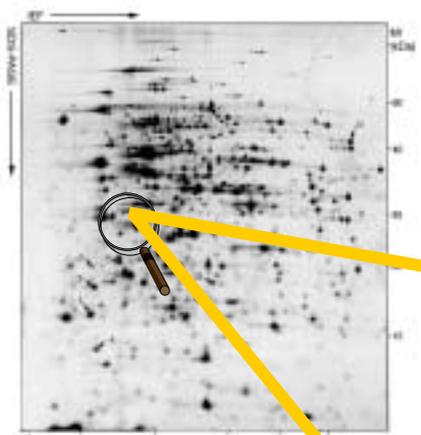


ANALYSE PROTEOMIQUE COMPARATIVE CHEZ *Bacillus subtilis* PAR GELS 2D, MALDI-TOF MS ET LC-MS/MS : EXPLOITATION DES RESULTATS ET ETUDE DE L'INFLUENCE DE L'ETALONNAGE

V. LABAS, J.VINH.
 Neurobiologie et Diversité Cellulaire, ESPCI, CNRS UMR 7637, F75231 Paris Cedex 05
 E. TURLIN, A. Skowska*
 Institut Pasteur. Unité de Génétique des Génomes Bactériens, F75724 Paris cedex 15
 *HKU Pasteur Research Center, Dexter HC Man Bldg, Pokfulam, Hong Kong.



Introduction

L'importance prise ces dernières années par l'analyse génomique engendre une somme croissante d'informations sur les séquences codantes à laquelle s'ajoutent les données quantitatives globales portant sur l'expression des gènes. A cet égard, le protéome reflète les répercussions des événements cellulaires au niveau tant traditionnel que post-traductionnel. De ce point de vue, seule une analyse protéique directe peut donner une image globale des systèmes macromoléculaires dans leur complexité. Une telle approche nous semblait prometteuse pour une étude différentielle.

L'identification par cartographie peptidique a été ici utilisée pour réaliser une analyse protéomique comparative de *B. subtilis* ¹ en fonction des conditions de croissance. Cette bactérie du sol sert de système modèle pour de nombreuses bactéries à Gram-positif proche phylogénétiquement. Nous avons comparé le protéome dans deux états physiologiques de cette bactérie : croissance en milieu solide ou liquide.

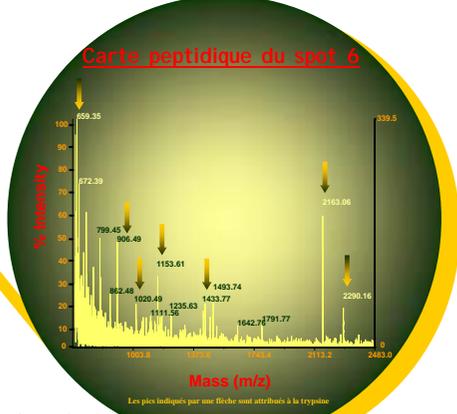
La spectrométrie de masse MALDI-TOF est une technologie puissante, permettant d'identifier un grand nombre de protéines par cartographie peptidique. Cependant cette approche présente quelques limitations : faible quantité de matériel (spots faiblement colorés au nitrate d'argent), problème de détection des protéines de faible (< 15 kDa) ou de très haut (> 100 kDa) poids moléculaires, copurification de plusieurs polypeptides d'intérêt ou contamination (kératines ou autres), extraction spectrale, modifications post-traductionnelles non répertoriées dans les banques, etc...

Il est montré ici que la spectrométrie de masse tandem couplée à la chromatographie liquide de phase inverse (LC-MS/MS) permet de retrouver des protéines qui n'avaient pas pu être identifiées par MALDI-TOF MS et recherche par empreinte peptidique dans les banques de données non-redondantes. Cette technique permet d'améliorer la limite de sensibilité des analyses et de travailler sur des protéines dans un mélange complexe ².

Dans cette étude, on observe également que le traitement manuel des données issues de l'analyse nanoHPLC-Q-TOFet l'utilisation d'un étalonnage interne permettent d'obtenir des précisions de masses qui évitent toute ambiguïté sur l'identité de la protéine en présentant un seul et unique candidat.

Gel d'électrophorèse bi-dimensionnelle

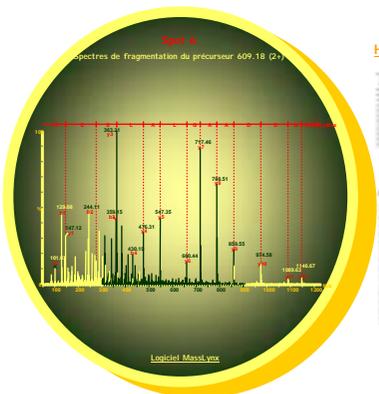
Les cellules sont prélevées au cours de la phase exponentielle. Après centrifugation et lavages, le culot bactérien est recueilli dans 1ml d'une solution Dnaase(10 µg ml⁻¹) / Rnase(5 µg ml⁻¹). Les cellules sont lysées à l'aide de microbilles et les débris cellulaires sont éliminés par ultracentrifugation. Environ 120 µg de protéines sont solubilisés dans 400 µL de solution de réhydratation et chargés sur des gels d'isoelectrofocalisation de gradient de pH, immobilisés 4.7 de 18 cm. Pour cette étude, la première dimension est réalisée à l'aide d'un système horizontal Multigrip II (Pharmacia). Après SDS-PAGE(11.5% à l'aide du système Proteom II et 2D Multigrip (BioRad)), les protéines sont colorées au nitrate d'argent. Les spots sont alors détectés et quantifiés à l'aide du logiciel PDQUEST (DJI, Huntington Station, NY).



Protéines identifiées

Spot	Analyse	ACC	M	pI	M	Nbre peptides Theo	Nbre peptides Obs	Nbre peptides 650-4000 Da	% de couverture	Nbre peptides fragments	Nbre spots
1	Lysozyme C (insectes)[1]	LC-MS/MS	P06909	10.5	8.2	16	16	2	100	2	2
2	Adonitol 1-phosphate[2]	LC-MS/MS	P14304	4.6	21	4	4	2	50	1	2
3	Adonitol 1-phosphate[2]	LC-MS/MS	P14304	4.6	21	4	4	2	50	1	2
4	Tricoumarine (insectes)[3]	LC-MS/MS	P19241	4.7	17	4	4	10	100	1	2
5	Flagellin bactérien	MALDI-TOF MS	SP_309311	4.9	27.49	5	5	15	22	1	2
6	Flagellin (FliC)	LC-MS/MS	P27968	5	32	3	3	3	100	2	2
7	Flagellin Microsporidien	MALDI-TOF MS	P21447	4.6	29.29	4	4	21	22	1	2
8	Oribolobosin[4]	MALDI-TOF MS	P21889	5	30.71	5	5	24	2	1	2
9	Flagellin (insecte)	LC-MS/MS	P12825	5.1	30	5	5	25	2	1	2
10	Homostreptolysine[5]	LC-MS/MS	P19192	5.1	47.4	5	5	25	0	2	2
11	Exopolysaccharide (EPS)	LC-MS/MS	P17796	5.2	26.21	5	5	26	2	1	2
12	Exopolysaccharide (EPS)	MALDI-TOF MS	Q11640	5.2	26.29	5	5	26	2	1	2
13	Exopolysaccharide (EPS)	MALDI-TOF MS	P16442	4.6	26.27	4	4	26	2	1	2

Etalonnage par défaut



Digestion dans le gel et analyses MALDI-TOF

Les spots d'intérêt sont découpés et digérés dans le gel par la trypsine (Roche Diagnostic) selon la méthode décrite par Shevchenko *et al.* ³. Après digestion des protéines, les extraits peptidiques sont dessalés par Zip Tip C18 (Millipore). L'échantillon est déposé sur la cible selon la méthode de la goutte séchée (v/v : 1 : 1) utilisant une solution saturée d'acide 2-5 dihydroxybenzoïque (DHB) en TFA 0.1%. Les cartes peptidiques sont alors obtenues à l'aide d'un spectromètre de masse MALDI-TOF Voyager DE-STR (Applied Biosystems, Framingham, MA), en mode réflectron positif avec une tension d'accélération de 20 000 V, une extraction retardée de 20 ns et environ 250 impulsions laser sont requises pour une acquisition spectrale. L'analyse et le retraitement des données sont effectués par le logiciel Data Explorer (Applied Biosystems). Tous les spectres obtenus sont étalonnés, d'une part automatiquement en externe et d'autre part, en interne sur les produits d'autodigestion de la trypsine. Une erreur de précision de masse inférieure à 25 ppm est tolérée, pour les recherches dans les banques de données non redondantes, par l'intermédiaire du logiciel ProFound.

Recherche par cartographie peptidique



Analyses par LC-MS/MS

L'analyse des protéines digérées par chromatographie liquide couplée à la spectrométrie de masse tandem (LC-MS/MS) a été réalisée à l'aide du système capillaire LC Ultimate (LC Packings, Amsterdam) couplé à un spectromètre de masse de type Quadrupôle-Temps de vol (Q-TOF2, Micromass, Manchester, UK). Le couplage entre la colonne analytique et la source d'ions électrostatron 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 pompage automatique entre les modes d'analyse MS et MS/MS : une seconde d'acquisition en mode MS est suivie de 4 de fragmentation des ions dont le lien 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 protéomiques SWISSPROT, 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 selon l'étalonnage, une tolérance sur la mesure des masses de 5 à 150 ppm en mode MS et de 100 à 300 ppm en mode MS/MS.

Homologie de séquence



Sequence Tag

Résultats et conclusion

Dans le cadre de cette étude, l'association de l'électrophorèse 2D et de la spectrométrie de masse MALDI-TOF a permis d'identifier par MFP (Mass Finger Print) 30% seulement des protéines d'intérêt. Ce résultat s'explique par la faible quantité de matériel présent (spots très légèrement colorés au nitrate d'argent) pour des protéines de masses apparentes de 14 à 45 kDa qui génèrent peu de peptides tryptiques et surtout par la présence de contaminants telles que les kératines. Malgré toutes les conditions de limitations de l'approche MFP, la bonne précision de mesure (tolérance inférieure à 25 ppm), les probabilités très élevées données par le logiciel de recherche ProFound [Long Form] (<http://profound.rockefeller.edu>) et le % de couverture de séquence (supérieur à 20%) ont permis d'identifier, par MALDI-TOF MS, 6 protéines sur 21 sans aucune ambiguïté. Les 15 autres protéines comme par exemple la protéine Flagelline présentée ci-dessus possèdent en dernier candidat par le logiciel ProFound ne sont pas clairement identifiées par cette approche.

L'utilisation de la spectrométrie de masse couplée à la chromatographie liquide de phase inverse (LC-MS/MS) a permis de caractériser 7 autres protéines. L'analyse par nanoHPLC-Q-TOF confirme la présence de la Flagelline, de la trypsine ainsi que des kératines par la fragmentation de 2 peptides ou plus, après une recherche (logiciel Mascot, <http://www.matrixscience.com>) sans aucune restriction de taxonomie ou de masse apparente.

D'une part, il est possible d'exploiter manuellement les spectres de fragmentation suite à l'analyse LC-MS/MS à l'aide du logiciel Mascot.Lyx (Micromass) pour utiliser les modes de recherche tels que l'homologie de séquence (<http://www.ebi.ac.uk/fasta3>) et la recherche par Sequence Tag () dont les résultats sont présentés ci-dessus.

D'autre part, il est montré ici que le traitement manuel des spectres MS (MS Survey) issus de l'analyse LC-MS/MS (avec une calibration par défaut) peuvent être combinés, listés et étalonnés en interne sur au minimum 3 pics de trypsine (m/z : 577.29, 717.36, 1082.03), améliorant ainsi de façon significative les scores d'identification.

L'utilisation d'un étalonnage interne permet d'obtenir une précision de masse < 5 ppm sur les ions précurseurs et une précision de mesure de masse < à 100 ppm sur les ions fragments, permettant ainsi de sélectionner un seul et unique candidat parmi toutes les séquences présentes dans l'organisme étudié.

Après étalonnage interne du MS Survey sur 3 pics de trypsine



Contacts : valerie.labas@espci.fr ; joelle.vinh@espci.fr ; eturlin@pasteur.fr

REMERCIEMENTS
 1 - Kunst F, Ogasawara N, Moszer I, Alberini AM, Alloni G, Azevedo M, Bessieres P, Bolotin A, Borchert S, Borriss R, Boursier L, Brans A, Braun M, Brignell SC, Bron S, Brouillet S, Bruschi CV, Caldwell B, Capuano V, Carter NM, Choi SK, Codan JJ, Connetton IF, Danchin A, et al. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature. 1997 Nov 20;390(6657):249-56.

2 - Pflieger D, Le Caer JP, Lemaire C, Bernard BA, Dujardin G, Rossier J. Anal. Chem. 74 : 2400-2406

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