manuscrit soumis. IFA : microscopie à immuno fluorescence.

#### **5.2.1.2.2. Séparation sur gel SDS-PAGE et analyse par Western blot.**

La séparation des protéines sur gel SDS-PAGE et leur analyse par Western Blot ont été effectuées comme décrit précédemment (Brosson *et al.*, 2005).

#### **5.2.1.2.3. Production de sera polyclonaux**

Des anticorps polyclonaux dirigés contre les protéines recombinantes ont été obtenus chez des souris BALB/c et SWISS après injection de bandes protéiques SDS-PAGE broyées. Toutes les deux semaines, les souris ont reçu (entre trois et cinq fois selon les protéines) intra-péritonéalement les échantillons protéiques homogénéisés avec de l'adjuvant de Freund (Sigma). Les sera ont été collectés deux semaines après la dernière injection, puis conservés à -20°C.

#### **5.2.1.2.4. Localisation par immunofluorescence indirecte (IFA)**

Des cellules HFF et MDCK infestées par *E. cuniculi* ont été fixées en méthanol à -80°C, perméabilisées en éthanol 70% / Triton X-100 0,5%, saturées en PBS lait écrémé 5% et incubées 1 heure avec une dilution appropriée d'anticorps primaires en PBS Triton X-100 0,1%. Après traitement pendant 1 heure avec une dilution au 1:1000 d'un anticorps anti-souris couplé à un fluorochrome (Alexia Fluor 488) les préparations ont été examinées grâce à un microscope Leica DMR à épifluorescence.

#### **5.2.1.2.5. Immunolocalisation par Microscopie Electronique à Transmission (MET)**

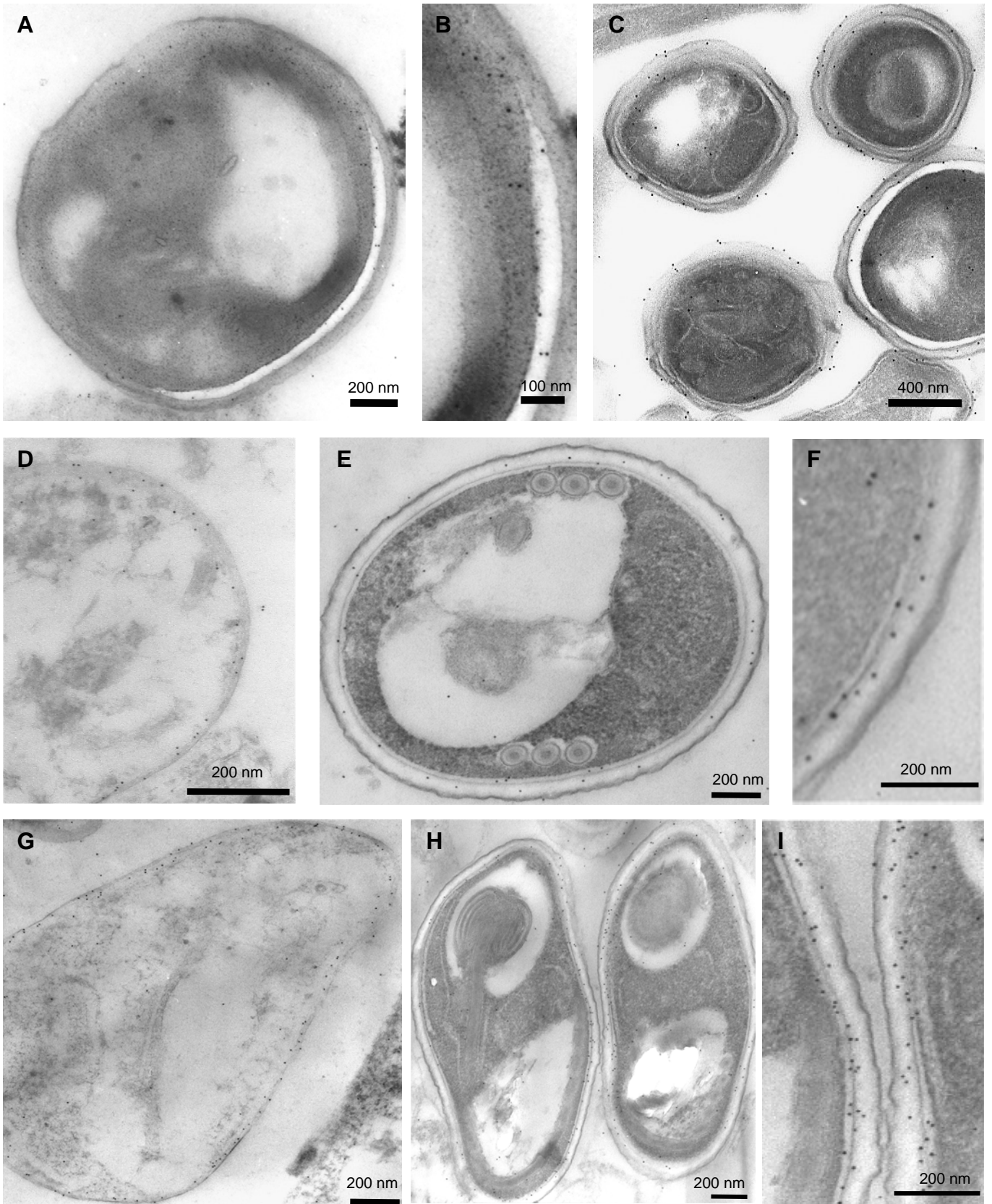
L'immunolocalisation des protéines d'intérêt a été effectuée comme décrite précédemment (Brosson *et al.*, 2005).

### **5.2.2. Résultats**

#### **5.2.2.1. Validation du crible bioinformatique sur des protéines expérimentalement identifiées comme pariétales**

Nous avons appliqué le crible bioinformatique devant permettre de repérer des protéines candidates à une localisation pariétale, aux deux seules protéines connues pour être localisées au niveau de la paroi, précisément à l'interface endospore - membrane plasmique pour *EcCDA* (Brosson *et al.*, 2005) et au niveau de l'exospore pour SWP1 (Bohne *et al.*, 2000). Pour chacune d'elle, SignalP prédit un peptide signal confirmé par TMPred. L'algorithme Psort II prédit une localisation extracellulaire avec une probabilité de 44% pour SWP1 et 66% pour *EcCDA*. De plus, la prédiction





**Figure 25.** Immunolocalisation des protéines pariétales par MET : ECU 11\_0510 (EcCDA) dans une spore mature (A,B) ; ECU02\_0150 (ExP1) chez plusieurs stades de fin de sporogenèse (C) ; ECU01\_0820 (EnP1) dans un sporoblaste (D) et dans une spore mature (E,F) ; ECU01\_1270 (EnP2) dans un sporonte (G) et dans des spores matures (H, I).

pour *EcCDA* d'une hélice transmembranaire en position C-terminale, peut expliquer sa localisation à proximité de la membrane plasmique.

Le bilan de la prédiction bioinformatique permet de conclure à une localisation extracellulaire prédite qui est concordante avec la localisation expérimentale.

#### 5.2.2.2. Recherche de candidates pariétales dans le protéome expérimental de *E. cuniculi*

Dans un premier temps, afin de travailler sur des protéines exprimées, nous avons recherché des protéines candidates à une localisation pariétale parmi les protéines à fonction inconnue identifiées dans le protéome expérimental. Ainsi, nous avons retenu parmi les 177 protéines identifiées dans l'analyse du protéome des stades tardifs d'*E. cuniculi* (Brosson *et al.* en préparation) les 45 protéines à fonction inconnue. Nous les avons criblées selon les critères bioinformatiques retenus et nous avons ainsi sélectionné neuf protéines candidates (tableau 9).

Afin de déterminer expérimentalement la localisation de ces protéines, nous avons cloné le gène d'intérêt en vecteur d'expression, purifié la protéine recombinante en fusion avec la GST avant de l'injecter à des souris. Les séras polyclonaux obtenus ont permis d'effectuer des marquages immunocytochimiques en microscopie optique à fluorescence et en MET (tableau 10) et figure 25. Parmi ces neuf candidates, l'immunolocalisation par MET a permis de conclure que quatre protéines sont clairement localisées au niveau de la paroi ou de la membrane plasmique (figure 25). Les protéines ECU01\_0820 (EnP1) et ECU01\_1270 (EnP2) sont localisées au sein de l'endospore (Peuvel-Fanget *et al.*, sous presse). La protéine ECU02\_0150 se situe au niveau de l'exospore (Brosson *et al.*, manuscrit soumis). Enfin, ECU11\_0510 ou *EcCDA* se situe au niveau de la membrane plasmique (cf chapitre 5.1. La chitine déacétylase...) (Brosson *et al.*, 2005).

Ces résultats permettent de confirmer la validité des hypothèses sur lesquelles nous avons construit notre crible. De plus, nous pouvons noter que ces quatre protéines présentent une probabilité supérieure à 40% de prédiction de localisation extracellulaire avec l'algorithme Psort II. Pour la protéine ECU08\_1730, cette probabilité n'est que de 21,7% et aucune localisation pariétale n'a été observée. Les critères de prédiction d'une protéine pariétale finalement retenus sont les suivants : prédiction d'un peptide signal avec Signal P et confirmation d'une hélice transmembranaire à l'extrémité N-terminale par TMPred, prédiction d'une localisation extracellulaire avec une probabilité supérieure à 40% par Psort II et recherche d'éventuel signaux d'ancrage GPI. Ce crible nous a donc permis de sélectionner cinq protéines, auxquelles nous pouvons ajouter la protéine SWP1 (cf ci-dessus). Sur ces six protéines, seule ECU07\_0820 se révèle être un faux positif avec une localisation cytoplasmique et non pariétale. Par conséquent, notre crible est validé

<i>Analyse</i>	<i>Nombre de protéines retenues</i>
<i>Protéines à fonction inconnue</i>	<i>1058</i>
<i>Analyse SignalP et TMPred</i>	<i>211</i>
<i>Analyse Psort (extracellulaire &gt;40%)</i>	<i>29</i>
<i>Nombre total de protéines candidates</i>	<i>11</i>

**Tableau 11.** Bilan des protéines retenues pour une localisation pariétale après criblage bioinformatique des protéines à fonction inconnue codées par le génome d'*E. cuniculi*.

	Nombre d'AA	SignalP	TMPred	Psort II	DGPI et Big-PI predictor	Homologie
<b>ECU01_1070</b>	400	1-17	1-15	44,4% extracellulaire 33,3% noyau	rien	
<b>ECU03_0100</b>	522	1-18	3-20	66,7% extracellulaire 11,1% vacuole	rien	similaire à la famille des phosphohydrolases
<b>ECU04_0410</b>	778	1-29	9-28	77,8% extracellulaire 11,1% vacuole	rien	
<b>ECU05_0420</b>	417	1-17	1-19	55,6% extracellulaire 22,2% réticulum 11,1% vacuole	rien	
<b>ECU05_1300</b>	962	1-20	2-21	55,6% extracellulaire 22,2% cytoplasme	rien	
<b>ECU06_0640</b>	127	1-17	1-20	55,6% extracellulaire 33,3% noyau	rien	
<b>ECU07_1430</b>	610	1-14	1-19	77,8% extracellulaire	rien	similarité avec la myosin-3
<b>ECU08_1370</b>	245	1-13	1-17	55,6% extracellulaire 11,1% réticulum	rien	
<b>ECU09_1340</b>	147	1-16	4-23	44,4% extracellulaire 22,2% mb plasmique	rien	Similarité avec la « flagellar GTP-binding protein »
<b>ECU09_1630</b>	321	1-35	13-31 189-211	44,4% extracellulaire 22,2% vacuole 22,2% cytoplasme	rien	
<b>ECU10_0650</b>	309	1-19	258-280	44,4% extracellulaire 33,3% cytoplasme 22,2% réticulum	rien	

**Tableau 12.** Bilan des protéines candidates à une localisation extracellulaire retenues parmi les protéines à fonction inconnue déduites du génome d'*E. cuniculi*. MB : membrane, AA : acides aminés.

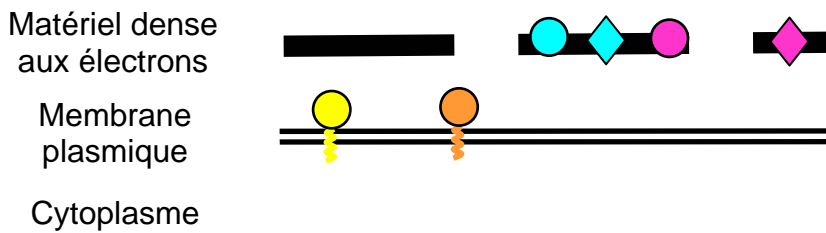
expérimentalement avec une probabilité supérieure à plus de 83% (cinq protéines sur les six vérifiant le crible sont effectivement pariétales).

### **5.2.2.3. Recherche de protéines candidates à une localisation pariétale dans l'ensemble des protéines d'*E. cuniculi***

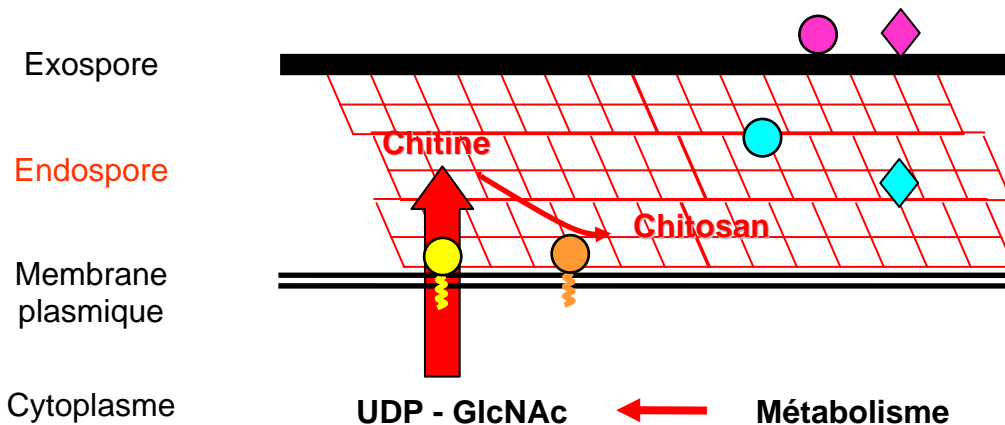
Nous avons appliqué ce crible bioinformatique à l'ensemble des séquences protéiques déduites du génome d'*E. cuniculi* pour laquelle aucune fonction n'a pu être attribuée. Ces protéines sont en effet supposées être, au moins en partie, spécifiques des microsporidies et constituent donc un réservoir intéressant pour rechercher des composants propres à la paroi microsporidienne.

Nous avons d'autre part exclu de notre sélection les protéines appartenant aux familles multigéniques situées aux extrémités des chromosomes. L'organisation génomique et l'expression de ces familles, complexe à analyser, fait l'objet de travaux spécifiques au sein de notre équipe. Les résultats de l'analyse des protéines retenues sont résumés dans le tableau 11. Nous avons donc finalement constitué un groupe de 11 protéines candidates à une localisation pariétale dont les caractéristiques sont présentées dans le tableau 12. Leur validation expérimentale est actuellement en cours.







**Sporonte : Dépôt de matériel pariétal en périphérie de la cellule**



**Sporoblaste – Spore : mise en place de l'endospore**



**Légende**

	<b>EcCDA</b>	ECU11_0510		<b>ExP1</b>	ECU02_0150
	<b>Chitine synthase</b>	ECU01_1390		<b>SWP1</b>	ECU10_1660
				<b>EnP1</b>	ECU01_0820
				<b>EnP2</b>	ECU01_1270

**Figure 26.** Modèle de morphogenèse de la paroi microsporidienne. Chez le jeune sporonte, la future paroi se met en place sous forme de dépôts de matériel dense à la périphérie. Expérimentalement la chitine déacétylase a été localisée à l'interface membrane plasmique-endospore alors que *Exp1*, *SWP1*, *EnP1* et *ENP2* ont été localisées à la surface de ce dépôt dense. Par analogie avec les champignons nous avons intégré à ce modèle une chitine synthase (déduite du génome d'*E. cuniculi*) localisée à proximité de la chitine déacétylase et ancrée à la membrane par sept segments transmembranaires. A partir du stade sporoblaste l'endospore s'épaissit par mise en place de la chitine. La chitine déacétylase « *EcCDA* » prendrait en charge directement certaines chaînes néoformées de chitine pour conduire à la formation de chitosan. En toute fin de sporogenèse, la paroi subit des remaniements structurels conduisant à la relocalisation de certaines protéines (*EnP1* et *EnP2*), pour atteindre leur localisation définitive chez la spore mature. Ces observations indiquent que le matériel dense déposé chez les sporontes ne constitue pas directement la future exospore.

### 5.3. Proposition d'un modèle de mise en place de la paroi microsporidienne

L'étude de la cinétique de mise en place de différentes protéines pariétales identifiées chez *E. cuniculi* : SWP1 (Bohne *et al.*, 2000), EcCDA (Brosson *et al.*, 2005), EnP1 et EnP2 (Peuvel-Fanget *et al.*, sous presse), ExP1 (Brosson *et al.* en préparation) nous permet de proposer un modèle de morphogenèse de la paroi microsporidienne (figure 26).

Lors de la transition entre les stades méronte et sporonte, des dépôts de matériel dense aux électrons organisés sous forme de lamelles sont mis en place à la périphérie du parasite. Alors que la chitine déacétylase est localisée à l'interface membrane plasmique-endospore, les autres protéines sont localisées à la surface de ce dépôt dense.

C'est à partir du stade sporoblaste que débute la mise en place de l'endospore qui s'épaissit progressivement par ajout de chitine jusqu'à la spore mature. Chez les champignons, la chitine déacétylase forme un complexe avec la chitine synthase, où la chitine déacétylase prend en charge directement les chaînes néoformées par la chitine synthase (Kafetzopoulos *et al.*, 1993). Nous postulons que la situation est similaire chez *E. cuniculi* : ces deux protéines seraient donc ancrées à la surface de la membrane plasmique et l'activité de leur tandem serait maximale durant toute la période conduisant à l'épaississement de l'endospore.

Parallèlement à cet épaississement de l'endospore les constituants protéiques de la paroi se réorganisent. Certains, en fin de sporogenèse (passage sporoblaste-spore), subissent en effet une réorganisation qui aboutit à leur localisation définitive. Les protéines SWP1 et ExP1 sont mises en place, chez le sporonte à la surface des dépôts de matériel dense et demeurent localisées à la surface de la paroi de la spore mature. Au contraire, les protéines EnP1 et EnP2, localisées à la surface de la paroi dès le stade sporonte, sont relocalisées au sein de l'endospore chez la spore mature. Ces observations indiquent que la paroi en toute fin de sporogenèse subit des remaniements structurels. Ces données originales permettent de conclure que le matériel dense déposé à la périphérie des sporontes n'est pas directement la future exospore. Ainsi, cette paroi microsporidienne semble, à l'image de la paroi des levures, bien plus dynamique qu'elle n'est décrite jusqu'à présent dans la bibliographie.



*Discussion*  
*générale et*  
*perspectives*





# 1. LE PROTEOME MICROSPORIDIEN

---

Ces travaux de thèse ont débuté alors que notre équipe venait de terminer le séquençage du génome d'*E. cuniculi*. Ainsi, c'est dans un contexte de post-génomique que nous avons effectué, pour la première fois chez les microsporidies, une analyse descriptive du protéome des stades tardifs du cycle de développement de *E. cuniculi*. Pour ces travaux, nous avons mené une double stratégie, d'une part une séparation des protéines par électrophorèse bidimensionnelle puis analyse des spots par spectrométrie de masse MS et MS/MS, et d'autre part une analyse par LC-MS/MS d'extraits protéiques digérés par la trypsine. Nous avons ainsi pu identifier un total de 177 protéines différentes, qui nous ont permis de dresser la première carte protéique d'*E. cuniculi*. Ces résultats nous éclairent sur la physiologie de la fin de la sporogenèse microsporidienne, notamment sur les protéines participant aux sécrétions et structures de la spore, aux métabolismes énergétiques et glucidiques, au contrôle du cycle cellulaire ainsi qu'à la survie de la spore dans l'environnement. Cette première analyse du protéome d'une microsporidie a également permis à notre équipe d'acquérir l'expérience nécessaire pour envisager une stratégie plus large de protéomique.

L'analyse du protéome microsporidien a montré que de nombreuses protéines étaient présentes sous de multiples isoformes, jusqu'à cinq dans le cas de l'actine. Ces modifications, comme les glycosylations et les phosphorylations, sont très peu connues chez les microsporidies, en dépit de leur très probable implication dans la régulation et l'activité de toute la physiologie du parasite. L'étude de ses modifications à l'échelle de l'organisme représente donc un nouveau domaine d'investigation qui permettrait d'obtenir une vision globale des mécanismes de régulation des protéines et part conséquent de la cellule entière. Ainsi, l'étude du phosphoprotéome des microsporidies pourrait conduire à l'obtention de données sur les cascades de régulation cellulaire. Cet axe de recherche vient de débiter au laboratoire, avec l'exploitation des cartes protéiques obtenues grâce à l'analyse descriptive. Celles-ci sont hybridées avec des anticorps spécifiques des acides aminés phosphorylés (sérine, thréonine, tyrosine) permettant de définir un ensemble de protéines ayant été modifiées. Ces protéines sont des candidates pour être des acteurs majeurs dans les voies de signalisation de la cellule, présidant par exemple aussi bien au changement de stade qu'à la régulation du cycle cellulaire. De plus, ces protéines phosphorylées sont des cibles de choix pour le développement d'applications médicales (cibles thérapeutiques).



L'investigation du glycoprotéome microsporidien est un sujet très riche en raison de la quasi -inexistence des données portant sur les motifs glycosylés présents sur les glycoprotéines microsporidiennes. Il semblerait, d'après la reconstruction du métabolisme à partir du génome de *E. cuniculi* et d'*Antonospora locustae*, que les microsporidies soient capables d'effectuer uniquement la O-glycosylation (V. Taupin, communication personnelle) mais cela n'a jamais été démontré. Chez les champignons pathogènes, les glycoprotéines jouent un rôle majeur dans les mécanismes d'interaction et de reconnaissance avec l'hôte. Ainsi, l'étude du rôle et de la nature des glycosylations chez les microsporidies pourraient permettre de mieux comprendre les mécanismes conduisant à la reconnaissance de la cellule hôte et au déclenchement des processus d'invasion (comme par exemple la dévagination du tube polaire). Des travaux portant sur l'étude du glycoprotéome ont débuté dans le laboratoire afin d'obtenir la nature exacte de l'ensemble des glycosylations présentes dans des extraits protéiques microsporidiens et de purifier des protéines glycosylées par chromatographies d'affinité avec des lectines spécifiques de glycosylations données.

L'un des axes majeurs de recherche en protéomique parasitaire est l'analyse comparative des protéomes de chaque stade de développement. En effet, cette approche, où les profils d'expression protéique de chaque stade sont comparés, constitue une véritable « pierre de rosette » pour l'étude de la physiologie de ces stades. Ces analyses permettent de définir aussi bien des protéines spécifiques à un stade donné, que des protéines ayant un taux d'expression différent entre deux stades. De même, disposer de données sur les protéines spécifiques à chaque stade permet de développer des outils de reconnaissance ayant des retombées aussi bien fondamentales qu'appliquées : meilleure compréhension du cycle parasitaire, outils de diagnostic, cibles biologiques pour le développement d'anti-parasitaires. Les variations dans les profils protéiques sont également le reflet de l'adaptation d'un stade de développement à son hôte (dans le cas de parasites hétéroxènes), et leurs études permettent de mieux appréhender les interactions hôte-parasite.

Cette analyse de protéomique comparative, appliquée aux microsporidies, permettrait d'élucider de nombreuses inconnues concernant les acteurs de l'organisation du cycle parasitaire, sur l'identification de « facteurs de virulences » tels que les protéines impliquées dans le déclenchement du mécanisme invasif, ainsi que toutes les protéines participant aux transferts entre le parasite et la cellule hôte. Malheureusement, chez *E. cuniculi*, nous sommes actuellement incapables de purifier et séparer de grandes quantités de cellules d'un stade donné, rendant une approche de protéomique classique par 2D-PAGE impossible. Cependant, la très grande sensibilité de nouvelles stratégies d'analyse comme la LC-MS/MS ou l'analyse MudPIT, permettent d'envisager une analyse



protéomique comparative avec les faibles quantités de matériel biologique que l'on peut obtenir avec les protocoles de purification qui sont actuellement développés au laboratoire.

Le protéome de *E.cuniculi* étudié ici présente une forte proportion de protéines à fonction inconnue. Ces protéines sont, pour certaines, parmi les plus exprimées du protéome, et jouent sans aucun doute un rôle important pour la biologie du parasite, ainsi l'identification de leur fonction permettrait d'effectuer de grandes avancées dans la connaissance du parasite. La fonction d'une protéine peut être vue comme un ensemble de données convergentes, dont l'interprétation permet d'obtenir une présomption de fonction. Ces données sont la cinétique d'expression de la protéine, sa localisation, les protéines avec lesquelles elle interagit, ses homologues de structures (acides aminés, domaines, repliements tridimensionnels), et dans le meilleur des cas une activité biologique prouvée expérimentalement. Cette dernière donnée étant rare, l'identification de la fonction nécessite donc l'obtention des autres données pour obtenir un faisceau de présomptions convergeant vers une fonction. L'une des données qui peut être obtenue est le réseau d'interaction dans lequel se place une protéine. En étudiant à l'échelle du protéome les interactions entre protéines. Cette analyse, qui constitue la protéomique fonctionnelle, donne des indices sur la fonction d'une protéine selon la fonction des protéines avec lesquelles elle interagit. Des travaux d'analyse de l'interactome d'*E.cuniculi* sont en cours dans notre équipe et sont principalement centrés sur les protéines de contrôle du cycle cellulaire, du métabolisme fer-soufre, ainsi que sur les protéines de structure. Un autre domaine permettant d'apporter des réponses sur ces protéines est une analyse par protéomique structurale. En effet, malgré le nombre croissant de séquences protéiques disponibles dans les banques de données, la simple homologie de séquence primaire n'a pas permis d'identifier la fonction de ces protéines. Il est donc nécessaire de changer de niveau d'organisation et d'identifier, grâce à la résolution de leur structure tridimensionnelle, des domaines ou repliements caractéristiques d'une activité biologique donnée. Notre équipe est impliquée dans l'évaluation de la faisabilité de l'analyse cristallographique des protéines microsporidiennes.



## 2. LES PROTEINES PARIETALES MICROSPORIDIENNES

---

Le second volet de ces travaux de thèse porte sur l'étude des protéines de paroi chez *E. cuniculi*. Il s'agit en effet d'une structure essentielle à la survie et aux processus d'invasion de la spore, mais sur laquelle peu de chose sont connues, particulièrement d'un point de vue moléculaire. Rappelons que les informations antérieures concernant cette structure ne mentionnaient la présence que d'une seule protéine chez *E. cuniculi* : SWP1. Nous avons exploité les données fournies par le protéome afin de rechercher parmi les protéines à fonction inconnue de nouvelles protéines pariétales. Grâce à la sélection d'un ensemble de critères bioinformatiques, nous avons déterminé un panel de protéines candidates à une localisation pariétale. Leur localisation expérimentale a permis d'identifier quatre nouvelles protéines de paroi EnP1, EnP2, ExP1 et *EcCDA*, une enzyme très probablement impliquée dans la mise en place de la paroi et de proposer le premier modèle dynamique de la morphogenèse de la paroi chez les microsporidies.

Nous avons appliqué le crible bioinformatique à l'ensemble des protéines à fonction inconnue codées par le génome de *E. cuniculi*, et ainsi retenu 11 protéines candidates. Les résultats de ces travaux laissent donc envisager l'existence d'un plus large répertoire de protéines pariétales. L'une des perspectives les plus immédiates de cette thèse est la localisation expérimentale de ces protéines candidates, selon la même stratégie que nous avons également développée : clonage des gènes, production de protéines recombinantes, production d'anticorps et immunolocalisation par IFA et MET. Il serait également intéressant de produire ces anticorps chez plusieurs organismes dans le but d'effectuer des expériences de colocalisation, entre protéines à fonction inconnue mais aussi entre elles et des enzymes impliquées dans la morphogenèse de la paroi (chitine synthase par exemple). Cela permettrait d'appréhender de possibles interactions entre les protéines localisées au sein d'une même structure : EnP1 et EnP2 dans l'endospore, SWP1 et ExP1 au niveau de l'exospore, la chitine synthase et la chitine déacétylase au niveau de la membrane plasmique. L'analyse de ces colocalisations serait un moyen de préciser leur cinétique de mise en place au cours du cycle de développement du parasite, et par conséquent de préciser la morphogenèse de la paroi. Ces résultats, grâce à leurs validations expérimentales, peuvent également venir pondérer des analyses à plus grande échelle en servant de témoins positifs comme par exemple dans le cas d'analyses de l'ensemble des réseaux d'interactions chez les microsporidies.





Concernant l'attribution de fonction à ces protéines pariétales, l'identification de leurs interactions est une donnée importante mais cependant limitée. En effet, si l'on peut proposer des rôles biologiques à une protéine au centre d'un réseau d'interaction avec des partenaires dont la fonction est définie, il n'en va pas de même dans le cas où la majorité des partenaires n'ont eux même pas de fonction attribuée. C'est pourtant, au vu du faible nombre de protéines pariétales dont l'activité est connue, à ce genre de situation que l'on risque d'être confronté chez les microsporidies. Ce problème peut être d'ailleurs posé de façon plus générale à l'ensemble des protéines à fonction inconnue.

Ces travaux ont également permis d'apporter des précisions sur la morphogénèse de la paroi microsporidienne qui semble bien plus complexe que précédemment décrite, avec le dépôt dans un premier temps de l'exospore, puis épaissement de l'endospore. Nos résultats montrent qu'il s'agit d'une structure beaucoup plus dynamique, à l'image de ce qui est décrit pour la paroi des levures. En effet, nous avons montré une réorganisation de la répartition de certaines protéines entre le début et la fin de la sporogénèse. Cela nous permet de conclure que la partie la plus externe de la paroi, en début de sporogénèse, n'est pas directement la future exospore. Il serait donc intéressant d'identifier les régulations de la morphogénèse pariétale, ainsi que les mécanismes conduisant à la relocalisation des protéines. Ces mécanismes pourraient être appréhendés par inactivation génique des protéines potentiellement impliquées dans cette régulation et l'observation des phénotypes résultants. Cependant cette approche n'est que théorique car les outils de génétique inverse ne sont disponibles actuellement pour aucune microsporidie. Une autre approche pourrait être basée sur l'utilisation de drogues conduisant à des modifications structurales au sein de la paroi, comme par exemple des inhibiteurs des enzymes du métabolisme de la chitine. Si les organismes sont viables en présence de ces drogues, les bouleversements et modifications structuraux qui en découlent pourraient nous éclairer sur les interactions entre constituants de la paroi.

La poursuite de ces travaux pourrait également être envisagée d'un point de vue biochimique dans le but de déterminer exactement l'ensemble des constituants de la paroi. En effet, la mise en évidence, au niveau de la paroi, d'une protéine qui est très probablement une enzyme impliquée dans la conversion de la chitine en chitosan révèle que la nature biochimique de tous les constituants pariétaux n'est pas encore totalement connue. Ainsi, la mise en évidence directe du chitosan au sein de la paroi serait un des moyens de confirmer la fonctionnalité de la protéine *EcCDA*. De même, la localisation exacte du chitosan permettrait de mieux comprendre l'étonnante résistance mécanique et chimique de cette paroi, ainsi que sa résistance aux digestions enzymatiques par les chitinases. La présence d'une chitine déacétylase, ainsi que les données de



phylogénie récentes, permettent d'établir un parallèle avec la paroi des levures. Dès lors, il serait intéressant de mener des investigations visant à vérifier la présence ou non des autres constituants classiquement retrouvés dans les parois des formes de résistance fongique (mannoprotéines structurales telles que le mannane, la dityrosine).

La mise en évidence d'homologues à la protéine SWP1 chez *Encephalitozoon intestinalis*, ainsi que, comme nous l'avons déjà signalé, la présence d'un gène homologue à celui codant pour la chitine déacétylase chez la microsporidie parasite d'insecte *A. locustae*, laissent penser qu'une certaine conservation peut être retrouvée entre protéines de paroi chez les microsporidies. Ainsi nous pourrions étendre la recherche de ces protéines à l'ensemble du genre *Encephalitozoon*. Une première approche pourrait être la recherche de réactions croisées grâce aux anticorps développés contre les protéines de *E. cuniculi*. Ces profils définis par l'ensemble des réactions croisées et des réactions spécifiques permettraient l'utilisation de ces anticorps pour le typage et l'identification des différentes souches et/ou espèces du genre *Encephalitozoon*.

Pour conclure, ces travaux de thèse apportent des éléments nouveaux sur la physiologie de la spore microsporidienne ainsi que sur la biochimie et la morphogenèse de sa paroi. Ils valident et proposent également de nouveaux outils (prédiction de protéines pariétales, cartes protéiques et analyses protéomiques), et permettent aux microsporidies, pathogènes opportunistes en émergence, de s'ancrer encore plus dans les champs d'investigation les plus actuels de la parasitologie.



*R é f é r e n c e s*  
*b i b l i o g r a p h i q u e s*



## REFERENCES BIBLIOGRAPHIQUES

- Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, Deng M, Liu C, Widmer G, Tzipori S, Buck GA, Xu P, Bankier AT, Dear PH, Konfortov BA, Spriggs HF, Iyer L, Anantharaman V, Aravind L, Kapur V. 2004.** Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* 304, 441-445.
- Acestor N, Masina S, Walker J, Saravia NG, Fasel N, Quadroni M. 2002.** Establishing two-dimensional gels for the analysis of *Leishmania* proteomes. *Proteomics*. 2, 877-879.
- Alberti-Segui C, Morales AJ, Xing H, Kessler MM, Willins DA, Weinstock KG, Cottarel G, Fechtel K, Rogers B. 2004.** Identification of potential cell-surface proteins in *Candida albicans* and investigation of the role of a putative cell-surface glycosidase in adhesion and virulence. *Yeast* 21, 285-302.
- Amster IJ. 1996.** Fourier Transform Mass Spectrometry. *Journal of Mass Spectrometry* 31, 1325-1337.
- Anderson NL, Anderson NG. 1998.** Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* 19, 1853-1861.
- Atwood JA, Weatherly DB, Minning TA, Bundy B, Cavola C, Opperdoes FR, Orlando R, Tarleton RL. 2005.** The *Trypanosoma cruzi* proteome. *Science* 309, 473-476.
- Barker R. 1975.** Ultrastructural observation on *Encephalitozoon cuniculi* Levaditi, Nicolau et Schoen, 1922, from mouse peritoneal macrophages. *Folia Parasitol. (Praha)* 22, 1-9.
- Beckers PJ, Derks GJ, Gool T, Rietveld FJ, Sauerwein RW. 1996.** *Encephalitozoon intestinalis*-specific monoclonal antibodies for laboratory diagnosis of microsporidiosis. *J. Clin. Microbiol.* 34, 282-285.
- Bente M, Harder S, Wiesgigl M, Heukeshoven J, Gelhaus C, Krause E, Clos J, Bruchhaus I. 2003.** Developmentally induced changes of the proteome in the protozoan parasite *Leishmania donovani*. *Proteomics*. 3, 1811-1829.
- Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Bohme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC, Arrowsmith C, Atkin RJ, Barron AJ, Bringaud F, Brooks K, Carrington M, Cherevach I, Chillingworth TJ, Churcher C, Clark LN, Corton CH, Cronin A, Davies RM, Doggett J, Djikeng A, Feldblyum T, Field MC, Fraser A, Goodhead I, Hance Z, Harper D, Harris BR, Hauser H, Hostetler J, Ivens A, Jagels K, Johnson D, Johnson J, Jones K, Kerhornou AX, Koo H, Larke N, Landfear S, Larkin C, Leech V, Line A, Lord A, Macleod A, Mooney PJ, Moule S,**





- Martin DM, Morgan GW, Mungall K, Norbertczak H, Ormond D, Pai G, Peacock CS, Peterson J, Quail MA, Rabbinowitsch E, Rajandream MA, Reitter C, Salzberg SL, Sanders M, Schobel S, Sharp S, Simmonds M, Simpson AJ, Tallon L, Turner CM, Tait A, Tivey AR, Van Aken S, Walker D, Wanless D, Wang S, White B, White O, Whitehead S, Woodward J, Wortman J, Adams MD, Embley TM, Gull K, Ullu E, Barry JD, Fairlamb AH, Opperdoes F, Barrell BG, Donelson JE, Hall N, Fraser CM, Melville SE, El Sayed NM. 2005.** The genome of the African trypanosome *Trypanosoma brucei*.  
*Science* 309, 416-422.
- Biderre C, Pages M, Metenier G, Canning EU, Vivarès CP. 1995.** Evidence for the smallest nuclear genome (2.9 Mb) in the microsporidium *Encephalitozoon cuniculi*.  
*Mol. Biochem. Parasitol.* 74, 229-231.
- Bigliardi E, Selmi MG, Lupetti P, Corona S, Gatti S, Scaglia M, Sacchi L. 1996.** Microsporidian spore wall: ultrastructural findings on *Encephalitozoon hellem* exospore.  
*J. Eukaryot. Microbiol.* 43, 181-186.
- Biron DG, Moura H, Marche L, Hughes AL, Thomas F. 2005.** Towards a new conceptual approach to "parasitoproteomics".  
*Trends Parasitol.* 21, 162-168.
- Bjellqvist B, Ek K, Righetti PG, Gianazza E, Gorg A, Westermeier R, Postel W. 1982.** Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications.  
*J Biochem. Biophys. Methods* 6, 317-339.
- Blackman MJ, Bannister LH. 2001.** Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation.  
*Mol. Biochem. Parasitol.* 117, 11-25.
- Blomberg A. 1997.** Osmoresponsive proteins and functional assessment strategies in *Saccharomyces cerevisiae*.  
*Electrophoresis* 18, 1429-1440.
- Bohne W, Ferguson DJ, Kohler K, Gross U. 2000.** Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cuniculi*.  
*Infect. Immun.* 68, 2268-2275.
- Boucherie H, Sagliocco F, Joubert R, Maillet I, Labarre J, Perrot M. 1996.** Two-dimensional gel protein database of *Saccharomyces cerevisiae*.  
*Electrophoresis* 17, 1683-1699.
- Bracker CE, Ruiz-Herrera J, Bartnicki-Garcia S. 1976.** Structure and transformation of chitin synthetase particles (chitosomes) during microfibril synthesis in vitro.  
*Proc. Natl. Acad. Sci. U. S. A* 73, 4570-4574.
- Braunfuchsova P, Salat J, Kopecky J. 2001.** CD8+ T lymphocytes protect SCID mice against *Encephalitozoon cuniculi* infection.  
*Int. J. Parasitol.* 31, 681-686.



- Briza P, Winkler G, Kalchhauser H, Breitenbach M. 1986.** Dityrosine is a prominent component of the yeast ascospore wall. A proof of its structure.  
*J Biol Chem* 261, 4288-4294.
- Briza P, Ellinger A, Winkler G, Breitenbach M. 1988.** Chemical composition of the yeast ascospore wall. The second outer layer consists of chitosan.  
*J Biol Chem* 263, 11569-11574.
- Briza P, Eckerstorfer M, Breitenbach M. 1994.** The sporulation-specific enzymes encoded by the DIT1 and DIT2 genes catalyze a two-step reaction leading to a soluble LL-dityrosine-containing precursor of the yeast spore wall.  
*Proc. Natl. Acad. Sci. U. S. A* 91, 4524-4528.
- Bromley E, Leeds N, Clark J, McGregor E, Ward M, Dunn MJ, Tomley F. 2003.** Defining the protein repertoire of microneme secretory organelles in the apicomplexan parasite *Eimeria tenella*.  
*Proteomics*. 3, 1553-1561.
- Brosson D, Kuhn L, Delbac F, Vivarès CP, Texier C.** Proteomic analysis of the eukaryotic parasite *Encephalitozoon cuniculi* (Microsporidia): a reference map for proteins expressed in late sporogonial stages.  
*Proteomics. Manuscrit soumis*.
- Brosson D, Kuhn L, Prensier G, Vivarès CP, Texier C. 2005.** The putative chitin deacetylase of *Encephalitozoon cuniculi*: a surface protein implicated in microsporidian spore-wall formation.  
*FEMS Microbiol. Lett.* 247, 81-90.
- Bruneau JM, Maillet I, Tagat E, Legrand R, Supatto F, Fudali C, Caer JP, Labas V, Lecaque D, Hodgson J. 2003.** Drug induced proteome changes in *Candida albicans*: comparison of the effect of beta(1,3) glucan synthase inhibitors and two triazoles, fluconazole and itraconazole.  
*Proteomics*. 3, 325-336.
- Bryan RT, Weber R, Schwartz DA. 1997.** Microsporidiosis in patients who are not infected with human immunodeficiency virus.  
*Clin. Infect. Dis.* 24, 534-535.
- Cabib E, Drgonova J, Drgon T. 1998.** Role of small G proteins in yeast cell polarization and wall biosynthesis.  
*Annu. Rev. Biochem.* 67, 307-333.
- Canning EU, Lom J. 1986.** *The Microsporidia of Vertebrates*. Academic Press, New York. 1-289.
- Canning EU. 1993.** Microsporidia. In : *Parasite Protozoa*, edited by Kreier J.P. p.299-370.
- Caro LH, Tettelin H, Vossen JH, Ram AF, Van Den EH, Klis FM. 1997.** In silico identification of glycosyl-phosphatidylinositol-anchored plasma-membrane and cell wall proteins of *Saccharomyces cerevisiae*.  
*Yeast* 13, 1477-1489.
- Cassone A. 1989.** Cell wall of *Candida albicans*: its functions and its impact on the host.  
*Curr. Top. Med. Mycol.* 3, 248-314.



- Castillo L, Martinez AI, Garcera A, Elorza MV, Valentin E, Sentandreu R. 2003.** Functional analysis of the cysteine residues and the repetitive sequence of *Saccharomyces cerevisiae* Pir4/Cis3: the repetitive sequence is needed for binding to the cell wall beta-1,3-glucan. *Yeast* 20, 973-983.
- Causier B. 2004.** Studying the interactome with the yeast two-hybrid system and mass spectrometry. *Mass Spectrom. Rev.* 23, 350-367.
- Cavalier-Smith T. 1998.** A revised six-kingdom system of life. *Biol. Rev. Camb. Philos. Soc.* 73, 203-266.
- Chapman JR. 1996.** Mass spectrometry. Ionization methods and instrumentation. *Methods Mol. Biol* 61, 9-28.
- Chaurand P, Schwartz SA, Caprioli RM. 2004.** Assessing protein patterns in disease using imaging mass spectrometry. *J Proteome. Res.* 3, 245-252.
- Chevallet M, Santoni V, Poinas A, Rouquie D, Fuchs A, Kieffer S, Rossignol M, Lunardi J, Garin J, Rabilloud T. 1998.** New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. *Electrophoresis* 19, 1901-1909.
- Choi W, Yoo YJ, Kim M, Shin D, Jeon HB, Choi W. 2003.** Identification of proteins highly expressed in the hyphae of *Candida albicans* by two-dimensional electrophoresis. *Yeast* 20, 1053-1060.
- Christodoulidou A, Bouriotis V, Thireos G. 1996.** Two sporulation-specific chitin deacetylase-encoding genes are required for the ascospore wall rigidity of *Saccharomyces cerevisiae*. *J Biol Chem* 271, 31420-31425.
- Christodoulidou A, Briza P, Ellinger A, Bouriotis V. 1999.** Yeast ascospore wall assembly requires two chitin deacetylase isozymes. *FEBS Lett.* 460, 275-279.
- Chuang JS, Schekman RW. 1996.** Differential trafficking and timed localization of two chitin synthase proteins, Chs2p and Chs3p. *J Cell Biol* 135, 597-610.
- Cid VJ, Duran A, del Rey F, Snyder MP, Nombela C, Sanchez M. 1995.** Molecular basis of cell integrity and morphogenesis in *Saccharomyces cerevisiae*. *Microbiol. Rev.* 59, 345-386.
- Cleves AE, Cooper DN, Barondes SH, Kelly RB. 1996.** A new pathway for protein export in *Saccharomyces cerevisiae*. *J. Cell Biol.* 133, 1017-1026.
- Cohen AM, Rumpel K, Coombs GH, Wastling JM. 2002.** Characterisation of global protein expression by two-dimensional electrophoresis and mass spectrometry: proteomics of *Toxoplasma gondii*. *Int. J. Parasitol.* 32, 39-51.



- Coluccio A, Bogengruber E, Conrad MN, Dresser ME, Briza P, Neiman AM. 2004.** Morphogenetic pathway of spore wall assembly in *Saccharomyces cerevisiae*. *Eukaryot. Cell* 3, 1464-1475.
- Coluccio A, Neiman AM. 2004.** Interspore bridges: a new feature of the *Saccharomyces cerevisiae* spore wall. *Microbiology* 150, 3189-3196.
- Cooper RA, Carucci DJ. 2004.** Proteomic approaches to studying drug targets and resistance in *Plasmodium*. *Curr. Drug Targets. Infect. Disord.* 4, 41-51.
- Cormack BP, Ghori N, Falkow S. 1999.** An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells. *Science* 285, 578-582.
- Cornillot E, Méténier G, Vivarès CP, Dassa E. 2002.** Comparative analysis of sequences encoding ABC systems in the genome of the microsporidian *Encephalitozoon cuniculi*. *FEMS Microbiol. Lett.* 210, 39-47.
- Couzinet S, Cejas E, Schittny J, Deplazes P, Weber R, Zimmerli S. 2000.** Phagocytic uptake of *Encephalitozoon cuniculi* by nonprofessional phagocytes. *Infect. Immun.* 68, 6939-6945.
- Curgy J, Vavra J, Vivarès CP. 1980.** Presence of ribosomal ARN with prokaryotic properties in Microsporidia, eukaryotic organisms. *Biol Cell* 38, 49-52.
- Dall DJ. 1983.** A theory for the mechanism of polar filament extrusion in the Microspora. *J Theor. Biol.* 105, 647-659.
- Dallies N, Francois J, Paquet V. 1998.** A new method for quantitative determination of polysaccharides in the yeast cell wall. Application to the cell wall defective mutants of *Saccharomyces cerevisiae*. *Yeast* 14, 1297-1306.
- de Groot PW et al. 2001.** Genomic approach for the identification and classification of genes involved in cell wall formation and its regulation in *Saccharomyces cerevisiae*. *Comparative and Functional Genomics* 2, 124-142.
- De Groot PW, de Boer AD, Cunningham J, Dekker HL, de Jong L, Hellingwerf KJ, de Koster C, Klis FM. 2004.** Proteomic analysis of *Candida albicans* cell walls reveals covalently bound carbohydrate-active enzymes and adhesins. *Eukaryot. Cell* 3, 955-965.
- Del Aguila C, Rueda C, De la CC, Fenoy S. 2001.** Seroprevalence of anti-*Encephalitozoon* antibodies in Spanish immunocompetent subjects. *J Eukaryot. Microbiol. Suppl.* 75S-78S.
- Delbac F, Duffieux F, David D, Méténier G, Vivarès CP. 1998a.** Immunocytochemical identification of spore proteins in two microsporidia, with emphasis on extrusion apparatus. *J. Eukaryot. Microbiol.* 45, 224-231.





- Delbac F, Peyret P, Méténier G, David D, Danchin A, Vivarès CP. 1998b.** On proteins of the microsporidian invasive apparatus: complete sequence of a polar tube protein of *Encephalitozoon cuniculi*. *Mol. Microbiol.* 29, 825-834.
- Delbac F, Peuvél I, Méténier G, Peyretailade E, Vivarès CP. 2001.** Microsporidian invasion apparatus: identification of a novel polar tube protein and evidence for clustering of ptp1 and ptp2 genes in three *Encephalitozoon* species. *Infect. Immun.* 69, 1016-1024.
- Deshaies RJ, Seol JH, McDonald WH, Cope G, Lyapina S, Shevchenko A, Shevchenko A, Verma R, Yates JR. 2002.** Charting the protein complexome in yeast by mass spectrometry. *Mol. Cell Proteomics.* 1, 3-10.
- Desportes-Livage I. 1996.** Human microsporidiosis and AIDS: recent advances. *Parasite* 3, 107-113.
- de Venevelles P, Chich JF, Faigle W, Loew D, Labbe M, Girard-Misguich F, Pery P. 2004.** Towards a reference map of *Eimeria tenella* sporozoite proteins by two-dimensional electrophoresis and mass spectrometry. *Int. J Parasitol.* 34, 1321-1331.
- Didier ES. 2005.** Microsporidiosis: an emerging and opportunistic infection in humans and animals. *Acta Trop.* 94, 61-76.
- Dlugonska H, Dytynska K, Reichmann G, Stachelhaus S, Fischer HG. 2001.** Towards the *Toxoplasma gondii* proteome: position of 13 parasite excretory antigens on a standardized map of two-dimensionally separated tachyzoite proteins. *Parasitol. Res.* 87, 634-637.
- Domer JE. 1989.** *Candida* cell wall mannan: a polysaccharide with diverse immunologic properties. *Crit Rev. Microbiol.* 17, 33-51.
- Doolan DL, Southwood S, Freilich DA, Sidney J, Graber NL, Shatney L, Bebris L, Florens L, Dobano C, Witney AA, Appella E, Hoffman SL, Yates JR, Carucci DJ, Sette A. 2003.** Identification of *Plasmodium falciparum* antigens by antigenic analysis of genomic and proteomic data. *Proc. Natl. Acad. Sci. U. S. A* 100, 9952-9957.
- Drummelsmith J, Brochu V, Girard I, Messier N, Ouellette M. 2003.** Proteome mapping of the protozoan parasite *Leishmania* and application to the study of drug targets and resistance mechanisms. *Mol. Cell Proteomics.* 2, 146-155.
- Drummelsmith J, Girard I, Trudel N, Ouellette M. 2004.** Differential protein expression analysis of *Leishmania major* reveals novel roles for methionine adenosyltransferase and S-adenosylmethionine in methotrexate resistance. *J. Biol. Chem.* 279, 33273-33280.



- Dunn AM, Smith JE. 2001.** Microsporidian life cycles and diversity: the relationship between virulence and transmission.  
*Microbes. Infect.* 3, 381-388.
- Edman P. 1949.** A method for the determination of the amino acid sequence of peptides.  
*Archives of Biochemistry and Biophysics* 22, 475-483.
- El Fakhry Y, Ouellette M, Papadopoulou B. 2002.** A proteomic approach to identify developmentally regulated proteins in *Leishmania infantum*.  
*Proteomics.* 2, 1007-1017.
- El Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberger SJ, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E, Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de Jong P, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasch AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, Mizuno Y, Mottram JC, Nelson S, Ochaya S, Osoegawa K, Pai G, Parsons M, Pentony M, Pettersson U, Pop M, Ramirez JL, Rinta J, Robertson L, Salzberg SL, Sanchez DO, Seyler A, Sharma R, Shetty J, Simpson AJ, Sisk E, Tammi MT, Tarleton R, Teixeira S, Van Aken S, Vogt C, Ward PN, Wickstead B, Wortman J, White O, Fraser CM, Stuart KD, Andersson B. 2005.** The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease.  
*Science* 309, 409-415.
- Enriquez FJ, Taren D, Cruz-Lopez A, Muramoto M, Palting JD, Cruz P. 1998.** Prevalence of intestinal encephalitozoonosis in Mexico.  
*Clin. Infect. Dis.* 26, 1227-1229.
- Erickson BW, Blanquet RS. 1969.** The occurrence of chitin in the spore wall of *Glugea weissenbergi*.  
*J Invertebr Pathol* 14, 358-364.
- Escott GM, Adams DJ. 1995.** Chitinase activity in human serum and leukocytes.  
*Infect. Immun.* 63, 4770-4773.
- Fabrega C, Hausmann S, Shen V, Shuman S, Lima CD. 2004.** Structure and mechanism of mRNA cap (guanine-N7) methyltransferase.  
*Mol. Cell* 13, 77-89.
- Fasshauer V, Gross U, Bohne W. 2005.** The parasitophorous vacuole membrane of *Encephalitozoon cuniculi* lacks host cell membrane proteins immediately after invasion.  
*Eukaryot. Cell* 4, 221-224.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. 1989.** Electrospray ionization for mass spectrometry of large biomolecules.  
*Science* 246, 64-71.
- Fernandez-Arenas E, Molero G, Nombela C, Diez-Orejas R, Gil C. 2004.** Contribution of the antibodies response induced by a low virulent *Candida albicans* strain in protection against systemic candidiasis.  
*Proteomics.* 4, 1204-1215.



- Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, Muster N, Sacci JB, Tabb DL, Witney AA, Wolters D, Wu Y, Gardner MJ, Holder AA, Sinden RE, Yates JR, Carucci DJ. 2002.** A proteomic view of the *Plasmodium falciparum* life cycle.  
*Nature* 419, 520-526.
- Florens L, Liu X, Wang Y, Yang S, Schwartz O, Peglar M, Carucci DJ, Yates JR, Wub Y. 2004.** Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes.  
*Mol. Biochem. Parasitol.* 135, 1-11.
- Foucault C, Drancourt M. 2000.** Actin mediates *Encephalitozoon intestinalis* entry into the human enterocyte-like cell line, Caco-2.  
*Microb. Pathog.* 28, 51-58.
- Fradin C, Kretschmar M, Nichterlein T, Gaillardin C, d'Enfert C, Hube B. 2003.** Stage-specific gene expression of *Candida albicans* in human blood.  
*Mol. Microbiol.* 47, 1523-1543.
- Franzen C, Muller A. 2001.** Microsporidiosis: human diseases and diagnosis.  
*Microbes. Infect.* 3, 389-400.
- Frevort U, Sinnis P, Cerami C, Shreffler W, Takacs B, Nussenzweig V. 1993.** Malaria circumsporozoite protein binds to heparan sulfate proteoglycans associated with the surface membrane of hepatocytes.  
*J Exp. Med.* 177, 1287-1298.
- Frieman MB, Cormack BP. 2004.** Multiple sequence signals determine the distribution of glycosylphosphatidylinositol proteins between the plasma membrane and cell wall in *Saccharomyces cerevisiae*.  
*Microbiology* 150, 3105-3114.
- Frixione E, Ruiz L, Cerbon J, Undeen AH. 1997.** Germination of *Nosema algerae* (Microspora) spores: conditional inhibition by D<sub>2</sub>O, ethanol and Hg<sup>2+</sup> suggests dependence of water influx upon membrane hydration and specific transmembrane pathways.  
*J. Eukaryot. Microbiol.* 44, 109-116.
- Frixione E, Ruiz L, Santillan M, de Vargas LV, Tejero JM, Undeen AH. 1992.** Dynamics of polar filament discharge and sporoplasm expulsion by microsporidian spores.  
*Cell Motil. Cytoskelet* 22, 38-50.
- Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Perteau M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. 2002.** Genome sequence of the human malaria parasite *Plasmodium falciparum*.  
*Nature* 419, 498-511.
- Gardner MJ, Bishop R, Shah T, de Villiers EP, Carlton JM, Hall N, Ren Q, Paulsen IT, Pain A, Berriman M, Wilson RJ, Sato S, Ralph SA, Mann DJ, Xiong Z, Shallom SJ, Weidman J, Jiang L, Lynn J, Weaver B, Shoaibi A, Domingo AR, Wasawo D,**



- Crabtree J, Wortman JR, Haas B, Angiuoli SV, Creasy TH, Lu C, Suh B, Silva JC, Utterback TR, Feldblyum TV, Pertea M, Allen J, Nierman WC, Taracha EL, Salzberg SL, White OR, Fitzhugh HA, Morzaria S, Venter JC, Fraser CM, Nene V. 2005.** Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* 309, 134-137.
- Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edlmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G. 2002.** Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141-147.
- Geissler S, Sokolowska-Kohler W, Bollmann R, Jungblut PR, Presber W. 1999.** *Toxoplasma gondii* infection: analysis of serological response by 2-DE immunoblotting. *FEMS Immunol. Med. Microbiol.* 25, 299-311.
- Germot A, Philippe H, Le Guyader H. 1997.** Evidence for loss of mitochondria in Microsporidia from a mitochondrial-type HSP70 in *Nosema locustae*. *Mol. Biochem. Parasitol.* 87, 159-168.
- Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG. 1996.** Life with 6000 genes. *Science* 274, 546, 563-546, 567.
- Goldman RC, Sullivan PA, Zakula D, Capobianco JO. 1995.** Kinetics of beta-1,3 glucan interaction at the donor and acceptor sites of the fungal glucosyltransferase encoded by the BGL2 gene. *Eur. J Biochem.* 227, 372-378.
- Good MF, Doolan DL. 1999.** Immune effector mechanisms in malaria. *Curr. Opin. Immunol.* 11, 412-419.
- Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W. 2000.** The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21, 1037-1053.
- Gorg A, Weiss W, Dunn MJ. 2004.** Current two-dimensional electrophoresis technology for proteomics. *Proteomics.* 4, 3665-3685.
- Griffin TJ, Gygi SP, Ideker T, Rist B, Eng J, Hood L, Aebersold R. 2002.** Complementary profiling of gene expression at the transcriptome and proteome levels in *Saccharomyces cerevisiae*. *Mol. Cell Proteomics.* 1, 323-333.
- Gunasekera AM, Patankar S, Schug J, Eisen G, Wirth DF. 2003.** Drug-induced alterations in gene expression of the asexual blood forms of *Plasmodium falciparum*. *Mol. Microbiol.* 50, 1229-1239.





- Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R. 2000.** Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology.  
*Proc. Natl. Acad. Sci. U. S. A* 97, 9390-9395.
- Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, Florens L, Janssen CS, Pain A, Christophides GK, James K, Rutherford K, Harris B, Harris D, Churcher C, Quail MA, Ormond D, Doggett J, Trueman HE, Mendoza J, Bidwell SL, Rajandream MA, Carucci DJ, Yates JR, III, Kafatos FC, Janse CJ, Barrell B, Turner CM, Waters AP, Sinden RE. 2005.** A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses.  
*Science* 307, 82-86.
- Hausmann S, Vivarès CP, Shuman S. 2002.** Characterization of the mRNA capping apparatus of the microsporidian parasite *Encephalitozoon cuniculi*.  
*J. Biol. Chem.* 277, 96-103.
- Hayman JR, Hayes SF, Amon J, Nash TE. 2001.** Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*.  
*Infect. Immun.* 69, 7057-7066.
- Hayman JR, Southern TR, Nash TE. 2005.** Role of sulfated glycans in adherence of the microsporidian *Encephalitozoon intestinalis* to host cells in vitro.  
*Infect. Immun.* 73, 841-848.
- Hazard EI, Brookbank JW. 1984.** Karyogamy and meiosis in an *Amblyospora* sp. (Microspora) in the mosquito *Culex salinarius*.  
*Journal of Invertebrate Pathology* 44, 3-11.
- Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C. 1993.** Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases.  
*Proc. Natl. Acad. Sci. U. S. A* 90, 5011-5015.
- Hernandez R, Nombela C, Diez-Orejas R, Gil C. 2004.** Two-dimensional reference map of *Candida albicans* hyphal forms.  
*Proteomics.* 4, 374-382.
- Hirt RP, Healy B, Vossbrinck CR, Canning EU, Embley TM. 1997.** A mitochondrial Hsp70 orthologue in *Vairimorpha necatrix*: molecular evidence that microsporidia once contained mitochondria.  
*Curr. Biol.* 7, 995-998.
- Holan Z, Pokorny V, Beran K, Gemperle A, Tuzar Z, Baldrian J. 1981.** The glucan-chitin complex in *Saccharomyces cerevisiae*. Precise localisation of chitin and glucan in bud scar and their physico-chemical characterisation.  
*Arch Microbiol.* 130, 312-318.
- Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, and O'Shea EK. 2003.** Global analysis of protein localization in budding yeast.  
*Nature* 425, 686-691.
- Ignatius R, Henschel S, Liesenfeld O, Mansmann U, Schmidt W, Koppe S, Schneider T, Heise W, Futh U, Riecken EO, Hahn H, Ullrich R. 1997.** Comparative evaluation of modified



trichrome and Uvitex 2B stains for detection of low numbers of microsporidial spores in stool specimens.

*J. Clin. Microbiol.* 35, 2266-2269.

- Ishihara R, Hayashin Y. 1968.** Some properties of ribosomes from the sporoplasm of *N. bonbycis*. *J Invertebr Pathol* 11, 377-385.
- Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchetti G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De Gaudenzi J, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasc AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Muller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'Neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabbinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schafer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart KD, Barrell B, Myler PJ. 2005.** The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442.
- James P, Quadroni M, Carafoli E, Gonnet G. 1993.** Protein identification by mass profile fingerprinting. *Biochem. Biophys. Res. Commun.* 195, 58-64.
- Johnson JR, Florens L, Carucci DJ, Yates JR. 2004.** Proteomics in malaria. *J Proteome. Res.* 3, 296-306.
- Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, Newport G, Thorstenson YR, Agabian N, Magee PT, Davis RW, Scherer S. 2004.** The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. U. S. A* 101, 7329-7334.
- Jong AY, Chen SH, Stins MF, Kim KS, Tuan TL, Huang SH. 2003.** Binding of *Candida albicans* enolase to plasmin(ogen) results in enhanced invasion of human brain microvascular endothelial cells. *J Med. Microbiol.* 52, 615-622.
- Jung JW, Lee W. 2004.** Structure-based functional discovery of proteins: structural proteomics. *J Biochem. Mol. Biol* 37, 28-34.
- Kafetzopoulos D, Martinou A, Bouriotis V. 1993.** Bioconversion of chitin to chitosan: purification and characterization of chitin deacetylase from *Mucor rouxii*. *Proc. Natl. Acad. Sci. U. S. A* 90, 2564-2568.
- Kamaishi T, Hashimoto T, Nakamura Y, Nakamura F, Murata S, Okada N, Okamoto K, Shimizu M, Hasegawa M. 1996.** Protein phylogeny of translation elongation factor EF-1 alpha suggests microsporidians are extremely ancient eukaryotes. *J. Mol. Evol.* 42, 257-263.



- Kapteyn JC, Hoyer LL, Hecht JE, Muller WH, Andel A, Verkleij AJ, Makarow M, Van Den EH, Klis FM. 2000.** The cell wall architecture of *Candida albicans* wild-type cells and cell wall-defective mutants.  
*Mol. Microbiol.* 35, 601-611.
- Karas M. 1996.** Matrix-assisted laser desorption ionization MS: a progress report.  
*Biochem. Soc. Trans.* 24, 897-900.
- Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prensier G, Barbe V, Peyretailade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivarès CP. 2001.** Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*.  
*Nature* 414, 450-453.
- Keeling PJ, Fast NM. 2002.** Microsporidia: biology and evolution of highly reduced intracellular parasites.  
*Annu. Rev. Microbiol.* 56, 93-116.
- Kenrick KG, Margolis J. 1970.** Isoelectric focusing and gradient gel electrophoresis: a two-dimensional technique.  
*Anal. Biochem.* 33, 204-207.
- Khan IA, Moretto M, Weiss LM. 2001.** Immune response to *Encephalitozoon cuniculi* infection.  
*Microbes. Infect.* 3, 401-405.
- Klis FM, de Groot P, Hellingwerf K. 2001.** Molecular organization of the cell wall of *Candida albicans*.  
*Med. Mycol.* 39 Suppl 1, 1-8.
- Klis FM, Mol P, Hellingwerf K, Brul S. 2002.** Dynamics of cell wall structure in *Saccharomyces cerevisiae*.  
*FEMS Microbiol. Rev.* 26, 239-256.
- Klose J. 1975.** Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals.  
*Humangenetik.* 26, 231-243.
- Kolkman A, Olsthoorn MM, Heeremans CE, Heck AJ, Slijper M. 2005.** Comparative proteome analysis of *Saccharomyces cerevisiae* grown in chemostat cultures limited for glucose or ethanol.  
*Mol. Cell Proteomics.* 4, 1-11.
- Kollar R, Reinhold BB, Petrakova E, Yeh HJ, Ashwell G, Drgonova J, Kapteyn JC, Klis FM, Cabib E. 1997.** Architecture of the yeast cell wall. Beta(1-->6)-glucan interconnects mannoprotein, beta(1-->3)-glucan, and chitin.  
*J Biol Chem* 272, 17762-17775.
- Koudela B, Kucerova S, Hudcovic T. 1999.** Effect of low and high temperatures on infectivity of *Encephalitozoon cuniculi* spores suspended in water.  
*Folia Parasitol. (Praha)* 46, 171-174.



- Krainer E, Stark RE, Naider F, Alagramam K, Becker JM. 1994.** Direct observation of cell wall glucans in whole cells of *Saccharomyces cerevisiae* by magic-angle spinning <sup>13</sup>C-NMR.  
*Biopolymers* 34, 1627-1635.
- Kreger-Van Rij NJ. 1978.** Electron microscopy of germinating ascospores of *Saccharomyces cerevisiae*.  
*Arch Microbiol.* 117, 73-77.
- Kucerova-Pospisilova Z, Ditrich O. 1998.** The serological surveillance of several groups of patients using antigens of *Encephalitozoon hellem* and *E. cuniculi* antibodies to microsporidia in patients.  
*Folia Parasitol. (Praha)* 45, 108-112.
- Kucerova-Pospisilova Z, Secor WE, Moura H, Desportes-Livage I, Datry A, Bern C, Leitch G, Visvesvara GS. 2001.** An ELISA test to detect human serum antibodies reactive with *Encephalitozoon intestinalis*.  
*J Eukaryot. Microbiol. Suppl.* 73S-74S.
- Kupiec M, Byers RE, Esposito RE, Mitchell AP. 1997.** Meiosis and sporulation in *Saccharomyces cerevisiae*. In : *The molecular biology of the yeast Saccharomyces*, edited by J. Pringle, J. R. Broach, and E. W. Jones p.889-1036.
- Lamanda A, Zahn A, Roder D, Langen H. 2004.** Improved Ruthenium II tris (bathophenanthroline disulfonate) staining and destaining protocol for a better signal-to-background ratio and improved baseline resolution.  
*Proteomics.* 4, 599-608.
- Langley RC, Cali A, Somberg EW. 1987.** Two-dimensional electrophoretic analysis of spore proteins of the Microsporida.  
*J. Parasitol.* 73, 910-918.
- Larsson R. 1986.** Ultrastructure, function and classification of microsporidia. In : *Progress in Prostistology*, edited by J. O. Corliss and D. J. Patterson (Biopress, Ltd, Bristol), p.325-390.
- Lasonder E, Ishihama Y, Andersen JS, Vermunt AM, Pain A, Sauerwein RW, Eling WM, Hall N, Waters AP, Stunnenberg HG, Mann M. 2002.** Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry.  
*Nature* 419, 537-542.
- Lee EG, Kim JH, Shin YS, Shin GW, Suh MD, Kim DY, Kim YH, Kim GS, Jung TS. 2003.** Establishment of a two-dimensional electrophoresis map for *Neospora caninum* tachyzoites by proteomics.  
*Proteomics.* 3, 2339-2350.
- Leulliot N, Tresaugues L, Bremang M, Sorel I, Ulryck N, Graille M, Aboufath I, Poupon A, Liger D, Quevillon-Cheruel S, Janin J, van Tilbeurgh H. 2005.** High-throughput crystal-optimization strategies in the South Paris Yeast Structural Genomics Project: one size fits all ?  
*Acta Crystallogr. D. Biol Crystallogr.* 61, 664-670.





- Levitz SM, Nong S, Mansour MK, Huang C, Specht CA. 2001.** Molecular characterization of a mannoprotein with homology to chitin deacetylases that stimulates T cell responses to *Cryptococcus neoformans*.  
*Proc. Natl. Acad. Sci. U. S. A* 98, 10422-10427.
- Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR. 1999.** Direct analysis of protein complexes using mass spectrometry.  
*Nat. Biotechnol.* 17, 676-682.
- Liu HL, Hsu JP. 2005.** Recent developments in structural proteomics for protein structure determination.  
*Proteomics.* 5, 2056-2068.
- Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, Pop M, Duchene M, Ackers J, Tannich E, Leippe M, Hofer M, Bruchhaus I, Willhoeft U, Bhattacharya A, Chillingworth T, Churcher C, Hance Z, Harris B, Harris D, Jagels K, Moule S, Mungall K, Ormond D, Squares R, Whitehead S, Quail MA, Rabbinowitsch E, Norbertczak H, Price C, Wang Z, Guillen N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, El Sayed NM, Petri WA, Clark CG, Embley TM, Barrell B, Fraser CM, Hall N. 2005.** The genome of the protist parasite *Entamoeba histolytica*.  
*Nature* 433, 865-868.
- Lom J, Nilsen F. 2003.** Fish microsporidia: fine structural diversity and phylogeny.  
*Int. J Parasitol.* 33, 107-127.
- Lores B, Lopez-Miragaya I, Arias C, Fenoy S, Torres J, Del Aguila C. 2002.** Intestinal microsporidiosis due to *Enterocytozoon bieneusi* in elderly human immunodeficiency virus-negative patients from Vigo, Spain.  
*Clin. Infect. Dis.* 34, 918-921.
- Lu CF, Montijn RC, Brown JL, Klis F, Kurjan J, Bussey H, Lipke PN. 1995.** Glycosyl phosphatidylinositol-dependent cross-linking of alpha-agglutinin and beta 1,6-glucan in the *Saccharomyces cerevisiae* cell wall.  
*J Cell Biol* 128, 333-340.
- Luche S, Santoni V, Rabilloud T. 2003.** Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis.  
*Proteomics.* 3, 249-253.
- Lujan HD, Conrad JT, Clark CG, Touz MC, Delbac F, Vivarès CP, Nash TE. 1998.** Detection of microsporidia spore-specific antigens by monoclonal antibodies.  
*Hybridoma* 17, 237-243.
- Mann M, Hojrup P, Roepstorff P. 1993.** Use of mass spectrometric molecular weight information to identify proteins in sequence databases.  
*Biol Mass Spectrom.* 22, 338-345.
- Matsubayashi H, Koike T, Mikata T, Hagiwara S. 1959.** A case of *Encephalitozoon*-like body infection in man.  
*Arch Pathol* 67, 181-187.



- Molano J, Bowers B, Cabib E. 1980.** Distribution of chitin in the yeast cell wall. An ultrastructural and chemical study.  
*J Cell Biol* 85, 199-212.
- Molloy MP. 2000.** Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients.  
*Anal. Biochem.* 280, 1-10.
- Montijn RC, Vink E, Muller WH, Verkleij AJ, Van Den EH, Henrissat B, Klis FM. 1999.** Localization of synthesis of beta1,6-glucan in *Saccharomyces cerevisiae*.  
*J Bacteriol.* 181, 7414-7420.
- Moor H, Muhlethaler K. 1963.** Fine structure in frozen-etched yeast cells.  
*J. Cell Biol.* 17, 609-628.
- Morris HR, Paxton T, Dell A, Langhorne J, Berg M, Bordoli RS, Hoyes J, Bateman RH. 1996.** High sensitivity collisionally-activated decomposition tandem mass spectrometry on a novel quadrupole/orthogonal-acceleration time-of-flight mass spectrometer.  
*Rapid Commun. Mass Spectrom.* 10, 889-896.
- Moura H, Visvesvara GS. 2001.** A proteome approach to the host-parasite interaction of the microsporidian *Encephalitozoon intestinalis*.  
*J. Eukaryot. Microbiol. Suppl*, 56S-59S.
- Moura H, Ospina M, Woolfitt AR, Barr JR, Visvesvara GS. 2003.** Analysis of four human microsporidian isolates by MALDI-TOF mass spectrometry.  
*J. Eukaryot. Microbiol.* 50, 156-163.
- Mrsa V, Seidl T, Gentzsch M, Tanner W. 1997.** Specific labelling of cell wall proteins by biotinylation. Identification of four covalently linked O-mannosylated proteins of *Saccharomyces cerevisiae*.  
*Yeast* 13, 1145-1154.
- Muller A, Ensley H, Pretus H, McNamee R, Jones E, McLaughlin E, Chandley W, Browder W, Lowman D, Williams D. 1997.** The application of various protic acids in the extraction of (1->3)-beta-D-glucan from *Saccharomyces cerevisiae*.  
*Carbohydr. Res.* 299, 203-208.
- Muller A, Bialek R, Kamper A, Fatkenheuer G, Salzberger B, Franzen C. 2001.** Detection of microsporidia in travelers with diarrhea.  
*J. Clin. Microbiol.* 39, 1630-1632.
- Nabieva E, Jim K, Agarwal A, Chazelle B, Singh M. 2005.** Whole-proteome prediction of protein function via graph-theoretic analysis of interaction maps.  
*Bioinformatics.* 21 Suppl 1, i302-i310.
- Nakai K, Horton P. 1999.** PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization.  
*Trends Biochem. Sci.* 24, 34-36.
- Neiman AM. 1998.** Prospore membrane formation defines a developmentally regulated branch of the secretory pathway in yeast.  
*J Cell Biol* 140, 29-37.



- Nielsen H, Engelbrecht J, Brunak S, von Heijne G. 1997.** Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.  
*Protein Eng* 10, 1-6.
- Nimrichter L, Rodrigues ML, Rodrigues EG, Travassos LR. 2005.** The multitude of targets for the immune system and drug therapy in the fungal cell wall.  
*Microbes. Infect.* 7, 789-798.
- Nischik N, Schade B, Dytnerka K, Dlugonska H, Reichmann G, Fischer HG. 2001.** Attenuation of mouse-virulent *Toxoplasma gondii* parasites is associated with a decrease in interleukin-12-inducing tachyzoite activity and reduced expression of actin, catalase and excretory proteins.  
*Microbes. Infect.* 3, 689-699.
- Nugent PG, Karsani SA, Wait R, Tempero J, Smith DF. 2004.** Proteomic analysis of *Leishmania mexicana* differentiation.  
*Mol. Biochem. Parasitol.* 136, 51-62.
- O'Farrell PH. 1975.** High resolution two-dimensional electrophoresis of proteins.  
*J Biol Chem* 250, 4007-4021.
- Okada M, Huston CD, Mann BJ, Petri WA, Kita K, Nozaki T. 2005.** Proteomic analysis of phagocytosis in the enteric protozoan parasite *Entamoeba histolytica*.  
*Eukaryot. Cell* 4, 827-831.
- Orlean P. 1997.** Biogenesis of yeast wall and surface components. In : *The molecular biology of the yeast saccharomyces*, edited by J. R. Pringle, J. R. Broach, and E. W. Jones (Cold Spring Harbor Laboratory, New York), p.229-362.
- Oshiro G, Wodicka LM, Washburn MP, Yates JR, Lockhart DJ, Winzeler EA. 2002.** Parallel identification of new genes in *Saccharomyces cerevisiae*.  
*Genome Res.* 12, 1210-1220.
- Osumi M. 1998.** The ultrastructure of yeast: cell wall structure and formation.  
*Micron.* 29, 207-233.
- Overdijk B, Van Steijn GJ, Odds FC. 1999.** Distribution of chitinase in guinea pig tissues and increases in levels of this enzyme after systemic infection with *Aspergillus fumigatus*.  
*Microbiology* 145 ( Pt 1), 259-269.
- Pain A, Renauld H, Berriman M, Murphy L, Yeats CA, Weir W, Kerhornou A, Aslett M, Bishop R, Bouchier C, Cochet M, Coulson RM, Cronin A, de Villiers EP, Fraser A, Fosker N, Gardner M, Goble A, Griffiths-Jones S, Harris DE, Katzer F, Larke N, Lord A, Maser P, McKellar S, Mooney P, Morton F, Nene V, O'Neil S, Price C, Quail MA, Rabinowitsch E, Rawlings ND, Rutter S, Saunders D, Seeger K, Shah T, Squares R, Squares S, Tivey A, Walker AR, Woodward J, Dobbelaere DA, Langsley G, Rajandream MA, McKeever D, Shiels B, Tait A, Barrell B, Hall N. 2005.** Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*.  
*Science* 309, 131-133.
- Pakes SP, Shadduck JA, Cali A. 1975.** Fine structure of *Encephalitozoon cuniculi* from rabbits, mice and hamsters.  
*J. Protozool.* 22, 481-488.



- Palma C, Serbousek D, Torosantucci A, Cassone A, Djeu JY. 1992.** Identification of a mannoprotein fraction from *Candida albicans* that enhances human polymorphonuclear leukocyte (PMNL) functions and stimulates lactoferrin in PMNL inhibition of candidal growth.  
*J Infect. Dis.* 166, 1103-1112.
- Pappin DJ, Hojrup P, Bleasby AJ. 1993.** Rapid identification of proteins by peptide-mass fingerprinting.  
*Curr. Biol* 3, 327-332.
- Pardo M, Monteoliva L, Pla J, Sanchez M, Gil C, Nombela C. 1999.** Two-dimensional analysis of proteins secreted by *Saccharomyces cerevisiae* regenerating protoplasts: a novel approach to study the cell wall.  
*Yeast* 15, 459-472.
- Parodi-Talice A, Duran R, Arrambide N, Prieto V, Pineyro MD, Pritsch O, Cayota A, Cervenansky C, Robello C. 2004.** Proteome analysis of the causative agent of Chagas disease: *Trypanosoma cruzi*.  
*Int. J. Parasitol.* 34, 881-886.
- Patterson SD, Latter GI. 1993.** Evaluation of storage phosphor imaging for quantitative analysis of 2-D gels using the Quest II system.  
*Biotechniques* 15, 1076-1083.
- Payne WE, Garrels JI. 1997.** Yeast Protein database (YPD): a database for the complete proteome of *Saccharomyces cerevisiae*.  
*Nucleic Acids Res.* 25, 57-62.
- Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. 2003.** Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.  
*J Proteome. Res.* 2, 43-50.
- Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999.** Probability-based protein identification by searching sequence databases using mass spectrometry data.  
*Electrophoresis* 20, 3551-3567.
- Peuvel I, Peyret P, Méténier G, Vivarès CP, Delbac F. 2002.** The microsporidian polar tube: evidence for a third polar tube protein (PTP3) in *Encephalitozoon cuniculi*.  
*Mol. Biochem. Parasitol.* 122, 69-80.
- Peuvel-Fanget I, Polonais V, Brosseau D, Texier C, Kuhn L, Vivarès C, Delbac F.** EnP1 and EnP2, two proteins associated with the *Encephalitozoon cuniculi* endospore, the chitin-rich inner layer of the microsporidian spore wall.  
*Int. J Parasitol. sous presse.*
- Peyretailade E, Broussolle V, Peyret P, Méténier G, Gouy M, Vivarès CP. 1998.** Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin.  
*Mol. Biol. Evol.* 15, 683-689.





- Pitarch A, Sanchez M, Nombela C, Gil C. 2002.** Sequential fractionation and two-dimensional gel analysis unravels the complexity of the dimorphic fungus *Candida albicans* cell wall proteome.  
*Mol. Cell Proteomics.* 1, 967-982.
- Pleshinger J, Weidner E. 1985.** The microsporidian spore invasion tube. IV. Discharge activation begins with pH-triggered Ca<sup>2+</sup> influx.  
*J. Cell Biol.* 100, 1834-1838.
- Prigneau O, Achbarou A, Bouladoux N, Mazier D, Desportes-Livage I. 2000.** Identification of proteins in *Encephalitozoon intestinalis*, a microsporidian pathogen of immunocompromised humans: an immunoblotting and immunocytochemical study.  
*J Eukaryot. Microbiol.* 47, 48-56.
- Rabilloud T, Adessi C, Giraudel A, Lunardi J. 1997.** Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients.  
*Electrophoresis* 18, 307-316.
- Rabilloud T, Strub JM, Luche S, van Dorsselaer A, Lunardi J. 2001.** A comparison between Sypro Ruby and ruthenium II tris (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels.  
*Proteomics.* 1, 699-704.
- Rabitsch KP, Toth A, Galova M, Schleiffer A, Schaffner G, Aigner E, Rupp C, Penkner AM, Moreno-Borchart AC, Primig M, Esposito RE, Klein F, Knop M, Nasmyth K. 2001.** A screen for genes required for meiosis and spore formation based on whole-genome expression.  
*Curr. Biol* 11, 1001-1009.
- Raynaud L, Delbac F, Broussolle V, Rabodonirina M, Girault V, Wallon M, Cozon G, Vivarès CP, Peyron F. 1998.** Identification of *Encephalitozoon intestinalis* in travelers with chronic diarrhea by specific PCR amplification.  
*J. Clin. Microbiol.* 36, 37-40.
- Ress DA, Morris ER, Thom D, and Madden JK. 1982.** Shapes and interactions of carbohydrate chains. In : *The Polysaccharides*, edited by G. O. Aspinall (Academic Press, New York), p.196-290.
- Sagi D, Kienz P, Denecke J, Marquardt T, Peter-Katalinic J. 2005.** Glycoproteomics of N-glycosylation by in-gel deglycosylation and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry mapping: application to congenital disorders of glycosylation.  
*Proteomics.* 5, 2689-2701.
- Santoni G, Lucciarini R, Amantini C, Jacobelli J, Spreghini E, Ballarini P, Piccoli M, Gismondi A. 2002.** *Candida albicans* expresses a focal adhesion kinase-like protein that undergoes increased tyrosine phosphorylation upon yeast cell adhesion to vitronectin and the EA.hy 926 human endothelial cell line.  
*Infect. Immun.* 70, 3804-3815.
- Scheele GA. 1975.** Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins.  
*J Biol Chem* 250, 5375-5385.



- Schwartz SA, Weil RJ, Johnson MD, Toms SA, Caprioli RM. 2004.** Protein profiling in brain tumors using mass spectrometry: feasibility of a new technique for the analysis of protein expression.  
*Clin. Cancer Res.* 10, 981-987.
- Schwikowski B, Uetz P, Fields S. 2000.** A network of protein-protein interactions in yeast.  
*Nat. Biotechnol.* 18, 1257-1261.
- Service RF. 2005.** Proteomics. Protein chips map yeast kinase network.  
*Science* 307, 1854-1855.
- Shadduck JA, Polley MB. 1978.** Some factors influencing the *in vitro* infectivity and replication of *Encephalitozoon cuniculi*.  
*J Protozool.* 25, 491-496.
- Shaw JA, Mol PC, Bowers B, Silverman SJ, Valdivieso MH, Duran A, Cabib E. 1991.** The function of chitin synthases 2 and 3 in the *Saccharomyces cerevisiae* cell cycle.  
*J Cell Biol* 114, 111-123.
- Shematek EM, Cabib E. 1980.** Biosynthesis of the yeast cell wall. II. Regulation of beta-(1 leads to 3)glucan synthetase by ATP and GTP.  
*J Biol Chem* 255, 895-902.
- Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schonfisch B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C. 2003.** The proteome of *Saccharomyces cerevisiae* mitochondria.  
*Proc. Natl. Acad. Sci. U. S. A* 100, 13207-13212.
- Simpson JC, Pepperkok R. 2003.** Localizing the proteome.  
*Genome Biol* 4, 240.
- Sokolova Y, Snigirevskaya E, Morzhina E, Skarlato S, Mironov A, Komissarchik Y. 2001a.** Visualization of early golgi compartments at proliferate and sporogenic stages of a microsporidian *Nosema grylli*.  
*J. Eukaryot. Microbiol. Suppl.* 86S-87S.
- Sokolova Y, Snigirevskaya ES, Skarlato SO, Komissarchik YY, Mironov AA. 2001b.** Unusual Golgi apparatus at the proliferative stage of microsporidian life cycle.  
*Dokl. Biol. Sci.* 378, 290-293.
- Sprinzak E, Sattath S, Margalit H. 2003.** How reliable are experimental protein-protein interaction data ?  
*J Mol. Biol* 327, 919-923.
- Tachikawa H, Bloecher A, Tatchell K, Neiman AM. 2001.** A Gip1p-Glc7p phosphatase complex regulates septin organization and spore wall formation.  
*J Cell Biol* 155, 797-808.
- Tajiri M, Yoshida S, Wada Y. 2005.** Differential analysis of site-specific glycans on plasma and cellular fibronectins: Application of a hydrophilic affinity method for glycopeptide enrichment.  
*Glycobiology, sous presse.*



- Takvorian PM, Cali A. 1994.** Enzyme histochemical identification of the Golgi apparatus in the microsporidian, *Glugea stephani*. *J. Eukaryot. Microbiol.* 41, 63S-64S.
- Tanaka K, Waki H, Ido Y, Akita S, Akita S, Yoshida Y, Yoshida T. 1988.** Protein and Polymer Analyses up to m/z 100,000 by Laser Ionization Time-of-flight Mass Spectrometry. *Mass Spectrometry* 2, 151-153.
- Teixeira MC, Santos PM, Fernandes AR, Sa-Correia I. 2005.** A proteome analysis of the yeast response to the herbicide 2,4-dichlorophenoxyacetic acid. *Proteomics*. 5, 1889-1901.
- Texier C, Brosseau D, El Alaoui H, Méténier G, Vivarès CP. 2005.** Post-genomics of microsporidia, with emphasis on a model of minimal eukaryotic proteome: a review. *Folia Parasitol. (Praha)* 52, 15-22.
- Thomarat F, Vivarès CP, Gouy M. 2004.** Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. *J Mol. Evol.* 59, 780-791.
- Trabalzini L, Paffetti A, Scaloni A, Talamo F, Ferro E, Coratza G, Bovalini L, Lusini P, Martelli P, Santucci A. 2003.** Proteomic response to physiological fermentation stresses in a wild-type wine strain of *Saccharomyces cerevisiae*. *Biochem. J* 370, 35-46.
- Tsigos I, Bouriotis V. 1995.** Purification and characterization of chitin deacetylase from *Colletotrichum lindemuthianum*. *J Biol. Chem.* 270, 26286-26291.
- Tsigos I, Martinou A, Kafetzopoulos D, Bouriotis V. 2000.** Chitin deacetylases: new, versatile tools in biotechnology. *Trends Biotechnol.* 18, 305-312.
- Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, Rothberg JM. 2000.** A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623-627.
- Undeen AH, Frixione E. 1991.** Structural alteration of the plasma membrane in spores of the microsporidium *Nosema algerae* on germination. *J. Protozool.* 38, 511-518.
- Undeen AH, Vander Meer RK. 1999.** Microsporidian intrasporal sugars and their role in germination. *J Invertebr. Pathol.* 73, 294-302.
- Urban C, Sohn K, Lottspeich F, Brunner H, Rupp S. 2003.** Identification of cell surface determinants in *Candida albicans* reveals Tsa1p, a protein differentially localized in the cell. *FEBS Lett.* 544, 228-235.



- Valperga SM, Jogna Prat SA, de Valperga GJ, Lazarte SG, de Trejo AV, Diaz N, Huttman HM. 1999.** Microsporidian spores in the stool specimens of toddlers, with or without diarrhea, from Tucuman, Argentina.  
*Rev. Argent Microbiol.* 31, 157-164.
- van de Peer Y, Ben Ali A, Meyer A. 2000.** Microsporidia: accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi.  
*Gene* 246, 1-8.
- van Gool T, Canning EU, Dankert J. 1994.** An improved practical and sensitive technique for the detection of microsporidian spores in stool samples.  
*Trans. R. Soc. Trop. Med. Hyg.* 88, 189-190.
- van Gool T, Vetter JC, Weinmayr B, Van Dam A, Derouin F, Dankert J. 1997.** High seroprevalence of *Encephalitozoon* species in immunocompetent subjects.  
*J. Infect. Dis.* 175, 1020-1024.
- van Gool T, Biderre C, Delbac F, Wentink-Bonnema E, Peek R, Vivarès CP. 2004.** Serodiagnostic studies in an immunocompetent individual infected with *Encephalitozoon cuniculi*.  
*J Infect. Dis.* 189, 2243-2249.
- Vavra J. 1968.** Ultrastructural features of *Caudospora simulii* Weiser (Protozoa, Microsporidia).  
*Folia Parasitol. (Praha)* 15, 1-9.
- Vavra J. 1976a.** Structure of the microsporidia. In : *Biology of the Microsporidia*, edited by L. A. B and T. C. C (Plenum Press, New York), p.1-86.
- Vavra J. 1976b.** The development of microsporidia. In : *Biology of the Microsporidia*, edited by L. A. B and T. C. C (Plenum Press, New York), p.87-110.
- Vavra J, Chalupsky J. 1982.** Fluorescence staining of microsporidian spores with the brightener Calcofluor White M2R.  
*J Protozool.* 29, 503.
- Vavra J, Vinckier D, torpier G, Porchet E, Vivier E. 1986.** A freeze-fracture study of microsporidia : The sporophorous vesicle, the spore wall, the spore plasma membrane.  
*Protistologica* 22, 143-154.
- Vavra J, Larsson JI. 1999.** Structure of the microsporidia. In : *The Microsporidia and Microsporidiosis*, edited by M. Wittner and L. Weiss (American Society for Microbiology, Washington, DC), p.7-84.
- Vivarès CP, Loubes C, Bouix G. 1976.** Further cytochemical researches on Microsporidia parasites of the Mediterranean shore crab, *Carcinus mediterraneus* Czerniavsky, 1884.  
*Ann. Parasitol. Hum. Comp* 51, 1-14.
- Vivarès CP, Bouix G, Manier JF. 1977.** *Ormiersia carcini* gen. n., sp. n., microsporidian of the Mediterranean crab, *Carcinus mediterraneus* Czerniavsky, 1884: developmental cycle and ultrastructural study.  
*J. Protozool.* 24, 83-94.
- Vivarès CP, Sprague V. 1979.** The fine structure of *Ameson pulvis* (Microspora, Microsporidia) and its implication regarding classification and chromosome cycle.





*J. Invertebr. Pathol.* 33, 40-52.

**Vivarès CP, Méténier G. 2001.** The microsporidian *Encephalitozoon*.  
*Bioessays* 23, 194-202.

**von Heijne G. 1983.** Patterns of amino acids near signal-sequence cleavage sites.  
*Eur. J. Biochem.* 133, 17-21.

**von Mering C, Krause R, Snel B, Cornell M, Oliver SG, Fields S, Bork P. 2002.** Comparative assessment of large-scale data sets of protein-protein interactions.  
*Nature* 417, 399-403.

**Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR. 1987.** Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes.  
*Nature* 326, 411-414.

**Wadstrom T, Ljungh A. 1999.** Glycosaminoglycan-binding microbial proteins in tissue adhesion and invasion: key events in microbial pathogenicity.  
*J Med. Microbiol.* 48, 223-233.

**Washburn MP, Wolters D, Yates JR. 2001.** Large-scale analysis of the yeast proteome by multidimensional protein identification technology.  
*Nat. Biotechnol.* 19, 242-247.

**Weidner E, Byrd W, Scarborough A, Pleshinger J, Sibley D. 1984.** Microsporidian spore discharge and the transfer of polaroplast organelle membrane into plasma membrane.  
*J Protozool.* 31, 195-198.

**Weidner E, Sibley LD. 1985.** Phagocytized intracellular microsporidian blocks phagosome acidification and phagosome-lysosome fusion.  
*J. Protozool.* 32, 311-317.

**Weiss LM, Cali A, Levee E, LaPlace D, Tanowitz H, Simon D, Wittner M. 1992.** Diagnosis of *Encephalitozoon cuniculi* infection by western blot and the use of cross-reactive antigens for the possible detection of microsporidiosis in humans.  
*Am. J. Trop. Med. Hyg.* 47, 456-462.

**Weiss LM, Vossbrinck C. 1999.** Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In : *The Microsporidia and Microsporidiosis*, edited by M. Wittner and L. Weiss (American Society for Microbiology, Washington, DC), p.129-171.

**Wright JH, Craighead EM. 1922.** Infectious motor paralysis in young rabbits.  
*J. Exp. Med.* 36, 135-140.

**Xu P, Widmer G, Wang Y, Ozaki LS, Alves JM, Serrano MG, Puiu D, Manque P, Akiyoshi D, Mackey AJ, Pearson WR, Dear PH, Bankier AT, Peterson DL, Abrahamsen MS, Kapur V, Tzipori S, Buck GA. 2004.** The genome of *Cryptosporidium hominis*.  
*Nature* 431, 1107-1112.

**Xu Y, Takvorian P, Cali A, Weiss LM. 2003.** Lectin binding of the major polar tube protein (PTP1) and its role in invasion.  
*J Eukaryot. Microbiol.* 50 Suppl, 600-601.



- Xu Y, Takvorian PM, Cali A, Orr G, Weiss LM. 2004.** Glycosylation of the major polar tube protein of *Encephalitozoon hellem*, a microsporidian parasite that infects humans. *Infect. Immun.* 72, 6341-6350.
- Xu Y, Weiss LM. 2005.** The microsporidian polar tube: A highly specialised invasion organelle. *Int. J Parasitol.* 35, 941-953.
- Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC, Roberts JR, Edgerton M, Gonzalez A, Nadaf S, Moore JH, Caprioli RM, Carbone DP. 2003.** Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 362, 433-439.
- Yates JR, Speicher S, Griffin PR, Hunkapiller T. 1993.** Peptide mass maps: a highly informative approach to protein identification. *Anal. Biochem.* 214, 397-408.
- Yin Z, Stead D, Selway L, Walker J, Riba-Garcia I, McLnerney T, Gaskell S, Oliver SG, Cash P, Brown AJ. 2004.** Proteomic response to amino acid starvation in *Candida albicans* and *Saccharomyces cerevisiae*. *Proteomics.* 4, 2425-2436.
- Zhang Z, Henzel WJ. 2004.** Signal peptide prediction based on analysis of experimentally verified cleavage sites. *Protein Sci.* 13, 2819-2824.
- Ziman M, Chuang JS, Schekman RW. 1996.** Chs1p and Chs3p, two proteins involved in chitin synthesis, populate a compartment of the *Saccharomyces cerevisiae* endocytic pathway. *Mol. Biol Cell* 7, 1909-1919.
- Zhou XW, Kafsack BF, Cole RN, Beckett P, Shen RF, Carruthers VB. 2005.** The opportunistic pathogen *Toxoplasma gondii* deploys a diverse legion of invasion and survival proteins. *J Biol Chem.* 280, 34233-44.



# *A n n e x e s*



Texier C, Brosson D, El Alaoui H, Méténier G, Vivarès CP.

Post-genomics of microsporidia, with emphasis on a model of minimal eukaryotic proteome: a review.

*Folia Parasitologica* 2005 52 :15-22.





## Post-genomics of microsporidia, with emphasis on a model of minimal eukaryotic proteome: a review

Catherine Texier, Damien Brosson, Hicham El Alaoui, Guy Méténier and Christian P. Vivarès

Equipe de Parasitologie Moléculaire et Cellulaire, Laboratoire de Biologie des Protistes, UMR CNRS 6023, Université Blaise Pascal (Clermont 2), Complexe Universitaire des Cézeaux, 63177–Aubière Cedex, France

Key words: Microsporidia, *Encephalitozoon*, intracellular parasite, post-genomics, protein size, deletions, proteome analysis

**Abstract.** The genome sequence of the microsporidian parasite *Encephalitozoon cuniculi* Levaditi, Nicolau et Schoen, 1923 contains about 2,000 genes that are representative of a non-redundant potential proteome composed of 1,909 protein chains. The purpose of this review is to relate some advances in the characterisation of this proteome through bioinformatics and experimental approaches. The reduced diversity of the set of *E. cuniculi* proteins is perceptible in all the compilations of predicted domains, orthologs, families and superfamilies, available in several public databases. The phyletic patterns of orthologs for seven eukaryotic organisms support an extensive gene loss in the fungal clade, with additional deletions in *E. cuniculi*. Most microsporidial orthologs are the smallest ones among eukaryotes, justifying an interest in the use of these compacted proteins to better discriminate between essential and non-essential regions. The three components of the *E. cuniculi* mRNA capping apparatus have been especially well characterized and the three-dimensional structure of the cap methyltransferase has been elucidated following the crystallisation of the microsporidial enzyme Ecm1. So far, our mass spectrometry-based analyses of the *E. cuniculi* spore proteome has led to the identification of about 170 proteins, one-quarter of these having no clearly predicted function. Immunocytochemical studies are in progress to determine the subcellular localisation of microsporidia-specific proteins. Post-translational modifications such as phosphorylation and glycosylation are expected to be soon explored.

### INTRODUCTION

Sequencing the small genome (2.9 Mbp) of *Encephalitozoon cuniculi* Levaditi, Nicolau et Schoen, 1923, a monokaryotic microsporidian parasite infecting mammals, has revealed a complement of ~2,000 protein-coding genes that are densely distributed over 11 small chromosomes (Katinka et al. 2001). Some metabolic potentialities deduced from gene annotation have been recently reviewed, mainly as regard to energy metabolism, lipid biosynthesis, protein trafficking (Vivarès and Méténier 2004) and a Fe-S cluster synthesis pathway that is thought to take place within a mitochondrion-derived compartment or “mitosome” (Méténier and Vivarès 2004, Van der Giezen and Tovar 2004).

Although sharing a common intracellular parasitic status and a basic invasion strategy involving spore polar tube extrusion, microsporidia are highly diversified species. Undoubtedly, comparative genomic studies are of great interest to examine this diversity at the gene level. Genome sequence surveys from *Vittaforma corneae* (Mittleider et al. 2002) infecting humans and *Antonosporea* (formerly *Nosema*) *locustae* (Slamovits et al. 2004) infecting grasshoppers, have yet revealed some important differences in gene content compared to *E. cuniculi* (see review by Keeling et al. 2005). A striking example of divergence is provided by the recent demonstration that *A. locustae* spores express a class II photo-

lyase, an enzyme involved in the repair of UV-induced DNA damage (Slamovits and Keeling 2004). Indeed, no photolyase homologue is encoded in *E. cuniculi* genome and no class II photolyase has been found in fungi, the organisms viewed as the closest relatives of microsporidia. The “*Nosema locustae* Genome Project” from the Marine Biological Laboratory at Woods Hole (funded by NSF award number 0135272) has released numerous contigs at the site <http://jbpc.mbl.edu/Nosema/Nannotation.html>. New interesting data are expected from large-scale analyses of this genome.

In the three last years, bioinformatic and/or experimental studies have used *E. cuniculi* gene sequences with different objectives such as the identification of additional orthologous proteins, the evaluation of gene losses during evolution and the exploration of the function of some key-enzymes (Table 1). Regrettably, very little is still known about the characterisation of microsporidia-specific proteins that may play crucial roles in parasitic processes of development and host invasion.

Here, we will review major features of the proteome of *E. cuniculi* and some published experimental data relevant to protein functionality. Mass spectrometry studies aimed to identify proteins expressed during spore differentiation, are in progress in our laboratory and will be finally evoked.



## THE REDUCED COMPLEXITY OF THE PREDICTED PROTEOME OF *ENCEPHALITOZOON CUNICULI*

Statistical analyses for completed genomes and predicted proteomes of several prokaryotic and eukaryotic organisms including *E. cuniculi*, are available in the EBI Integr8 web site (<http://www.ebi.ac.uk/integr8/EBI-Integr8-HomePage.do>). The *E. cuniculi* non-redundant proteome is represented by only 1,909 potential “proteins” (polypeptide chains, more exactly), which is much less than in the free-living yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (6,224 and 4,929 proteins, respectively) or in the pathogenic filamentous fungus *Ashbya* (or *Eremothecium*) *gossypii* with a 9.2-Mbp genome (4,713 proteins). The number of proteins with InterPro matches, i.e. with signatures characteristic for known families, domains or functional sites, attains only 1,294 in *E. cuniculi*, contrasting with 3,713 in *A. gossypii* and 3,366 in the apicomplexan parasite *Plasmodium falciparum* 3D7. A reduced diversity of the potential microsporidian proteins sharing common features with known proteins is also evident through their distribution at the superfamily level (Madera et al. 2004, <http://supfam.mrc-lmb.cam.ac.uk/SUPERFAMILY>). Only 364 superfamilies with an assignment are distinguished in *E. cuniculi*. This is far from the numbers in two *Plasmodium* species (566 in *P. falciparum* and 567 in *P. yoelii*) and in various fungi (from 620 in *Candida glabrata* to 774 in *Fusarium graminearum*).

The updated version of the NCBI COG database contains a compilation of 4,852 clusters of orthologs derived from seven sequenced eukaryotic genomes (*Homo sapiens*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *S. cerevisiae*, *S. pombe* and *E. cuniculi*) and named “eukaryotic orthologous groups” or KOGs (Tatusov et al. 2003, <http://www.ncbi.nlm.nih.gov/COG/new/shokog.cgi>). About 60% of *E. cuniculi* proteins are assigned to KOGs. A massive gene loss in the fungal clade, with additional elimination in *E. cuniculi*, has been inferred from the reconstruction of parsimonious evolutionary scenarios using either a subset of KOGs (Tatusov et al. 2003) or the complete collection (Koonin et al. 2004). The common ancestor of *E. cuniculi* and two yeast species is predicted to contain 3,048 KOGs, and the branch leading to the microsporidian would be characterized by 586 gene gains and up to 1,969 gene losses! This dramatic elimination concerns 269 genes of unknown function, 241 genes with only a general functional prediction and various genes assigned to all major functional categories: 167 for post-translational modification, protein turnover and chaperones, 122 for translation, 116 for intracellular trafficking and secretion, 110 for energy production and conversion, among others.

Strikingly, the *E. cuniculi* proteome shows a tendency to the shortening of individual protein sequences, in relation with a high frequency of intragenic deletions.

The average protein length is close to 360 amino acid (aa) residues and only 6 sequences encompass 2,000 aa (Katinka et al. 2001). Numerous proteins are frequently smaller than most eukaryotic counterparts. Typical examples of compacted proteins are offered by two conserved members of the AAA ATPase family: cytoplasmic dynein heavy chain (Vivarès et al. 2002) and midasin (Garbarino and Gibbons 2002). The latter is a nuclear protein encoded by the largest open reading frame in *S. cerevisiae*. Yeast midasin (4,910 aa; ~560 kDa) is smaller than human ortholog (5,596 aa; ~632 kDa) but *E. cuniculi* midasin has been prone to a severe reduction (2,832 aa; ~324 kDa). As illustrated in Fig. 1A, so far the microsporidian protein is the smallest, compared to those from animals, fungi or protists. Analysis by multiple sequence alignment has revealed the reduction of all five domains at different extents (Garbarino and Gibbons 2002). Taking *S. cerevisiae* midasin as a reference, the authors have provided the following degrees of reduction: 90% for the N-terminal domain, 19% for the AAA domain retaining all six tandem AAA protomers, 50% for the lengthy linker domain as for the acidic D/E-rich domain, and 20% for the C-terminal M-domain retaining all three MIDAS motifs. This is consistent with the critical roles of AAA- and M-domains, and indicates the non-essential character of large parts of the three other domains. We have examined the case of ribonucleotide reductase (RNR), a ubiquitous enzyme required to provide the 2'-deoxyribonucleotide precursors for both DNA replication and DNA repair. Like in all other eukaryotes, the *E. cuniculi* enzyme can be assigned to the oxygen-dependent class Ia and should exist as a tetramer  $\alpha_2\beta_2$ . The large ( $\alpha$ ) and small ( $\beta$ ) subunits contain 768 aa and 325 aa, respectively. These correspond to the lowest RNR lengths among eukaryotic organisms whereas most fungi display the highest values (Fig. 1B). The large subunit of *E. cuniculi* RNR has a reduced C-terminal domain, justifying a length close to R1 chain in *Escherichia coli* (761 aa) or in several eukaryote viruses. Conversely, the small subunit has a reduced N-terminal domain and is closer in size to *E. coli* R2F chain (319 aa).

After assignment of conserved domains from the NCBI CDD database to orthologous sets of eukaryotic proteins, Koonin et al. (2004) have considered that “...among the unicellular eukaryotes, *Encephalitozoon* had by far the least complex domain architectures, which reflects the general reduction in this intracellular parasite”. The number of orthologous groups in which the average number of domains is greater in *E. cuniculi* ranges only from 10 (*vs Homo sapiens*) to 19 (*vs S. pombe*). In contrast, the number of groups in which the domain number is lower in *E. cuniculi* varies from 332 (*vs H. sapiens*) to 140 (*vs S. pombe*). Thus, the *E. cuniculi* proteome appears as a package of compact proteins containing a significant proportion of orthologs with simplified domain organisation.



**Table 1.** A list of studies on the characterisation of *Encephalitozoon cuniculi* gene sequences or products, in a three-year period (2002–2004).

<i>E. cuniculi</i> proteins or RNAs	Main findings	References
Kinesin-like proteins (KLPs 2, 3, 4, 5 and 6)	– Expression and localisation of five <i>E. cuniculi</i> GFP-KLPs in the protist <i>Leishmania major</i>	Dubessay et al. 2004*
Protein kinase A, catalytic subunit (ECU07_0520; EcPKAc)	– Recombinant EcPKAc is inactive in phosphorylating a usual synthetic substrate of PKAs. Two gene copies in <i>E. intestinalis</i>	Equinet et al. 2004*
RNA polymerase II subunit A C-terminal domain phosphatase (ECU07_0890; EcFcp1)	– EcFcp1 is a monomeric enzyme that has CTD phosphatase activity <i>in vitro</i> and can replace yeast Fcp1 <i>in vivo</i>	Hausmann et al. 2004*
Orthologous protein clusters (see COG database)	– Scenarios of loss and emergence of genes during eukaryotic evolution, based on comparative analyses of KOGs from seven eukaryotes including <i>E. cuniculi</i>	Krylov et al. 2003 Tatusov et al. 2003 Koonin et al. 2004
Signal recognition particle RNA	– An SRP RNA sequence is present on chromosome X	Rosenblad et al. 2004
ATP/ADP translocase (4 genes)	– All ATP/ADP translocases would be of deep origin	Amiri et al. 2003
Methionine aminopeptidase-2 (ECU10_0750, EcMetAP2)	– Homology modelling; potential binding sites of EcMetAP2 for Co <sup>2+</sup> , fumagillin and TNP-470 – TNP-470 inhibits the activity of EcMetAP2 <i>in vitro</i>	Bontems et al. 2003* Weiss et al. 2003*
Homeobox (12 genes)	– Clustering of two homeobox genes is suggestive of a mating-type locus (ECU10_1470, EcHd1; ECU10_1480, EcHd2)	Burglin 2003*
Midasin (ECU08_1900)	– Large AAA ATPase characterized by a strong size reduction in <i>E. cuniculi</i>	Garbarino and Gibbons 2002
20S proteasome (14 genes)	– Seven $\beta$ -type subunits are assigned to $\beta$ 1– $\beta$ 7 subtypes but $\alpha$ subunits are highly divergent	Gille et al. 2003
tRNAs (46 genes)	– Comparative study of 60 genomes to detect tRNA paralogs	Xue et al. 2003
Pot1-like protein (ECU05_1050)	– Similar to a single-stranded telomeric DNA-binding protein	Baumann et al. 2002
ABC transporters (13 genes)	– Only four families of ABC proteins are distinguished	Cornillot et al. 2002*
mRNA 5'-triphosphatase (ECU04_1550, EcCet1)	– EcCet1 is shown to be an homodimeric metal-dependent RNA triphosphatase and lacks a non-essential N-terminal extension	Hausmann et al. 2002*
mRNA guanylyltransferase (ECU09_0400, EcCeg1)	– EcCeg1 exhibits guanylyltransferase activity <i>in vitro</i> and can sustain the growth of yeast <i>ceg1<math>\Delta</math></i> mutants	Hausmann et al. 2002*
mRNA cap (guanine-N7) methyltransferase (ECU10_0380, Ecm1)	– Ecm1 can sustain the growth of yeast <i>abd1<math>\Delta</math></i> mutants – Resolution of crystal structures of this enzyme in complex with S-adenosylmethionine, S-adenosylhomocysteine and the cap guanylate	Hausmann et al. 2002* Fabrega et al. 2004*
ORMDL (ECU11_1150)	– Assignment to a novel family of ER membrane proteins	Hjelmqvist et al. 2002
NARF-like protein (ECU05_0970)	– Similar to nuclear prelamin A recognition factor and Fe-hydrogenases	Horner et al. 2002
Ribosomal proteins	– Comparative analysis in sequenced genomes; <i>E. cuniculi</i> lacks S21e, L14e, L28e and L38e homologs	Lecompte et al. 2002
Polar tube protein 3 precursor (ECU11_1440, PTP3)	– Identification and characterisation of a microsporidia-specific protein associated with the polar tube	Peuvel et al. 2002*
Ferredoxin (ECU07_0600)	– Assignment to type-II mitochondrial ferredoxin	Seeber 2002

\*Stars indicate reports dealing with only microsporidial genes.

Most gene gains presumably arose from duplication-divergence events. However, a small set of KOGs represented in *E. cuniculi* but absent in the two yeast species, has led to propose two explanations: (1) loss of ancestral archaeo-eukaryotic genes in yeasts, for the orthologs that are also found in all archaea but not in bacteria, and (2) acquisition of genes by *E. cuniculi* through horizontal gene transfer, for those showing high similarities to some bacterial orthologs, e.g. 3-methyl-adenine DNA glycosylase and ATP/ADP translocase (Tatusov et al.

2003, Koonin et al. 2004). Horizontal gene transfer has been also invoked to justify the presence of class II catalase and photolyase in *A. locustae* (Fast et al. 2003, Slamovits and Keeling 2004). As non-mitochondrial ATP/ADP translocases (ATP-importing) are distributed in plant chloroplasts, microsporidia and intracellular parasitic bacteria (*Rickettsia*, Chlamydiaceae), a relatively recent acquisition by microsporidia and plants via independent horizontal gene transfers from bacteria has been hypothesized (Koonin et al. 2004, Méténier and



Vivarès 2004). An alternate hypothesis has been proposed in a phylogenetic study supporting a very ancient origin for all mitochondrial and non-mitochondrial (“plastid/parasite”) ATP/ADP translocases (Amiri et al. 2003). The “plastid/parasite” gene would have been present in a bacterial ancestor of mitochondria and *Rickettsia*, and then transferred into the nuclear genome of the early eukaryotic cell. This gene would have been subsequently lost in various eukaryotic lineages, except in plants for cytosol-to-plastid ATP transport and in microsporidia for host ATP import. Whatever the evolutionary scenario, discovering the membrane location and substrate specificity of the four ATP/ADP translocase paralogs in *E. cuniculi* will be of interest for the knowledge of host-microsporidia energetic relationships.

### SOME INSIGHTS INTO THE FUNCTIONALITY OF ENCEPHALITOOZON CUNICULI PROTEINS

When examining the list of reports in Table 1, it is clear that the deepest post-genomic experimental investigations of *E. cuniculi* proteins deal with the biochemical and genetic dissection of the three enzymes responsible for mRNA capping: mRNA 5'-triphosphatase (EcCet1), RNA guanylyltransferase (EcCeg1) and RNA (guanine-N7) methyltransferase (Ecm1). One reason of the interest in the use of *E. cuniculi* sequences was to determine whether the triphosphatase component was of cysteine-phosphatase or metal-dependent type. Another reason was to test the functionality of three typically compacted microsporidial enzymes using yeast expression plasmids. Lacking the large N-terminal extensions observed in fungal homologs, EcCet1 contains only 221 aa. The purified recombinant enzyme can hydrolyze ATP in the presence of  $Mn^{2+}$  or  $Co^{2+}$  ions, and the replacement of conserved glutamate residues required for metal coordination with alanine residues abolishes its activity, revealing a metal-dependent triphosphatase like in fungi, DNA viruses and *Plasmodium* (Hausmann et al. 2002). Intermediate in size between *Chlorella* virus and *S. pombe* guanylyltransferases, EcCeg1 (364 aa) has been demonstrated to be specific for GTP and active in cap formation *in vivo* in *S. cerevisiae* cells deleted for the endogenous enzyme. Likewise, Ecm1 without N-terminal extensions and reduced to 298 aa (*vs* 476 aa in *H. sapiens* and 402 aa in *S. pombe*), has been verified to be active by complementation of yeast mutants without endogenous cap methyltransferase (Hausmann et al. 2002). In summary, the components of the *E. cuniculi* mRNA capping apparatus provide a nice illustration of eukaryotic enzymes having preserved just a little more than the minimum catalytic domains. The X-ray structure of guanine-N7-methyl transferase has been successfully determined for the first time after co-crystallisation of Ecm1 with three different ligands, thus completing the suite of protein structures for mRNA capping enzymes (Fabrega et al. 2004). This confirms that the use

of minimal *E. cuniculi* orthologs can help to elucidate the three-dimensional (3-D) structure of important enzymes and to identify critical regions for ligand binding.

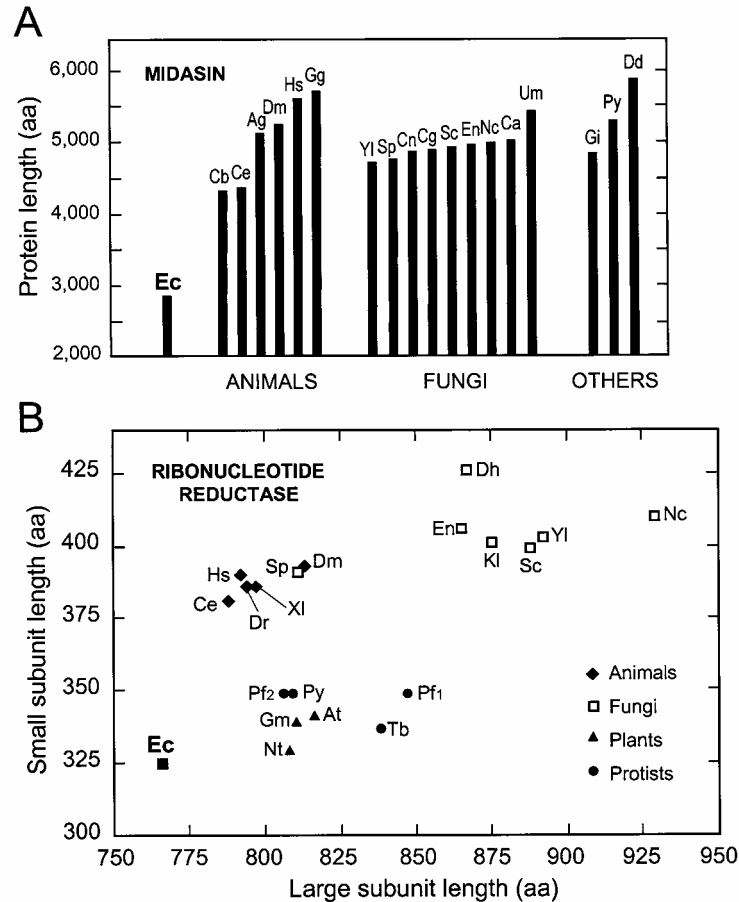
Because of its somewhat surprising results, we will also mention a recent study on the expression of green fluorescent protein (GFP)-fused versions of *E. cuniculi* genes in an original heterologous system represented by the parasitic flagellate *Leishmania major* (Dubessay et al. 2004). The selected genes were those coding for a family of six kinesin-like proteins (EcKLPs) and five of these were successfully expressed in *Leishmania*. Three EcKLPs exhibited a diffuse cytoplasmic localisation. The surprise arose from the very curious but precise localisation of the two other GFP-fused proteins. EcKLP3 was confined at the posterior end of the *Leishmania* cell, a region characterized by the termination of subpellicular microtubules, while EcKLP6 was targeted to the tip of the flagellum. Even if these two opposite locations cannot be extrapolated to the *Encephalitozoon* cell, it seems likely that the different EcKLPs play specific roles in intracellular transport processes required for microsporidian life. The comparison with data of immunolocalisation in *E. cuniculi* cells is expected to be informative.

A protein that has drawn the attention of workers in the field of the chemotherapy of microsporidiosis is methionine aminopeptidase type 2 (MetAP2). Fumagillin or its synthetic derivative TNP-470 has been shown to bind covalently to mammalian MetAP2 (Griffith et al. 1998, Liu et al. 1998). These compounds are efficient antimicrosporidial agents and it is conceivable that microsporidial MetAP2 also represents fumagillin/TNP-470 target. An explanation is offered by the lack of MetAP1 gene in the *E. cuniculi* genome (Katinka et al. 2001). The effect of a drug inhibiting the unique MetAP2 might be sufficient to strongly alter the viability of microsporidia. Data from two different approaches support specific interactions between *E. cuniculi* MetAP2 (EcMetAP2) and fumagillin/TNP-470: (1) *in silico* modelling of EcMetAP2 3-D structure indicates the conservation of critical residues for these interactions (Bontems et al. 2003), and (2) a purified recombinant EcMetAP2 is strongly inhibited by TNP-470 (Weiss et al. 2003). Note, however, that depletion of MetAP2 by small interfering RNA failed to inhibit the proliferation of endothelial cells and to modify their sensitivity to fumagillin (Kim et al. 2004). It cannot be excluded that other fumagillin targets exist in such cells, and perhaps also in microsporidia.

Investigating microsporidia-specific proteins is a major challenge. Two polar tube proteins (PTP1, PTP2), extracted in the presence of a thiol-reducing compound, are encoded by two neighbouring genes in *Encephalitozoon* species (Delbac et al. 1998, 2001, Keohane et al. 1998). A large-sized protein (PTP3; 1,256 aa) has been subsequently demonstrated to be associated with the polar tube of *E. cuniculi* (Peuvel et al. 2002). Chemical







**Fig. 1.** Lengths of midasin (A) and of ribonucleotide reductase subunits (B) in *Encephalitozoon cuniculi* (Ec) and other eukaryotic organisms. Species abbreviations are: Ag – *Anopheles gambiae*; At – *Arabidopsis thaliana*; Ca – *Candida albicans*; Cb – *Caenorhabditis briggsae*; Ce – *Caenorhabditis elegans*; Cg – *Candida glabrata*; Cn – *Cryptococcus neoformans*; Dd – *Dicystostelium discoideum*; Dh – *Debaryomyces hansenii*; Dm – *Drosophila melanogaster*; Dr – *Danio rerio*; En – *Emericella nidulans*; Gg – *Gallus gallus*; Gi – *Giardia intestinalis*; Gm – *Glycine max*; Hs – *Homo sapiens*; K1 – *Kluyveromyces lactis*; Nc – *Neurospora crassa*; Nt – *Nicotiana tabacum*; Pf<sub>1</sub> – *Plasmodium falciparum* 3D7; Pf<sub>2</sub> – *Plasmodium falciparum* Dd2; Py – *Plasmodium yoelii*; Sc – *Saccharomyces cerevisiae*; Sp – *Schizosaccharomyces pombe*; Tb – *Trypanosoma brucei*; Um – *Ustilago maydis*; XI – *Xenopus laevis*; Yl – *Yarrowia lipolytica*.

cross-linking experiments support the involvement of the three proteins in the formation of a large multimeric complex. In contrast with the two other PTPs, PTP3 has no cysteine residues, indicating that this protein cannot interact via disulfide linkages. The most abundant PTP is PTP1 (~70% of the mass of the polar tube) and evidence for O-mannosylation of this protein has been recently provided (Xu et al. 2004). Sequence data about spore wall proteins are restricted to one *E. cuniculi* protein (SWP1, Bohne et al. 2000) and two related *Encephalitozoon intestinalis* proteins having divergent C-terminal domains (SWP1 and SWP2, Hayman et al. 2001), all localised to the outer wall layer (exospore). Two *E. cuniculi* proteins associated with the inner wall layer (endospore) are currently investigated in our laboratory (Delbac F., unpublished data).

It should be stressed that a significant conservation of gene order exists between two distantly related micro-

sporidians such as *E. cuniculi* and *Antonospora locustae* (Slamovits et al. 2004). The compilation of the groups of neighbouring genes common to the two species should provide a good basis to further investigate undefined proteins that are encoded by homologous genes within similar chromosomal contexts, shared by at least two different genera of the microsporidian phylum.

#### FIRST MASS SPECTROMETRY-BASED PROTEOMICS APPROACHES AND PERSPECTIVES

An important area of proteomics research implies the use of mass spectrometry (MS) as a powerful tool to identify small amounts of proteins from complex mixtures, frequently after separation by 2-D gel electrophoresis procedures. Even in the case of species cultured *in vitro*, there are very few reports of 2-D patterns of microsporidian proteins, mainly because of difficulties in the development of reproducible protocols for the ex-



traction of proteins from spores equipped with a resistant wall and for the fractionation of intracellular stages including wall-lacking meronts. However, an interesting approach has been realized with *E. intestinalis* maintained in monkey kidney E6 cells, consisting in the comparison of 2-D protein profiles of four kinds of samples: purified *E. intestinalis* spores, E6 cells with estimated infection levels of either ~50% or ~90%, and control uninfected E6 cells (Moura and Visvesvara 2001). Of more than 200 protein spots resolved in mid-sized gels, 95 were found to be immunogenic on the basis of their reactivity with a rabbit anti-*E. intestinalis* serum, some of these spots providing potential markers for either merogonial or sporogonial stages. In another study, whole spores and soluble fractions of three *Encephalitozoon* species and *Brachiola algerae* were directly analyzed by matrix-assisted laser desorption ionisation time-of-flight (MALDI-TOF) MS, providing species-specific spectral markers within the molecular size range of 2–8 kDa (Moura et al. 2003).

Within the context of the exploitation of *E. cucurbitae* genome sequence and with the help of a laboratory specialised in protein identification by automated MS techniques, we have undertaken a first large-scale study of the proteome representative of the latest *E. cucurbitae* sporogonial stages (immature and mature spores) (Texier et al. 2004). Approximately 350 Coomassie blue-stained spots, in amounts accessible to MALDI-TOF and MS/MS analyses, were resolved with molecular masses between 10 and 250 kDa and within the pI range of 3–10, after 2-D electrophoresis in large-sized gels. Independently, whole protein extracts were subjected to trypsin digestion, then liquid chromatography coupled to MS/MS identification, revealing especially proteins with multiple transmembrane domains. More than 170 different proteins have been assigned, at the present time. Their distribution among large functional categories is consistent with some major requirements for the protection against environmental stresses. Several components of the translation machinery were also identified. Whether this reflects a late protein synthesis during sporogony or a pool of essential factors that will be used after sporoplasm discharge for initiating merogony, or both, is an open question.

About 25% of identified proteins have no clearly predicted function, suggesting that a large proportion of *E. cucurbitae* genes encoding the so-called “hypothetical proteins” can be expressed. Some of these genes were cloned and corresponding recombinant proteins were used for the production of polyclonal antibodies in order to determine the localisation of the antigenic proteins at the ultrastructural level. Currently, we have obtained evidence for four proteins targeted to the spore envelope. One of these proteins has a predicted domain characteristic for the polysaccharide deacetylase family and was found to be localised at the endospore/plasma membrane interface (Brosson D., Kuhn L., Prensier G., Vivarès C.P. and Texier C., manuscript in preparation). This is reminiscent of the localisation of chitin deacetylase near the inner face of the cell wall (periplasmic space) in the mycelium of the zygomycete *Absidia coerulea* (Gao et al. 1995). Given the chitin richness of the microsporidian endospore, it seems very likely that the *E. cucurbitae* enzyme acts as a chitin deacetylase, hydrolyzing the N-acetamido groups of N-acetyl-D-glucosamine residues in chitin chains and producing chitosan. Thus, the construction of a highly rigid cell wall during sporogony presumably depends on the combined functioning of chitin synthase and chitin deacetylase.

A usual characteristic of protozoan genomes is the placement of gene families encoding surface proteins near the ends of multiple chromosomes. Putative gene families associated with subtelomeric regions of *E. cucurbitae* chromosomes are currently tested for a possible expression. In a next future, we expect to evaluate the subsets of microsporidian proteins subjected to major post-translational modifications such as phosphorylation and glycosylation (“phosphoproteome” and “glycoproteome”), in complement of specific studies on proteins required for cell cycle progression or for spore wall and polar tube organisation. As already stressed in a previous review (Vivarès and Méténier 2004), it is important to test whether *E. cucurbitae* is really incapable of N-glycosylation. Thorough biochemical investigations of carbohydrate content and glycosidic linkages are therefore required.

## REFERENCES

- AMIRI H., KARLBERG O., ANDERSSON S.G.E. 2003: Deep origin of plastid/parasite ATP/ADP translocases. *J. Mol. Evol.* 56: 137–150.
- BAUMANN P., PODELL E., CECH T.R. 2002: Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. *Mol. Cell. Biol.* 22: 8079–8087.
- BOHNE W., FERGUSON D.J., KOHLER K., GROSS U. 2000: Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cucurbitae*. *Infect. Immun.* 68: 2268–2275.
- BONTEMS F., LE FLOCH P., DUFFIEUX F., BIDERRE C., PEYRET P., LALLEMAND J.Y. 2003: Homology modeling and calculation of the cobalt cluster charges of the *Encephalitozoon cucurbitae* methionine aminopeptidase, a potential target for drug design. *Biophys. Chem.* 105: 29–43.
- BURGLIN T.R. 2003: The homeobox genes of *Encephalitozoon cucurbitae* (Microsporidia) reveal a putative mating-type locus. *Dev. Genes Evol.* 213: 50–52.



- CORNILLOT E., MÉTÉNIER G., VIVARÈS C.P., DASSA E. 2002: Comparative analysis of sequences encoding ABC systems in the genome of the microsporidian *Encephalitozoon cuniculi*. FEMS Microbiol. Lett. 210: 39–47.
- DELBAC F., PEUVEL I., MÉTÉNIER G., PEYRETAILLADE E., VIVARÈS C.P. 2001: Microsporidian invasion apparatus: identification of a novel polar tube protein and evidence for clustering of *ptp1* and *ptp2* genes in three *Encephalitozoon* species. Infect. Immun. 69: 1016–1024.
- DELBAC F., PEYRET P., MÉTÉNIER G., DAVID D., DANCHIN A., VIVARÈS C.P. 1998: On proteins of the microsporidian invasive apparatus: complete sequence of a polar tube protein of *Encephalitozoon cuniculi*. Mol. Microbiol. 29: 825–834.
- DUBESSAY P., PAGÈS M., DELBAC F., BASTIEN P., VIVARÈS C., BLAINEAU C. 2004: Can heterologous gene expression shed (a torch) light on protein function? Trends Biotechnol. 22: 557–559.
- EQUINET L., BAPTESTE E., THELLIER M., OUARZANAMARA M., VIVARÈS C.P., DESPORTES-LIVAGE I., DOERIG C. 2004: The genes encoding cAMP-dependent protein kinase catalytic subunit homologues of the microsporidia *Encephalitozoon intestinalis* and *E. cuniculi*: molecular characterisation and phylogenetic analysis. Parasitol. Int. 53: 277–285.
- FABREGA C., HAUSMANN S., SHEN V., SHUMAN S., LIMA C.D. 2004: Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. Mol. Cell 13: 77–89.
- FAST N.M., LAW J.S., WILLIAMS B.A., KEELING P.J. 2003: Bacterial catalase in the microsporidian *Nosema locustae*: implications for microsporidian metabolism and genome evolution. Eukaryot. Cell 2: 1069–1075.
- GAO X.D., KATSUMOTO T., ONODERA K. 1995: Purification and characterization of chitin deacetylase from *Abisidia coerulea*. J. Biochem. (Tokyo) 117: 257–263.
- GARBARINO J.E., GIBBONS I.R. 2002: Expression and genomic analysis of midasin, a novel and highly conserved AAA protein distantly related to dynein. BMC Genomics 3: 18.
- GILLE C., GOEDE A., SCHLOETELBURG C., PREISSNER R., KLOETZEL P.M., GOBEL U.B., FROMMEL C. 2003: A comprehensive view on proteasomal sequences: implications for the evolution of the proteasome. J. Mol. Biol. 326: 1437–1448.
- GRIFFITH E.C., SU Z., NIWAYAMA S., RAMSAY C.A., CHANG Y.H., LIU J.O. 1998: Molecular recognition of angiogenesis inhibitors fumagillin and ovalicin by methionine aminopeptidase 2. Proc. Natl. Acad. Sci. USA 95: 15183–15188.
- HAUSMANN S., SCHWER B., SHUMAN S. 2004: An *Encephalitozoon cuniculi* ortholog of the RNA polymerase II carboxyl-terminal domain (CTD) serine phosphatase Fcp1. Biochemistry 43: 7111–7120.
- HAUSMANN S., VIVARÈS C.P., SHUMAN S. 2002: Characterization of the mRNA capping apparatus of the microsporidian parasite *Encephalitozoon cuniculi*. J. Biol. Chem. 277: 96–103.
- HAYMAN J.R., HAYES S.F., AMON J., NASH T.E. 2001: Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*. Infect. Immun. 69: 7057–7066.
- HJELMQVIST L., TUSON M., MARFANY G., HERRERO E., BALCELLS S., GONZALEZ-DUARTE R. 2002: ORMDL proteins are a conserved new family of endoplasmic reticulum membrane proteins. Genome Biol. 3: R27.
- HORNER D.S., HEIL B., HAPPE T., EMBLEY T.M. 2002: Iron hydrogenases – ancient enzymes in modern eukaryotes. Trends Biochem. Sci. 27: 148–153.
- KATINKA M.D., DUPRAT S., CORNILLLOT E., MÉTÉNIER G., THOMARAT F., PRENSIER G., BARBE V., PEYRETAILLADE E., BROTTIER P., WINCKER P., DELBAC F., EL ALAOUI H., PEYRET P., SAURIN W., GOUY M., WEISSENBAACH J., VIVARÈS C.P. 2001: Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414: 450–453.
- KEELING P.J., FAST N.M., LAW J.S., WILLIAMS B.A.P., SLAMOVITS C.H. 2005: Comparative genomics of microsporidia. Folia Parasitologica 52: XXX–XXX.
- KEOHANE E.M., ORR G.A., ZHANG H.S., TAKVORIAN P.M., CALI A., TANOWITZ H.B., WITTNER M., WEISS L.M. 1998: The molecular characterization of the major polar tube protein gene from *Encephalitozoon hellem*, a microsporidian parasite of humans. Mol. Biochem. Parasitol. 94: 227–236.
- KIM S., LAMONTAGNE K., SABIO M., SHARMA S., VERSACE R.W., YUSUFF N., PHILLIPS P.E. 2004: Depletion of methionine aminopeptidase 2 does not alter cell response to fumagillin or bengamides. Cancer Res. 64: 2984–2987.
- KOONIN E.V., FEDOROVA N.D., JACKSON J.D., JACOBS A.R., KRYLOV D.M., MAKAROVA K.S., MAZUMDER R., MEKHEDOV S.L., NIKOLSKAYA A.N., RAO B.S., ROGOZIN I.B., SMIRNOV S., SOROKIN A.V., SVERDLOV A.V., VASUDEVAN S., WOLF Y.I., YIN J.J., NATALE D.A. 2004: A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biol. 5: R7.
- KRYLOV D.M., WOLF Y.I., ROGOZIN I.B., KOONIN E.V. 2003: Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. Genome Res. 13: 2229–2235.
- LECOMPTE O., RIPP R., THIERRY J.-C., MORAS D., POCH O. 2002: Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acids Res. 30: 5382–5390.
- LIU S., WIDOM J., KEMP C.W., CREWS C.M., CLARDY J. 1998: Structure of human methionine aminopeptidase-2 complexed with fumagillin. Science 282: 1324–1327.
- MADERA M., VOGEL C., KUMMERFELD S.K., CHOTHIA C., GOUGH J. 2004: The SUPERFAMILY database in 2004: additions and improvements. Nucleic Acids Res. 32 Database issue: D235–239.
- MÉTÉNIER G., VIVARÈS C.P. 2004: Genomics of microbial parasites: the microsporidian paradigm. In: R.P. Hirt and D.S. Horner (Eds.), Organelles, Genomes and Eukaryote



- Phylogeny. An Evolutionary Synthesis in the Age of Genomics. CRC Press, Boca Raton, pp. 207–236.
- MITTLEIDER D., GREEN L.C., MANN V.H., MICHAEL S.F., DIDIER E.S., BRINDLEY P.J. 2002: Sequence survey of the genome of the opportunistic microsporidian pathogen, *Vittaforma corneae*. *J. Eukaryot. Microbiol.* 49: 393–401.
- MOURA H., OSPINA M., WOOLFITT A.R., BARR J.R., VISVESVARA G.S. 2003: Analysis of four human microsporidian isolates by MALDI-TOF mass spectrometry. *J. Eukaryot. Microbiol.* 50: 156–163.
- MOURA H., VISVESVARA G.S. 2001: A proteome approach to the host-parasite interaction of the microsporidian *Encephalitozoon intestinalis*. *J. Eukaryot. Microbiol.* 48 (Suppl.): 56S–59S.
- PEUVEL I., PEYRET P., MÉTÉNIER G., VIVARÈS C.P., DELBAC F. 2002: The microsporidian polar tube: evidence for a third polar tube protein (PTP3) in *Encephalitozoon cuniculi*. *Mol. Biochem. Parasitol.* 122: 69–80.
- ROSENBLAD M.A., ZWIEB C., SAMUELSSON T. 2004: Identification and comparative analysis of components from the signal recognition particle in protozoa and fungi. *BMC Genomics* 5: Article 5.
- SEEBER F. 2002: Eukaryotic genomes contain a [2Fe-2S] ferredoxin isoform with a conserved C-terminal sequence motif. *Trends Biochem. Sci.* 27: 545–547.
- SLAMOVITS C.H., FAST N.M., LAW J.S., KEELING P.J. 2004: Genome compaction and stability in microsporidian intracellular parasites. *Curr. Biol.* 14: 891–896.
- SLAMOVITS C.H., KEELING P.J. 2004: Class II photolyase in a microsporidian intracellular parasite. *J. Mol. Biol.* 341: 713–721.
- TATUSOV R.L., FEDOROVA N.D., JACKSON J.D., JACOBS A.R., KIRYUTIN B., KOONIN E.V., KRYLOV D.M., MAZUMDER R., MEKHEDOV S.L., NIKOLSKAYA A.N., RAO B.S., SMIRNOV S., SVERDLOV A.V., VASUDEVAN S., WOLF Y.I., YIN J.J., NATALE D.A. 2003: The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4: 41.
- TEXIER C., BROSSON D., KUHN L., GARIN J., VIVARÈS C. 2004: Proteomics of the microsporidian *Encephalitozoon cuniculi*: application to the identification of new spore wall proteins. *J. Eukaryot. Microbiol.* 51 (Suppl.): 28A.
- VAN DER GIEZEN M., TOVAR J. 2004: Mitosomes, hydrogenosomes and mitochondria: variations on a theme? In: R.P. Hirt and D.S. Horner (Eds.), *Organelles, Genomes and Eukaryote Phylogeny. An Evolutionary Synthesis in the Age of Genomics*. CRC Press, Boca Raton, pp. 289–308.
- VIVARÈS C.P., GOUY M., THOMARAT F., MÉTÉNIER G. 2002: Functional and evolutionary analysis of a eukaryotic parasitic genome. *Curr. Opin. Microbiol.* 5: 499–505.
- VIVARÈS C.P., MÉTÉNIER G. 2004: The microsporidia genome: living with minimal genes as an intracellular eukaryote. In: D.S. Lindsay and L.M. Weiss (Eds.), *Opportunistic Infections: Toxoplasma, Sarcocystis and Microsporidia*. World Class Parasites. Vol. 9. Kluwer Academic Publishers, Boston, pp. 215–242.
- WEISS L.M., ZHOU G.C., ZHANG H. 2003: Characterization of recombinant microsporidian methionine aminopeptidase type 2. *J. Eukaryot. Microbiol.* 50 (Suppl.): 597–599.
- XU Y., TAVKORIAN P.M., CALI A., ORR G., WEISS L.M. 2004: Glycosylation of the major polar tube protein of *Encephalitozoon hellem*, a microsporidian parasite that infects humans. *Infect. Immun.* 72: 6341–6350.
- XUE H., TONG K.-L., MARCK C., GROSJEAN H., WONG J.T. 2003: Transfer RNA paralogs: evidence for genetic code-aminoacid biosynthesis coevolution and an archaean root of life. *Gene* 310: 59–66.

Received 8 December 2004

Accepted 2 February 2005





Brosson D, Kuhn L, Prensier G, Vivarès CP, Texier C.

The putative chitin deacetylase of *Encephalitozoon cuniculi*: a surface protein implicated in the microsporidian spore wall formation.

FEMS Microbiology Letter 2005 247:81-90.



# The putative chitin deacetylase of *Encephalitozoon cuniculi*: A surface protein implicated in microsporidian spore-wall formation

Damien Brosseau<sup>a</sup>, Lauriane Kuhn<sup>b</sup>, Gérard Prensier<sup>a</sup>, Christian P. Vivarès<sup>a</sup>,  
Catherine Texier<sup>a,\*</sup>

<sup>a</sup> *Equipe Parasitologie Moléculaire et Cellulaire, LBP, UMR CNRS 6023, Université Blaise Pascal, 63177 Aubière CEDEX, France*

<sup>b</sup> *Laboratoire de Chimie des Protéines, DBMS, CEA, 38054 Grenoble, France*

Received 2 March 2005; received in revised form 20 April 2005; accepted 21 April 2005

First published online 3 May 2005

Edited by R.C. Staples

## Abstract

Microsporidia are fungal-like unicellular eukaryotes which develop as obligate intracellular parasites. They differentiate into resistant spores that are protected by a thick cell wall composed of glycoproteins and chitin. Despite an extensive description of the fibrillar structure of this wall, very little is known about its protein components and deposit mechanisms. In this study on the human pathogen *Encephalitozoon cuniculi*, we identify by mass spectrometry the target of polyclonal antibodies previously raised against a 33-kDa protein located at the outer face of the parasite plasma membrane. This 254-amino acid protein is encoded by the ECU11\_0510 open reading frame and presents two isoforms of 33 and 55 kDa. Sequence analysis supports an assignment to the polysaccharide deacetylase family with a suspected chitin deacetylase activity (*EcCDA*). As demonstrated by TEM studies, *EcCDA* is present at the plasma membrane of the early stages of *E. cuniculi* life-cycle. At the sporoblast stage, the enzyme accumulates especially in paramural bodies which are convolutions of the plasma membrane opened to the wall. The identification of an *EcCDA* homologue in the insect parasite *Antonospora locustae* (ex *Nosema locustae*) suggests a widespread distribution of this enzyme among Microsporidia. This characterization of a new microsporidian surface protein creates new perspectives to understand spore wall formation and spore resistance.

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

**Keywords:** Microsporidia; *Encephalitozoon cuniculi*; Chitin deacetylase

## 1. Introduction

The obligate intracellular Microsporidia are spore-forming amitochondriate unicellular eukaryotes that can parasitize a wide variety of animals ranging from protists to mammals. Some species, including *Encephalitozoon cuniculi*, are opportunist human pathogens [1] with a high seroprevalence in immunocompetent popu-

lations [2–4]. Phylogenetic analyses based on some conserved proteins suggest that Microsporidia are related to fungi [5,6]. Complete sequencing of the *E. cuniculi* genome [7] has also provided data consistent with a fungal origin of Microsporidia [8].

Microsporidian spores display an original invasion mechanism, involving a coiled structure (the polar tube) that can be quickly extruded to transfer the spore content or sporoplasm into a target cell. The parasite enters a proliferation phase or merogony and the resulting meronts then undergo a differentiation step (sporogony). They transform into sporonts (beginning of cell wall

\* Corresponding author. Tel.: +33 4 73 40 74 34; fax: +33 4 73 40 76 70.

E-mail address: catherine.texier@univ-bpclermont.fr (C. Texier).



formation), then into sporoblasts (beginning of polar tube elaboration) and finally into mature spores protected by a thick wall. The microsporidian spore wall plays a major role in the protection against environmental stresses, allowing long-term survival of the parasite after its release from the host cell [9,10]. While the classical concept of the parasite penetration mode involves the polar tube acting as a needle–syringe system, microsporidian spores may also be internalized via attachment to the host plasma membrane then endocytosis [11]. If this occurs, the induced endocytosis process should depend on interactions with the wall.

The microsporidian spore wall usually comprises a proteinaceous outer electron-dense layer (exospore) and a chitinous inner electron-transparent layer (endospore) [12–14], which is of fibrillar nature and connected to the plasma membrane [13,15]. The stippled deposit of an electron-dense material on the outer side of the parasite plasma membrane is the main visible feature of the transition from merogony to sporogony [16,17]. This deposit turns into a continuous layer in young sporoblasts, and should be at the origin of the exospore. The electron-transparent endospore is built in later stages and thus appears as a cytological marker of the mature spore.

Few studies have been dedicated to the characterisation of spore wall proteins. Immunostaining with polyclonal antibodies raised against partially purified *E. cucurbitae* proteins has revealed a 30-kDa antigen within the exospore and a 33-kDa one close to the spore plasma membrane [18]. Several monoclonal antibodies have also been reported to be specific to spore wall antigens [19–21]. To date, the knowledge of the primary structure of such antigens is limited to two exospore components with unknown function: SWP1 (50 kDa) in both *E. cucurbitae* [22] and *Encephalitozoon intestinalis* [23], and SWP2 (150 kDa) in *E. intestinalis* only [23]. Here, we provide evidence for an *E. cucurbitae* surface protein belonging to the polysaccharide deacetylase family that should be active in chitin deacetylation during all sporogonic stages.

## 2. Materials and methods

### 2.1. Microsporidian spore production and protein extraction

The *E. cucurbitae* spores were produced on Madin–Darby Canine Kidney (MDCK) or Human Foreskin Fibroblast (HFF) cells as previously described [18].

Parasites ( $10^9$ – $10^{10}$ ) were disrupted in lysis buffer containing 100 mM DTT, 4% CHAPS and 0.2% SDS, by repeated cycle of freezing–thawing and sonication (Deltasonic 1320, 300 W, 28 kHz). Disrupted spores were then incubated in the extraction solution (7 M urea, 2 M thiourea, 100 mM DTT, 4% CHAPS, 0.2% SDS)

for 6 h at room temperature. Soluble proteins were collected and stored at  $-20\text{ }^{\circ}\text{C}$ .

### 2.2. Recombinant ECU11\_0510 protein expression

The ECU11\_0510 protein was expressed as a truncated form in fusion with Glutathione-S-transferase (GST) at the N-terminal and an 8xHis-tag at the C-terminal extremity. PCR primers were designed (5'-CGGAATTCTTTGTGGACGGTCCCTGTC-3' and 5'-CGCTCGAGTCCCTATGCTCTCAACGCC-3' containing an *EcoRI* and a *XhoI* restriction site, respectively) to amplify a 624-bp genomic DNA fragment corresponding to the amino acid region 32–239 of the 254-amino acid ECU11\_0510 protein. PCR reactions were performed according to standard conditions with an annealing temperature of  $55\text{ }^{\circ}\text{C}$ , and the purified product was cloned in frame with GST into the prokaryotic expression vector pGEX-4T1 modified to provide a C-terminal 8xHis-tag after the *XhoI* site. The resulting recombinant plasmid was introduced in the *Escherichia coli* BL21+ strain. After induction with IPTG, bacterial proteins were extracted with Laemmli solution and analysed by SDS–PAGE.

### 2.3. Gel electrophoresis and Western blot analysis

Protein samples were analysed by SDS–PAGE on 12% polyacrylamide gels. For two-dimensional (2-D) gel electrophoresis, protein isoelectrofocalisation was performed along linear immobilized pH gradient strips of 7 and 17 cm (BioRad) using a rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS) supplemented with 2 mM tributyl phosphine (TBP) and 0.5% ampholytes (BioRad) with the IPGPhor apparatus (Amersham). Before standard SDS–PAGE on 12% polyacrylamide gel, strips were equilibrated with 50 mM Tris–HCl pH 8.8, 6 M urea, 30% glycerol, 2% SDS containing 100 mM DTT, and then 135 mM iodoacetamide. For mass spectrometry analysis, gels were fixed in a 7.5% acetic acid/30% ethanol solution, stained with Coomassie Blue (BioRad) and destained with 30% ethanol. For immunoblotting studies, proteins were transferred onto polyvinylidene difluoride (PVDF) membrane (Millipore). After blocking for 1 h in PBS–5% skim milk, membranes were incubated for 3 h with appropriate dilutions of polyclonal mouse antibodies in PBS 0.1% Triton X-100, treated with a 1:10000 dilution of alkaline phosphatase-conjugated goat anti-mouse IgG (H + L) (Promega) for 1 h at room temperature and revealed with NBT/BCIP (Promega).

### 2.4. Tryptic digestion for mass spectrometry analysis

Protein spots were manually excised from Coomassie blue-stained 2-D gels and transferred in 96-well-microtitration plates. Sample preparation was fully



automated (Mass Prep Station Multiprobe II, Micro-mass). Briefly, excised gels were washed with destaining solutions (25 mM  $\text{NH}_4\text{HCO}_3$  for 15 min and 50% (v/v) acetonitrile containing 25 mM  $\text{NH}_4\text{HCO}_3$  for 15 min). After dehydration with 100% acetonitrile and drying, proteins were reduced in 25 mM  $\text{NH}_4\text{HCO}_3$ , 10 mM DTT for 1 h at 37 °C, and alkylated in 25 mM  $\text{NH}_4\text{HCO}_3$ , 55 mM iodoacetamide for 30 min at 37 °C. After washing with destaining solutions and pure water, gels were shrunk again with 100% acetonitrile. Depending on protein amount, 2–3  $\mu\text{L}$  of 0.1  $\mu\text{g}/\mu\text{L}$  modified trypsin (Promega, sequencing grade) in 25 mM  $\text{NH}_4\text{HCO}_3$  were added over the gel spots for 30 min. Seven  $\mu\text{L}$  of 25 mM  $\text{NH}_4\text{HCO}_3$  were then added to cover the gel spots and digestion allowed to continue for 5 h at 37 °C.

### 2.5. MALDI-TOF-MS analysis and database search

For MALDI-TOF-MS analysis, 0.5  $\mu\text{L}$  of trypsin peptide mixture was mixed with 0.5  $\mu\text{L}$  of  $\alpha$ -cyano-4-hydroxycinnamic acid at half saturation in 60% acetonitrile/0.1% (v/v) trifluoroacetic acid (TFA). The resulting solution was automatically spotted on the target plate and rinsed by 2  $\mu\text{L}$  of 0.1% TFA for 30 s, liquid being then blown off by pressurized air. Peptide mixtures were analysed with a MALDI-TOF mass spectrometer (Autoflex, Bruker Daltonik) in reflector mode over a mass range of 0–4200 Da. For each sample, spectrum acquisition and annotation were done automatically. Spectra from 200 shots at several positions were combined to generate a peptide mass fingerprint for database searches. Spectra were calibrated using a mixture of angiotensin II (1046.54 Da), substance P (1347.74 Da), bombesin (1619.82 Da) and ACTH (18–39) (2465.20 Da) as an external standard. The *E. cucurbitae* specific database was explored with an intranet version of Mascot 1.7. All peptide masses were assumed to be monoisotopic and  $[\text{M} + \text{H}]^+$  (protonated molecular ions). Peptide modifications allowed during the search were cysteine carbamidomethylation, protein N-acetylation and methionine oxidation. The maximum number of missed cleavages was set to 1 and the mass tolerance to  $\pm 150$  ppm.

### 2.6. Polyclonal antibody production

Antiserum to *E. coli* expressed recombinant protein was produced in BALB/c mice from crushed SDS-PAGE protein bands. Every two weeks, mice were injected intraperitoneally with samples homogenized with Freund's adjuvant (Sigma). Serum was collected 2 weeks after the last injection and stored at  $-20$  °C.

### 2.7. Transmission electron microscopy immunolabelling

*E. cucurbitae*-infected MDCK cells were pelleted, then fixed with 4% paraformaldehyde and 0.1% glutaralde-

hyde in 0.1 M phosphate buffer pH 7.2. After washing, the pellet was impregnated with a cryoprotective agent (25% glycerol, 5% DMSO in phosphate buffer) and frozen in pasty nitrogen before storage in liquid nitrogen. The 85 nm ultrathin sections were performed with the FCS system on the Leica UltraCut S ultramicrotome. The sections were picked on collodion-coated nickel grids and stored on 2.3 M sucrose before immunolabelling. After blocking for 30 min with 2% BSA in PBS, sections were incubated for 1 h30 with various dilutions of polyclonal antibodies, and then for 1 h with a 1:100 dilution of protein A or goat anti-mouse antibody conjugated with 5 nm colloidal gold particles (Sigma). After washing, the grids were stained and embedded with 0.5% aqueous uranyl acetate 1.6% methylcellulose for 10 min prior to examination with a JEOL 1200EX transmission electron microscope.

### 2.8. Sequence analysis

Physico-chemical parameters of the protein sequence were calculated with ProtParam ([www.expasy.org](http://www.expasy.org)). The SignalP program version 2 ([www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)), the TMPRED program ([www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)) and the PSORT II algorithm ([www.psорт.nibb.ac.jp/form2.html](http://www.psорт.nibb.ac.jp/form2.html)) were used to predict signal peptide, transmembrane helices and protein localisation, respectively. Consensus secondary structure prediction were performed on NPS@ Network Protein Sequence Analysis (<http://npsa-pbil.ibcp.fr>) [24] using the DSC, GOR4, PHD and SOPMA algorithms. Search for protein homologies and protein domains were done using the BLAST program ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and the Pfam database ([www.sanger.ac.uk/Software/Pfam/search.shtml](http://www.sanger.ac.uk/Software/Pfam/search.shtml)), respectively. Multiple sequence alignments were performed with Dialign program ([bibiserv.techfak.uni-bielefeld.de/dialign/](http://bibiserv.techfak.uni-bielefeld.de/dialign/)) and analyzed for amino acid identity and similarity with Boxshade ([www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

## 3. Results and discussion

Mouse polyclonal antibodies (PAb) were previously raised against several SDS-PAGE separated *E. cucurbitae* spore proteins and analysed by Western blotting and indirect immunofluorescence [18]. One of these, directed against a 33-kDa protein band (anti-33 kDa PAb), showed a labelling close to the outer face of the parasite plasma membrane. This prompted us to identify the protein target of this anti-33 kDa PAb.

### 3.1. The anti-33 kDa PAb recognizes the ECU11\_0510 protein

To get a better separation of the targeted protein, we performed 2-D electrophoresis of proteins extracted





from *E. cuniculi* spores (Fig. 1(a)). After Western blotting, the anti-33 kDa PAb highlighted two major regions at 33 and 55 kDa (Fig. 1(b)). MALDI-TOF mass spec-

trometry data obtained with the 33-kDa region (Fig. 1(c)) matched the ECU11\_0510 protein (SWISS-PROT: Q8SU65) in the *E. cuniculi* proteome database deduced

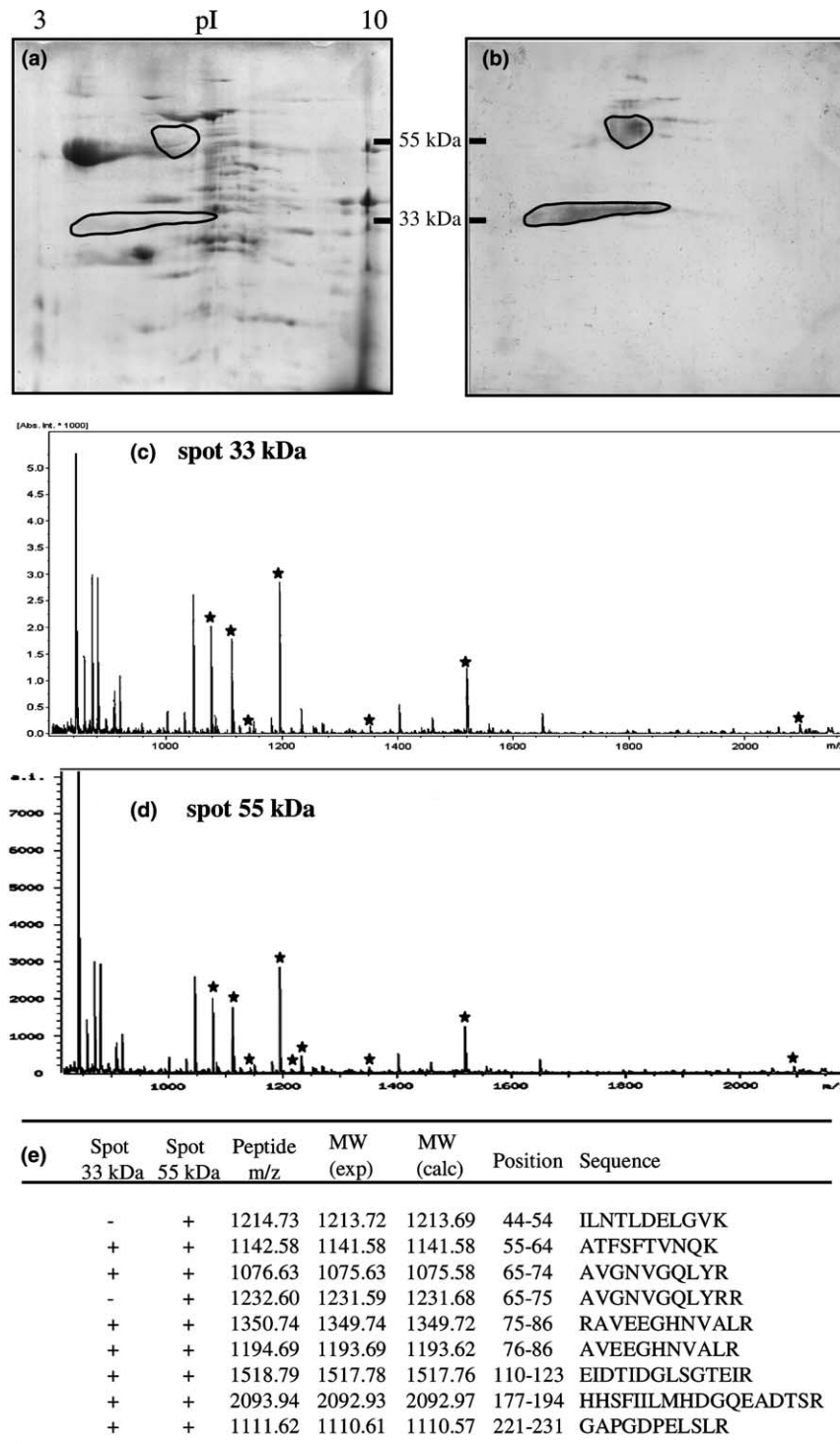


Fig. 1. Migration on 2-D gel and identification by mass spectrometry of the anti-33 kDa PAb target. *E. cuniculi* spore proteins were separated by 2-D electrophoresis and either stained by Coomassie Blue (a) or subjected to Western blot immunodetection with a 1:4000 dilution of the anti-33 kDa PAb (b). Trypsin digests of the 33-kDa spot (c) and the 55-kDa spot (d) was analyzed by MALDI-TOF-MS. Seven (nine) experimental peptides for the 33-kDa spot (55-kDa one), labelled with stars in (c) (in (d)), matched the ECU11\_0510 protein with a score of 60 (100) and sequence coverage of 29% (33.8%) (e). Only one miscleavage was observed. Differences between experimental and calculated molecular weights (MW) do not exceed 0.09 Da.



from genome sequencing [7] (Fig. 1(e)). Identical data were obtained with the 55-kDa spot (Fig. 1(d) and (e)). We produced a recombinant form of the *EcCDA* protein which was slightly truncated to exclude predicted transmembrane helices (see below). The purified recombinant protein reacted with the anti-33 kDa PAb (Fig. 2) and was also used to produce specific PAb in BALB/c mice (anti-11\_0510 PAb). These new antibodies exhibited a similar reaction pattern as the anti-33 kDa PAb either on the recombinant ECU11\_0510 protein or on *E. cucuruli* extract (Fig. 2).

### 3.2. The ECU11\_0510 protein is predicted to be a chitin deacetylase (*EcCDA*)

As supported by genome data [7], the ECU11\_0510 open reading frame is a 765-bp single-copy gene located on chromosome 11 which codes for a 254-amino acid protein (28.1 kDa, pI 4.56). As summarized in Fig. 3(a), the extreme N-terminal region displays the characteristics of a signal peptide with a predicted cleavage site between positions 15 (Gly) and 16 (Ser) corroborated by the TMPred prediction of a transmembrane helix (TMH) in the 1–18 region. An additional TMH should be located at the 234–253 C-terminal extremity in an outside–inside orientation, suggesting that the protein attached to the membrane via this TMH has a large extracellular domain. This is in accordance with the efficient recovery of ECU11\_0510 (not shown) when using an improved membrane protein solubility mixture characterized by the amidosulfobetaine ASB14 detergent and high chaotrope molarity [25,26]. PSORT program

also predicts an extracellular location including cell wall (66.7%,  $k = 9/23$ ).

The ECU11\_0510 protein contains the PfamA conserved domain 15022 representative of the polysaccharide deacetylase family (Fig. 3(b)). This family includes *Rhizobium* NodB (nodulation protein B), endoxylanases and fungal chitoooligosaccharide deacetylases. The presence of chitin in the microsporidian spore wall has been unambiguously demonstrated [12,13]. Moreover, chitin is the sole polysaccharide that metabolism reconstruction predicts to be synthesized in *E. cucuruli* [7]. *N*-acetylglucosamine 6-P is converted into *N*-acetylglucosamine 1-P, then into UDP-*N*-acetylglucosamine via the activities of phosphoacetylglucosamine mutase and UDP-GlcNAc pyrophosphorylase. Chitin synthase carries out the polymerisation of *N*-acetylglucosamine into chitin chains. We thus assume that the ECU11\_0510 protein may be a chitin deacetylase enzyme (*EcCDA*) which deacetylates chitin into chitosan. Enzymatic assays with a purified recombinant protein that was expressed in *E. coli*, were unsuccessful. The recombinant protein may lack some post-translational modifications required for biological activity as it is the case for the yeast CDA2p which loses its activity after deglycosylation [27].

Interestingly, strong homologues of the chitosan pathway enzymes, including CDA, were identified in the partially sequenced genome of an insect Microsporidia, *Nosema locustae* renamed *Antonosporea locustae* [28] (“*Nosema locustae* Genome Project, Marine Biological Laboratory at Woods Hole, funded by NSF award number 0135272”, <http://jbcp.mbl.edu/Nosema/index.html>). The two microsporidian CDA sequences are of nearly identical size and exhibit 40% identity (Fig. 3(b)).

The presence of CDA in Microsporidia provides an additional support to the fungal origin of these intracellular parasites. Chitin deacetylases have been characterized in several fungi [29], including zygomycetes (*Mucor rouxii* [30,31]) and ascomycetes (*Colletotrichum lindemuthianum* [32], *Saccharomyces cerevisiae* [33,34]). These enzymes are monomeric glycoproteins ranging from 12 to 150 kDa in size, with a carbohydrate content ranging from 30% to 67% by weight in *Colletotrichum lindemuthianum* [32]. A well-defined role in cell-wall chitosan biosynthesis has been demonstrated for CDA secreted in the periplasm. In *M. rouxii*, chitin synthase first polymerizes chitin chains, then the *N*-acetamido bonds are hydrolysed by CDA [29]. In *S. cerevisiae*, two CDAs are required for the biosynthesis of a chitosan-containing ascospore wall [35], resistant to stress conditions such as glucosylase or heat treatment [34,36]. In the microsporidian spore wall, chitosan could provide interactions with the other wall components and could thus account for the extreme rigidity and resistance of the parasite wall to chitinase treatments and to different stresses such as freezing, heating and extreme pH [9,10]. In some pathogen fungi, CDA is extracellular

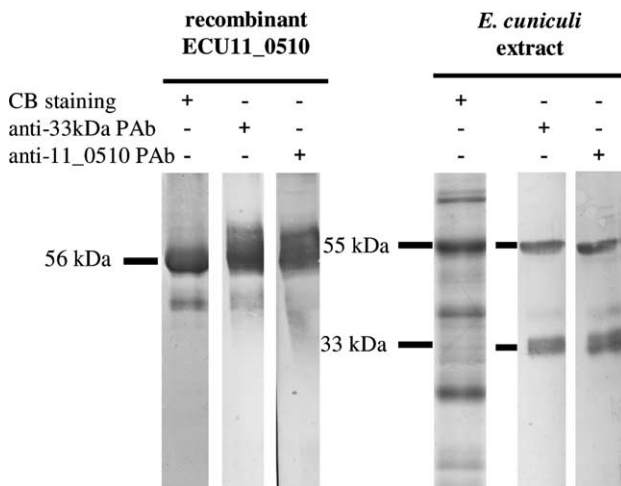


Fig. 2. Immunoblotting with anti-33 kDa PAb and anti-11\_0510 PAb against both the recombinant ECU11\_0510 protein and *E. cucuruli* extract. The 56-kDa recombinant form of the ECU11\_0510 protein (fused in N-terminal with GST and in C-terminal with 8xHis tag) was injected into BALB/c mice to produce polyclonal antibodies (anti-11\_0510 PAb). Both antibodies reacted in a similar way (CB: Coomassie Blue).



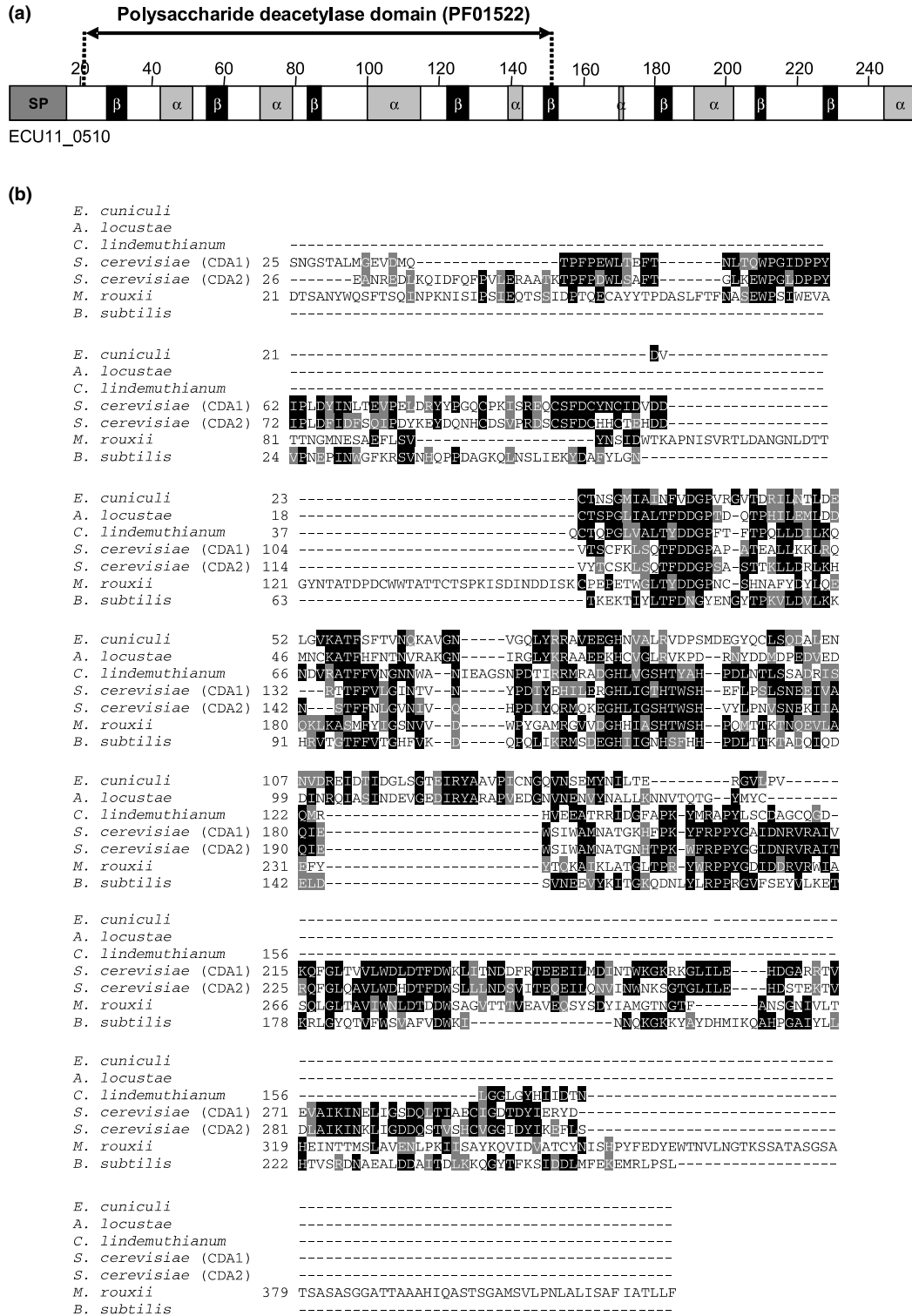


Fig. 3. Sequence analysis of the ECU11\_0510 protein (a) and alignment with chitin deacetylase domains from other organisms (b). (a) The full-length protein (254 amino acids) is represented with different predicted elements: signal peptide (SP), secondary structure consensus with beta strand in black and alpha helix in grey. The ECU11\_0510 polysaccharide deacetylase domain encompasses the position 21–150 and is included in the pfamA01522 domain (score = 11.1,  $E_{\text{value}} = 3.8e - 05$ ). The predicted  $\alpha\beta$  class structure fits with the crystal three-dimensional model of *Bacillus subtilis* polysaccharide deacetylase (PDB: 1NY1). (b) Sequence alignment of chitin deacetylase domains from the mammalian Microsporidia *Encephalitozoon cuniculi* (ECU11\_0510), the insect Microsporidia *Antonospora* (formerly *Nosema*) *locustae* (sequence from contig\_159\_1742:2506, <http://jbcpc.mbl.edu/Nosema/index.html>, pfam: score = 56.4,  $E_{\text{value}} = 8.3e - 14$ ), the fungi *Colletotrichum lindemuthianum*, *Saccharomyces cerevisiae*, *Mucor rouxii* and the bacterium *Bacillus subtilis*. Identical and similar residues are highlighted in black and grey respectively. The microsporidian domain is markedly reduced but presents a strong homology with the other proteins on its 60–70 first amino acids.



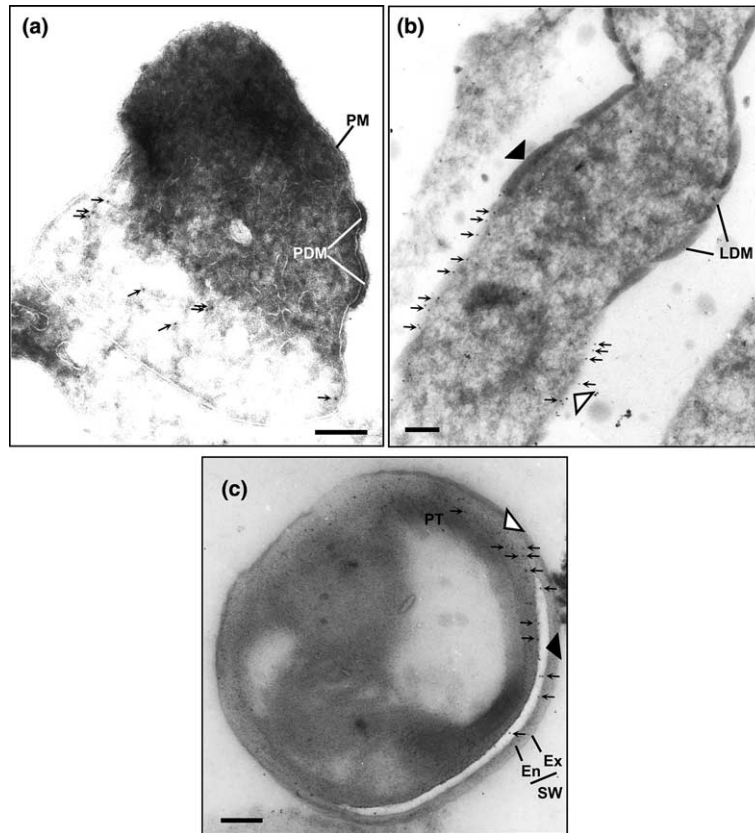


Fig. 4. Immunogold electron microscopy of cryosections of different *E. cuniculi* developmental stages using the anti-11\_0510 PAb and secondary antibody conjugated with 5 nm colloidal gold (bars = 200 nm). The labelling (small arrows) is observed as early as the meront-to-sporont transition in the cytoplasm (a). In dividing sporonts (b), gold particles are found at the periphery of the cell. Finally, the antibodies react with the innermost part of the endospore of the mature spore (c). Black arrowheads highlight transversal sections of the spore wall, while the white ones show grazing section planes. PM: plasma membrane. PDM: patches of dense material. LDM: lamellar dense material. SW: spore wall. Ex: exospore. En: endospore. PT: polar tube.

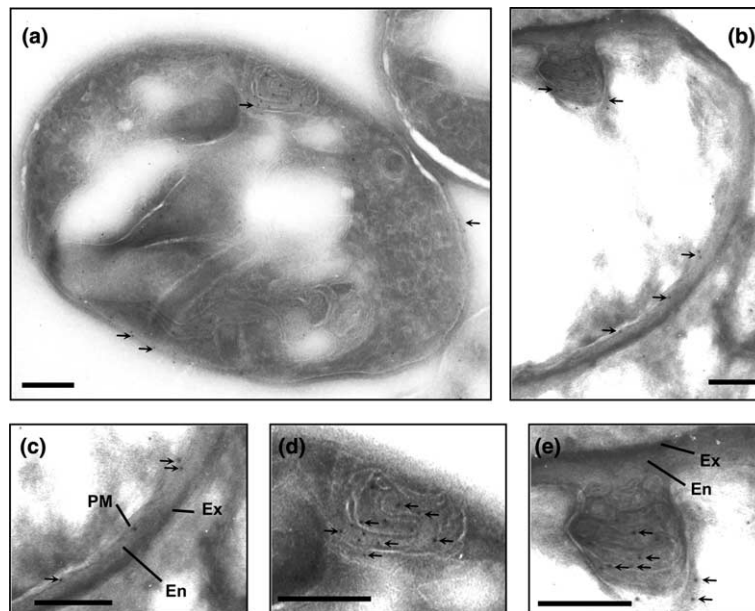


Fig. 5. Immunogold electron microscopy of cryosections of *E. cuniculi* sporoblasts using the anti-11\_0510 PAb and secondary reagent labelled with 5 nm colloidal gold (bars = 200 nm). In sporoblasts (a,b), the labelling (small arrows) is associated with the constructing wall (c) and concentrates in peripheral structures (d,e). PM: plasma membrane. Ex: exospore. En: endospore.





and is thought to be implicated in the bypass of host-defence mechanisms during *C. lindemuthianum* plant invasion [32] or *Aspergillus fumigatus* mammalian infestation [37].

### 3.3. The EcCDA protein presents two isoforms

In the *E. cucuruli* extract, both PAb recognized the two ECU11\_0510 isoforms at 33 and 55 kDa (Fig. 2). Several fungi also exhibit multiple CDA glycoproteins: two, three, four and up to five isoforms were described in *Mucor rouxii* [31], *Metarhizium anisopliae* [38], *Rhizopus nigricans* [39] and *Uromyces viciae-fabae* [40], respectively. The *E. cucuruli* genome sequence predicts only few genes involved in protein glycosylation, all of them being related to O-mannosylation [7]. Sixteen serine and eleven threonine residues are potential O-glycosylation sites in EcCDA. While O-mannoproteins were recently detected in the spore wall of *E. cucuruli* and in the SDS-PAGE 55-kDa region (Méténier, G., Mazet, M. and Taupin, V., personal communication), concanavalin-A labelling did not co-localize with the ECU11\_0510 protein on 2-D gels (not shown).

### 3.4. EcCDA is located at the plasma membrane–endospore interface of the parasite

Immunocytochemical experiments at the ultrastructural level with the anti-11\_0510 PAb, revealed a slight gold labelling of the cytoplasm of stages representative of the meront-to-sporont transition (Fig. 4(a)). Such cells are indeed still tightly attached to the parasitophorous vacuole membrane by one side, and their engagement in early sporogony is marked by patches of dense material abutting on the plasma membrane. As sporogony progresses, the labelling tends to accumulate at the periphery of the parasitic cells. In longitudinal sections of dividing sporonts (Fig. 4(b)), gold particles are scattered just beneath the lamellar dense material deposit on the plasma membrane. In mature spores (Fig. 4(c)), gold particles are mainly seen at the plasma membrane–endospore interface. In sporoblasts, the EcCDA is also present in the inner part of the cell wall tightly associated with the plasma membrane (Fig. 5(a)–(c)). This location is consistent with biochemical and sequence data, and suggests a role in cell-wall formation. Late sporogony is characterized by the thickening of the chitin-containing endospore [12], but the mechanism of chitin deposition is unknown. Since the EcCDA enzyme is regularly distributed at the plasma membrane throughout sporogony, chitin may be produced all around the cell.

Late sporoblasts frequently harbour a membrane-rich structure (300 × 200–250 nm) which reacts with EcCDA-specific PAb (Fig. 5(a), (b), (d) and (e)), and which opens to the wall. Gold particles are clearly asso-

ciated with membranes of this structure that is filled with an electron-dense material similar to that of the endospore. Fig. 5(e) suggests that this material is flowing between the plasma membrane and the future exospore. In different Microsporidia such as *Pleistophora* sp., *Stempellia* sp. and *Ameson pulvis*, a similar membranous whorl termed “scindosome” or “paramural body” has been observed at the furrowing sites during division stage of the parasite [13,41].

Since the sporoblast decreases in size as its differentiation proceeds, the *E. cucuruli* paramural body could be interpreted as an endocytic recuperation of membrane/endospore material. Alternatively, chitin might be synthesized and deacetylated within the paramural body before being incorporated to the wall by an exocytosis-like process. As a result, this structure would be the equivalent of the fungal chitosome, an organelle which is specialized for chitin biosynthesis [42,43]. The chitosome hypothesis should be tested through further studies on the expression and location of microsporidian chitin synthase (ECU01\_1390).

### Acknowledgments

D. Brosson was supported by a grant from “Ministère de l’Éducation, de la Recherche et de la Technologie”. We thank G. Méténier for helpful discussion and critical reading of this manuscript. We also thank B. Chebance and I. Wawrzyniak for *E. cucuruli* cell culture and R. Guerry and A. Voldoire for technical advice.

### References

- [1] Kotler, D.P. and Orenstein, J.M. (1999) Clinical syndromes associated with microsporidiosis. In: *The Microsporidia and Microsporidiosis* (Wittner, Ed.), pp. 258–292. ASM Press, Washington, DC.
- [2] van Gool, T., Vetter, J.C., Weinmayr, B., Van Dam, A., Derouin, F. and Dankert, J. (1997) High seroprevalence of *Encephalitozoon* species in immunocompetent subjects. *J. Infect. Dis.* 175, 1020–1024.
- [3] Raynaud, L., Delbac, F., Broussolle, V., Rabodonirina, M., Girault, V., Wallon, M., Cozon, G., Vivarès, C.P. and Peyron, F. (1998) Identification of *Encephalitozoon intestinalis* in travelers with chronic diarrhea by specific PCR amplification. *J. Clin. Microbiol.* 36, 37–40.
- [4] van Gool, T., Biderre, C., Delbac, F., Wentink-Bonnema, E., Peek, R. and Vivarès, C.P. (2004) Serodiagnostic studies in an immunocompetent individual infected with *Encephalitozoon cucuruli*. *J. Infect. Dis.* 189, 2243–2249.
- [5] Keeling, P.J., Luker, M.A. and Palmer, J.D. (2000) Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. *Mol. Biol. Evol.* 17, 23–31.
- [6] Keeling, P.J. (2003) Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. *Fungal. Genet. Biol.* 38, 298–309.
- [7] Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretailade, E., Brottier, P.,



- Wincker, M., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J. and Vivarès, C.P. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414, 450–453.
- [8] Thomarat, F., Vivarès, C.P. and Gouy, M. (2004) Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. *J. Mol. Evol.* 59, 780–791.
- [9] Shadduck, J.A. and Polley, M.B. (1978) Some factors influencing the in vitro infectivity and replication of *Encephalitozoon cuniculi*. *J. Protozool.* 25, 491–496.
- [10] Koudela, B., Kucerova, S. and Hudcovic, T. (1999) Effect of low and high temperatures on infectivity of *Encephalitozoon cuniculi* spores suspended in water. *Folia Parasitol. (Praha)* 43, 171–174.
- [11] Vivarès, C.P. and Méténier, G. (2001) The microsporidian *Encephalitozoon*. *Bioessays* 23, 194–202.
- [12] Vavra, J. and Larson, J. (1999) Structure of the microsporidia In: *The Microsporidia and Microsporidiosis* (Wittner, Ed.), pp. 7–84. ASM Press, Washington, DC.
- [13] Vavra, J. (1976) Structure of the microsporidia In: *Comparative Pathology* (B, L.A. and C, T.C., Eds.). *Biology of the Microsporidia*, vol. 1, pp. 1–86. Plenum Press, New York.
- [14] Vivarès, C.P., Loubes, C. and Bouix, G. (1976) Recherches cytochimiques approfondies sur les Microsporidies parasites du Crabe Vert de la Méditerranée *Carcinus mediterraneus* Czerniavsky, 1884. *Ann. Parasitol. Hum. Comp.* 51, 1–14.
- [15] Bigliardi, E., Selmi, M.G., Lupetti, P., Corona, S., Gatti, S., Scaglia, M. and Sacchi, L. (1996) Microsporidian spore wall: ultrastructural findings on *Encephalitozoon hellem* exospore. *J. Eukaryot. Microbiol.* 43, 181–186.
- [16] Pakes, S.P., Shadduck, J.A. and Cali, A. (1975) Fine structure of *Encephalitozoon cuniculi* from rabbits, mice and hamsters. *J. Protozool.* 22, 481–488.
- [17] Hamilton, R.C., Cox, J.C. and Pye, D. (1977) Wall structure of the sporonts of *Encephalitozoon cuniculi* grown in human fibroblasts. *J. Gen. Microbiol.* 98, 305–307.
- [18] Delbac, F., Duffieux, F., David, D., Méténier, G. and Vivarès, C.P. (1998) Immunocytochemical identification of spore proteins in two microsporidia, with emphasis on extrusion apparatus. *J. Eukaryot. Microbiol.* 45, 224–231.
- [19] Beckers, P.J., Derks, G.J., van Gool, T., Rietveld, F.J. and Sauerwein, R.W. (1996) *Encephalocytozoon intestinalis*-specific monoclonal antibodies for laboratory diagnosis of microsporidiosis. *J. Clin. Microbiol.* 34, 282–285.
- [20] Lujan, H.D., Conrad, J.T., Clark, C.G., Touz, M.C., Delbac, F., Vivarès, C.P. and Nash, T.E. (1998) Detection of microsporidia spore-specific antigens by monoclonal antibodies. *Hybridoma* 17, 237–243.
- [21] Visvesvara, G.S., Leitch, G.J., da Silva, A.J., Croppo, G.P., Moura, H., Wallace, S., Slemenda, S.B., Schwartz, D.A., Moss, D., Bryan, R.T. and Pieniazek, N.J. (1994) Polyclonal and monoclonal antibody and PCR-amplified small-subunit rRNA identification of a microsporidian, *Encephalitozoon hellem*, isolated from an AIDS patient with disseminated infection. *J. Clin. Microbiol.* 32, 2760–2768.
- [22] Bohne, W., Ferguson, D.J., Kohler, K. and Gross, U. (2000) Developmental expression of a tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cuniculi*. *Infect. Immun.* 68, 2268–2275.
- [23] Hayman, J.R., Hayes, S.F., Amon, J. and Nash, T.E. (2001) Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*. *Infect. Immun.* 69, 7057–7066.
- [24] Combet, C., Blanchet, C., Geourjon, C. and Deleage, G. (2000) NPS@: network protein sequence analysis. *Trends Biochem. Sci.* 25, 147–150.
- [25] Chevallet, M., Santoni, V., Poinas, A., Rouquie, D., Fuchs, A., Kieffer, S., Rossignol, M., Lunardi, J., Garin, J. and Rabilloud, T. (1998) New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. *Electrophoresis* 19, 1901–1909.
- [26] Molloy, M.P. (2000) Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. *Anal. Biochem.* 280, 1–10.
- [27] Martinou, A., Koutsoulis, D. and Bouriotis, V. (2002) Expression, purification, and characterization of a cobalt-activated chitin deacetylase (Cda2p) from *Saccharomyces cerevisiae*. *Protein Expr. Purif.* 24, 111–116.
- [28] Slamovits, C.H., Williams, B.A. and Keeling, P.J. (2004) Transfer of *Nosema locustae* (Microsporidia) to *Antonospora locustae* n. comb. based on molecular and ultrastructural data. *J. Eukaryot. Microbiol.* 51, 207–213.
- [29] Tsigos, I., Martinou, A., Kafetzopoulos, D. and Bouriotis, V. (2000) Chitin deacetylases: new, versatile tools in biotechnology. *Trends Biotechnol.* 18, 305–312.
- [30] Kafetzopoulos, D., Thireos, G., Vournakis, J.N. and Bouriotis, V. (1993) The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products. *Proc. Natl. Acad. Sci. USA* 90, 8005–8008.
- [31] Kafetzopoulos, D., Martinou, A. and Bouriotis, V. (1993) Bioconversion of chitin to chitosan: purification and characterization of chitin deacetylase from *Mucor rouxii*. *Proc. Natl. Acad. Sci. USA* 90, 2564–2568.
- [32] Tsigos, I. and Bouriotis, V. (1995) Purification and characterization of chitin deacetylase from *Colletotrichum lindemuthianum*. *J. Biol. Chem.* 270, 26286–26291.
- [33] Mishra, C., Semino, C.E., McCreath, K.J., de la Vega, H., Jones, B.J., Specht, C.A. and Robbins, P.W. (1997) Cloning and expression of two chitin deacetylase genes of *Saccharomyces cerevisiae*. *Yeast* 13, 327–336.
- [34] Christodoulidou, A., Bouriotis, V. and Thireos, G. (1996) Two sporulation-specific chitin deacetylase-encoding genes are required for the ascospore wall rigidity of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 271, 31420–31425.
- [35] Briza, P., Ellinger, A., Winkler, G. and Breitenbach, M. (1988) Chemical composition of the yeast ascospore wall. The second outer layer consists of chitosan. *J. Biol. Chem.* 263, 11569–11574.
- [36] Christodoulidou, A., Briza, P., Ellinger, A. and Bouriotis, V. (1999) Yeast ascospore wall assembly requires two chitin deacetylase isozymes. *FEBS Lett.* 460, 275–279.
- [37] Overdijk, B., Van Steijn, G.J. and Odds, F.C. (1999) Distribution of chitinase in guinea pig tissues and increases in levels of this enzyme after systemic infection with *Aspergillus fumigatus*. *Microbiology* 145, 259–269.
- [38] Nahar, P., Ghormade, V. and Deshpande, M.V. (2004) The extracellular constitutive production of chitin deacetylase in *Metarhizium anisopliae*: possible edge to entomopathogenic fungi in the biological control of insect pests. *J. Invertebr. Pathol.* 85, 80–88.
- [39] Trudel, J. and Asselin, A. (1990) Detection of chitin deacetylase activity after polyacrylamide gel electrophoresis. *Anal. Biochem.* 189, 249–253.
- [40] Deising, H. and Siegrist, J. (1995) Chitin deacetylase activity of the rust *Uromyces viciae-fabae* is controlled by fungal morphogenesis. *FEMS Microbiol. Lett.* 127, 207–212.
- [41] Vivarès, C.P. and Sprague, V. (1979) The fine structure of *Ameson pulvis* (Microspora, Microsporidia) and its implication regarding classification and chromosome cycle. *J. Invertebr. Pathol.* 33, 40–52.



- [42] Bracker, C.E., Ruiz-Herrera, J. and Bartnicki-Garcia, S. (1976) Structure and transformation of chitin synthetase particles (chitosomes) during microfibril synthesis in vitro. *Proc. Natl. Acad. Sci. USA* 73, 4570–4574.
- [43] Ruiz-Herrera, J., Sentandreu Ramon, R. and Martinez, J.P. (1992) Chitin biosynthesis in fungi In: *Handbook of Applied Mycology* (Arora, D.K., Elander, R.P. and Mukerji, K.G., Eds.). Fungal Biotechnology, vol. 4. Marcel Dekker, New York.



Peuvel-Fanget I, Polonais V, Brosseau D, Texier C, Kuhn L, Vivarès C, Delbac F.

EnP1 and EnP2, two proteins associated with the *Encephalitozoon cuniculi* endospore, the chitin-rich inner layer of the microsporidian spore wall.

International Journal for Parasitology. *Sous presse*.







# EnP1 and EnP2, two proteins associated with the *Encephalitozoon cuniculi* endospore, the chitin-rich inner layer of the microsporidian spore wall

Isabelle Peuvel-Fanget<sup>a,1</sup>, Valérie Polonais<sup>a</sup>, Damien Brosseau<sup>a</sup>, Catherine Texier<sup>a</sup>,  
Lauriane Kuhn<sup>b</sup>, Pierre Peyret<sup>a</sup>, Christian Vivarès<sup>a</sup>, Frédéric Delbac<sup>a,\*</sup>

<sup>a</sup> Equipe Parasitologie Moléculaire et Cellulaire, LBP, UMR CNRS 6023, Université Blaise Pascal, 24 Avenue des Landais, 63177 Aubière Cedex, France

<sup>b</sup> Laboratoire de Chimie des Protéines, DBMS, CEA, 38054 Grenoble, France

Received 28 July 2005; received in revised form 28 September 2005; accepted 25 October 2005

## Abstract

Microsporidia are obligate intracellular parasites forming environmentally resistant spores that harbour a rigid cell wall. This wall comprises an outer layer or exospore and a chitin-rich inner layer or endospore. So far, only a chitin deacetylase-like protein has been shown to localize to the *Encephalitozoon cuniculi* endospore and either one or two proteins have been clearly assigned to the exospore in two *Encephalitozoon* species: SWP1 in *E. cuniculi*, SWP1 and SWP2 in *Encephalitozoon intestinalis*. Here, we report the identification of two new spore wall proteins in *E. cuniculi*, EnP1 and EnP2, the genes of which are both located on chromosome I (ECU01\_0820 and ECU01\_1270, respectively) and have no known homologue. Detected by immunoscreening of an *E. cuniculi* cDNA library, *enp1* is characterized by small-sized 5' and 3' untranslated regions and is highly expressed throughout the whole intracellular cycle. The encoded basic 40 kDa antigen displays a high proportion of cysteine residues, arguing for a significant role of disulfide bridges in spore wall assembly. EnP2 is a 22 kDa serine-rich protein that is predicted to be O-glycosylated and glycosylated phosphatidyl inositol-anchored. Although having been identified by mass spectrometry of a dithiothreitol-soluble fraction, this protein contains only two cysteine residues. Mouse polyclonal antibodies were raised against EnP1 and EnP2 recombinant proteins produced in *Escherichia coli*. Our immunolocalisation data indicate that EnP1 and EnP2 are targeted to the cell surface as early as the onset of sporogony and are finally associated with the chitin-rich layer of the wall in mature spores.

© 2005 Australian Society for Parasitology Inc Published by Elsevier Ltd. All rights reserved.

**Keywords:** Microsporidia; *Encephalitozoon cuniculi*; Spore wall; Endospore; Cysteine; Immunolocalization

## 1. Introduction

Microsporidia are obligate intracellular parasites found in a wide variety of animal hosts (Weiss, 2001). Species from several genera, including *Encephalitozoon*, have been recognized as opportunistic human pathogens, some of these causing severe diarrhoea in AIDS patients (Bryan and Schwartz, 1999). However, serological studies indicated that microsporidiosis may also occur in immunocompetent hosts (van Gool et al., 1997; 2004). Microsporidia have 70S ribosomes and lack typical eukaryotic organelles such as mitochondria and peroxysomes. This likely results from a highly regressive

evolution, given that phylogenetic analyses support the conception that microsporidia are closely related to fungi (Keeling, 2003; Thomarat et al., 2004).

The extracellular infective stage of microsporidia is the spore which is surrounded by a rigid cell wall and contains a remarkable invasion apparatus. A structure coiled within the spore, called the polar tube, can be suddenly extruded then used to transfer the sporoplasm into the cytoplasm of a host cell. The first part of the intracellular development corresponds to a proliferative phase (merogony), during which meronts divide by binary or multiple fission. The second part called sporogony begins when meronts transform into sporonts that are characterized by the deposition of an electron-dense material on the plasma membrane. Sporonts divide into sporoblasts in which the formation of the thick wall is associated with the differentiation of the invasion apparatus. In *Encephalitozoon* species, all the development takes place inside a parasitophorous vacuole. After completion of the maturation of sporoblasts

\* Corresponding author. Tel.: +33 4 73 40 78 68; fax: +33 4 73 40 76 70.  
E-mail address: frederic.delbac@univ-bpclermont.fr (F. Delbac).

<sup>1</sup> Present address: Biologie Cellulaire et Moléculaire de la Sécrétion CNRS UPR 1929, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France.



into spores, the rupture of host cells may lead to the release of spores in the environment.

The spore wall consists of two major layers: (i) an electron-dense outer layer of 25–30 nm, the exospore, which is principally proteinaceous, and (ii) an electron-lucent inner layer of 30–35 nm, the endospore, which contains chitin and proteins (Bigliardi et al., 1996). This wall makes the spore resistant to various environmental stresses and its aqueous permeability may play a role in the control of the polar tube discharge activation (Frixione et al., 1997).

The microsporidian spore wall has been shown to be extremely resistant to dissociation in the absence of thiol-reducing agents such as dithiothreitol (DTT) or 2-mercaptoethanol (Keohane et al., 1996; Hayman et al., 2001). Few data about the protein components of this structure are available, although a variety of antibodies directed against microsporidia were frequently found to be specific to spore wall antigens (Beckers et al., 1996; Enriquez et al., 1997; Delbac et al., 1998a; Lujan et al., 1998; Prigneau et al., 2000). The sequences of either one or two exospore proteins have been identified in two *Encephalitozoon* species: spore wall protein (SWP)1 and SWP2 in *Encephalitozoon intestinalis* (Hayman et al., 2001) and only SWP1 in *Encephalitozoon cuculiculi* (Bohne et al., 2000). These are cysteine-rich polypeptides with similar N-terminal domains and with repeating amino acid units at their C-terminus. The localization of a potential chitin deacetylase at the endospore-plasma membrane interface has been recently demonstrated in *E. cuculiculi* (Brosson et al., 2005). Taking advantage of the availability of the full sequence of the *E. cuculiculi* nuclear genome (2.9 Mbp; Katinka et al., 2001), we report here the characterization of two different *Encephalitozoon* Endospore Proteins (EnPs) after either immunoscreening of a cDNA library for EnP1 or mass spectrometry analysis of a DTT-soluble protein extract for EnP2.

## 2. Materials and methods

### 2.1. Microsporidian spore production and purification

*Encephalitozoon cuculiculi* was grown in vitro on MDCK (Madin-Darby canine kidney) or HFF (human foreskin fibroblast) cells, in minimum essential medium (MEM) supplemented with 5% FCS and 2 mM glutamine at 37 °C in a 5% CO<sub>2</sub> atmosphere. Spore-containing supernatants were collected every 72 h. Spores were harvested (18,000×g, 2 min), washed and stored in PBS at 4 °C.

### 2.2. Antibody production

Polyclonal antibodies to *Escherichia coli* expressed recombinant proteins were produced in BALB/c mice from SDS-PAGE separated protein bands. Protein bands were excised from Coomassie blue-stained gel and crushed in PBS with a Potter apparatus. Mice were then injected intraperitoneally with samples homogenized with FCA for the first injection and Freund incomplete adjuvant for the next

injections (at days +14, +21 and +28). Sera were collected 1 week after the last injection and stored at –20 °C.

The monoclonal antibody Ec102 was raised against a 55 kDa band separated by SDS-PAGE (Delbac et al., 1998a) and was shown to be directed against both *E. cuculiculi* polar tube protein (PTP)1 (Delbac et al., 1998b) and PTP2 (Delbac et al., 2001).

### 2.3. SDS-PAGE and Western blotting

Total *E. cuculiculi* proteins were extracted in a lysis buffer containing 2.5% SDS and 100 mM DTT. Spores were disrupted by 10 steps of freezing and thawing in liquid nitrogen, followed by sonications (10×30 s on ice) and boiled for 15 min. After centrifugation for 5 min at 18,000×g, the residual insoluble pellet was incubated for a further 72 h in 200 mM DTT, 2.5% SDS and boiled for 15 min. Protein samples were then analyzed by SDS-PAGE. After electrophoresis, proteins were stained with Coomassie blue or transferred onto polyvinylidene difluoride (PVDF) membrane (Millipore). For immunological detection, membranes were saturated in PBS, 5% milk and incubated for 3 h with an appropriate dilution of mouse antibodies (1:100–1:1,000). After washing, membranes were reacted with horseradish peroxidase-conjugated goat anti-mouse IgG, A and M (Sigma) and developed with the ECL Western blot detection kit (Amersham).

### 2.4. Indirect immunofluorescence assay

*Encephalitozoon cuculiculi* infected-HFF cells grown on glass slides were fixed with 4% paraformaldehyde, 0.05% glutaraldehyde for 20 min at room temperature then washed in PBS, 0.1 M glycine. Cells were permeabilized with PBS, 0.5% Triton X-100 and blocked with 5% skim milk in PBS. Slides were incubated for 1 h with mouse sera diluted in PBS, 0.1% Triton X-100, washed and incubated further 1 h with a 1:1000 dilution of Alexa 488 conjugated goat anti-mouse IgG (Molecular Probes) or a 1:400 dilution of TRITC-conjugated goat anti-mouse IgG (Sigma). Preparations were examined with a DMR Leica epifluorescence microscope.

### 2.5. Transmission electron microscopy immunolabeling

Intracellular parasites grown on MDCK cells were fixed with 4% paraformaldehyde, 0.05% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.2. Cells were then scraped off, dehydrated and embedded in Unicryl resin. Ultra-thin sections were blocked with PBS, 0.1% BSA, incubated for 3 h with mouse sera diluted in PBS, 0.25% BSA, washed and incubated further 1 h with a 1:100 dilution of goat anti-mouse IgG gold-conjugated (Sigma). After staining with 4% uranyl acetate, preparations were examined with a JEOL 1200X transmission electron microscope.



## 2.6. *EnP1 and EnP2 expression in Escherichia coli*

*Encephalitozoon cuniculi* DNA was released by boiling purified spores at 100 °C for 10 min. *Enp1* and *enp2* coding sequences were amplified by PCR to introduce in 5' and 3' of the amplification products one *Bam*H I and one *Eco*R I restriction site, respectively. PCR amplifications were performed using direct primers P1expD (5'-gtccggatccccatagtgccgtgtag-3') and P2expD (5'-cgggatccagcagtgccctacc-3') and reverse primers P1expR (5'-gtcggaaatcggcattgat-tacatgcag-3') and P2expR (5'-cggaaatccatcacaattgagaacagagc-3') according to standard conditions (Eurobio). PCR products were analyzed by electrophoresis on 1% agarose gel, purified using a Qiaquick gel extraction kit (Qiagen), digested with the appropriate restriction endonucleases *Bam*H I and *Eco*R I, then cloned into the prokaryotic expression vector pGEX-4T1 (Pharmacia). This vector, which contains a glutathione-S-transferase (GST) gene, was previously modified to include an 8-His tag at the C-terminus of the expressed protein useful to purify the recombinant protein by affinity chromatography on Ni-NTA columns (Qiagen). PCR products were thus cloned in frame with both the GST and the (His)<sub>8</sub> tag. The resulting recombinant plasmids (pGEX-4T1-His-EnP1 or pGEX-4T1-His-EnP2) were introduced into the *E. coli* BL21<sup>+</sup> strain. Expression of the recombinant proteins was induced with 2 mM IPTG (isopropyl-β-D-thiogalactopyranoside) and bacteria were cultured for 4 h at 37 °C. Total *E. coli* proteins were then solubilized in 2.5% SDS, 100 mM DTT. Recombinant proteins were also purified on Ni-NTA column according to the recommendations of the manufacturer (Qiagen). Protein samples were then analyzed on SDS-PAGE gels.

## 2.7. cDNA library screening and sequencing

An *E. cuniculi* cDNA library constructed in the ZAP Express vector (Stratagene) was kindly provided by Dr W. Bohne (Dept. of Bacteriology, Univ. of Göttingen). The immunoscreening of the library was done with a monoclonal antibody (Ec102) directed against the *E. cuniculi* polar tube. Clones of interest were subjected to 'in vivo excision' according to the manufacturer's protocol (Stratagene) and sequenced by the Sanger method using the ABI Prism Dye Terminator Cycle sequencing according to the recommendations of the manufacturer (Perkin-Elmer). DNA sequences from the cDNA 3' end were then compared with the whole *E. cuniculi* genome sequence database (Katinka et al., 2001) to identify the complete open reading frame (ORF).

## 2.8. Tryptic digestion and MALDI-TOF-MS analysis

The four protein bands from the 200 mM DTT-soluble fraction were manually excised from Coomassie blue-stained gels and washed several times with destaining solutions (25 mM NH<sub>4</sub>HCO<sub>3</sub> for 15 min and then 50% (v/v) acetonitrile containing 25 mM NH<sub>4</sub>HCO<sub>3</sub> for 15 min). After dehydration with 100% acetonitrile and drying, the gels were incubated

with a reducing solution (25 mM NH<sub>4</sub>HCO<sub>3</sub>, 10 mM DTT) for 1 h at 37 °C, and subsequently with an alkylating solution (25 mM NH<sub>4</sub>HCO<sub>3</sub>, 55 mM iodoacetamide) for 30 min at 37 °C. Gels were then washed several times with the destaining solutions and a final wash with pure water for 15 min, before being shrunk again with 100% acetonitrile. Depending on protein amount, 2–3 μL of 0.1 μg/μL modified trypsin (Promega, sequencing grade) in 25 mM NH<sub>4</sub>HCO<sub>3</sub> was added to the gel. After 30 min of incubation, 7 μL of 25 mM NH<sub>4</sub>HCO<sub>3</sub> was added to cover the gel for a further incubation of 5 h at 37 °C. For MALDI-TOF-MS analysis, a 0.5 μL aliquot of peptide mixture was mixed with 0.5 μL matrix solution (α-cyano-4-hydroxycinnamic acid at half saturation in 60% acetonitrile/0.1% trifluoroacetic acid). The resulting solution was automatically spotted on the target plate and then rinsed with 2 μL of 0.1% TFA for 30 s, after which liquid was blown off by pressurized air. Peptides mixtures were then analysed with a MALDI-TOF mass spectrometer (Autoflex, Bruker Daltonik) in reflector mode over a mass range of 0–4200 Da. For each sample, spectrum acquisition and annotation were done automatically. Spectra from 200 shots at several positions were combined to generate a peptide mass fingerprint (PMF) for database searches. Spectra were calibrated using a mixture of angiotensin II (1046.54 Da), Substance P (1347.74 Da), bombesin (1619.82 Da) and ACTH (18–39) (2465.20 Da) as an external standard. Consecutive searches against the specific database of *E. cuniculi* were performed for each sample, with an intranet version of Mascot 1.7. All peptide masses were assumed to be monoisotopic and [M+H]<sup>+</sup> (protonated molecular ions). Peptide modifications allowed during the search were: carbamidomethylation of cysteines, *N*-acetylation of proteins and oxidation of methionines. The maximum number of missed cleavages was set to one and the mass tolerance was limited to ±150 ppm. In most cases, proteins showing a score higher than 70 and an average coverage of 20% were automatically validated. Proteins with scores lower than 70 and with a coverage lower than 20% were checked manually.

## 2.9. 5' rapid amplification of cDNA ends (RACE)

Total RNAs were isolated from *E. cuniculi* infected MDCK cells, using the RNeasy Midi Kit (Qiagen). *Enp1* mRNAs were reverse transcribed with the antisense primer *enp1R* (5'-gatgcacactcggagcattgggc-3') using the SMART RACE cDNA Amplification Kit (Clontech) as described (Peuvél et al., 2002). PCR products were then cloned into the pGEM-T Easy Vector System (Promega) and sequenced.

## 2.10. Sequence analyses

Gene and protein statistical analysis, molecular mass and isoelectric point were calculated using FREXSQ, SAPS and MWCALC, available on the French molecular biology server INFOBIOGEN (<http://www.infobiogen.fr>). Signal peptide was predicted using PSORT (<http://psort.nibb.ac.jp>). Search for homologous proteins in databases was done using BLAST



(<http://www.ncbi.nlm.nih.gov/BLAST/>). Search for potential glycosylated phosphatidyl inositol (GPI)-anchorage was done using the DGPI algorithm ([http://129.194.185.165/dgpi/index\\_en.html](http://129.194.185.165/dgpi/index_en.html)). *N*- and *O*-glycosylation potential sites were determined using NetOglyc and NetNglyc servers (<http://us.expasy.org>).

### 3. Results

#### 3.1. Identification of a cysteine-rich protein by immunoscreening of an *E. cuniculi* cDNA library

We previously showed that a monoclonal antibody (Ec102) raised against the polar tube reacts with three different proteins in Western blots, with apparent sizes of 55, 35 and 28 kDa (Delbac et al., 1998a). The 55 and 35 kDa bands correspond to the known PTP1 (Delbac et al., 1998b) and PTP2 (Delbac et al., 2001). In an attempt to characterize the protein associated with the lowest molecular weight band, immunoscreening of an *E. cuniculi* cDNA library was performed with Ec102. Three kinds of clones were identified, two of these corresponding to *ptp1* and *ptp2* genes. The other recombinant phages containing inserts of about 1 kbp were sequenced and compared to the *E. cuniculi* genome sequence database. The resulting cDNA sequence maps to a single ORF located on chromosome I (ECU01\_0820, CAD24952). Considering the first in-frame methionine as the initiation codon, this ORF of 1071 bp encodes a protein of 357 amino acids (40,550 Da) with a high cysteine content (Fig. 1). This protein is basic (pI 9.14), the most abundant charged residue being lysine (9.52%), and has a von Heijne-predicted 16 amino acid signal sequence at the amino terminus. The cysteine residues (6.44%) are dispersed along the sequence but are absent in the last 40 amino acids of the C-terminal region. One potential *N*-glycosylation site (NGSN) is found in position 47–50 but no *O*-glycosylation site is predicted. Devoid of large repeated motifs, the protein comprises some stretches of three to five identical amino acids (Fig. 1). Comparisons of ECU01\_0820 with protein databases failed to reveal significant homology. Sequencing of the cDNA 3' end from four different clones showed that the 3' untranslated region (UTR) is AT-rich (70%) and only 20 nucleotides (nt) long. A potential polyadenylation site

(AAATAAA) is found 4 nt downstream of the stop codon. The 5' cDNA end from recombinant phages started a few nt downstream of the putative initiation codon, suggesting that the clones were truncated. The transcription start site of ECU01\_0820 was thereby mapped by 5' rapid amplification of cDNA ends (5' RACE). Sequencing of three different clones revealed that transcription should be initiated at the -1 position relative to the translational start codon.

#### 3.2. The cysteine-rich protein is destined to the endospore (EnP1)

For heterologous expression of a recombinant ECU01\_0820 polypeptide in *E. coli*, a 800 bp region corresponding to amino acids 32–300 was PCR amplified and cloned into the pGEX-4T1 vector, in frame with the GST and a 8-histidine epitope at the N- and C-terminal parts, respectively. SDS-PAGE analysis from IPTG-induced bacteria indicated high expression of a 55 kDa fusion protein (not shown). Surprisingly, the antibody Ec102 used for the cDNA library immunoscreening did not detect the recombinant protein. This prompted us to produce specific antibodies against this recombinant protein by injection into mice. The antisera reacted with a 40 kDa band in *E. cuniculi* sporul lysates (Fig. 2, lane 2), not with the 28 kDa band that was recognized by the Ec102 antibody. No polar tube labeling was observed by immunofluorescence assay (IFA). In contrast, a significant fluorescent signal was associated with the cell surface of some developmental stages within the parasitophorous vacuole (Fig. 3(A)–(D)). The labeled parasites are elongated cells that may correspond to meronts and sporonts. Immunoelectron microscopy (IEM) was performed to test whether the expression of ECU01\_0820 was restricted to early stages, or not. Indeed, the absence of reactivity of late sporogonic stages in IFA may be an artefact due to an insufficient permeabilization of the thick wall. Parasitic stages at the onset of sporogony clearly showed a gold labeling of the cell periphery (Fig. 4(A)–(D)). During spore differentiation, the labeling was specifically concentrated within the electron-lucent inner layer of the wall, designated as the endospore (Fig. 4(E)–(H)). Thus, the protein named EnP1 is in fact present in all sporogonic stages including mature spores.

```

MKLLGFLIVGLSAISALKTKALHLTEQELRPYSAVVDANMAFALNGSN      50
IHEAIKYLQAMNIKKAYVLYWHDHLRGTMPVLYDNGALAPFDPYTNTAK      100
YVLVEAEPPGSKAASVGGFQAATSSEKIYVEGSARPAQSEVIEFPVE      150
RRPHYKKIVVNPSPSNIPEPEYDSSSSSENKKRKTFPRIKEKEG      200
SRRRGPRKVEVLKSQKTYTFDIEKYRRRGEVVVVVSKDSKEKFERFIL      250
SRNGEIRGNNNNILEPLPKLRPGQLHKLKHHIERKVQEVMYINA      300
KDIFVLVGDDFYRVVVNDRRRYRNHLHLKKVRGHKLRELKHLGFLGVEF      350
GPLDLDR                                                       357

```

Fig. 1. Amino acid sequence of EnP1 (GenBank accession no. CAD24952). The predicted signal peptide is underlined and the cysteine residues are shaded. The potential *N*-glycosylation site is double underlined. Stretches of identical residues are in bold type. Amino acids are numbered on the right.





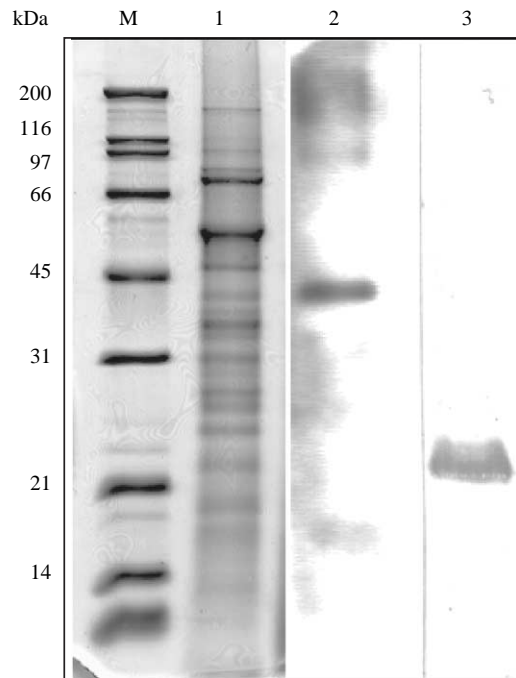


Fig. 2. Mouse antibodies raised against the recombinant EnP1 and EnP2 detect some *Encephalitozoon cuniculi* proteins of 40 and 22 kDa in size, respectively. SDS-PAGE (12%) and Western blot of total *E. cuniculi* sporal proteins (Buffer lysis 2.5% SDS/100 mM dithiothreitol). Lane 1: Coomassie blue stained profile. Lane 2: Immunoblotting reactivity of the anti-recombinant EnP1 antiserum (1:200 dilution) with a 40 kDa *E. cuniculi* protein band. Lane 3: Immunoblotting reactivity of the anti-recombinant EnP2 antiserum (1:500 dilution) with a 22 kDa *E. cuniculi* protein band. The blots were developed using ECL (Amersham). Molecular mass standards (M) are in kilodaltons.

### 3.3. Identification of a potentially GPI-anchored 22-kDa protein by mass spectrometry analysis of a 200 mM DTT fraction

As known microsporidian PTPs and SWPs are cysteine-rich and solubilized in the presence of high concentrations of thiol-reducing agents, disulfide bridges may be thought to exist and to play an essential role in the assembly and stabilization of both the polar tube and the spore wall. After boiling *E. cuniculi*

spores in 2.5% SDS, 100 mM DTT, the insoluble fraction was treated for a further 72 h in a lysis buffer containing 200 mM DTT. The two protein samples (100 and 200 mM DTT-treated protein extracts) were then analyzed by SDS-PAGE (Fig. 5). Coomassie blue staining revealed four major protein bands in the 200 mM DTT fraction (55, 35, 30 and 22 kDa) (Fig. 5, lane 2). Two bands (55 and 35 kDa) were stained with Ec102 antibody and should represent PTP1 and PTP2 (data not shown). To better identify the constituents of the four bands, mass spectrometry analyses were performed and experimental peptide masses were compared to the values predicted from the *E. cuniculi* proteome database. As expected, the 55 kDa band contained tryptic peptides originating from both PTP1 and SWP1. Only PTP2 derived-peptides were found in the 35 kDa band. The 30 and 22 kDa bands were assigned to two proteins encoded by genes that are, respectively, located on the chromosome XI (ECU11\_0510, CAD25961) and chromosome I (ECU01\_1270, CAD25000). ECU11\_0510 was annotated as a chitooligosaccharide deacetylase whereas ECU01\_1270 had no known homologue. Both proteins are predicted to be GPI-anchored. These data encouraged us to study the localization of the ECU01\_1270 protein, a potential microsporidia-specific protein.

### 3.4. EnP2, a serine-rich protein, is also localized in the inner layer of the spore wall

Two in-frame ATG codons are located 33 and 120 nt downstream of the first ATG in the ECU01\_1270 sequence. However, only the first ATG is followed by a hydrophobic von Heijne peptide, as expected for a secreted protein. Flanking regions are AT-rich with 65% of AT nucleotides in the 100-bp sequence upstream the ATG codon and 58% of AT nt in the 3' flanking region. A potential polyadenylation site AATAAA is found to overlap the stop codon. The ECU01\_1270 ORF is 663-bp long and encodes a 221-amino acid protein, with calculated molecular mass of 22.5 kDa and pI of 8.74 (Fig. 6). A major acidic part of the protein (pI of 4.76 for the 165 first amino acids) contrasts with the C-terminal region characterized by a very basic KS-rich domain (pI 10.52). The N-terminal

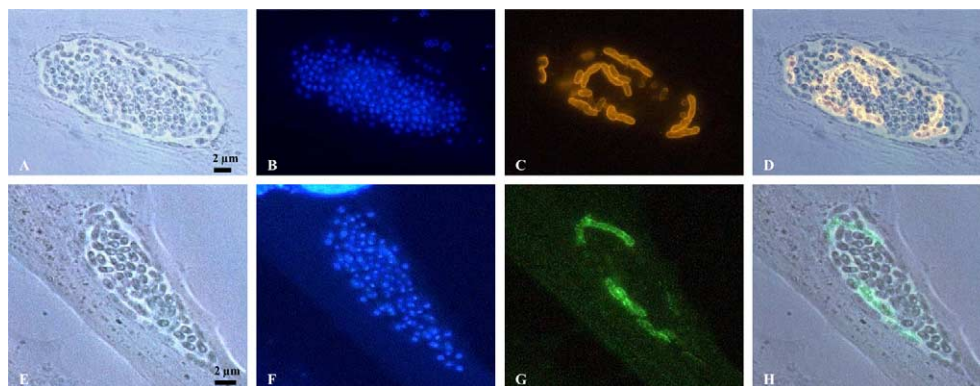


Fig. 3. Indirect-immunofluorescence assay with mouse anti-EnP1 and anti-EnP2 antibodies showing cell surface labelling of some intracellular parasitic stages. *Encephalitozoon cuniculi* infected human foreskin fibroblast cells are visualized in phase contrast (A, E), after DAPI staining (B, F) and after anti-EnP1 (C) and anti-EnP2 (G) labeling. The overlays of A and C (D), and E and G (H) confirm that only a few intracellular parasites are stained with both antibodies. Anti-EnP1 and -EnP2 antisera were used at a 1:100 dilution. Secondary antibody was TRITC-conjugated goat anti-mouse IgG (C) or Alexa 488-conjugated goat anti-mouse IgG (G).



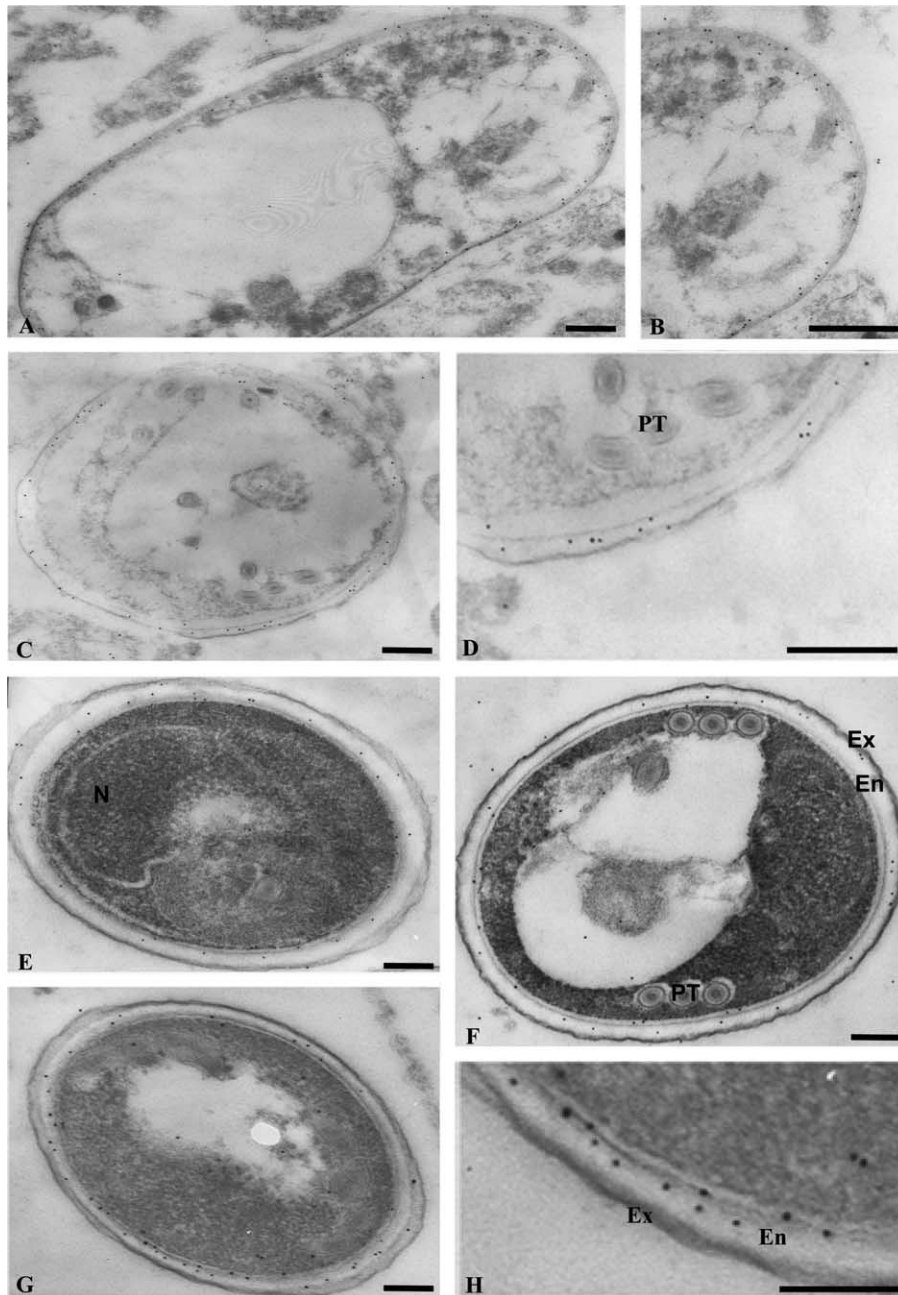


Fig. 4. Immunoelectron microscopy of different *E. cuniculi* developmental stages with anti-EnP1 antibodies reveals labelling in the endospore. (A–D): sporoblasts with gold particles at the periphery of the cells. (E–H): spores labeled in the chitin-rich inner layer of the spore wall, i.e. the endospore. Anti-EnP1 antisera were used at a 1:20 dilution. Secondary antibody was goat anti-mouse IgG conjugated with 10 nm colloidal gold particles (Sigma). B, D and H correspond to higher magnifications of A, C and G, respectively. Ex, exospore; En, endospore; PT, polar tube; N, nucleus. Bars, 200 nm.

signal peptide would be cleaved between two serines in position 17 and 18. An extracellular (including cell wall) or plasma membrane localization was predicted with the PSORT program (<http://psort.nibb.ac.jp>). The use of the DGPI algorithm indicated a GPI-anchorage signal with a potential cleavage site at 193. No *N*-glycosylation sites were identified whereas the NetOGlyc algorithm (<http://www.cbs.dtu.dk/services/NetyOGlyc/>) predicts 24 *O*-glycosylation sites, mainly within the central region. That the protein could be *O*-glycosylated is supported by the fact that serine is the most abundant residue (15.84%). Two tetrapeptides are repeated

twice: SSGL (18–21 and 54–57) and PAST (120–123 and 152–155) (Fig. 6).

The ECU01\_1270 sequence without its potential signal peptide was cloned into the pGEX-4T1 expression vector. After IPTG induction, the bacterial lysate was found to contain a 50 kDa recombinant protein (not shown). Mouse polyclonal antibodies raised against the fusion protein recognized a 22 kDa protein band in *E. cuniculi* extract (Fig. 2, lane 3), in agreement with the size predicted from the sequence. In IFA, these antibodies stained a parasite antigen having a peripheral location similar to EnP1 in some elongated intracellular



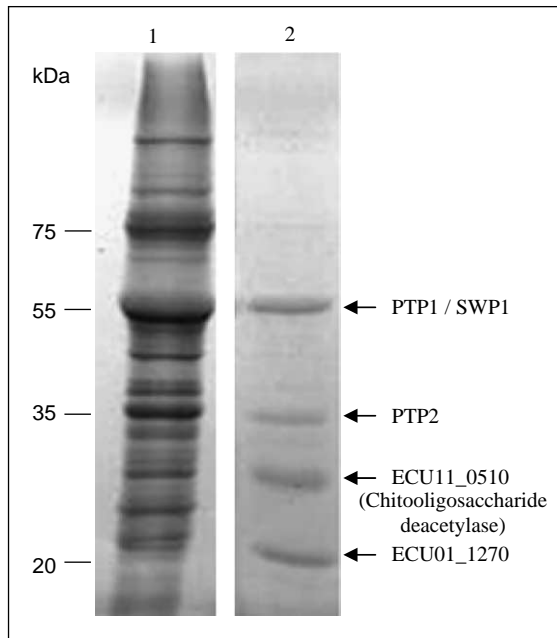


Fig. 5. SDS-PAGE (10%) analysis of the 200 mM dithiothreitol (DTT)-soluble *E. cucurbitae* protein fraction. Lane 1 corresponds to total sporal proteins solubilized with 2.5% SDS, 100 mM DTT. After treatment with 2.5% SDS, 100 mM DTT and centrifugation, the insoluble pellet was incubated for further 72 h in 200 mM DTT, revealing four major protein bands (lane 2). The five proteins identified by mass spectrometry are indicated.

developmental stages (Fig. 3(E)–(H)). Double immunostaining with anti-EnP1 and anti-ECU01\_1270 supported the colocalization of the two proteins (data not shown). A surface-specific labeling of differentiating stages was demonstrated by IEM (Fig. 7(A)–(C)). Here again, the reactivity of the inner wall layer with the anti-ECU01\_1270 was evident in mature spores (Fig. 7(D) and (E)). Gold particles were especially abundant over the inner most region overlapping the plasma membrane (Fig. 7(E)). We conclude that the antigen exists in all sporogonic stages and represents an additional protein targeted to the endospore, justifying its further designation as EnP2.

#### 4. Discussion

Microsporidia are fungal-like organisms whose infective stage is a spore surrounded by a thick cell wall that comprises an outer electron-dense exospore composed primarily of proteins and an inner electron-lucent endospore which contains

chitin and protein components (reviewed by Bigliardi and Sacchi, 2001). The protection against environmental stresses is illustrated by the capacity of survival of the microsporidian spore to a pH of 9 or 4 for 24 h, heating at 56 °C for 60 min, or storage at 4 °C for 2 years without losing infectivity (Koudela et al., 1999). Moreover, the rigidity of the wall would maintain a high intrasporal hydrostatic pressure and aqueous permeability might play a role in the control of the polar tube extrusion (Frixione et al., 1997). It cannot be ruled out that an interaction between some component(s) of the cell wall and some receptor(s) of the host is also a prerequisite for polar tube extrusion. Two different strategies based on cDNA library immunoscreening and mass spectrometry analysis of a DTT-soluble fraction were used in the present identification and characterization of two *E. cucurbitae* proteins located to the chitin-rich endospore. Although located on the same chromosome (ECU01\_0820, and ECU01\_1270), the corresponding genes are not tandemly arranged unlike those encoding two polar tube proteins (*ptp1* and *ptp2*) on chromosome VI (Delbac et al., 2001). The *enp1* and *enp2* genes may be considered to be microsporidia-specific since no homology with sequences were detected in databases including cell wall proteins of yeasts and filamentous fungi.

Very little is currently known about the translation and transcription start sites of microsporidian genes. Our 5' RACE PCR analysis indicates that the 5' UTR of the *enp1* mRNA is reduced to a single nucleotide. Short 5' UTRs have been previously observed in *E. cucurbitae swp1* (8 nt) and *ptp1* (3 nt) mRNA (Bohne et al., 2000; Peuvel et al., 2002), contrasting with an average length close to 90 nt in mammals and 52 nt in yeast (Kozak, 1987; Cigan and Donahue, 1987). However, short 5' UTRs also exist in amitochondrial protozoa of the genera *Entamoeba*, *Trichomonas* and *Giardia*. In spite of an extreme reduction in *Giardia*, a capped mRNA with a single nt leader has been demonstrated to be sufficient for recruiting ribosome and initiating translation (Li and Wang, 2004). Looking at the regions around start codons of the *E. cucurbitae* genome, Fuglsang (2005) recently identified a highly cytosine-enriched region 5–10 nt upstream of start codons that could serve as a signal for initiation of translation in *E. cucurbitae*. An alternative mechanism involving an element downstream of the start codon ('downstream box') cannot be excluded (Duffieux et al., 1998). The *E. cucurbitae* mRNAs studied so far are polyadenylated and display rather reduced 3' UTRs with

```

MVRRLYFLAVMGVVRSSGLYIPSVVLQELGIASSQGMLMIAETNNGNF    50
GIVSSGLENPVYITESPQGHREVSQPVVRGEDRAVPIYAPSAEEVRESIS    100
SVRGSEPGQYIAPQPTGFVPASTPVFGTIESASTAGAAVPVEGVFVASTE    150
NPASTGSSTSTPPPKGTAGTTDNKGKAGGAAADDKSKSSSSSSKSKKKK    200
GAKSLVALGAVATTALFSIVM                                     221

```

Fig. 6. Amino acid sequence of EnP2 (GenBank accession no. CAD25000). The predicted 17 amino acid signal peptide is underlined and the two cysteine residues are shaded. The two repetitive motifs of four amino acids (SSGL and PAST) are in bold and the KS-rich domain is boxed. The potential glycosylated phosphatidyl inositol-anchor signal is double underlined and the predicted cleavage site is indicated by an arrow. Amino acids are numbered on the right.



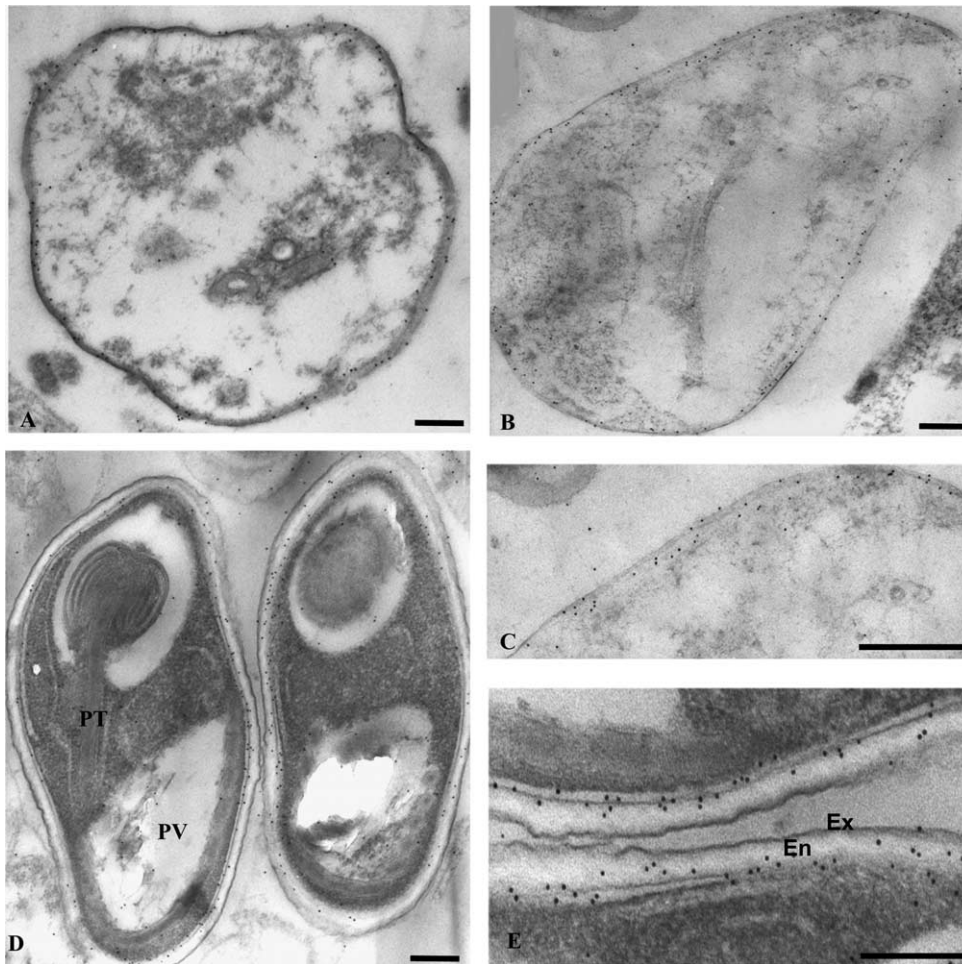


Fig. 7. Immunoelectron microscopy of different *E. cuniculi* developmental stages with anti-EnP2 antibodies reveals labelling in the endospore. Similar to EnP1, EnP2 is detected at the cell surface of some sporogonic stages (A–C) and in the endospore of mature spores (D–F). C and E correspond to higher magnifications of B and D, respectively. Anti-EnP2 antisera were used at a 1:20 dilution. Secondary antibody was goat anti-mouse IgG conjugated with 10 nm colloidal gold particles (Sigma). Ex, exospore; En, endospore; PT, polar tube; PV, posterior vacuole. Bars, 200 nm.

sizes ranging from 25 to 175 nt (Broussolle et al., 1999; Bohne et al., 2000; Delbac et al., 2001). This is the case for *enp1* mRNA with an AT-rich 3' UTR of only 20 nt. The highly reduced length of 5' and 3' UTRs fits with the remarkable small nuclear genome in microsporidia, particularly in *E. cuniculi* in which the mean distance between two adjacent coding DNA sequences is only 129 bp (Katinka et al., 2001). An overlapping of 3' UTRs has been suggested for two opposite transcription units separated by only 21 nt (Peuvél et al., 2002).

Curiously, the endospore protein EnP1 was identified by cDNA library immunoscreening using the monoclonal antibody Ec102 known to be directed against polar tube components including PTP1, PTP2 and an undefined 28 kDa protein. As EnP1 does not correspond to the 28 kDa protein, the *enp1*-containing clones are probably 'false positives'. Numerous clones selected by cDNA immunoscreening with Ec102 and other antibodies were found to correspond to *enp1* (not shown). This indicates an enrichment of the *E. cuniculi* cDNA library with *enp1* that is probably a highly expressed gene. Reverse transcriptase (RT)-PCR analyses from RNA extracted at different times p.i. confirmed that *enp1* is

expressed at a high level during the intracellular cycle (not shown). In contrast to *swp1* where mRNA appears to be upregulated during spore differentiation (Bohne et al., 2000), initiation of *enp1* mRNA expression begins during merogony (not shown).

Immunofluorescence imaging demonstrated that EnP1 and EnP2 colocalized to the cell surface of parasite stages at the onset of sporogony. However, immunoelectron microscopy revealed that both proteins were present in all sporogonic stages. After accumulation at the cell surface of sporonts then sporoblasts, EnP1 and EnP2 become clearly associated with the inner spore wall layer. Thus, IFA data in microsporidia should be cautiously interpreted because of limitations due not only to the very small size of these organisms but also to the thickening of a cell wall that hinders the penetration of antibodies.

Both EnP1 and EnP2 contain a N-terminal hydrophobic leader sequence consistent with a translocation of these proteins into the endoplasmic reticulum (ER). During trafficking through the ER-Golgi compartments, EnP2 glycosylation might occur, this protein being predicted to be O-glycosylated and GPI-anchored. So far, only one study has provided





evidence for *O*-mannosylation of PTP1, the major polar tube protein in *Encephalitozoon hellem* (Xu et al., 2004). In contrast, EnP1 would not be linked to carbohydrates because no *O*-glycosylation site is found and *N*-glycosylation machinery is probably lacking (Katinka et al., 2001). Yeasts including *Saccharomyces cerevisiae* and the opportunistic pathogen *Candida albicans* have also a rigid cell wall outside their plasma membrane that consists of  $\beta$ -glucans, chitin and various kinds of mannoproteins (Kapteyn et al., 2000; Klis et al., 2002). Major parts of mannoproteins are synthesized as GPI-anchored proteins and are then transferred to cell wall  $\beta$ -1,6 glucans. Whether EnP2 is GPI-anchored and then translocated to the cell wall endospore remains to be investigated.

We showed that EnP1 and EnP2 are also solubilized in the presence of DTT, arguing for a significant role for disulfide bridges in spore wall assembly. This is in agreement with the presence of a high number of cysteine residues (23) in EnP1. At the difference, although EnP2 is found in the 200 mM DTT-soluble protein extract, this protein only contains two cysteines. Among the four additional proteins found in the 200 mM DTT-soluble fraction, three correspond to previously reported cysteine-rich proteins from either the polar tube (PTP1 and PTP2, Delbac et al., 1998b; 2001) or the exospore (SWP1, Bohne et al., 2000), and one to a possible chitin deacetylase that is predicted to be GPI-anchored and localized to the outer face of the parasite plasma membrane (Brosson et al., 2005). The fact that EnP2 is particularly insoluble could be due to its possible GPI-anchorage or because of some strong interactions with other DTT-soluble proteins from the spore wall.

Looking at the sequences of *Antonosporea* (*Nosema*) *locustae*, a microsporidia parasite of insects for which a genome sequencing project is in progress (*'Nosema locustae* Genome Project, Marine Biological Laboratory at Woods Hole, funded by NSF award number 0135272', <http://jbpc.mbl.edu/Nosema/Nannotation.html>), was helpful in the identification of an EnP1-like protein (contig 271, positions 1139–2281). This protein is longer (381 amino acids in *A. locustae* versus 357 aa in *E. cuniculi*) and only exhibits 21.6% identity with that of *E. cuniculi*. Most notable is the high number of cysteine residues (29). The position and the spacing of 13 of these are conserved, suggesting that these proteins have similar secondary structures and functions. We also identified by PCR an EnP1 homologue in the microsporidium *Paranosema grylli* a species closely related to *A. locustae* arguing for the presence of EnP homologues in most microsporidian species (not shown). In contrast, a BlastP search for a homologue of EnP2 in *A. locustae* failed. The absence of hits could be due to incompleteness of genome sequencing but we cannot exclude the presence of divergent homologues or a real absence. No precise function has yet been attributed to EnP1 and EnP2, further work is now needed to identify the precise role of both proteins in spore wall biosynthesis and organization.

## Acknowledgements

This work was supported by a grant from 'Recherche fondamentale en microbiologie, et maladies infectieuses et

parasitaires'. I.P., V.P. and D.B. were supported by grants from 'Ministère de l'Éducation Nationale de la Recherche et de la Technologie'. We are grateful to W. Bohne (Dept. of Bacteriology, Univ. of Göttingen) for kindly providing the *E. cuniculi* cDNA library. We also thank G. Méténier for helpful discussion, B. Chebance and A. Voldoire for technical assistance.

## References

- Beckers, P.J.A., Derks, G.J.M.M., van Gool, T., Rietveld, F.J.R., Sauerwein, R.W., 1996. *Encephalitozoon intestinalis*-specific monoclonal antibodies for laboratory diagnosis of microsporidiosis. *J. Clin. Microbiol.* 34, 282–285.
- Bigliardi, E., Sacchi, L., 2001. Cell biology and invasion of the microsporidia. *Microbes Infect.* 3, 373–379.
- Bigliardi, E., Selmi, M.G., Lupetti, P., Corona, S., Gatti, S., Scaglia, M., Sacchi, L., 1996. Microsporidian spore wall: ultrastructural findings on *Encephalitozoon hellem* exospore. *J. Eukaryot. Microbiol.* 43, 181–186.
- Bohne, W., Ferguson, D.J.P., Kohler, K., Gross, U., 2000. Developmental expression of tandemly repeated, glycine- and serine-rich spore wall protein in the microsporidian pathogen *Encephalitozoon cuniculi*. *Infect. Immun.* 68, 2268–2275.
- Brosson, D., Kuhn, L., Prensier, G., Vivarès, C.P., Texier, C., 2005. The putative chitin deacetylase of *Encephalitozoon cuniculi*: a surface protein implicated in microsporidian spore-wall formation. *Fed. Eur. Microbiol. Soc. Microbiol. Lett.* 247, 81–90.
- Broussolle, V., Fumel, S., Peyret, P., Vivarès, C.P., 1999. Application of differential display RT-PCR to the analysis of gene expression in a host cell-microsporidian *E. cuniculi* interaction. *J. Eukaryot. Microbiol.* 46, 25S–26S.
- Bryan, R.T., Schwartz, D.A., 1999. Epidemiology of microsporidiosis. In: Wittner, M., Weiss, L.M. (Eds.), *The Microsporidia and Microsporidiosis*. American Society for Microbiology, Washington, DC, pp. 502–516.
- Cigan, A.M., Donahue, T.F., 1987. Sequence and structural features associated with translational initiator regions in yeast—a review. *Gene* 59, 1–18.
- Delbac, F., Duffieux, F., David, D., Méténier, G., Vivarès, C.P., 1998a. Immunocytochemical identification of spore proteins in two microsporidia, with emphasis on extrusion apparatus. *J. Eukaryot. Microbiol.* 45, 224–231.
- Delbac, F., Peyret, P., Méténier, G., David, D., Danchin, A., Vivarès, C.P., 1998b. On proteins of the microsporidian invasive apparatus: complete sequence of a polar tube protein of *Encephalitozoon cuniculi*. *Mol. Microbiol.* 29, 825–834.
- Delbac, F., Peuvél, I., Méténier, G., Peyretailade, E., Vivarès, C.P., 2001. Microsporidian invasion apparatus: identification of a novel polar tube protein and evidence for clustering of *ptp1* and *ptp2* genes in three *Encephalitozoon* species. *Infect. Immun.* 69, 1016–1024.
- Duffieux, F., Peyret, P., Roe, B.A., Vivarès, C.P., 1998. First report on the systematic sequencing of the small genome of *Encephalitozoon cuniculi* (Protozoa, Microspora): gene organization of a 4.3 kbp region on chromosome I. *Microb. Comp. Genomics* 3, 1–11.
- Enriquez, F.J., Ditrich, O., Palting, J.D., Smith, K., 1997. Simple diagnosis of *Encephalitozoon* sp. microsporidian infections by using a panspecific antiexospore monoclonal antibody. *J. Clin. Microbiol.* 35, 724–729.
- Frixione, E., Ruiz, L., Cerbon, J., Undeen, A.H., 1997. Germination of *Nosema algerae* (Microspora) spores: conditional inhibition by D<sub>2</sub>O, ethanol and Hg<sup>2+</sup> suggests dependence of water influx upon membrane hydration and specific transmembrane pathways. *J. Eukaryot. Microbiol.* 44, 109–116.
- Fuglsang, A., 2005. A cytosine-rich region upstream of start codons serving as a signal for initiation of translation in *Encephalitozoon cuniculi*? *Mol. Biochem. Parasitol.* 140, 69–73.
- Hayman, J.R., Hayes, S.F., Amon, J., Nash, T.E., 2001. Developmental expression of two spore wall proteins during maturation of the microsporidian *Encephalitozoon intestinalis*. *Infect. Immun.* 69, 7057–7066.
- Kapteyn, J.C., Hoyer, L.L., Hecht, J.E., Muller, W.H., Andel, A., Verkleij, A.J., Makarow, M., Van Den Ende, H., Klis, F.M., 2000. The cell wall architecture of *Candida albicans* wild-type cells and cell wall-defective mutants. *Mol. Microbiol.* 35, 601–611.



- Katinka, M.D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretailade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., Vivarès, C.P., 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414, 450–453.
- Keeling, P.J., 2003. Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. *Fungal Genet. Biol.* 38, 298–309.
- Keohane, E.M., Takvorian, P.M., Cali, A., Tanowitz, H.B., Wittner, M., Weiss, L.M., 1996. Identification of a microsporidian polar tube protein reactive monoclonal antibody. *J. Eukaryot. Microbiol.* 43, 26–31.
- Klis, F.M., Mol, P., Hellingwerf, K., Brul, S., 2002. Dynamics of cell wall structure in *Saccharomyces cerevisiae*. *Fed. Eur. Microbiol. Soc. Microbiol. Rev.* 26, 239–256.
- Koudela, B., Kucerova, S., Hudcovic, T., 1999. Effect of low and high temperatures on infectivity of *Encephalitozoon cuniculi* spores suspended in water. *Folia Parasitol.* 46, 171–174.
- Kozak, M., 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15, 8125–8148.
- Li, L., Wang, C.C., 2004. Capped mRNA with a single nucleotide leader is optimally translated in a primitive eukaryote, *Giardia lamblia*. *J. Biol. Chem.* 279, 14656–14664.
- Lujan, H.D., Conrad, J.T., Clark, C.G., Touz, M.C., Delbac, F., Vivarès, C.P., Nash, T.E., 1998. Detection of microsporidia spore-specific antigens by monoclonal antibodies. *Hybridoma* 17, 237–243.
- Peuvél, I., Peyret, P., Méténier, G., Vivarès, C.P., Delbac, F., 2002. The microsporidian polar tube: evidence for a third polar tube protein (PTP3) in *Encephalitozoon cuniculi*. *Mol. Biochem. Parasitol.* 122, 69–80.
- Prigneau, O., Achbarou, A., Bouladoux, N., Mazier, D., Desportes-Livage, I., 2000. Identification of proteins in *Encephalitozoon intestinalis*, a microsporidian pathogen of immunocompromised humans: an immunoblotting and immunocytochemical study. *J. Eukaryot. Microbiol.* 47, 48–56.
- Thomarat, F., Vivarès, C.P., Gouy, M., 2004. Phylogenetic analysis of the complete genome sequence of *Encephalitozoon cuniculi* supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. *J. Mol. Evol.* 59, 780–791.
- van Gool, T., Vetter, J.C., Weinmayr, B., Van Dam, A., Derouin, F., Dankert, J., 1997. High seroprevalence of *Encephalitozoon* species in immunocompetent subjects. *J. Infect. Dis.* 175, 1020–1024.
- van Gool, T., Biderre, C., Delbac, F., Wentink-Bonnema, E., Peek, R., Vivarès, C.P., 2004. Serodiagnostic studies in an immunocompetent individual infected with *Encephalitozoon cuniculi*. *J. Infect. Dis.* 189, 2243–2249.
- Weiss, L.M., 2001. Microsporidia: emerging pathogenic protists. *Acta Trop.* 78, 89–102.
- Xu, Y., Takvorian, P.M., Cali, A., Orr, G., Weiss, L.M., 2004. Glycosylation of the major polar tube protein of *Encephalitozoon hellem*, a microsporidian parasite that infects humans. *Infect. Immun.* 72, 6341–6350.



