

Published in final edited form as:

Wiley Interdiscip Rev Syst Biol Med. 2011 May ; 3(3): 255–268. doi:10.1002/wsbm.118.

Functional Protein Microarray Technology

Shaohui Hu^{1,2}, Zhi Xie³, Jiang Qian^{3,4}, Seth Blackshaw⁵, and Heng Zhu^{1,2,4,*}

¹Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

²The Center for High-Throughput Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

³Department of Ophthalmology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

⁴The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

⁵The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

Abstract

Functional protein microarrays are emerging as a promising new tool for large-scale and high-throughput studies. In this article, we will review their applications in basic proteomics research, where various types of assays have been developed to probe binding activities to other biomolecules, such as proteins, DNA, RNA, small molecules, and glycans. We will also report recent progress of using functional protein microarrays in profiling protein posttranslational modifications, including phosphorylation, ubiquitylation, acetylation, and nitrosylation. Finally, we will discuss potential of functional protein microarrays in biomarker identification and clinical diagnostics. We strongly believe that functional protein microarrays will soon become an indispensable and invaluable tool in proteomics research and systems biology.

Keywords

protein microarray; high-throughput; proteomics; post-translational modification; biomarker identification

The fundamental principle of microarray technology was first put forward by Ekins et al over 20 years ago [1]. Their ambient analyte theory stated that a tiny spot of purified antibody (or any other macromolecule) provides substantially better sensitivity than when used in conventional immunoassay formats. Fueled by large-scale genome sequencing projects, DNA microarray technology became the first application of this theory and has been widely used in gene expression profiling [2-6]. However, biological functions are carried out primarily by proteins rather than nucleic acids. Furthermore, RNA expression levels do not always correlate with protein expression levels, and it is almost impossible to predict biochemical properties of a protein encoded by a given gene simply based on its expression profiles [7,8]. Therefore, by focusing on studies of protein structures, functionalities, and protein-protein interactions one can more directly characterize biological function of a given gene.

*Correspondence to: heng.zhu@jhmi.edu (H. Z.).

Large-scale protein-centered analyses of gene function, however, have not been generally as fruitful as their DNA-centered counterparts for several reasons. First, the biochemical properties of proteins are far more diverse and complex than those of nucleic acids. Second, there is no ready method to amplify proteins for analysis, unlike PCR-based amplification of nucleic acids. Third, many proteins are prone to denature or degrade in standard buffer conditions and at ambient temperature, making them substantially more challenging to study. Therefore, technology for systematically assaying protein function that is both high-throughput and highly flexible is urgently needed. The past success of the DNA microarray technology highlights the power of a highly parallel, high-throughput platform that allows profiling of thousands of molecular targets in a single experiment. By the same token, protein microarray technology is now emerging as a promising new tool that can push proteomic studies to a new level. In the past decade, many methodologies based on the protein microarray technology have been successfully developed and applied to proteomic studies, including protein identification, quantification, and functional analysis of signaling pathways and networks, as well as clinical diagnostics and antibody characterizations.

A protein microarray, also known as a protein chip, is a solid surface (typically glass) on which thousands of different proteins (e.g., antigens, antibodies, enzymes, substrates, etc) are immobilized in discrete spatial locations, forming a high density protein dot matrix. Depending on their applications, protein microarrays can be classified into two types: the analytical and functional protein microarrays. Analytical protein microarrays are usually composed of well-characterized biomolecules with specific binding activities, such as antibodies, to analyze the components of complex biological samples (e.g., serum and cell lysates) or to determine whether a sample contains a specific protein of interest. They have been used for protein expression profiling, biomarker identification, cell surface marker/glycosylation profiling, clinical diagnosis, and environmental/food safety analysis. On the other hand, functional protein microarrays are constructed by printing a large number of individually purified proteins, and are mainly used to comprehensively query biochemistry properties and activities of those immobilized proteins. In principle, it is feasible to print arrays comprised of virtually all annotated proteins of a given organism, effectively comprising a whole proteome microarray. Functional protein microarrays have been successfully applied to identify protein-protein, protein-lipid, protein-antibody, protein-small molecules, protein-DNA, protein-RNA, lectin-glycan, and lectin-cell interactions, and to identify substrates or enzymes in phosphorylation, ubiquitylation, acetylation, and nitrosylation, as well as to profile immune response. In this review, we will mainly focus on the fabrication and application of functional protein microarrays.

Fabrication of Functional Protein Microarrays

Capture molecules with high specificity and selectivity are the essential prerequisite for the design and fabrication of both protein and DNA microarrays. Regardless of their sequences, the biochemical properties of DNA molecules are essentially the same, which allows the same chemistries to be applied to either immobilize to or synthesize *in situ* DNA strands on a solid surface. Therefore, the design and construction of oligonucleotide DNA microarrays are relatively straightforward. However, the protein world is much more complicated than DNA because of the vast differences among the structures, charge and hydrophobicity of individual proteins. This implies that the fabrication and analysis of protein microarrays is substantially less straightforward and standardizable than that of DNA microarrays. Unlike DNA or even RNA molecules, full-length proteins cannot be directly synthesized *in vitro* at high efficiency because of their complex chemistry. Although *in vitro* synthesis of peptides has been feasible for decades, it still suffers from low yield, high cost, and effective limitation to short sequences. Moreover, the vast majority of proteins must be correctly folded and modified to be functional after translation, which may require a complex

molecular machinery of chaperones and other accessory molecules that cannot be fully recapitulated *in vitro*. Thus far, proteins used to construct high-content proteome microarrays have been all individually expressed and purified from live cells (see below).

Because proteins must fold correctly in order to be active, proteins are prone to inactivation due to loss of their native conformations when directly immobilized on a solid surface. Proteins vary greatly in categories and properties, while a carrier surface can only be modified by one or two kinds of chemical or biological group. All these factors pose challenges for optimizing the physical attachment of proteins used in microarray construction to the slide surface.

Surface Chemistry

Choosing a proper surface for protein immobilization is crucial to the success of any assay performed using protein microarrays. An ideal surface should be able to retain protein functionality with relatively high signal-to-noise ratios, and possess both high protein-binding capacity and long shelf-life [9]. Glass slides covered with polyvinylidene fluoride (PVDF), nitrocellulose membrane, or polystyrene were popular for protein microarray fabrication in the early days of the technology [10-12]. However, PVDF and polystyrene are relatively soft, allowing lateral spread of printed proteins, and hence limited density of proteins to be printed. Nitrocellulose membranes, in addition, tend to generate high background and low signal-to-noise ratio for most applications.

To bypass these shortcomings, researchers developed three dimensional matrix arrays, in which glass slides are coated with polyacrylamide or agarose to form a porous hydrophilic matrix in which proteins or antibodies are trapped within the pores and lateral diffusion is restricted, reducing the size of printed protein spots and thus increasing the maximal complexity of the array [13,14]. Protein activity is generally well-preserved in such matrix arrays, and their protein binding capacity is relatively high. For instance, Zhu et al utilized soft lithography to generate nanowells on a polydimethylsiloxane (PDMS) sheet placed on top of microscope slides [15]. These nanowell chips were used to immobilize substrate proteins to profile phosphorylation specificity of 119 kinases encoded by budding yeast. The open structure of nanowells provides physical barriers and allows for sequential adding of different buffers, which is critical for multi-step experiments. The main disadvantage of this method is the requirement of specialized equipment needed to load nanowells at high density.

Other researchers printed proteins, antigens, or antibodies directly onto plain glass slides, which are usually coated with a bifunctional cross-linker with two functional groups, one reacting with the glass surface and the other with the desired proteins. For example, Schweitzer et al demonstrated in their study that protein microarrays fabricated on glass surface possess high sensitivities, wide dynamic range, and decent spot-to-spot reproducibility [16]. MacBeath and Schreiber demonstrated with three proteins that thousands of protein spots could be immobilized to aldehyde-activated plain glass surfaces to form a high-density protein microarray that was suitable for a range of different classes of assays [17].

High-Throughput Protein Production

Although the technologies of arraying proteins on various types of surfaces at high-density were starting to mature by the end of last millennium, the main hurdle to their more general use remained the difficulty in producing the large number of different proteins needed for construction of a truly high-content array. Obviously, a readily useable high-throughput protocol for parallel production of thousands of different proteins is the key to this.

An early attempt led by the Lehrach group was to express human proteins in bacteria using a library consisting of random cDNAs [10]. Individual cDNA clones of this library were robotically arrayed onto PVDF membrane laid on top of agar media and allowed to grow to full size. These cells were then lysed *in situ* to extract proteins, and then by incubation with a labeled test protein to identify interacting partners. Strictly speaking, human proteins bound to the nitrocellulose membranes were not purified – the vast majority of the proteins in every spot were bacterial proteins. Furthermore, the proteins were neither unique nor in native conformation, given the redundancy of the library and denaturing conditions used to break cells open. Though powerful as a screening technique in early days, this particular experimental strategy had limited general application.

To overcome these hurdles, the Snyder group created a high-throughput protein purification protocol in the budding yeast [18]. Using a homologous recombination-based strategy, more than 5,800 full-length yeast open reading frames (ORFs) were cloned into a yeast expression vector that, upon galactose induction, produces glutathione-S-transferase (GST)-tagged N-terminal fusion proteins. The purification protocol took advantage of both a 96-well format and immobilized affinity chromatography. This strategy allowed parallel purification of unprecedented numbers of proteins – up to 1,152 per day. The success of this approach is built upon several unique aspects: First, it utilizes a eukaryotic expression system that both generates high levels of recombinant proteins and tends to produce a high fraction of soluble proteins. Compared with bacterial expression systems, in which a large fraction of recombinant proteins end up in inclusion bodies, this is a huge advantage when a large number of eukaryotic proteins are being generated. Second, the expression of recombinant proteins is only induced over about two total cell cycles, which greatly reduces toxicity and cell death. Third, a foreign eukaryotic protein purified from yeast is more likely to be active because post-translational modifications (PTMs) necessary for function are more likely to occur correctly than in either bacteria or a cell-free system. Fourth, the use of an N-terminal GST tag helps protein fold correctly and therefore, improve its stability and solubility. Other commonly used tags include the so-called TAP-tag, MPB, and 6xHis, to name a few. In fact the same group later went on to build a TAP-tagged yeast ORF collection and purified >5,000 yeast proteins [19].

Another commonly used expression system is *Escherichia coli*. The procedures for automatic high-throughput protein expression/purification using the 6xHis tag has been developed [20]. The subsequent protein purification takes advantage of immobilized Ni-NTA affinity chromatography [21]. The 6xHis tag usually does not alter the properties of the fusion proteins, and the increment of molecular weight is less than 1 kDa.

Furthermore, it is selective and stable even under severe denaturing conditions [22,23]. Our group has recent reported a high-throughput protein purification protocol for 6xHis-tagged proteins in *E. coli* [24].

Despite the fact that high-throughput protein production in both prokaryotes and eukaryotes is now increasingly feasible, these protocols are still both labor-intensive and costly. Aside from the cost of protein production, fabrication of a proteome microarray requires construction of an expressible collection of full-length ORFs, which can be both challenging and expensive when dealing with higher eukaryotes with a large number of genes, such as humans. To explore alternative approaches, several groups have attempted to test the *in vitro* transcription/translation systems, such as the *E. coli*, wheat germ, and rabbit reticulocyte systems. In these systems, proteins can be expressed directly from cDNA templates [25], which can be obtained through PCR amplification without the lengthy and costly process of subcloning. For example, the *E. coli* cell-free protein expression system has been used to synthesize proteins in a 96-well format [26], and the improved wheat germ cell-free protein

synthesis system has been applied to the *in vitro* expression of 13,364 human proteins [27]. Furthermore, these systems can significantly decrease the reaction volume required for generation of recombinant proteins [28], which is also one of the advantages because the cost of *in vitro* expression system is rather high.

Such systems can also be applied to directly synthesize proteins on glass slides to fabricate so-called “in situ protein microarrays.” In the PISA (Protein In Situ Array) method, proteins are expressed directly from DNA *in vitro* and become attached to the array surfaces through recognition of a sequence that serves as an affinity tag [29]. Similarly, in the NAPPA (Nucleic Acid Programmable Protein Array) technology, biotinylated cDNA plasmids encoding proteins as GST fusions are printed onto avidin-coated slides, together with anti-GST antibodies as the capture molecules [30]. The cDNA array is then incubated with rabbit reticulocyte lysate to express the proteins, which become trapped by the antibodies adjacent to each DNA spot. Recently, NAPPA has been successfully expanded to high density arrays of 1000 different proteins [31]. In addition, Tao et al developed a different method in which ribosomes are installed at the end of an RNA template to allow for the capture of the nascent polypeptides by a puromycin moiety that is grafted at one end of an oligonucleotide immobilized on a solid surface [32].

Another similar method is called DAPA (DNA Array to Protein Array), in which proteins are synthesized between two glass slides, one of which is arrayed with DNA while the other carries a specific affinity reagent to capture the proteins [33]. In this approach, tagged proteins are synthesized in parallel from the DNA array, spread across the gap between the two slides, and then bind to the tag-capturing reagents on the other slide to form a protein array. Unlike the NAPPA method in which proteins are present together with DNA and the DNA array can only be used once, DAPA generates multiple copies of ‘pure’ protein arrays on a separate surface from the same DNA template, with at least 20 copies capable of being produced from a single template.

Protein Printing

With regard to spotting proteins the two major mechanisms are contact and non-contact printing. Adapted from DNA microarray fabrication, the robotic contact printing tool is the most suitable for producing protein microarrays of high content due to the requirement to array large numbers of different proteins. Metal pins with solid or quill tips are used in contact printers to deliver sub-nanoliter of protein samples to the slide surface. Quill pins, which have a larger sample capacity, can print hundreds of spots continuously after each sample loading. The printed spots are typically circular and the size depends largely on the pin tip dimension, surface chemistry, and the printing buffer. A significant advantage of this type of microarrayers is their speed and throughput - up to 48 pins can be loaded and more than 200 slides printed at a time. However, the pins are very fragile and expensive, and the pin tips may damage the slide surface, especially complex 3D substrates (e.g., nitrocellulose-coated slides). Furthermore, some proteins are sticky to metal and the general washing steps may not clean them completely from the pins, and thus give rise to cross-contamination of protein samples and carry-over problem.

To address these issues, non-contact dispensing techniques have been developed for printing protein microarrays, by which a small droplet of protein sample is delivered to the slide surface without touching it. Droplets can be generated by conventional ink-jet, piezoelectric pulsing, or electrospray deposition [34-36]. Unlike contact printing, the amount of liquid deposited by non-contact printers is not dependent on surface properties of the slide, and significant better spots morphology has been observed on hydrophobic surfaces using non-contact printing compared to contact printing [37]. In addition to standard glass slides, non-contact printers can also print on membranes. However, such instruments usually suffer

from longer printing time and fewer pins, which is a significant drawback when printing a large number of protein samples. Moreover, non-contact printers can sometimes misplace spots and/or generate satellite spots, resulting in a high failure rate [38]. An additional disadvantage of non-contact printers is that they usually require a larger sample volume, which is challenging and expensive for high-throughput protein production.

Protein Immobilization

The physical and chemical properties of different proteins vary greatly, and protein activities are closely related to their structures. Therefore, the development of a stable, universal immobilization method that does not change protein structures is one of the difficulties of protein microarray fabrication. So far several different methods have been used for protein immobilization on solid carrier surfaces, such as noncovalent adsorption, covalent binding, and affinity capture.

Noncovalent adsorption provides both high protein capacities and low impact on protein structures, but cannot control the amount and orientation of immobilized proteins. Thus the reaction efficiency, accuracy, and reproducibility of arrays produced in this manner are variable. Covalent binding, on the other hand, results in chemically cross-linked proteins via reactive residues (e.g., lysine and cystine) to surface-grafted ligands, such as aldehyde, epoxy, reactive ester, etc [17,39,40]. Lee et al developed novel calixcrown derivatives as a ProLinker that permits efficient immobilization of captured proteins on solid matrixes, and the immobilized proteins showed both consistent directionality and functionality [41]. Covalent binding is suitable for immobilization of a wide range of proteins with strong conjunctions to the carrier surfaces. However, the modification of chemical groups can sometimes both alter the activities of target proteins and their binding to specific ligands.

Affinity capture is an attractive way to immobilize proteins that avoids many of the shortcomings of the previously detailed approaches. For example, biotinylated proteins have been used for protein immobilization to streptavidin-coated slides. The use of genetically encoded affinity tags, which can be fused to target proteins and bind to a specific slide surface, is an analogous approach. For example, 6xHis-tags have been utilized to immobilize proteins on nickel-NTA coated glass slides [18]. Presumably, affinity-based protein immobilization should result in immobilization of proteins in relatively uniform orientation with minimum interruption of protein structure, and thus may be the best approach to for preserving the structure and function of printed. One important caveat to bear in mind, however, is that the incorporation of affinity tags may alter the protein structures.

One way to deal with this challenge was demonstrated by Zhang et al., who developed a flexible polypeptide scaffold consisting of a surface immobilization domain and a protein capture domain, which allows much greater flexibility in the immobilization of proteins on a microarray [42]. Wacker et al compared the DNA-directed immobilization (DDI) method with both direct spotting and with biotin-streptavidin affinity immobilization for antibodies [43]. DDI is based on the self-assembly of semisynthetic DNA-streptavidin conjugates that converts a DNA oligomer array into an antibody array [44]. DDI and direct spotting showed the highest fluorescence intensities. DDI also performed the best in spot homogeneity and intra- and inter-experimental reproducibility. Moreover, DDI required the lowest amount of antibodies, at least 100-fold less than direct spotting. The drawback of DDI is that proteins have to be linked to DNA prior to immobilization, which increases the workload involved in generating microarrays.

The orientation of immobilized proteins may influence both their activity and their affinity for specific ligands. Peluso et al compared randomly versus specifically oriented capture agents based on both full-sized antibodies and Fab' fragments [45]. The specific orientation

of capture agents consistently increased the analyte-binding capacity of the surfaces up to 10-fold relative to surfaces with randomly oriented capture agents. When specifically oriented, Fab' fragments formed a dense monolayer and 90% of them were active, while randomly attached Fab's both packed at lower density and had lower specific activity.

Signal Detection

In addition to optimized surface modification and optimized reaction condition, the detection sensitivity of samples bound on microarrays is another key parameter in the design of protein microarray assays. There are two basic detection methods: label-dependent and label-free detections.

Label-Dependent Detection Methods

Radioisotopes and fluorescent dyes are the two most common labeling methods for signal detection in protein microarray assays. Fluorescent dyes, such as Cy-3/5 and their equivalent, have been used as a popular labeling method. Because most good dyes have relatively narrow excitation and emission spectra, multi-color scheme can be readily implemented for simultaneous detection and direct comparison of different samples, both reducing cost and avoiding chip-to-chip variation. Semiconductor quantum dot labeling, which is brighter and more stable than organic dyes, has also been applied to protein microarrays [46,47].

In addition to fluorescent labeling, Huang et al detected multiple cytokines on an antibody array with enhanced chemiluminescence (ECL), providing an alternative detection method [48]. Enzymatic signal amplification is also a valuable labeling method. Rolling circle amplification (RCA) has been developed for protein microarray assays. For low abundance protein samples, the sensitivity of traditional fluorescence or chemiluminescence detection is relatively low, while RCA can detect captured proteins at fmol level and is promising to improve the sensitivity of fluorescent detection [16,49-52]. Tyramide signal amplification (TSA) is another way to amplify signals with enzymes, which utilizes the horseradish peroxidase conjugated on secondary antibodies to convert the labeled substrates (tyramide) into short-lived, extremely reactive intermediates, which then very rapidly react with and covalently bind to adjacent proteins [53].

For some biochemical assays, especially enzymatic reactions, use of radioisotopes is the only detection method available (see below for more details). They still offer the most sensitive and reliable detection of PTM events when there is a lack of high quality and high affinity detection reagents, such as antibodies. We and others have successfully applied ^{32}P -, ^{33}P -, and ^{14}C -labeled substrates to detect protein phosphorylation and acetylation events [54,55].

Label-Free Detection Methods

One obvious disadvantage of label-dependent detection is the requirement of either manipulating structure of a probe or a specific antibody. It is not amenable to real-time detection, which can provide important information when analyzing reaction dynamics. Therefore, label-free detection methods have also been investigated for protein microarrays. Surface Plasmon resonance (SPR) is a label-free technology to analyze biomolecular interactions in real-time, and has been adapted for protein microarray signal detection [56,57]. Based on the principle that incident light can resonate with plasma on a metal surface in total internal reflection, the resonance signals will change when ligands bind to (and dissociate from) ligands on the array surface. Binding event can thus be monitored and the kinetic parameters calculated in real-time. Mass spectrometry has also been used for detecting ligands bound to individual proteins printed on protein microarrays, with such

approaches as MALDI-MS, SELDI-TOF-MS, and MALDI-TOF-MS used for this purpose [58-60]. The analysis is rapid and simple, requires small sample amount, and can be used for direct detection of analytes bound from complex samples, such as urine, serum, plasma, and cell lysates. Atomic force microscopy (AFM) uses surface topological changes to identify the analytes bound on the array [61,62]. More specifically, AFM detects the increase in height of the proteins/antibodies on the array, and thus is able to measure binding interactions.

Applications of Protein Microarrays

Unlike the DNA/oligo microarray or analytical protein microarrays, functional protein microarrays provide a flexible platform that allows development and detection of a wide range of protein biochemical properties. To date, well-developed assays include detection of various types of protein-ligand interactions, such as protein-protein, protein-DNA, protein-RNA, protein-lipid, protein-drug, and protein-glycan interactions [17,18,24,63-69], and identification of substrates of various classes of enzymes, such as protein kinase, ubiquitin E3 ligase, and acetyltransferase, to name a few [15,54,55,70,71]. Application of these assays has had a profound impact on a wide range of research areas. This is especially true when they are used in large-scale, high-throughput projects, exemplified in both network construction and biomarker identification (see below and Table 1).

Protein-Protein and -Lipid Interactions

Among the first applications of protein microarrays was in the analysis of protein-protein and protein-lipid interactions, where test ligands were directly or indirectly labeled with fluorescent dyes. For example, Zhu et al developed the first proteome microarray composed of ~5800 recombinant yeast proteins (>85% of the yeast proteome) and identified binding partners of calmodulin and phosphatidylinositides (PIPs) [18]. They first incubated the microarrays with biotinylated bovine calmodulin and discovered 39 new calmodulin binding partners. In addition, using liposomes as a carrier for various PIPs, they identified more than 150 binding proteins, >50% of which were known membrane-associated proteins. Popescu et al developed a protein microarray containing 1,133 *Arabidopsis thaliana* proteins and also used it to globally identify proteins bind to calmodulins or calmodulin-like proteins in *Arabidopsis* [63]. A large number of previously known and novel targets were identified, including transcription factors (TFs), receptor and intracellular protein kinases, F-box proteins, RNA-binding proteins, and proteins of unknown function. Alternative approaches to identifying protein-protein interactions, such as the yeast two-hybrid system and protein complex purification coupled with mass spectrometry analysis, are well-established, however, and are used as standard high-throughput methods to detect protein-protein interactions in higher eukaryotes [72,73]. Thus, while protein microarray-based approaches provide a rapid approach to characterizing protein-protein interactions, they have much competition in this arena.

Protein-Peptide Interaction

MacBeath and colleagues fabricated protein domain microarrays to investigate protein-peptide interactions in a semi-quantitative fashion that might play an important role in signaling [74]. They constructed an array by printing 159 human Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains on the aldehyde-modified glass substrates, and incubated the arrays with 61 peptides representing tyrosine phosphorylation sites on the four ErbB receptors. Eight concentrations of each peptide (10 nM to 5mM) were tested in the assay, allowing quantitative measurement of the binding affinity of each peptide to its protein ligand.

Protein-DNA Interaction

Protein microarrays have also been applied extensively and productively to characterize protein-DNA interactions (PDIs). In an earlier study, Snyder and colleagues screened for novel DNA-binding proteins by probing the yeast proteome microarrays with fluorescently labeled yeast genomic DNA [64]. Of the ~200 positive proteins, half were not previously known to bind to DNA. By focusing on a single yeast gene, *ARG5,6*, encoding two enzymes involved in arginine biosynthesis, they discovered that it bound to a specific DNA motif and associated with specific nuclear and mitochondrial loci *in vivo*.

In a later report, the Snyder and Johnston groups constructed a protein microarray with 282 known and predicted yeast TFs to identify their interactions with 75 evolutionarily conserved DNA motifs [65]. Over 200 specific PDIs were identified and >60% of them are previously unknown. The binding site of a previously uncharacterized DNA-binding protein, Yj1103p, was defined and a number of its target genes were identified, many of which are involved in stress response and oxidative phosphorylation.

Our team developed a bacterial proteome microarray composed of 4,256 proteins encoded by the *Escherichia coli* K12 strain (~99% coverage of the proteome) using a bacterial high-throughput protein purification protocol [24]. To demonstrate the usefulness, end-labeled, double-stranded (ds) DNA probes carrying abasic or mismatched base pairs were used to identify proteins involved in DNA damage recognition. A small number of proteins were specifically recognized by each type of the probes with high affinity. Two of them, YbaZ and YbcN, were further characterized to encode base-flipping activity using biochemical assays.

Recently, our group also undertook a large-scale analysis of human PDIs using a protein microarray composed of 4,191 unique human proteins in full-length, including ~90% of the annotated TFs and a wide range of other protein categories, such as RNA-binding proteins, chromatin-associated proteins, nucleotide-binding proteins, transcription co-regulators, mitochondrial proteins and protein kinases [66]. The protein microarrays were probed with 400 predicted and 60 known DNA motifs and a total of 17,718 PDIs were identified. Many known PDIs and a large number of new PDIs for both well characterized and predicted TFs were recovered, and new consensus sites for over 200 TFs were determined, which doubled the number of previously reported consensus sites for human TFs [66,75]. Surprisingly, over 300 proteins that were previously unknown to specifically interact with DNA showed sequence-specific PDIs, suggesting that many human proteins may bind specific DNA sequences as a moonlighting function. To further investigate whether the DNA-binding activities of these unconventional DNA binding proteins (uDBPs) were physiologically relevant, we carried out in-depth analysis on a well studied protein kinase, Erk2, to determine the potential mechanism behind its DNA-binding activity. Using a series of *in vitro* and *in vivo* approaches, such as EMSA, luciferase assay, mutagenesis, and ChIP, we demonstrated that the DNA-binding activity of Erk2 is independent of its protein kinase activity and it acts as a transcription repressor of transcripts induced by interferon gamma signaling [66]. Our study suggests that moonlighting functions of uDBPs based on their sequence-specific DNA-binding activity may be a widespread phenomenon in humans.

Protein-Small Molecule Interaction

Discovering new drug molecules and drug targets is another field in which protein microarrays have shown its potential. For example, Huang et al incubated biotinylated small-molecule inhibitors of rapamycin (SMIRs) on the yeast proteome microarrays, and obtained the binding profiles of the SMIRs across the entire yeast proteome [68]. They identified candidate target proteins of the SMIRs, including Tep1p, a homolog of the

mammalian PTEN tumor suppressor, and Ybr077cp (Nir1p), a protein of previously unknown function, both of which are validated to associate with PI(3,4)P₂, suggesting a novel mechanism by which phosphatidylinositides might modulate the TOR (target of rapamycin) pathway.

Protein-RNA Interaction

The yeast proteome microarray has been used to identify specific RNA-binding proteins for antiviral activities [67]. In these experiments, arrays were incubated with a fluorescently tagged small RNA hairpin containing a clamped adenine motif (CAM), which is required for the replication of Brome Mosaic Virus (BMV), a plant-infecting RNA virus that can also replicate in the budding yeast. Two of the candidate proteins, Pseudouridine Synthase 4 (Pus4) and the Actin Patch Protein 1 (App1), were further characterized in *Nicotiana benthamiana*. Both of them modestly reduced BMV genomic plus-strand RNA accumulation and dramatically inhibited the spread of BMV in plants.

Protein-Glycan Interaction

Protein glycosylation, a general PTM of proteins involved in cell membrane formation, is crucial to dictate proper conformation of many membrane proteins, retain stability on some secreted glycoproteins, and play a role in cell-cell adhesion. To further understand the roles of protein glycosylation in yeast, the Zhu and Snyder teams reasoned that since proteins on the yeast proteome microarrays are expressed in their original host and therefore, which should maintain most of their PTMs, these arrays can be used to profile glycosylation using fluorescently labeled lectins, such as Concanavalin A (ConA) and Wheat-Germ Agglutinin (WGA) [69]. A total of 534 proteins were identified, 406 of which were previously not known to be glycosylated. Many proteins in the secretory pathway were identified, as well as other functional classes of proteins, including TFs and mitochondrial proteins. Upon treatment with tunicamycin, an inhibitor of N-linked protein glycosylation, two of the four mitochondrial proteins identified showed partial distribution to the cytosol and reduced localization to the mitochondria, suggesting a new role of protein glycosylation in mitochondrial protein function and localization.

Protein Phosphorylation

Protein phosphorylation plays a central role in almost, if not all, aspects of cellular processes. The application of protein microarray technology to protein phosphorylation was first demonstrated by Zhu et al [15]. They immobilized 17 different substrates on a nanowell protein microarray, followed by individual kinase assays with almost all of the yeast kinases (119/122). This approach allowed them to determine the substrate specificity of the yeast kinome and identify new tyrosine phosphorylation activity.

In a later report, Snyder's group accomplished a large scale "Phosphorylome Project" using the yeast proteome microarrays [54]. Eighty seven purified yeast kinases or kinase complexes were individually incubated on the yeast proteome arrays in a kinase buffer in the presence of ³³P-γ-ATP and a total of 1,325 distinct protein substrates were identified, representing a total of 4,129 phosphorylation events. These results provided a global network that connect kinases to their potential substrates and offered a new opportunity to identify new signaling pathways or cross-talk between pathways.

Several smaller scale studies of kinase-substrate interactions have been reported. For instance, Popescu et al probed 10 *Arabidopsis* mitogen-activated protein kinases (MPKs) to protein microarrays containing 2158 *Arabidopsis* proteins and identified 570 putative MPK phosphorylation targets, which were enriched in transcription factors involved in the regulation of development, defense, and stress responses [76]. A commercially available

human protein microarray comprised of approximately 3,000 individual proteins was used to identify substrates of cyclin-dependent kinase 5 (Cdk5), a serine/threonine kinase that plays an important role during CNS development [70].

Protein Ubiquitylation

Ubiquitylation is one of the most prevalent PTMs and controls almost all types of cellular events in eukaryotes. To establish a protein microarray-based approach for identification of ubiquitin E3 ligase substrates, Lu et al developed an assay for yeast proteome microarrays that utilizes a HECT-domain E3 ligase, Rsp5, in combination with the E1 and E2 enzymes [71]. More than 90 new substrates were identified, eight of which were validated as *in vivo* substrates of Rsp5. Further *in vivo* characterization of two substrates, Sla1 and Rnr2, demonstrated that Rsp5-dependent ubiquitylation affects either posttranslational process of the substrate or subcellular localization.

Protein Acetylation

Histone acetylation and deacetylation, which are catalyzed by histone acetyltransferases (HATs) and histone deacetylases (HDACs), respectively, are emerging as critical regulators of chromatin structure and transcription. However, it has been hypothesized that many HATs and HDACs might also modify non-histone substrates. For example, the core enzyme, Esa1, of the essential nucleosome acetyltransferase of H4 (NuA4) complex, is the only essential HAT in yeast, which strongly suggested that it may target additional non-histone proteins that are crucial for cell to survive. To identify non-histone substrates of the NuA4 complex, Lin et al established and performed acetylation reactions on the yeast proteome microarrays using the NuA4 complex in the presence of [¹⁴C]-Acetyl-CoA as a donor [55]. Surprisingly, 91 proteins were found to be readily acetylated by the NuA4 complex on the array. To further validate these *in vitro* results, 20 of them were randomly chosen and 13 of them showed Esa1-dependent acetylation in cells. One of them, phosphoenolpyruvate carboxykinase (Pck1p), was further characterized to explore the possible link between acetylation and metabolism. Mass spectrometry assay revealed Lys19 and 514 as the acetylation sites of Pck1p, and mutagenesis analyses demonstrated that acetylation on K514 is critical to enhance Pck1p's enzyme activity and results in longer life span for yeast cells growing under starvation. This study offers a molecular link between the HDAC Sir2 and yeast longevity.

S-Nitrosylation

S-nitrosylation is independent of enzyme catalysis but an important PTM that affects a wide range of proteins involved in many cellular processes. Recently, Foster et al developed a protein microarray-based approach to detect proteins reactive to S-nitrosothiol (SNO), the donor of NO⁺ in S-nitrosylation, and to investigate determinants of S-nitrosylation [77]. S-nitrosocysteine (CysNO), a highly reactive SNO, was added to the yeast proteome microarray and the nitrosylated proteins were then detected using a modified biotin switch technique. The top 300 proteins with the highest relative signal intensity were further analyzed and the results revealed that proteins with active-site Cys thiols residing at N termini of alpha-helices or within catalytic loops were particularly prominent. However, substantial variations of S-nitrosylation were observed even within these protein families, indicating that secondary structure or intrinsic nucleophilicity of Cys thiols was not sufficient to interpret the specificity of S-nitrosylation. Further analyses revealed that NO-donor stereochemistry and structure had significant impact on S-nitrosylation efficiency.

Biomarker Identification

Though the applications described above are most useful in basic research, functional protein microarrays may have enormous impacts on clinical diagnosis and prognosis. When proteins on a functional protein microarray are viewed as potential antigens that may or may not be associated with a particular disease, it becomes a powerful tool in biomarker identification. The principle is straightforward: when an auto-antibody presented in human sera associated with a human disease (e.g., auto-immune diseases) recognizes a human protein spotted on the array, it can be readily detected with fluorescently labeled anti-human immunoglobulin antibodies (e.g., anti-IgG) and a profile of auto-antibodies associated with a disease thus created, providing a rapid approach to identifying potential disease biomarkers. For example, Robinson et al reported the first application of protein microarray technology to profile multiple human disease sera [78]. They constructed a microarray with 196 biomolecules shown to be autoantigens in eight human autoimmune diseases, including proteins, peptides, enzyme complexes, ribonucleoprotein complexes, DNA, and post-translationally modified antigens. The arrays were incubated with patient sera to study the specificity and pathogenesis of autoantibody responses, and were used to identify and define relevant auto-antigens in human autoimmune diseases, including systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA).

Hu et al reported a new approach for high-throughput characterization of monoclonal antibody (mAb) target specificity using a protein microarray composed of 1058 unique human liver proteins [79]. They immunized mice with live cells from human livers, isolated 54 hybridomas with binding activities to human cells, and identified the corresponding antigens for 5 mAbs via screening on the protein microarray. Expression profiles of the corresponding antigens of the 5 antibodies were characterized by using tissue microarrays and one of the antigens, eIF1A, was found to be expressed in normal human liver but not in hepatocellular carcinoma. Other applications include biomarker identification for ovarian cancer [80], inflammatory bowel disease [81], alopecia areata [82], and autoimmune hepatitis [83], etc.

Protein microarrays can also be used for detection of infectious diseases. Zhu et al developed a coronavirus protein microarray for the diagnosis of severe acute respiratory syndrome (SARS), which included 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) [84]. These microarrays could quickly distinguish patient serum samples as SARS-positive or SARS-negative based on the presence of human IgG and IgM antibodies against SARS-CoV proteins, with a 94% accuracy compared to standard diagnostic methods. Patients carrying antibodies against other coronavirus proteins were also identified. The advantages of this microarray-based assay to standard ELISA-based diagnostic methods include at least 100-fold higher sensitivity and the need for substantially less sample for analysis.

Outlook

Recent years have witnessed a rapid growth in using functional protein microarrays for basic research [85]. Although the technology is still at a relatively early stage of development, it has become obvious that the protein microarray platform can and will act as a versatile tool suitable for the large-scale, high-throughput biology, especially in the areas of profiling PTMs and in analysis of signal transduction networks and pathways [54,66]. As another crucial proteomics technology, recent progress in mass spectrometry has allowed global profiling of PTMs using a shotgun approach. For example, the Zhao, Mann, and Guan groups recently identified numerous acetylated lysine residues in metabolic enzymes in mice and human cells without knowing the upstream HATs [86-88]. In parallel, our team also

identified many yeast metabolic enzymes as substrates of the NuA4 acetylation complex without knowing the actual modified sites. Therefore, we envision that the combination of the two technologies will have enormous potential to both identify critical regulatory PTMs at the resolution of modified individual amino acids and to identify the enzymes that mediate these effects. Another emerging direction is in the forefront of understanding the molecular mechanisms of pathogen-host interactions. In the same manner in which we identified host proteins that recognized the SLD loop of the BMV virus, functional protein microarrays (e.g., a human protein microarray) can be used to discover those host proteins targeted by pathogens (e.g., HIV, HCV, and SARS-CoV). The identification of the host targets of a virus will provide alternative therapeutics that cannot be rapidly evaded *via* mutation of the viral genomes. In conclusion, the potential of functional protein microarrays is only just now starting to reveal itself. It is expected that it will become an indispensable and invaluable tool in proteomics and systems biology research.

Acknowledgments

We thank the NIH for funding support.

References

1. Ekins RP. Multi-analyte immunoassay. *J Pharm Biomed Anal.* 1989; 7(2):155–68. [PubMed: 2488616]
2. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci U S A.* 1994; 91(11):5022–6. [PubMed: 8197176]
3. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* 1995; 270(5235):467–70. [PubMed: 7569999]
4. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science.* 1997; 278(5338):680–6. [PubMed: 9381177]
5. Schadt EE, Monks SA, Drake TA, Lusk AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH. Genetics of gene expression surveyed in maize, mouse and man. *Nature.* 2003; 422(6929):297–302. [PubMed: 12646919]
6. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG. Genetic analysis of genome-wide variation in human gene expression. *Nature.* 2004; 430(7001):743–7. [PubMed: 15269782]
7. Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol.* 1999; 19(3):1720–30. [PubMed: 10022859]
8. Lueking A, Cahill DJ, Mullner S. Protein biochips: A new and versatile platform technology for molecular medicine. *Drug Discov Today.* 2005; 10(11):789–94. [PubMed: 15922937]
9. Kumble KD. Protein microarrays: new tools for pharmaceutical development. *Anal Bioanal Chem.* 2003; 377(5):812–9. [PubMed: 12851736]
10. Bussow K, Cahill D, Nietfeld W, Bancroft D, Scherzinger E, Lehrach H, Walter G. A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucleic Acids Res.* 1998; 26(21):5007–8. [PubMed: 9776767]
11. Lueking A, Horn M, Eickhoff H, Bussow K, Lehrach H, Walter G. Protein microarrays for gene expression and antibody screening. *Anal Biochem.* 1999; 270(1):103–11. [PubMed: 10328771]
12. Holt LJ, Bussow K, Walter G, Tomlinson IM. By-passing selection: direct screening for antibody-antigen interactions using protein arrays. *Nucleic Acids Res.* 2000; 28(15):E72. [PubMed: 10908365]
13. Guschin D, Yershov G, Zaslavsky A, Gemmell A, Shick V, Proudnikov D, Arenkov P, Mirzabekov A. Manual manufacturing of oligonucleotide, DNA, and protein microchips. *Anal Biochem.* 1997; 250(2):203–11. [PubMed: 9245440]
14. Afanassiev V, Hanemann V, Wolf S. Preparation of DNA and protein micro arrays on glass slides coated with an agarose film. *Nucleic Acids Res.* 2000; 28(12):E66. [PubMed: 10871389]

15. Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M. Analysis of yeast protein kinases using protein chips. *Nat Genet.* 2000; 26(3):283–9. [PubMed: 11062466]
16. Schweitzer B, Wiltshire S, Lambert J, O'Malley S, Kukanskis K, Zhu Z, Kingsmore SF, Lizardi PM, Ward DC. Inaugural article: immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. *Proc Natl Acad Sci U S A.* 2000; 97(18):10113–9. [PubMed: 10954739]
17. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science.* 2000; 289(5485):1760–3. [PubMed: 10976071]
18. 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 RA, Gerstein M, Snyder M. Global analysis of protein activities using proteome chips. *Science.* 2001; 293(5537):2101–5. [PubMed: 11474067]
19. Gelperin DM, White MA, Wilkinson ML, Kon Y, Kung LA, Wise KJ, Lopez-Hoyo N, Jiang L, Piccirillo S, Yu H, Gerstein M, Dumont ME, Phizicky EM, Snyder M, Grayhack EJ. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev.* 2005; 19(23):2816–26. [PubMed: 16322557]
20. Scheich C, Sievert V, Bussow K. An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. *BMC Biotechnol.* 2003; 3:12. [PubMed: 12885298]
21. Hochuli E, Dobeli H, Schacher A. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J Chromatogr.* 1987; 411:177–84. [PubMed: 3443622]
22. Mukhija R, Rupa P, Pillai D, Garg LC. High-level production and one-step purification of biologically active human growth hormone in *Escherichia coli*. *Gene.* 1995; 165(2):303–6. [PubMed: 8522194]
23. Joshi B, Janda L, Stoytcheva Z, Tichy P. PkWA, a WD-repeat protein, is expressed in spore-derived mycelium of *Thermomonospora curvata* and phosphorylation of its WD domain could act as a molecular switch. *Microbiology.* 2000; 146 Pt 12:3259–67. [PubMed: 11101684]
24. Chen CS, Korobkova E, Chen H, Zhu J, Jian X, Tao SC, He C, Zhu H. A proteome chip approach reveals new DNA damage recognition activities in *Escherichia coli*. *Nat Methods.* 2008; 5(1):69–74. [PubMed: 18084297]
25. Allen SV, Miller ES. RNA-binding properties of in vitro expressed histidine-tagged RB69 RegA translational repressor protein. *Anal Biochem.* 1999; 269(1):32–7. [PubMed: 10094772]
26. Murthy TV, Wu W, Qiu QQ, Shi Z, LaBaer J, Brizuela L. Bacterial cell-free system for high-throughput protein expression and a comparative analysis of *Escherichia coli* cell-free and whole cell expression systems. *Protein Expr Purif.* 2004; 36(2):217–25. [PubMed: 15249043]
27. Goshima N, Kawamura Y, Fukumoto A, Miura A, Honma R, Satoh R, Wakamatsu A, Yamamoto J, Kimura K, Nishikawa T, Andoh T, Iida Y, Ishikawa K, Ito E, Kagawa N, Kaminaga C, Kanehori K, Kawakami B, Kenmochi K, Kimura R, Kobayashi M, Kuroita T, Kuwayama H, Maruyama Y, Matsuo K, Minami K, Mitsubori M, Mori M, Morishita R, Murase A, Nishikawa A, Nishikawa S, Okamoto T, Sakagami N, Sakamoto Y, Sasaki Y, Seki T, Sono S, Sugiyama A, Sumiya T, Takayama T, Takayama Y, Takeda H, Togashi T, Yahata K, Yamada H, Yanagisawa Y, Endo Y, Imamoto F, Kisu Y, Tanaka S, Isogai T, Imai J, Watanabe S, Nomura N. Human protein factory for converting the transcriptome into an in vitro-expressed proteome. *Nat Methods.* 2008; 5(12):1011–7. [PubMed: 19054851]
28. Angenendt P, Nyarsik L, Szaflarski W, Glokler J, Nierhaus KH, Lehrach H, Cahill DJ, Lueking A. Cell-free protein expression and functional assay in nanowell chip format. *Anal Chem.* 2004; 76(7):1844–9. [PubMed: 15053642]
29. He M, Taussig MJ. Single step generation of protein arrays from DNA by cell-free expression and in situ immobilisation (PISA method). *Nucleic Acids Res.* 2001; 29(15):E73–3. [PubMed: 11470888]
30. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, Lau AY, Walter JC, LaBaer J. Self-assembling protein microarrays. *Science.* 2004; 305(5680):86–90. [PubMed: 15232106]

31. Ramachandran N, Raphael JV, Hainsworth E, Demirkan G, Fuentes MG, Rolfs A, Hu Y, LaBaer J. Next-generation high-density self-assembling functional protein arrays. *Nat Methods*. 2008; 5(6): 535–8. [PubMed: 18469824]
32. Tao SC, Zhu H. Protein chip fabrication by capture of nascent polypeptides. *Nat Biotechnol*. 2006; 24(10):1253–4. [PubMed: 17013375]
33. He M, Stoevesandt O, Palmer EA, Khan F, Ericsson O, Taussig MJ. Printing protein arrays from DNA arrays. *Nat Methods*. 2008; 5(2):175–7. [PubMed: 18204456]
34. Roda A, Guardigli M, Russo C, Pasini P, Baraldini M. Protein microdeposition using a conventional ink-jet printer. *Biotechniques*. 2000; 28(3):492–6. [PubMed: 10723562]
35. Sloane AJ, Duff JL, Wilson NL, Gandhi PS, Hill CJ, Hopwood FG, Smith PE, Thomas ML, Cole RA, Packer NH, Breen EJ, Cooley PW, Wallace DB, Williams KL, Gooley AA. High throughput peptide mass fingerprinting and protein macroarray analysis using chemical printing strategies. *Mol Cell Proteomics*. 2002; 1(7):490–9. [PubMed: 12239277]
36. Avseenko NV, Morozova TY, Ataullakhanov FI, Morozov VN. Immunoassay with multicomponent protein microarrays fabricated by electrospray deposition. *Anal Chem*. 2002; 74(5):927–33. [PubMed: 11924994]
37. Dufva M. Fabrication of high quality microarrays. *Biomol Eng*. 2005; 22(5-6):173–84. [PubMed: 16242381]
38. Hartmann M, Sjodahl J, Stjernstrom M, Redeby J, Joos T, Roeraade J. Non-contact protein microarray fabrication using a procedure based on liquid bridge formation. *Anal Bioanal Chem*. 2009; 393(2):591–8. [PubMed: 19023564]
39. Ziauddin J, Sabatini DM. Microarrays of cells expressing defined cDNAs. *Nature*. 2001; 411(6833):107–10. [PubMed: 11333987]
40. Templin MF, Stoll D, Schrenk M, Traub PC, Vohringer CF, Joos TO. Protein microarray technology. *Trends Biotechnol*. 2002; 20(4):160–6. [PubMed: 11906748]
41. Lee Y, Lee EK, Cho YW, Matsui T, Kang IC, Kim TS, Han MH. ProteoChip: a highly sensitive protein microarray prepared by a novel method of protein immobilization for application of protein-protein interaction studies. *Proteomics*. 2003; 3(12):2289–304. [PubMed: 14673779]
42. Zhang K, Diehl MR, Tirrell DA. Artificial polypeptide scaffold for protein immobilization. *J Am Chem Soc*. 2005; 127(29):10136–7. [PubMed: 16028902]
43. Wacker R, Schroder H, Niemeyer CM. Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavidin-biotin attachment: a comparative study. *Anal Biochem*. 2004; 330(2):281–7. [PubMed: 15203334]
44. Niemeyer CM, Sano T, Smith CL, Cantor CR. Oligonucleotide-directed self-assembly of proteins: semisynthetic DNA-streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates. *Nucleic Acids Res*. 1994; 22(25):5530–9. [PubMed: 7530841]
45. Peluso P, Wilson DS, Do D, Tran H, Venkatasubbaiah M, Quincy D, Heidecker B, Poindexter K, Tolani N, Phelan M, Witte K, Jung LS, Wagner P, Nock S. Optimizing antibody immobilization strategies for the construction of protein microarrays. *Anal Biochem*. 2003; 312(2):113–24. [PubMed: 12531195]
46. Shingyoji M, Gerion D, Pinkel D, Gray JW, Chen F. Quantum dots-based reverse phase protein microarray. *Talanta*. 2005; 67(3):472–8. [PubMed: 18970191]
47. Zajac A, Song D, Qian W, Zhukov T. Protein microarrays and quantum dot probes for early cancer detection. *Colloids Surf B Biointerfaces*. 2007; 58(2):309–14. [PubMed: 17408931]
48. Huang RP. Detection of multiple proteins in an antibody-based protein microarray system. *J Immunol Methods*. 2001; 255(1-2):1–13. [PubMed: 11470281]
49. Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC, Ward DC. Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat Genet*. 1998; 19(3): 225–32. [PubMed: 9662393]
50. Schweitzer B, Roberts S, Grimwade B, Shao W, Wang M, Fu Q, Shu Q, Laroche I, Zhou Z, Tchernev VT, Christiansen J, Velleca M, Kingsmore SF. Multiplexed protein profiling on microarrays by rolling-circle amplification. *Nat Biotechnol*. 2002; 20(4):359–65. [PubMed: 11923841]

51. Shao W, Zhou Z, Laroche I, Lu H, Zong Q, Patel DD, Kingsmore S, Piccoli SP. Optimization of Rolling-Circle Amplified Protein Microarrays for Multiplexed Protein Profiling. *J Biomed Biotechnol.* 2003; 2003(5):299–307. [PubMed: 14688416]
52. Zhou H, Bouwman K, Schotanus M, Verweij C, Marrero JA, Dillon D, Costa J, Lizardi P, Haab BB. Two-color, rolling-circle amplification on antibody microarrays for sensitive, multiplexed serum-protein measurements. *Genome Biol.* 2004; 5(4):R28. [PubMed: 15059261]
53. Varnum SM, Woodbury RL, Zangar RC. A protein microarray ELISA for screening biological fluids. *Methods Mol Biol.* 2004; 264:161–72. [PubMed: 15020788]
54. Ptacek J, Devgan G, Michaud G, Zhu H, Zhu X, Fasolo J, Guo H, Jona G, Breikreutz A, Sopko R, McCartney RR, Schmidt MC, Rachidi N, Lee SJ, Mah AS, Meng L, Stark MJ, Stern DF, De Virgilio C, Tyers M, Andrews B, Gerstein M, Schweitzer B, Predki PF, Snyder M. Global analysis of protein phosphorylation in yeast. *Nature.* 2005; 438(7068):679–84. [PubMed: 16319894]
55. Lin YY, Lu JY, Zhang J, Walter W, Dang W, Wan J, Tao SC, Qian J, Zhao Y, Boeke JD, Berger SL, Zhu H. Protein acetylation microarray reveals that NuA4 controls key metabolic target regulating gluconeogenesis. *Cell.* 2009; 136(6):1073–84. [PubMed: 19303850]
56. Wegner GJ, Lee HJ, Marriott G, Corn RM. Fabrication of histidine-tagged fusion protein arrays for surface plasmon resonance imaging studies of protein-protein and protein-DNA interactions. *Anal Chem.* 2003; 75(18):4740–6. [PubMed: 14674449]
57. Unfricht DW, Colpitts SL, Fernandez SM, Lynes MA. Grating-coupled surface plasmon resonance: a cell and protein microarray platform. *Proteomics.* 2005; 5(17):4432–42. [PubMed: 16222719]
58. Gavin IM, Kukhtin A, Glesne D, Schabacker D, Chandler DP. Analysis of protein interaction and function with a 3-dimensional MALDI-MS protein array. *Biotechniques.* 2005; 39(1):99–107. [PubMed: 16060374]
59. Diamond DL, Zhang Y, Gaiger A, Smithgall M, Vedvick TS, Carter D. Use of ProteinChip array surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS) to identify thymosin beta-4, a differentially secreted protein from lymphoblastoid cell lines. *J Am Soc Mass Spectrom.* 2003; 14(7):760–5. [PubMed: 12837598]
60. Evans-Nguyen KM, Tao SC, Zhu H, Cotter RJ. Protein arrays on patterned porous gold substrates interrogated with mass spectrometry: detection of peptides in plasma. *Anal Chem.* 2008; 80(5):1448–58. [PubMed: 18254611]
61. Lee KB, Park SJ, Mirkin CA, Smith JC, Mrksich M. Protein nanoarrays generated by dip-pen nanolithography. *Science.* 2002; 295(5560):1702–5. [PubMed: 11834780]
62. Yan H, Park SH, Finkelstein G, Reif JH, LaBean TH. DNA-templated self-assembly of protein arrays and highly conductive nanowires. *Science.* 2003; 301(5641):1882–4. [PubMed: 14512621]
63. Popescu SC, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP. Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays. *Proc Natl Acad Sci U S A.* 2007; 104(11):4730–5. [PubMed: 17360592]
64. Hall DA, Zhu H, Zhu X, Royce T, Gerstein M, Snyder M. Regulation of gene expression by a metabolic enzyme. *Science.* 2004; 306(5695):482–4. [PubMed: 15486299]
65. Ho SW, Jona G, Chen CT, Johnston M, Snyder M. Linking DNA-binding proteins to their recognition sequences by using protein microarrays. *Proc Natl Acad Sci U S A.* 2006; 103(26):9940–5. [PubMed: 16785442]
66. Hu S, Xie Z, Onishi A, Yu X, Jiang L, Lin J, Rho HS, Woodard C, Wang H, Jeong JS, Long S, He X, Wade H, Blackshaw S, Qian J, Zhu H. Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell.* 2009; 139(3):610–22. [PubMed: 19879846]
67. Zhu J, Gopinath K, Murali A, Yi G, Hayward SD, Zhu H, Kao C. RNA-binding proteins that inhibit RNA virus infection. *Proc Natl Acad Sci U S A.* 2007; 104(9):3129–34. [PubMed: 17360619]
68. Huang J, Zhu H, Haggarty SJ, Spring DR, Hwang H, Jin F, Snyder M, Schreiber SL. Finding new components of the target of rapamycin (TOR) signaling network through chemical genetics and proteome chips. *Proc Natl Acad Sci U S A.* 2004; 101(47):16594–9. [PubMed: 15539461]

69. Kung LA, Tao SC, Qian J, Smith MG, Snyder M, Zhu H. Global analysis of the glycoproteome in *Saccharomyces cerevisiae* reveals new roles for protein glycosylation in eukaryotes. *Mol Syst Biol*. 2009; 5:308. [PubMed: 19756047]
70. Schnack C, Hengerer B, Gillardon F. Identification of novel substrates for Cdk5 and new targets for Cdk5 inhibitors using high-density protein microarrays. *Proteomics*. 2008; 8(10):1980–6. [PubMed: 18491313]
71. Lu JY, Lin YY, Qian J, Tao SC, Zhu J, Pickart C, Zhu H. Functional dissection of a HECT ubiquitin E3 ligase. *Mol Cell Proteomics*. 2008; 7(1):35–45. [PubMed: 17951556]
72. Vidal M, Brachmann RK, Fattaey A, Harlow E, Boeke JD. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc Natl Acad Sci U S A*. 1996; 93(19):10315–20. [PubMed: 8816797]
73. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MH, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*. 2006; 440(7084):637–43. [PubMed: 16554755]
74. Jones RB, Gordus A, Krall JA, MacBeath G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature*. 2006; 439(7073):168–74. [PubMed: 16273093]
75. Xie Z, Hu S, Blackshaw S, Zhu H, Qian J. hPDI: a database of experimental human protein-DNA interactions. *Bioinformatics*. 2010; 26(2):287–9. [PubMed: 19900953]
76. Popescu SC, Popescu GV, Bachan S, Zhang Z, Gerstein M, Snyder M, Dinesh-Kumar SP. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Dev*. 2009; 23(1):80–92. [PubMed: 19095804]
77. Foster MW, Forrester MT, Stamler JS. A protein microarray-based analysis of S-nitrosylation. *Proc Natl Acad Sci U S A*. 2009; 106(45):18948–53. [PubMed: 19864628]
78. Robinson WH, DiGennaro C, Hueber W, Haab BB, Kamachi M, Dean EJ, Fournel S, Fong D, Genovese MC, de Vegvar HE, Skrinier K, Hirschberg DL, Morris RI, Muller S, Pruijn GJ, van Venrooij WJ, Smolen JS, Brown PO, Steinman L, Utz PJ. Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nat Med*. 2002; 8(3):295–301. [PubMed: 11875502]
79. Hu S, Li Y, Liu G, Song Q, Wang L, Han Y, Zhang Y, Song Y, Yao X, Tao Y, Zeng H, Yang H, Wang J, Zhu H, Chen ZN, Wu L. A protein chip approach for high-throughput antigen identification and characterization. *Proteomics*. 2007; 7(13):2151–61. [PubMed: 17549792]
80. Hudson ME, Pozdnyakova I, Haines K, Mor G, Snyder M. Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. *Proc Natl Acad Sci U S A*. 2007; 104(44):17494–9. [PubMed: 17954908]
81. Chen CS, Sullivan S, Anderson T, Tan AC, Alex PJ, Brant SR, Cuffari C, Bayless TM, Talor MV, Burek CL, Wang H, Li R, Datta LW, Wu Y, Winslow RL, Zhu H, Li X. Identification of novel serological biomarkers for inflammatory bowel disease using *Escherichia coli* proteome chip. *Mol Cell Proteomics*. 2009; 8(8):1765–76. [PubMed: 19357087]
82. Lueking A, Huber O, Wirths C, Schulte K, Stieler KM, Blume-Peytavi U, Kowald A, Hensel-Wiegel K, Tauber R, Lehrach H, Meyer HE, Cahill DJ. Profiling of alopecia areata autoantigens based on protein microarray technology. *Mol Cell Proteomics*. 2005; 4(9):1382–90. [PubMed: 15939964]
83. Song Q, Liu G, Hu S, Zhang Y, Tao Y, Han Y, Zeng H, Huang W, Li F, Chen P, Zhu J, Hu C, Zhang S, Li Y, Zhu H, Wu L. Novel autoimmune hepatitis-specific autoantigens identified using protein microarray technology. *J Proteome Res*. 2010; 9(1):30–9. [PubMed: 19545157]
84. Zhu H, Hu S, Jona G, Zhu X, Kreiswirth N, Willey BM, Mazzulli T, Liu G, Song Q, Chen P, Cameron M, Tyler A, Wang J, Wen J, Chen W, Compton S, Snyder M. Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray. *Proc Natl Acad Sci U S A*. 2006; 103(11):4011–6. [PubMed: 16537477]

85. Tao SC, Chen CS, Zhu H. Applications of protein microarray technology. *Comb Chem High Throughput Screen*. 2007; 10(8):706–18. [PubMed: 18045082]
86. Kim SC, Sprung R, Chen Y, Xu Y, Ball H, Pei J, Cheng T, Kho Y, Xiao H, Xiao L, Grishin NV, White M, Yang XJ, Zhao Y. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell*. 2006; 23(4):607–18. [PubMed: 16916647]
87. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*. 2009; 325(5942):834–40. [PubMed: 19608861]
88. Zhao S, Xu W, Jiang W, Yu W, Lin Y, Zhang T, Yao J, Zhou L, Zeng Y, Li H, Li Y, Shi J, An W, Hancock SM, He F, Qin L, Chin J, Yang P, Chen X, Lei Q, Xiong Y, Guan KL. Regulation of cellular metabolism by protein lysine acetylation. *Science*. 2010; 327(5968):1000–4. [PubMed: 20167786]

Table 1
Application of functional protein microarrays in large-scale projects

Assay Type	Array Content	Type of Probe	No. of Probe	Application	Ref
Protein-Peptide Interaction	159 human SH2 and PTB domains	Peptide	61	Protein interaction network	[74]
Protein-DNA Interaction	282 yeast TFs	DNA motif	75	Protein-DNA interaction network	[65]
	4191 human proteins	DNA motif	460	Protein-DNA interaction network	[66]
Kinase Assay	Yeast proteome	Protein kinase	87	Phosphorylation network	[54]
Antigen-Antibody Interaction	196 human autoantigens	Autoimmune disease patient sera	>50	Antibody profiling	[78]
	82 coronavirus proteins	SARS patient sera	602	Antibody profiling	[84]
	<i>E. coli</i> proteome	IBD patient sera	134	Biomarker identification	[81]
	5005 human proteins	Ovarian cancer patient sera	60	Biomarker identification	[80]
	5011 human proteins	AIH patient sera	278	Biomarker identification	[83]