



Protein microarrays as a discovery tool for studying protein–protein interactions

Ulrike Korf[†] and Stefan Wiemann

CONTENTS

Classification of protein microarrays

Fabrication of protein microarrays

Generating the content for protein–function microarrays

Remaining challenges

Detection of protein–protein interactions, protein–small molecule interactions & protein phosphorylation

Summary & conclusions

Expert opinion

Five-year view

Key issues

References

Affiliations

[†]Author for correspondence
DKFZ Heidelberg,
Division of Molecular Genome
Analysis (B050), Im Neuenheimer
Feld 580, 69120 Heidelberg,
Germany
Tel.: +49 622 142 4765
Fax: +49 622 142 3454
U.Korf@DKFZ.de

KEYWORDS:
cDNA library, interaction,
protein chip, protein domain,
protein microarray,
protein–protein,
recombinant protein

Exploring the function of the genome and the encoded proteins has emerged as a new and exciting challenge in the postgenomic era. Novel technologies come into view that promise to be valuable for the investigation not only of single proteins, but of entire protein networks. Protein microarrays are the innovative assay platform for highly parallel *in vitro* studies of protein–protein interactions. Due to their flexibility and multiplexing capacity, protein microarrays benefit basic research, diagnosis and biomedicine. This review provides an overview on the basic principles of protein microarrays and their potential to multiplex protein–protein interaction studies.

Expert Rev. Proteomics 2(1), 13–26 (2005)

Biologic systems are complex, and regulation via modular and frequently crosstalking signaling systems is a common scheme for their organization [1–3]. Historically, biochemical analysis has been restricted to the examination of one or a few proteins at a time. Studying only a minority of components would negate influence of other components of the signaling network. Experimental approaches considering the complex character of cellular signal transduction networks are therefore mandatory for comprehensive basic biology and drug research. Most drugs target proteins directly, which puts them into the focus of target identification.

The transmission of information in and between cells is largely controlled by protein–DNA and protein–protein interactions. There are manifold ways for proteins to network. For example, proteins can associate due to the presence of dedicated domains mediating protein–protein interactions [4]. The introduction of so-called post-translational modifications into proteins can also be responsible for the initiation of other protein–protein interactions. In the case of protein phosphorylation, the modification is often temporary. To date, more than 300 different post-translational modifications have been described that may occur in different combinations, on different

splice variants and on proteolytically processed forms of proteins. As a result, the number of different components in the human proteome vastly exceeds the number of different genes (~20–40,000) in the human genome. It is therefore estimated that the number of components in the proteome add up to approximately 1 million. In addition, single nucleotide polymorphisms (SNPs) can lead to subtle differences in protein sequence and protein structure, but result in a high impact on protein function [5,6]. For certain proteins, somatic mutations occur during disease (e.g., cancer), which explains the fundamentally different response upon drug treatment [7]. Understanding protein function at the molecular level is a prerequisite for the development of more effective drugs that might include a tailored treatment strategy to match individual genotypes in the future.

Characterization of the proteome would ideally involve measuring the abundance, post-translational modifications, interactions and intracellular localization of all proteins in a single sample. Currently available technologies allow only the determination of selected parameters, and focus either on measuring protein abundance or elucidating protein networks (TABLE 1). For example, mass spectrometry (MS)

has evolved as an extremely valuable tool for the reliable identification of proteins and protein modifications [8,9]. The combination of 2D gel electrophoresis (2DE) and MS is widely used to describe changes of protein abundance for several thousand proteins in parallel [10–13]. Generally, proteins are identified based on their altered abundance relative to a reference sample. However, the low dynamic range of gel electrophoresis identifies only the most prominent proteins by 2DE. The unbiased 2DE approach has nevertheless proven to be a true discovery tool. MS has further evolved to sophisticated methods such as the identification and quantification of proteins from cell extracts based on the labeling of two biologic samples with different isotope-coded affinity tags (ICAT) [14].

Currently, the yeast two-hybrid (Y2H) system is the method of choice to map protein–protein interactions. This approach has significantly contributed to drafting a scheme of protein–protein interactions in the cellular network [15–19]. However, it solely identifies protein interaction partners and does not deliver information on the dynamic character of protein–protein interactions. The Y2H system is adaptable to a high-throughput format. However, proteins studied using the Y2H approach lack their natural post-translational modification. Protein–protein interactions can also be mapped based on the homologous expression of epitope-tagged bait proteins in yeast. Practically, protein complexes are affinity-purified, separated by gel electrophoresis and identified by MS [20,21]. All proteins associating with a certain bait protein are pulled down within a large protein complex. However, these proteins

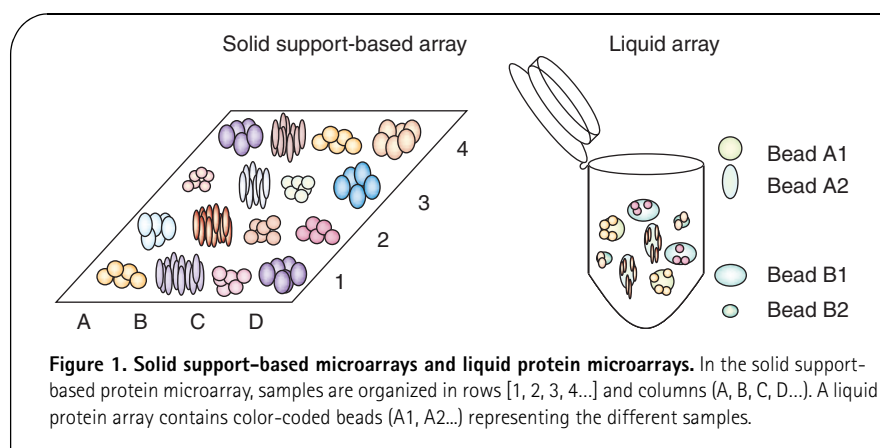
may never interact with the bait at the same time and in the same place *in vivo*. For this reason, protein networks give the impression of being more elaborate than they actually are.

Overall, there is poor overlap between protein–protein interaction maps identified by the different approaches. For example, only 7% of the interactions identified with the tandem affinity purification (TAP)/MS technology were also observed by Y2H assays [20]. Generally, data from both approaches are difficult to compare since Y2H studies portray binary protein–protein interactions whereas the TAP/MS technology analyzes complex protein–protein mixtures. This suggests that the assembly of protein complexes is more than the sum of binary protein–protein interactions, and comprehensive Y2H studies are therefore not particularly suited to elucidate protein–protein interaction networks. Y2H analysis and TAP/MS technology are equally complementary approaches addressing very different aspects of protein–protein interactions.

Short-range interactions control the assembly of individual protein complexes. None of the established technologies allow systematic studies of post-translational modifications, binding of partner proteins, cofactors and metabolites, and the molecular assembly of protein complexes. As a recently developed technology, protein microarrays have the potential to fill the knowledge gap left by the previously established proteomic technologies. Based on protein microarray technology, protein–protein interaction studies can be performed with less material, higher throughput, better reproducibility and in less time compared with standard biochemical methods [22,23]. Due

Table 1. Overview on established and novel technologies in proteomics.

Technique	Question studied
Yeast two-hybrid technologies	Identification of novel protein–protein interactions Drafting of protein–protein interaction maps Advantage: automation, genome-wide screening and discovery tool Disadvantage: high number of false positives and proteins lack post-translational modifications
Tandem affinity purification technology Epitope tag-based purification of protein complexes, identification by mass spectrometry	Drafting of protein–protein interaction maps Advantage: proteins expressed under natural promoter, native processing and discovery tool Disadvantage: protein complexes appear larger than they actually are
Protein–function arrays	Identification of enzyme–substrate and protein–protein interactions Advantage: automation, little material consumption and easy-to-perform assays Disadvantage: access to a large number of recombinant proteins required, arrays are not yet commercially available and array design requires experimental hypothesis
Isotope-coded affinity tag Mass spectrometry-based methods to determine protein abundance	Protein abundance Advantage: global tool for protein quantification in cells and tissues Disadvantage: limited to cysteine-containing proteins, method needs skilled and experienced laboratory personnel
2D gel electrophoresis and mass spectrometry	Protein abundance Advantage: discovery tool Disadvantage: time and material consuming, does not cover the entire proteome
Protein-detection arrays	Protein abundance Advantage: automation, little material consumption, easy-to-perform assays Disadvantage: access to a large number of highly specific antibodies required and array design requires experimental hypothesis



to their capacity to study multiple time points in parallel, protein microarrays are also a promising tool to monitor time-resolved dynamic changes of protein metabolism in a cell, tissue or organism.

This versatile platform is expected to impact our understanding of biology and signal transduction networks, to speed up diagnosis and drug discovery and to advance the development of personalized medicine.

Classification of protein microarrays

A protein microarray contains a defined set of proteins immobilized in high density in an addressable format. Typically, glass slides (2.5×7.5 cm) serve as a solid support to harbor the immobilized proteins as probes in a geometric pattern. Consequently, a protein spot can be addressed similar to the organization of samples in a multiwell plate. In an alternative set-up, protein probes are immobilized to different color-coded microbeads, thus forming a liquid protein microarray (FIGURE 1) [24]. The sample also carries a color-label and the analysis is based on sorting beads according to their color-label.

The diverse character of proteins is reflected by the development of diverse assay platforms and formats. Protein microarrays are available that measure protein abundance in cell lysates and body fluids [25,26]. Other applications focus on finding substrates for certain enzymes or protein interaction partners [24,27]. Protein microarrays with proteins or antibodies present in an addressable pattern are referred to as forward-phase protein arrays (TABLE 2). These arrays can be probed with cell lysates or with purified and

well-characterized proteins. Alternatively, in so-called reverse-phase protein arrays, different cell extracts can be spotted onto the array. These arrays are probed with different antibodies (e.g., to monitor the regulation of signaling pathways) [28–30].

Forward-phase protein microarray applications can be divided into two principal types [31]. Microarrays containing capture agents to assess the concentration of proteins are known as protein-detection arrays, where antibodies are typically immobilized as capture agents. This kind of protein microarray

corresponds to the technology employed for RNA expression profiling. Nevertheless, as RNA expression profiling data do not necessarily correspond to the actual protein concentration in a cell, protein-detection microarrays are a necessary and valuable tool whenever the real protein concentration has to be determined. Protein expression profiling requires highly specific antibodies with minimal crossreactivity towards all other sample proteins. Current applications are limited to the availability of suitable antibodies that are highly specific and bind their cognate antigens with comparable affinity constants. The fact that the concentration of different proteins in a cell covers several orders of magnitude affects the performance of antibody arrays. Proteins present at high molar concentrations might always be captured by antibodies directed against proteins present at lower molar concentration. Currently, only the quantification of cytokines and chemokines has been realized in a multiplexed system, and kits are already commercially available from different suppliers. Functional properties of a large collection of immobilized recombinant proteins are usually addressed by the second category of protein microarray, the protein-function array.

This review will focus exclusively on this category of protein microarray and its potential to study protein–protein interactions and protein–small molecule interactions.

Large filter arrays (e.g., 22×22 cm) were also employed as protein arrays. The advantage of the miniaturized format (e.g., small sample consumption) is certainly lost by switching to a macro format. Nevertheless, large filter sets accommodating

Table 2. Overview on different categories of protein microarrays.

Category	Subtype	Array content	Assay
Forward phase	Protein-function array	Recombinant proteins or peptides and peptide libraries Antibodies	Screening for novel substrates Enzymatic activities Protein–protein interactions Protein–lipid interactions Protein–small molecule interactions
	Protein-detection array	Protein capture agents (e.g., antibodies)	Protein quantification
Reverse phase		Cell extracts	Antibody-based detection of proteins

thousands of different expression library clones have been successfully employed for screening applications (e.g., the characterization of antibody specificities) [32]. As far as they are instructive for potentially miniaturized applications, results from dot-blot arrays and large filter arrays are included in this review. Other issues relevant for protein microarray experimentation, such as the instrumentation required for production of protein microarrays, labeling of the sample and methods for signal detection, will also be introduced. Approaches to provide a high number of proteins as content for protein microarrays will also be discussed. Finally, the perspectives and needs for future protein-function microarray experimentation will be assessed.

Fabrication of protein microarrays

The fabrication of solid support-based protein microarrays requires that different probes are delivered to the slide surface in a geometric pattern. Usually, the probe is delivered as a small droplet containing a presynthesized protein or a purified antibody. The sample is either absorbed into a microporous matrix or bound covalently to the activated slide surface. Unbound probe is removed by washing and the slide surface is blocked to minimize nonspecific binding (FIGURE 2). Alternatives to spotting proteins are also explored, such as printing expression vectors that encode different proteins. In this case, protein expression can be performed right on the spot. As with any other strategy, *in situ* fabrication of protein microarrays has advantages and disadvantages. *In situ* production of proteins is not limited to the use of *Escherichia coli* extracts, wheat germ and reticulocyte lysate work. Nevertheless, *in situ*-produced proteins are synthesized when required and individual differences of the long-term stability properties of different proteins do not present an issue. In addition, toxic proteins can also be produced by *in situ* translation. Currently, the deposition of presynthesized proteins is more commonly used. This approach allows a direct control on the amount of protein deposited in contrast to the *in situ* synthesis of proteins.

Proteins can be immobilized to the solid surface by covalent attachment or adsorption. A range of different surface options are offered by different suppliers (TABLE 3). For example, a thin pad can be attached to the glass slide, thereby trapping proteins into a 3D structure. Frequently applied materials include ultrathin polyacrylamide gel pads or nitrocellulose layers, or carbohydrate-based matrices. These coatings generate a 3D environment to accommodate proteins. It is assumed that the tertiary structure of proteins is better maintained within a 3D matrix than on

a regular glass surface. A polycarylamide-based hydrogel and nitrocellulose-based coatings absorb proteins into their 3D meshwork. When the protein drop is delivered to the dry hydrophilic coating, the matrix takes up water and simultaneously the dissolved molecules, which are trapped within the meshwork. Other 3D surface coatings possess functional groups for covalent attachment of the ligand to the matrix. General issues such as choice of microarray type, blocking of nonspecific binding, improving the signal-to-noise ratio, increasing sensitivity and signal range have been intensely investigated [33,34].

Generating the content for protein-function microarrays

Microarray-based protein function-assays require access to recombinant proteins. Preparing a high number of proteins that are of sufficient purity and stability is a challenging step. Proteins are commonly produced with N- or C-terminal fusion tags, which can subsequently be used in downstream processes such as affinity purification and immobilization. Proteins can be expressed in a bacterial host such as *E. coli* or in eukaryotic expression systems such as insect cells or yeast. Each expression system has its own limitations and benefits. For example, bacterial protein expression is fast, inexpensive and easy to perform. This is the system of choice for small and soluble proteins. The size and hydrophobicity of many human proteins indeed makes it difficult to express them as functionally active proteins in *E. coli*. This fact aside, eukaryotic proteins are, in most cases, not processed correctly and therefore lack their native post-translational modifications. Eukaryotic hosts, such

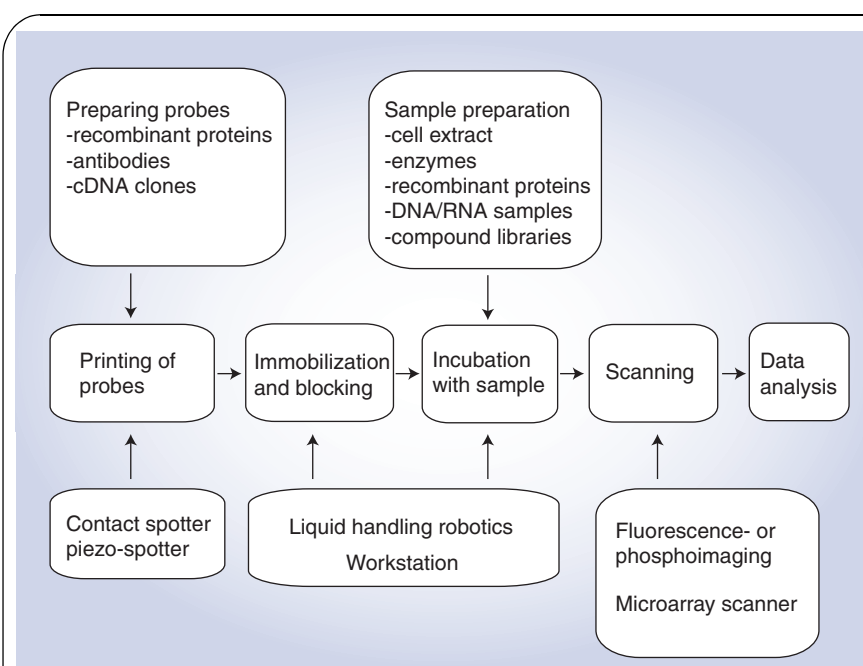


Figure 2. The different steps of a protein microarray experiment. The principal steps are listed in the center. Information on the instrumentation required for production of protein microarrays, microarray processing and microarray analysis is listed in the lower part. The upper part informs on steps such as probe and sample preparation. cDNA: Complementary DNA.

Table 3. Overview on surface coating designed for printing protein microarrays.

Structure	Surface	Slide	Company
3D	Polyacrylamide gel	Hydrogel	PerkinElmer
3D	Organic hydrogel	Nexterion H	Schott Nexterion
3D	Nitrocellulose	FAST	Schleicher & Schuell
3D	Nitrocellulose	Vivid	PALL
3D	Hydrophilic polymer	CodeLink™	Amersham/Surmodics
3D	Dendrimeric polymer	ProteoBind	VBC Genomics
2D	Coatings with functionalized silane	Various types	Greiner, Schott/Quantifoil

as yeast and insect cells, provide a way to produce hydrophobic proteins and those with post-translational modifications. Furthermore, it is also possible to coexpress different proteins in insect cells. This facilitates, for example, the production of binary protein complexes. Thus, it is also feasible to coexpress a modifying enzyme, such as a kinase, to ensure the presence of the desired post-translational modification. This is particularly important when the activity of an enzyme is frequently controlled by post-translational modification. Although a eukaryotic host mostly guarantees correct processing, eukaryotic protein expression is more tedious and a much lower expression level is usually achieved. This is particularly true in the case of insect cells where protein expression requires experience and hands-on time. Therefore, the advantages and disadvantages of a certain expression system have to be evaluated carefully and matched to the requirements of downstream experimentation.

Although concerns have been raised regarding correct folding, protein stability and the presence of the correct post-translational modification of heterologously expressed proteins, recombinant human proteins have successfully been used for microarray-based *in vitro* studies. For instance, 48 purified recombinant human proteins of the basal transcription machinery were arrayed onto nitrocellulose and successfully probed for protein–protein, protein–DNA and protein–small ligand interactions [35]. Furthermore, even large filter arrays of complementary DNA (cDNA) expression libraries, where each spot contains a background of *E. coli* proteins, can serve as protein array. For example, such a filter array was screened with the nuclear protein kinase DYRK1A. The screen, which had close to 38,000 double-spotted plasmid clones of a human fetal brain expression library, identified a previously unknown substrate of the cyclin family, cyclin L2, as a substrate of the kinase DYRK1A [36]. Today, microarray-based screens have not been performed for such a large number of purified proteins. A chip with close to 6000 different glutathione S-transferase (GST)-tagged proteins is the most comprehensive protein chip created thus far. This protein microarray represents nearly 80% of the yeast proteome. Each protein was produced as a GST fusion protein and affinity-purified from yeast [37]. Due to the size and complexity of

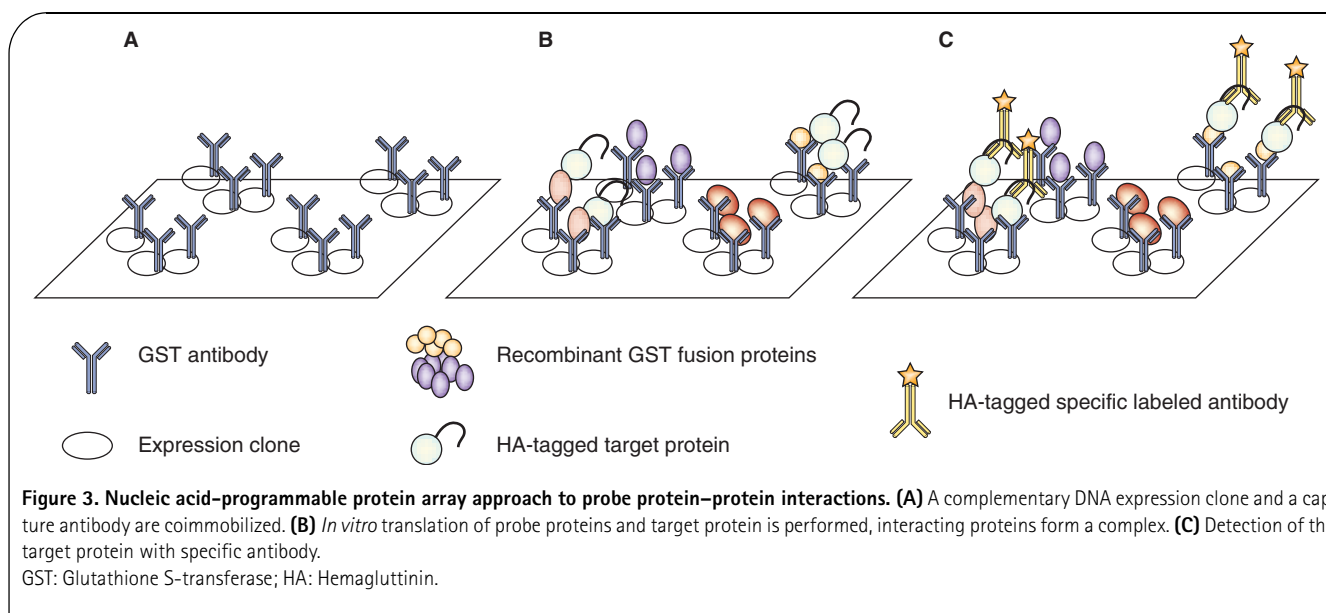
the human genome, the production and purification of a large set of recombinant human proteins is even more challenging [38–40]. Fortunately, the human open reading frame (ORF)ome is also approaching completion and protein-encoding cDNAs are widely available [41,42].

A few companies have embarked on the task to produce protein microarrays. For example, a protein chip representing the yeast proteome has already been announced. This yeast protein chip will soon be followed by arrays representative of the human proteome [43]. The proteome protein microarrays are expected to advance the discovery of disease signaling pathways, identify novel drug targets and help to close down on mechanisms leading to drug side effects.

LaBaer and coworkers envisaged the challenges associated with high-throughput protein production and purification. Instead of spotting proteins, cDNAs encoding the target proteins were delivered to the microarray [44]. The actual proteins were produced with a cell-free expression system and captured on the array via epitope tags fused to the protein coding sequence. Using this method, an array of different proteins can be produced and trapped to the slide surface by using a suitable immobilization strategy. In an approach referred to as nucleic acid-programmable protein array (NAPPA), binary protein–protein interactions of the human replication complex were examined (FIGURE 3).

Instrumentation

In order to generate addressable patterns on the microarray surface, a high-precision instrument, a microarray spotter, is employed as a printing device. This instrument delivers a tiny droplet of the dissolved protein (subnanoliter volume) to the microarray surface. This results in the formation of microscopic spots approximately 100–250 µm in diameter. The distance between the spots varies from 200–500 µm. In principle, proteins can be arrayed in high spatial density (1000 spots/cm²) and a standard microscope slide (25 × 75 mm) can accommodate thousands of samples. The droplet can be delivered to the surface by direct contact between the tip and the solid support. In surface contact printing, the pins are dipped into the sample solution, resulting in sample uptake into split pins or transfer of a small volume to the tip of a solid pin



(FIGURE 4). The volume delivered to the microarray surface by direct contact depends on parameters such as pin type, sample viscosity, temperature, humidity and the type of microarray surface. Drop delivery in a noncontact drop-on-demand fashion is accomplished with higher spot-to-spot reproducibility using piezo element-based arrayers. However, up to now the use of piezo element-based arrayers has been restricted by the

comparatively low throughput. A focus on the instrumentation needed for protein microarray manufacturing is provided in the review by Todd Martinsky [45].

Signal detection & amplification

Signals on protein microarrays were detected with phospho-, fluorescence- or chemoluminescence-imaging tools. The instrument should match the needs of microarray-based experiments by providing a high resolution ($>10 \mu\text{m}$), dynamic range and scanning speed. Microarray scanners originally developed for the field of RNA expression profiling fulfill these requirements. Most microarray scanners are equipped with lasers for the excitation of fluorescent dyes in the 500–700 nm range. Currently, only microarray scanners provide the scanning speed and resolution required in microarray experiments.

The detection of a protein in microarray-based experiments requires a label that can be identified in the subsequent scanning process. Frequently, fluorescent fluorophores are used to label a protein ahead of the experiment. There are a few general concerns regarding external labeling of proteins. Most of all, the labeling itself may change the properties of the sample and thus influence the interaction with the probe. In order to compensate for this effect, a double labeling strategy employing a dye switch can be used. This way, the individual effect of two different dyes on signal intensity and background can be normalized. Typically, a primary amino function, which is generally present in all proteins, is modified by the reaction with a fluorescent dye or, instead, with a hapten such as biotin or fluorescein [46,47]. The fluorescent dye can be directly detected on the array. Detection of the hapten requires a hapten-directed antibody tagged with, for example, a fluorescent marker (FIGURE 5). Alternatives to labeling amino groups were also explored. For example, a label can be introduced as a sulfur-coordinated metal complex through interactions with methionine or cysteine side chains [48]. The number of accessible functional

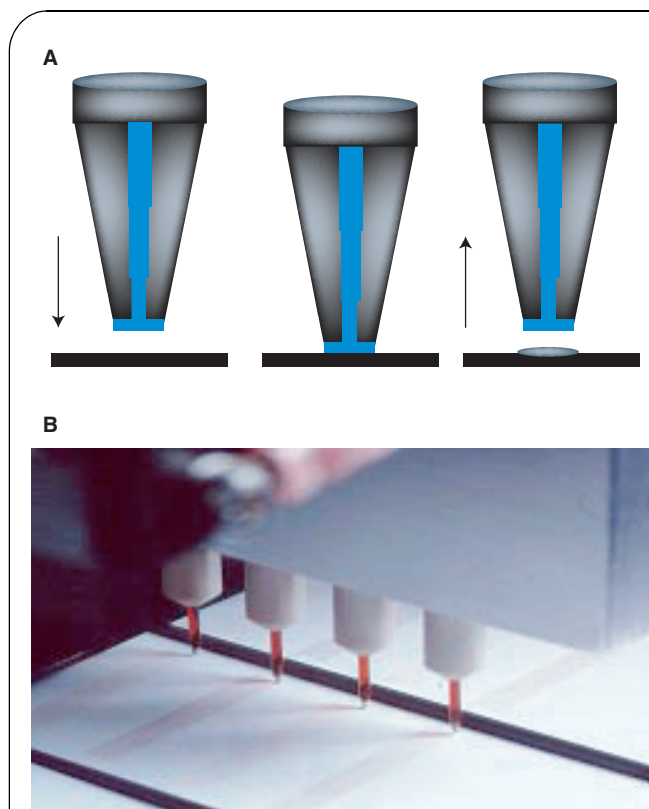
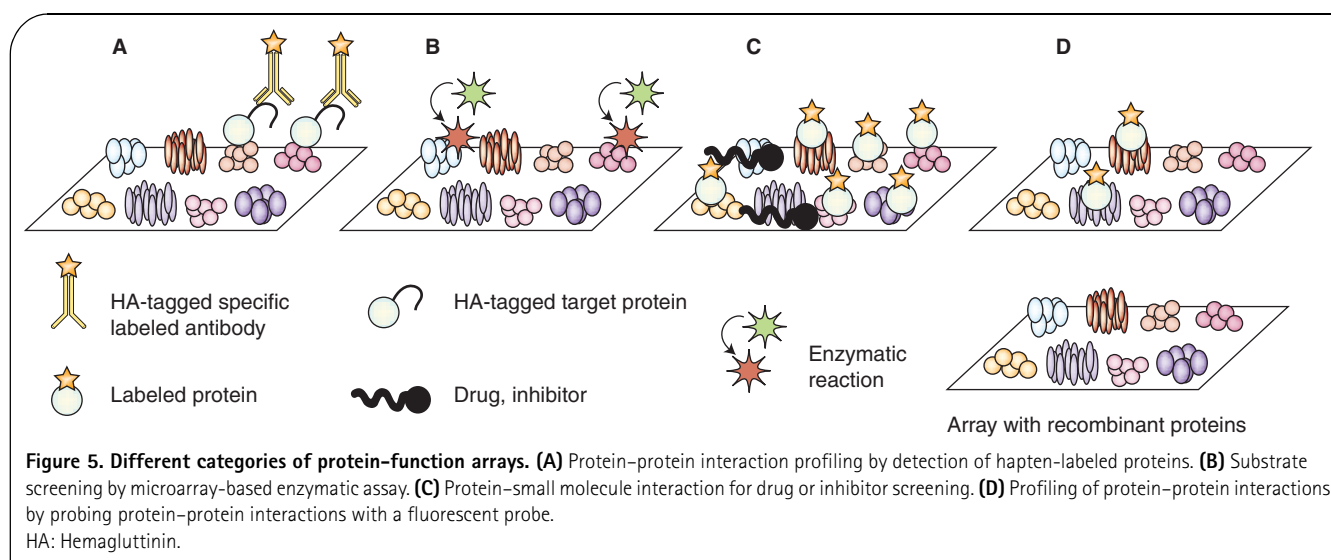


Figure 4. (A) Telechem's split-pin spotting pins. **(B)** Drop-on-demand delivery with a piezo-spotter.



groups of a certain protein (e.g., number of methionines, cysteines and primary amino functions) determines labeling efficiency. Consequently, different proteins are labeled with different efficiency.

In an ideal experimental setting, the detection of a protein on the microarray does not require an external labeling step ahead of the experiment. This can be realized in two different ways. First, a detectable label is produced in the microarray-based assay itself to ensure that a direct readout is possible. For example, kinase activities can be monitored when radiolabeled ATP is used in the phosphorylation reaction. The kinase-catalyzed enzymatic reaction introduces a phosphate group into an amino acid motif-defined site of the protein. Proteins carrying the radioactive phosphate can be detected among the array of nonphosphorylated substrate proteins, by exposure to autoradiography film or by phosphoimaging. However, in order to monitor the protein-protein interaction itself, the sample must carry a tag, such as an epitope tag.

Second, surface plasmon resonance (SPR) is a promising method for the direct detection of protein-protein interactions. In SPR, the binding of small molecules or proteins to array-immobilized proteins is detected by the change of the refractive index in the interface between the liquid phase and the solid carrier. SPR instruments were developed to analyze multiplexed samples in a microtiter plate format. This application has not been exploited in the field of protein microarrays. A further drawback is that the method is not as sensitive as fluorescence.

Remaining challenges

Protein microarrays are considered a promising tool for biologic discovery and drug development. However, studying proteins and protein function on a large scale is still associated with a number of technical challenges. First, the dynamic range of proteins in a cell or in an organism can vary by ten orders of magnitude [49]. As a consequence, proteins present in high concentrations can crossreact with array components, leading to background problems and reducing signal significance. Second,

the stability and functional activity of proteins throughout the microarray experiment must be guaranteed. Certain surface coatings may potentially stabilize different proteins with different efficiencies. Third, printing of protein microarrays requires access to recombinant proteins. The complex character of proteins still requires skills and experience. Currently, the timeframe of preparing a large number of recombinant proteins exceeds the timeframe of the protein-function microarray experiment.

A variety of proof-of-concept studies were performed with a limited number of human proteins. Proteins were produced in a bacterial expression system or in yeast. According to these results, the protein content for protein-protein interaction studies can be produced by simple expression systems. The following chapter provides an overview on the current status of protein-function arrays regarding human proteins as well as the yeast proteome (TABLE 4).

Table 4. Different categories of protein-function arrays and their applications.

Assay type	Examples
Protein-protein interaction profiling	Protein-protein interaction mapping Identification of direct protein-protein interactions for large protein complexes Discovery of novel protein-protein interactions Substrate screening
Protein-DNA interaction profiling	Discovery of DNA-binding proteins and of their target sequences
Protein-small molecule interaction profiling	Metabolomics Drug discovery Inhibitor screening Identification of alternative targets Optimization of binding specificities
Protein-lipid interaction profiling	Determination of protein function

Detection of protein–protein interactions, protein–small molecule interactions & protein phosphorylation

In a proof-of-concept study, the group of MacBeath and Schreiber spotted proteins of the nuclear factor (NF)- κ B pathway to demonstrate that purified proteins retain their activity when immobilized on glass slides [57]. Three pairs of proteins that were well known to interact were spotted on glass slides, protein G and immunoglobulin G (IgG), and two proteins of the NF- κ B complex. The interaction of the third pair, the human immunophilin FKB12 and an appropriate FKBP12-binding domain, depends on the presence of a small molecule, rapamycin. All proteins were recognized specifically by their counterpart. In addition, three different kinase substrates, kemptide, protein phosphatase inhibitor 2 (I-2) and Elk1, were specifically phosphorylated by protein kinase A, casein kinase II and p42 mitogen-activated protein kinase (Erk2). Furthermore, protein–small molecule interactions were tested with synthetic ligands of FKB12 and could be detected even for those molecules with binding constants in the micromolar range.

Protein kinase assays

In a pioneering study by Zhu and coworkers, 119 yeast kinases were expressed and purified as GST-fusion proteins, and assayed with 17 different substrates [56]. The reaction was performed in a 0.5 μ l volume in a self-fabricated silicone elastomer protein chip. The self-fabricated protein chip revealed a tenfold better signal-to-noise ratio when compared with a standard 384-well microplate. The analysis was performed by exposure to x-ray film with a phosphoimage scanner. This group found that particular kinases preferentially phosphorylate particular substrates. The results from *in vitro* assays were generally in accordance with *in vivo* data.

Screening phospholipid & calmodulin interacting proteins

A comprehensive yeast protein microarray was fabricated by Zhu and coworkers containing an almost complete set of proteins encoded by the 5800 ORFs from yeast [37]. The proteins were prepared as GST-fusion proteins and printed on glass slides using contact spotting.

This yeast chip was screened for interactions with five different types of phosphatidylinositol (PI)-containing liposomes, which were spiked with a biotinylated lipid for detection. These studies revealed that the different types of phospholipid mixtures bound specifically to certain proteins. The yeast chip was also probed with calcium/calmodulin for calmodulin-binding proteins. This study is the most comprehensive published microarray-based investigation to date on protein–protein and protein–lipid interactions.

Protein domain studies

Protein microarray technology was used to profile protein domains mediating protein–protein interactions through short peptide motifs. This is one of the mechanisms controlling the assembly of large multiprotein complexes *in vivo*. The group of Bedford spotted 212 GST-fusion proteins representing

145 different protein domains, such as Src homology (SH)-2 and SH-3, and the PDZ motif, on nitrocellulose film [53]. Fluorescently labeled peptides known to bind specifically to these domains were used as target and unique binding profiles were detected for different samples. Another mechanism directing protein–protein interactions are post-translational modifications. The protein-domain microarray also distinguished between targets with differing post-translational modifications (e.g., between an arginine-methylated peptide and its non-methylated counterpart). Furthermore, the array was examined with unlabeled cell lysate and microarray-bound proteins were detected with appropriate antibodies. In contrast to traditional blot-overlay assays, the microarray approach offers the option to screen a large number of proteins in parallel in less time and with higher sensitivity.

Systems-approach to map the protein–protein interactions of the human DNA replication complex

Ramachandran and coworkers have used their NAPPA assay to query a collection of 29 cDNAs representing different proteins of the human DNA replication complex [44]. The proteins were produced by *in vitro* translation as GST-fusion proteins. All cDNAs were coimmobilized on simple glass slides with the tag-specific antibody. Protein translation was induced in the presence of a cDNA coding for a potential interaction factor. Both the immobilized probes and target were produced in parallel by *in vitro* translation, and probes with the potential to interact with each other could assemble on the array. Employing this strategy, the 29 proteins were probed against each other (TABLE 5). Protein–protein interactions were monitored with the help of a hemagglutinin (HA)-tagged antibody since the target was translated as an HA-tagged protein. Results from the NAPPA strategy were compared with Y2H results and published data. The NAPPA approach detected 110 protein–protein interactions, 63 of which were previously unknown.

Mapping interactions of p53 mutations

Blackburn and coworkers used a different approach to probe mutants of the p53 tumor suppressor protein with its regulatory oncoprotein MDM2 [46]. They produced 45 variations of the p53 protein encoding established missense and nonsense germline mutations and spotted the crude cell lysates containing the recombinant proteins on coated microarray slides. The array was first probed with purified recombinant MDM2 labeled with Cy3 fluorophore. Two p53 mutants showed greatly reduced binding with MDM2. In addition, replicate p53 arrays were probed with double-stranded DNA corresponding to the GADD45 promoter element. This approach identified high-affinity binding proteins (wild type p53) as well as low-affinity binding and nonbinding p53 mutants. Furthermore, a casein kinase II assay was performed and phosphorylated proteins were identified by staining with a phosphospecific antibody. Truncated versions of p53 lacking the relevant serine phosphorylation site were easily identified.

Table 5. Published applications to probe protein–protein interactions using protein arrays: the most recent examples and early filter array applications.

Assay	Principle	Array density	Ref.
Protein–protein interactions of the human DNA replication initiation complex	Coexpression of GST-tagged probe and target, probe is captured to array surface by GST antibody, target carries HA tag, detection (e.g., with HA tag-specific antibody)	29 proteins	[47]
Analysis of 45 germline mutations of p53	Spotting of cell lysates containing recombinant p53 MDM2 interaction assay GADD45 promoter element binding Phosphorylation assay	45 proteins	[46]
G-protein-coupled receptor microarrays	Chemically activated surface to immobilize lipids as model membrane with lateral fluidity and high mechanical stability, immobilization of recombinant receptors, ligand and inhibitor binding studies	Three receptors	[50]
Identification of chromatin-related protein interactions	Proteins, antibodies and DNA immobilized to functionalized glass slides, probed with DNA repair proteins, antibody-based identification of novel protein–protein interactions	Eight different protein samples	[51]
Specificity mapping of coiled-coil proteins (bZIP's)	Array of recombinant bZIP proteins, incubation with fluorescently labeled probe	59 bZIP proteins from human and yeast	[52]
Probing protein–protein interaction-mediating domains	Array of recombinant GST fusion proteins probed with fluorescently labeled peptides	212 proteins, protein domains and motifs	[53]
Antibody-based fluorescent detection of phosphoproteins	Detection of phosphorylation sites with phosphospecific antibodies as an alternative to radioisotope labeling	Two peptide substrates	[54]
Peptide chips to map kinase specificity	Self-assembled monolayer to determine K_i value for three c-src kinase	10 peptide substrates	[55]
Identification of calmodulin- and lipid-binding proteins	Array of recombinant GST-fusion proteins prepared from yeast, incubation with biotinylated liposome preparations and with biotinylated calmodulin	5800 yeast proteins	[37]
Kinase specificity mapping	Immobilization of substrate protein to a self-fabricated silicone-elastomer-based protein chip	17 substrate proteins assayed with 119 kinases Autophosphorylation of kinases	[56]
Kinase assay Protein–protein interaction Small molecule–protein interaction	Immobilization of probes to glass slides by contact printing	Three kinase substrates Three pairs of proteins well known to interact with each other Three small molecule targets	[57]
Protein–protein Protein–DNA Protein–ligand interactions	Dot-blot assay on nitrocellulose	48 proteins	[35]
Protein–protein interactions	Antibody array on PVDF membrane	100 antibodies	[58]

GST: Glutathione S-transferase; HA: Hemagglutinin; K_i : Inhibition constant; PVDF: Polyvinylidene difluoride.

G-protein-coupled receptor microarrays

A class of membrane proteins, G-protein-coupled receptors (GPCRs), are interesting drug targets and are therefore of major interest for the pharmaceutical industry. Assays to screen ligands against GPCRs are typically performed in a low-throughput format. To date, approximately 190 GPCRs with known natural ligands have been described. In addition, 700 orphan receptors, which are GPCRs with unidentified ligands, exist. Multiplexing different GPCRs in a protein microarray is considered

to be a potentially powerful tool to streamline ligand identification. Typically, membrane proteins need to be embedded in a membrane environment to maintain their native conformation and functional activity. Fang and coworkers have identified γ -aminopropylsilane-modified glass surfaces to be ideal for printing membrane arrays [50]. The immobilized lipids maintain their mechanical stability and lateral fluidity. In a proof-of-concept study, three different subtypes of adrenergic receptors were probed with three different compounds.

Identification of chromatin-related protein interactions using protein microarrays

In a study by Coleman and coworkers, proteins, antibodies and DNA were immobilized to functionalized glass slides and screened with a small set of DNA repair proteins, RAD51B, RAD51C and SMARCA1 [51]. Proteins retained on the microarray were identified with specific antibody. The DNA repair protein RAD51B, but not its cognate partner RAD51C, has affinity for free histones, but not for histones complexed with DNA (nucleosomes).

Prefabricated arrays to probe protein-protein interactions

Whole-proteome microarrays have the potential to transform target discovery, drug discovery and basic research. These arrays could be screened with recombinant proteins, small molecules, DNA and RNA. In addition, protein microarrays can be reproducibly quantitated and reach a high level of sensitivity to allow detection of interactions that standard technologies do not detect. A few companies have announced that prefabricated proteome microarrays will soon be on the market. A yeast proteome array based on the work of the Snyder and coworkers has already been released [37]. This array comprises 4088 different yeast proteins [43]. Since 30% of the yeast proteome has significant homology to human proteins, this chip will be of interest beyond the yeast research community. Protein microarrays covering the human proteome or parts of it have also been announced. Thus, access to recombinant human proteins is offered by Protagen [101]. Their proprietary UNICLONE collection contains 11,000 different clones coding for human proteins and purified proteins are offered in milligram scale.

Summary & conclusions

Protein microarrays present a versatile *in vitro* platform open for various types of biochemical analyses. Diverse assays have been adapted to the microarray format. These range from measuring protein abundance to screening a library of unknown proteins for enzymatic activities. All approaches exploit the idea of multiplexed miniaturized assays. In proof-of-concept studies, it has been shown that miniaturization of assays indeed holds the promise of low sample consumption and increased sensitivity. The development of new assays requires a suitable detection strategy (e.g., fluorescence or radiolabeling) and access to a high-resolution scanner.

However, the protein microarray content presents a key issue for the experimental design of functional assays. General access to clone collections, which is available through various sequencing projects, presents a valuable asset for the development of new applications. Currently, the main bottleneck lies within the preparation of a high number of different recombinant proteins in a pure form. Alternative splicing increases the complexity of the human genome even further. The structural diversity of proteins requires that different expression strategies are employed in parallel to result in a reasonable number of successfully produced proteins. In particular, the large-scale production of human proteins requires a tailored

approach paying attention to the properties of different classes of proteins. Thus, a considerable number of human proteins associate with membrane-based compartments. Those proteins are generally difficult to express in a soluble form.

Nevertheless, studies focusing on a limited number of human proteins have generated valuable results and new insights into biologic questions. For example, functional properties of well-known somatic p53 mutants were screened in a microarray format. A different study recently focused on interactions between different proteins of the human DNA replication complex. Taken together, both studies show that the protein microarray format presents a useful platform for systems-approaches performed for a limited number of proteins.

Expert opinion

Protein microarray technology certainly has considerable potential for the large-scale characterization of *in vitro* properties of proteins. Nevertheless, a general utilization of this technology has been held back by the fact that preparing a large number of suitable probes is a laborious and challenging process. Large-scale expression of protein requires skills and experience and access to clone resources. It has been announced that comprehensive microarrays representative of a eukaryotic proteome will become available for the biotechnology market, but are as yet unavailable. Therefore, the preparation of a suitable content still lies within the hands of the investigator. Most proteins are of a diverse and delicate structure. Thus, large-scale experimentation requires that the 3D structure is maintained in the time span between probe preparation and data analysis for all proteins of the array. Initial studies focused on identifying suitable solid supports and spotting conditions. Among the different surface coatings, those forming a microporous surface have turned out to accommodate a large number of different proteins with adequate reproducibility. In addition, the subsequent quantitative analysis of microarray data requires that all probes are immobilized homogeneously with respect to spot size and morphology. Noncontact or drop-on-demand printing approaches are well suited for microporous materials since the sample can be delivered without damaging the microarray surface.

Historically, the first array of human proteins was generated as a large filter array by spotting *E. coli* expression clones of a randomly cloned cDNA library. Protein expression was induced directly on the filter. The bacteria were lysed and all bacterial as well as recombinant proteins were bound to the filter membrane. This approach includes most proteins present in the human proteome. Nevertheless, protein expression based on randomly cloned cDNA libraries is burdened with a number of shortcomings. First, many of the proteins are not cloned in frame with their coding sequence. The recombinant protein is therefore artificial. Second, many of the encoded proteins are not full length. Third, it is not possible to control the yield of recombinant protein on these filter arrays, which renders these as well as comparable approaches unsuitable for quantitative experimentation. Fourth, bacterially expressed

human proteins are not processed appropriately. Nevertheless, filter arrays have successfully been used as a screening tool to characterize antibody specificities and to identify novel kinase targets.

Through worldwide sequencing efforts, verified expression libraries and clone collections are currently becoming available. The cDNA libraries need to be normalized and screened for sequences coding for full-length proteins. Proteins of low abundance are not presented by cDNA libraries. However, whole genome data are helpful to complete cDNA collections. Current systematic strategies aim at the production of indexed libraries.

Exploiting these clone resources is the first prerequisite to produce a suitable content for protein-function arrays. A first step has been made by preparing a unigene set of the human genome. Recent advances in the field of cell-free protein expression, which potentially offer access to proteins that evade protein expression due to their toxicity, are also promising. In addition, the production of recombinant membrane proteins remains a challenge. Since many drug target proteins belong to this class (e.g., GPCRs), strategies for the expression of membrane proteins are especially required. The technical problems associated with large proteins and membrane-associated proteins can be partly circumvented by the production of certain protein domains in a soluble form. Due to the intricacy of producing and purifying a large number of proteins in parallel, the most recent applications of protein function arrays have been performed for only a limited number of human proteins (TABLE 4). Signal transduction networks are controlled by the assembly and disassembly of protein complexes. Even subtle differences in the primary sequence of similar proteins, or the presence of a certain post-translational modification can relate to huge functional differences. Protein microarray technology allows the examination of multiple proteins in parallel. For example, with the help of protein-domain microarrays, interaction maps were produced for the protein of interest [52,53]. In this manner, not only were binding partners of proteins identified, but qualitative differences between different probes were also evaluated. The comparative character of biochemical analysis cannot be achieved with established biochemical methodology, or with other large-scale proteomic methods such as TAP/MS technology or Y2H.

As with any method, protein microarrays are likely to produce false-positive results. The occurrence of false positives is the consequence of disturbing the natural biologic balance, and the mechanisms are relatively well understood for Y2H and TAP/MS technology. The development of protein microarray technology is surely in its early years, and therefore the occurrence of false-positives can only be hypothetically approached. First, *in vitro* experimentation can bring together proteins that would never meet *in vivo*. If these proteins are identified as binding partners, the subsequent *in vivo* follow-up will identify this result as an *in vitro* artifact. This type of false-positive result can especially be expected for the *in vitro* examination of as yet uncharacterized

proteins. On the contrary, when known proteins are studied in a comparative study, the rate of false positives will be significantly lower. Second, protein concentration in a cell covers several orders of magnitude. Exposing proteins to each other at nonphysiologic concentrations will potentially result in a number of false positives. Again, in this case the biologic relevance of the observed protein-protein interaction must also be confirmed *in vivo*.

From a technological standpoint, advances in the field of signal detection could further help to promote protein microarray applications. SPR certainly presents a promising option. SPR instruments were developed for detailed kinetic studies, but not as a high-throughput discovery tool. The ideal microarray platform would combine label-free detection with the identification of unknown binding partners by MS.

Five-year view

Preparing a large number of soluble, pure and stable recombinant proteins is currently the largest bottleneck for extensive applications of protein-function microarrays. This is especially true for studies focusing on mammalian proteomes.

The number of yeast genes has been debated for an extended time following completion of the sequence of the yeast genome. With the help of comparative genomics, this number has decreased to 5600 different genes at present. The number of human genes and proteins is considerably higher, but the exact number is still unknown. In addition, the complexity of the human proteome is further increased by a high, but still unknown, number of splice variants. The high number of splice variants also effects the generation of ORF resources. Cloning yeast genes has been simple, as only 10% of yeast genes contain one intron, and eight yeast genes have two introns. Cloning of yeast genes was therefore possible by amplifying genomic DNA. However, most human genes are spliced, and a different strategy based on reverse transcription PCR cloning is therefore required. The complex development of human beings results in a large number of different tissues and organs, which adds a level of complexity with respect to the resources that have to be analyzed for ORF identification. Consequently, approximately 30–40% of human genes have not yet been cloned. The structure of yeast proteins is generally simpler than that of human proteins. The lower number of protein domains in yeast is reflected by the smaller size of the yeast proteins. Generally, large proteins are difficult to express and require adequate expression strategies. The NAPPA expression system proves a promising technology as an alternative strategy to classical bacterial protein methods. However, the complex character of proteins suggests that eukaryotic hosts, such as insect cells, will also be explored.

Normalized cDNA libraries compatible with a recombination-based cloning system are a prerequisite for the production of tissue-specific proteome microarrays. Currently available human cDNA collections comprise several thousand full-length cDNAs that can immediately be utilized for protein expression. However, protein solubility and stability is not

necessarily indicative for correct processing and folding. Protein folding and processing is a complex and highly regulated process in mammalian cells. Despite its shortcomings, the fast and inexpensive bacterial expression systems will continue to be a resource for the production of eukaryotic proteins. It has been shown that approximately 60–80% of human proteins can be produced in full length in bacteria. Protein expression in eukaryotic hosts, such as insect cells, will contribute to enhancing the content value of protein microarrays.

The equipment to manufacture protein microarrays is relatively expensive and consumes space and maintenance time. Preparing the probes remains the biggest challenge. On the contrary, the instrumentation required for the analysis of protein microarrays (e.g., fluorescent scanners and phospho-imaging systems) are readily available and easy to use. Premanufactured protein microarrays could easily be analyzed in most academic laboratories.

In the next 5 years, protein microarrays will become valuable for the identification of human serum biomarkers. Human serum is a very informative source for the analysis of health and disease states. Due to the number of potential disease markers, protein microarrays will replace standard enzyme-linked immunosorbent assay technologies being restricted to the analysis of a single marker protein. Within the next 5 years, premade protein microarrays focusing on a certain clinical or biologic question will become available. However, emphasis on protein microarray design should be in probe quality rather than on probe numbers. Here, it is especially important to determine purity and integrity of the probes, and to normalize the protein concentration.

Acknowledgements

The authors would like to thank Silke Argo, Andreas Bunness, DKFZ Heidelberg, and the referees for their helpful comments on the review. The authors were supported by grant 01GR0101 of the National Genome Research Network, funded by the Bundesministerium für Bildung und Forschung (BMBF).

Key issues

- Protein–protein interactions are key to understanding cellular function.
- Traditional methods for mapping protein–protein interactions on a high-throughput scale produce a significant number of false positives and negatives.
- Established methods provide a static view of the dynamic regulation of biologic systems.
- Protein microarrays are ideal tools to monitor the dynamic changes that occur in a cell.
- Protein microarrays have a high capacity for multiplexing.
- The large amount of data available from microarray-based experiments requires statistic tools for the analysis and development of software tools.
- Protein microarrays consume little material of precious samples.
- Protein microarrays are versatile and can be adapted to different formats of *in vitro* assays.
- Protein microarrays are useful for inhibitor screening.
- Access to recombinant proteins is mandatory for performing protein function experiments.

References

Papers of special note have been highlighted as:

- of interest
 - of considerable interest
- 1 Pawson T, Saxton TM. Signaling networks – do all roads lead to the same genes? *Cell* 97, 675–678 (1999).
 - 2 Pawson T. Specificity in signal transduction: from phosphoryrosine-SH2 domain interactions to complex cellular systems. *Cell* 116, 191–203 (2004).
 - 3 Xenarios I, Salwinski L, Duan XQJ *et al.* DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res.* 30, 303–305 (2002).
 - 4 Pawson T, Nash P. Protein–protein interactions define specificity in signal transduction. *Genes Dev.* 14, 1027–1047 (2000).
 - 5 Nordborg M, Tavaré S. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 18, 83–90 (2002).
 - 6 Emahazion T, Feuk L, Jobs M *et al.* SNP association studies in Alzheimer's disease highlight problems for complex disease analysis. *Trends Genet.* 17, 407–413 (2001).
 - 7 Paez JG, Janne PA, Lee JC *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304, 1497–1500 (2004).
 - 8 Tyers M, Mann M. From genomics to proteomics. *Nature* 422, 193–197 (2003).
 - 9 Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 422, 198–207 (2003).
 - 10 O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021 (1975).
 - 11 Klose J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 26, 231–243 (1975).
 - 12 Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16, 1034–1059 (1995).
 - 13 Klose J. Genotypes and phenotypes. *Electrophoresis* 20, 643–652 (1999).
 - 14 Gygi SP, Rist B, Gerber SA *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnol.* 17, 994–999 (1999).
 - 15 Ito T, Chiba T, Ozawa R *et al.* A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl Acad. Sci. USA* 98, 4569–4574 (2001).
 - 16 Uetz P, Giot L, Cagney G *et al.* A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623–627 (2000).
 - 17 Rain JC, Selig L, de Reuse H *et al.* The protein–protein interaction map of *Helicobacter pylori*. *Nature* 409, 211–215 (2001).
 - 18 Giot L, Bader JS, Brouwer C *et al.* A protein interaction map of *Drosophila melanogaster*. *Science* 302, 1727–1736 (2003).

- 19 Li SM, Armstrong CM, Bertin N *et al.* A map of the interactome network of the metazoan *C-elegans*. *Science* 303, 540–543 (2004).
- 20 Gavin AC, Bosche M, Krause R *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147 (2002).
- 21 Ho Y, Gruhler A, Heilbut A *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415, 180–183 (2002).
- 22 Ekins R, Edwards P. On the meaning of 'sensitivity'. *Clin. Chem.* 43, 1824–1831 (1997).
- 23 Templin MF, Stoll D, Schwenk JM *et al.* Protein microarrays: promising tools for proteomic research. *Proteomics* 3, 2155–2166 (2003).
- 24 Vogel V, Sohnlein P, Jager C, Schade C, Steinert K. LiquiChip bead-based arrays: a flexible and efficient alternative to planar protein arrays. *Shock* 21, 2–2 (2004).
- 25 Stillman BA, Parker BO, Tonkinson J, Harvey MA. Applying multiplexed microspot immunoassays. *Genet. Eng. News* 24, 38 (2004).
- 26 Huang RP, Huang RC, Fan Y, Lin Y. Simultaneous detection of multiple cytokines from conditioned media and patient's sera by an antibody-based protein array system. *Anal. Biochem.* 294, 55–62 (2001).
- 27 Fetchko M, Studer I, Hort J, Stagljar I. Elucidation of the interaction network between yeast DNA processing proteins using endogenously tagged ORFs and the LiquiChip technology. *Yeast* 20, S285–S285 (2003).
- 28 Wulfschuhle JD, Aquino JA, Calvert VS *et al.* Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays. *Proteomics* 3, 2085–2090 (2003).
- 29 Pawletz CP, Charboneau L, Bichsel VE *et al.* Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 20, 1981–1989 (2001).
- 30 Grubb RL, Calvert VS, Wulfschuhle JD *et al.* Signal pathway profiling of prostate cancer using reverse phase protein arrays. *Proteomics* 3, 2142–2146 (2003).
- 31 Kodadek T. Development of protein-detecting microarrays and related devices. *Trends Biochem. Sci.* 27, 295–300 (2002).
- 32 Lucking A, Possling A, Huber O *et al.* A nonredundant human protein chip for antibody screening and serum profiling. *Mol. Cell. Proteomics* 2, 1342–1349 (2003).
- 33 Kusnezow W, Jacob A, Walijew A, Diehl E, Hoheisel JD. Antibody microarrays: an evaluation of production parameters. *Proteomics* 3, 254–264 (2003).
- 34 Angenendt P, Glöckler J, Murphy D, Lehrach H, Cahill DJ. Toward optimized antibody microarrays: a comparison of current microarray support materials. *Anal. Biochem.* 309, 253–260 (2002).
- 35 Ge H. UPA, a universal protein array system for quantitative detection of protein–protein, protein–DNA, protein–RNA and protein–ligand interactions. *Nucleic Acids Res.* 28, E3 (I–VII) (2000).
- 36 de Graaf K, Hekerman P, Spelten O *et al.* Characterization of cyclin L2, a novel cyclin with an arginine/serine-rich domain: phosphorylation by DYRK1A and colocalization with splicing factors. *J. Biol. Chem.* 279(6), 4612–4624 (2004).
- 37 Zhu H, Bilgin M, Bangham R *et al.* Global analysis of protein activities using proteome chips. *Science* 293, 2101–2105 (2001).
- Describes the production and application of a protein array covering almost the entire yeast genome.
- 38 Korf U, Kohl T, van der Zandt H *et al.* Large scale protein expression for proteome research. *Proteomics* (2005) (In Press).
- 39 Braun P, Hu YH, Shen BH *et al.* Proteome-scale purification of human proteins from bacteria. *Proc. Natl Acad. Sci. USA* 99, 2654–2659 (2002).
- 40 Hammarstrom M, Hellgren N, van den Berg S, Berglund H, Hard T. Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci.* 11, 313–321 (2002).
- 41 Wiemann S, Weil B, Wellenreuther R *et al.* Toward a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs. *Genome Res.* 11, 422–435 (2001).
- 42 Strausberg RL, Feingold EA, Grouse LH *et al.* Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl Acad. Sci. USA* 99, 16899–16903 (2002).
- 43 Chua J. Protomatrix readies whole-proteome array. *Scientist* 17, 42–42 (2003).
- 44 LaBaer J. Harnessing the proteome. *Biopolymers* 71, 282–282 (2003).
- 45 Martinsky T. Protein microarray manufacturing. *PharmaGenomics* 42 (2004).
- Profound discussion of the topic.
- 46 Boutell JM, Hart DJ, Godber BLJ, Kozlowski RZ, Blackburn JM. Functional protein microarrays for parallel characterisation of p53 mutants. *Proteomics* 4, 1950–1958 (2004).
- Describes a protein microarray application to analyze functional differences between allelic variants of the p53 oncoprotein.
- 47 Ramachandran N, Hainsworth E, Bhullar B *et al.* Self-assembling protein microarrays. *Science* 305, 86–90 (2004).
- First report of the *in-situ* production of a protein microarray based on the use of a mammalian reticulocyte lysate.
- 48 Heetebrij RJ. Universal linkage system (ULS): a versatile target labeling tool in genomics and proteomics. *Clin. Chem.* 48, 2091–2091 (2002).
- 49 Liotta LA, Espina V, Mehta AI *et al.* Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* 3, 317–325 (2003).
- 50 Fang Y, Frutos AG, Lahiri J. G-protein-coupled receptor microarrays. *ChemBiochem.* 3, 987–991 (2002).
- 51 Coleman MA, Miller KA, Beernink PT, Yoshikawa DM, Albala JS. Identification of chromatin-related protein interactions using protein microarrays. *Proteomics* 3, 2101–2107 (2003).
- 52 Newman JRS, Keating AE. Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science* 300, 2097–2101 (2003).
- Instructive protein microarray application based on mapping protein–protein interactions between human basic-region leucine zipper transcription factors.
- 53 Espejo A, Cote J, Bednarek A, Richard S, Bedford MT. A protein-domain microarray identifies novel protein–protein interactions. *Biochem. J.* 367, 697–702 (2002).
- Protein microarray with characterized modular protein domains to generate domain binding profiles of cellular proteins.
- 54 Lesaichere ML, Uttamchandani M, Chen GYJ, Yao SQ. Antibody-based fluorescence detection of kinase activity on a peptide array. *Bioorg. Med. Chem. Lett.* 12, 2085–2088 (2002).
- 55 Houseman BT, Huh JH, Kron SJ, Mrksich M. Peptide chips for the quantitative evaluation of protein kinase activity. *Nature Biotechnol.* 20, 270–274 (2002).
- 56 Zhu H, Klemic JF, Chang S *et al.* Analysis of yeast protein kinases using protein chips. *Nature Genet.* 26, 283–289 (2000).
- Pioneering experimental approach demonstrating the usefulness of protein microarray technology to characterize kinase activities.

- 57 MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science* 289, 1760–1763 (2000).
- **First study showing the innovative potential of protein microarrays.**
- 58 Wang YJ, Wu TR, Cai SY, Welte T, Chin YE. Stat1 as a component of tumor necrosis factor alpha receptor 1-TRADD signaling complex to inhibit NF-kappa B activation. *Mol. Cell. Biol.* 20, 4505–4512 (2000).

Website

- 101 Protagen
www.protagen.de
(Viewed January 2005)

Affiliations

- *Ulrike Korf, PhD*
DKFZ Heidelberg, Division of Molecular Genome Analysis (B050), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
Tel.: +49 622 142 4765
Fax: +49 622 142 3454
U.Korf@DKFZ.de

- *Stefan Wiemann, PhD*
DKFZ Heidelberg, Division of Molecular Genome Analysis (B050), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
Tel.: +49 622 142 4702
Fax: +49 622 142 3454
S.Wiemann@DKFZ.de