

**AMBASSADE DE FRANCE AUX ETATS-UNIS
MISSION POUR LA SCIENCE ET LA TECHNOLOGIE**

**'MICROARRAYS', (PUCES), APPLIQUES
A L'ANALYSE DES PROTEINES.
DEVELOPPEMENTS TECHNOLOGIQUES
ET APPLICATIONS,
UN APERÇU AUX ETATS-UNIS.**

Février 2005

Armand RENUCCI

Attaché pour la Science et la Technologie

-
- Mission pour la Science et la Technologie, Consulat Général de France, 530 Bush Street, San Francisco, CA 94108, USA.
 - Tel : 415-397-4440
 - Fax : 415-591-4830
 - Email : attache.sdv@consulfrance-sanfrancisco.org

Sommaire

I. Problématiques, enjeux et définitions. p2

II Les technologies. p4

-1 Les agents de capture.

1-1 Agents de capture de type protéine/peptide

1-1-1 Anticorps

1-1-2 Une alternative aux Anticorps : les PAM 'Peptide Antibody Mimetic' p6

1-1-3 Utilisation de peptides dans le 'Molecular Imprinting'

1-2 Agents de capture non protéique p7

- 2 Dépôt et attachement des agents de capture.

2-1 Choix du substrat p8

2-2 Dépôt

2-3 Immobilisation de l'agent de capture.

2-3-1 Couplage non covalent

2-3-2 Couplage covalent ou quasi-covalent p9

-3 Détection et lecture des interactions. p10

3-1 Détection basée sur des modifications des partenaires ou l'utilisation d'agents secondaires marqués.

3-1-1 Marquage du partenaire

3-1-2 Marquage de l'agent capture

3-1-3 Marquage d'un agent de détection p11

3-2 Détection sans modifications des agents de capture ou des molécules analysées

3-2-1 Détection directe de l'interaction

3-2-2 Détection indirecte de l'interaction p12

3-3 Détection d'une activité fonctionnelle

III Applications p13

-1 'microarrays capture': recherche de biomarqueurs, diagnostic et recherche pharmaceutique.

1-1 Diagnostic

1-1-1 Application à des pathologies humaines complexes

1-1-2 Détection d'agents infectieux p14

1-1-3 Détection (bio)organique

1-2 Applications pharmacologiques

1-2-1 Recherche de cibles thérapeutiques

1-2-2 Validation des cibles

1-2-3 Analyse pharmacoprotéomique p15

- 2 'microarrays fonctionnels' p16

2-1 Dissection mécanistique des processus cellulaires : caractérisation de réseaux d'interactions

2-2 Recherche et caractérisation de nouvelles molécules pharmacologiques

IV Conclusion p17

Annexes : - Liste de sociétés engagées dans le développement des 'microarrays' appliqués aux protéines p19

- Figures 1 et 2 p22

L'objectif de ce rapport est de présenter les grandes lignes des approches technologiques, en cours de développement ou commercialisées, dans le domaine des 'microarrays' (puce) appliqués à l'analyse des protéines. Il s'agit d'illustrer la dynamique américaine dans le secteur académique mais surtout en terme d'activité R&D industrielle sur laquelle l'accent sera mis avec également une présence européenne notable dans ce domaine (allemande, britannique, scandinave).

Une partie des informations présentées dans ce rapport est issue d'une conférence organisée par le Cambridge Health Institute à San Diego (Californie) en janvier 2004.

Des références plus précises, relatives aux sources des technologies évoquées (en particulier pour l'académique), peuvent être demandées (coordonnées page en-tête) si l'information présentée s'avère insuffisante pour approfondir un point d'intérêt. Le rapport est disponible sous forme de fichier sur le site : <http://www.france-science.org> à la rubrique rapport d'études.

I. Problématiques, enjeux et définitions.

Les processus physiologiques et physiopathologiques, depuis le niveau cellulaire jusqu'à l'organisme complet où ils sont intégrés, reposent sur l'activité de macromolécules essentielles les protéines. La description individuelle de chaque protéine en terme d'activité et de fonctions apparaît comme une tâche considérable, en effet leur nombre est estimé entre 500000 et un million à cause en particulier des modifications post-traductionnelles qu'elles subissent (le couplage transitoire ou permanent de un ou plusieurs groupements chimiques sur la protéine : phosphate, sucre, acide gras... induit ou modifie son activité). La notion de fonction est elle-même complexe puisqu'elle dépend en partie de la position de la protéine dans les réseaux d'interactions protéine/protéine spécifiques auxquels elle participe dans les différents types cellulaires.

L'appréhension de ces processus a beaucoup progressé grâce aux outils de la biologie moléculaire (en particulier la disponibilité des séquences génomiques, génome humain et modèles animaux, et l'utilisation des 'microarrays' ADN) où l'ARN messenger qui code la protéine analyse reste un outil et un objet central d'analyse. L'approche protéomique même partielle reste limitée; la détection (électrophorèses 2D combinée à la spectrométrie de masse) et les investigations fonctionnelles directes des protéines (eg. interactions petite molécule/protéine, protéine/protéine) in vivo et in vitro reposent en général sur des manipulations lourdes et des quantités de matériel biologique relativement importantes à

mobiliser accompagnés de nombreuses limitations (protéines membranaires et du cytosquelette, avec des propriétés physico-chimiques frontières et/ou présentes en faible abondance avec les nombreux artefacts de la technologie du double hybride utilisée pour détecter les interactions entre protéines). Une analyse en parallèle à haut débit reste difficile à mettre en œuvre dans ce contexte.

Le format 'microarray' ADN qui combine miniaturisation (mise en jeu de quantités de matériels limitées) et automatisation a permis d'analyser avec succès de nombreux processus cellulaires. Des milliers de séquences ADN (spécifiques pour chaque gène) ordonnées à haute densité permettent en particulier une analyse en parallèle et à haut débit du transcriptome cellulaire en autorisant la détection semi-exhaustive des ARN messagers exprimés dans un type cellulaire. L'analyse des protéines avec un format 'microarray' a été réalisée initialement dans une optique d'immunodiagnostic (interaction antigène/anticorps). Une analyse à l'échelle du protéome, nécessaire pour répondre aux problématiques évoquées incluant de nombreuses applications en diagnostic et recherche pharmaceutique, a motivé un ensemble de développements technologiques très variés liés à la complexité des protéines et aux problèmes originaux soulevés par leur intégration sur des 'microarrays'.

- Ce rapport est centré sur le format 'microarray' planaire où les agents de capture (protéines, peptides) ordonnés sur une telle surface peuvent interagir avec un partenaire (protéine, petite molécule) et générer un signal permettant la détection de cette interaction (où d'une activité (bio)chimique liée à cette interaction), **Figure1/2**. Il est à noter que le terme 'microarray' en ce qui concerne les protéines est étendu à des systèmes présentant une échelle de taille supérieure à la taille standard des 'microarrays' ADN avec des densités éventuellement beaucoup plus faibles (quelques dizaines de spots). De même, les 'microarrays' où l'agent de capture correspond à des acides nucléiques permettant de détecter des protéines seront également considérés. D'autres formats : billes ('beads') en suspension, microcanaux etc... seront évoqués dans la conclusion au travers des sociétés qui développent ces technologies.

- Bien qu'en terme de fabrication il n'y ait pas de différence majeure on peut schématiquement classer les 'microarrays' en deux catégories en fonction de leur application : les 'microarrays capture' où l'objectif est de détecter et quantifier une protéine spécifique et des 'microarrays fonctionnels' où des paramètres plus complexes sont analysés : activité fonctionnelle (eg. enzymatique), caractéristiques des interactions (affinité etc...) avec des partenaires variés, **Figure1/2**.

II Les technologies :

-1 Les agents de capture.

Ils correspondent essentiellement à des agents protéiques, protéines et peptides, avec une exception les 'aptamers' (acides nucléiques), associés au substrat du 'microarray'. Dans la catégorie protéine, les anticorps sont actuellement les plus utilisées. Quelque soit la nature des protéines utilisées, l'ingénierie et la production mis en jeu à un stade quelconque du processus se heurtent, malgré des systèmes de production bactériens et eukaryotes performants (production de masse à haut débit) à de multiples problèmes : construction des vecteurs d'expression et purification de protéines possédant une conformation adéquate et suffisamment stables.

1-1 Agents de capture de types protéines et peptides

1-1-1 Anticorps

Ce sont les molécules de choix pour détecter une protéine, il est en effet, au moins théoriquement, possible de générer contre la plupart des protéines existantes des anticorps présentant une spécificité et une affinité compatible avec les contraintes du 'microarray'.

Les enjeux actuels sont :

- L'obtention d'anticorps avec une haute affinité, une stabilité suffisante et de manière essentielle très spécifique (indispensable pour une application de type diagnostique).
- Générer de nouvelles collections d'anticorps de qualité en haut débit et à des coûts réduits dans un contexte où beaucoup d'anticorps commerciaux sont considérés d'une qualité insuffisante pour répondre aux besoins requis pour la fabrication et l'utilisation des 'microarrays'. La lourdeur des procédures traditionnels et les coûts importants de production associés constituent donc un frein à cet objectif.

Obtention in vivo.

- Les Anticorps conventionnels polyclonaux :

-- Ils sont traditionnellement obtenus après injection d'une protéine purifiée chez un animal. La réaction immunitaire obtenue génère des anticorps qui présentent une structure et des propriétés similaires à celles des anticorps naturellement obtenus après une maladie. Plusieurs limitations existent : l'immunisation qui requière plusieurs microgrammes de protéine purifiée, la protéine est traditionnellement synthétisée à partir d'un système cellulaire (E.Coli ou cellules eucaryotes) avec un vecteur d'expression mais l'obtention et la purification en quantité adéquate d'une protéine possédant une conformation satisfaisante est difficile à

mettre en œuvre à haut débit en considérant en particulier la variété de leur propriétés physico-chimiques; l'étape de purification par affinité des anticorps, à partir du sérum, est lourde et peu reproductible enfin la qualité des anticorps obtenue apparaît très variable en terme de spécificité et d'affinité. Il est à noter que la société Milagen a entrepris de générer avec cette approche, en haut débit, une collection exhaustive d'anticorps, 'Antibiomix', reliés à une très large gamme de pathologies; cette société restée très discrète sur ses protocoles met cette collection à la disposition de société biotech. ou pharma dans le cadre de partenariats spécifiques.

-- L'immunisation génétique, plus récente, est réalisée, par injection d'un vecteur d'expression de la protéine choisie au niveau de la peau,. Elle permet d'éviter, par rapport à l'immunisation conventionnelle, d'avoir à produire et purifier l'antigène qui étant exprimé in vivo présente une conformation native correcte limitant les risques de générer des anticorps non spécifiques. Le protocole global reste néanmoins lourd à mettre en œuvre avec la nécessité d'optimiser les codons du vecteur d'expression pour améliorer l'efficacité de la traduction.

- Les Anticorps conventionnels monoclonaux.

Produits à partir d'hybridomes (fusion myélome et lymphocytes B), ils présentent l'avantage d'une très bonne spécificité et d'une affinité allant du nano au picomolaire (sensibilité suffisante pour détecter des protéines sériques eg. cytokines). Ils permettent d'éviter le bruit de fond dû à la réactivité croisée des anticorps polyclonaux mais néanmoins, présentent parfois, de manière inattendue, cet inconvénient imposant également une validation systématique en amont. Enfin, comme pour les anticorps polyclonaux, les protocoles actuels restent difficilement adaptables à une production à haut débit.

Obtention in vitro.

Cette approche est maintenant dominante, elle est basée sur l'utilisation de bibliothèques d'expression 'display libraries' d'anticorps synthétiques ou naturels. Les anticorps clonés correspondent habituellement à un fragment monovalent soit à une seule chaîne d'immunoglobuline où les domaines variables (contenant le site de reconnaissance d'un épitope) V_H et V_L sont associés par un linker ('single chain Fv', scFv) soit à un fragment Fab incluant les domaines V et C. Des bibliothèques dont la complexité atteint 100 milliards d'anticorps permettent de sélectionner des anticorps recherchés avec une affinité suffisante. La sélection est basée, soit sur la transfection des phages dans des bactéries et l'expression de l'anticorps cloné mais ce processus de sélection est limité par le nombre de clones analysables après transfection, soit avec des anticorps produits et sélectionnés par un processus totalement

in vitro : le 'ribosome/-mRNA display', après transcription de l'ARN messager à partir du clone dépourvu de codon stop, l'anticorps est directement synthétisé in vitro et reste complexé au ribosome. Cette approche est très prometteuse puisque il n'y a pas de limitations de la taille de la librairie analysable, elle permet également d'isoler des anticorps contre des protéines qu'il n'est pas possible de produire dans un système cellulaire, à des fins d'immunisation par exemple, à cause de leur toxicité en cas de surexpression. De plus elle autorise un certain degré d'automatisation nécessaire à une production d'anticorps à haut débit.

Après cette phase de sélection il est possible, en particulier par génie génétique, de produire des anticorps non conventionnels, de taille inférieure à celle d'un anticorps natif, avec des qualités supérieures en terme d'affinité, (anticorps multivalents) ou des propriétés originales (anticorps bifonctionnels). Plusieurs sociétés s'appuient sur ce type de plateforme Dyax et Morphosys, Cambridge Antibody Technology, Domantis (Grande-Bretagne), Bioinvent (Suède).

Une étape supplémentaire est réalisée par Affibody (Suède) qui a créé une collection de domaines, 3.10 puissance 9 'Affibody', qui possèdent les mêmes propriétés que les régions complémentaires des anticorps. Ces molécules de 58 acides aminés outre l'intérêt de leur petite taille présentent une haute affinité pour des protéines humaines, virales et bactériennes, elles sont obtenues à partir de la mutagenèse combinatoire sur 13 sites d'un domaine structural de la protéine A du staphylocoque (technologie utilisée par la société Gyros). Phylos, avec la même stratégie, développe les 'Trinectins' à partir de la mutagenèse dirigée sur un domaine de la fibronectine.

L'ensemble de ces outils permettent une production à haut débit d'anticorps ou de molécules similaires, ils restent néanmoins, en amont, lourd à mettre en œuvre avec la construction de bibliothèques et en aval la validation des molécules obtenues.

1-1-2 Les PAM 'Peptide Antibody Mimetic'.

Ces peptides de 20 acides aminés sont une alternative aux anticorps. Conçus avec une assistance bioinformatique, ils présentent la même affinité et la même spécificité que les anticorps conventionnels avec deux avantages, leur très petite taille et une manipulation plus aisée.

1-1-3 Utilisation de peptides dans le 'molecular imprinting'

Cette technologie est une alternative à la conception classique des 'microarrays' où les agents de capture sont ordonnés et fixés sur la surface. La polymérisation de monomères spécifiques autour d'un peptide correspondant au domaine d'intérêt de la protéine étudiée, permet

d'obtenir une matrice présentant des 'cavités' correspondantes à la séquence d'intérêt (procédé 'ProteinPrint') Aspira Biosystems. Les protéines qui présentent le motif ou la structure d'intérêt peuvent être ainsi spécifiquement capturées dans ces cavités. Cette méthode permet dans le principe de générer un agent de capture pour n'importe quelle structure ce que ne permet pas l'utilisation d'anticorps.

1-2 Agent de capture non protéique

- Un des problèmes pour les agents de nature protéique correspond à leur stabilité limitée rencontrée lors de la fabrication, de l'utilisation et du stockage du 'microarray'. L'utilisation de molécules ARN ou ADN simple brin, de type 'aptamers', présente un avantage majeur en terme de stabilité. Ils sont sélectionnés in vitro sur la base de leur affinité pour des protéines spécifiques (protocole SELEX 'systematic evolution of ligands by exponential enrichment'). Des collections d'aptamers ont été sélectionnées à partir de bibliothèques présentant une complexité de 10^{14} molécules (la synthèse est réalisée in situ sur le substrat, voir partie Technologie II 2-3-2).

- Un développement de cette technologie est réalisé par Somalogic qui permet d'augmenter son efficacité. Les photoaptamers sont des aptamers combinés au groupement 5'-bromodeoxyuridine (BrdU) photoactivable. L'illumination par des UV génère une liaison covalente avec la protéine cible fixée ('crosslinking') augmentant de manière très importante la sensibilité de détection (dans la gamme picomolaire pour des facteurs de croissance sériques). Les limitations résident d'une part dans la sensibilité aux nucléases présentes dans les extraits protéiques à tester et d'autre part une spécificité moins élevée qu'avec des agents de type anticorps.

- 2 Dépôt et attachement des agents protéiques.

La qualité des étapes de dépôt et d'attachement des agents de capture sur le substrat du 'microarray' conditionne bien sûr son efficacité. La conformation des agents protéiques, très sensible aux conditions physico-chimiques de leur environnement, doit être préservée lors de ces processus incluant le stockage du 'microarray'. Par ailleurs plusieurs paramètres doivent être optimisés incluant l'orientation (pour une meilleure interaction), la solidité de la fixation sur le support (résistance aux différents traitements, lavages etc...). A ces aspects se superpose la contrainte de la reproductibilité dans le cadre d'une production industrielle par exemple, contrôle de la quantité déposée indispensable à la quantification des échantillons testés.

Une large combinaison de choix au niveau des substrats, des techniques de dépôt et des stratégies d'immobilisation permettent de répondre à divers degrés à ces objectifs.

2-1 Choix du substrat

La nature du substrat n'est pas l'élément le plus critique vis-à-vis de ces objectifs car différents types peuvent être utilisés avec succès pour une même problématique. Le substrat le plus répandu est la lame de verre traitée avec diverses chimies permettant une fixation covalente ou non de l'agent de capture (Pierce Boston Technology); d'autres substrats sont utilisés : plaque de microtitrage en plastique, membranes de nature diverse (nitrocellulose (ProteinOne), polyvinylidene difluoride (PVDF)), matrices poreuses de natures diverses, et particules (voir conclusion). On peut citer l'utilisation potentielle future des nanotubes de carbone (l'anticorps est greffé sur le nanotube, l'élongation du tube liée à la fixation de la protéine correspondante modifie la conductivité de la structure et permet une détection directe de l'interaction).

2-2 Dépôt

- 'Spotting'. Utilisé pour la plupart des microarrays planaires, le dépôt de microgouttes à la surface ('spotting') peut se faire avec différentes techniques ('Contact printing', 'Pinprinting', 'Ink thermal jetting', 'Piezoelectric spotting'). Récemment une optimisation de l'outil spectromètre de masse, au niveau contrôle de l'ionisation afin d'éviter la détérioration des protéines et la perte d'activité qui en résulte traditionnellement, permet de l'utiliser à des fins de dépôt. Des systèmes de centrifugation (Gyros), microcapillaires (Living Chips et Biotrove) sont aussi utilisés.

- Synthèse in situ. Des collections de protéines recombinantes munie d'un 'tag' GST sont synthétisées in situ et immobilisées sur le substrat préalablement traité avec un anticorps GST (détails II 2.3.2).

2-3 Immobilisation sur le substrat

Deux grandes options de couplage : covalent / non-covalent.

2-3-1 Couplage non covalent.

L'intérêt majeur de cette option réside dans l'absence de modifications de l'agent de capture, permettant l'utilisation directe d'une protéine native.

- *Couplage basé sur l'adsorption passive de l'agent de capture sur la surface*
Méthodologiquement simple, assez reproductible conduisant à un bruit de fond faible lors de la lecture, il présente plusieurs inconvénients : absence d'un contrôle de l'orientation et de la quantité de molécules déposées, une perte de signal peut être observé pendant le test à cause de la désorption de la protéine et enfin l'adsorption peut s'accompagner d'une dénaturation

partielle qui autorise la fixation d'un partenaire mais limite la possibilité de tester d'une activité fonctionnelle. L'adsorption peut être basée sur des forces électrostatique, l'hydrophobicité de la protéine, des forces de Van der Waals, la liaison hydrogène sur des supports de verre siliconé ou traités à la poly-lysine, des membranes nitrocellulose éventuellement traitée au PVDF et des plaques de microtitration traitées au cyanosilane.

- Couplage basé sur la diffusion de l'agent de capture dans une matrice poreuse

Avec l'HydroGel (PerkinElmer), un gel de polyacrylamide ou d'agarose sur verre, le bruit de fond observé est faible avec une bonne capacité de rétention de la protéine sous sa forme native mais toujours avec un problème de contrôle quantitatif et d'orientation de l'agent de capture.

2-3-2 Couplage covalent ou quasi-covalent

L'utilisation d'une liaison forte assure bien entendu un meilleur maintien de l'agent de capture sur la puce mais apporte également avec la plupart des méthodes utilisées un contrôle de l'orientation spatial sur le support augmentant ainsi la sensibilité de la détection de l'interaction ainsi qu'une meilleure reproductibilité en terme de fabrication.

- **Fixation directe** : Des surfaces de type lame de verre traitées avec des amino-silane ou des aldéhyde-silane permettent la fixation covalente d'anticorps via les résidus lysine, quand ils possèdent un domaine carbohydate celui-ci peut être utilisé moyennant l'activation de l'anticorps. La réactivité du N-hydroxyde succimide sur les groupements amine de la protéine est aussi utilisée.

- **Fixation indirecte** : des stratégies plus élaborées évitent de fixer directement l'agent de capture sur le substrat en utilisant un agent de couplage incorporé sur la surface qui interagit avec un groupement chimique greffé sur la protéine ou l'introduction d'un 'tag' dans la protéine recombinée. Les principales inconvénients de cette approche sont les conséquences potentielles sur la structure et l'activité de la protéine et la lourdeur de la production de ces agents. Avec ces technologies il est également nécessaire de minimiser les interactions non spécifiques (limitation du bruit de fond) de l'échantillon analysé avec la surface qui peut être traitée avec la BSA ou du Polyéthylène glycol (PEG).

- Modification de la protéine avec un groupement dérivé de l'acide phenyldiboronic (système Versalinx, Prolinx) qui permet l'immobilisation de la protéine sur la lame traitée à l'acide salicylhydroxamique.

- Greffe de la streptavidine sur la protéine immobilisée avec la biotine conjuguée à une couche de polylysine sur une surface de type titanium dioxide (Zyomyx) ou tantalum pentoxide, (Zeptosens).

- Introduction d'une séquence de polyhistidine dans la protéine, l'interaction avec l'agent de capture est réalisée sur la surface traitée avec du nickel.

- Introduction d'un 'tag' GluthationSTransferase (GST) qui permet d'immobiliser l'agent de capture avec un anticorps de couplage antiGST fixé au substrat. Cette approche a été récemment validé dans un système où l'agent de capture est synthétisé in situ à partir d'un vecteur d'expression incluant le 'tag' GST'. L'immobilisation du vecteur d'expression (selon une distribution ordonnée) sur le substrat du 'microarray' est basée sur une interaction biotine/avidine permettant (avec un extrait cellulaire) une synthèse locale de l'agent qui une fois la traduction terminée est immobilisé grâce à l'interaction immédiate avec l'anticorps antiGST. Un des grands avantages réside dans l'absence de manipulation des agents de capture en tant que tels.

-3 Détection et lecture des interactions

La détection est généralement réalisée in situ (sauf 3-2-2), les deux enjeux majeurs sont d'une part la sensibilité (en particulier pour la détection de protéines peu abondantes) et d'autre part la quantification des signaux générés par l'interaction de l'agent de capture avec un partenaire.

3-1 Détection basée soit sur une modification de l'agent de capture ou de son partenaire soit sur l'utilisation d'agents secondaires

3-1-1 Marquage du partenaire

- Avec un élément radioactif : cette approche très sensible reste difficile à automatiser et apparaît plutôt limitée à des développements expérimentaux amont, en laboratoire.

- Le couplage avec un fluorophore (eg. Cy3, Cy5) est très utilisé. Il permet une lecture avec les mêmes systèmes et logiciels utilisés avec les puces à ADN et procure une sensibilité allant jusqu'à 1pg/ml dans l'analyse d'un échantillon. Néanmoins le couplage d'un fluorophore à des protéines en particulier tend à réduire leur solubilité et peut donc interférer sur la reconnaissance stérique et l'intensité du marquage, de plus la qualité du couplage varie pour chaque protéine et nécessite une étape de contrôle importante en particulier dans le cas où des échantillons dont le marquage diffère doivent être comparés. Cette stratégie pose un problème majeur de reproductibilité.

3-1-2 Marquage de l'agent de capture

- L'agent de capture est couplé à deux groupements fluorescents l'interaction avec un partenaire, en modifiant la distance entre les deux groupes, définit un donneur et un accepteur qui en fluoresce de manière spécifique (effet FRET 'Fluorescence Resonance Energy Transfer') cette approche a été appliquée aux 'Affibody'.

- L'agent de capture peut être couplé à un 'microcantilever' (Protiveris).

3-1-3 Marquage d'un agent de détection

L'agent de détection est un anticorps (dit secondaire) qui reconnaît le partenaire complexé avec l'agent de capture. Il peut être marqué de plusieurs manières, le signal obtenu, amplifié, procure une détection beaucoup plus sensible et une quantification plus aisée. L'inconvénient majeur réside dans le déploiement d'un autre système à haute affinité à côté de celui défini par l'agent de capture fixé sur la puce, **Figure 2**.

-- Un fluorophore peut être directement couplé à l'anticorps de détection, ou associé à la streptavidine qui interagit avec la biotine couplée à l'anticorps de détection **Figure 2**.

-- Une enzyme peut être couplée à l'anticorps secondaire sur le principe développé en immunocytochimie, l'amplification du signal est basée sur l'accumulation locale d'un composé, le substrat utilisé peut être fluorescent. L'alkaline phosphatase est très utilisée, la tyramide biotinilée est utilisée avec l'unité catalytique de la peroxydase du raifort (horseradish) le signal est amplifié avec un composant fluorescent couplé à la streptavidine, (Perkin Elmer).

-- D'autres systèmes d'amplification ont été développés, en particulier le 'RCA' (Rolling Circle Amplification), (Molecular Staging). Une amorce ADN est accrochée à l'anticorps de détection, à l'aide d'une ADN polymérase l'amorce permet de générer circulairement une molécule très longue et d'obtenir une sensibilité de détection de l'ordre de 100fg/ml (détection de cytokines et d'anticorps de type IgE).

-- On peut citer un protocole de détection ultrasensible le 'biobarcode', (ref JM Nam et al. Science, 2003, 301, 1884-1886) qui permet de détecter une protéine à l'échelle attomolaire.

Bien que permettant d'atteindre une grande sensibilité, la complexité de ces protocoles restent un frein à une application haut débit.

3-2. Détection de l'interaction sans modification des agents de capture ou des molécules analysées

3-2-1 Détection directe de l'interaction sur le 'microarray'

La détection est basée sur la modification intrinsèque des propriétés physico-chimiques du complexe formé lors de l'interaction. Plusieurs avantages : pas de manipulations supplémentaires des agents de capture ou des échantillons testés, une grande sensibilité (qui peut être supérieure à la fluorescence) et une bonne adaptation à l'analyse à haut débit de puces à haute densité.

- '*Surface Plasmon Resonance*', (SPR), (HTS Biosystems, Intrinsic Bioprobes, Biacore). Un rayonnement laser émis avec un angle défini peut créer un champ électromagnétique (dans les conditions de SPR) de part et d'autre de la couche métallique (définissant le plan de la puce). La propagation de ce champ dépend des caractéristiques physico-chimiques de la solution. Les interactions des partenaires (eg protéine-protéine) modifient les conditions de SPR et donc l'angle d'émission nécessaire, c'est la mesure de cette variation qui permet la détection de l'interaction.
- '*Raman scattering*'. Cette méthode très sensible est basée sur l'utilisation d'un laser et la mesure de la réémission énergétique spécifique au niveau atomique, une molécule unique peut être détectée (un pixel), (Biopraxis). L'objectif est d'analyser des milliers de pixels (100000) simultanément. Les spectres obtenus (comparés avec des banques spectrales) permettent d'aborder de multiples paramètres : constante d'affinité, constantes cinétiques permettant l'analyse de tests de compétition, de déplacement et de dénaturation.
- '*Resonance light scattering*' (Genicon Sciences).
- '*Atomic Force Microscopy*' (AFM), (BioForce Laboratories).
- *Ellipsométrie*, (Zyomix).

3-2-2 Détection indirecte de l'interaction

L'interaction est indirectement évaluée après désorption de la molécule associée à l'agent de capture et analyse en spectrométrie de masse (Applied Biosystem Group). Une limitation existe au niveau de l'identification des protéines qui doivent être préalablement fragmentées, aux très faibles concentrations mises en jeu, les protéases classiques sont inopérantes. Des procédés chimiques ont été développés, leur miniaturisation et intégration dans le spectromètre de masse sont actuellement des enjeux importants. Ciphergen développe cette technologie, les peptides adsorbés sur une surface sont désorbés par laser et directement analysés après ionisation : 'Surface Enhanced Laser Desorption Ionisation' (SELDI), néanmoins l'analyse de l'échantillon biologique n'est pas lié en amont à l'utilisation d'un agent de capture, le support permettant seulement une séparation liée après dépôt des molécules et ne rentre pas strictement dans le cadre des 'microarrays' définies dans ce rapport.

3-3 Détection d'une activité fonctionnelle

Il s'agit pour l'heure de tests d'activité enzymatique où les problèmes technologiques sont similaires à ceux mis en jeu dans l'immunocytochimie (voir 3-1-3) L'utilisation d'une surface plane hydratée nécessite, en particulier, que le produit de la réaction soit insoluble ou couplé à un composant de la puce. La société Kinexus (Canada) teste les substrats de protéines kinase, Ces enzymes ont un rôle central dans la transduction de signaux intra et extra-cellulaire en

phosphorylant des protéines cibles modifiant ou induisant ainsi leur activité, elles sont impliquées dans de multiples pathologies (une cinquantaine d'inhibiteurs de kinase sont en cours d'essais cliniques).

III Applications.

1- 'microarray à capture' : recherche de biomarqueurs dans le domaine du diagnostic et de la recherche pharmacologique.

Le développement du format puce a été initialement basé sur une application diagnostique qui reste, pour raisons commerciales, le principal moteur en terme d'innovation.

1-1 Diagnostic

1-1-1 Application à des pathologies humaines complexes

La recherche de biomarqueurs diagnostiques au niveau moléculaire reste un enjeu majeur. Les tentatives de caractérisation au niveau cellulaire et moléculaire de nombreuses pathologies à l'aide de 'microarrays' ADN (basées sur l'expression de combinaisons de marqueurs par quantification de l'ARN messager du gène correspondant) montrent une hétérogénéité et une complexité liée en partie au profil génétique original de chaque individu mais aussi à une limitation intrinsèque de l'utilisation d'un marqueur au niveau ARN. En effet les variations de niveau d'ARN apparaissent mal corrélées à ceux de la protéine correspondante limitant l'exploitation de ces résultats comme le montre l'exemple de l'alpha-foeto-protéine (AFP) qui est un bon marqueur de cancer hépatique au niveau protéique mais pas de son ARN. L'étude de différents types de cancer illustre, en particulier, la difficulté de définir des signatures spécifiques avec des combinaisons minimales de marqueurs. L'alternative du 'microarray' appliqué aux protéines représente ainsi dans le domaine du diagnostic une voie plus prometteuse, leur application est aujourd'hui centrée sur l'analyse sérologique qui avec ses avantages intrinsèques (prélèvement facilité, analyse de protéines solubles) permet d'aborder des pathologies très diverses, incluant des cancers, avec une sensibilité suffisante (de l'ordre du picomolaire, eg cytokines).

-- Zyomix a commercialisé une puce début 2003 avec 30 cytokines. La technologie mise en oeuvre autorise une quantification précise des protéines détectées dans l'échantillon.

-- Molecular Staging développe plusieurs puces : cytokines avec 75 anticorps validés, chemokines, facteurs de croissance, 133 marqueurs dans le champ de l'oncologie et la maladie d'Alzheimer.

-- Telechem : propose aussi des cytokines et développe des marqueurs liés aux maladies auto immunes, cancer, et apoptose.

-- Hypromatrix a développé et commercialisé trois puces au format membrane concernant les voies de transduction, l'apoptose et le cycle cellulaire.

-- Protagen (Allemagne) possède une collection de 2413 protéines recombinantes purifiées (tag polyhistidine), les agents de capture correspondent à des antigènes capables de détecter des anticorps sériques liés à des infections, des allergies (détection des IgE qui médie la réaction) et des maladies auto-immunes (selon la pathologie le patient sécrète différents anticorps contre ses propres protéines).

-- Somalogic développe une puce avec 75 aptamers liée à l'angiogenèse et aux réactions inflammatoires.

-- Ely Lilly, 150 cytokines dont 62 uniques

-- Zeptosens

1-1-2 Détection d'agents infectieux

Un grand nombre de maladies infectieuses (eg. salmonelle, malaria....) sont bien caractérisés en terme de réponse immunitaire. La validation des anticorps sériques produits en réponse à l'infection permet une nouvelle approche diagnostique grâce au format puce particulièrement bien adapté au développement d'un équipement portable pour une analyse in situ dans des zones isolées (pays tiers monde) ou nécessitant la mise en place de mesures sanitaires et d'un traitement rapide (cas d'une menace bioterroriste).

Une autre application se situe au niveau de la caractérisation des antigènes cibles du système immunitaire permettant le développement de nouveaux médicaments et vaccins (voir section III 1-2).

1-1-3 Détection (bio)organique

Des applications nombreuses en particulier dans le domaine de l'environnement (détection de polluants organiques), plus anecdotiquement un format 'microarray' devrait être utilisé pour la détection d'une activité biotique sur Mars (sondes 2007).

1-2 Applications pharmacologiques

1-2-1 Recherche de cibles thérapeutiques

Cette démarche correspond à la recherche de bio-marqueurs pathologiques basée sur la comparaison des profils d'expression, qualitative et/ou quantitative, des protéines entre un tissu malade et un tissu sain. La technologie Ciphergen a, en particulier, été utilisée avec succès pour caractériser des marqueurs du cancer de la prostate. L'approche protéomique donne en particulier accès à un plus grand nombre de cibles potentielles en considérant les variants des protéines qui reflètent leur activité physio(patho)logique (modifications post-traductionnelles, eg. phosphorylations)

1-2-2 Validation des cibles

Les approches génétiques et moléculaires récemment développées ont permis de caractériser un très grand nombre de cibles potentielles et parallèlement contribué aux nouvelles stratégies de validation fonctionnelle de ces cibles avec des modèles animaux et des systèmes de culture (évaluation des potentialités thérapeutiques de la protéine dans un processus physiopathologique sur la base de la modulation positive ou négative de son activité). L'étape de validation est aujourd'hui cruciale à double titre, d'une part afin d'exploiter pleinement cette richesse d'informations et d'autre part afin de fournir dans un temps minimal une réponse claire qui conditionne l'engagement de la suite du processus (caractérisation de composants pharmacologiques etc...). L'optimisation des temps et des coûts afférents de mise au point d'un nouveau composant sont, en effet, une priorité pour les sociétés concernées.

Dans ce contexte, le format 'microarray' peut être utilisée d'une part pour tester le niveau d'expression de la protéine analysée dans un test de validation comme cible et d'autre part réaliser une première évaluation d'effets adverses potentiels en testant les variations d'expression associées de familles de protéines clés.

1-2-3 Analyse pharmacoprotéomique

L'industrie biotech/pharmaceutique, auréolée des avancées et des promesses de l'ère génomique/post-génomique, amène de nouvelles générations de médicaments sur le marché (eg.cancer, SIDA...) et se trouve ainsi confrontée à une double attente de plus en plus grande des patients en terme d'efficacité et de sécurité. Cette dernière question a, en particulier, été mise en relief lors de la commercialisation et de la diffusion à très grande échelle de nouvelles molécules où la manifestation d'effets adverses observés chez un petit nombre de patients a mis en lumière l'hétérogénéité des patients face à un traitement standard d'une même pathologie.

La caractérisation au niveau moléculaire des effets thérapeutiques et adverses observés chez un patient lors d'un traitement est sur le principe possible grâce à l'intégration des outils génomiques et protéomiques au format 'microarray' : obtention d'une signature moléculaire spécifique individuelle à l'aide de familles de biomarqueurs. Cette approche, base du concept de 'médecine personnalisée', offre de nouvelles perspectives en terme de définition des propriétés thérapeutiques d'une molécule avec des bénéfices accrus pour les malades (demandeurs de nouvelles thérapies) et les sociétés impliquées qui pourraient réduire la prise de risque sur leurs investissements. Le principe est de définir au cours des phases initiales d'essais cliniques des petits groupes de patients (allant des effets adverses maximaux à un gain thérapeutique optimum) en reliant la réponse à un profil moléculaire, les phases ultérieures

des tests sont effectuée des cohortes préselectionnées sur la base de ces profil et non plus seulement de leur état pathologique augmentant largement les chances d'obtenir la validation d'une nouvelle molécule non seulement en tant que telle mais éventuellement, sur ces bases, en combinaison avec d'autres composants susceptibles de réduire, par exemple, les effets adverses observés dans les autres groupes de patients. In fine des traitements optimisés pour différents groupes de malades pronostiqués avant application seraient ainsi offerts.

2- 'microarrays fonctionnels'.

2-1. Dissection mécanistique des processus cellulaires : caractérisation des réseaux d'interactions

Les 'microarrays' permettent d'aborder dans un contexte in vitro, à l'échelle de fractions du protéome, les bases mécanistiques de processus cellulaires en permettant d'identifier les protéines partenaires dont les interactions conditionnent ces processus. Ces questions essentiellement abordées in vivo avec la levure par la technologie du double-hybride restent lourdes à mettre en oeuvre avec une capacité de contrôle limitée qui ne permet pas d'éliminer les très nombreuses interactions non spécifiques induites. Le format 'microarray' peut contribuer à des objectif larges comme l'établissement de cartes d'interactions protéiques dans différents tissus et situations physiologiques.

Un 'microarray' portant une fraction significative des protéines (définies par les cadre de lecture ouverts) de la levure soit 5800 protéines a permis une caractérisation semi-exhaustive des partenaires de la calmoduline, une protéine clé dans le contrôle des processus cellulaires contrôlé par le calcium (Protometrix, récemment racheté par InVitrogen). Une autre application importante concerne l'analyse des cascades de phosphorylations liées à l'activité des protéines kinase centrales dans la signalisation cellulaire (description du 'kinéome'). A côté de l'aspect haut débit, le format 'microarray' permet d'aborder in vitro des classes d'interactions difficiles à détecter par des approches classiques, ProteinOne commercialise des membranes avec des collections spécialisées de famille de protéines (eg facteurs de transcription) qui permettent par exemple de détecter des interactions très labiles entre des cofacteurs de type TATA et l'ARN polymérase. InVitrogen (qui a établi un partenariat privilégié avec Zyomix) vient de commercialiser un 'microarray' humain avec 1800 protéines recombinantes GST sur un support nitrocellulose. Couvrant une large représentation d'importantes classes fonctionnelles de protéines, la détection des interactions est basée sur la biotinylation des échantillons testés et l'utilisation de streptavidine couplée à un fluorophore.

2-2 Recherche et caractérisation de nouvelles molécules pharmacologiques

Ce format est bien adapté à l'identification à haut débit de nouvelles molécules à valeur pharmacologique :

- Recherche de partenaires dans des bibliothèques combinatoires de composants, eg. analyse de substrats de kinase, recherche d'inhibiteurs spécifiques (voir section II 3-3).
- Caractérisation de molécules sélectionnées : estimation de l'affinité, évaluation de leurs propriétés (eg. agonistes/antagonistes). Afin d'explorer ces questions avec les 'G protein coupled receptors' (les GPCRs sont des protéines membranaires d'importance stratégique puisqu'elles représentent 45% des cibles pharmacologiques actuelles et 20% des 200 médicaments les plus vendus) Corning Inc. Life Sciences a développé un 'microarray' qui permet de les tester dans un environnement lipidique compatible avec le maintien de leur conformation et de leur mobilité, les protéines sont déposées dans des microgouttes lipidiques ('pinprinter') sur une surface traitée avec un dérivé de silane qui offre un bon compromis entre fluidité et résistance mécanique de ces microgouttes exposées brièvement à l'air entre les incubations et les lavages.

IV Conclusion.

Un large spectre d'applications des 'microarrays' appliqués aux protéines motive l'ensemble des développements technologiques présentés (agents de capture, stratégies de fabrication de la puce, détection et traitement des signaux). Ce marché est évalué à 400 millions \$ en 2007 (100 millions en 2002), avec plus de 70 sociétés aux USA engagées sur une ou plusieurs des composantes évoquées. Le développement des 'microarrays capture' appliqué en particulier à l'immunodiagnostic constitue toujours le principal moteur de cette activité concrétisée par la commercialisation d'une première génération (Zyomyx, Pierce Boston Technology, Zeptosens, BD Biosciences Clontech, Biorad et Sigma). Néanmoins les coûts plus élevés de développement et de fabrication nécessaires ne rendent pas forcément cette approche compétitive par rapport aux tests déjà disponibles sur le marché. Par ailleurs, en terme d'innovation pharmaceutique le format 'microarray fonctionnel' peut contribuer à accélérer de nombreuses étapes dans le processus de développement d'une nouvelle molécule (InVitrogen, Corning, ProteinOne).

L'hétérogénéité et la complexité des protéines en terme de conformation et d'activité restent toujours au cœur des problèmes posés lors de la mise en oeuvre des 'microarrays'. La production et la validation de la spécificité des anticorps sont un problème central ainsi que celui de la reproductibilité expérimentale : instabilité des protéines, liée en particulier à la

complexité des échantillons analysés et des conditions de tests, corrélée à la genèse de signaux artefactuels et détection de protéines en très faible concentration dans les échantillons (miniaturisation accrue nécessaire). Les nombreux développements dans le domaine de la 'microfluidique' (circulation des échantillons) et l'association avec la spectrométrie de masse seront certainement des éléments de réponse à ces problèmes, d'autres problèmes restent cependant difficiles à résoudre comme la caractérisation de protéines physiologiquement intégrées dans des complexes multiprotéiques.

La capacité à réaliser des 'microarrays' à haute densité (testée par Bioforce) et la production à haut débit d'anticorps par Milagen illustrent la possibilité d'une approche à grande échelle, au niveau du protéome, avec de nombreuses applications potentielles en recherche fondamentale et pharmacologique. Néanmoins l'ensemble des problèmes évoqués associé aux contraintes de coût et de rentabilisation des investissements laisse à penser qu'un nombre limité de plateformes technologiques devrait émerger dans un proche avenir. Elles ne devraient mettre en œuvre que des sous-fractions du protéome et des familles spécifiques de protéines pour des applications ciblées en diagnostic et recherche pharmaceutique (différents types de cancer, diabète, obésité etc...)

Plus prospectivement l'obtention d'anticorps contre les variants post-traductionnels qui reflètent l'état fonctionnel des protéines constitueront un développement important dans l'exploration de ces questions à l'aide de l'outil 'microarray', d'autre part de nombreuses sociétés développent des 'microarrays' non-planaires (MetriGenix), 'beads arrays' avec Luminex, BD Biosciences, Bang Laboratories, Illumina, PharmaSeq, Smartbeads et Nanoplex, ces approches ne permettent pour l'instant qu'une analyse parallèle numériquement plus limitée par rapport au format planaire, néanmoins avec l'application de technologies de type 'biobarcode' aux 'beads', elles pourraient devenir plus compétitives dans le futur.

Sociétés impliquées dans le développement des 'microarrays' appliqués aux protéines. La société est américaine par défaut si le pays d'appartenance n'est pas indiqué, les sociétés dont l'activité est évoquée dans le rapport sont soulignées.

- A
 - Accelr. Technology
 - Aclara Biosciences
 - Advantix (Allemagne)
 - Advion Biosciences
 - Affibody
 - Affitech (Norvège)
 - Akceli
 - Alpha Innotech
 - Amersham Biosciences (Suède, USA)
 - Applied Biosystems
 - Applied Precision
 - Archemix
 - Aspira Biosystems
 - Axon Instruments
- B
 - Bangs Laboratory
 - BD Biosciences Clontech
 - Berthold Technologies (Grande-Bretagne)
 - Biacore (Suède)
 - BioCat GmbH (Allemagne)
 - BioArray Solutions
 - Biocept
 - Biodot
 - BioForce Nanosciences
 - BioInsights
 - BioInvent (Suède)
 - BioMedTech Laboratories
 - Biopraxis
 - Bio-Rad
 - Biosite
 - Biotrace
 - Biotrove
- C
 - Cambridge Antibody Technology (Grande-Bretagne)
 - Cartesian Technologie
 - Ciphergen
 - Combimatrix
 - Corning Life Sciences
 - Covalys (Suisse)
 - CSEM (Suisse)
 - CyBio (Allemagne)
- D
 - Denovo Biolabels (Allemagne)
 - DiagnoSwiss (Suisse)
 - Discerna (Grande-Bretagne)
 - Domantis (Grande-Bretagne)
 - Dyax (Grande-Bretagne)
 - Dyomics (Allemagne)
- E
 - Ely Lilly

EMD (Allemagne).
 Eurogentec (Belgique)
 EngeneOS
 Everest Biotech (Grande-Bretagne-Népal)
 Exigon (Danemark)
 Expression Pathology
 F Fluidigm
 G Genescan (Allemagne)
 Genetix (Grande-Bretagne)
Genicon Sciences
 Genomic Solutions
 Genosens Diagnostics (Autriche)
 GenTel Biosurfaces
 GeSim (Allemagne)
 Glycodata (Israël)
 Glycominds (Israël)
 GPC-Biotech (Allemagne)
Gyros (Suède)
 H HTS Biosystems
 Human Genome Sciences
Hypromatrix
 I Illumina
 Imaging Research (Canada)
 Ingenium Pharmaceuticals (Allemagne)
Intrinsic Bioprobes
Invitrogen
 Iobioninformatics
 J Jerini AG (Allemagne)
 JPT Technologies (Allemagne)
 K Kinexus (Canada)
 Kreatech Biotechnology (Hollande)
 L Large Scale Biology
 Li-Cor Biosciences
Lilly
Living Chips
 LumiCyte
Luminex Corporation
 M Merrimack Pharmaceuticals
 Meso Scale Discovery
MetriGenix
 Microsurfaces
Milagen
 MiraiBio (USA-Japon)
Molecular Staging
Morphosys
 N Nanoliter
 Nanolytics
Nanoplex Technologies
Nanotype (Allemagne)
 Neupro Technology (Taiwan)

Nextgen (Grande-Bretagne)
 NoAb Biodiscoveries (Canada)
 O Oxford Glycosciences (Grande-Bretagne)
 P Packard BioScience
 Pall Corporation
 Pamgene (Hollande)
 Panomics
 Pepsan (Hollande)
PerkinElmer Life Sciences
PharmaSeq
Phylos
Pierce Boston Technology
 Pieris (Allemagne)
 Procognia (Grande-Bretagne)
 Prolinx
 Promega
Protagen (Allemagne)
 Protea Biosciences
ProteinOne
 ProteomTech
 Proteom System (Australie)
Protiveris
Protometrix
 Q Qiagen
 Quantifoil
 Quantum Dot
 R Radius Biosciences
 Radox (Grande-Bretagne)
 RayBiotech
 Roche Diagnostics
 Rules-Based Medicine
 S Scandinavian Micro Biodevices (Danemark)
 Schleicher & Schuell Bioscience
 Schott Nexterion (Allemagne)
 Serologicals
Sigma-Aldrich
SomaLogic
Smartbead Technologies (Grande-Bretagne)
 Spendlove Research Foundation
 SRU Biosystems
 Surromed
 T Tecan
 TekCel
 Telechem International
 U Upstate
 V V&P Scientific
 VBC Genomics (Autriche)
 X Xantec (Allemagne)
 X-Mine
 Xenopore

Z Zyomyx
Zeptosens (Suisse)

DEUX TYPES DE 'MICROARRAYS' PROTEINES

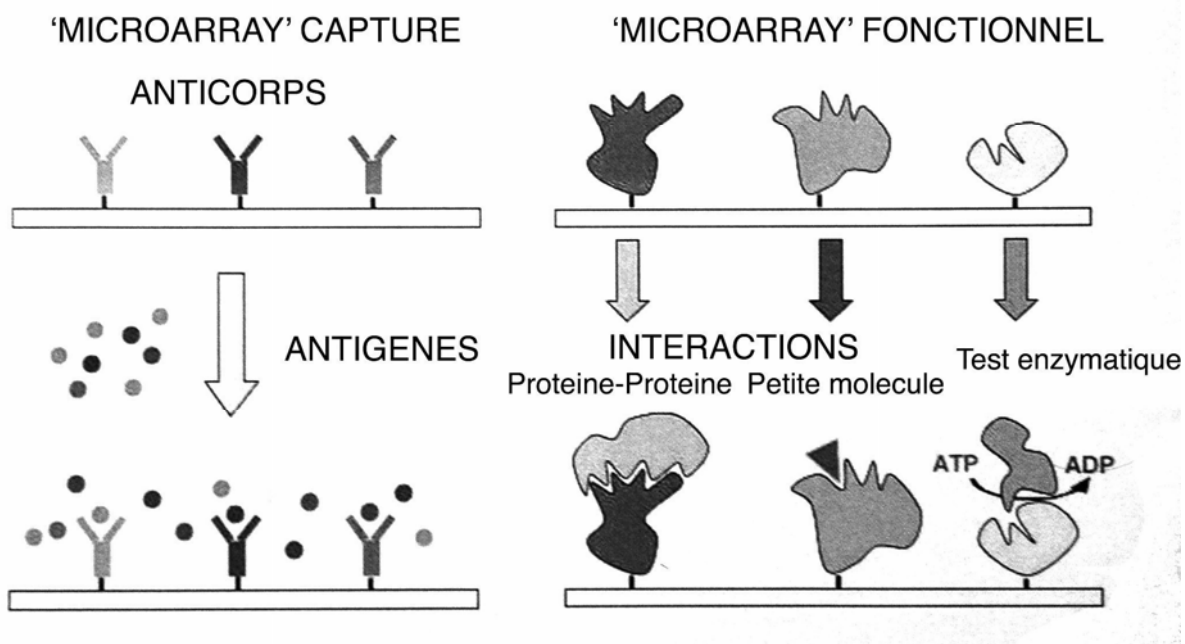


Figure 1

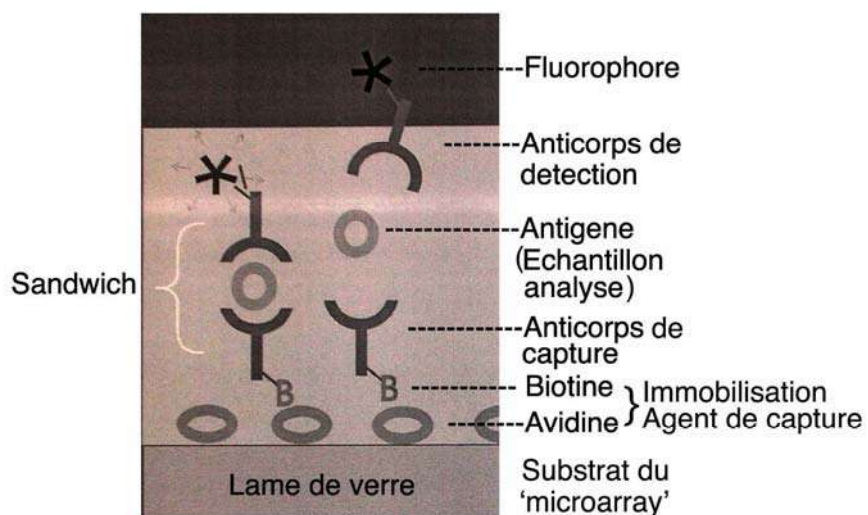


Figure 2