Review

Virginia Espina[1]
Arpita I. Mehta[1, 3]
Mary E. Winters[1]
Valerie Calvert[2]
Julia Wulfkuhle[1]
Emanuel F. Petricoin III[2]
Lance A. Liotta[1]

[1]FDA-NCI Clinical Proteomics Program, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA
[2]FDA-NCI Clinical Proteomics Program, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, MD, USA
[3]Howard Hughes Medical Institute, Bethesda, MD, USA

Protein microarrays: Molecular profiling technologies for clinical specimens

Proteomics, the study of protein function within biologic systems, will further our understanding of cancer pathogenesis. Coupled with transcript profiling, proteomics can herald the advent of molecular therapy tailored to the individual patient's neoplasm. Protein microarrays, one emerging class of proteomic technologies, have broad applications for discovery and quantitative analysis. This technology is uniquely suited to gather information about the post-translational modifications of proteins reflecting the activity state of signal pathways and networks. Protein microarrays now make it feasible to conduct signal network profiling within cellular samples. Nevertheless, to be successful, design and use of protein microarrays must take into consideration enormous analytical challenges. A subclass of protein microarrays, Reverse Phase Arrays, created to meet these challenges, has been optimized for use with tissue specimens, and is now in use for the analysis of biopsy samples for clinical trial research.

Keywords: Catalyzed reporter deposition / Individual targeted therapy / Laser capture microdissection / Molecular profiling / Protein microarray / Review PRO 0592

Contents

Correspondence: Virginia Espina, National Cancer Institute, Building 10 Room B1B53, 9000 Rockville Pike, Bethesda, MD 20892, USA
E-mail: espinav@mail.nih.gov
Fax: +1-301-402-0043

Abbreviation: LCM, Laser capture microdissection

1 Introduction

A comprehensive analysis of the molecular basis of cancer and other disease states requires integration of the distinct, but complementary knowledge guided from the fields of genomics and proteomics. Genetic defects ultimately lead to tumor cell survival by altering the functional proteins that confer survival advantages for the tumor [1, 2]. Cancer growth, invasion and metastasis may be the result of dominations or perturbations in the normal network of cell signaling proteins [3]. Genomics and DNA microarrays point toward potential genetic defects that

may cause disruptions in cell signaling pathways. Protein microarrays, on the other hand, provide crucial information about the functional state of these disrupted pathways. Development of the cellular proteomic network for an individual patient's tumor will hopefully reveal potential drug targets and diagnostic molecules for disease prognosis or treatment [4–8] . This molecular integration provides a complete view of the disrupted cellular machinery governing disease.

Realigning concepts and techniques from genomics and applying them to proteomics has led to the development of protein microarrays. In the simplest sense, protein microarrays are immobilized protein spots [9–12]. The spots may be homogeneous or heterogeneous and may consist of a bait molecule, such as an antibody, a cell or phage lysate, a nucleic acid, drug or a recombinant protein or peptide [11, 13–21]. Detection of the array is achieved by probing with a tagged antibody, ligand or serum/cell lysate. The signal-generating tagging molecule generates a pattern of positive and negative spots. The signal intensity of each spot is proportional to the quantity of applied tagged molecules bound to the bait molecule [16].

The post-translational modifications of protein networks can be profiled employing protein microarrays by comparing the proportion of total (activated and non-activated) protein to the phosphorylated (activated) protein. This information in general reflects the state of information flow through a protein network. Monitoring the total and phosphorylated proteins over time, before and after treatment, or between disease and nondisease states may allow us to infer the activity levels of the proteins in a particular pathway in real time [22, 23]. The utility of protein microarrays lies in their ability to provide a map of known cell signaling proteins. Identification of critical nodes, or interactions, within the network is a potential starting point for drug development and/or design of individual therapy regimens [22, 23].

2 Protein microarray technology

Significant analytical challenges exist for protein microarrays that do not exist for gene arrays [6, 11, 13]. The initial challenge is developing a system capable of detecting a vast range of analyte concentrations. Proteins of interest may exist in a broad dynamic range (up to a factor of 10^{10}) in any cell. The second challenge is identifying query molecules with adequate sensitivity and specificity for detecting low abundance proteins. Low abundance proteins exist in a complex biologic milieu containing high abundant contaminating proteins. A system with 99% detection specificity may exhibit unacceptable signal to background ratios if the there are cross-reacting contaminating molecules. Systems with femtomolar range sensitivity are required for low abundant cell signaling proteins [24, 25].The third challenge is the ability to block endogenous molecules such as peroxidases, biotin, immunoglobulins, alkaline phosphatases, fluorescent proteins and avidin. These molecules may interfere with the amplification chemistries [26–28].

2.1 Array formats

There are currently two approaches to producing arrays capable of generating this type of network information. The first format is the forward phase array, in which a labeled bait molecule is immobilized on the substratum. The bait molecule is typically an antibody, with each spot representing only one type of bait molecule. The array is incubated with only one test sample that contains several different analytes of interest. The captured analytes are detected with a second tagged molecule or by labeling the analyte directly.

Reverse phase arrays consist of immobilized analyte molecules, with multiple analytes in each spot (Fig. 1). Each spot represents an individual test sample, allowing an array to be comprised of multiple, different samples. The reverse phase array is probed with a single detection molecule and a single analyte is measured for each spot on the array, across multiple samples. This format allows multiple samples to be analyzed under the same experimental conditions for any given analyte. The arrays are probed separately with two different classes of antibodies to specifically detect the total and phosphorylated forms of the protein of interest. Comparison of the proportion of phosphorylated proteins across patients on the same array provides insights into the cellular signaling network for individual patients. Because nitrocellulose can not be stripped and reprobed, each slide array is probed with a different antibody, generating a set of array slides for each set of probe antibodies [29]. Construction of reverse phase microarrays follows these general protocols: sample collection, sample preservation, preparation of frozen tissue sections, isolation of pure cell populations, protein extraction, microarray printing and finally protein detection. We will discuss the construction and detection of reverse phase arrays for cellular proteome profiling.

2.2 Microarray applications

Protein microarrays may be used for drug discovery, biomarker identification and molecular profiling of cellular material. Technologies such as MS, 2-DE, bead capture and micro-ELISA are currently successfully employed

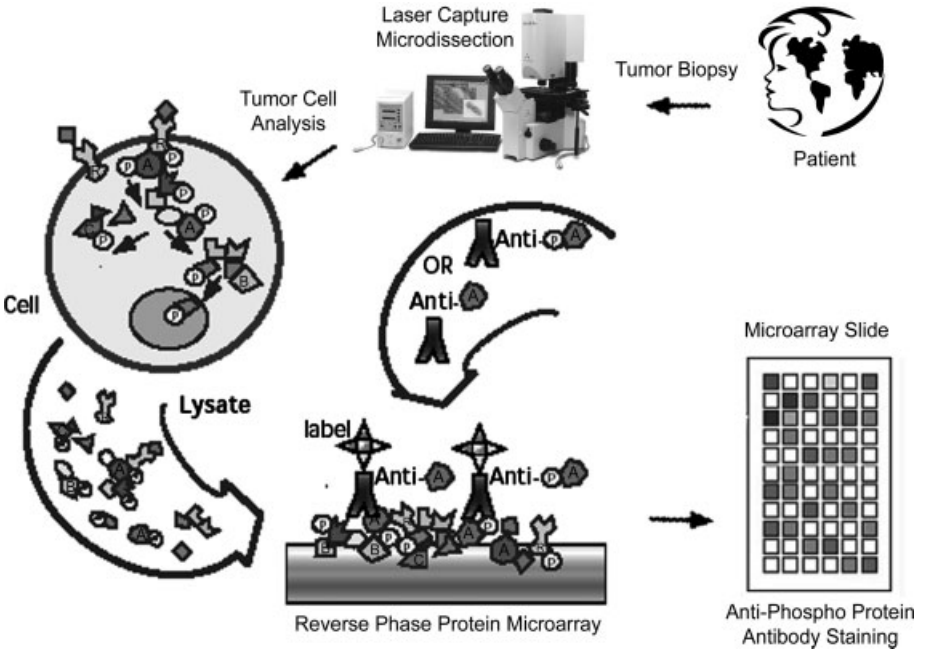

Figure 1. Molecular profiling protocols for patient biopsy samples have been integrated into IRB approved clinical trials at the National Cancer Institute. The protocols focus on biopsy collection and preservation, microdissection of pure cell populations, and reverse phase protein microarrays. After tissue microdissection, the isolated cell populations are lysed and the entire cellular proteome is immobilized on a glass-backed nitrocellulose slide. The immobilized proteins represent total and phosphorylated forms of cell signaling proteins. The proteins of interest are probed with antibodies directed against the total and phosphorylated forms of the protein. After signal amplification, development and imaging, a pattern or circuit diagram of cell signaling pathways is revealed. This pattern can be compared across patient samples on the array.

for drug discovery and biomarker identification [6, 30–32]. The utility of protein microarrays lies in their ability to develop molecular network profiles of cellular samples [16, 23, 33]. The benefit of reverse phase microarrays is their high throughput capabilities using low sample volumes. Typical reverse phase microarrays for patient biopsy material are printed with 30–60 uL of whole cell lysate, yielding 50 or more microarrays. These microarrays may be stored and probed at a later date or as additional proteins of interest in the signaling network are identified, or as additional antibody probes are developed. These benefits are perfectly matched for molecular profiling of clinical patient samples. Frequently only a small amount of patient material is available for molecular analysis. The protein microarray provides a means of assaying small volumes of tumor cells.

2.3 Cellular samples from tissue

Protein microarray technology can be applied to biopsies, tissue cell aspirates or body fluid specimens [16, 33–35]. The amount of cellular material may be small in volume and low in analyte concentration. The total number of cells required for the array is dependent on the number of analyte molecules, the sensitivity of the system and the number of molecules/mole. This may be expressed by the formula:

$$T = \frac{(A^{*}s)}{X} \quad (1)$$

Where T is the threshold for cell procurement (cells/volume); A is Avogadro's number (6.023×10^{23} molecules/mole); s is the sensitivity of the detection system (moles/volume); and X is the number of analyte molecules per cell (molecules/cell).

Core needle biopsies, obtained with a 16 or 18 gauge needle, typically provide a 2 mm $\times$ 5 mm core of tissue. This is usually a sufficient tissue volume for adequate cell procurement of 2000–8000 cells.

2.4 Sample preservation

Tissue preservation techniques, such as formalin fixation and paraffin embedding, lead to cross-linking of protein molecules, and consequently render the proteins unfit for analysis by protein microarray techniques [36–38]. Snap freezing in liquid nitrogen or prompt embedding in Optimal Cutting Temperature medium (Sakura FineTek,

Torrance, CA, USA) with storage at −80°C is recommended for reduction in protein degradation. Additional preservation steps may be undertaken such as adding protease and/or phosphatase inhibitors to the tissue staining baths and/or the extraction buffer [39].

2.5 Laser capture microdissection (LCM)

Cellular signaling networks are dependent on the tissue's cell type and microenvironment, demanding isolation of pure cell populations for accurate network profiling (Fig. 1) [3]. Microdissection is an established technique for isolating pure cell populations from heterogeneous tissue [37, 40]. LCM performed with the PixCell II or AutoPix instruments from Arcturus Engineering (Mountain View, CA, USA) provides cellular material for array construction from Hematoxylin and Eosin (H&E) stained frozen tissue sections. Depending on the cellularity of the sample, multiple tissue sections may need to be microdissected to generate an adequate number of cells for analysis. Typical array construction using microdissected cells requires ideally a minimum of 5000 cells/10 uL of lysis buffer.

2.6 Frozen tissue section preparation

Frozen tissue sections, cut at 5–10 μm thickness are readily amenable to LCM [36]. H&E stained tissue sections are prepared following standard tissue staining protocols. Other tissue stains capable of providing nuclear and cytoplasmic details may be substituted for the H&E stain. Successful microdissection hinges on complete dehydration of the tissue section with an alcohol gradient and a final clearing solution of xylene or xylene substitute (Sub-X, SurgiPath, Richmond, IL, USA). Minimum staining and fixation times, for whatever staining method is used, are recommended for maximizing the ease of microdissection and limiting sample exposure to potential chemical reactants.

2.7 Validation of antibodies

Successful detection of proteins by microarray requires specific, high affinity antibodies and ligands [14]. DNA probes have been manufactured with known specific affinity constants. On the other hand, antibodies, aptamers and ligands utilized for protein microarrays can not be manufactured with predictable affinity or specificity. This is the limiting factor for the successful application of microarray technology [14]. This necessitates the validation of antibodies by Western blot, with a sample similar to that used on the array. Each antibody should reveal a single, specific band at the expected molecular weight, validating its specificity.

2.8 Protein extraction

Molecular profiling for elucidation of cell signaling networks requires liberation of the entire cellular proteome. Previously unidentified cell signaling partners could potentially be overlooked if a subcellular compartment was excluded from the extraction protocol. Lysis buffer consisting of equal volumes of T-PER® (Pierce Chemical, Rockford, IL, USA), and a 2–4% solution of β-mercaptoethanol (Sigma, St. Louis, MO, USA) in 2X Tris-glycine-SDS sample buffer (Invitrogen, Leek, The Netherlands) has been used successfully for protein extraction of cells procured by LCM. This procedure entails extracting proteins for 30 min to 2 h at 70°C, with vortexing before and after incubation. Samples are spun after incubation in a microcentrifuge at 14 000 rpm for 1–2 min to pellet any cellular debris. The supernatant extract is used for printing the microarray. Lysates may be stored at −80°C prior to microarray printing.

2.9.1 Arraying technology and substratum

Protein microarray printing formats are printed using the same technology as DNA microarrays, but the protein array layout is vastly different than a typical DNA array. Both printing technologies transfer sample fluid from a microtiter plate onto a high surface area substratum, usually a coated glass slide. DNA arrays are generally printed as high-density spots of the same sample and a single array is probed with multiple antibodies. The substratum for DNA arrays is typically amine or lysine coated glass, permitting adherence of the negatively charged DNA to the positively charged coating [41]. Protein array substrata may be nylon, nitrocellulose or silanized silica [29, 42, 43]. The substratum requirements for protein arrays are 1) high binding capacity 2) it should not alter the protein structure and 3) low background. Nitrocellulose coated glass slides are a common substratum for protein arrays (FAST slides, Schleicher & Schuell BioSciences, Dassel, Germany).

Proteins bind to nitrocellulose *via* electrostatic interactions in an irreversible manner, limiting the number of probes that can be used with any one set of immobilized proteins [29, 42]. The nitrocellulose slide format allows multiple slides to be printed for each set of samples, thus maximizing the number of different probes (antibodies) that may be used for analyzing a given set of arrays

(Fig. 1). The nitrocellulose polymer coating of FAST slides permits protein binding capacities of 75–150 μg/cm^2 in a volume of 0.3–2 nL/spot, meeting the criteria for a high binding capacity substratum. Chromogenic, fluorometric and luminescent detection methods may be used with FAST slides with an adequate signal/noise ratio. Microarray printing needs to be reproducible, high capacity and automated for high-throughput testing. The pinhead, and type of printing assembly, is key to successful protein array printing. Sample volume, viscosity, number of arrays required and substratum are parameters to be considered prior to selecting a printing system. Protein array formats consist of multiple patients on a single array, often representing samples obtained before and after treatment or normal and malignant tissue lysates, allowing comparison of analytes across samples on an array. Each array is probed, or stained, with a single antibody of interest. Protein arrays may also be printed in sector formats. A sector array consists of multiple small pads of substratum on a slide. A reservoir placed around each sector permits a different antibody to be used for probing the samples. The sector format miniaturizes the array, providing an increased signal/noise ratio.

2.9.2 Arraying devices

Microarrays may be printed in a variety of configurations using several different types of printing methods [44]. Each method has inherent advantages and disadvantages based on the type of sample to be printed, sample volume, capacity of array and production capacity required for array analysis. Printing technology currently exists in two forms: contact and noncontact devices. Contact printing is accomplished by direct contact between the print head and the substratum. Noncontact printing dispenses a minute volume of sample above the substratum. Examples of contact printing formats are solid pin, quill, and pin and ring assemblies. Noncontact printing technology utilizes piezoelectric or syringe solenoid devices (Table 1).

Table 1. Comparison of array printing technologies

	Contact printing	Noncontact printing
Technology	Direct contact with substratum	Fluid contact only
Types	Quill, ink jet, pin and ring	Piezoelectric, solenoid
Determination of spot size	Pin diameter	Droplet volume
Sample delivery volume	0.3–2.0 nL	0.1–0.3 nL piezoelectric 4.0–8.0 nL solenoid

The maximum print rate/spot is integral for high throughput capacity. Most printing assemblies are able to print 1–2 spots/second. The spot resolution is a function of the XYZ positioning capabilities of the print head assembly. Repeatability, accuracy and resolution of the print assembly determine the final quality of the array printing [44].

2.9.3 Contact printing devices

Contact pin arrayers rely on a solid pin printing assembly, composed of a solid pin, with a flat end, for contact with the substratum and transfer of fluid. Submerging the pin in a liquid sample transfers a small volume of sample to the tip of the pin. Contact of the pin with the substratum transfers the fluid to the substratum. The pin diameter and fluid properties determine the volume of fluid deposited, and thus the spot size. The solid pin format does not lend itself to duplicate or triplicate printing due to the necessity of the pin to be re-submerged in the sample for each spot printing [44].

A quill type printing assembly consists of a flat pinhead with a defined hollow bore, similar to quill-style writing instruments (MicroGrid II; Biorobotics, Cambridge, UK). Sample fluid wicks into the hollow space and is deposited on the substratum when the pinhead touches the substratum. The quill style formats allow multiple spot printing from each sample.

Pin and ring assemblies are a combination of a ring that holds microliter quantities of sample and a flat head pin (GMS417, Affymetrix/MWG Biotech). The pin travels through the fluid retained in the ring and deposits the sample on the substratum. The pin and ring assembly is capable of replicate spot printing. Typical sample delivery volumes of contact printing devices are 0.3–2 nL. Pin sizes are available in 125, 300 or 500 μm diameters.

Ink jet style printing devices are similar in theory to ink jet paper printers [44]. A small volume of fluid is transferred to a hollow bore pin. Touching the pin to the substratum causes wicking of a minute sample volume from the pin onto the substratum due to the attractive forces between the substratum and the fluid. Pins with various tip diameters are available for printing spots from 75–360 μm in diameter.

2.9.4 Noncontact printing devices

Noncontact printing devices utilize a sensor for depositing fluid above the substratum (BioChip; Perkin Elmer, Norwalk, CT, USA). This sensor may be a piezoelectric crystal or a solenoid. Piezoelectric devices consist of a

glass capillary tube surrounded by a deformable piezoelectric material [44]. Piezoelectric material is typically a ceramic that changes form in the presence of an electrical charge. The deformation of the piezoelectric material provides pressure on the glass capillary containing the sample, causing fluid to be dispensed from the tip of the glass capillary. Picoliter quantities of fluid may be dispensed with a piezoelectric tip. Typical sample delivery volumes for these devices are 0.1–0.3 nL.

Syringe solenoid systems utilize pressure supplied by a syringe to aspirate fluid into a sample tip. Opening the solenoid valve allows droplets of fluid to be ejected from the tip. The dispensed droplet volume is 4–8 nL.

2.9.5 Factors affecting success of microarray printing

Temperature, humidity and dust/debris affect array printing quality. Increased temperature and decreased humidity may cause evaporation of the sample from the microtiter plate. On the other hand, humidity greater than 75% may cause water vapor condensation on the metal parts of the printing device or the substratum [44]. Dust on the substratum causes interference with fluorescent detection methods. Debris on the printing surface causes the substratum surface to be uneven, resulting in missing spots or damaged substratum due to the pins compressing the surface.

The substratum surface should be inspected for obvious defects prior to printing. Holes, bubbles, and varying thickness of the substratum affect printing and detection quality.

Pin washing is a critical step for reproducible printing. A series of wash solutions and suction or vacuum drying steps prevents unintentional sample dilution from the wash solutions and prevents damage to the nitrocellulose. Wash solutions of deionized water and 70% v/v ethanol in water are adequate for removal of protein samples from the pins between each sample. Pin drying times of 10–20 seconds are recommended for adequate drying of the pins to prevent dilution effects. A word of caution if using ethanol solutions in the wash baths – ethanol solutions will damage the nitrocellulose surface, preventing successful array printing.

A pin calibration sets the level of contact between the substratum and the pin for contact style printing devises. A pin calibration is recommended for each lot number or series of slides to be printed. Too little contact with the slide surface results in missing spots. Too much contact compresses the nitrocellulose, resulting in less than expected protein binding capacity.

Ring calibration, of pin and ring style assemblies, ensures adequate sample loading on the ring for maximal array capacity. Under filling of the ring limits the total number of spots and/or arrays that may be printed. Overextension of the pin and ring assembly into the sample causes wicking of the sample along the ring support, leading to poor reproducibility.

2.9.6 Quality control and array design

Incorporation of a recombinant protein antigen and a reference lysate or peptide of known performance enhance the quality control aspects and potential clinical utility of each array. The recombinant protein antigen is serially diluted in the same manner as test samples, providing a quantitative comparison of signal intensity for each antibody. The reference or control lysate typically consists of a pool of tissue and cell culture protein lysates of known origin and staining characteristics. The control lysate serves as a standard for validating the performance characteristics of new lot numbers of reagents and materials used in the microarray assay.

A typical array layout consists of a series of 4–6 horizontal spots, representing a mini-dilution curve of each sample (Fig. 2). Multiple samples may be printed on each slide, representing before and after treatment samples or normal and malignant tissue lysates. Each sample is printed in triplicate, permitting statistical calculations for each sample. The use of multiple sample dilutions ensures the protein of interest is within the dynamic range of the assay, based on the antibody sensitivity and affinity.

2.10 Detection strategies

The reverse phase protein microarray format has been optimized for detection of cell signaling proteins in clinical cellular material [45]. Each spot on an array is a few hundred microns in diameter, containing a representative sample of the cellular proteome. Detection strategies for these minute samples require the use of amplification strategies [24, 25, 46–48]. Signal amplification is a prerequisite for achieving the sensitivity required for analysis of low abundance proteins. A reliable method capitalizes on the catalyzed reporter desposition technologies developed for clinical immunoassays [26–28, 46, 49]. This technology is based on the enzyme mediated deposition of biotin-tyramide conjugates at the site of a biotinylated antibody-ligand complex (CSA kit; DakoCytmation, Carpinteria, CA, USA).

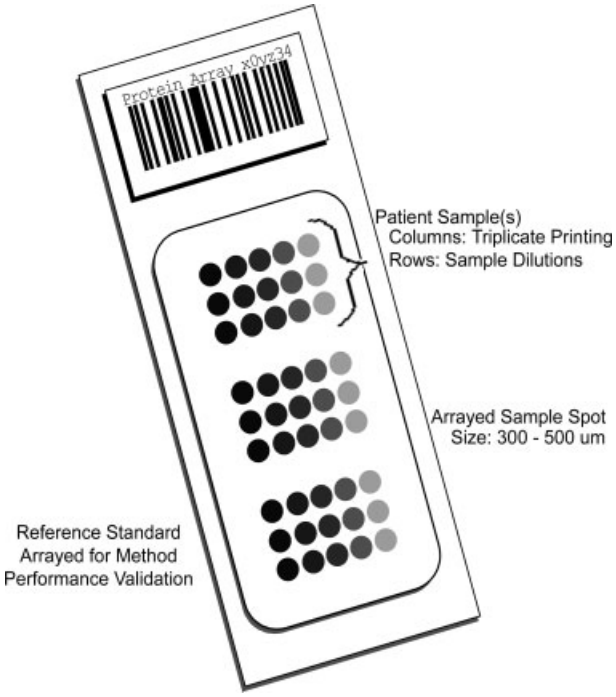

Figure 2. A typical protein microarray used for molecular profiling of patient biopsy samples consists of patient samples printed in triplicate and a reference lysate or control sample printed on a nitrocellulose coated glass slide. Each patient sample represents a before or after treatment specimen, or microdissected normal, stromal or malignant tissue cells. Each sample is printed in a dilution curve representing neat, 1:2, 1:4, 1:8, and 1:16 dilutions. A negative control, consisting of extraction buffer, is printed as the sixth spot. The reference and control lysates are used for monitoring assay performance. The spot size is dependent on the arraying device, while the density of spots may be operator controlled.

2.11.1 Bioinformatic analysis of reverse phase microarrays

There are currently a variety of methods for microarray analysis [50–54]. Reverse phase protein microarray analysis requires consideration of 1) spot placement on the array 2) background intensity and 3) a sufficient number of data points. We have developed an analysis technique for reverse phase microarrays using the open-source program, PCSAN [51] to address these considerations.

The spot placement on the array is determined to some extent by the pin configuration of the arrayer. The placement of serial dilutions is operator controlled and may be in a vertical or horizontal format. Spot intensity is integrated over a fixed area. This area is fixed for each spot, minimizing intensity variation due to variations in the integration area.

Background intensity is determined by programming "empty" spots on the array layout. These are areas lacking sample, but the nitrocellulose will have been exposed to all blocking and detection reagents in the same manner as the experimental areas. The density of each background area is integrated in the same manner as the sample spot. Ideally the signal should be 2 SD above background.

A sufficient number of data points is achieved by analyzing each spot in the dilution curve and printing each sample in duplicate or triplicate (Fig. 2).

2.11.2 Normalizing spot intensity

Total protein concentration variability exists between samples on the array due to the variability in the number of cells/sample obtained by LCM. This variability is controlled for each spot on the array by normalization of the total protein/spot. Total protein on a reverse phase array may be obtained by using a Sypro Ruby Blot staining protocol (Molecular Probes, Eugene, OR, USA). Spot intensities on the array are reduced to a single value corresponding to the amount of protein in the sample [51].

The intensity versus log dilution plot should appear as an exponential decay curve (see section 2.11.3). The assumption is made that intensity above background is proportional to protein concentration. Curve fitting techniques may be applied for determining overall curve fit, background intensity and decay constants.

2.11.3 Method for microarray analysis

Each array is scanned, spot intensity analyzed, data is normalized, and a standardized, single data value is generated for each sample on the array. This single data point may then be used for comparison to every other spot on the array. This data set may be used for Bayesian clustering analysis and generation of traditional "heat maps" for generation of network profiles across patient samples [45, 51].

Step 1. PSCAN

Images are analyzed with PSCAN yielding a raw intensity for each spot on each array. Special purpose MATLAB scripts are then used to combine the intensity data for each array with annotations. The data are then imported into the JMP statistical package (www.JMP.com, www.mathworks.com) for further analysis and visualization.

Step 2. MATLAB

Each sample appears in several dilutions on the array. A binding score for each sample is determined with the following formula, where P1 represents the neat spot and P2, P3, *etc.* represent the dilution spots, and b is a bias term set to 50:

$$\text{Score} = (P1 - P3)/[(P2 + P3)/2 + b] \quad (2)$$

This score attempts to reconstruct the specific binding (total minus non-specific) and expresses the result relative to the non-specific binding in that assay. The bias term increases the statistical readability of the score by preventing very small denominator values.

Step 3. JMP

Empirically, the score shows stability across different arrays, and is largely resistant to variations in overall slide intensity due to variations in staining efficiency, *etc.* If trends are observed across the slide a corrected score may be calculated by subtracting the local-average score for left, middle, and right portions of each array from the score as follows:

$$\text{Corrected score} = \text{score} - \text{local average score} \quad (3)$$

Step 4. Standardization of score

For some antibodies, the spread of the corrected score values for the "empty" (background) spots may overlap that of the real samples, while in others, the real sample scores may be markedly above those of the empty spots. These empty spots may serve as an indicator of precision for each antibody stained array. To combine scores from different antibodies, the standard deviation of the empty spots may be used to standardize the corrected score as follows:

$$\text{Standardized score} = \text{corrected score} - [\text{average(corrected score)}/\text{SD(empty corrected scores)}] \quad (4)$$

Step 5. Single Data Matrix

The results of the each antibody labeled array in standardized score units is combined into a single data matrix. The maximum standard score for each array can be interpreted as a signal-to-noise ratio.

Step 6. Two-way hierarchical clustering

Ward's method may be used for generation of clustering data by samples (rows) and antibodies (columns) [57]. Color coding according to the standardized score allow data to be displayed as traditional heat maps.

Step 7. Discriminant analysis

Discrimination of tumor from true normal samples may be performed with Fisher's discriminant analysis [58].

3 Microarray performance characteristics

The specificity and sensitivity of protein microarrays can be analyzed using the ambient analyte assay theory of Ekins and coworkers [55, 56]. This theory demonstrates the analysis of a small amount of target in a small area is more sensitive than the same amount of target in a larger area (Fig. 3). The sensitivity of the system is increased because the binding reaction occurs at the highest possible concentration and the bait-detection complex is localized in the microarray spot. This occurs because the target molecule is not the limiting factor in a small spot. Signal intensity will remain virtually constant below a certain spot size even if there is a further reduction in spot size. As the spot size increases, for a given amount of target, the signal intensity decreases.

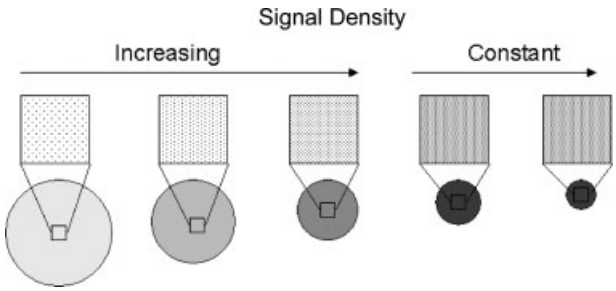

Figure 3. As the arrayed spot size increases for a given analyte concentration, the density of the spot reaches a maximum. In other words for a given concentration of analyte occupying ever increasing spot diameters, the intensity (density) of the spot will decrease as the spot diameter increases. This is due to the same number of target molecules occupying an increasingly larger area. In contrast, as the spot size decreases, the density increases for a decreasing concentration of analyte. There is a point at which the spot intensity will remain constant for decreasing spot size. Incorporation of a sample dilution curve into the printed array format permits the analysis of each analyte in its linear dynamic range without limitations due to the unmatched antibody association/dissociation constants. This figure was adapted from [14].

3.1 Sensitivity and specificity of protein microarrays

Reverse phase protein arrays exhibit excellent within-run and between-run precision with greater than 6 cell equivalents (3–10%) using a pin and ring style arrayer [16]. Incorporating a miniature dilution curve in the array

design monitors linearity for each sample (Fig. 2). Each sample is prepared in a series of dilutions and each dilution is printed on the array. This allows matching of the antibody concentration with the analyte concentration, permitting analysis of each analyte in a linear region.

4 Concluding remarks

The current limitation of the protein microarray is the lack of availability of antibodies with high affinity and specificity for posttranslational modified proteins and disease related gene products. Fortunately this issue is being addressed by the Human Proteome Organization as well as the Alliance for Cellular Signaling (www.signaling-gateway.org).

Future goals for the enhancement of the microarray technology would be 1) to increase the concentration of protein loaded *per* spot by increasing the surface area and concentrating the input sample and 2) to reduce the substratum variability from lot to lot and across the slide. These enhancements can impact sensitivity and reproducibility.

Molecular profiling of clinical samples has the potential to change the face of prognosis and diagnosis of cancer. The utility of reverse phase protein arrays for protein network profiling has been demonstrated for serum and tissue samples [35, 53]. Incorporation of common molecular biology and clinical laboratory techniques in the reverse phase microarray protocols has made the technology widely available and reproducible. On-going clinical trials at the National Cancer Institute have currently incorporated reverse phase microarray technology for evaluation of the sample procurement and processing technologies, as well as a means to elucidate hypotheses concerning treatment effects on cell signaling networks. As more cell signaling information is gleaned from protein microarrays, we envision an enhanced ability to develop/identify new combinatorial therapies or targeted therapies. The ability to discern treatment effectiveness, *via* protein microarrays, at an early time point during treatment can lead to early intervention and positive patient outcomes.

The authors wish to thank John Milia of Arcturus Engineering, Inc. for helpful, consistent service and advice. AIM is an NIH-Howard Hughes Research Scholar.

Received February 23, 2002

5 References

[1] Hanahan, D., Weinberg, R. A., *Cell.* 2000, *1*, 57–70.

[2] Hunter, T., *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 1998, *1368*, 583–605.

[3] Liotta, L. A., Kohn, E. C., *Nature* 2001, *6835*, 375–379.

[4] Blume-Jensen, P., Hunter, T., *Nature* 2001, *6835*, 355–365.

[5] Bowden, E. T., Barth, M., Thomas, D., Glazer, R. I. *et al.*, *Oncogene* 1999, *31*, 4440–4449.

[6] Celis, J.E., Gromov, P., *Cancer Cell.* 2003, *1*, 9–15.

[7] Hunter, T., *Cell.* 2000, *1*, 113–127.

[8] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. *et al.*, *Nature* 2000, *6804*, 651–654.

[9] Liotta, L., Petricoin, E., *Nat. Rev. Genet.* 2000, *1*, 48–56.

[10] MacBeath, G., *Nat. Genet.* 2002, *32*, 526–532.

[11] Zhu, H., Snyder, M., *Curr. Opin. Chem. Biol.* 2003, *1*, 55–63.

[12] Haab, B. B., Dunham, M. J., Brown, P. O., *Genome Biol.* 2001, *2*, RESEARCH0004.

[13] Lal, S. P., Christopherson, R. I., dos Remedios, C. G., *Drug Discov. Today* 2002, *18 Suppl*, S143–9.

[14] Templin, M. F., Stoll, D., Schrenk, M., Traub, P. C. *et al.*, *Trends Biotechnol.* 2002, *4*, 160–166.

[15] Wilson, D.S., Nock, S., *Angew. Chem. Int. Ed. Engl.* 2003, *5*, 494–500.

[16] Paweletz, C. P., Charboneau, L., Bichsel, V. E., Simone, N. L. *et al.*, *Oncogene* 2001, *16*, 1981–1989.

[17] MacBeath, G., Schreiber, S. L., *Science. 2000, 5485, 1760–1763.*

[18] Humphery-Smith, I., Wischerhoff, E., Hashimoto, R., *Drug Discovery World* 2002, *1*, 17–27.

[19] Petach, H., Gold, L., *Curr. Opin. Biotechnol.* 2002, 309–314.

[20] Schaeferling, M., Schiller, S., Paul, H., Kruschina, M. *et al.*, *Electrophoresis* 2002, *18*, 3097–3105.

[21] Weng, S., Gu, K., Hammond, P. W., Lohse, P. *et al.*, *Proteomics* 2002, *1*, 48–57.

[22] Liotta, L.A., Kohn, E. C., Petricoin, E. F., *Jama* 2001, *18*, 2211–2214.

[23] Petricoin, E.F., Zoon, K. C., Kohn, E. C., Barrett, J. C. *et al.*, *Nat. Rev. Drug Discov.* 2002, *9*, 683–695.

[24] Morozov, V. N., Gavryushkin, A. V., Deev, A. A., *J. Biochem. Biophys. Methods* 2002, *1*, 57–67.

[25] Schweitzer, B., Roberts, S., Grimwade, B., Shao, W. *et al.*, *Nat. Biotechnol.* 2002, *4*, 359–365.

[26] Hunyady, B., Krempels, K., Harta, G., Mezey, E., *J. Histochem. Cytochem.* 1996, *12*, 1353–1362.

[27] Bobrow, M. N., Harris, T. D., Shaughnessy, K. J., Litt, G. J., *J. Immunol. Methods* 1989, *1–2*, 279–285.

[28] Bobrow, M. N., Shaughnessy, K. J., Litt, G. J., *J. Immunol. Methods* 1991, *1*, 103–112.

[29] Tonkinson, J. L., Stillman, B. A., *Front. Biosci.* 2002, c1–12.

[30] Chen, G., Gharib, T. G., Huang, C. C., Thomas, D. G. *et al.*, *Clin. Cancer Res.* 2002, *7*, 2298–2305.

[31] Delehanty, J. B., Ligler, F. S., *Anal. Chem.* 2002, *21*, 5681–5687.

[32] Pandey, A., Mann, M., *Nature* 2000, *6788*, 837–846.

[33] Charboneau, L., *Briefings in Functional Genomics and Proteomics* 2002, *3*, 305–315.

[34] Jessani, N., Liu, Y., Humphrey, M., Cravatt, B. F., *Proc. Natl. Acad. Sci. USA.* 2002, *16*, 10335–10340.

[35] Knezevic, V., Leethanakul, C., Bichsel, V. E., Worth, J. M. *et al.*, *Proteomics* 2001, *10*, 1271–1278.

[36] Gillespie, J. W., Best, C. J., Bichsel, V. E., Cole, K. A. *et al.*, *Am. J. Pathol.* 2002, *2*, 449–457.

[37] Emmert-Buck, M. R., Strausberg, R. L., Krizman, D. B., Bonaldo, M. F. *et al.*, *J. Mol. Diagn.* 2000, *2*, 60–66.

[38] Ahram, M., Flaig, M. J., Gillespie, J. W., Duray, P. H. *et al.*, *Proteomics* 2003, *4*, 413–421.

[39] Simone, N. L., Remaley, A. T., Charboneau, L., Petricoin, E. F. 3rd, *et al.*, *Am. J. Pathol.* 2000, *2*, 445–452.

[40] Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F. *et al.*, *Science* 1996, *5289*, 998–1001.

[41] Miller, L. D., Long, P. M., Wong, L., Mukherjee, S. *et al.*, *Cancer Cell* 2002, *5*, 353–361.

[42] Stillman, B. A., Tonkinson, J. L., *Biotechniques* 2000, *3*, 630–635.

[43] Nashat, A. H., Moronne, M., Ferrari, M., *Biotechnol. Bioeng.* 1998, *2*, 137–146.

[44] Schena, M., Microarray Biochip Technology, Eaton, Natick 2000.

[45] Liotta, L. A., Espina, V., Mehta, A. I., Calvert, V. *et al.*, *Cancer Cell* 2003, *4*, 317–325.

[46] King, G., Payne, S., Walker, F., Murray, G. I., *J. Pathol.* 1997, *2*, 237–241.

[47] Kukar, T., Eckenrode, S., Gu, Y., Lian, W. *et al.*, *Anal. Biochem.* 2002, *1*, 50–54.

[48] Wiese, R., *Luminescence* 2003, *1*, 25–30.

[49] Graf, R., Friedl, P., *Anal. Biochem.* 1999, *2*, 291–297.

[50] Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, A. *et al.*, *Nat. Genet.* 2001, *4*, 365–371.

[51] Carlisle, A. J., Prabhu, V. V., Elkahloun, A., Hudson, J. *et al.*, *Mol. Carcinog.* 2000, *1*, 12–22.

[52] Cutler, P., *Proteomics* 2003, *1*, 3–18.

[53] Miller, J. C., Zhou, H., Kwekel, J., Cavallo, R. *et al.*, *Proteomics* 2003, *1*, 56–63.

[54] Sreekumar, A., Nyati, M. K., Varambally, S., Barrette, T. R. *et al.*, *Cancer Res.* 2001, *20*, 7585–7593.

[55] Ekins, R., Chu, F., Biggart, E., *Clin. Chim. Acta.* 1990, *1*, 91–114.

[56] Ekins, R. P., *J. Pharm. Biomed. Anal.* 1989, *2*, 155–168.

[57] Holland, J. H. (Ed.), *Adaption in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence,* 3rd Edition, MIT, Cambridge 1994.

[58] Chernick, M. R. and Friis, R. H., *Introductory Biostatistics for the Health Sciences,* Wiley, Hoboken 2003, pp. 204–210.