

Applications of Protein Microarray Technology

Sheng-Ce Tao, Chien-Sheng Chen and Heng Zhu*

Department of Pharmacology and Molecular Sciences/High Throughput Biology Center, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

Abstract: Protein microarrays, an emerging class of proteomic technologies, are quickly becoming essential tools for large-scale and high throughput biochemistry and molecular biology. Recent progress has been made in all the key steps of protein microarray fabrication and application, such as the large-scale cloning of expression-ready prokaryotic and eukaryotic ORFs, high throughput protein purification, surface chemistry, protein delivery systems, and detection methods. Two classes of protein microarrays are currently available: analytical and functional protein microarrays. In the case of analytical protein microarrays, well-characterized molecules with specific activity, such as antibodies, peptide-MHC complexes, or lectins, are used as immobilized probes. These arrays have become one of the most powerful multiplexed detection platforms. Functional protein microarrays are being increasingly applied to many areas of biological discovery, including drug target identification/validation and studies of protein interaction, biochemical activity, and immune responses. Great progress has been achieved in both classes of protein microarrays in terms of sensitivity and specificity, and new protein microarray technologies are continuing to emerge. Finally, protein microarrays have found novel applications in both scientific research and clinical diagnostics.

INTRODUCTION

Protein microarrays, also known as protein chips, are miniaturized, parallel assay systems that contain small amounts of purified proteins in a high-density format [1, 2]. They allow the simultaneous determination of a variety of analytes from small amounts of samples in a single experiment. Early attempts to screening large number of protein/protein fragments for biochemical activities included spotting down bacterial strains and/or bacterial lysates of protein expression cDNA libraries on nylon membranes [3, 4]; however, both suffered from fragmentation and impurity of the proteins of interest, as well as redundancy and insufficient coverage of the cDNA libraries. The major breakthrough came from a report by Zhu *et al.* in which proteome-wide protein microarrays that bear all adjectives of a protein microarray was fabricated and applied to identify calmodulin- and phospholipids-binding proteins [1]. Up to now, the Snyder group and others have demonstrated numerous applications of protein chips in various biochemical assays, including target identification of small molecules, as described later in this review.

Protein microarrays are typically prepared by immobilizing proteins onto a microscope slide using a standard contact [1, 5] or non-contact microarrayer [6-8]. A variety of slide surfaces can be used. Popular types include aldehyde- and epoxy-derivatized glass surfaces, Fullmoon slides and Schott NHS-derivatized slides for random attachment through amines [5, 9], nitrocellulose [10, 11] or gel-coated slides [12, 13] for attachment through diffusion and adsorption/absorption, and nickel-coated slides for affinity attachment of Hisx6-tagged proteins. The nickel slide was

to provide 10-fold better sensitivity than had been obtained with other random attachment methods [1]. After proteins are immobilized on the slides, they can be probed for a variety of functions/activities. The reaction signals are usually measured and recorded by means of fluorescent or radioisotope labeling. A typical protein microarray is shown in Fig. 1.

PROTEIN MICROARRAY FABRICATION

The basic concepts behind the design of protein microarrays are the same as those for DNA microarrays [14-17]: miniaturization, parallelism, and automation to achieve high throughput screening. The general procedures for the fabrication of DNA and protein chips can be divided into similar steps, including probe (protein/DNA) preparation, array substrate selection, probe (protein/DNA) delivery, sample probing, and signal detection. Many tools developed for DNA microarrays can be directly adapted to protein microarrays, such as microarray printing robots, laser-based microarray scanners, and the software needed for data analysis. In the case of DNA microarray construction, probe preparation is simple: Probes can either be synthesized (e.g., in-situ-synthesized oligonucleotide microarrays from Affymetrix, Inc. or NimbleGen Systems, Inc.) or PCR-amplified (e.g., cDNA microarrays). However, the same step is much more complicated in the case of protein chip fabrication, and especially for the so-called functional protein chips (see below), because the proteins have to be translated from mRNAs, necessitating sequential steps of gene cloning, protein expression, and protein purification in most cases. Another significant difference is the selection of microarray substrate: Unlike oligos/DNA fragments that share the same biochemical properties, proteins are broadly heterogeneous in their biochemical properties and/or activity, as well as in their size, shape, and stability. Therefore, the fabrication for protein microarray is much more challenging. Because proteins are the major driving force in carrying out cellular functions, it is clear that studies performed at the protein level can pro-

*Address correspondence to this author at the Department of Pharmacology and the HiT Center, Johns Hopkins University School of Medicine, 333 BRB, 733 N Broadway, Baltimore, MD 21205, USA; Tel: (410) 502-0878; Fax: (410) 502-1872; E-mail: heng.zhu@jhmi.edu

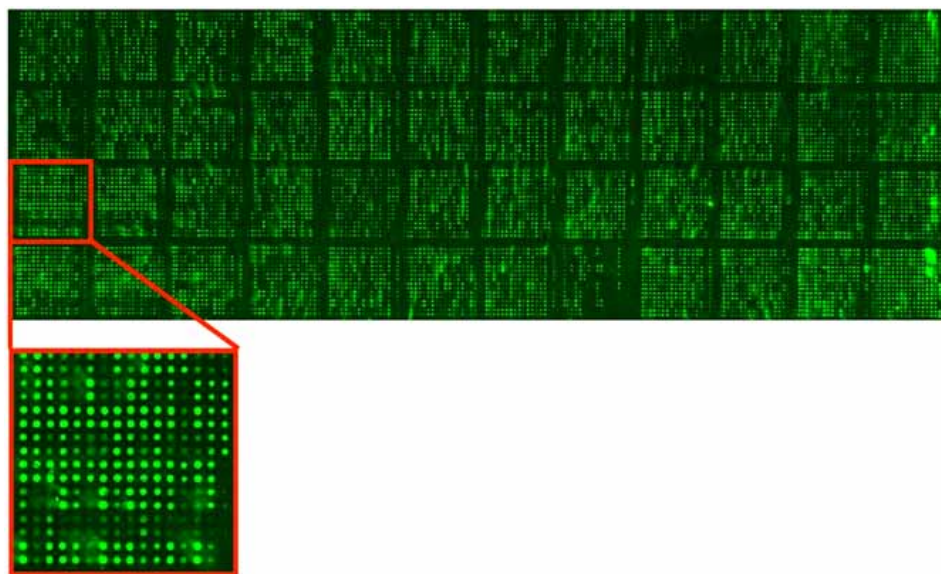


Fig. (1). Image of a typical protein microarray. A yeast protein microarray has been probed with an anti-GST antibody, followed by a Cy3-conjugated secondary antibody. An enlarged image of one of the 48 blocks is shown below the protein chip.

vide more direct information to elucidate the molecular mechanisms behind various pathways and networks. Later in this paper, we will focus on how different types of protein microarrays, especially functional proteome microarrays, have been applied to investigate a wide variety of biological functions, with applicability to studies of small molecule-protein interactions. The outcomes of these studies should be very useful in building a global network of pathways.

Expression-Ready ORF Collections and High Throughput Production of Proteins

The first two steps in fabricating a proteome microarray are the construction of proteome-wide, expression-ready ORF collections and the development of high throughput strategies for the production of a large number of proteins. In recent years, many research groups and companies have put tremendous effort into these fundamental steps.

Phizicky and colleagues first attempted to clone all 6,000 yeast ORFs into an expression vector using the yeast recombination cloning strategy and then applied this collection to screening novel biochemical activities in yeast [18]. This first ORFeome collection in yeast, however, was later found to be unstable after repetitive replications because the recombinant clones were not rescued, and thus each expression yeast strain was not clonal. Furthermore, the identity and extent of the sequence error in the recombinant clones were not checked. In 2001, Snyder and colleagues [1] applied the same cloning strategy to clone 5,800 yeast ORFs into an expression vector that over-produces N-terminal glutathione-S-transferase (GST)::Hisx6-tagged fusion proteins in yeast upon induction with galactose. Every recombinant clone was rescued into bacteria and sequenced from one end to confirm its identity and to determine whether any frame-shift mutation had occurred. The verified recombinant clones were then reintroduced into yeast for protein purification. Although this process was more laborious, it guaranteed a high-quality ORFeome collection that can be stably stored for

many generations. Later, the Snyder and Phizicky groups together constructed a C-terminal TAP tag collection in yeast by using the same strategy [19].

Based on the genomic sequence data of the *Escherichia coli* K-12 strain, Mori and coworkers [20, 21] have constructed two sets of 4,364 cloned individual genes encoding His-tagged proteins, with or without fused GFP, for use in functional genomic analysis. Each clone encodes a protein product of the predicted ORF attached to a His-tag at the N-terminal end and/or GFP at the C-terminal end. The expression of the cloned ORF is under the control of an IPTG-inducible promoter.

The total number of annotated human ORFs is expected to be about 22,000 (including 2,188 predicted) [22], and the expected number of transcripts, including splice variants, is expected to be 1.54 times that number. Based on these figures, the total number of full-length cDNA clones required to cover the entire human genome is about 33,000. Several projects are underway in both the industrial and academic spheres, with an ultimate goal of cloning all the human genes. OriGene's TrueClone Collection covers over 70% of all known loci and contains more than 24,000 unique human full-length cDNA clones. All TrueClones are in a uniform mammalian expression vector (pCMV) with 5'- and 3'-untranslated regions (UTRs). Invitrogen's Ultimate Human ORFTM collection represent 16,240 human ORFs, cloned in the Gateway[®] Entry vector, which makes it pretty handy to shuttle the inserts into any Gateway[®] Destination vectors for expression and functional analysis of the target protein in a variety of hosts, including *E. coli*, yeast, baculovirus, CHO cells, and mammalian cell lines, as well as cell-free transcription and translation coupling systems, making the collection suitable for a variety of different applications. In addition, GeneCopeia's ORFEXPRESSTM Gateway[®] Shuttle Clone product line currently contains over 16,000 unique clones derived from sequence-validated, full-length cDNA clones or human tissue cDNA libraries. In the academic set-

ting, the collection from the Mammalian Gene Collection (MGC) contains 24,759 full-length ORF human clones, encompassing 14,534 non-redundant ORFs. Vidal and colleagues recently created a collection of ORF-only clones (full-length cDNAs without UTRs) in the Gateway system; at present this collection contains 12,212 ORFs representing 10,214 unique genes [23, 24], with future updates being planned.

After the successful construction of these expression-ready collections, the next crucial step in protein chip fabrication is to express and purify these proteins simultaneously in a high throughput manner while retaining their biological activity.

Based on 1) the fact that the combination of recombinant proteins and affinity purification has been widely used to express and purify proteins from various hosts, including *E. coli*, yeast, insects, and human cell lines [1, 3, 25-27], and 2) the advantage that this combination offers almost one-step purification and the high purity of the resulting proteins, affinity purification was chosen for the high throughput production of fusion proteins. Affinity tags were fused to either C-terminal or N-terminal ends of all the collections described above: Hisx6 [20, 21], GST [1], and ZZ domains [19], respectively, as tags, and Ni-NTA-Sepharose, glutathione-Sepharose, and IgG-Sepharose as matrices for the affinity purifications. By combining affinity purification with a 96-well format that handled all the steps from culture to protein elution, and with the aid of several semi-automatic devices, thousands of proteins can be purified per day [1].

Labaer and colleagues [25] have created a system (FLEXP) that carries out all the steps from cDNA cloning to protein production (from *E. coli*) in a fully automated fashion. In a test case, ~80% of 336 random cDNA clones could be successfully purified as full-length fusion proteins. Because the protocol was automated, at least 1,000 proteins could be purified in one day. However, eukaryotic proteins expressed in *E. coli* lack their native post-translational modifications. As a result, the *in vivo* activity of these proteins may be lost. In contrast, the proteins in the yeast collections described above were expressed in their own (homologous) system, and because the proteins were in their native environment, they were subject to the usual post-translational modification and were able to interact with their natural partners. Therefore, these proteins are well-suited for biochemical analysis on the protein microarrays [1, 19, 28].

Surface Chemistry and Chip Substrate

To achieve strong and specific chip surface-protein attachment while retaining the activity and function of the proteins, several groups have created covalently reactive surfaces on glass that can covalently cross-link to proteins [1, 5, 29]. In general, a bifunctional cross-linker is used to modify the chip surface; this cross-linker has one functional group that reacts with the hydroxyl groups on the glass surface and another that can either react directly with the primary amine groups of the desired proteins (i.e., aldehyde, epoxy, or NHS-ester groups) or can be further chemically modified to achieve maximum specificity [30, 31]. A specific microarray surface can be derivatized with one functional cross-linker or a combination of different functional cross-linkers in order to

improve the protein-binding capacity, as in the case of the so-called Fullmoon slides.

The gold-coated glass surface is another variation [32, 33]. To form a self-assembled monolayer (SAM) on the gold surface, a bifunctional thio-alkylene is usually used; this reagent has an SH group that reacts with the gold and another group that reacts with capture molecules. The advantage of using a gold-coated surface is that surface plasmon resonance (SPR) and mass spectrometry can potentially be integrated as detection methods to monitor the dynamics of the reaction and to identify the captured molecules from a mixture, respectively [32-34]. This approach offers the opportunity to study the dynamics of biochemical reactions in a high throughput fashion and has great potential for use in drug and drug-target discovery and biomedical research.

Although most groups are currently adhering proteins to the microarray surface in a random fashion through covalent binding or physical attachment, it is technically feasible to attach proteins using affinity tags. Affinity attachment is likely to result in a higher proportion of active protein and homogeneous attachment of all the immobilized protein molecules [1,32,33,35]. Proteins fused with a high-affinity tag at their C- or N- terminus are linked to the surface of the chip through this tag; hence, all of the attached proteins should be uniformly oriented away from the surface [1, 19, 32, 33]. When this method is used, immobilized proteins are more likely to remain in their native conformation, and the analytes have easier access to the active sites of the proteins. Several affinity tag systems have already been developed and applied to the fabrication of protein microarrays. For example, the affinity binding system involving a His-tag and nickel has been used to attach 5,800 fusion proteins containing a His tag onto a nickel-coated glass slide [1, 19]. Recently, Taussig and coworkers [36] designed a double-His-tag sequence, comprising two His-tags separated by an 11-amino acid spacer, which shows at least 1 order of magnitude stronger binding to Ni-NTA-modified surfaces than a conventional single His-tag or two single-His-tags at the N- and C-termini and may further increase the usefulness of this approach. Other affinity tags may also be adopted for protein microarray fabrication.

A variety of types of microarray substrates have been designed, including 3D surface structures and plain glass chips. Polyacrylamide gel packet and agarose thin film microarrays, prepared by using photolithography technology on a glass surface, have been created by Guschin *et al.* [37] and Afanassiev *et al.* [38], respectively. Because both acrylamide and agarose form highly porous and hydrophilic matrices, capture molecules such as DNA, proteins, and antibodies can readily diffuse into the porous structure and are immobilized by cross-linking to the reactive ligands modified in the matrices. Biochemical assays can then be carried out by adding analytes to these 3D arrays. However, the processes involved in making these 3D substrates are sophisticated.

A simpler way to make the 3D matrices is to place a nitrocellulose membrane (widely used for protein binding) on top of a solid support, usually a glass slide. This type of substrate is commercially available and has been used, for example, in the SuperNitro microarray substrate from TeleChem Inc. The traditional nitrocellulose membrane is

quite thick and prone to producing high background in protein microarray applications. To reduce the background and improve the sensitivity, other slides coated with a much thinner membrane have been developed, including the FastTM slide, the Fullmoon C slide, and the PathTM slide. Because of the formation of 3D matrices on the glass surface, the capacity for protein immobilization is much higher than that on a 2D surface; the homogeneous water environment also minimizes protein denaturation and thereby helps keep proteins in an active state. However, in addition to the sophisticated processes required to create such 3D matrixes, a major disadvantage of the 3D arrays is the higher level of difficulty involved in changing buffers and recovering trapped molecules from the matrix microarrays [39].

Protein Delivery

The standard protein delivery system used for protein microarrays is the same as that used for DNA microarrays: the high throughput printing robot [17], which can deliver a tiny amount of protein onto the appropriate substrate in an addressable fashion. In order to maintain the native conformation and activity of the printed protein, relatively high humidity is required, and cold room printing is preferred. There are two main types of robotic printers, the contact printer and the ink-jet printer. The contact printer delivers sub-nanoliter sample volumes directly to the surface by using pins with tiny tips, with or without capillary quills. The TeleChem StealthTM steel pin is very popular because of its flexibility and accuracy. However, there are several disadvantages associated with robotic systems equipped with these pins: carry-over contamination is still an issue, and handling of the pins is a big challenge because they are very fragile.

To improve the quality of the contact printing, the so-called silicon microarrayTM technology has been developed. Instead of the pins being fabricated individually, the silicon pins are fabricated in parallel on a single crystal silicon wafer by using high-precision micromachining. This process improves pin-to-pin uniformity, eliminates pre-spotting, and reduces dead volume, carryover, and "missing spot" phenomena. Non-contact robotic printers, which use ink-jet technology, have been used to deposit nanoliter-to-picoliter protein droplets onto glass slides, polyacrylamide gel packets [40], and nanowells [29]. Although the current Packard ink-jet microarrayer can be slower when many different samples are spotted, and the shearing force during drop formation can damage some samples [41], the advantage of ink-jet printing is that it is not restricted in terms of the surface structure and is well suited for more complicated biochemical assays. The disadvantages of the ink-jet microarrayer approach are that a larger protein volume is required, and the size of the spots usually is bigger than those produced by a contact printer. Recently, electrospray deposition technology has been applied to delivering dry proteins onto a dextran-grafted surface [42]. This technology further reduced the spot size from ~150 μm to ~30 μm .

Detection Methods

Fluorescence detection-based methods are generally preferred for detection because they are simple, safe, sensitive

and offer high resolution [35]. Typically, a chip is either directly probed with a fluorescently labeled molecule or is probed in two steps, either by first using a tagged probe (e.g., biotin) that can then be detected in a second step using a fluorescent affinity reagent (e.g., streptavidin), or by using specific fluorescent antibodies. Other detection methods can also be used. For example, ELISAs have been used to detect proteins on both nitrocellulose-coated arrays [3, 43] and glass arrays [37]. Ge *et al.* [44], and Zhu *et al.* [1], respectively, have used radioisotope labeling to study protein-protein, protein-DNA, and protein-drug interactions on nitrocellulose-coated arrays, as well as kinase-substrate interactions in nanowells.

Because labeling molecules can sometimes affect protein activity and are restricted to the available detection channels, non-labeling methods have an advantage as a direct detection approach for antibody microarrays. Surface-enhanced laser desorption/ionization (SELDI) mass spectrometry has been used to detect low-density arrays of captured proteins [45]. The captured proteins on a metal surface array (SELDI protein array) were vaporized using a laser beam, followed by the analysis of mass spectrometry data to reveal the identities of the proteins. In addition to being label-free, another major advantage of the combined MS-protein microarray approach is that a mixture of samples can be probed on the slide all at the same time. Also, MS can quantitatively decode the bound molecules one by one, a decided advantage for the study of protein-small molecule interactions, including the process of drug screening [46]. A second approach, the atomic force microscopy (AFM) method, takes advantage of surface topological changes to identify the captured proteins on an antibody array [8]. When the immobilized rabbit IgG on a gold surface binds its complementary antibody, goat anti-rabbit IgG, AFM can detect the height increase and therefore reveal the binding activities.

To study the kinetics of antigen-antibody interactions, however, real-time detection methods are useful. SPR has matured as a versatile detection tool for studying the kinetics of receptor-ligand interactions with a wide range of molecular weights, affinities, and binding rates [47-49]. Efforts have been made in both the academic and industrial settings to explore SPR's potential application to the field of protein microarrays. Myszkowski and Rich [50] have described a sensor surface with 64 individual immobilization sites in a single flow cell. Biacore has two SPR systems available for array users: the Flexchip and the A100. The Flexchip platform allows 400 probes to be spotted on a slide and then screened against a single sample. In contrast, the A100 system is limited to 20 immobilized probes. When combined with a parallel flow system, however, it can perform up to four simultaneous screens with large numbers of samples. Alternatively, Sapsford *et al.* [51] have developed an antibody array biosensor to study the kinetics of antigen binding using a planar waveguide as the detection method. More importantly, they have demonstrated that significant signal intensity can be obtained from spots as small as 200 μm in diameter. The latter approach is therefore expected to be well suited for high throughput and parallel kinetics studies. Another relatively new technological approach is illustrated by the arrays based on microcantilevers that have been developed by Protiveris. Microcantilevers are extremely sensitive nano-

mechanical surface stress and mass sensors and respond to the minute changes in mass and stress through their deflection and resonance behaviors. They can be used to detect both the molecular-scale stresses that are caused by molecular recognition of biomolecules or the adsorption of chemicals, as well as the change in mass resulting from these binding events.

ANALYTICAL PROTEIN MICROARRAYS

Protein microarrays can be classified into two types, depending on their application: analytical protein microarrays and functional protein microarrays. Analytical protein microarrays use well-characterized molecules that have known specific activities as immobilized probes, such as antibodies, peptide-MHC complexes, or lectins. They have become one of the most powerful multiplexed detection platforms and can be used for protein expression monitoring, biomarker identification, cell surface marker/glycosylation profiling, clinical diagnosis, or environmental/food safety analysis. Traditionally, many different proteins, or even the total proteome of an organism, are spotted individually on a functional protein microarray. Often these immobilized proteins are not well characterized. Functional protein microarrays are mainly used to screen for various types of protein activities, including protein-protein, protein-lipid, protein-DNA, protein-drug and protein-peptide interactions; to identify enzyme substrates; and to profile immune responses.

Applications of both the analytical and functional protein microarrays are shown in Fig. 2. In the following sections, we will provide examples of various applications of both types of microarray, focusing mainly on functional protein microarrays. We will also address the development of some

new protein microarray technologies.

Perhaps the most representative class of analytical microarrays is the antibody microarray, in which antibodies are arrayed at a high density on glass surfaces. The biggest challenge associated with antibody microarrays is the production of antibodies that are able to identify the proteins of interest with high specificity and affinity in a high throughput fashion. Because the traditional method for generating monoclonal antibodies is time-consuming and laborious, researchers have recently sought alternative approaches. For example, phage antibody-display, ribosome display, systematic evolution of ligands by exponential enrichment (SELEX), mRNA display, and affibody display have been developed to expedite the production of antibodies with high specificity [35, 41, 52, 53]. All of these methods involve the construction of large library of viable regions with potential binding activity, which can then be selected by multiple rounds of affinity purification. The binding affinity of the resulting candidate clones can be further improved using maturation strategies. However, the ideal selection system has yet to be fully developed: one that is not only fast, robust, sensitive, and of low cost but also automated and minimized.

Despite the challenge involved in obtaining specific antibodies, many studies using antibody microarrays have recently been reported. In a pioneer study by Haab and colleagues [54], the first high-density antibody microarrays were used to test whether a linear relationship could be detected between an antibody and antigen pair in an array format. They investigated the ability of 115 well-characterized antibody-antigen pairs to react in high-density microarrays on modified glass slides: 30% of the pairs showed the expected linear relationships, indicating that a fraction of the antibodies were suitable for quantitative analysis. Sreekumar

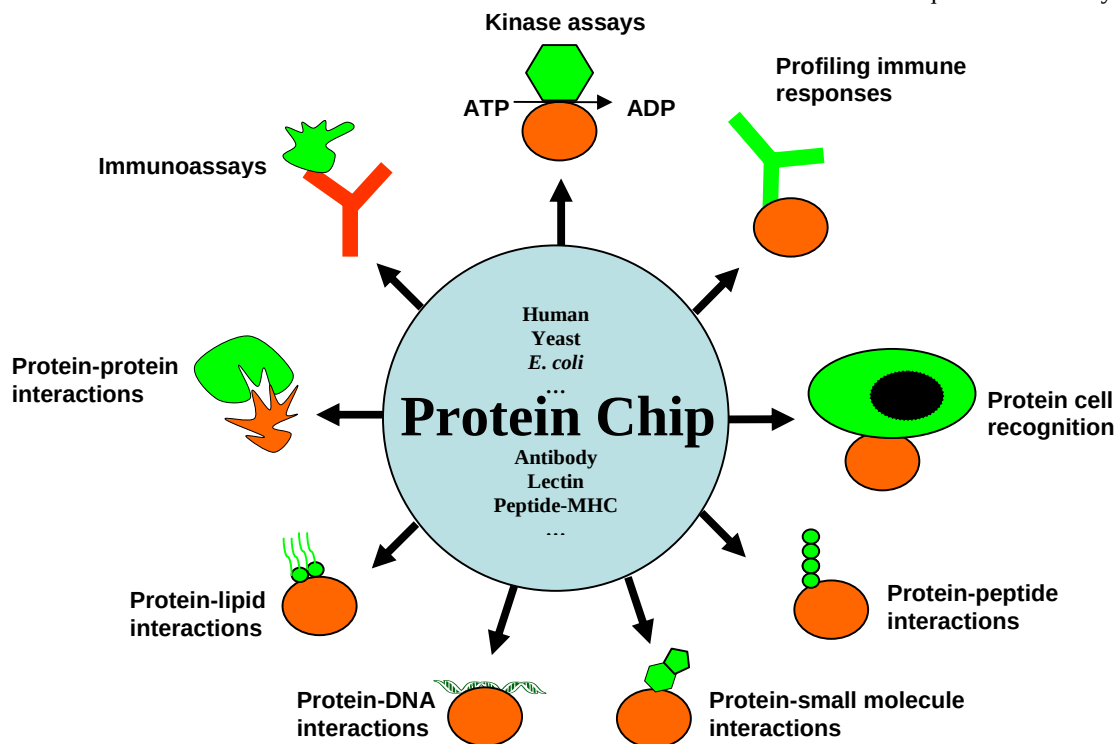


Fig. (2). Applications of protein microarrays. Antibody arrays can be used for clinical diagnosis or environmental/food safety analysis. Functional protein arrays are mainly used to study various types of protein activity, including protein-protein, protein-lipid, protein-DNA, protein-drug, protein-peptide, and protein-cell interactions; to identify enzyme substrates; and to profile immune responses.

et al. [55] created antibody arrays with 146 distinct antibodies against proteins involved in the stress response, cell cycle progression, and apoptosis and used these arrays to monitor the alterations in protein quantity in LoVo colon carcinoma cells. The reference standards and samples were labeled separately using either Cy5 or Cy3 dyes, and the fluorescent signals of the bound proteins were detected with a confocal microarray scanner. The authors were able to obtain differential expression profiles, with radiation-induced up-regulation of apoptotic regulators such as p53, DNA fragmentation factors, and tumor necrosis factor-related ligand.

In order to increase the affinity and specificity, analytical microarrays usually employ a signal amplification system and sandwich assay format, in which the first antibody is spotted on the array and then a captured antigen on the chip is detected with a second antibody that recognizes a different part of the antigen. An antibody microarray system combining both methods has been shown to be capable of simultaneously detecting 75 cytokines with high specificity, femtomolar sensitivity and a 3-log quantitative range, all with minimal sample consumption [56]. Although the sandwich format dramatically increases the specificity of the antigen detection, it requires at least two high-quality antibodies for each antigen that is to be detected.

One of the latest and most interesting developments in the category of analytical protein microarray is the cell surface-specific protein microarray (Fig. 3). These microarrays have the potential to elucidate the behavior of both normal and abnormal cells by allowing researchers to characterize cells from a complex mixture at the molecular level.

Soen *et al.* [57] have fabricated an analytical microarrays using peptide–MHC complexes to detect and characterize antigen-specific T-cell populations, an area of knowledge that is critical for our understanding of the development and physiology of the immune system and its responses in health and disease. Phycoerythrin (PE)-labeled peptide–MHC tetramers with the MHC class I (H-2k^b) murine antigen OVA and MHC class II (I-E^k) murine antigen moth cytochrome C (MCC) were printed onto polyacrylamide film-coated glass slides using a piezo-electric non-contact arrayer. The microarrays were then directly probed with T cells, and after

incubation and wash steps, binding between the antibody and the MHC tetramer spots was observed. To further define the cell-binding specificity, OVA-specific OT-1 and MCC-specific 5c.c7 lymphocytes (6×10^5) were mixed at a 1:1

ratio and added to the preprinted peptide–MHC array: While anti-CD3 and anti-CD28 mAb spots captured both lymphocyte populations, the OVA or MCC tetramers captured only the OVA-specific or MCC-specific cells, respectively. The authors also showed that this peptide–MHC microarray could be used to detect particular antigen-specific or rare T cells within a complex population. This study is the first to report the successful development of cellular microarrays with single-cell resolution.

By co-spotting mixtures of capture molecules, detector probes, and effectors in a microarray format, Chen *et al.* have developed a new analytical protein microarray [58]. Blood, lymph nodes, and surgical biopsies from patients with malignancies were dissociated into single-cell suspensions and tested on this microarray. In addition to offering rapid molecular and functional profiling, this analytical microarray was also able to detect the presence of unexpected therapeutic targets, such as over-expression of *c-kit*, on the surface of tumor cells from some of the clinical samples.

Kuschel and coworkers [59] have generated an extracellular matrix (ECM) protein microarray by using a piezoelectric microarray printer to print ECM protein solutions onto nitrocellulose-coated glass slides. Five thousand or fewer cells could be applied to an array of 4 x 4 mm, consisting of 64 microspots. By using this ECM microarray, differences in the adhesive properties of three cell lines to 14 different ECM proteins were identified.

Glycosylation of bacterial cell surfaces is a critical factor in symbiosis, pathogenesis, cell-cell interactions, and immune evasion. Hsu *et al.* have fabricated a lectin chip with 21 lectins for use in profiling the surface lipopolysaccharides (LPS) in bacterial cells [60]. When labeled *E. coli* cells were incubated on the lectin chips, the lectins were able to capture the bacterial cells onto the chip surfaces *via* the interactions with LPSs attached to the repeating sugar units, the O-antigens, on the bacteria. Two closely related *E. coli* strains,

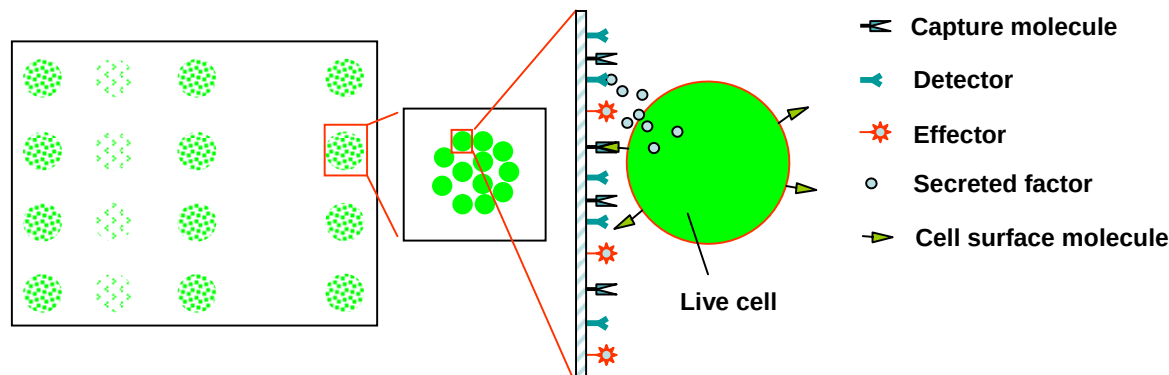


Fig. (3). Analytical protein microarrays for live cell studies. A cartoon of a typical analytical protein microarray after live cell binding is shown. Each spot is printed with a single reagent or a mixture of multiple reagents. The classes of reagents used to print analytical protein microarrays include capture molecules, detectors, and effectors. Capture probes recognize cell-surface molecules, detection probes recognize molecules secreted by cells, and effector probes can signal a change in cell behavior.

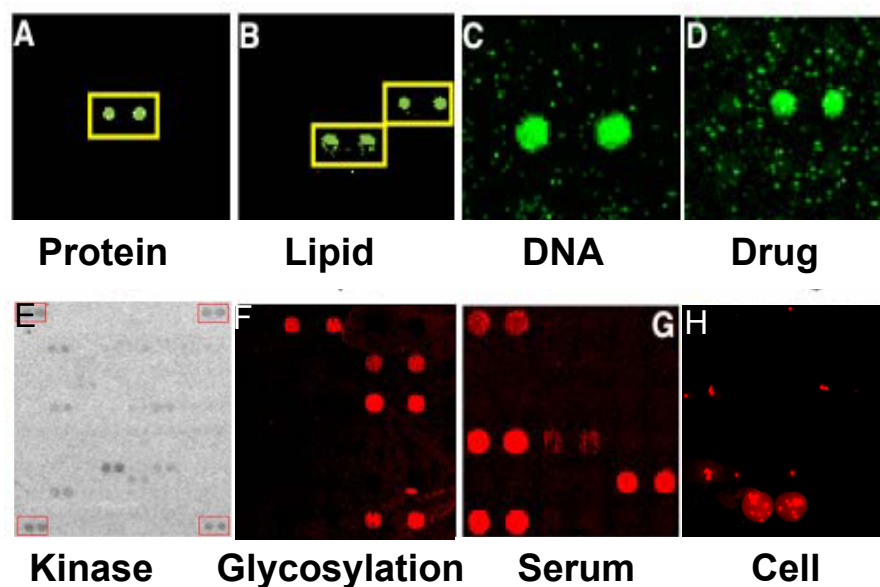


Fig. (4). Examples of different assays on functional protein microarrays. Different types of biochemical assays were carried out on chips, including assays of protein-protein (A), protein-DNA (C), protein-cell (H), protein-lipid (B), and protein-drug (D) interactions. The chips can also be used to monitor immune responses in patients (G) or posttranslational modifications of proteins, such as phosphorylation (E) and glycosylation (F). These assays achieved high signal-to-noise ratios and were very informative in terms of characterizing previously uncharacterized genes.

JM109 and HB101, could be distinguished by their differences in cell-surface glycosylation. In addition, dynamic changes in the surface glycosylation of the neonatal meningitis-associated *E. coli* strain RS218 were observed.

FUNCTIONAL PROTEIN MICROARRAYS

A variety of types of functional microarray have been developed, such as the human protein microarray [61], yeast proteome microarray [1, 19], and *E. coli* proteome microarray (unpublished). These microarrays, and in particular yeast proteome microarrays, have recently been applied to many aspects of discovery-based biology, including protein-protein, protein-lipid, protein-DNA, protein-drug, and protein-peptide interactions. Although we have attempted to describe all the major developments and applications of functional protein microarrays, it is impossible to cover all the instances in which they have been explored. Therefore, we will mainly focus here on examples of yeast proteome microarrays (Fig. 4).

Protein-Protein and Protein-Lipid Interactions

Zhu *et al.* [1] reported the construction and application of the first proteome microarrays, which contained >5,800 individually purified yeast proteins, or 85% of the yeast proteome. These proteome chips were first used to study protein-protein interactions, with the chips being incubated with biotinylated calmodulin in order to identify new binding partners. In addition, protein microarrays have also been used to examine interactions with various phospholipids that are known to act as secondary messengers. When biotinylated liposomes containing various phospholipids of interest were used as probes, more than 150 phospholipid binding proteins were identified, and a wide range of proteins was found to bind to the lipid vesicles; over 50% of these were previously known to be associated with membranes.

Protein-Small Molecule Interactions

One of the major hurdles in the pharmaceutical industry is to discover new drugs and to fully understand the molecular mechanism(s) of a lead compound coming out of a large-scale screening process. Protein microarrays also have great potential for use in drug discovery and the identification of drug targets. When performed on a proteome chip, the binding profile of a drug of interest can be simultaneously obtained across an entire proteome and allows for monitoring its specificity and/or off-target effects. The results not only provide the potential target identities of small molecules, but also serve as readouts for improving drug design by eliminating unwanted binding activities [62].

Although the design of the small molecule-binding assays is straightforward, the detection is not. To effectively measure the binding signals of small molecules on a protein chip, two commonly used labeling methods are illustrated in Fig. 5A,B. In the first methods, the small molecule of interest is first cross-linked to a carrier protein (BSA in this case), which is then labeled with fluorophores to saturation. The major benefit of this approach is that more fluorophores per small molecule can be achieved through the carrier protein. However, adding a bulky protein/fluorophore moiety might reduce the binding affinity and/or introduce nonspecific binding events. For the second approach, a fluorophore or tag is directly tethered to the small molecule *via* a flexible linker. With a less bulky tag, it is expected that the activity of the small molecule is less affected by the tag. As demonstrated by Huang *et al.* [63], two biotinylated small-molecule inhibitors of rapamycin (SMIRs) were applied to the yeast proteome chips to identify their targets that might function in the rapamycin (TOR) signaling network. They identified candidate drug targets of the SMIRs and validated a previously unknown protein as the target of the SMIRs. Interest-

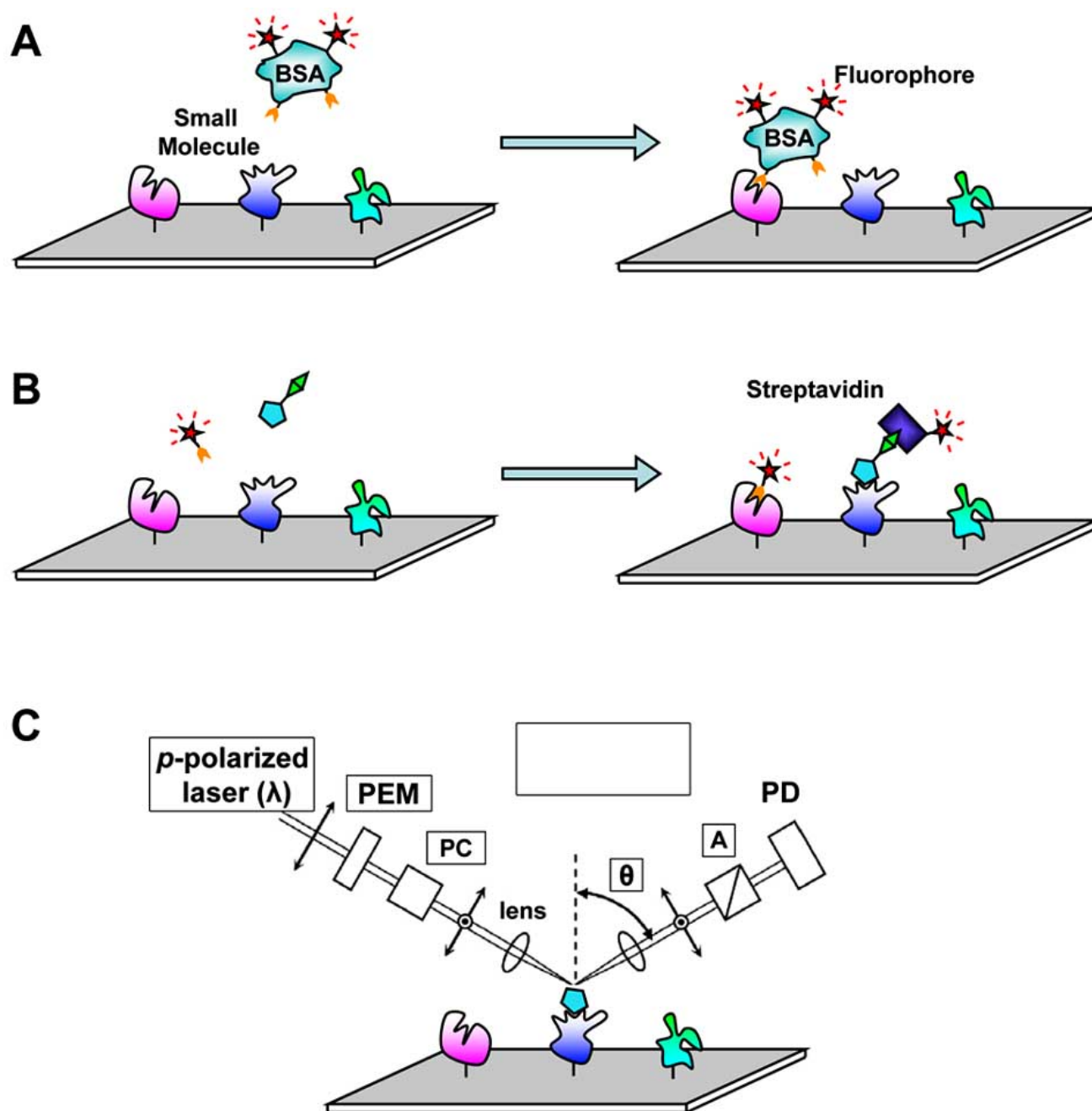


Fig. (5). Drug target identification on protein chips. **(A)** A small molecule can be tethered to a carrier protein, which is then labeled with fluorescent dyes as detection tags. **(B)** Alternatively, a small molecule can also be directly tethered to a fluorophore or a tag, such as biotin. When biotin is used, a second step of incubation with fluorescently labeled streptavidin is required for the detection. **(C)** The OI-RD microscope imaging system. As a label-free system, parallel real-time events can be monitored and recorded. PEM, photoelastic modulator; PC, Pockels cell; A, analyzer; PD, photodiode.

ingly, in an independent study, Claudio De Virgilio and colleagues [64] also identified the same protein in the TOR signaling pathway using a different approach.

An interesting twist of screening protein-small molecule interactions was reported recently by Kim *et al.* [65]. The authors employed a protein microarray approach for exploring ligand-initiated nuclear hormone receptor (NHRs) pharmacology, receptor specificity, and heterodimer functional-

ity. When NHRs bind their cognate hormones or hormone analogs, they adopt distinctive conformations that enable or disable the binding of coregulator proteins in response to changes in circulating levels of hormones. By spotting down the estrogen and thyroid hormone receptors, they managed to monitor the ability of various ligands to increase or decrease the activation state of ER homo- and/or heterodimer ligand binding domains by quantifying the interaction of the NHR-ligand complex with different coactivators. In addition, these arrays were used to screen for synthetic compounds that reduce ER activity by either competing with agonist ligands or with coactivator proteins, compounds that could lead to improved treatment strategies of hormone-sensitive or -resistant breast carcinomas.

The major disadvantages of using labeled small molecules in a binding assay on protein chips are: 1) most of the existing small molecule libraries and FDA-approved drugs do not have any labels or tags; 2) adding a tag to an existing compound requires *de novo* synthesis of the compound; and 3) Labeling reagents sometimes affect activity of the drugs. Therefore, label-free detection has advantages as a direct detection approach for protein microarrays. Also, non-label methods use simpler protocols and can provide real-time measurement. In addition to the SPR system described above, the oblique-incidence optical reflectivity difference (OI-RD) microscope has gained some heat in recent years as an alternative real-time, label-free detection system as demonstrated by Zhu and colleagues (Fig. 5C) [66]. One of the major advantages of this new system is its flexibility with surface chemistry—any types of transparent surfaces can be used. Because both systems can detect signals from protein spots smaller than 300 microns, they are well suited for high throughput, real time kinetics studies.

Protein-DNA Interactions

The identification of DNA-binding proteins and the DNA elements to which they bind is a fundamental problem in biology. Such protein-DNA binding information is necessary for uncovering gene regulatory networks that control cellular and developmental processes. In a later report, Snyder's group also used yeast proteome chips to screen for novel DNA binding activities using fluorescently labeled yeast genomic DNA [28]. A total of 200 proteins that reproducibly bound DNA were identified. Half of these had not previously been shown to have DNA-binding activity; these new proteins fell into a wide variety of functional categories. The most surprising discovery in this category was the identification of Arg5,6 as a DNA-binding protein. The *ARG5,6* gene encodes two mitochondrial enzymes that mediate two key steps in the biosynthesis of ornithine (a precursor to arginine). Follow-up experiments revealed that these enzymes associate with specific mitochondrial loci *in vivo*, and this information was used to define a DNA-binding motif for these proteins. Thus, a novel DNA binding activity was found to be associated with a well-characterized protein, thereby identifying a novel function for that protein.

Recently, Snyder and colleagues have constructed a protein chip with 282 known and potential yeast transcriptional factors [67]. Forty Cy3-end-labeled ds-DNA oligos, containing 75 DNA motifs previously identified on the basis of their

evolutionary conservation, were probed with the protein chips. In all, 211 specific DNA-protein interactions were identified, 131 of them are previously unknown. Yjl103, a previously uncharacterized DNA-binding protein, was characterized, and its binding site on DNA was identified as two

direct repeats of CGG separated by eight or nine nucleotides. Among the 22 genes tested *in vivo*, Yjl103 was found to be able to bind 19 of them upstream, and all 22 contained the CGGN8CGG motif. In many cases, the known or predicted functions of the proteins encoded by these genes are primarily associated with the metabolism of carbon compounds and carbohydrates, a finding that is consistent with the proposed role of Yjl103 in energy utilization.

Protein-Peptide Interactions

In a very recent report, Jones and colleagues [68] demonstrated that they could measure quantitative interactions between proteins and peptides in an array format. They cloned, expressed, and purified almost all the human Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains. They then printed the 159 domains on the aldehyde-modified glass substrates and incubated 61 peptides representing physiological sites of tyrosine phosphorylation on the four ErbB receptors with the protein chips. For quantitative measurement, eight concentrations of each peptide, ranging from 10 nM to 5 μ M, were used in the assay, allowing the binding affinity of each peptide to be measured. With this microarray, 43 of the 65 previously reported interactions were detected, and 116 new interactions were identified. Also, ErbB1 and ErbB2 were found to become more promiscuous with increasing concentration, whereas ErbB3 did not. Because ErbB1 and ErbB2 are overexpressed in many human cancers, the authors suggested that this potential for increased promiscuity might contribute to the oncogenic potential of receptor tyrosine kinases.

Identification of Kinase Substrates on Protein Chips

Since phosphorylation is known to be involved in almost every aspect of cell processes, identification of the downstream substrates of protein kinases is a critical step in understanding the effects of phosphorylation on protein functions. In order to demonstrate that the protein chip approach is suitable for such investigations, Zhu *et al.* [29] first analyzed the substrate specificity of 119 yeast kinases on 17 different substrates using nanowell protein chips. Recently, as an extension of the same idea but on a much larger scale, the so-called "Phosphorylome Project" was accomplished using yeast proteome microarrays [69]. The goal was to identify all the potential protein substrates of each yeast kinase. *In vitro* kinase reactions were carried out on the yeast proteome chips using 87 individually purified kinases/kinase complexes in the presence of ^{33}P -ATP. In all, 4,129 phosphorylation events involving 1,325 different proteins were identified. To ensure that the signals resulted from phosphorylation events, 5% SDS was used to denature the proteins on the chips to remove the signals from binding of kinase proteins or ^{33}P -ATP. These phosphorylation results have been assembled into a first-generation, global kinase signaling network in yeast.

Protein Glycosylation Analysis

Glycosylation is one of the most important protein post-translational modifications and is highly related to the regulation and function of many proteins. Although 20% to 50% of eukaryotic proteins are predicted to be glycosylated, there are only a few proteins with known carbohydrates linkages. To

evaluate the feasibility of using protein microarrays to study protein glycosylation on the proteome scale, Gelperin and coworkers [19] constructed the so-called "MORF" yeast proteome microarray. This microarray consisted of 5,573 C-terminally tagged yeast proteins, which were purified from yeast. A polyclonal antibody that recognized yeast glycans was incubated with the proteome chip, resulting in the identification of 599 putative glycosylated proteins. The chip results were validated by gel-shift assays using proteins treated with the deglycosylation enzymes EndoH and PNGase. Of the 49 known glycoproteins, 33 exhibited a mobility shift after deglycosylation treatment, including 21 of the 25 known N-linked glycoproteins. Among 344 candidates, 109 new glycoproteins were confirmed, nearly twice the number of previously identified N-linked glycoproteins in yeast.

Profiling Immune Responses

The microarray-based identification of the autoantigens targeted by autoantibodies during the immune response has considerable potential for use in diagnosis, classification, and prognosis [61]. Robinson and colleagues [70] published the first report of the simultaneous analysis of multiple human disease sera *via* protein microarray. They arrayed 196 distinct biomolecules involved in eight distinct human autoimmune diseases, including proteins, peptides, enzyme complexes, ribonucleoprotein complexes, DNA, and post-translationally modified antigens, onto glass slides to form the autoantigen microarrays. These arrays were incubated with patient serum samples as a means of defining the pathogenesis of autoantibody responses in human autoimmune diseases. Recently, Cahill and colleagues [71] constructed a protein array consisting of polypeptides translated from 37,200 random human cDNA clones in *E. coli* and used this array to identify potential autoantigens involved in the pathogenesis of alopecia areata. Eight autoantigens were identified and successfully confirmed by Western blotting. Likewise, the human protein chips from the same group [61] were incubated with sera from alopecia and arthritis patients. Sera from clinically non-remarkable persons were used as controls. High intensities for sera from diffuse alopecia and arthritis were obtained for autoantigen p69. The sera of patients with rheumatoid arthritis recognized a protein of the RAS association domain family 1 and diglycerol kinase ζ . The alopecia serum recognized two additional proteins, the tumor suppressor p33 ING homologue and a protein with unknown function.

Zhu *et al.* [72] have also fabricated protein chips that allowed them to rapidly and sensitively distinguish the immune responses of SARS-infected and healthy people. These protein chips harbored all the SARS-CoV proteins as well as proteins from five additional coronaviruses that can infect humans (HCoV-229E and HCoV-OC43), cows (BCV), cats (FIPV), and mice (MHVA59). The presence of human IgG and IgM antibodies against SARS-CoV was detected on the

chips with labeled anti-human IgG and IgM antibodies. Sera from patients could quickly be clustered as SARS-positive or SARS-negative on the basis of the serum-probing signals. Comparison to other methods (e.g., ELISA and IFA) indicated that the origin of 94% of the sera could be correctly predicted on the microarrays. The chip-based assay was at least 100-fold more sensitive than the ELISA/IFA assays and required a smaller amount of sample.

DEVELOPMENT OF NEW PROTEIN MICROARRAY TECHNOLOGIES

As indicated above, tremendous efforts have been made during the past few years to develop a variety of different analytical and functional protein microarrays as well as to explore their applications. In addition to these, there are important protein microarray technologies that cannot be simply classified as either analytical or functional. Here we describe two very interesting examples: the so-called "reverse-phase protein microarray", which can be used to analyze hundreds to thousands of samples simultaneously on a single chip [73], and the new protein microarray fabrication technologies that make it possible to generate protein microarrays directly on the glass slides at a low cost in a highly efficient fashion [73-75].

Reverse-Phase Protein Microarrays

In contrast to previous protein arrays that immobilize the probe, Liotta and coworkers [76] have developed a so-called "reverse-phase" protein array that relies on immobilized lysates, which represent the state of individual tissue cell populations undergoing disease transitions. Using the reverse protein microarray, they analyzed the pro-survival checkpoint proteins in patient-matched samples at the microscopic transition stage from histologically normal prostate epithelium to prostate intraepithelial neoplasia (PIN) and then to invasive prostate cancer. Cancer progression was associated with an increased phosphorylation of Akt, suppression of apoptosis pathways, and decreased phosphorylation of ERK. At the transition from histologically normal epithelium to PIN, a significant increase in phosphorylated Akt and a concomitant suppression of downstream apoptosis pathways were observed, preceding the transition into invasive carcinoma. Use of the same strategy, revealed two promising diagnostic markers (villin and moesin) for distinguishing colon from ovarian adenocarcinomas [77]. Similarly, the same group constructed another reverse protein microarray [78] on which the lysates represented 60 human cancer cell lines (NCI-60). Fifty-two mouse monoclonal antibodies were probed individually to test the expression of 52 proteins on the microarrays. By using this reverse-protein microarray, the authors found that cell-structure-related proteins almost invariably showed a high correlation between mRNA and protein levels across the NCI-60 cell lines, while non-cell-structure-related proteins showed a poor correlation.

Making Protein Microarrays Without Large-Scale Cloning and Expressions

Protein microarray technology has not been widely adopted, mainly because it is time-consuming and labor-intensive to purify the thousands of proteins required [1, 79]. The general process of protein chip fabrication currently

requires large-scale gene cloning and transformation, protein induction, and high throughput protein purification. In recent years, several strategies have been introduced in an attempt to overcome the tedious protein preparation procedure. One strategy has taken advantage of cell-free systems that have been developed to synthesize proteins *in vitro*, including the *E. coli*, wheat germ agglutinin, and rabbit reticulocyte systems. For example, the *E. coli* cell-free protein synthesis system has been used to synthesize proteins in a 96-well format [80]. After protein synthesis, a further protein purification step is still required.

One of the most important developments in protein chip fabrication is the so-called nucleic acid-programmable protein array (NAPPA) [75]. LaBaer and colleagues have immobilized plasmids that could express target proteins fused to GST using an *in vitro* transcription/translation system. The *de novo*-synthesized proteins were captured locally through binding of the GST moiety to anti-GST antibodies spotted together with the plasmid DNAs. Prior to *in vitro* transcription/translation, such chips can be easily stored because DNA and antibodies are very stable. Since the proteins are translated and immobilized locally, no complicated purification step is needed. This technology can be easily applied to expression-ready cDNA libraries as long as they express tagged fusion proteins. However, this strategy requires the cloning of cDNAs in such expression-ready vectors, and because transcription/translation reactions occur on the chip, both the nascent mRNAs and proteins are free in solution and are therefore sensitive to vibration, which can cause diffusion among the adjacent spots. Therefore, to prevent cross-contamination, the adjacent spots must be kept at a certain distance from each other, and this requirement limits the density of the features on these chips.

An alternative approach to synthesizing proteins is to use the so-called PROFusion technology, which is based on the mRNA display strategy [81-83]. A poly(dA)₂₇CCA oligo labeled with a puromycin molecule at its 3'-end is ligated to the 3'-end of an mRNA molecule, which carries a ribosome-binding site at its 5'-end. Upon incubation with an *in vitro* translation system, a ribosome translates the encoded protein until it is stalled at the RNA-DNA junction. As a result, the puromycin molecule tethered at the 3'-end of the RNA-DNA hybrid has sufficient time to enter the A site of the ribosome and form a covalent linkage to the nascent peptide before the dissociation of the ribosome mRNA complex. Therefore, the synthesized peptide is tagged with its coding sequence. Because of this unique feature, a large number of such peptides can be screened for a particular activity in a mixture, and the selected peptides can be later decoded and retrieved based on the tagged mRNA sequences using RT-PCR. Such selection can continue through multiple rounds until the desired activity is obtained [84]. Furthermore, the decoding process can be performed on an addressable microarray containing the complementary oligos to form a high-density peptide/protein microarray, the PROFusion microarray. This strategy obviates the necessity for a purification procedure, and different proteins can be translated in a single tube. However, before protein immobilization, a ligation step is required to form a RNA-DNA hybrid in order to incorporate puromycin molecules individually for every mRNA species to be translated. Because it is cumbersome to ligate single-stranded RNA and

DNA molecules and the yield is somewhat low, this approach is hard to adapt to high throughput peptide/protein production. Furthermore, in order to enhance the quantities of the proteins immobilized to a glass surface, the untranslated RNA-DNA hybrids have to be removed to eliminate any competition with RNA-tagged peptides/proteins.

Building upon these earlier approaches, our group [73] has developed two new strategies that greatly simplify the process of making peptide/protein chips using *in vitro* translation systems. Specifically, we have developed new methods to stall ribosomes during protein translation to allow for the capture of the nascent polypeptides by puromycin that is grafted at the end of an oligo immobilized on a solid surface. As a proof of concept, a peptide chip with three commonly used epitopes/peptides, namely Hisx6, FLAG, and StrepTag II, was constructed. On the peptide chip, Hisx6 and FLAG could be sensitively and specifically detected by their corresponding antibodies, while StrepTag II could only be detected by streptavidin. To demonstrate that high-density protein chips can be fabricated, Strategy 2 was applied to produce several proteins of various families with a wide range of molecular size, from 27 to 57 kDa, including GST; two protein kinases, Tpk1 and Hrr25; two transcription factors, Mcm1 and Cbf1; and one auto-fluorescent protein m-RFP. As revealed by antibody detection, all of these proteins could be translated and captured in full-length at high density. Protein-DNA motif interactions have been successfully carried out on the chip, strongly suggesting that protein chips of this variety can be used to analyze protein activity.

OUTLOOK

Protein microarray technology has been shown to be a useful tool for multiplexed detection and proteomics studies. Femtomolar sensitivity has been achieved in analytical protein microarrays, and the number of applications of functional protein microarrays has grown dramatically. Novel applications utilizing protein microarrays and new protein microarray technologies are continually emerging. However, there are still several issues that need to be resolved before protein chip technology can be widely applied. First, the traditional cloning-expressing-purification-printing approach is still the gold standard procedure for making protein microarrays, especially for proteome microarrays. Because of the sophisticated expertise required and the high cost of production, it is almost impossible for most laboratories to make their own microarrays, and the price for the commercial protein chips is unacceptably high. A variety of promising strategies have already been tested to bypass the traditional procedure. However, none of them are close to being used for the large-scale fabrication of protein microarrays. Thus, to make protein microarray technology more applicable, a simpler and more powerful strategy is needed. Second, there is no widely accepted experimental standardization protocol. Fortunately, this issue is now being investigated by the Human Proteome Organization (HUPO), which is developing guidelines for experimental design and data annotation. Third, at present most the protein microarray results are only semi-quantitative. In order to reach the goal of accurate quantification, new technology is needed and is eagerly anticipated. Though still in their infancy, protein microarrays will no doubt prove to be one of the most powerful tools in

both basic science and clinical research. Improvements in our ability to generate large sets of high-quality proteins and antibodies or their mimetics will play a key role in quantitative analysis and promote the extension of this technology to many model organisms.

ACKNOWLEDGEMENTS

We thank the NIH (U54RR020839-01) for financial support and Dr. Deborah McClellan for editorial assistance.

REFERENCES

- [1] Zhu, H.; Bilgin, M.; Bangham, R.; Hall, D.; Casamayor, A.; Bertone, P.; Lan, N.; Jansen, R.; Bidlingmaier, S.; Houfek, T.; Mitchell, T.; Miller, P.; Dean, R.A.; Gerstein, M.; Snyder, M. *Science*, **2001**, 293, 2101.
- [2] Chen, C.S.; Zhu, H. *Biotechniques*, **2006**, 40, 423.
- [3] Bussow, K.; Cahill, D.; Nietfeld, W.; Bancroft, D.; Scherzinger, E.; Lehrach, H.; Walter, G. *Nucleic Acids Res.*, **1998**, 26, 5007.
- [4] Lueking, A.; Horn, M.; Eickhoff, H.; Bussow, K.; Lehrach, H.; Walter, G. *Anal. Biochem.*, **1999**, 270, 103.
- [5] MacBeath, G.; Schreiber, S.L. *Science*, **2001**, 289, 1760.
- [6] Delehanty, J.B.; Ligler, F.S. *Biotechniques*, **2003**, 34, 380.
- [7] Delehanty, J.B. *Methods Mol. Biol.*, **2004**, 264, 135.
- [8] Jones, V.W.; Kenseth, J.R.; Porter, M.D.; Mosher, C.L.; Henderson, E. *Anal. Chem.*, **1998**, 70, 1233.
- [9] Kusnezow, W.; Jacob, A.; Walijew, A.; Diehl, F.; Hoheisel, J.D. *Proteomics*, **2003**, 3, 254.
- [10] Kramer, A.; Feilner, T.; Possling, A.; Radchuk, V.; Weschke, W.; Burkle, L.; Kersten, B. *Phytochemistry*, **2004**, 65, 1777.
- [11] Stillman, B.A.; Tonkinson, J.L. *Biotechniques*, **2000**, 29, 630 (2000).
- [12] Angenendt, P.; Glokler, J.; Murphy, D.; Lehrach, H.; Cahill, D.J. *Anal. Biochem.*, **2002**, 309, 253.
- [13] Charles, P.T.; Goldman, E.R.; Rangasamy, J.G.; Schauer, C.L.; Chen, M.S.; Tait, C.R. *Biosens. Bioelectron.*, **2004**, 20, 753.
- [14] Fodor, S.P.; Read, J.L.; Pirrung, M.C.; Stryer, L.; Lu, A.T.; Solas, D. *Science*, **1991**, 251, 767.
- [15] Schena, M.; Shalon, D.; Heller, R.; Chai, A.; Brown, P.O.; Davis, R.W. *Proc. Natl. Acad. Sci. USA*, **1996**, 93, 106149.
- [16] Derisi, J.; Penland, L.; Brown, P.O.; Bittner, M.L.; Meltzer, P.S.; Ray, M.; Chen, Y.; Su, Y.A.; Trent, J.M. *Nat. Genet.*, **1996**, 14, 457.
- [17] Schena, M.; Shalon, D.; Davis, R.W.; Brown, P.O. *Science*, **1995**, 270, 467.
- [18] Martzen, M.R.; McCraith, S.M.; Spinelli, S.L.; Torres, F.M.; Fields, S.; Grayhack, E.J.; Phizicky, E.M. *Science*, **1999**, 286, 1153.
- [19] Gelperin, D.M.; White, M.A.; Wilkinson, M.L.; Kon, Y.; Kung, L.A.; Wise, K.J.; Lopez-Hoyo, N.; Jiang, L.; Piccirillo, S.; Yu, H.; Gerstein, M.; Dumont, M.E.; Phizicky, E.M.; Snyder, M.; Grayhack, E.J. *Genes Dev.*, **2005**, 19, 2816.
- [20] Kitagawa, M.; Ara, T.; Arifuzzaman, M.; Ioka-Nakamichi, T.; Inamoto, E.; Toyonaga, H.; Mori, H. *DNA Res.*, **2005**, 12, 291.
- [21] Arifuzzaman, M.; Maeda, M.; Itoh, A.; Nishikata, K.; Takita, C.; Saito, R.; Ara, T.; Nakahigashi, K.; Huang, H.C.; Hirai, A.; Tsuzuki, K.; Nakamura, S.; Altaf-Ul-Amin, M.; Oshima, T.; Baba, T.; Yamamoto, N.; Kawamura, T.; Ioka-Nakamichi, T.; Kitagawa, M.; Tomita, M.; Kanaya, S.; Wada, C.; Mori, H. *Genome Res.*, **2006**, 16, 686.
- [22] International Human Genome Sequencing Consortium Finishing the euchromatic sequence of the human genome. *Nature*, **2004**, 431, 931.
- [23] Rual, J.F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G.F.; Gibbons, F.D.; Dreze, M.; Ayivi-Guedehoussou, N.; Klitgord, N.; Simon, C.; Boxem, M.; Milstein, S.; Rosenberg, J.; Goldberg, D.S.; Zhang, L.V.; Wong, S.L.; Franklin, G.; Li, S.; Albala, J.S.; Lim, J.; Fraughton, C.; Llamasas, E.; Cevik, S.; Bex, C.; Lamesch, P.; Sikorski, R.S.; Vandenhaute, J.; Zoghbi, H.Y.; Smolyar, A.; Bosak, S.; Sequerra, R.; Doucette-Stamm, L.; Cusick, M.E.; Hill, D.E.; Roth, F.P.; Vidal, M. *Nature*, **2005**, 437, 1173.
- [24] Rual, J.F.; Hirozane-Kishikawa, T.; Hao, T.; Bertin, N.; Li, S.; Dricot, A.; Li, N.; Rosenberg, J.; Lamesch, P.; Vidalain, P.O.; Clingsmith, T.R.; Hartley, J.L.; Esposito, D.; Cheo, D.; Moore, T.; Simmons, B.; Sequerra, R.; Bosak, S.; Doucette-Stamm, L.; Le Peuch, C.; Vandenhaute, J.; Cusick, M.E.; Albala, J.S.; Hill, D.E.; Vidal, M. *Genome Res.*, **2004**, 14, 2128.
- [25] Braun, P.; Hu, Y.; Shen, B.; Halleck, A.; Koundinya, M.; Harlow, E.; LaBaer, J. *Proc. Natl. Acad. Sci. USA*, **2002**, 99, 2654.
- [26] Hall, D.A.; Ptacek, J.; Snyder, M. *Mech. Ageing Dev.*, **2007**, 128, 161.
- [27] Albala, J.S.; Franke, K.; McConnell, I.R.; Pak, K.L.; Folta, P.A.; Rubinfeld, B.; Davies, A.H.; Lennon, G.G.; Clark, R. *J. Cell Biochem.*, **2000**, 80, 187.
- [28] Hall, D.A.; Zhu, H.; Zhu, X.; Royce, T.; Gerstein, M.; Snyder, M. *Science*, **2004**, 306, 482.
- [29] Zhu, H.; Klemic, J.F.; Chang, S.; Bertone, P.; Casamayor, A.; Klemic, K.G.; Smith, D.; Gerstein, M.; Reed, M.A.; Snyder, M. *Nat. Genet.*, **2000**, 26, 283.
- [30] MacBeath, G.; Koehler, A.N.; Schreiber, S.L. *J. Am. Chem. Soc.*, **1999**, 121, 7967.
- [31] Heyssel, S.; Vogel, H.; Sanger, M.; Sigrist, H. *Protein Sci.*, **1995**, 4, 2532.
- [32] Bieri, C.; Ernst, O.P.; Heyse, S.; Hofmann, K.P.; Vogel, H. *Nat. Biotechnol.*, **1999**, 17, 1105.
- [33] Houseman, B.T.; Huh, J.H.; Kron, S.J.; Mrksich, M. *Nat. Biotechnol.*, **2002**, 20, 270.
- [34] Rich, R.L.; Day, Y.S.N.; Morton, T.A.; Myszk, D.G. *Anal. Biochem.*, **2001**, 296, 197.
- [35] Templin, M.F.; Stoll, D.; Schrenk, M.; Traub, P.C.; Vohringer, C.F.; Joos, T.O. *Trends Biotechnol.*, **2002**, 20, 160.
- [36] Khan, F.; He, M.; Taussig, M.J. *Anal. Chem.*, **2006**, 78, 3072.
- [37] Guschin, D.; Yershov, G.; Zaslavsky, A.; Gemmell, A.; Shick, V.; Proudnikov, D.; Arenkov, P.; Mirzabekov, A. *Anal. Biochem.*, **1997**, 250, 203.
- [38] Afanassiev, V.; Hanemann, V.; Wolf, S. *Nucleic Acids Res.*, **2000**, 28, E66.
- [39] Zhu, H.; Snyder, M. *Curr. Opin. Chem. Biol.*, **2003**, 7, 55.
- [40] Arenkov, P.; Kukhtin, A.; Gemmell, A.; Voloshchuk, S.; Chuppeva, V.; Mirzabekov, A. *Anal. Biochem.*, **2000**, 278, 123.
- [41] Haab, B.B. *Curr. Opin. Drug Discov. Devel.*, **2001**, 4, 116.
- [42] Avseenko, N.V.; Morozova, T.Y.; Ataulakhov, F.I.; Morozov, V.N. *Anal. Chem.*, **2002**, 74, 927.
- [43] Joos, T.O.; Schrenk, M.; Hopfl, P.; Kroger, K.; Chowdhury, U.; Stoll, D.; Schorner, D.; Durr, M.; Herick, K.; Rupp, S.; Sohn, K.; Hammerle, H. *Electrophoresis*, **2000**, 21, 2641.
- [44] Ge, H. *Nucleic Acids Res.*, **2000**, 28, e3.
- [45] Davies, H.; Lomas, L.; Austen, B. *Biotechniques*, **1999**, 27, 1258.
- [46] Finnskog, D.; Ressine, A.; Laurell, T.; Marko-Varga, G. *J. Proteome Res.*, **2004**, 3, 988.
- [47] McDonnell, J.M. *Curr. Opin. Chem. Biol.*, **2001**, 5, 572.
- [48] Salamon, Z.; Brown, M.F.; Tollin, G. *Trends Biochem. Sci.*, **1999**, 24, 213.
- [49] Nieba, L.; Nieba-Axmann, S.E.; Persson, A.; Hamalainen, M.; Edebratt, F.; Hansson, A.; Lidholm, J.; Magnusson, K.; Karlsson, A.F.; Pluckthun, A. *Anal. Biochem.*, **1997**, 252, 217.
- [50] Myszk, D.G.; Rich, R.L. *Pharm. Sci. Technol. Today*, **2000**, 3, 310.
- [51] Sapsford, K.E.; Liron, Z.; Shubin, Y.S.; Ligler, F.S. *Anal. Chem.*, **2001**, 73, 5518.
- [52] Stoll, D.; Templin, M.F.; Schrenk, M.; Traub, P.C.; Vohringer, C.F.; Joos, T.O. *Biosci.*, **2002**, 7, c13-c32.
- [53] Cahill, D.J. *J. Immunol. Methods*, **2001**, 250, 81.
- [54] Haab, B.B.; Dunham, M.J.; Brown, P.O. *Genome Biol.*, **2001**, 2, RESEARCH0004.
- [55] Sreekumar, A.; Nyati, M.K.; Varambally, S.; Barrette, T.R.; Ghosh, D.; Lawrence, T.S.; Chinnaiyan, A.M. *Cancer Res.*, **2001**, 61, 7585.
- [56] Schweitzer, B.; Roberts, S.; Grimwade, B.; Shao, W.; Wang, M.; Fu, Q.; Shu, Q.; Laroche, I.; Zhou, Z.; Tchernev, V.T.; Christiansen, J.; Velleca, M.; Kingsmore, S.F. *Nat. Biotechnol.*, **2002**, 20, 359.
- [57] Soen, Y.; Chen, D.S.; Kraft, D.L.; Davis, M.M.; Brown, P.O. *PLoS Biol.*, **2003**, 1, E65.
- [58] Chen, D.S.; Soen, Y.; Davis, M.M.; Brown, P.O. *J. Clin. Oncol.*, **2004**, 22, 9507.
- [59] Kuschel, C.; Steuer, H.; Maurer, A.N.; Kanzok, B.; Stoop, R.; Angres, B. *Biotechniques*, **2006**, 40, 523.

- [60] Hsu, K.L.; Pilobello, K.T.; Mahal, L.K. *Nat. Chem. Biol.*, **2006**, 2, 153.
- [61] Lueking, A.; Possling, A.; Huber, O.; Beveridge, A.; Horn, M.; Eickhoff, H.; Schuchardt, J.; Lehrach, H.; Cahill, D.J. *Mol. Cell Proteomics*, **2003**, 2, 1342.
- [62] Huang, Y.H.; Li, D.; Winoto, A.; Robey, E.A. *Proc. Natl. Acad. Sci. USA*, **2004**, 101, 4936.
- [63] Huang, J.; Zhu, H.; Haggarty, S.J.; Spring, D.R.; Hwang, H.; Jin, F.; Snyder, M.; Schreiber, S.L. *Proc. Natl. Acad. Sci. USA*, **2004**, 101, 165949.
- [64] Dubouloz, F.; Deloche, O.; Wanke, V.; Cameron, E.; De Virgilio, C. *Mol. Cell*, **2005**, 19, 15.
- [65] Kim, S.H.; Tamrazi, A.; Carlson, K.E.; Katzenellenbogen, J.A. *Mol. Cell Proteomics*, **2005**, 4, 267.
- [66] Landry, J.P.; Zhu, X.D.; Gregg, J.P. *Opt. Lett.*, **2004**, 29, 581.
- [67] Ho, S.W.; Jona, G.; Chen, C.T.; Johnston, M.; Snyder, M. *Proc. Natl. Acad. Sci. USA*, **2006**, 103, 9940.
- [68] Jones, R.B.; Gordus, A.; Krall, J.A.; Macbeath, G. *Nature*, **2005**, 439, 168.
- [69] Ptacek, J.; Devgan, G.; Michaud, G.; Zhu, H.; Zhu, X.; Fasolo, J.; Guo, H.; Jona, G.; Breitkreutz, A.; Sopko, R.; McCartney, R.R.; Schmidt, M.C.; Rachidi, N.; Lee, S.J.; Mah, A.S.; Meng, L.; Stark, M.J.; Stern, D.F.; De Virgilio, C.; Tyers, M.; Andrews, B.; Gerstein, M.; Schweitzer, B.; Predki, P.F.; Snyder, M. *Nature*, **2005**, 438, 679.
- [70] Robinson, W.H.; DiGennaro, C.; Hueber, W.; Haab, B.B.; Kamachi, M.; Dean, E.J.; Fournel, S.; Fong, D.; Genovese, M.C.; de Vegvar, H.E.; Skriver, K.; Hirschberg, D.L.; Morris, R.I.; Muller, S.; Puijn, G.J.; van Venrooij, W.J.; Smolen, J.S.; Brown, P.O.; Steinman, L.; Utz, P.J. *Nat. Med.*, **2002**, 8, 295.
- [71] Lueking, A.; Huber, O.; Wirths, C.; Schulte, K.; Stieler, K.M.; Blume-Peytavi, U.; Kowald, A.; Hensel-Wiegel, K.; Tauber, R.; Lehrach, H.; Meyer, H.E.; Cahill, D.J. *Mol. Cell Proteomics*, **2005**, 4, 1382.
- [72] Zhu, H.; Hu, S.; Jona, G.; Zhu, X.; Kreiswirth, N.; Willey, B.M.; Mazzulli, T.; Liu, G.; Song, Q.; Chen, P.; Cameron, M.; Tyler, A.; Wang, J.; Wen, J.; Chen, W.; Compton, S.; Snyder, M. *Proc. Natl. Acad. Sci. USA*, **2006**, 103, 4011.
- [73] Tao, S.C.; Zhu, H. *Nat. Biotechnol.*, **2006**, 24, 1253.
- [74] Angenendt, P.; Kreutzberger, J.; Glokler, J.; Hoheisel, J.D. *Mol. Cell Proteomics*, **2006**, 5, 1658.
- [75] Ramachandran, N.; Hainsworth, E.; Bhullar, B.; Eisenstein, S.; Rosen, B.; Lau, A.Y.; Walter, J.C.; LaBaer, J. *Science*, **2004**, 305, 86.
- [76] Paweletz, C.P.; Charboneau, L.; Bichsel, V.E.; Simone, N.L.; Chen, T.; Gillespie, J.W.; Emmert-Buck, M.R.; Roth, M.J.; Petricoin, E.F. III; Liotta, L.A. *Oncogene*, **2001**, 20, 1981.
- [77] Nishizuka, S.; Chen, S.T.; Gwadry, F.G.; Alexander, J.; Major, S.M.; Scherf, U.; Reinhold, W.C.; Waltham, M.; Charboneau, L.; Young, L.; Bussey, K.J.; Kim, S.; Lababidi, S.; Lee, J.K.; Pittaluga, S.; Scudiero, D.A.; Sausville, E.A.; Munson, P.J.; Petricoin, E.F. III; Liotta, L.A.; Hewitt, S.M.; Raffeld, M.; Weinstein, J.N. *Cancer Res.*, **2003**, 63, 5243.
- [78] Nishizuka, S.; Charboneau, L.; Young, L.; Major, S.; Reinhold, W.C.; Waltham, M.; Kouros-Mehr, H.; Bussey, K.J.; Lee, J.K.; Espina, V.; Munson, P.J.; Petricoin, E. III; Liotta, L.A.; Weinstein, J.N. *Proc. Natl. Acad. Sci. USA*, **2003**, 100, 142294.
- [79] Zhu, H.; Bilgin, M.; Snyder, M. *Annu. Rev. Biochem.*, **2003**, 72, 783.
- [80] Murthy, T.V.; Wu, W.; Qiu, Q.Q.; Shi, Z.; LaBaer, J.; Brizuela, L. *Protein Expr. Purif.*, **2004**, 36, 217.
- [81] Liu, R.; Barrick, J.E.; Szostak, J.W.; Roberts, R.W. *Methods Enzymol.*, **2000**, 318, 268.
- [82] Roberts, R.W.; Szostak, J.W. *Proc. Natl. Acad. Sci. USA*, **1997**, 94, 122972.
- [83] Weng, S.; Gu, K.; Hammond, P.W.; Lohse, P.; Rise, C.; Wagner, R.W.; Wright, M.C.; Kuimelis, R.G. *Proteomics*, **2002**, 2, 48.
- [84] Wilson, D.S.; Keefe, A.D.; Szostak, J.W. *Proc. Natl. Acad. Sci. USA*, **2001**, 98, 3750.

Copyright of Combinatorial Chemistry & High Throughput Screening is the property of Bentham Science Publishers Ltd. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.