

Protéomique du cancer du sein : du différentiel au fonctionnel

par Hubert Hondermarck

JE 2488 ERI-8 INSERM « Signalisation des facteurs de croissance dans le cancer du sein », IFR-118, Université des Sciences et Technologies de Lille, 59650 Villeneuve d'Ascq. Tél. : 03 20 43 40 97.

Email : hubert.hondermarck@univ-lille1.fr

Reçu le 3 mars 2006

RÉSUMÉ

De l'approche différentielle destinée à l'identification de marqueurs à l'analyse fonctionnelle pour la mise en évidence de cibles thérapeutiques, la protéomique apporte de nouvelles données fondamentales pour la compréhension des mécanismes de la cancé-

rogenèse mammaire et des perspectives pour la clinique. Néanmoins la protéomique, comme l'ensemble des approches de post-génomique, doit faire face à la complexité cellulaire et moléculaire de la glande mammaire qui reste un défi majeur.

SUMMARY Proteomics of breast cancer: from differential to functional analysis

From differential analysis to identify biomarkers, to functional analysis for finding new therapeutic targets, proteomics bring new comprehensive information for a better understanding of the molecular basis of oncology and new perspectives for the clinic. How-

ever the major limitation of proteomic investigations, more generally of post-genomic approaches, remains the molecular and cellular complexity of the mammary gland that is still a major challenge.

LE PROTÉOME : NOUVEAU CHAMP D'EXPLORATION BIOMÉDICAL

Il y a encore quelques années, il était difficile de défendre l'idée que l'étude de l'ensemble des protéines (protéome) de cellules tumorales, de biopsies ou de fluides biologiques était une façon pertinente d'identifier de nouveaux marqueurs ou de nouvelles cibles thérapeutiques pour le cancer. En effet, la génétique moléculaire était triomphante et l'analyse des acides nucléiques était censée permettre d'établir les causes des maladies (mutations) qui pourraient éventuellement être corrigées (thérapie génique) pour que tout rentre en ordre. Chaque période scientifique a ses mythes et avec la fin du séquençage du génome humain il est apparu que la seule connaissance du matériel génétique ne permettait pas la compréhension absolue ni le traitement des mécanismes moléculaires menant aux pathologies comme le cancer du sein. Aujourd'hui il est devenu nécessaire de dire que les protéines sont en fait les effecteurs fonctionnels des gènes et que la fonction des gènes ne peut se comprendre qu'à travers l'étude des protéines correspondantes. Le protéome est ainsi devenu un centre d'intérêt primordial pour la recherche de marqueurs pathologiques et de cibles thérapeutiques (Hondermarck, 2004). Les développements technologiques qu'a connus la protéomique ces dernières années, notamment avec la spectrométrie de masse, la chimie et la bioinformatique, permettent

l'exploration fine du protéome et l'analyse des interactions protéine-protéine qui sont à la base des phénomènes de régulation du métabolisme et du phénotype cellulaire (Hondermarck, 2003 ; Celis *et al.*, 2005).

DU GÉNOME AU PROTÉOME : REDÉCOUVERTE DE LA COMPLEXITÉ

Il semble maintenant clair que le nombre de gènes humains est proche de 25 000. Ces gènes sont constitués de quatre éléments chimiques de base (les nucléotides ATCG) et sont alignés le long d'une molécule unique, l'ADN (essentiellement contenu dans un compartiment cellulaire le noyau). Les nucléotides de l'ADN ne subissent que de rares modifications chimiques (méthylation notamment). Il y a deux allèles par gène dans toutes les cellules et toutes les cellules de l'organisme possèdent le même génome qui reste stable et ne varie pas en fonction des conditions physiologiques dans lesquelles les cellules se trouvent. Enfin, les gènes n'interagissent que peu physiquement les uns avec les autres. Avec le recul, que tout cela paraît simple. Les protéines sont constituées de 20 éléments chimiques de base (les acides aminés majoritaires) qui peuvent être modifiés de multiples façons : environ 200 types de modifications post-traductionnelles (phosphorylation, glycosylation, amidation, ubiquitination, etc.) ont été décrites à ce jour. Les pro-

téines sont présentes en multiples copies et sont distribuées dans l'ensemble des compartiments cellulaires et nul ne sait dire actuellement combien d'espèces protéiques existent chez l'Homme. Les évaluations vont de 500 000 à 5 millions et nous sommes à peu près dans l'état d'incertitude qui était de mise sur le nombre de gènes au moment du démarrage du séquençage du génome. A cela s'ajoute que le protéome est une entité variable d'un type cellulaire à l'autre (il existe chez l'humain environ 200 types cellulaires de base, donc 200 protéomes de base) et d'une condition physiologique à une autre puisque les cellules ajustent leur contenu protéique en fonction des conditions dans lesquelles elles se trouvent (en cas de stress par exemple). Enfin, les protéines interagissent physiquement les unes avec les autres dans le cadre de réseaux d'interactions (maintenant appelés interactomes) et ces interactions se traduisent par des modifications telles que la phosphorylation qui changent la conformation et donc l'activité des protéines. La conséquence de cette complexité est que la protéomique doit employer un arsenal de technologies complémentaires allant des méthodes séparatives comme l'électrophorèse bi-dimensionnelle aux approches physicochimiques comme la spectrométrie de masse, la cristallographie et la RMN notamment. De ce fait, l'analyse à haut débit reste plus un objectif théorique qu'une réalité en protéomique.

PROTÉOMIQUE DIFFÉRENTIELLE DU CANCER DU SEIN

L'idée est simple de comparer des échantillons normaux et cancéreux pour définir les spécificités moléculaires propres à la cancérisation. Dans le sein, l'exploration du contenu protéique avait commencée bien avant la protéomique et des marqueurs ont été identifiés par électrophorèse bidimensionnelle, comme certaines cytokératines ou tropomyosines, dès les années 80. La célèbre cathepsine D, protéine marqueur des cancers du sein hormono-dépendants fait partie de cette série de marqueurs initialement identifiés. Cependant, il faut reconnaître que depuis lors, peu de nouvelles protéines ont été identifiées et les récepteurs aux estrogènes constituent toujours les marqueurs de référence dans le cancer du sein. De fait, il est maintenant établi que le protéome de cellules cancéreuses ne subit qu'un nombre limité de modifications par rapport au protéome des cellules normales. Ceci traduit probablement le fait que la cancérisation d'une cellule nécessite un nombre limité de modifications moléculaires. Au niveau génique, il a d'ailleurs été montré que l'on peut induire la cancérisation avec seulement quelques modifications (mutations de quelques oncogènes et gènes suppresseurs) et le même profil semble se dessiner au niveau protéique. Autre constatation, il existe une hétérogénéité, une diversité des tumeurs mammaires ainsi qu'une diversité des mécanismes menant à la cancérisation. Des différences en nombre limité et une variabilité d'une tumeur à l'autre, voilà les difficultés principales de l'analyse différentielle. Cependant de nouveaux marqueurs ont été identifiés plus récemment. C'est le cas notamment de la chaperonne moléculaire 14-

3-3 σ . Le taux de protéine 14-3-3 σ est 10 fois moindre dans les cellules de cancer du sein que dans les cellules mammaires normales et son niveau apparaît très faible dans les biopsies de cancer du sein (Vercoutter *et al.*, 2001). De façon complémentaire à l'analyse protéomique, les méthodes de la génomique ont montré une hyperméthylation du gène de 14-3-3 σ dans les cellules de cancer du sein, entraînant une perte d'expression des ARNm de 14-3-3 σ dans les tumeurs mammaires (Hermeking *et al.*, 2003). 14-3-3 σ est maintenant considérée comme un marqueur précoce de la cancérisation mammaire et cette protéine suscite également de l'intérêt pour son rôle potentiel dans le contrôle du phénotype tumoral. En définitive, la mise en évidence de 14-3-3 σ comme suppresseur de tumeur illustre la complémentarité des approches de la génomique et de la protéomique pour la mise en évidence de modifications moléculaires caractéristiques du processus de cancérisation. Un autre exemple est constitué par le facteur de croissance NGF (nerve growth factor). Facteur neurotrophique par excellence, le NGF est bien connu pour son rôle stimulateur de la survie et la différenciation des neurones au cours du développement du système nerveux. De façon inattendue, il a été montré que ce facteur de croissance est fortement exprimé dans les cellules de cancer du sein alors qu'il n'est pas trouvé dans les cellules épithéliales mammaires normales (Dollé *et al.*, 2003). Il est encore trop tôt pour savoir si 14-3-3 et le NGF seront des marqueurs d'intérêt clinique et éventuellement des cibles thérapeutiques, mais l'exemple du récepteur membranaire Erb-B2 et son utilisation en clinique (ciblage de l'Herceptine) montre que les protéines de faible niveau d'expression présente un intérêt considérable. Or la sensibilité est précisément encore l'une des limites des méthodes actuelles en protéomique. Bien que la spectrométrie de masse permette une sensibilité de l'ordre du picogramme sur une protéine pure, l'analyse d'échantillons réels comme des extraits cellulaires ou tissulaires contenant plusieurs centaines de milliers d'espèces protéiques différentes exprimées à des niveaux très différents s'avère difficile. Un aspect incontournable, quelle que soit la méthodologie utilisée, est que la présence de protéines majoritaires gêne ou empêche la détection des protéines minoritaires. Les solutions pour contourner ces difficultés peuvent consister à enlever par chromatographie les protéines majoritaires quand cela est possible (par exemple l'albumine quand on analyse des échantillons sanguins) ou fractionner les compartiments subcellulaires (par centrifugation différentielle par exemple). Malheureusement ces solutions ne constituent que des améliorations à portée limitée à cause de la gamme de concentrations observée pour les protéines qui s'étale sur plus de 10 log.

Une difficulté supplémentaire dans l'identification de marqueurs de cancérisation est apportée par l'hétérogénéité cellulaire des tumeurs. Les tumeurs du sein sont principalement composées de cellules épithéliales cancéreuses, mais elles contiennent également d'autres types cellulaires comme des cellules épithéliales normales, des cellules myoépithéliales, des fibroblastes, des adipocytes, des cellules endothéliales et des micro-fibres nerveuses. La présence de ces différentes catégories cellulaires complique

sérieusement l'analyse des profils protéomiques obtenus. En théorie, les techniques de microdissection laser rendent maintenant possible la séparation des cellules tumorales avant analyse. Par ailleurs, il ne faut pas négliger la possibilité que les cellules non tumorales d'une tumeur puissent exprimer des marqueurs d'intérêt. Prenons l'exemple des cellules endothéliales : on sait que l'angiogenèse (formation de vaisseaux et capillaires sanguins) est un phénomène crucial pour le développement tumoral et l'identification de marqueurs endothéliaux dans les tumeurs mammaires pourrait s'avérer utile à l'établissement du pronostic. Grâce à la microdissection il est maintenant envisageable à terme de réaliser l'analyse protéomique de presque tous les types cellulaires présents dans les tumeurs mammaires. Cependant, là encore les difficultés liées à la complexité du monde des protéines diminuent la réalité des potentialités théoriques. Première difficulté, il n'existe pas de système équivalent à la PCR pour amplifier les protéines et l'analyse protéomique suppose des quantités de matériels biologiques incompressibles. A titre d'exemple, il faut au minimum environ 50 000 cellules pour établir un profil protéomique de base (par électrophorèse 2D ou LC-MS/MS). Mais à partir de tels profils, seules 1 000 protéines majoritaires peuvent être identifiées et le fractionnement nécessaire à l'exploration des protéines plus minoritaires requiert des quantités supplémentaires qui deviennent rapidement incompatibles avec la capacité de la microdissection. En conclusion, sans système d'amplification l'analyse protéomique d'échantillons tumoraux microdisséqués reste difficile.

PROTÉOMIQUE FONCTIONNELLE POUR LA DISSECTION DES VOIES DE SIGNALISATION

Au-delà de l'analyse différentielle, la protéomique offre maintenant la possibilité d'explorer les interactions protéine-protéine et les réseaux de signalisation intracellulaires qui gouvernent le phénotype et plus généralement le comportement des cellules cancéreuses. Ce domaine d'activité que l'on appelle protéomique fonctionnelle est actuellement en plein essor et est d'un intérêt primordial pour l'analyse des mécanismes moléculaires de la cancérisation. Des interactions ligand-récepteur, à l'activation de voies de signalisation comme les MAP-kinases ou la régulation de l'activité des facteurs de transcription, rien n'échappe dorénavant à l'outil protéomique avec comme apport essentiel de permettre une approche sans *a priori*, dans laquelle de nouveaux partenaires non suspectés pourront être découverts après analyse en spectrométrie de masse. Ces potentialités font maintenant appel à des méthodologies spécifiques de mise en évidence des interactions, des marquages par isotope stables dont la détection en spectrométrie de masse peut permettre de mettre en évidence une protéine mais aussi de suivre ses modifications post-traductionnelles (Aebersold *et al.*, 2003). Dans le domaine du cancer du sein, le décryptage des voies de signalisation des facteurs de croissance et des hormones constitue une

piste pour la mise en évidence de nouvelles cibles thérapeutiques. Une illustration de la potentialité de ces approches est fournie par les travaux réalisés avec le facteur de croissance FGF. La cycline D s'est ainsi révélée une cible de la signalisation de ce facteur de croissance (Vercoutter *et al.*, 2000) et il a été mis en évidence que les protéines de stress, en particulier HSP-90 (*Heat Shock Protein 90*) voient leur synthèse augmentée de façon importante dans les premières minutes, de l'activation des cellules de cancer du sein par le FGF-2 (Vercoutter *et al.*, 2001). De façon intéressante, l'inhibition de l'activité de HSP90 (par la drogue geldanamycine) entraîne un blocage de l'effet mitogène du FGF-2 et se traduit par un blocage de la prolifération et de la migration des cellules de cancer du sein. Ainsi, HSP90 est nécessaire à l'activation de la croissance des cellules de cancer du sein et cette protéine constitue une cible thérapeutique potentielle pour la mise en place de traitements plus ciblés du cancer du sein. Plus récemment, la protéine thioredoxine a été montrée être activée par l'expression d'oncogènes (Lottin *et al.*, 2002), TRADD a été identifié comme partenaire du récepteur Trk dans les cellules de cancer du sein par une démarche reposant au départ sur la protéomique (El-Yazidi *et al.*, 2003) et des travaux plus récents montrent l'identification de formes particulières des kinases IKK et AKT dans la signalisation de la survie des cellules de cancer du sein (Vandermoere *et al.*, 2005, 2006). Il est important de souligner ici que l'utilisation de la protéomique pour l'analyse de la signalisation nécessite la complémentarité avec d'autres approches de biologie cellulaire et moléculaire. La réalité d'une interaction devra toujours être confirmée en microscopie confocale afin de confirmer dans la cellule ce qui est observé en tube à essai et l'effet d'une inhibition de la protéine d'intérêt devra être étudié après inhibition de la synthèse et/ou de l'activité de cette protéine par ARN interférence ou emploi de dominants-négatifs. Il est également important de toujours vérifier de façon ultime l'impact sur le phénotype cellulaire (prolifération, migration, différenciation, croissance sans ancrage). Ainsi, dans le domaine de l'analyse de la signalisation des cellules cancéreuses, la protéomique est un outil qui s'inscrit nécessairement dans une démarche plus globale et qui complète l'arsenal technologique traditionnel.

DE L'ÉPOQUE DES PIONNIERS À LA CONSTITUTION DE PROGRAMMES INTERNATIONAUX

Dans le cas du cancer du sein mais aussi dans les autres domaines biomédicaux, la protéomique n'est plus maintenant l'apanage de quelques spécialistes mais commence à être plus largement utilisée. Cela correspond tout d'abord à une démocratisation des méthodes et un accès plus facile aux bases de données. Bien sûr les plateformes technologiques vont continuer à concentrer les équipements lourds de spectrométrie de masse, mais les méthodes séparatives peuvent maintenant s'employer de façon plus simple et à cet égard le développement de

l'approche de chromatographie de surface SELDI (*surface enhanced laser desorption ionisation*) aura été essentiel au développement de la protéomique en milieu clinique. À côté des équipements se pose également la question de la formation des personnels et il est important de favoriser les formations de protéomique pour permettre un essaimage des personnes formées dans les laboratoires. Du point de vue de la formation et de l'assistance, la France bénéficie de l'existence d'une société savante active, la SFEAP (Société Française d'Électrophorèse et d'Analyse Protéomique, www.sfeap.free.fr) qui organise des journées scientifiques régulières, des ateliers de formation et congrès annuels. Au niveau mondial, le programme international HUPO (*HUMAN Proteome Organization*, www.hupo.org) a pour vocation de stimuler et coordonner l'analyse du protéome humain et, comme HUGO à son époque, il est sous-tendu par les potentialités d'applications biomédicales : la mise en évidence de marqueurs et cibles thérapeutiques. Comment les choses vont s'organiser au sein de HUPO est une question en cours de débat, mais l'étude des cancers y tiendra bien sûr une place importante. La standardisation des procédures de collecte des échantillons et d'analyse, la définition de guides des bonnes pratiques, la centralisation bioinformatique des données et la coordination au niveau mondial seront des questions prioritaires. En France, le Réseau National des Génopoles et l'Inserm réfléchissent à ces questions et mettent en place une articulation avec HUPO.

La protéomique du cancer du sein doit encore effectuer le passage en phase d'application clinique. Les méthodes actuellement employées pour la recherche de marqueurs et cibles, comme l'électrophorèse bidimensionnelle, le marquage isotopique et la spectrométrie de masse ne peuvent être utilisées pour l'analyse de grandes séries tumorales et il apparaît nécessaire de développer des méthodes d'analyse plus légères pour une utilisation plus routinière. L'espoir dans ce domaine vient du développement des biopuces à protéines. Le concept des puces à protéines est basé sur un principe analogue aux puces à acides nucléiques avec l'imprégnation de substrats (comme des anticorps) sur un support permettant une quantification de différentes protéines d'un échantillon; la mise en place et la lecture de "microarrays" permettant théoriquement une analyse en haut débit de plusieurs milliers de protéines par échantillons. Il faut bien reconnaître que les puces à protéines sont encore actuellement en phase de validation technologique, mais il est clair que leur potentiel d'application dans le domaine biomédical est important. Grâce à ces technologies, on peut imaginer pouvoir bientôt établir, pour chaque tumeur mammaire (donc pour chaque patiente), une « carte d'identité moléculaire » au niveau protéique qui déboucherait sur une classification précise de la tumeur. La conséquence serait une prise en charge personnalisée de chaque patiente, avec en particulier le choix de cibles thérapeutiques appropriées et donc la mise en place de traitements spécifiques. L'intégration des données issues des analyses génomiques, protéomiques et cliniques (Hanash *et al.*, 2004; Celis *et al.*, 2004) nécessitera également des développements bio-

informatiques importants et le passage de la protéomique de la paillasse des laboratoires spécialisés au milieu clinique est certainement un défi majeur des années à venir.

BIBLIOGRAPHIE

- Aebersold R. & Mann M., Mass spectrometry-based proteomics. *Nature*, 2003, 422, 198-207.
- Celis J. E., Moreira J. M., Gromova I., Cabezon T., Ralfkiaer U., Guldberg P., Straten P. T., Mouridsen H., Friis E., Holm D., Rank F. & Gromov P., Towards discovery-driven translational research in breast cancer. *FEBS J.*, 2005, 272, 2-15.
- Celis J. E., Gromov P., Cabezon T., Moreira J. M., Ambartsoumian N., Sandelin K., Rank F. & Gromova I., Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. *Mol. Cell. Proteomics*, 2004, 3, 327-344.
- Dollé L., El-Yazidi I., Adriaenssens E., Nurcombe V. & Hondermarck H., Nerve growth factor overexpression and autocrine loop in breast cancer cells. *Oncogene*, 2003, 22, 5592-5601.
- El Yazidi- Belkoura I., Adriaenssens E., Dollé L., Descamps S. & Hondermarck H., Tumor necrosis factor receptor-associated death domain protein (TRADD) is involved in the neurotrophin receptor (p75^{NTR})-mediated anti-apoptotic activity of nerve growth factor in breast cancer cells. *J. Biol. Chem.*, 2003, 278, 16952-16956.
- Hanash S., Integrated global profiling of cancer. *Nat. Rev. Cancer*, 2004, 4, 638-644.
- Hermeking H., The 14-3-3 cancer connection. *Nat. Rev. Cancer*, 2003, 3, 931-943.
- Hondermarck H., Breast cancer: when proteomics challenges biological complexity. *Mol. Cell. Proteomics*, 2003, 2, 281-291.
- Hondermarck H., Biomedical and pharmaceutical applications of proteomics, Kluwer-Springer ed., Dordrecht, 2004. The Netherlands. ISBN : 1-4020-2322-7.
- Lottin S., Vercoutter-Edouard A. S., Adriaenssens E., Czeszak X., Lemoine J., Roudbaraki M., Coll J., Hondermarck H., Dugimont T. & Cury J. J., Thioredoxin post-transcriptional regulation provides a new function to non-coding mRNA. *Oncogene*, 2002, 21, 1625-1631.
- Vandermoere F., El Yazidi-Belkoura I., Adriaenssens E., Lemoine J. & Hondermarck H., The antiapoptotic effect of fibroblast growth factor-2 is mediated through nuclear factor- κ B activation induced via interaction between Akt and I κ B kinase- β in breast cancer cells. *Oncogene*, 2005, 24 (35), 5482-5491.
- Vandermoere F., El Yazidi-Belkoura I., Slomianny C., Demont Y., Bidaux G., Adriaenssens E., Lemoine J. & Hondermarck H., The valosin-containing protein (VCP) is a target of AKT signaling required for cell survival. *J. Biol. Chem.*, 2006, 281, 14307-14313.
- Vercoutter-Edouard A., Lemoine J., Smart C. E., Nurcombe V., Boilly B., Peyrat J. & Hondermarck H., The mitogenic signaling pathway for fibroblast growth factor-2 involves the tyrosine phosphorylation of cyclin D2 in MCF-7 human breast cancer cells. *FEBS Lett.*, 2000, 478, 209-215.
- Vercoutter-Edouard A. S., Lemoine J., Le Bourhis X., Louis H., Boilly B., Nurcombe V., Revillion F., Peyrat J. P. & Hondermarck H., Proteomic analysis reveals that 14-3-3 σ is down-regulated in human breast cancer cells. *Cancer Res.*, 2001, 61, 76-80.
- Vercoutter-Edouard A. S., Czeszak X., Crepin M., Lemoine J., Boilly B., Le Bourhis X., Peyrat J. P. & Hondermarck H., Proteomic detection of changes in protein synthesis induced by fibroblast growth factor-2 in MCF-7 human breast cancer cells. *Exp. Cell. Res.*, 2001 262, 59-68.