# Analyse protéomique par LC-MS/MS

D. Pflieger, J.-P. Le Caer, J. Vinh, V. Labas, J. Rossier,

*Laboratoire de Neurobiologie CNRS UMR 7637, ESPCI, 10 rue Vauquelin, 75005 Paris*

B. Bernard,

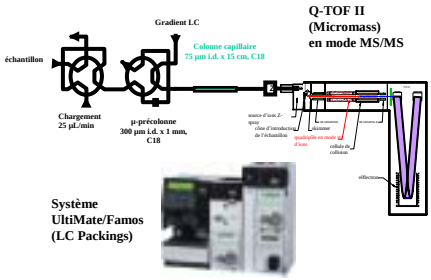
*Groupe Biologie du Cheveu, l'Oréal, Centre Charles Zviak, 90 rue du Général Roguet, 92583 Clichy*

## Introduction

La stratégie classique de l'étude du protéome débute par une séparation des extraits protéiques par électrophorèse sur gel bidimensionnelle (gel 2D). Si cette technique révèle un grand nombre de protéines exprimées dans un type cellulaire, elle présente un certain nombre de limitations et d'inconvénients : impossibilité de visualiser les protéines trop peu abondantes, séparation sur une gamme de pI limitée, mauvaise pénétration et séparation des protéines hydrophobes dans la première dimension, durée importante de l'obtention du gel, etc.

Depuis quelques années se développe une stratégie alternative, qui repose sur l'analyse de mélanges de protéines par le couplage entre la chromatographie en phase liquide et la spectrométrie de masse en tandem (LC-MS/MS). Brièvement, des échantillons protéiques complexes sont digérés par une endoprotéase spécifique, puis le mélange de peptides résultant est analysé par LC-MS/MS. Les informations de séquence obtenues sur les peptides détectés permettent de déterminer l'identité des protéines présentes dans l'échantillon initial. Une étude comparative de l'expression protéique est en outre possible par le marquage des polypeptides par des isotopes stables.

L'objectif de ce poster est d'évaluer les performances et les limitations de l'approche LC-MS/MS dans l'analyse de mélanges complexes de protéines.. La présente étude s'attache à cerner les possibilités d'identification de protéines du modèle cellulaire *Saccharomyces cerevisiae*, avant d'envisager des analyses quantitatives.




## Matériel et méthodes

**Extraction des protéines de levure** : La source de protéines utilisée a été la souche de *Saccharomyces cerevisiae* CW04, cultivée en milieu complet jusqu'à la phase de croissance exponentielle ($10^8$ cellules/mL). Les protéines ont été extraites par action d'une solution 1,85 M NaOH, 7,4 % β-mercaptoéthanol (v:v) sur un culot de cellules, puis précipitation des polypeptides solubilisés dans TCA 50 % (v:v) et lavage à l'acétone du précipité. Les protéines ont ensuite été resolubilisées dans un tampon 5 % SDS, 50 mM Tris non tamponné.

**Traitements des échantillons protéiques** : Les protéines de l'extrait total ont été séparées sur un gel SDS-PAGE à 12,5% d'acrylamide. 25 µg de protéines ont été déposés par puits. Les protéines ont été révélées par coloration au bleu de Coomassie colloïdal (Biosafe, Biorad). 37 bandes de gel ont été découpées sur la gamme de masse 5-95 kDa. Les protéines contenues dans chaque bande ont été réduites et alkylées par action de DTT et d'iodoacétamide, puis digérées par la trypsine, selon le protocole de Shevchenko *et al.*, 1996. Typiquement, l'équivalent du quart d'une bande digérée a été injecté pour une analyse en LC-MS/MS.

**Analyses par LC-MS/MS** : Les protéines digérées ont été analysées par le couplage entre la chromatographie sur colonne capillaire (système FAMOS-UltiMate, LC Packings) et la spectrométrie de masse en tandem (nanoESI-Q-TOF II, Micromass). Le couplage entre la colonne analytique et la source d'ions électrospray non coaxiale est assuré via un capillaire métallisé (Picotip, NewObjective).

***Conditions de séparation chromatographique*** :

*Colonne capillaire* : Pepmap C18, diamètre interne 75 µm, L = 15 cm, granulométrie 3 µm, porosité 100 Å, LC Packings

*Débit* : 220 nL/min.

*Solvants* : A = 0,1 % acide formique, 4 % acétonitrile, 96 % eau v:v

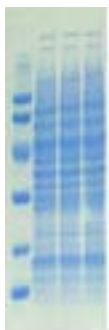
B = 0,085 % acide formique, 90 % acétonitrile, 10 % eau, v:v

*Gradient* : 100 % A => 45% B en 50 min, puis 45% B => 100 % B en 10 min.

***Conditions d'acquisition en masse*** : Les données de spectrométrie de masse ont été obtenues par passage automatique entre les modes d'analyse MS et MS/MS : une seconde d'acquisition en mode MS est suivie de 4 s de fragmentation des huit ions donnant lieu aux signaux détectés les plus intenses.

**Identification de protéines** : Les identifications de protéines ont été obtenues par confrontation des informations de masse à la banque de données protéiques NCBInr, par l'intermédiaire du programme de calcul Mascot (MatrixScience) disponible sur Internet (www.matrixscience.com). Les recherches ont été conduites sans indication de type taxonomique, avec une tolérance sur la mesure des masses de 0,15 Da en mode MS et de 0,25 Da en mode MS/MS.

Les caractéristiques des protéines (masse moléculaire, point isoélectrique et biais de codon) ont été obtenues par interrogation du site http://yeast.genome.jp/cgi-bin/genome_search.cgi, qui contient toutes les données concernant le génome de S. *cerevisiae*. Les grandeurs considérées (masse moléculaire et point isoélectrique) sont celles des protéines matures (déterminées expérimentalement, ou, le plus souvent, prédites).



**Etalons de masse**

94 kDa = phosphorylase b

67 kDa = albumine

43 kDa = ovalbumine

30 kDa = anhydrase carbonique

20,1 kDa = inhibiteur de trypsine

14,4 kDa = α-lactalbumine

## Résultats

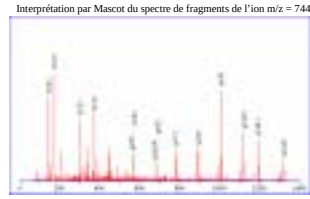
Au total, **408** protéines de levure différentes ont été identifiées. Elles sont de localisations et fonctions très variées (protéines cytosoliques, mitochondriales, nucléaires, ribosomales, solubles ou membranaires).

Les performances analytiques du couplage LC-MS/MS sont résumées dans le tableau suivant :

| | |
|---|---|
| **Nombre moyen de protéines identifiées en une analyse LC-MS/MS** | **18** |
| **Nombre maximal de protéines identifiées en une analyse LC-MS/MS** | **30** |
| **Nombre moyen de peptides contribuant aux identifications de protéines par analyse LC-MS/MS** | **44** |
| **Nombre maximal de peptides contribuant aux identifications de protéines en une analyse LC-MS/MS** | **92** |

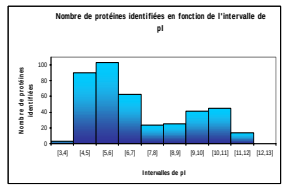
**Exemple d'identification de protéine par le programme de calcul Mascot**

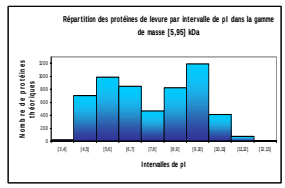
Interprétation par Mascot du spectre de fragments de l'ion m/z = 744,88

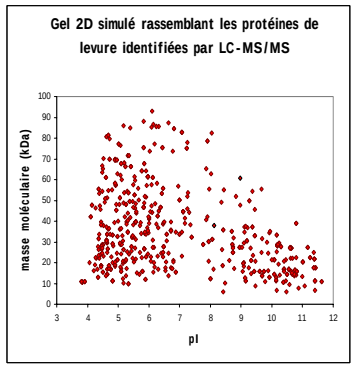


En particulier, 52 protéines du ribosome 80S sur 78 ont été identifiées dans les bandes de gel digérées. L'approche gel 1D –LC-MS/MS a donc permis de caractériser un nombre important de protéines ribosomales, sans avoir recours à une purification préalable du complexe macromoléculaire.

### Répartition des protéines identifiées en fonction du pI

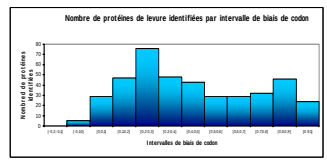


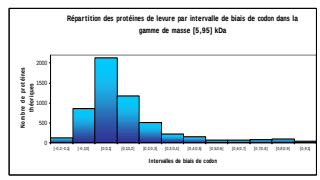



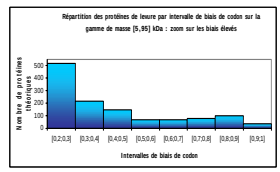




### Répartition des protéines en fonction du biais de codon

Le biais de codon d'un polypeptide est une grandeur permettant d'évaluer son taux d'expression au sein de la cellule. La répartition des protéines identifiées en fonction du biais de codon donne une idée de la gamme dynamique de concentration en protéines analysables par l'approche gel 1D –LC-MS/MS.




L'identification de protéines de biais inférieur à 0,1 est un avantage de la stratégie analytique gel 1D LC-MS/MS, puisque les approches protéomiques classiques reposant sur l'utilisation de gels 2D ne permettent généralement de caractériser que des protéines de biais supérieurs à cette valeur.

## Conclusion

Les potentialités de la technique LC-MS/MS en termes d'analyse de mélanges complexes de protéines ont été testées sur un extrait total de protéines de *Saccharomyces cerevisiae* préalablement séparées sur un gel SDS-PAGE. L'approche gel 1D - LC-MS/MS permet l'identification rapide d'un grand nombre de protéines grâce à des outils informatiques puissants (acquisitions automatiques des spectres de masses, traitement des données expérimentales par l'algorithme Mascot). Cette stratégie présente un certain nombre d'avantages par rapport à la technique classique de caractérisation de mélanges protéiques, qui repose sur l'électrophorèse bidimensionnelle. Notamment, l'analyse des protéines est faite sans restriction sur le point isoélectrique. Les protéines trop hydrophobes pour pénétrer dans la première dimension (isoélectrofocalisation) deviennent accessibles. Enfin, les résultats obtenus montrent que certaines protéines faiblement exprimées, souvent non identifiables à partir de gels 2D, sont repérées même au sein de mélanges contenant des protéines nettement plus abondantes. Cette approche analytique SDS-PAGE-LC-MS/MS peut donc être appliquée à l'étude des protéines présentes dans un extrait cellulaire brut, ou des protéines localisées dans un organelle préalablement purifié. Elle est aussi tout à fait adaptée à l'analyse des composants de complexes multiprotéiques, intervenant dans une cascade de signalisation par exemple.

**Remerciements** : A Geneviève Dujardin et Claire Lemaire, Centre de Génétique Moléculaire, CNRS Gif-sur-Yvette, pour la fourniture de l'échantillon de levure et leur soutien scientifique dans l'interprétation biologique des résultats.

A l'Oréal, Centre Charles Zviak, Clichy, pour le soutien financier