

N. Royer<sup>1</sup>, V. Labas<sup>1</sup>, E. Turlin<sup>2</sup>, J. Rossier<sup>1</sup>, J. Vinh<sup>1</sup>

<sup>1</sup> Laboratoire de Neurobiologie et Diversité Cellulaire, CNRS UMR7637, ESPCI, 10 rue Vauquelin, 75231 Paris Cedex 05

<sup>2</sup> Institut Pasteur. Unité de Génétique des Génomes Bactériens, F75724 Paris cedex 15

## Introduction

Actuellement, l'analyse protéomique est une composante essentielle à la compréhension des systèmes biologiques. L'identification des protéines doit donc pouvoir être réalisée pour des niveaux d'expression les plus faibles possibles. L'analyse protéomique classique implique, après séparation par électrophorèse (1D ou 2D) et/ou par chromatographie en phase liquide, une protéolyse des protéines suivie de l'analyse des produits de digestion par spectrométrie de masse et finalement la recherche dans les banques de données de séquences. Ceci peut s'avérer fastidieux. C'est pour cette raison que les premières étapes ont été peu à peu robotisées. La mise au point des protocoles pour l'ensemble des étapes est cruciale. Elle doit permettre d'une part de diminuer le temps de manipulation ainsi que les contaminations externes de type kératines, et d'autre part, la reproductibilité de l'analyse. Cette étude porte sur la mise en place d'un protocole robotisé permettant de s'affranchir des interventions humaines pour les étapes de digestion des protéines dans le gel ainsi que la préparation des échantillons pour une analyse en spectrométrie de masse. Afin d'étudier les potentialités de cet automate, différents tests ont été effectués sur des protéines de référence. Ceci nous a permis de valider l'application « protéomique » sur l'automate et d'obtenir un niveau de sensibilité compatible avec l'analyse des protéines colorées au nitrate d'argent. Pour cela, nous avons validé pas à pas l'importance des diverses étapes du protocole. Ces résultats ont été validés dans le cadre de l'analyse protéomique comparative d'*Escherichia coli* pathogène et non pathogène à partir d'extraits protéiques cytoplasmiques et membranaires. Celle-ci a été réalisée sur une gamme de pH allant de 4 à 7, sur des protéines de faible abondance, colorées au nitrate d'argent.

## Protocole manuel de digestion protéique

### La digestion et l'analyse MALDI-TOF

Après un gel SDS-PAGE et une coloration protéique au bleu de Coomassie (BioRad) ou au nitrate d'argent (Amersham Pharmacia Biotech), les protéines sont soumises à une stratégie de préparation d'échantillons (Tableau 1). Les différentes étapes en mode manuel sont des lavages, une réduction/albylation suivie d'une digestion dans le gel par la trypsine bovine modifiée (Roche) à 0,25 ou 12,5 ng/ $\mu$ L selon la méthode de Shevchenko *et al.* (1996<sup>6</sup>). Le temps d'incubation a été porté à 2 heures à 37°C. Après digestion, les extraits peptidiques sont dessalés sur des micro-colonnes de phase greffée octadécyle (Zip-Tip™ C18, Millipore). L'échantillon est préparé en goutte séchée (v/v: 1:1) avec l'acide  $\alpha$ -cyano-4-hydroxyméthylsuccinimide (CHCA) saturé en TFA 0,1% (solvent/vol) (v/v: 1:1). Les cartes peptidiques sont obtenues grâce à un spectromètre de masse type MALDI-TOF Voyager DE-STR (Applied Biosystems, Applied), en mode réflexion positif avec une tension d'accélération de 20 000 V, une extraction retardée de 210 ns. Environ 800 impulsions laser sont moyennées pour une acquisition. Tous les spectres obtenus sont échantillonnés en externe (ACTH 1-17, ACTH 18-39, ACTH 7-38 et angiotensine) puis en interne sur les produits d'autolyse de la trypsine.

### Recherche dans les banques de données

Les identifications cartographiques peptidiques sont menées par recherche dans la banque de données de séquence NCBI<sup>3</sup>, à l'aide du logiciel ProFound. Elles ont été conduites selon les paramètres suivants : taxonomie : selon l'espèce étudiée et manuscrites pour la recherche des kératines, enzymes utilisées : trypsine, 2 masses/valeurs autorisées, modifications variables : carboxyméthylation des Cys, oxydation des Met, masses monoisotopiques MH<sup>+</sup>, tolérance de masse : 50ppm.

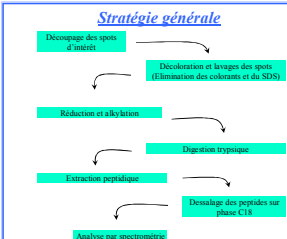


Tableau 1. Stratégie générale

EN MANUEL	EN ROBOTIQUE
Découpage de bandes au scalpel stérile	-
Agitations mécaniques	Aspirations / refoulements
Séchage par centrifugation sous vide (Speed-Vac)	Séchage en hôtels thermostatisés (37 ou 60°C)
Volumes réactionnels de 1.5ml. (tubes Eppendorf)	Volumes réactionnels de 20µl. (micro-plaque de 96 puits)
Viscosité des liquides appréciée visuellement	Création de classes liquides par chaque réactif
Volume de solution enzymatique 30µl.	Volume de solution enzymatique 30µl.
Sonication des tubes	Sonication des plaques
Dépit de 1/6 de l'échant	Dépit de 1/2 de l'échant

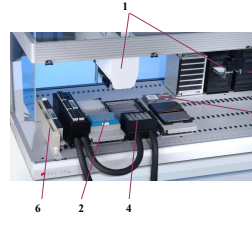
Tableau 2. Paramètres variables entre la digestion manuelle et celle robotique

### Test d'optimisation supplémentaire

Afin d'extraire le plus de matériel biologique des morceaux de gels, nous avons inclus une étape d'extraction peptidique utilisant de l'acétoxytrypsin 100%. Cette étape permet de déshydrater totalement le morceau de gel. Le suragant enrichi est transféré dans une seconde plaque qui sera incubée 1h30 à 50°C pour l'évaporation des solvants. Une fois mis à sécher, les peptides sont réplacés en acide formique 1%, puis fixés sur Zip-Tip™ C18 et élus dans une troisième plaque.

Nous avons par ailleurs étudié les conditions de dépôt sur cible MALDI, en testant plusieurs rapports matrice CHCA / échantillon (1:1, 1:2 et 2:1; v/v).

## Automatisation du protocole



Configuration de l'automate

La configuration du robot (Robot Genesis Proteom 150, TECAN S.A (2)) présent au laboratoire est doté de deux bras manipulateurs (1). Le premier sert aux aspirations et distributions des réactifs. Celui-ci est constitué de 4 agillettes couplées à des serres-joints de 1ml, et de 4 ports cône type P10 couplés à des serres-joints de 20µl. Les ports cône permettent de manipuler les Zip-Tip™ (2) pour traiter 4 échantillons simultanément. Ces Zip-Tip™ sont véhiculés, manipulés et remis sur leur support, permettant ainsi d'être réutilisés après reconditionnement grâce à une mémoire temporaire. Le second bras permet de véhiculer les micro-plaques 96 ou 384 puits sur le plan de travail. Ce bras nommé "table" est constitué d'une griffe à force variable. Le robot, par action de son bras Robo, enlève les plaques traitées de la surface du plan de travail et les véhicule dans un incubateur (3) à température contrôlée (les températures utilisées ici sont 37°C, 50°C et température ambiante).

Les enzymes et autres réactifs sont stockés à la température désirée, généralement 4°C et température ambiante, grâce à un support de micro-tubes et de micro-plaques réfrigérés (4) par un système de refroidissement liquide.

L'automate est doté d'un support pour une cible MALDI (5) afin d'effectuer les dépôts directement sur cible MALDI. Au cours de ces dépôts, de très faibles volumes sont utilisés. Ils sont gérés par un système d'aspersion standard. Le robot est muni d'une station de lavage (6) dotée de deux compartiments permettant successivement le lavage de l'extrémité de l'agillette et l'autre permettant de nettoyer l'agillette sur une grande partie de sa hauteur. Au centre de cette station de lavage se situe une poubelle, permettant de rejeter les réactifs usagés contenus dans les agillettes.

### Paramètres mis au point et optimisés sur le robot

Pour la mise au point, nous avons travaillé avec trois protéines standards: la Sérurum Albumine Bovine (BSA), l'Ovalbumine (OVAL) et le Lysozyme de poulet (LYSO). Ces protéines ont été, comme pour les analyses manuelles, colorées au bleu de Coomassie et au nitrate d'argent.

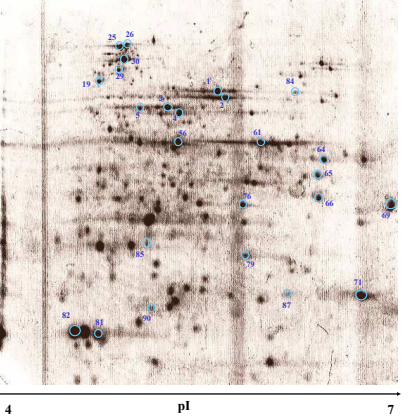
L'automatisation du protocole exige quelques adaptations (Tableau 2). Les paramètres qui ont été modifiés et adaptés au robot ne sont pas des paramètres d'ordre biochimique mais plutôt de manipulation physique. En effet les problèmes se rencontrent surtout au niveau des actions portées sur l'échantillon lors de l'expérience. Les différents paramètres ayant fait l'objet d'une étude sont les suivants :

- le type de support de travail. Avec le protocole manuel, les échantillons sont traités dans des micro-tubes, alors qu'avec le robot le traitement se fait ici dans des micro-plaques 96 puits. Les plaques utilisées sont identiques dans leur nature (polyéthylène) mais légèrement différents en terme de traitement de surface (mouillabilité, adsorption des échantillons, forme et taille des réactifs...)
- l'agitation des échantillons. En mode manuel, un agitateur rotatif est utilisé. Les étapes d'agitation robotisées ont été remplacées par des aspirations-refoulements répétés.
- le séchage. Afin d'automatiser entièrement le protocole de digestion enzymatique et de supprimer toute intervention humaine (gain de temps et contamination), les étapes de séchage effectuées manuellement au speed-vac ont été remplacées par des étapes d'incubation dans les hôtels thermostatisés. Les temps d'incubation dans les hôtels thermostatisés ont été optimisés pour chaque étape demandant normalement séchage.
- le volume de trypsine distribué. Manuellement, le morceau de gel est mis en présence d'excès de trypsine. Celui-ci est retiré après une incubation de 40 minutes à 4°C. Le protocole robotisé utilise 5µl de trypsine (température ambiante), volume de solution enzymatique minimal requérant pour le réhydratation d'un morceau de gel.
- la viscosité. La viscosité des réactifs diffère au cours du protocole et il faut en tenir compte. Différentes « classes liquides » ont été répertoriées, pour déterminer des vitesses d'aspersion et de distributions ainsi que d'autres paramètres pour un réactif donné.

## Validation biologique du protocole de digestion robotique

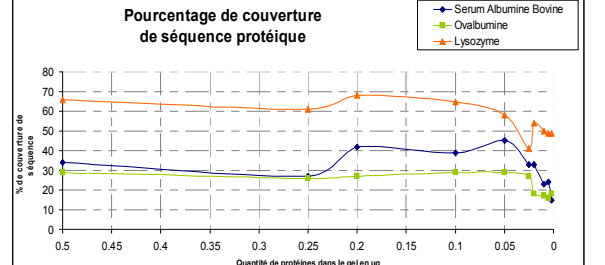
### Gel d'électrophorèse bi-dimensionnelle

Les cellules sont prélevées au cours de la phase exponentielle. Après centrifugation et lavages, le culot bactérien est resuspendu dans 1ml. d'une solution DNase (10 mg ml<sup>-1</sup>) / RNase (5 mg ml<sup>-1</sup>). 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 immobilisé 4-7 de 18 cm. Pour cette étude, la première dimension est réalisée à l'aide d'un système horizontal Multiphor II (Pharmacia). Après SDS-PAGE 11.5%, les protéines sont colorées au nitrate d'argent.



N° ECH	N° accession NCBI	Nom de la protéine	PI	MW	PI exp.	PI app.	% de couv.	Nbre peptides identifiés
1	gi198912945 g 198912945	Lysine-tartrate lyase	5.7	31.04	5.6	5.6	6	5
2	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
3	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
4	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
5	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
6	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
7	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
8	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
9	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
10	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
11	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
12	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
13	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
14	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
15	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
16	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
17	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
18	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
19	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
20	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
21	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
22	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
23	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
24	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
25	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
26	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
27	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
28	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
29	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
30	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
31	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
32	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
33	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
34	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
35	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
36	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
37	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
38	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
39	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
40	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
41	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
42	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
43	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
44	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
45	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
46	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
47	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
48	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
49	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
50	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
51	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
52	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
53	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
54	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
55	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
56	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
57	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
58	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
59	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31
60	gi198912945 g 198912945	Chaperon chaperonin 60	5.5	41.15	5.7	5.4	31	31

Tableau 3. Résultats des identifications par MALDI-TOF



µg	0,5	0,25	0,1	0,05	0,025	0,01	0,005	0,002
BSA	7500	3750	3000	1500	750	375	300	150
OVAL	11500	5750	4600	2300	1150	575	460	230
LYSO	11000	5500	4400	2200	1100	550	440	220

Graphique 1: Résultats obtenus sur l'optimisation de la digestion de protéines avec le robot

### Observation des résultats de l'optimisation

Le protocole utilisé pour digérer les trois protéines BSA, OVAL et LYSO colorées au nitrate d'argent ou au bleu de Coomassie, ne présente aucune différence à l'exception de la concentration de la trypsine. La gamme de quantité de protéines utilisées (0,5µg à 2 ng) a permis de déterminer les limites de sensibilité pour une bonne identification. En effet, les analyses ont montré que pour chacune des protéines standards identifiées, nous obtenons un pourcentage de couverture de séquence relativement élevé, entre 20 et 45% pour la BSA, 20 et 30% pour OVAL et entre 50 et 70% pour le LYSO (Graph 1). Le nombre de peptides générés par la digestion tryptique est également assez élevé (de 7 à 31 peptides pour la BSA, 4 à 11 pour OVAL et 7 à 15 pour le LYSO). Il a été montré que la diminution de la concentration de trypsine dans le milieu permet, pour les faibles quantités de protéines, d'obtenir des spectres dans lesquels les pics d'autodigestion sont minoritaires. Les phénomènes d'extinction spectrale sont ainsi atténués. Ces expériences ont révélées également qu'il est indispensable d'inclure des étapes de sonication afin d'éviter les phénomènes de collages et ainsi de récupérer l'ensemble des peptides et d'identifier les protéines en faibles quantités.

Les pourcentages de couverture de séquence protéique sont sensiblement les mêmes dans la gamme suivante : 0,5 à 0,1 µg. En effet, lors d'une automatisation, à l'étape des dépôts, tous les échantillons sont traités à l'identique alors que les quantités de peptides varient en fonction de la gamme. Par conséquent, les différents tests effectués avec différents rapports matrice / échantillon ont montré qu'il est préférable d'utiliser le rapport matrice / échantillon (1:2; v/v) qui s'est révélé être optimale pour l'identification des protéines présentes en faibles quantités. Ce mode de dépôt utilise également pour des fortes quantités de peptides permet d'identifier facilement les protéines par cartographie peptidique mais n'est pas toujours adapté à certaines recherches de modifications post-traductionnelles qui demandent des conditions de préparation plus fines. Les tests supplémentaires portant sur une extraction peptidique présentent une trop grande variabilité qui ne nous permet pas de conclure à l'heure actuelle sur sa nécessité. En effet, l'utilisation d'une troisième plaque entraîne des pertes peptidiques par collage difficilement contrôlables.

### Résultats de la validation biologique

Le protocole robotisé mis au point pour le traitement d'échantillons colorés au nitrate d'argent a été appliqué dans le cadre de cette étude. La réalisation d'une carte de référence sur un organisme bien connu tel que *Escherichia coli* nous a permis d'évaluer la sensibilité d'analyse atteinte. En effet, les analyses ont été réalisées sur 90 spots choisis aléatoirement de forte à faible intensité de coloration, sur des points isoelectriques (pI) et des masses moléculaires (MW) variés. Le tableau ci-dessous (Tableau 3) comporte quelques exemples d'identifications. Une grande partie des protéines ont été identifiées sans aucune ambiguïté, avec un nombre élevé de peptides attribués (jusqu'à 15) et un pourcentage de couverture supérieur à 20%. Cependant, certaines identifications présentent un faible pourcentage de couverture restant à confirmer par d'autres analyses de type MS/MS. Par ailleurs, certaines cartographies présentant de nombreux pics non attribués suggèrent que l'échantillon est très hétérogène. Sachant que l'analyse de protéines en mélange est l'une des limitations de l'analyse par cartographie peptidique, il sera par conséquent nécessaire d'analyser ces échantillons par des techniques de type LC-MS/MS. Cette validation a également révélé qu'il est nécessaire d'inclure des étapes de sonication afin d'éviter les phénomènes de collage des peptides et ainsi de récupérer l'ensemble des peptides et d'identifier les protéines présentes en faibles quantités dans le gel.

### Conclusion sur cette automatisation

L'analyse des 3 protéines standards sur une gamme de 500 à 2 ng a permis de mettre en évidence une bonne reproductibilité et répétabilité pour un protocole donné et d'atteindre une sensibilité suffisante pour être compatible avec les outils de spectrométrie de masse tel que le MALDI-TOF. En effet la gamme de protéines utilisées pour les tests correspond à celles généralement rencontrées dans les applications protéomiques. L'automatisation permet de diminuer le coût de manipulation; la diminution est croissante avec le nombre d'échantillons traités (20 échantillons: réduction de 1% du coût; 96 échantillons: réduction de 23%). Le protocole de digestion robotisé est plus long que le protocole manuel mais il nécessite pas d'intervention humaine! L'automatisation du protocole a permis d'éliminer la quasi totalité des contaminations par les kératines durant une manipulation de digestion. Les travaux effectués sur