

## L'APPROCHE PROTEOMIQUE ET SES APPLICATIONS

CARMEN PALII<sup>1\*</sup>, LAURA MITROFAN<sup>1</sup>,  
JEAN MONTREUIL<sup>2</sup>, VLAD ARTENIE<sup>1</sup>

**Mots – clés** : protéome, solubilisation, électrophorèse, spectrométrie de masse.

**Résumé** : L'analyse protéomique constitue une science jeune qui permette l'étude des protéines que ce soit constituantes, enzymes du métabolisme comme marqueurs de l'état cellulaire, ou médiateurs des régulations liées à la dynamique des différents états comme la pathologie ou la différenciation. La génomique a permis d'accroître nos connaissances en cancérologie moléculaire et dans le domaine de la prédisposition génétique. L'utilisation de la protéomique permet de compléter ces acquis, en particulier pour une meilleure compréhension des mécanismes physiopathologiques de la fonction des gènes et pour la recherche des marqueurs diagnostiques ou des cibles thérapeutiques. Les avancées de la biochimie des protéines dépendent des avancées technologiques de cette approche: amélioration des performances de la 2D-PAGE et des outils de détection qui y sont associés comme l'automatisation de la manipulation des échantillons, nouveaux développements de la spectrométrie de masse et de la bioinformatique.

### INTRODUCTION

Le mot « protéome » est apparu dans la littérature scientifique en 1995. Sa définition officielle, telle qu'énoncée par Marc Wilkins il y a près de huit ans, évoque « le complément protéinique d'un génome ». Aujourd'hui de très nombreux articles sont référencés dans les banques de données bibliographiques sous ce mot-clé. En 2001, Science Direct en comptait 2190.

Les progrès de la Biologie Moléculaire amènent aujourd'hui à considérer les applications pratiques des études post-génomiques. En effet, un gène unique peut générer une protéine unique, aucune protéine (pseudo-gène), ou plusieurs protéines (modifications post-traductionnelles, différences d'épissage des exons et de clivages protéolytiques). Cette situation a conduit depuis quelques années à l'émergence du concept de protéome. Le terme de protéome désigne l'ensemble des protéines exprimées par un génome à un moment précis en réponse à un environnement donné (Wilkins *et al.*, 1997). L'analyse protéomique permet de caractériser et de quantifier les protéines dans des conditions déterminées.

La protéomique peut donc être définie comme la discipline de la « comparaison quantitative de protéomes similaires sous différents stimuli permettant une meilleure compréhension des processus biologiques complexes ». Le concept d'analyse protéomique (proteomics, en anglais), est né *a posteriori* pour désigner l'utilisation de la quantification de la protéine comme mesure objective de l'expression génique caractérisant un processus biologique donné (effets des médicaments, de l'environnement etc.) et comme moyen de décodage des mécanismes contrôlant cette expression (Anderson *et al.*, 1998). L'analyse protéomique, dans l'acceptation de cette définition, met l'accent sur la description dynamique de la régulation des gènes et ce faisant sur la régulation moléculaire étudiée au travers de l'expression protéique. Il s'agit, au travers de la comparaison des cartographies protéiques, d'établir une analyse différentielle et sans *a priori* des modifications survenant lors d'un certain processus physiologique ou pathologique.

L'analyse d'un protéome nécessite donc un ensemble de compétences et d'instruments regroupés sous le terme de "plate-forme protéomique". Elle résulte de la convergence de différentes disciplines: la biochimie, la microchimie des protéines, la bioinformatique. Les méthodologies mises en oeuvre peuvent schématiquement être séparées en deux grands groupes (Hochstrasser *et al.*, 1998). Dans le premier peuvent être regroupées les méthodes expérimentales réalisées dans le laboratoire "réel" ("wet laboratory"), conduisant à l'utilisation de l'électrophorèse bidimensionnelle en gel de polyacrylamide (2D-PAGE), à la caractérisation des certains polypeptides par des méthodes classiques de la chimie des protéines (immunodétection après transfert, microséquençage de spots excisés du gel) ou de plus en plus de méthodes spectrométriques, comme la spectrométrie de masse. Un second groupe de méthodes, mises en oeuvre dans le laboratoire « virtuel » (« dry laboratory »), fait appel à la bioinformatique à l'aide de logiciels spécialisés utilisés au cours des différentes étapes de l'analyse: validation et interprétation des gels, utilisation des masses déterminées en spectromètre pour la consultation des banques de séquences protéiques ou nucléotidiques afin d'identifier des protéines candidates et construire des banques de données.

## PREPARATION DES ECHANTILLONS : UNE ETAPE DETERMINANTE DANS L'ETUDE DES PROTEOMES

Le but de la préparation des échantillons est de permettre une séparation efficace du plus grand nombre de protéines. Cette étape est déterminante quant aux résultats de la 2D-PAGE. Le soin qui lui est apporté conditionne notamment le nombre d'artefacts et la reproductibilité de la séparation. Son importance est telle qu'elle fait l'objet de travaux de recherche (Reymond *et al.*, 1997) visant à optimiser la préparation des échantillons en vue d'applications particulières de la protéomique, par exemple l'étude des cellules cancéreuses.

**Préparation des échantillons pour SDS – PAGE.** Du fait de l'efficacité du SDS, la préparation des échantillons biologiques en vue d'une électrophorese SDS est particulièrement simple. Il suffit de chauffer l'échantillon biologique dans un milieu tamponné en présence de SDS et d'un agent réducteur qui casse les ponts disulfures et permet une séparation parfaite des polypeptides. L'électrophorèse de zone se déroulant dans des milieux tamponnés de force ionique moyenne (0,1M), l'influence des sels naturellement présents dans l'échantillon biologique est le plus souvent négligeable. Le chauffage de l'échantillon en présence du SDS est un moyen d'accélérer la dénaturation des protéines et de faciliter la liaison uniforme du SDS. Cette dénaturation accélérée présente l'avantage d'inactiver rapidement les activités lytiques de l'échantillon comme les protéases, les phosphatases et les glycosidases. Les polypeptides ainsi préparés sont donc représentatifs de la réalité cellulaire.

Les seuls composants cellulaires capables d'interférer avec ce type de préparation sont les lipides et les acides nucléiques. Les lipides interfèrent en entrant en compétition avec les protéines pour la liaison du SDS, formant des micelles mixtes lipides-SDS. Quant aux acides nucléiques, ils posent problème par la viscosité importante qu'ils induisent et par leur propension à entrer dans les gels de polyacrylamide et à boucher leurs pores. Dans le cas d'échantillons riches en acides nucléiques, comme les préparations de noyaux, la dénaturation par le SDS empêche toute élimination ultérieure des acides nucléiques. Une solution plus radicale, peu utilisée, consiste à renverser la polarité de l'analyse et à utiliser un détergent cationique comme le chlorure de benzalkonium qui a des propriétés dénaturantes proches de celles du SDS, et l'avantage secondaire de précipiter les acides nucléiques. L'inconvénient majeur de ce type de méthode est la nécessité de réaliser l'électrophorèse à pH acide, ce qui impose des systèmes de polymérisation différents et souvent difficiles à contrôler.

Il importe de souligner la robustesse et la grande polyvalence de cette approche. Ses deux seuls points faibles, mineurs, sont parfois des erreurs dans l'estimation des masses moléculaires et les modifications induits par le gel lui-même. L'utilisation des systèmes de polymérisation plus performants pourrait diminuer ces artefacts (Rabilloud *et al.*, 1996). En revanche, comme toutes les approches basées sur un critère unique de séparation, cette technique manque de résolution dès que la complexité de l'échantillon augmente. Dans le cas d'un échantillon complexe, chaque bande de protéines se compose d'un mélange de plusieurs protéines. En outre, comme le critère de séparation est la masse moléculaire, avec une résolution modérée, les mutants ponctuels et les variantes posttraductionnelles d'une même protéine co-migrent quasiment systématiquement, ce qui complique l'analyse des isoformes protéiques.

**Préparation des échantillons pour l'électrophorèse bidimensionnelle (2D-PAGE).** Les méthodes efficaces et reproductibles pour la préparation des échantillons sont la clé pour la réussite de la 2D-PAGE (Molloy, 2000). En développant une stratégie de préparation de l'échantillon, il est important d'avoir une idée claire de ce qu'on cherche. Le but est-il de chercher autant de protéines possibles, ou seulement un sous-ensemble des protéines d'intérêt potentiel? Les étapes additionnelles de la préparation de l'échantillon peuvent améliorer la qualité du résultat final, mais peuvent avoir comme conséquence la perte sélective de certaines protéines. Le choix entre la qualité améliorée de l'échantillon et la représentation complète des protéines doit donc être bien considérée.

La préparation des échantillons consiste à extraire les protéines, puis à les traiter de manière à favoriser leur séparation au cours de l'isoélectrofocalisation (IEF). Elle doit donc permettre de solubiliser le maximum de ces molécules et empêcher leur agrégation, afin de faciliter leur pénétration dans le gel, tout en évitant leur dégradation. En pratique, les échantillons subissent, après lavage, un traitement chimique (utilisation d'un tampon de lyse) ou physique (French press, ultrasonication). Il est impératif que ces traitements soient compatibles avec l'étape de séparation ultérieure et que l'intégrité des protéines soit maintenue. Cette contrainte impose de choisir prudemment les réactifs de solubilisation. On utilise préférentiellement des détergents zwitterioniques dont la charge globale est neutre. Ces derniers n'apportent aucune charge nette aux protéines, et se révèlent totalement compatibles avec l'IEF. Quel que soit le soin apporté à la séparation, il n'est pas rare qu'une quantité plus ou moins importante de protéines difficiles à solubiliser (hydrophobicité importante, liaison à la membrane plasmique) soit perdue.

La quantité globale de protéines contenue dans l'échantillon doit elle aussi être adaptée aux étapes de séparation et d'analyse utilisées en aval. Une quantité insuffisante de protéines peut rendre la détection difficile, et leur surabondance entache la résolution de l'électrophorese.

La protection contre la protéolyse est assurée en solubilisant l'échantillon directement dans des dénaturants forts tels que l'urée 8 M, l'acide trichloracétique 10%, ou SDS 2% ainsi que la préparation de l'échantillon à une température aussi bas possible. On peut utiliser aussi des inhibiteurs de protéases ou des mélanges d'inhibiteurs.

La préparation de l'échantillon en vue de l'IEF est donc beaucoup plus complexe et moins performante que la préparation de l'échantillon en vue d'une électrophores SDS. En effet, l'IEF impose deux contraintes: de respecter la charge de la protéine et de travailler à basse force ionique. Si la première contrainte est évidente du fait du principe de séparation appliqué, et exclut tout détergent chargé (comme le SDS) qui, par sa liaison aux protéines, modifierait leur charge, la deuxième contrainte est moins évidente. Dans le cas d'une électrophorèse SDS, la protéine à séparer porte une charge électrique importante et peut donc être séparée avec des champs relativement faibles (~10V/cm). En revanche, dans le cas d'une IEF, la charge électrique des protéines est faible au départ et décroît d'autant plus qu'elles se rapprochent de leurs points isoélectriques. En conséquence, des champs nettement plus forts doivent être utilisés (100-300 V/cm), et pendant des temps assez longs (>10 heures). A ces valeurs de champs, la présence de sels induit un effet joule énorme, ce qui impose donc la préparation de l'échantillon dans des milieux de faible force ionique.

Afin de dénaturer, réduire et solubiliser au mieux les différentes protéines rencontrées dans un échantillon biologique complexe, ces milieux de solubilisation contiennent un agent réducteur des ponts disulfure, un chaotrope neutre (urée ou mélange urée-thiouree) à forte concentration, un surfactant électriquement neutre, et un système tampon à faible concentration, qui doivent répondre chimiquement et électriquement aux exigences de l'IEF (Rabilloud, 2000). Ce système tampon contient souvent des ampholytes porteurs qui présentent l'avantage d'augmenter la solubilité des protéines et de lisser la conductivité des gradients de pH immobilisés. Il peut aussi contenir des agents précipitants les acides nucléiques, comme la spermine (Rabilloud, 1994).

Si les détergents ioniques conduisent à une solubilisation optimale des protéines en empêchant toute réaggrégation du fait de la répulsion électrostatique qu'ils induisent entre les molécules, ce phénomène n'existe pas pour les détergents non-ioniques, qui sont donc nettement moins efficaces. Les contraintes théoriques induites par les détergents neutres laissent penser qu'il sera impossible de réaliser une solubilisation complète et une analyse par IEF des protéines dans l'absence des tensioactifs chargés.

**Enrichissement en protéines d'intérêt.** Seule une partie des protéines cellulaires sont suffisamment exprimées pour pouvoir être mises en évidence lors d'une analyse protéomique classique. Cette limite peut être attribuée aux différences de concentrations protéiques entre les différents compartiments de la cellule et à la présence des protéines de forte concentration dont les spots de taille importante masquent ceux des protéines minoritaires. Ceci pose un problème lorsqu'il s'agit de mettre en évidence des modifications portant sur les voies de signalisation. Explorer cette partie non visible de l'iceberg protéomique constitue une nécessité parce que ces protéines assurent des fonctions extrêmement importantes en terme de régulation. Un tri sélectif des molécules protéiques permet d'éliminer les protéines les plus abondantes qui en masquent d'autres, tout en sélectionnant certaines familles de protéines présentant un intérêt particulier.

Le tri par différences de solubilité est l'une des méthodes et peut s'effectuer à différents niveaux. On peut choisir de ne pas s'intéresser à la cellule entière, mais seulement à l'un de ses compartiments ou organites. On peut également utiliser le fait que chaque protéine, en fonction de sa composition chimique, ait une solubilité particulière. Ainsi, en extrayant les protéines sans utiliser du détergent, des protéines très solubles, principalement cytoplasmiques, seront privilégiées. L'utilisation des solutions ayant un pouvoir d'extraction croissant peut ensuite permettre la solubilisation des protéines plus hydrophobes. Afin d'enrichir les extraits protéiques en molécules d'intérêt, il sera alors nécessaire d'intercaler une étape de fractionnement entre l'extraction et la séparation par 2D-PAGE. Cette étape utilisera généralement des approches dérivées des méthodes classiques de chromatographie et d'électrophorèse.

En théorie, toutes les méthodes utilisées pour la chromatographie des protéines peuvent être transposées pour constituer une étape de préfractionnement lors d'une analyse protéomique. Les limites à leur utilisation se situent essentiellement au niveau de l'échelle (la protéomique utilisant souvent de faibles volumes d'échantillons biologiques précieux) et de la salinité des tampons utilisés lors de la séparation. Le choix d'une méthode ou d'une autre sera en fonction du but recherché. On peut se contenter de fractionner l'extrait protéique selon un critère physico-chimique (charge, hydrophophilie, etc.), chaque fraction étant ensuite analysée individuellement. Une séparation des protéines à partir d'un extrait brut peut ainsi être réalisée par chromatographie en phase inverse, permettant ainsi l'obtention des différentes fractions analysables par 2D-PAGE après l'évaporation du solvant organique utilisé pour l'éluion. La chromatographie d'échange d'ions est moins adaptée, à cause de l'utilisation des tampons de forte salinité ;

Un autre approche est de sélectionner une famille de protéines d'intérêt, c'est-à-dire un sous-protéome : un « phospho-protéome », un « glyco-protéome ». Des méthodes de micro-purifications sur des billes avec des anticorps anti-phosphotyrosine permettent d'étudier les modifications du phospho-protéome lors de la transduction intracellulaire des signaux biologiques (Imam-Sghiouar *et al.*, 2002). Des approches voisines peuvent être utilisées pour l'étude spécifique du glyco-protéome en utilisant un pré-fractionnement sur des lectines immobilisées. D'autres types de micro-chromatographies d'affinité, par exemple sur héparine immobilisée permettent d'enrichir des extraits bactériens dans certaines fractions protéiques.

## DU GEL AU PROTEOME

L'analyse protéomique se décompose en deux sousdisciplines, la **protéomique d'interaction** où les conclusions dérivent de la mise en évidence des interactions physiques entre les protéines étudiées et la **protéomique d'expression** où les conclusions sont tirées des variations dans les quantités des protéines exprimées au cours du phénomène biologique étudié. Ces approches protéomiques supposent donc de pouvoir identifier les protéines, si possible, de façon quantitative. Du fait de la complexité des échantillons biologiques il est presque toujours nécessaire de séparer les analytes biologiques avant l'identification.

Deux méthodes électrophorétiques sont utilisées dans l'analyse protéomique: **l'électrophorèse de zone** et la **focalisation isoélectrique**, toutes les deux réalisées en gel de polyacrylamide. Du fait qu'elles séparent les protéines selon des critères différents, ces deux sont souvent couplées pour réaliser une séparation bidimensionnelle à haute résolution.

**Electrophorèse SDS.** L'analyse protéomique utilise une seule variante d'électrophorèse de zone, l'électrophorèse discontinue en présence de SDS (sodium dodecylsulphate). Les deux paramètres critiques sont l'utilisation du SDS comme agent dénaturant des protéines et la réalisation d'une séparation basée uniquement sur leur masse moléculaire. Cette séparation suppose un effet de tamisage important et, de ce fait, des gels d'acrylamide relativement concentrés (7 à 15 % de solides) sont utilisés. Elle suppose, en plus, de s'affranchir de l'influence de la charge électrique de la protéine. C'est une des fonctions du SDS. Ce tensioactif anionique présente la particularité de se lier aux protéines indépendamment de leur séquence, à la concentration de 1,4 g de SDS/ g de protéine. Cette concentration aboutit à la transformation des protéines en édifices moléculaires composés d'une chaîne polypeptidique unique portant une charge négative importante où l'influence de la charge initiale de la protéine soit négligeable, à l'exception des certaines protéines très basiques comme les histones. En conséquence, chaque protéine est séparée en ses différentes chaînes polypeptidiques.

De plus, la charge négative importante portée par ces polypeptides dénaturés par le SDS conduit à une répulsion électrostatique entre les polypeptides, ce qui abolit toute interaction secondaire. Enfin, cette forte charge négative permet de rendre soluble dans l'eau n'importe quel polypeptide, y compris les plus hydrophobes comme les protéines membranaires.

**Electrophorèse bidimensionnelle.** Les travaux de Patrick O'Farrel, à l'Université du Colorado (Boulder, Etats-Unis) conduisent à la mise au point d'une technique qui, aujourd'hui encore est au cœur de toutes les études du protéome : l'électrophorèse bidimensionnelle en gel de polyacrylamide ou 2D-PAGE. Elle permet de séparer, en deux temps, les protéines d'un mélange complexe, par exemple un extrait cellulaire, en fonction de deux propriétés distinctes de ces molécules : leur charge et leur masse moléculaire.

Dans un premier temps, les protéines sont soumises à une électrophorèse dans un gel présentant un gradient de pH continu. Au cours de cette étape, dénommé isoélectrofocalisation (IEF), les protéines soumises à un champ électrique constant migrent dans un gel, jusqu'à la position où le pH du milieu est égal à une valeur, dénommé point isoélectrique ou pI, pour laquelle la charge globale de la protéine est nulle.

Ensuite, une deuxième séparation orientée perpendiculairement à l'IEF – d'où la notion de « deux dimensions » – est réalisée dans un gel réticulé constitué de polyacrylamide, en présence d'un agent chimique dénaturant, le SDS. La réticulation du gel lui confère la propriété de tamiser les molécules, et le SDS interagit avec les protéines, les délie et donne à chacune le même rapport charge/masse. Les molécules sont séparées ainsi en fonction de leur masse relative. La combinaison des deux techniques permet de cumuler le pouvoir résolutif de chacune d'entre elles et confère à la 2D-PAGE une capacité encore inégalée à séparer les protéines d'un mélange complexe.

L'électrophorèse bidimensionnelle reste une technique dont la capacité à effectivement séparer les protéines, c'est-à-dire la résolution, laisse encore à désirer. Ainsi, le plus souvent les taches protéiques ne correspondent pas à une seule protéine. Une solution consiste à segmenter le gradient de pH (par exemple : 3,5 - 4,5; 4 - 5; 5,5 - 6,7). Ce rétrécissement de la largeur du gradient pour des gels dont la taille reste identique (18 cm) entraîne une très nette amélioration de la résolution; c'est le concept des gels contigus. La séparation est réalisée sur plusieurs fenêtres de pH contiguës, et les profils obtenus sont ensuite assemblés sur ordinateur pour reconstruire la carte 2D de l'ensemble des protéines de l'organisme étudié. Une autre approche permet d'augmenter la résolution tout en réduisant les temps de migration (8 heures, contre 18 à 24 heures pour une séparation classique) et en augmentant la tension à 8000 V, contre 3500 pour les systèmes classiques.

Si l'électrophorèse bidimensionnelle est encore aujourd'hui irremplaçable, de nouvelles méthodes de séparation pourraient s'avérer des outils intéressants. Le système qui semble le plus avancé repose sur un appareil d'électrophorèse capillaire miniaturisé, une puce à électrophorèse, directement connecté à un spectromètre de masse. Il est capable à réaliser un nombre important d'études dans un minimum de temps : 96 capillaires sont assemblés en parallèle, chacun étant lié à un spectromètre de masse.

Une technique différente est fondée sur l'utilisation d'une puce à protéine ou ProteinCmp et vise à séparer et analyser les protéines sur un seul système dont l'élément central est la puce à protéine. Cette dernière est une surface miniaturisée fonctionnalisée au moyen de différents ligands: anticorps, récepteurs, acides nucléiques. Le dépôt d'un échantillon sur celle-ci permet de séparer les protéines spécifiquement reconnues par les ligands. Après lavage, ces

protéines capturées et analysées directement à la surface de la puce par une méthode de spectrométrie de masse dénommée SELDI (*surface-enhanced laser desorption/ionization*). Ce système présente l'avantage de permettre à lui seul la réalisation d'études variées, aussi structurales (simple détermination de masse moléculaire, séquençage) que fonctionnelles (interactions protéines-protéines, protéines-ADN).

## LA REVELATION DES ELECTROPHOREGRAMES

Une fois séparées, les protéines doivent être mises en évidence, afin d'obtenir la carte à deux dimensions du protéome de l'organisme ou du type cellulaire étudié. Les profils ainsi obtenus sont ensuite digitalisés, et l'outil informatique permet leur comparaison et leur stockage dans des banques de données spécifiques.

**Détection des taches protéiques.** Il existe aujourd'hui de nombreuses méthodes pour la détection des protéines séparées (Wirth *et al.*, 1995). Sur le gel, les protéines peuvent être révélées par des techniques de coloration utilisant le bleu de Coomassie, l'argent et les réactifs fluorescents, ou par autoradiographie et fluographie. Lorsque les protéines ont été transférées sur une membrane de polyvinylidène (PVDF), s'ajoutent d'autres techniques, telles que la coloration au rouge Ponceau ou à l'or colloïdal. Le choix de la méthode de révélation dépend de la résolution recherchée et du type d'analyse qui sera effectué sur les protéines après leur séparation. Certaines protéines qui sont peu, ou pas du tout révélées, par une méthode peuvent l'être par une autre. Les colorations au bleu de Coomassie et à l'argent une bonne corrélation entre la quantité de protéines retrouvées au niveau des taches protéiques et l'intensité obtenue après coloration (Giometti *et al.*, 1991). La haute sensibilité de la coloration à l'argent permet la révélation du maximum de protéines et s'avère très intéressante pour la constitution d'images de référence (ou images maîtres). La méthode avec bleu de Coomassie est moins sensible, mais peut s'appliquer aux membranes de PVDF destinées au séquençage, à l'analyse de la composition en acides aminés et à la spectrométrie de masse. Dans ce cas, les taches de protéines doivent être décolorées avant l'analyse, ce qui ralentit globalement l'étude et entraîne la perte d'une certaine quantité de protéines. L'analyse par SM d'un gel coloré superficiellement à l'argent implique une décoloration longue et une digestion enzymatique dans le gel. Cette opération est plus fastidieuse que le traitement d'une membrane, mais présente l'avantage d'éviter toute perte de matériel consécutive au transfert sur membrane.

**Analyse des gels.** La première étape de l'analyse est la digitalisation des gels, c'est-à-dire la transformation de l'image expérimentale en une information numérique utilisable par l'ordinateur. Plusieurs types d'appareils permettent de réaliser cette acquisition d'image: cameras CCD, densitomètres lasers, phospho- ou fluoro-imageurs et scanners. Ces outils découpent l'image en pixels et déterminent la différence de ton entre chacun d'entre eux. Ils reproduisent l'intensité, le périmètre, l'aire et l'orientation des taches protéiques.

Divers logiciels (Tycho, Gellab Quest, Bio-image, PDQuest, Phoretix 2D Melanie I) permettent ensuite de traiter les images obtenues afin d'éliminer le bruit de fond, les artefacts de migration et de comparer les taches résolues. La détection des protéines présentes sur un gel n'a en elle-même que peu d'intérêt. Les informations pertinentes sont tirées de la comparaison des taches obtenues sur différents gels. Toute la difficulté de l'analyse repose, donc, sur la capacité de comparer, d'un gel à un autre, les taches de même origine.

L'analyse d'images à l'aide de logiciels adaptés répond donc à la nécessité de détecter, puis de quantifier l'ensemble des polypeptides séparés afin d'identifier les différences par rapport à une modification du protéome au cours du processus biologique étudié.

**Approche comparatifs multigels.** Pour une analyse différentielle, le nombre de gels de chaque série doit être suffisant pour qu'une analyse statistique puisse permettre de définir de façon objective les variations au cours du processus biologique étudié. Deux gels de la même série ne sont pas parfaitement identiques. Les artefacts les plus courants portent sur les différences de migration des polypeptides et sur l'intensité des colorations (Joubert-Caron *et al.*, 1999a).

Une analyse différentielle fiable nécessite d'établir une comparaison entre les séries d'au moins trois à quatre gels. Des outils statistiques comme l'analyse heuristique ou l'analyse par correspondance (Pun *et al.*, 1988) permettent objectivement de déterminer la dispersion entre les gels des différentes séries. De multiples opérations vont conduire de l'image brute numérisée à la mise en évidence des variants. La digitalisation des images produit un bruit de fond qui doit être éliminé afin de ne garder que des informations pertinentes. De plus, en fonction de la quantité de protéines déposées les protéines majoritairement exprimées vont avoir la tendance de précipiter et apparaissent sur le gel sous forme de traînées horizontales ou verticales.

La quantification de l'expression se réalise le plus souvent en utilisant une valeur normalisée (% de volume) qui correspond au rapport entre le volume d'un spot considéré et le volume de tous les spots détectés sur l'image x 100. Afin de comparer l'expression protéique, les gels doivent être alignés, comparés deux à deux et les spots en commun sont appariés. Une paire de spots représente le même polypeptide présent sur deux gels et un groupe de

spots représente le même spot retrouvé sur un ensemble N de gels. Des représentations graphiques sous forme d'histogrammes permettent de visualiser les différences.

**Analyse différentielle sur un seul gel par la technologie DIGE (differential in gel-electrophoresis).** Le principe repose sur le marquage covalent à l'aide de cyanines fluorescentes (Cy3 et Cy5) des protéines contenues dans deux extraits à analyser (Unlu *et al.*, 1997). Les deux extraits ainsi marqués sont mélangés à quantité égale et les protéines sont séparées ensemble sur le même gel 2D. Le rapport protéine/colorant doit être assez faible (< 5 %) afin que les protéines visualisées ne contiennent qu'une seule molécule fluorescente. Il est conseillé de réaliser les expériences en faisant des réactions croisées des échantillons avec des sondes afin de standardiser la quantification. L'analyse d'images d'un gel par DIGE est simplifiée du fait que les deux échantillons ont migré sur le même gel. Les images acquises aux deux longueurs d'onde sont superposées et comparées quantitativement. Les profils de quantification sont fondés sur les intensités de fluorescence. L'étude différentielle aboutit à la mise en évidence des spots correspondant aux polypeptides présents dans les deux échantillons à des degrés différents d'expression ainsi que seulement dans l'un des deux. Bien que relativement coûteuse la DIGE est une méthode séduisante, qui permet théoriquement de diminuer le nombre de gels 2D nécessaires à une analyse différentielle fiable.

## IDENTIFICATION DES PROTEINES PAR SPECTROMETRIE DE MASSE

La spectrométrie de masse (SM) permet de déterminer avec une précision et une sensibilité extrêmes un paramètre intrinsèque des molécules: leur masse. Vu ses avantages (sensibilité, rapidité, aptitude à l'automatisation, relative facilité d'utilisation, faible coût des analyses – notamment par rapport au séquençage d'Edman, qui consomme beaucoup de réactifs – cette technique constitue la méthode de choix pour l'identification des protéines, la caractérisation de leurs éventuelles modifications post-traductionnelles et leur séquençage.

Deux catégories d'appareils de spectrométrie de masse sont utilisés pour cette identification : les spectromètres équipés d'une source de type Maldi et d'un analyseur à temps de vol ou Maldi-TOF (*matrix assisted laser desorption ionization-time of flight*), et les spectromètres de masse en tandem (SM/SM, couplage de 2 analyseurs au moins) associés à une source de type électrospray ou dérivée (nanoélectrospray). Ces machines ne sont pas concurrentes, mais complémentaires. Les spectromètres Maldi-TOF présentent un débit relativement important jusqu'à plusieurs dizaines de protéines par jour pour les systèmes automatisés, et sont pratiques pour effectuer une première analyse. Celle-ci peut à elle seule permettre une identification de la protéine analysée, à condition qu'elle soit déjà répertoriée dans une base de données. Si ce n'est pas le cas, la SM/SM prend le relais à un rythme d'identifier les protéines moins soutenu (3 à 4 protéines par jour en moyenne), mais elle permet d'obtenir une analyse plus fine qui peut aller jusqu'à la caractérisation d'éventuelles modifications post-traductionnelles.

Les différentes approches reposent toutes sur le même principe. Premièrement, une acquisition de données expérimentales spécifiques de la protéine : sa masse et/ou la masse des peptides qui résultent de son hydrolyse enzymatique ou chimique, dans des conditions définies et la séquence partielle de ces peptides. Ensuite, ces données éventuellement enrichies en d'autres paramètres caractéristiques de la protéine recherchée (son pI) sont traitées par des programmes bioinformatiques. Ces logiciels criblent les banques de données (protéines, EST) et permettent de déterminer si une molécule déjà répertoriée correspond aux données expérimentales. La masse moléculaire ne constitue pas en elle-même un indice suffisant. En pratique, l'identification d'une protéine est principalement fondée sur l'analyse de l'ensemble de peptides obtenus à partir de l'hydrolyse. C'est le cas de l'empreinte de masse peptidique, ou *peptide mass finger printing*.

Concrètement une molécule étudiée en spectrométrie de masse suit tout d'abord, au niveau d'un organe de l'appareil dénommé la source, une ionisation qui la «transforme» en ion moléculaire. Ce dernier représente un ensemble d'ions primaires caractérisés par les rapports de leur masse sur leur charge ( $m/z$ ). Ces ions sont par la suite séparés en fonction de ce rapport à l'intérieur d'un analyseur. Ensuite, un système de détection permet de les quantifier et d'évaluer leur masse. La mesure des rapports  $m/z$  fournit à l'expérimentateur la masse moléculaire des composés étudiés et constitue un indice de leur structure.

Ce schéma très général de fonctionnement des spectromètres de masse est quelque peu modifié pour les appareils de spectrométrie en tandem (SM/SM). Les molécules subissent, dans ce cas, deux analyses de masses entrecoupées d'une étape d'activation dans une cellule de collision. Un ion précurseur est sélectionné en fonction de sa masse dans un premier analyseur, il est ensuite activé/cassé dans la cellule de collision, puis ses produits de fragmentation sont évalués par un deuxième analyseur. Cette approche est très efficace pour l'étude structurale des molécules.

Outre ses qualités analytiques intrinsèques, la SM, surtout Maldi-TOF, présente l'avantage de pouvoir être automatisée, l'automatisation constituant un facteur clef de l'évolution de l'étude du protéome. L'Australian Proteome Analysis Facility a mis au point un système dans lequel deux automates réalisent les différentes opérations nécessaires à l'analyse : le prélèvement des taches protéiques, l'hydrolyse des protéines et le chargement des matrices destinées à la SM-Maldi-TOF (Quadroni *et al.*, 1999).

Pour l'électrospray SM/SM, le mode d'ionisation, la création d'un brouillard électriquement chargé à partir d'un flux de liquide, permet de coupler le spectromètre à différentes techniques séparatives. Ainsi, grâce aux débits très faibles utilisés par les sources électrospray évoluées comme le nanoélectrospray, il est possible d'associer directement la SM/SM à la nano-/microchromatographie liquide de haute performance, ou à l'électrophorèse capillaire. Cette dernière association offre l'avantage d'une sensibilité inférieure à 10 femtomoles et la capacité de produire des spectres de masse d'une qualité exceptionnelle. Aujourd'hui, les systèmes associant, d'une part, la préparation automatisée de l'échantillon et d'autre part, un bloc séparation-analyse directement connecté au système informatique de criblage de banque de données, constituent les outils d'analyse du protéome les plus évolués.

Si la SM représente clairement l'avenir de la science protéomique, elle a l'inconvénient de ne fournir que des données brutes: des chiffres. Le véritable défi de l'analyse des protéines par SM repose donc sur l'élaboration de programmes bioinformatiques encore plus efficaces pour traiter les données expérimentales. L'approche qui consiste à découper une protéine en fragments peptidiques pour la caractériser est finalement très proche du principe fondamental de la spectrométrie de masse : casser des molécules pour les étudier. En pratique, l'analyse d'un mélange de peptides peut être réalisée selon deux stratégies : la détermination de la masse des fragments peptides et/ou leur séquençage partiel.

La première repose sur le fait que l'ensemble constitué par la valeur des masses des peptides obtenus avec une protéase donnée est unique (Blackstock *et al.*, 1999). Il est un indicateur permettant une identification spécifique. Une fois cette «empreinte de masse» déterminée par SM MalDI-TOF, sa comparaison avec les empreintes de masse théoriques d'une banque de données de protéines permet de «pêcher» la protéine recherchée. Cette méthode souffre néanmoins d'un inconvénient majeur : une protéine non répertoriée ou pour laquelle l'information reste partielle ne peut être identifiée.

Dans ce cas, il est possible de faire appel à deux méthodes qui utilisent la nanospray-SM/SM. On part de l'idée que la séquence des peptides issus de la fragmentation d'une protéine peut aussi être considérée comme une signature spécifique, un *peptide tag* (Reymond *et al.*, 1997). Les spectres sont interprétés «manuellement» de façon à déterminer la séquence d'un ou plusieurs petits peptides (de 2 à 4 acides aminés). Ensuite, des banques de données de protéines ou d'étiquettes de séquences de gènes exprimées (EST) sont criblées avec cette information, éventuellement enrichie en d'autres indices, par exemple une empreinte de masse. L'autre approche ne demande pas l'interprétation des spectres expérimentaux, mais ceux-ci sont comparés directement par l'algorithme Sequest à des bases de données contenant des spectres théoriques (Rabilloud *et al.*, 1997).

Toutes ces méthodes reposent sur le criblage de bases de données. Leur succès n'est pas donc uniquement du aux performances des spectromètres de masse, mais tient aussi du développement de logiciels bioinformatiques.

L'identification des protéines à partir de mélanges biologiques complexes sans séparation devrait permettre une analyse rapide des complexes multiprotéiques, des interactions protéine-protéine ainsi qu'une vue générale des protéines et des complexes protéiques situés dans des différents compartiments subcellulaires et tout ça avec un traitement limité de l'échantillon. Afin d'identifier les peptides on utilise un système totalement automatisé constitué du couplage microcolonne de chromatographie liquide – spectrométrie de masse en tandem, et le criblage de bases de données par Sequest. Par cette approche, les mélanges de protéines subissent une digestion enzymatique sans séparation avant d'être analysés.

A cause de sa similarité avec la stratégie appliquée au séquençage génomique, cette méthode est dénommée «identification shotgun». Elle combine reproductibilité et sensibilité et se prête à l'analyse d'échantillons dont les concentrations et les masses moléculaires des composés varient très largement. On peut caractériser des protéines enrichies par immunoaffinité, des interactions avec des protéines de fusion, des interactions avec des complexes macromoléculaires.

## APPLICATIONS DE L'ANALYSE PROTEOMIQUE

L'étude du complément protéique total d'un génome offre aux biologistes les clés des portes que le séquençage de l'ADN ne peut pas les ouvrir. En effet, la génomique ne permet pas de saisir toute la complexité du fonctionnement cellulaire. Les données qui en découlent ne suffisent pas à déterminer la raison pour laquelle un gène particulier est exprimé à un instant donné, la concentration relative des produits d'expression, l'importance des modifications post-traductionnelles, les effets de la surexpression ou de la sous-expression d'un gène, la présence de gènes de petite taille, ou encore les phénotypes dus aux phénomènes multigéniques (Humphery – Smith *et al.*, 1997). Elucider chacun de ces points requiert une analyse des produits des gènes, ARN messagers (ARNm) ou protéines. Toutefois, un ARNm n'a qu'une seule

fonction : celle de porter temporairement une information. De plus, son expression ne conduit pas directement à l'expression des molécules dont dépend la vie cellulaire.

Ainsi, seule une étude ciblée directement sur les protéines peut aller au-delà des limites de la génomique et vise à déterminer les niveaux et le mode d'expression des protéines, leurs éventuelles modifications post-traductionnelles, les relations qui se nouent entre elles. La protéomique trouve ainsi deux champs d'application (Blackstock *et al.*, 1999). Premièrement, la connaissance des niveaux d'expression des protéines et l'impact sur ceux-ci des événements perturbateurs tels que la maladie, un traitement médicamenteux etc. Deuxièmement, l'exploration et la cartographie de la machine cellulaire, c'est-à-dire l'inventaire des protéines, la détermination de leur localisation au niveau des organites et leurs interactions.

Technique d'analyse aux applications variées, la SM permet, par sa grande précision de mesure, de détecter les modifications mineures comme l'absence d'un acide aminé ou la formation d'un pont de disulfure. En effet, quelques minutes suffisent pour déterminer une masse moléculaire à partir d'un échantillon de  $10^{-12}$  à  $10^{-15}$  mol selon la technique d'ionisation et le type d'analyseur. L'ajout ou l'élimination des groupements chimiques sur la chaîne polypeptidique permettent à la cellule d'effectuer la transduction des signaux biologiques ou peuvent constituer un signal d'adressage. Il s'agit, dans ce cas, de modifications majeures qui interviennent dans les régulations épigénétiques impliquées dans la pathologie. Ces modifications incluent des phosphorylations, des N- et O-glycosylations, des acétylations, des nitrosylations (Resing *et al.*, 1999).

La spectrométrie de masse pyrolytique et la spectrométrie MALDI-TOF permettent d'envisager des protocoles d'identification bactériologiques réalisés sur les fractions cellulaires ou, de façon plus intéressante, directement sur les cellules entières, dont la rapidité correspond parfaitement aux exigences du secteur agroalimentaire, hospitalier et militaire. La confirmation de l'intégrité chimique et la vérification des structures primaire et secondaire des protéines recombinantes à des fins thérapeutiques ne peuvent être réalisées que par la combinaison de plusieurs méthodes: la chromatographie liquide (CLHP), l'électrophorèse (sur gel ou capillaire), la focalisation isoélectrique, le microséquençage, l'analyse d'acides aminés et la SM de la protéine entière ou de ses fragments obtenus par hydrolyse enzymatique.

Pour les entreprises pharmaceutiques, la protéomique constitue un domaine potentiellement générateur de baisses de coûts, de développement de nouveaux médicaments et d'amélioration des essais cliniques. L'étude de la structure peut renseigner sur les mécanismes (fonction biochimique) et de là, sur le rôle (fonction biologique) des protéines.

La possibilité d'analyser le protéome d'une cellule (saine, pathologique, traitée par une drogue, modifiée génétiquement ou soumise à un stress) est un outil puissant au niveau d'études fondamentales dans la biochimie des protéines qu'au niveau d'études appliquées (Patton, 1999). Elle est devenue ainsi un élément important de la compréhension du dysfonctionnement cellulaire et l'élaboration des stratégies de diagnostic et de traitement. En effet, l'analyse des protéomes permet de prendre en compte l'ensemble des modifications phénotypiques, depuis les régulations transcriptionnelles jusqu'aux régulations post-traductionnelles. La possibilité d'étudier séparément les différents compartiments cellulaires en fait également une méthode d'étude de l'adressage des protéines (Lutomski *et al.*, 1997).

Aujourd'hui, un enjeu majeur est l'analyse protéomique des cellules des mammifères et, en particulier, du protéome des cellules humaines. Deux difficultés majeures sont ici à surmonter: l'absence du séquençage complet du génome et le très grand nombre des modifications post-traductionnelles. Quoi qu'il en soit, de nombreux travaux fournissent déjà des informations



précieuses sur les variations induites par la différenciation cellulaire, l'action des cytokines, les molécules utilisées en chimiothérapie, la recherche de protéines spécifiques aux différentes pathologies (Page *et al.*, 1999). Parmi les domaines où des analyses protéomiques systématiques sont en cours, on peut citer : la configuration des cartes des fluides biologiques à visée diagnostique, la recherche d'une vision globale des modifications induites par le vieillissement, la tumorigenèse, ou par l'action des agents chimiques ou biologiques. L'approche protéomique consiste alors à effectuer une analyse différentielle des protéomes d'un échantillon de tissu normal et du même tissu pathologique, après le traitement par un médicament (Joubert-Caron *et al.*, 1999a). Les études en cancérologie se prêtent particulièrement à ce type d'approche, du fait de la possibilité d'obtenir à la fois du tissu sain et pathologique. D'autres types de projets au niveau biomédical sont de trouver dans les liquides biologiques des marqueurs de diagnostic précoce pour la maladie d'Alzheimer, l'accident vasculaire cérébral et la maladie de Creutzfeld-Jacob.

L'analyse protéomique s'applique en pharmacologie à l'estimation *in vitro* de l'efficacité des nouveaux médicaments et à l'étude comparative des relations structure/activité d'une série de composés (Wang *et al.*, 1999). La caractérisation complète du protéome mitochondrial peut mener au développement de nouveaux médicaments, en particulier dans le domaine des anti-cancéreux qui viendraient alors enrichir la liste des principes actifs protéiniques, comme l'érythropoïétine et l'interféron.

Tandis que la génomique tente d'identifier l'ensemble d'éléments du mécanisme de la maladie, l'élément diagnostique peut être un élément complètement parallèle au mécanisme pathologique. Du moment qu'une molécule apparaît ou disparaît entre deux états physiologiques (qu'elle soit impliquée dans la physiopathologie ou qu'elle soit un simple produit de dégradation), elle peut en principe être détectée par la protéomique, et devenir par conséquent une molécule intéressante pour l'industrie du diagnostic.

## CONCLUSIONS

La protéomique est sûrement un domaine en devenir. Du décryptage des mécanismes fins au niveau des synapses à la détection des marqueurs de tumeurs, les champs d'application sont immenses. La recherche en protéomique est passionnante parce que, pluridisciplinaire, elle rassemble autour d'un même projet des biologistes, des chimistes, des physiciens et des bioinformaticiens. Elle devrait conduire à des développements technologiques encore plus pointus, notamment dans le domaine de la quantification des protéines (en vue de l'analyse systématique de la dynamique des protéomes) et de la miniaturisation. Des « *Lab-on-a-chip* » dédiés aux analyses protéomiques permettront la miniaturisation, l'intégration, et la parallélisation des étapes qui se trouvent en amont de l'analyse par spectrométrie de masse ce qui rend l'approche protéomique encore plus sensible et rapide.

## BIBLIOGRAPHIE

1. Anderson N.L., Anderson N.G., 1998. *Electrophoresis*, 19:1853-1861.
2. Blackstock W.P., Weir M.P., 1999. *Trends Biotechnol.*, 17/3: 121-127.
3. Giometti C.S., Gemmill M.A., Tollaksen S.L., Taylor J., 1991. *Electrophoresis*, 12/7-8:536-543.
4. Hochstrasser D.F., 1998. *Clin. Chem. Lab.Med.*, 825-836.
5. Humphery-Smith I., Cordwell S.J., Blackstock W.P., 1997. *Electrophoresis*, 18/8:1217-1242.
6. Imam-Sghiour N., Laude-Lemaire I., Labas V., Pflieger D., Le Caer J.P., Caron M., Nabias D.K., Joubert-Caron R., 2002. *Proteomics*, 2/7 : 828-838.

7. Joubert-Caron R., Feuillard J., Kohanna S., Poirier, F., Le Caer J.P., Schuhmacher M., Bornkamm G.W., Polack A., Carro M., Bladier D., Raphael M., 1999a. *Electrophoresis*, 20:1017-1026.
8. Lutomski D., Fouillit M, Bourin P, Mellottee D., Denize N., Pontet M., Bladier D., Caron M., Joubert-Caron R., 1997. *Glycobiology*, 7: 1193-1199.
9. Molloy M.P., 2000. *Anal. Biochem.*, 280: 1-10.
10. Page M.J., Amess B., Rohlf C. Stubberfield C, and Parekb R., 1999. *Drug Discovery Today*, 4: 56-62.
11. Patton W.F., 1999. *J. Chromatogr. B*, 722: 203-223.
12. Pun T., Hochstrasser D.F., Appel R.D., Funk M., Villars-Augsburger V., Pellegrini C., 1988. *Electrophoresis*, 1/1: 3-9.
13. Quadroni M., 1999, *Electrophoresis*, 20/4-5: 664-677.
14. Rabilloud T., 1994. *Cellular and Molecular Biology*, 40(1): 57-75.
15. Rabilloud T., 1996. *Electrophoresis*, 17: 813-829.
16. Rabilloud T., 2000. *Proteome Research: Two Dimensional Gel Electrophoresis and Identification Tools*, Springer, Berlin.
17. Rabilloud T., Adessi C., Giraudel A., Lunardi J., 1997. *Electrophoresis*, 18/3-4: 307-316.
18. Resing K.A., Ahn N.G., 1999, *Prog. Biophys. Molec. Biol.*, 71: 501-523.
19. Reymond M.A., Sanchez J.C., Schneider C., Rohwer P., Tortola S., Hohenberger W., Kirchner T., Hochstrasser D.F., Kockerling F., 1997. *Electrophoresis*, 18/3-4: 622-624.
20. Unlu M., Morgan M.E., Minden J.S., 1997. *Electrophoresis*, 18/11: 2071-2077.
21. Wang J. H., Hewick R.M., 1999. *Drug Discovery Today*, 4: 129-133.
22. Wilkins M. R., Williams K.X., Appel R.D., Hochstrasser D.F., 1997. *Proteomic Research: New Frontiers in Functional Genomics (Principles and Practice)*. Berlin: Springer Verlag.
23. Wirth P.J., Romano A., 1995. *J. Chromatogr. A*, 698/1-2: 123-143.

1 - Université "Al. I. Cuza" Iasi, Faculté de Biologie

2 - Université des Sciences et Technologies de Lille

\* - georgiana\_palii@yahoo.com

## LE RÔLE DES PROTÉINES BRCA1 ET BRCA2 DANS LA RÉPARATION DES ALTÉRATIONS MOLÉCULAIRES DE L'ADN

LUCIAN NEGURĂ<sup>1\*</sup>, ANCA-MIHAELA HUMĂ<sup>1</sup>,  
VLAD ARTENIE<sup>1</sup>, EUGEN CARASEVICI<sup>2</sup>

**Mots-cléf :** BRCA1, BRCA2, lésions double brin (DSB), recombinaison homologue (HR), réparation ADN.

**Résumé :** Les gènes BRCA1 et BRCA2 sont des suppresseurs tumoraux dont les phénotypes mutants prédisposent au cancer mammaire et ovarien. Ce sont des gènes nouveaux et peu d'informations concernant leurs fonctions ont été apportées par leur séquence. Cependant, de nombreuses études sur les protéines BRCA et leurs partenaires moléculaires ont montré l'implication dans une multitude de processus cellulaires fondamentaux, incluant la réponse cellulaire aux altérations de l'ADN. Nous présentons quelques unes des plus récentes connaissances dans ce domaine.

### INTRODUCTION

Révéler les fonctions normales de BRCA1 et de BRCA2 est depuis plus de dix ans un problème fascinant pour les scientifiques. Par contraste avec la compréhension génétique très claire des gènes codant pour BRCA1 et BRCA2, les fonctions cellulaires normales de ces deux protéines se sont avérées difficiles à définir. Les recherches concernant les fonctions protéiques ont mis en évidence des interactions entre les protéines BRCA et un grand nombre d'autres protéines régulatrices. Jusqu'à l'heure actuelle, un nombre limité de fonctions cellulaires est clairement attribuable aux gènes BRCA:

- 1) On a pu constater que BRCA1 et BRCA2 sont nécessaires à un développement prolifératif normal lors de l'embryogenèse précoce, étant régulées en association avec prolifération cellulaire de l'épithélium mammaire durant la puberté, la gestation et la lactation. Ce rôle paraît assez paradoxal, étant donné que dans l'épithélium mammaire et ovarien adulte, la perte de BRCA1 ou de BRCA2 conduit à la tumorigenèse par prolifération cellulaire [Rajan, 1997].
- 2) Le domaine N-terminal en doigt RING (*RING finger domain*) de BRCA1, ainsi que les régions adjacentes, facilitent le transfert de l'ubiquitine vers les protéines cellulaires destinées à la dégradation [Lorick, 1999].
- 3) BRCA1 et BRCA2 jouent un rôle fondamental BRCA dans la régulation transcriptionnelle de certains gènes impliqués dans la réparation ADN, le cycle cellulaire et l'apoptose. BRCA1 forme un complexe avec la RNA polymérase II [Scully, 1997], alors que BRCA1 et BRCA2 interagissent toutes les deux avec des régulateurs transcriptionnels [Deng 2000].
- 4) Chose la plus importante, les deux protéines maintiennent la stabilité du génome [Welch, 2000] à travers leur implication dans la recombinaison homologue (*Homologous Recombination* – HR) et dans la réparation des lésions double brin (*Double Strand Breaks* – DSB) couplée à la transcription [Chen, 1998]. Ce rôle est suggéré en égale mesure par les interactions de BRCA1 et/ou BRCA2 avec d'autres protéines connues comme impliquées dans la réparation de l'ADN, et par la déficience de réparation associée aux défauts des points de contrôle du cycle cellulaire affichée par les cellules mutantes *BRCA1<sup>-/-</sup>* et *BRCA2<sup>-/-</sup>*.

Le but de notre travail consiste à essayer une synthèse intégrative des éléments démontrant les rôles de BRCA1 et de BRCA2 dans la réparation de l'ADN, ainsi que de cibler les questions fondamentales qui restent non élucidées à ce sujet.

### LES CLEFS STRUCTURAUX DES FONCTIONS BIOLOGIQUES

Quoique les phénotypes des cancers mammaire et ovarien associés aux mutations de BRCA1 et de BRCA2 sont similaires, les deux gènes ne sont pas sensiblement apparentés par leurs séquences. Étant donné les importantes différences entre les structures primaires, une parallèle génomique entre BRCA1 et BRCA2 s'avère particulièrement frappante : les deux gènes sont très larges, s'étendant sur environ 80 kbases d'ADN génomique ; tous les deux possèdent des exons centraux extrêmement longs codant pour plus de 50% de la protéine ; et tous les deux codent pour de très grosses protéines, formées respectivement de 1863 et de 3418 acides aminés. La comparaison entre les séquences en acides aminés des deux protéines humaines avec leurs équivalentes murines indique une proportion d'identité d'environ 60% ; il faut cependant préciser que la plupart des gènes suppresseurs tumoraux sont beaucoup plus conservés entre les espèces que le sont BRCA1 et BRCA2.