Protein microarrays: challenges and promises

Anupam Talapatra[1], Richard Rouse[2] & Gary Hardiman[2]

[1]Department of Vascular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA, 92037, USA
[2]Biomedical Genomics Microarray Facility, Division of Biology, University of California San Diego, 9500 Gilman Drive, La Jolla CA 92093-0349, USA
Tel: +1 858 822 3792; Fax: +1 858 822 3021; E-mail:ghardiman@ucsd.edu

Genomics and proteomics are playing increasingly important roles as discovery tools in basic biological sciences and as diagnostic and rational therapeutic aids in the clinical arena. In recent years, high-density arrays of specified DNA sequences have gained popularity. Protein microarrays are at the forefront of this biochip revolution and promise the parallel examination of large numbers of proteins. These miniaturized arrays are currently being developed to facilitate high analytical resolution, detection sensitivity and sample throughput. Many challenges are presented by proteome scale manipulation of proteins, as there is currently no methodological equivalent to the gene chip for comparative proteomics.

Keywords: antibodies, comparative proteomics, expression analyses, protein microarrays, proteome

Ashley Publications Ltd
www.ashley-pub.com

In the strictest sense of the definition of 'proteome', only one proteome exists in each organism, namely the full complement of proteins encoded by its genome. A more common usage of the term 'proteome' has crept into the literature in recent years and this definition refers to all of the proteins expressed by a particular cell or tissue at a specific time. Proteomics is therefore a global study of gene expression at the protein level, encompassing comparative, functional and structural proteomics. Comparative proteomics thus seeks to quantify the relative abundance of each protein species present in two or more proteomes. Quantitative comparison of these proteomes is a major goal of proteomics, analogous to gene expression profiling.

DNA microarray and oligonucleotide genechips have emerged as powerful tools for gene expression profiling on a genomic scale and for establishing functional relationships between the large number of genes involved in distinct cellular processes [1-4]. In addition to detection of DNA copy-number and localization of transcription factor binding, nucleic acid arrays have been extensively utilized for the detection of gene transcription [5,6]. mRNA abundance in a cell often correlates poorly with the amount of protein synthesized, and proteins rather than mRNA transcripts are the major effector molecules in the cell [7]. DNA microarrays have little utility in identifying physiologically relevant post-translational modifications of proteins, which influence protein function. Finally, only protein expression analyses are possible in the case of samples lacking mRNA, such as bodily fluids like urine. Thus the question often raised is why comparative proteomics is not the method of choice for the global analysis of gene expression.

Development of protein microarrays

Biochemical studies of protein activity have traditionally focused on the analyses of single molecular species. The rapid pace of discovery of new gene products by large-scale genomic and proteomic initiatives has required the development of high-throughput strategies to elucidate their function [8]. There have primarily been two approaches to characterize multiple proteins in biological samples. The first approach, 2D-gels, has been widely used to separate and visualize 2000–10,000 proteins in a single experiment. Proteins can subsequently be excised and identified by mass spectrometry (MS). This approach is both time consuming and unsuitable for the analysis of low abundant proteins. Due to limitations with 2D-gel separation technology, increasing attention is being focused on a second approach, the development of protein microarrays as an alternative and complementary approach [9-11].

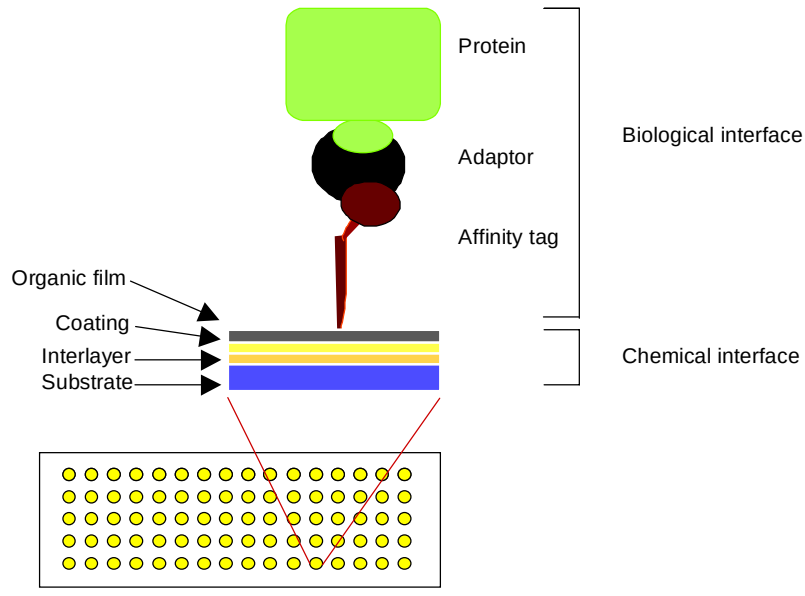
Protein arrays are comprised of a library of proteins immobilized in a 2D addressable grid on a chip. Protein microarray biochips extract and retain targets from liquid media and remain distinct from microfluidic biochips, which separate and process proteins in a transport medium *in situ* using microfluidic devices [12,13]. Different chip formats currently exist, including glass and matrix slides and nanowells. A typical array may contain 10^3–10^4 spatially distinct elements within a total area of 1 cm^2 [14].

Array architecture and surface chemistry

Generally, all protein arrays are composed of a substrate, which constitutes the underlying core material of the array and a protein attachment layer (**Figure 1**). Preferred substrates include materials based on the following:

 ISSN 1462-2416

Figure 1. Microarray architecture.



The protein is immobilized directly or via an affinity tag (optional) onto an organic layer that forms a thin film on the array substrate. Each array is comprised of a precise, micrometer-scale, 2D addressable grid of biological molecules immobilized on an organic thin film coated on a substrate.

- glass
- silicon (Zyomyx, Hayward, CA, USA; Lumicyte, Fremont, CA, USA)
- plastic (Large Scale Biology, Vacaville, CA, USA)
- synthetic polymers (e.g., polystyrene, polyvinylidenedifluoride [PVDF], nitrocellulose)

Array substrates may have a coating layer of gold, aluminum or other metals. These coatings are applied using thin-film technologies, such as physical vapor deposition (PVD) or chemical vapor deposition (CVD) [15]. Chip surfaces may have adhesion interlayers (e.g., epoxy glue) that bind coatings to the substrate. An interlayer of titanium or chromium, may be used to fix a gold coating to a silicon wafer substrate.

The protein attachment layer on the chip surface is typically an organic film < 20 nm thick. The thickness varies with the nature of application. The choice of material selected for attachment is based on the ability of a particular chemistry to immobilize proteins without denaturing them. Generally, hydrophilic materials perform well as they promote protein stability and binding. Materials that have been studied include agarose [16], dextran-based hydrogel (Biacore, Uppsala, Sweden) [17], porous polyacrylamide hydrogel (Packard Bioscience, Meriden, CT, USA), hydrophilic polymers and polyamino acids (e.g., poly-L-lysine coated microscopic slides) [18].

There is considerable interest in using self-assembled monolayers for the purpose of protein attachment. Monolayers investigated were composed of aldehyde-containing silanes [19] and alkylthiols (Interactiva Biotechnology, Ulm, Germany). Aldehyde groups readily react with primary amines on proteins to form Schiff's base linkage. The thin film of the protein array XNA on Gold, developed by Interactiva, is based on a self-assembling monolayer of long chain alkylthiols adsorbed onto a 100 nm thick coating of 24-carat gold [20]. The film surface can also be derivatized with thiol-reactive maleimidyl groups, amine-reactive N-hydroxysuccinimidyl (NHS) groups or other functionalities that capture proteins.

Amino acids or carbohydrate moieties inherent to proteins have been used to tether proteins directly to organic thin films. However, in many instances, affinity tags and adaptor molecules offer specific advantages in terms of protein immobilization. It should be noted that although affinity tags do not necessarily promote accessibility of the active or binding sites in pro-

Table 1. Affinity tags utilized for constructing protein microarrays.

Affinity tags	Mode of attachment
Poly-amino acid	
Poly-His	Nickel resin
Poly-lysine	Amide & Schiff base linkages
Poly-cysteine	Thioether linkage
Biotin	Streptavidin
Protein G	Antibody Fc
Protein A	Antibody Fc
Recombinant fusion proteins	Antibody
GST	Anti-GST
MBP	Anti-MBP
TRX	Anti-TRX
GFP	Anti-GFP
Poly-His	Nickel resin

GFP: Green-fluorescent protein; GST: Glutathione-S-transferase; MBP: Maltose-binding protein; TRX: Thioredoxin.

teins, they confer enhanced site-specific attachment and oriented immobilization of capture proteins by binding to reactive groups of organic thin films. Table 1 outlines various tags used for protein capture. An affinity or epitope tag is usually an intact protein, a polypeptide or poly-amino acid (e.g., a poly-His tag for nickel binding sites, poly-lysine for amide or Schiff base linkages, poly-cysteine for thioether linkages). For example, streptavidin based immobilization schemes have widely been employed to attach any biotinylated biological element to array surfaces [21]. Biacore International Ltd and Interactiva Biotechnology have developed chip surfaces that employ streptavidin based immobilization schemes for the capture of biotinylated proteins. Protein G and protein A have been used as tags to immobilize the Fc regions of antibodies. Recombinant proteins have been genetically fused with glutathione-S-transferase (GST), maltose-binding protein (MBP), thioredoxin, green fluorescent protein (GFP) and poly-His for convenient attachment, mediated by binding of the respective affinity tags to the corresponding antibodies imprinted on arrays. The affinity tag–thin film combination allows for gentle immobilization conditions that maintain protein stability and function. Affinity tags permit a common immobilization strategy that can be applied to a variety of different proteins.

Protein microarray surfaces can be loosely classified as chemical or biological in nature. Chemical surfaces have been designed to chromatographically separate proteins [22]. The chip surface is comprised of a stationary phase that consists of a hydrophobic, charged or metal affinity surface. These surfaces selectively extract femtomole quantities of proteins directly from biological samples (Ciphergen Biosystems, Fremont, CA, USA). The principle is based on ion-exchange, reverse-phase or affinity chromatography. Biological surfaces are comprised of arrays of different capture molecules that utilize highly specific molecular recognition mechanisms for protein capture from complex biological fluids or libraries. These typically include antibodies or antibody fragments.

Generation of microarrays

Microarray manufacture is a highly automated process that involves imprinting of 'capture' molecules on a bio-reactive film on an array or slide surface in a 2D array format [23]. The sample to be immobilized can also be introduced via microfluidic channels in a continuous flow stream [24]. The spacing between spots depends on the size of the capture agents, antibody arrays typically have 375 μm feature spacing [18]. Alternatively, peptides have been synthesized on planar surfaces in an array format using photolithography [25] or SPOT technology [26]. Using photolithography, Zyomyx has etched miniature wells on the surface of silicon arrays. This format has enabled the manufacturing of high-density antibody arrays capable of detecting up to 10,000 proteins in parallel.

Protein array design difficulties

For nucleic acid biochips, miniaturization has led to an enormous increase in throughput, a decrease in bio-reagent cost and a highly parallel data collection process. In contrast, the inherent structural diversity and complexity in proteins has made the development of protein arrays technically very difficult. RNA based analysis offers many technical advantages over protein analysis in that mRNA molecules are relatively homogenous and possess high affinity and high specificity binding partners. Proteins in comparison do not possess straightforward binding partners and each exhibits diverse biochemical features. There is no polymerase chain reaction (PCR) equivalent that facilitates rapid production of proteins. Several major technical hurdles thus face protein array technology namely the acquisition, arraying and stable attachment of proteins to array surfaces and detection of interacting proteins.

High-throughput expression and purification of proteins

Thousands of purified proteins are required for the generation of high-density protein microarrays. Cloning systems utilizing site-specific recombination are routinely employed for high-throughput cloning and expression of protein sets [27]. Time consuming cloning, DNA sequence confirmation and gene identification procedures remain, however, cumbersome precursors to protein expression. An approach that expedites this process is the generation of protein microarrays from histidine (His)-tagged expression cDNA libraries. Bussow *et al.* arrayed His epitope tagged recombinant proteins from a human fetal brain cDNA expression library on nitrocellulose filters [28]. Expressed proteins can be readily immobilized with anti-His capture antibodies and cDNA inserts from clones of interest can be sequenced and identified. Comprehensive production of functional recombinant proteins is still, however, impeded by poor expression, aggregation and degradation of many eukaryotic proteins in bacterial systems. Zhu *et al.* employed yeast *Saccharomyces cerevisiae* as an alternative to bacteria and expressed 5800 yeast open reading frames (ORFs) as His tagged proteins [14]. Rapid protein purification was enabled through the use of a nickel resin. A novel cell free 'protein *in situ* array' (PISA) format has recently been described in which arrays are rapidly generated directly from DNA templates by rabbit reticulocyte expression. Proteins are simultaneously immobilized *in situ* via His residues to nickel coated surfaces [29]. This one step cell free system permits processing of a variety of post-translational modifications and is capable of assembling multisubunit functional protein complexes. An advantage over cell based gene expression systems is that PISA allows reliable expression of toxic proteins.

Stable attachment of proteins

Proteins are highly sensitive to the physiochemical properties of the chip support material. Polar arrays, for example, are chemically treated to bind hydrophilic proteins. Such surfaces are unsuitable for cell membrane proteins, such as G-protein-coupled receptors, possessing exposed hydrophobic patches. Unlike nucleic acids, proteins do not all behave in a similar fashion when exposed to the same surface chemistry. Surface chemistries may therefore promote the retention of some proteins and cause denaturation or loss in activity of others. Proteins that are soluble in their native environments may precipitate on chip surfaces. Consequently, it can be difficult to select a surface chemistry that permits diverse proteins to retain their native folded conformation and biological activity. Affinity tags are utilized to counteract such surface problems by providing gentle immobilization conditions that maintain protein stability and function.

Miniaturization of assays and protein dehydration

Miniaturization of assays to sub-microliter volumes often leads to problems of evaporation, denaturation and inactivation of proteins at the liquid-well and liquid-air interfaces, due to a substantial increase in surface to volume ratio. It is critical that proteins remain hydrated at ambient temperatures. The presence of a humidifier during printing and the inclusion of glycerol in the samples further helps to prevent sample evaporation. Nanoliter droplets of 40% glycerol remain completely hydrated at ambient temperatures, even when exposed overnight to the laboratory environment [30]. Matrix slides and nanowell array formats help reduce evaporation compared to glass slides. Glass slides are compatible with standard microarrayer and detection equipment and are relatively inexpensive but they are prone to high evaporation rates and sample cross-contamination. Matrix slides and nanowells reduce evaporation and minimize cross contamination but are more expensive, particularly matrix slides as they are fabricated by expensive photolithography.

Availability of highly specific antibodies

Historically, antibodies have represented the best all-purpose high affinity, high selectivity protein-binding reagent as they can be produced in both high quantity and purity for arraying purposes. The biggest challenge to antibody arrays is to obtain antibodies against the ≥ 100,000 proteins that comprise the human proteome. Currently there are antibodies available for a mere fraction of the proteome. Furthermore the specificity of many of these antibodies is poorly documented and additional antibodies may be required to allow the detection of post-translational modifications. A second fluorescently tagged antibody may be required for detection, raising the number of antibodies per arrayed protein to a minimum of two. Many antibody arrays are thus limited and contain a few well-defined capture agents directed at a particular class of protein markers.

Cross-reactivity of antibodies

Many antibodies are glycosylated and contain large protein-based supporting structures. Consequently they often cross-react with more than one target protein. This can contribute to large numbers of false positives. Smaller antibody fragments prepared using phage display (Cambridge Antibody Technology, Cambridge, UK; Dyax, Cambridge, MA, USA) may minimize the interaction of non-target proteins with a particular antibody. A predetermined knowledge of target antigen specificity can be imparted to array antibody candidates by preselection from antibody expression libraries [31-33]. Preselection enables the generation of highly specific antibody microarrays. Protein binding molecules other than antibodies, namely aptamers (Somalogic, Boulder, CO, USA) and fibronectin based peptide scaffolds (Phylos, Lexington, MA, USA) are useful alternatives to antibodies. Aptamers are protein binding RNA molecules, which are relatively easy to select for, synthesize and array. Purified protein target is however a requirement for their use and aptamers can exhibit biased binding as RNAs tend to be highly negatively charged.

Detection of bound target is more complex than with DNA microarrays

Analyte-selective binding and specific retention on the array surface proceeds via thermodynamically driven binding mechanisms similar to hybridization of nucleic acid targets to probes. Detection of bound target is considerably more complex though than with nucleic acid arrays. Although the proteomes under comparison can be labelled in a comparable fashion with different fluorophores, the reproducibility of these chemical reactions is poor and interference with the protein–antibody interactions presents an additional complexity. ELISA based detection methods suffer from the non-specificity of protein–antibody interactions, leading to many false positives. Non-uniform labeling of proteins can be addressed by performing a dual color ratiometric assay, where an internal standard is present for each target protein to be measured [18]. Labeling proteins with fluorophores has, however, the disadvantage of reducing the quantitative accuracy of such assays, as incorporation of a label can alter protein binding to other molecules. Although direct label detection methods are still widely used, the inherent problems with their use has resulted in the application of label free detection methods to protein microarrays. These schemes typically employ MS [34] or surface plasmon resonance (SPR) [35].

Dynamic range for detection

The range in cells of protein concentrations is several orders of magnitude greater than that for mRNAs. Protein microarray detectors require a greater dynamic range of operation, up to a factor of 10^8, as compared to 10^4 for mRNAs. An antibody with nanomolar affinity to a particular target is saturated by the presence of this target at micromolar concentrations and fails to detect pico- or femtomolar target levels. This presents a problem in the design of microarrays as separate arrays need to be created for the detection of rare and abundant proteins, respectively. Alternatively, multiple antibodies with varying affinities for the target can be positioned in different locations on the array. However, studies to date have shown that only 20% of arrayed antibodies provide measurements of proteins at low concentrations [18].

Non-specific binding

Non-specific binding of non-target proteins to the chip surface needs to be minimized. Following attachment of the capture molecules to the array, unreacted groups are typically quenched and subsequent non-specific binding minimized by immersing the arrays in a bovine serum albumin based (BSA) buffer [30].

Applications of protein arrays

Arrays can be engineered to address protein identification, quantitation and affinity studies. A profiling array quantitates levels of specific proteins on a global scale. An affinity array probes the interaction of peptides, proteins, oligonucleotides, oligosaccharides or small molecules with immobilized proteins, which are typically receptors, enzymes or antibodies. Arrays can be further classified as finite or chemical and self-replicating or biological [36]. Finite (chemical) arrays are involved in specific binding analyses, enzymatic assays or in high-throughput screening for discrete biochemical activities [14]. Biological (self-replicating) affinity arrays utilize living cells and facilitate functional studies of complex biological processes in a cellular environment.

Cagney *et al.* employed yeast cells to construct the first genome-scale 'living' protein array, which enabled systematic two-hybrid genetic screening of each ORF for interacting proteins [37]. In total, 192 ORFs were tested, 87 yielded putative interacting partners, ultimately generat-

ing 281 novel protein–protein pairs. Variations of two-hybrid based arrays have been used for DNA–protein, RNA–protein and small–molecule protein interaction screens and for the rapid identification of protein–protein interaction networks within a cell [38,39]. Arraying and screening of live cells has the advantage of avoiding protein unfolding and denaturation issues that arise when proteins are in contact with chemically treated surfaces. However, this system relies on protein interactions within the yeast nucleus, which is not the natural habitat of cytosolic or membrane proteins. An improvement to this system was the yeast affinity microarray created by Zhu *et al.* to uncover novel protein–protein and protein–lipid interactions [14]. The yeast proteome consisting of 6566 His-tagged yeast ORFs, representing ~ 5800 different yeast proteins, was probed respectively with biotinylated calmodulin and biotinylated phosphoinositides and bound molecules were detected with Cy3-labeled streptavidin. In total, 33 novel calmodulin binding proteins and 52 phosphoinositide binding proteins were discovered. Live array formats facilitate screening for an activity with a visual readout, such as radioactivity, fluorescence or chemiluminescence and detection of new protein functions in a relatively unbiased manner. A disadvantage with this format is that the arrays are relatively biased in terms of expression. Few membrane or secreted proteins, for example, are functionally expressed.

Chemical binding surfaces, such as the protein chip from Ciphergen Biosystems are useful alternatives to antibody arrays. These surfaces are non-specific and permit the binding of many diverse proteins. Although the mass spectrophotometric readout is rapid, the limited spatial separation of bound proteins can lead to a heavy bias towards the detection of abundant proteins. Comparison of mass profiles of proteins from normal and variant samples can, however, reveal differential protein expression in the two samples. Cai *et al.* used this chip to monitor the effects of various drugs on amyloid precursor protein (APP) fragmentation patterns [40]. Austen *et al.* also utlized this system to examine the effects of cholesterol on the biosynthesis and secretion of 1–40 and 1–42 forms of β-amyloid [41]. An anti-β-amyloid antibody that recognizes only cleaved forms of APP was immobilized on the chip via an amide linkage and surface enhanced laser desorption/ionization (SELDI) MS was used to analyze both peptides captured from culture medium by the antibody. The study demonstrated that treatment with cholesterol reduced secretion of both peptides from cultured cells. A recent study successfully applied this technology to analyse proteomic patterns in serum and facilitate the correct identification of ovarian cancer samples. This technology can be used as a screening tool for all stages of ovarian cancer in high risk and general populations [42].

Commercial protein chips

Many protein arrays are being developed in a commercial setting. **Table 2** outlines selected companies that are developing protein microarrays. Ciphergen Biosystems has developed protein chip arrays that contain chromatographic and biological surfaces. The Ciphergen protein array contain eight or 16 different spots per array and bound protein molecules at each spot are detected by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS. LumiCyte employ a similar technology. One major advantage of SELDI or 'seldiography' is that protein identities are not required prior to experimentation, unlike antibody based approaches. Seldiography permits the analysis and comparison of multiple proteins that are increasing and decreasing relative to quantities of other proteins.

SPR-based protein biosensor arrays (Biacore) containing four spots per array have been in use for several years [24]. The CM5 biosensor chip consists of a glass substrate coated with a 50 nm thick gold film, to which a carboxymethylated dextran matrix is attached via an inert thiohydroxy alkane linker layer. Several alternative surfaces have been developed, including hydrophobic surfaces, metal chelation surfaces and carboxymethylated dextran surfaces modified with streptavidin. SPR is currently being developed for protein microarray applications and SPR biosensors have been integrated with MALDI-TOF MS [43,44]. The thin film of protein array biochips (XNA on Gold) Interactiva Biotechnology utilizes long chain alkylthiol monolayers on gold. Biotin is covalently coupled to the film surface and saturated with streptavidin. The stability of the chemically bonded biotinylated monolayer combined with the high-affinity of the biotin–streptavidin linker layer facilitates attachment of biotinylated biological elements. The SPR phenomenon of light being coupled to the surface plasmon can be accomplished in a few different ways, the most common by use of a prism or a grating on the metal surface. HTS Biosystems

Table 2. Selected companies developing microarray technology.

Company	Chip material	Capture agents	Signal detection technology
Ciphergen Biosystems (Fremont, CA, USA)	Aluminum coated silicon wafer	Metal affinity, charged or hydrophobic chromatographic surface, antibodies	SELDI and time-of-flight MS
Lumicyte (Fremont, CA)	Silicon wafer	Chemical, biochemical or biological affinity surface	SELDI and time of-flight-MS
Biacore (Uppsala, Sweden)	Glass slide coated with layer of gold, dextran hydrogel surface	NHS/EDC activated surface, Ni^{2+} surface, antibodies, streptavidin	Prism based SPR technology
HTS Biosystems (Hopkinton, MA, USA)	Plastic with fine grating molded on surface, coated with a thin layer of gold	Antibodies and antibody fragments	Grating-coupled SPR technology
Large Scale Biology (Vacaville, CA, USA)	Plastic	Antibodies	Fluorescence
Biosite Diagnostics (San Diego, CA, USA)	Plastic	Antibodies	Fluorescence
Zyomyx (Hayward, CA, USA)	Silicon	Antibodies and antibody fragments (selected using phage display technology)	Fluorescence
Phylos (Lexington, MA, USA)	N/A	Fibronectin based polypeptide scaffold molecules (antibody mimics)	Fluorescence
Somalogic (Boulder, CO, USA)	N/A	Aptamers (protein binding RNA molecules)	Fluorescent
Akceli (Cambridge, MA, USA)	Glass slide	cDNA expressed by by cultured human embryonic kidney cells	Fluorescence
Packard Bioscience (Meriden, CT, USA) Oxford Glycosciences (Oxford, UK)	Polyacrylamide hydrogel	Antibodies	Fluorescence
Interactiva Biotechnology (Ulm, Germany)	Long chain alkylthiols on a gold surface, biotinylated surface	Streptavidin mediated capture of biotinylated biomolecules	Fluorescence
Protometrix (Guilford, CT, USA)	Glass slide	Ni^{2+} coated surface for capture of His-tagged proteins	Fluorescence
BD Biosciences Clontech (Palo Alto, CA)	Glass slide	Antibodies	Fluorescence
Molecular Staging Inc. (Haven, CT)	Glass slide	Antibodies and rolling circle amplification	RCAT-based amplification of associated DNA tag

MS: Mass spectrometry; NHS: N-hydroxysuccinimide; SELDI: Surface enhanced laser desorption/ionization; SPR: Surface plasmon resonance.

(Hopkinton, MA, USA) have developed Grating-coupled surface plasmon resonance (GCSPR). This system utilizes a disposable plastic chip with a fine grating moulded onto the surface. The grating is coated with a very thin layer of gold that is further modified with appropriate surface chemistries. Gold is typically used because it does not oxidize like other metals, which can affect attachment.

Many commercial arrays are based on antibodies or antibody fragments (HTS Biosystems, Large Scale Biology, Biosite Diagnostics (San Diego, CA, USA), Zyomyx, Packard Bioscience, Oxford Glycosciences (Oxford, UK) and BD Clontech (Palo Alto, CA, USA)). The BD Clontech commercial protein array contains 378 monoclonal antibodies immobilized onto a glass surface, allowing for the comparison of the cor-

Highlights

- There have been primarily two approaches to characterize multiple proteins in biological samples – 2D-gels and protein microarrays. Protein arrays are comprised of a library of proteins immobilized on a chip in a 2D addressable grid format.
- Development of protein arrays is technically very difficult due to the structural diversity and complexity in proteins.
- Several major technical hurdles face protein array technology namely the acquisition, arraying, and stable attachment of proteins to array surfaces and detection of interacting proteins.
- Antibody arrays represent a popular format. The greatest challenge is obtaining antibodies against all the proteins that comprise the human proteome. At this time antibodies are available for a mere fraction of the proteome.
- Detection of bound target is considerably more complex than with nucleic acid arrays. Proteomes under comparison can be labeled with different fluorophores, however reproducibility of these chemical reactions is poor and interference with the protein–antibody interactions poses an additional complexity
- Arrays can be engineered to address protein identification, quantitation and affinity studies.
- Profiling disease-related tissues using protein microarrays will facilitate discovery of novel biomarkers and potential drug targets.
- Although protein arrays hold many promises, the greatest limitation to their utility is that there is not always a direct correlation between protein abundance and activity.

responding antigens in two discrete biological samples. This open array-platform architecture is compatible with most commercially available fluorescent scanners. The arrayed antibodies encompass a broad range of biological functions from signal transduction, cancer, cell cycle regulation, cell structure and apoptosis, to neurobiology. As an alternative to antibodies, libraries of fibronectin-based polypeptide scaffold molecules are utilized by Phylos and light sensitive photoaptamers by Somalogic. Aptamers are short stretches of nucleotides that can bind and covalently crosslink target proteins. This facilitates higher stringency wash conditions which promotes detection of specific binding.

Molecular Staging, Inc. (Haven, CT, USA) have developed a highly sensitive and efficient rolling circle amplification method (immuno-RCAT™) that aids detection of target molecules in a wide array of testing formats including microarrays. In immuno-RCAT, a unique DNA sequence tag is associated with a particular antibody via covalent linkage. Antibodies bound to antigen are measured by subsequent RCAT-based amplification of the associated DNA tag.

Akceli (Cambridge, MA, USA) utilize a novel reverse transfection method coupled to DNA microarrays. Plamid DNA molecules are resuspended in a gelatin transfection reagent mix and printed on glass slides. After drying, the slides are covered with a layer of the cells which spread across each slide, enabling clusters of cells to react with individual DNA spots, ingest the plasmid and express the corresponding protein.

Outlook

Technologies are being developed to enable large-scale parallel detection of proteins and facilitate comparative surveys of proteomes. Profiling disease-related tissues using protein microarrays will help uncover novel biomarkers and potential drug targets. Although these arrays promise manipulation on a proteome scale, the greatest limitation to their utility is that there is not always a direct correlation between protein abundance and activity. Many proteins, for example proteases, kinases and phosphatases are synthesized and secreted as inactive zymogens and converted to an active form by enzymatic cleavage. Additionally, the expression and analysis of proteins is carried out in artificial environments relative to the cell. These environments may not have physiological relevance to the biological system under study. This technology will generate vast amounts of data related to protein interactions, enzymatic activities and other biochemical properties. Effective analyses and mining of this data will require sophisticated bioinformatics tools. Protein microarrays nevertheless represent novel efforts to bridge the information gap that exists between genomics and proteomics.

Acknowledgements

The authors would like to thank Dr Darren Woodside and Dr David Calderwood of The Scripps Research Institute, La Jolla, CA, USA for critical reading of the manuscript and helpful discussions. GH is supported by the Howard Hughes Medical Institute.

Bibliography

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Brown PO, Botstein D: Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* 21(1 Suppl.), 33-37 (1999).
2. Debouck C, Goodfellow PN: DNA microarrays in drug discovery and development. *Nat. Genet.* 21(1 Suppl.), 48-50 (1999).
3. Chee M, Yang R, Hubbell E *et al.*: Accessing genetic information with high-density DNA arrays. *Science* 274, 610-614 (1996).
4. Hardiman G: Microarray technologies. *Pharmacogenomics* 3(3), 293-297 (2002).
5. Pollack JR, Perou CM, Alizadeh AA *et al.*: Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* 1, 41-46 (1999).
6. Iyer VR, Horak CE, Scafe CS, Botstein D, Snyder M, Brown PO: Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409, 533-538 (2001).
7. Gygi SP, Rochon Y, Franza B, Abersold R: Correlation between protein and mRNA abundance in yeast. *Mol. Cell Biol.* 19, 1720-1730 (1999).
8. Pandey A, Mann M: Proteomics to study genes and genomes. *Nature* (Insight) 405, 837-8469 (2000).
9. Mitchell P: A perspective on protein microarrays. *Nat .Biotech.* 20, 225-229 (2002).
10. Service RF: Searching for recipes for protein arrays. *Science* 294, 2080-2082 (2001).
11. Cahill D: Protein arrays: a high-throughput solution for proteomics research? *Proteomics*: *A Trends Guide* 47-51 (2000).
12. Figeys D: Array and lab on an array technology for protein characterization. *Curr. Opin. Mol. Ther.* 1(6), 685-694 (1999).
13. Sanders GHW, Manz A: Array-based microsystems for genomic and proteomic analysis. *Trends in Analytical Chemistry* 19, 364-378 (2000).
14. Zhu H, Bilgin M, Bangham R *et al.*: Global analysis of protein activities using proteome arrays. *Science* 293, 2101-2105 (2001).

•• **Elegant study utilizing yeast as an alternative to bacterial expression. The authors expressed and arrayed 5800 yeast ORFs as His tagged proteins.**

15. Gates SM: Surface chemistry in the chemical vapor deposition of electronic materials. *Chem. Rev.* 96, 1519-1532 (1996).
16. Afanassiev V, Hanemann V, Wolf S: Preparation of DNA and protein microarrays on glass slides coated with an agarose film. *Nucleic Acids Res.* 28 12, e66 (2002).
17. Lofas S, Johnsson B, Tegendahl K, Ronnberg I: Dextran modified gold surfaces for surface plasmon resonance sensors; immunoreactivity of immobilized antibodies and antibody-surface interaction studies. *J. Colloid Interface Sci.* 65, 423-431 (1993).
18. Haab BB, Dunham MJ, Brown PO: Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biology* 2, 0004.1-0004.13 (2001).

•• **Methodology for fabrication of protein arrays and utilization of a comparative fluorescence assay to measure the abundance of many specific proteins in complex solutions.**

19. MacBeath G, Schreiber S: Printing proteins as microarrays for high-throughput function determination. *Science* 289, 1760-1763 (2000).

•• **Methodology for fabrication of protein arrays and demonstration of three specific applications, screening for protein–protein interactions, identification of protein kinase substrates and identification of the protein targets of small molecules.**

20. Mecklenburg M: XNA on Gold: A versatile microarray platform. *Nat. Genet.* 23(Suppl.), 61-62 (1999).
21. Burgener M, Sanger M, Candrian U: Synthesis of a stable and specific surface plasmon resonance biosensor surface employing covalently immobilized peptide nucleic acids. *Bioconjug. Chem.* 11, 749-754 (2000).
22. Weinberger SR, Morris TS, Pawlak M: Recent trends in protein bioarray technology. *Pharmacogenomics* 1, 395-416 (2000).

•• **Comprehensive review of three protein biochip technologies, Biomolecular Interaction Analysis (Biacore), Surface Enhanced Laser Desorption/Ionization (SELDI) Ciphergen Protein Chip Arrays and Fluorescent Planar Wave Guide.**

23. Eisen MB, Brown PO: DNA arrays for analysis of gene expression. *Methods Enzymol.* 303, 179-205 (1999).
24. Nice EC, Catimel B: Instrumental biosensors: new perspectives for the analysis of biomolecular interactions. *Bioessays* 21, 339-352 (1999).
25. Fodor SPA, Read JL, Pirrung MC, Stryer L, Liu AT, Solas D: Light directed, spatially addressable parallel chemical synthesis. *Science* 251, 767-773 (1991).
26. Frank R: SPOT synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* 48, 9217-9232 (1992).
27. Walhout AJ, Temple GF, Brasch MA *et al.:* (2000) Gateway recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol.* 328, 575-592.
28. Bussow K, Nordhoff E, Lubbert C, Lehrach H, Walter G: A human cDNA library for high-throughput protein expression screening. *Genomics* 65, 1-8 (2002).

• **His epitope tagged recombinant proteins from a human fetal brain cDNA expression library were immobilized with anti-His capture antibodies.**

29. He M, Taussig MJ: Single step generation of protein arrays from DNA by cell-free expression and *in situ* immobilisation (PISA method). *Nucl Acids Res.* 29, e73 (2001).

•• **Elegant cell free system which permits rapid generation of protein microarrays directly from DNA templates via a rabbit reticulocyte expression system. Advantageous in that it allows reliable expression of toxic proteins.**

30. MacBeath G, Schreiber S: Printing proteins as microarrays for high-throughput function determination. *Science* 289 1760-1763 (2000).
31. Winter G, Griffiths AD, Hawkin RE, Hoogenboom HR: Making antibodies by phage display technology. *Ann. Rev. Immun.* 12, 433-455 (1994).
32. Wagner P, Nock S, Ault-Riche D, Itin C: Arrays of protein-capture agents and methods of use thereof. PCT Publication WO * 00/4389A2 (Zyomix, Inc., Hayward, CA, USA) (2002).
33. Watkins JD, Beuerlein G, Wu H, McFadden PR, Pancook JD, Huse WD: Discovery of human antibodies to cell surface antigens by capture lift screening of phage-expressed antibody libraries. *Anal. Biochem.* 256, 169-177 (1998).
34. Roepstorff P: Mass spectrometry in protein studies from genome to function. *Curr. Opin. Biotechnol.* 8, 6-13 (1997).
35. Schuck P: Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules. *Ann. Rev. Biophys. Biomol. Struct.* 26, 541-566 (1997).
36. Emili AQ, Cagney G: Large-scale functional analysis using peptide or protein arrays. *Nat. Biotechnol.* 18, 393-397 (2000).

37. Cagney G, Uetz P, Fields S: High-throughput screening for protein-protein interactions using two-hybrid assay. *Methods Enzymol.* 328, 3-14 (2000).

•• **First genome-scale 'living' protein array, which enabled systematic two-hybrid genetic screening of each yeast ORF for interacting proteins.**

38. Kraemer B, ZhangB, Sengupta D, Fields S, Wickens M: Using the yeast three hybrid system to detect and analyse RNA-protein interactions. *Methods Enzymol.* 328, 297-321 (2000).

39. Yuan OY, Valenzuela D, Rapeijko P, Hoffmann H, Higgins KA: Proteomic interaction arrays. PCT Publication WO046406A2. (Alphagene, Inc., Woburn, MA, USA) (2000).

40. Cai H, Wang Y, McCarthy D *et al.*: BACE2 is the major β-secretase for generation of β - amyloid peptides by neurons. *Nat. Neurosci.* 4, 233-234 (2001).

• **Application of Ciphergen protein chip to monitor the effects of various drugs on APP fragmentation patterns.**

41. Austen B, Davies H, Stephens DJ, Frears ER, Walters CE: The role of cholesterol in the biosynthesis of β-amyloid. *Neuroreport* 10, 1699-1705 (1999).

• **Application of Ciphergen protein chip containing an immobilized anti-β-amyloid antibody. This study demonstrated that cholesterol treatment reduced secretion of 1-40 and 1-42 forms of β-amyloid peptides from cultured cells.**

42. Petricoin EF, Ardekani AM, Hitt BA *et al.*: Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359, 9306, 572-577 (2002).

•• **Application of Ciphergen protein chips to analyse proteomic patterns in serum and facilitate the identification of ovarian cancer samples. This technology can be used as a screening tool for all stages of ovarian cancer in high risk and general populations.**

43. Nelson RW, Nedelkov D, Tubbs KA: Biosensor chip mass spectrometry: a chip based proteomics approach. *Electrophoresis* 21, 1155-1163 (2000).

44. Williams C, Addona TH: The integration of SPR biosensors with mass spectrometry: possible applications for proteome analysis. *Trends Biotechnol.* 18, 45-48 (2000).