Multiplexed Protein Array Platforms for Analysis of Autoimmune Diseases

Imelda Balboni,[1,3,*] Steven M. Chan,[1,*] Michael Kattah,[1,*] Jessica D. Tenenbaum,[1,2,*] Atul J. Butte,[2,4] and Paul J. Utz[1,2]

[1]Department of Medicine, Division of Immunology and Rheumatology; [2]Department of Medicine, Stanford Medical Informatics; [3]Department of Pediatrics, Division of Rheumatology; [4]Department of Pediatrics, Division of Endocrinology; Stanford University School of Medicine, Stanford, California 94305; email: pjutz@stanford.edu

Annu. Rev. Immunol.
2006. 24:391–418

First published online as a Review in Advance on January 16, 2006

The *Annual Review of Immunology* is online at immunol.annualreviews.org

This article's doi:
10.1146/annurev.immunol.24.021605.090709

*The first four authors (I.B., S.M.C., M.K., and J.D.T.) contributed equally to this manuscript and are listed in alphabetical order.

0732-0582/06/0423-0391$20.00

Key Words

autoimmunity, proteomics, autoantibodies, cytokines, signaling

Abstract

Several proteomics platforms have emerged in the past decade that show great promise for filling in the many gaps that remain from earlier studies of the genome and from the sequencing of the human genome itself. This review describes applications of proteomics technologies to the study of autoimmune diseases. We focus largely on biased technology platforms that are capable of analyzing a large panel of known analytes, as opposed to techniques such as two-dimensional gel electrophoresis (2DIGE) or mass spectroscopy that represent unbiased approaches (as reviewed in 1). At present, the main analytes that can be systematically studied in autoimmunity include autoantibodies, cytokines and chemokines, components of signaling pathways, and cell-surface receptors. We review the most commonly used platforms for such studies, citing important discoveries and limitations that exist. We conclude by reviewing advances in biomedical informatics that will eventually allow the human proteome to be deciphered.

INTRODUCTION TO PROTEIN MICROARRAYS

Protein array/microarray: any platform in which a large number of proteins are immobilized on a solid support in a spatially (planar arrays) or spectrally (bead-based arrays) addressable manner

ELISA: enzyme-linked immunosorbent assay

Protein microarrays represent a validated platform for profiling protein levels and their post-translational modifications at a scale that is beyond what traditional techniques such as Western immunoblotting or enzyme-linked immunosorbent assays (ELISA) can realistically achieve. Protein microarrays have the potential to impact not only the broadly defined field of proteomics, but also other more defined biological disciplines, including immunology. As we describe here, the use of protein microarrays in immunology has largely been limited to the profiling of two main classes of proteins: (*a*) secreted factors (e.g., cytokines, chemokines, or growth factors) and (*b*) autoantibodies. Multiplexed analysis of intracellular proteins in a microarray format has proven to be a much more challenging task. The reverse-phase lysate (RPL) microarray platform stands out from other microarray platforms as the one with the greatest potential to achieve that goal. In the following sections, we review the work that has been published using each of the above mentioned technologies (summarized in **Table 1**) and speculate on their potential impact in the study of autoimmunity.

ANTIGEN MICROARRAYS

The relative frequency of autoreactive T and B cells in circulation is low, estimated at less than 1:10,000 lymphocytes (2–5), making the isolation and study of individual autoreactive

Table 1 Proteomics technologies for autoimmune disease research

Method	Format	Advantages	Limitations	Applications (Potential or Reported)	References
Autoantigen microarrays	Print peptides, purified/recombinant autoantigens, lipids, carbohydrates, DNA; probe with serum/fluid, detect autoantibodies	Profile many markers on small amounts of sample	Biased, currently limited to known autoantigens, controversial correlation with pathogenesis	Antigen-specific therapy, disease classification, monitor response to therapy	(8, 33)
Reverse-phase (lysate) protein microarrays	Print whole-cell lysates; profile with specific antibodies	Profile multiple samples with ≥100 antibodies, need only small amounts of lysate, monitor signal transduction events	Antibody cross-reactivity, multiple slides	Signaling defects in autoimmune cells (Tregs, autoreactive lymphocytes)	(1–3, 88–90)
Forward-phase (lysate) protein microarrays	Print/coat beads with antibodies against intracellular target, probe with lysate	Profile one or two samples simultaneously with many antibodies	Poor performance; denaturation of antibody; detection difficult (modify sample or generate two antibodies for every target); antibody cross-reactivity	Analysis of intracellular markers during apoptosis, could be used to profile intracellular disease markers	(135)

cells problematic. However, a hallmark of many autoimmune diseases is the production of high-titer, high-specificity autoantibodies directed against a variety of evolutionarily conserved molecules (6). These autoantibodies are readily detectable with various immunoassays, including ELISA, Western blotting, and immunoprecipitation. Although these assays do not directly provide data regarding the specificity of autoreactive T cells, there is a high degree of concordance between autoreactive B cell and T cell responses (7), and therefore the specificity of the autoantibody response is likely representative of the overall autoimmune response. Although such immunoassays are readily available and relatively easy to perform for a given autoantigen, analysis of multiple autoantibodies is costly, is time and labor intensive, and requires significant amounts of sera. Proteomic technologies including antigen array platforms enable the large-scale characterization of immune responses against foreign and self-antigens that may be involved in the development and progression of autoimmune disease.

Antigen microarrays allow the comprehensive analysis of autoantibodies directed against hundreds to thousands of antigens, including proteins, peptides, nucleic acids, macromolecular complexes (8–10), and, more recently, lipids (11). Probing these arrays requires microliter volumes of sera. Joos et al. (12) were the first group to describe a large-scale antigen microarray-based assay to detect serum autoantibodies. Eighteen autoantigens known to be serologic markers for several autoimmune diseases, including systemic lupus erythematosus (SLE), mixed connective tissue disease (MCTD), Sjögren's syndrome, scleroderma, and polymyositis were deposited onto nitrocellulose membranes or derivatized slides using a robotic microarrayer. Arrays were incubated with autoimmune sera followed by horseradish peroxidase (HRP)-conjugated secondary antibody. Chemiluminescent measurements were obtained following application of luminol substrate. These studies demonstrated that the microarray could be used to determine multiple autoantibody titers simultaneously in a single assay, and this assay was both sensitive and specific for autoantigen recognition (12).

In our laboratory, Robinson et al. (8) adapted the methods of others (13, 14) to design a 1152-feature connective tissue disease (CTD) array. The CTD array included 196 putative autoantigens targeted in several autoimmune diseases, including SLE, Sjögren's syndrome, rheumatoid arthritis (RA), polymyositis, scleroderma, and primary biliary cirrhosis. Antigens were spotted onto coated glass slides in an ordered array using a robotic arrayer. Arrays were incubated with either monoclonal antibodies or highly characterized autoimmune serum samples, washed, and probed with fluorescently labeled secondary antibodies. Arrays were scanned and fluorescence intensity measured. These studies demonstrated specific autoantibody binding that was linear over a 1000-fold range and was four to eight times more sensitive than ELISAs. To date, hundreds of highly characterized autoimmune serum samples have been analyzed and disease-specific autoantibody patterns detected. Using immunoglobulin G (IgG) subclass-specific secondary antibodies, the authors also demonstrated that the arrays could be used to characterize autoantibody subclasses that may be important in disease pathogenesis (8). The potential applications of antigen microarrays include (*a*) improved diagnosis of autoimmune and other diseases, (*b*) identification of autoantibody signatures that may represent subgroups of disease or have prognostic value, (*c*) monitoring of disease progression or response to therapy, (*d*) development of antigen-specific therapy, and (*e*) discovery of novel autoantigens or epitopes.

Diagnosis

Several autoimmune diseases are characterized by specific autoantibodies that are important in diagnosis. These include anti-double-stranded DNA (dsDNA) antibodies and

Proteomics: the use of techniques in molecular biology, biochemistry, and genetics to analyze the abundance, modifications, and interactions of a large number of proteins

Cytokines: secreted and soluble proteins that mediate communication between immune cells and their surrounding cells

Autoantibodies: any immunoglobulin that binds to self-antigens

RPL: reverse-phase lysate

Autoimmunity: a state in which an organism's immune system mounts a response to self-antigens, causing inflammation and damage to tissues and organs

CTD array: connective tissue disease array

RAST: radioallergosorbent test

anti-Smith antibodies in SLE, anti-U1-snRNP antibodies in MCTD, antiacetylcholine receptor antibodies in myasthenia gravis, and antithyroid stimulating hormone receptor antibodies in Graves' disease. Measurement of serum antibodies specific for multiple epitopes of a particular antigen have proven useful in identifying dominant linear epitopes in SLE antigens (18–21) and in increasing the sensitivity of detection of anticitrulline antibodies in RA by combining the frequencies of reactivity to a panel of citrulline-containing peptides (22, 23). Studies using antigen arrays allow for the simultaneous detection of hundreds of autoantibodies on one array with 1 μl or less of patient serum and therefore are ideal for profiling autoantibodies in SLE and other autoimmune diseases.

In addition to diagnosis of autoimmune disease, antigen microarrays have been designed to determine and monitor IgE reactivity profiles in patients with seasonal allergies. Hiller et al. (24) developed a microarray containing 94 purified allergens. They demonstrated that the array results were consistent with patients' known sensitization profiles based on skin testing or radioallergosorbent test (RAST)-based assays (24). Additional studies showed that these microarrays had a dynamic range comparable to RAST assays, that the sensitivity was similar to ELISA and exceeded that of RAST, and that no significant cross-reactivity was observed (25).

Antigen arrays have also been applied for serodiagnosis of infectious diseases. Mezzasoma et al. (26) designed arrays with antigens from multiple perinatal pathogens, including *Toxoplasma gondii*, cytomegalovirus, herpes simplex virus type 1 and 2, and rubella virus. Using a panel of characterized human sera, they validated the arrays, demonstrating a detection limit of 0.5 pg of antibody and sensitivity similar to ELISA (26). The arrays also included internal calibration curves for IgM and IgG, allowing quantification of individual immune responses (26). A comprehensive array for simian human immunodeficiency virus has also been used to dissect the B cell response in monkeys enrolled in an antigen-specific DNA vaccine trial (27).

With each of these microarray systems, the advantages are that the time to run the assay, the cost, and the amount of serum remain the same regardless of the number of analytes. In contrast, with ELISA each analyte must be assayed individually, increasing time, cost, and amount of serum required. In fact, clinical laboratories typically request at least 0.5 mL of serum for autoantibody studies, with a bare minimum of 0.15 mL per assay. Therefore, the amount of serum needed to test all the antigens on an array could be prohibitive, especially in seriously ill or pediatric patients.

Classification of Autoantibody Biosignatures and Prognostication

For many autoimmune diseases, autoantibody profiles not only provide diagnostic information but prognostic information as well, allowing patients to be divided into subgroups based on organ system involvement or disease severity. For example, in SLE anti-dsDNA antibodies are associated with active disease and nephritis; antiribosomal P antibodies are associated with neuropsychiatric lupus; and anti-Ro antibodies are associated with cutaneous lupus, photosensitivity, and the neonatal lupus syndrome (6, 28). Analogous examples can be found in scleroderma, polymyositis, and other diseases (6). Clinicians often use these limited profiles to determine prognosis and risk for disease flares.

Some reports suggest that serum from patients obtained prior to disease onset contained antibodies predictive of future disease. Analysis of serum samples from 130 people in the Department of Defense Serum Repository who eventually developed SLE revealed that 88% of patients had at least one autoantibody present in serum prior to the diagnosis of SLE. Certain autoantibodies, including antinuclear antibodies, antiphospholipid, anti-Ro, and anti-La antibodies, developed earlier than others (29). In addition, type 1 diabetes

(T1D) autoantibody profiles can be used to diagnose and predict future development of T1D. The presence of two or more autoantibodies directed against islet antigens, including insulin, glutamic acid decarboxylase, and tyrosine phosphatase-like protein IA-2, correlates with disease (30, 31).

Our laboratory, the Robinson laboratory, and others have developed several disease-specific antigen arrays containing putative antigens from the tissues affected by these diseases. These include the previously described CTD array, with almost 200 autoantigens from several rheumatic diseases (8); a synovial proteome array, with ~650 candidate RA autoantigens (32); a myelin proteome array, with ~500 myelin peptides and proteins (33); a vasculitis array (G. Alemi and P.J. Utz, unpublished data); and an islet cell proteome array, with pancreatic islet-derived autoantigens (34). Arrays of other biomolecules such as carbohydrates and lipids have also been created (11, 35). These arrays are being used to screen large cohorts of sera from patients and from animal models of disease to predict organ system involvement and to determine prognosis. Such arrays and other proteomic technologies will likely allow the identification of autoantibody biosignatures associated with a multitude of other autoimmune diseases in the future. In addition, these assays have the potential to enable better standardization and interpretation of clinical trials by allowing researchers to better classify patients prior to enrollment.

Relative levels of specific autoantibody isotypes may be important in the development and progression of SLE and other autoimmune diseases. Serum derived from patients with lupus nephritis often contains high-affinity IgG antibodies directed against dsDNA, although certain IgM antibodies against dsDNA have been associated with nephritis (36). Analysis of the antiribosomal P antibody response in two SLE patients showed that patients were well when their peak antiribosomal P response was of the IgM isotype and that the development of disease flares coincided with a switch to high-titer IgG antiribosomal P antibodies (37). Similarly, IgG subclasses may play a role in autoimmunity. Increases in IgG4 responses to desmoglein-1 are associated with onset of clinical disease in pemphigus foliaceus (38). Antigen arrays can be used to determine autoantibody isotype and subclass by probing the arrays with isotype-specific secondary antibodies differentially labeled with spectrally resolvable fluorophores (8). A preliminary study in our laboratory using antigen arrays showed that the presence of IgG2a antibodies against certain lupus autoantigens correlated with more severe nephritis on histopathologic evaluation in a murine model of lupus (K.L. Graham, unpublished observation). Using antigen arrays to characterize autoantibody isotypes may aid in the identification of important antigens for which pathogenic responses are generated and help discriminate other reactivities that do not impact disease progression or activity.

Monitoring Disease Progression and Response to Antigen-Specific Therapeutic Interventions

Following the initial stimulation of the immune system, autoantibody diversification occurs by the process of epitope spreading (7). Epitope spreading has been demonstrated in several autoimmune disease models (21, 39–44) and in patients with SLE (45–47) and T1D (48, 49). Investigators believe that the accrual of reactivities to different epitopes over time plays a role in the pathogenesis of these diseases, and epitope spreading has been associated with disease progression in lupus (46) and T1D (48).

Antigen arrays printed with overlapping peptides of known autoantigens provide a format for efficiently studying epitope spreading. Such studies could be useful in monitoring an individual patient's response to therapy and in determining which patients are at risk for progression of disease. Robinson et al. (33) have used myelin proteome arrays to study

experimental autoimmune encephalomyelitis (EAE), a murine model of multiple sclerosis (MS), and they demonstrated that the frequency of disease relapse was increased in mice with evidence of epitope spreading.

Current treatments for most autoimmune diseases are nonspecific and use agents that globally suppress the immune system. Such treatments have significant risks, including systemic infections and secondary malignancies, and there are many diseases for which response to such treatment is marginal at best. By determining the specific epitopes driving the autoimmune response, therapies that target only those cells that are reactive to these epitopes could be designed, leaving the rest of the immune system intact to function in its role of defense and surveillance. Robinson et al. (33) identified the early dominant epitopes targeted in EAE and used this information to develop tolerizing DNA vaccines encoding these epitopes. The DNA vaccines encoding myelin sheath components prevented epitope spreading and reduced relapses in mice, particularly when the vaccine was coadministered with a plasmid encoding interleukin-4 (IL-4) (33). Whether similar results can be obtained in human patients treated with antigen-specific or other immunomodulatory therapies remains to be determined.

Autoantigen Discovery

For many autoimmune diseases, including juvenile idiopathic arthritis, inflammatory bowel disease, psoriatic arthritis, and several vasculitides, the inciting autoantigen(s) is (are) currently unknown. cDNA expression libraries, peptide libraries, or arrayed fractions of tissues can be used to screen serum from autoimmune disease patients or animal models to discover novel autoantigens in these diseases. Once autoantigens are identified, further analyses can be performed to determine the sensitivity and specificity of the autoantibodies for a given disease. Arrays composed of bacterially expressed proteins have recently been used to identify candidate antigens in alopecia areata, an autoimmune disease leading to baldness (50). Newer methodology that uses in situ synthesis of proteins by mammalian ribosomes from cDNAs deposited on the surface of glass microscope slides may also enable such studies (51).

Limitations and Future Directions

Given the complex nature of proteins, optimal conditions for antigen arrays have not been established, and variation is seen using different slide surfaces and printing conditions (52, 53). In addition, following attachment of antigens to a planar surface, epitopes may be altered, resulting in lack of detection of autoantigens on the array (8, 13). Our laboratory has evaluated multiple (more than 25) commercially available slide surfaces and assay parameters to optimize arraying conditions for printing whole antigens and linear peptides. We have determined that FAST® slides (manufactured by Whatman® Schleicher & Scheull) consistently have lower coefficient of variance than other surfaces we analyzed, generally less than 25% and, under certain conditions, as low as 6%–8% (I. Balboni, unpublished observations). We are currently investigating other methods that will allow better internal control of the arrays. Internal control is particularly important for assays that will be used to monitor changes within patients over time, which may be quite subtle. Finally, arrays will need to be validated with many well-characterized serum samples before this platform will be ready for use in the clinical setting.

CYTOKINE ARRAYS

Proteomics platforms hold particular promise for the study of cytokines in autoimmunity. This is an important application because therapies that either increase or decrease cytokine levels have proven useful in the treatment of many autoimmune diseases (54–56). The therapeutic benefit of interferon-β in the treatment of relapsing-remitting MS is well

established (57), and therapies that inhibit IL-1, IL-6, IL-15, and tumor necrosis factor α (TNF-α) have been shown to improve RA (54, 55, 58–60). In addition to guiding the design of novel pharmacologic strategies, knowledge of cytokine levels adds to our understanding of disease pathogenesis and aids in identifying markers of disease. For example, in SLE, interferon signatures have been identified at the genetic level (61–63), and the increase in type I interferons has been verified at the protein level (64). In other cases, the role of cytokines is more complex. The controversial role of TNF-α as an inflammatory or immunomodulatory cytokine in MS is one example (65). In these situations in particular, the global analysis of multiple cytokines could provide a more complete understanding of complicated cytokine biology. Antibody arrays are a logical proteomics approach to studying serum proteins such as cytokines and chemokines in autoimmune disease. Some of the technological variations on this theme are highlighted here, as are their implications for the study of disease susceptibility.

Planar antibody arrays for analyzing cytokines in biological fluids were developed as a logical extension of the ELISA platform. The most common format involves a variation of the sandwich ELISA (66–68). This method requires two high-affinity cytokine binders, typically commercially available monoclonal or polyclonal antibodies. The capture reagent is spotted on a slide surface, a sample is applied and washed, and then a cocktail of labeled secondary reagents is added to bind available epitopes on the captured species. Detection typically involves fluorescent- or chemiluminescent-based methods (67–69). Some initial reports employing these assays were disappointing (70), but others have demonstrated reliable detection of multiple cytokines on planar arrays (71).

Because sensitivity is a concern in cytokine detection, rolling-circle amplification (RCA) has emerged as a well-suited detection method. RCA provides a means of linear signal amplification that remains localized to the microarray feature (72, 73). Arrays using this method operate exactly like other multiplex sandwich ELISAs, with the exception that the final step involves an oligonucleotide elongation reaction from a circular template. The first application of RCA to cytokine arrays was impressive. Schweitzer et al. (72) reported simultaneous detection of 75 human cytokines, femtomolar sensitivity, and a time-course analysis of cytokine secretion by mature dendritic cells (DCs) that confirmed and more importantly extended previously described results. Although RCA is slightly more elaborate than other detection methods, investigators have been able to simultaneously measure up to 180 cytokines using planar RCA arrays (74) and have developed two-color methods for analyzing two serum samples on the same array (74, 75).

RCA: rolling-circle amplification

Early studies established immuno-RCA as a powerful detection system, and a subsequent study demonstrated the utility of this microarray technology for investigating autoimmune disease in human patients (76). Kader and colleagues (76) investigated the cytokine profiles of pediatric patients with both Crohn's disease and ulcerative colitis, comparing clinical remission versus active disease. Although the authors anticipated increases in proinflammatory cytokines in patients with active disease, instead they detected an increase in so-called regulatory cytokines during remission. Placental growth factor, transforming growth factor-β1, IL-7, and IL-12p40 were all upregulated in serum from patients who were in clinical remission (76). These unexpected findings implicated a role for immune regulation in suppressing disease activity. Furthermore, the authors reported that a 10-cytokine classifier was better than a 4-cytokine classifier, suggesting that increased multiplexing of cytokine levels could better discriminate between disease states. With any developing technology, verification of important findings by an alternative measure is essential, and in this study data from conventional ELISA could have strengthened the findings. Nevertheless, this study exemplifies the advantages

of cytokine microarrays in studying autoimmune disease.

While the planar array format identifies capture antibodies by a coordinate on the slide, optically encoded microspheres can also serve as unique identifiers of the capture antibodies (bead arrays, or bead-based assays). Although adding additional features to suspension arrays is less trivial than for planar arrays, current bead array technologies can easily handle 100 analytes simultaneously and could be developed to handle thousands or tens of thousands of probes (77). The main advantages of this technology include easier quality control because beads are coated in bulk, higher density of array features, faster analysis of each sample, and higher throughput for multiple samples when compared with planar arrays (77). These advantages have made bead-based arrays useful for the study of cytokines in autoimmune disease. A study by Szodoray and colleagues (78) investigated Sjögren's syndrome using cytokine bead-based arrays. Serum was collected from Sjögren's syndrome patients, along with clinical and laboratory data. Differences among patients and controls were observed, including levels of IL-12p40, TNF-α, IL-6, and TNF receptor I (TNF-RI) and TNF-RII, some of which had not been previously linked to Sjögren's syndrome. Furthermore, subsets of Sjögren's syndrome patients showed differential cytokine levels. Levels of IL-2, epidermal growth factor (EGF), macrophage inflammatory protein-1α (MIP-1α), and TNF-RI were associated with elevations in erythrocyte sedimentation rate, but only IL-12p40 differed according to extraglandular manifestations. These studies demonstrate that quantitative, multiplexed analysis of serum cytokines in human autoimmune disease is indeed possible. Determining the scientific and clinical impact of these observations will be the next major step for cytokine arrays.

Planar arrays and bead-based suspension arrays identify capture antibodies by the use of encoded solid supports, whether that support is a microarray slide or a microsphere. An entirely separate approach to multiplex cytokine assays is the proximity assay, in which all reagents remain free in solution. Proximity assays have the advantages that no wash steps are necessary and no immobilization of capture reagents is required. Monogram BioSciences' eTAG™ system is an example of a proximity assay involving cleavable tags, although oligonucleotides are also useful for proximity-dependent cytokine detection (79). Fredriksson and colleagues (80) demonstrated the power of this oligonucleotide technology in highly sensitive cytokine detection using a pair of aptamers. Aptamers are probes made entirely of oligonucleotides that are selected through rounds of in vitro evolution for the ability to bind a target with high affinity. The aptamers used by Fredriksson et al. (80) were specific for two separate epitopes on platelet-derived growth factor. When the aptamers bound their target, the local concentration of the complementary aptamer ends increased by orders of magnitude, allowing detection of ligation by quantitative real-time polymerase chain reaction. After extensive optimization, zeptomole (10^{-21}) concentrations could be detected, which is far more sensitive than most cytokine assays. Multiplexing an assay of this kind is possible, although more difficult than planar or bead-based arrays. Subsequent application of proximity ligation employed oligonucleotide-conjugated anti-IL-2 and anti-IL-4 mono- and polyclonal antibodies, affirming the generalizability of this proteomics approach (81). Proximity assays require more sophisticated reagents and assay conditions than planar or bead-based arrays, but they can be developed for exquisite sensitivity, specificity, and throughput.

A variety of technological advances have made quantitative, multiplex cytokine profiling possible. The main advantages of these techniques over existing single analyte ELISA include small sample requirement, high-throughput capability, and reliable data on many markers simultaneously. The multiplexing ability of planar, bead-based, and

proximity assays will most likely not be the limiting factor in the development of these assays. One thing all the above technologies have in common, however, is absolute dependence on highly specific, high-affinity binders, sometimes relying on pairs of binders. Intense commercial and academic effort will be necessary to generate and evaluate all these reagents for every cytokine of interest. As these proteomics platforms evolve, studying cytokines should prove invaluable for investigating autoimmunity. These data could provide new therapeutic targets, a better understanding of pathogenesis, improved disease classification, and novel markers for measuring response to therapies. Because many therapies are beginning to target cytokines, and because cytokine levels can now be studied in a high-throughput fashion, in the future cytokine arrays may be able to guide patient-tailored cytokine therapy.

REVERSE-PHASE LYSATE (RPL) MICROARRAYS

The concept of RPL microarrays was originally described by Paweletz et al. (82) in a paper showing activation of the prosurvival signaling protein Akt and suppression of extracellular regulated kinases (ERK)1/2 phosphorylation in the transition from normal prostate epithelium to prostate intraepithelial neoplasia, and then to invasive prostate cancer. Although only a small number of patient samples and signaling proteins were analyzed, the potential to scale up this microarray platform for high-throughput analysis of intracellular proteins was clear. The ability to profile the abundance and activation state of a large panel of signaling proteins is particularly exciting for immunology research because most, if not all, immunological decisions depend on the signaling pathways that are turned on and their degree of activation.

Before we focus on the technical aspects of RPL microarrays, it is helpful to first distinguish between forward-phase and reverse-phase approaches and to discuss why RPL arrays have unique advantages over other array formats for protein profiling, especially for the study of intracellular signaling networks. In the forward-phase approach, analytes of interest are captured from solution phase by an array of immobilized antibodies. The bound analytes, usually proteins, are subsequently detected and quantified. A major challenge to the development of forward-phase arrays is the lack of antibodies that function in this format. In one study involving 115 antibody/antigen pairs, less than 20% of the antibodies provided specific and accurate measurements (13). Notably, in this study a 3% nonfat dry milk solution spiked with purified antigens was used to probe the arrays. Other groups have reported that only about 5% of intracellular antibodies are suitable for forward-phase arrays when the arrays are probed with a complex solution such as a cell lysate (83, 84). The forward-phase approach is further complicated by the issue of detection. Direct labeling of the protein sample with fluorescent tags is one approach and has been tried by several groups for protein microarray studies (13, 84, 85). However, the tags can sterically interfere with antibody binding if they are located within binding epitopes (83). Furthermore, the efficiency of protein labeling reactions is notoriously variable, making quantitative comparisons between samples difficult (13). Currently, the "sandwich immunoassay" approach appears to be the best option for forward-phase microarrays. However, antibody pairs are not available for many intracellular proteins.

The reverse-phase approach immobilizes lysate samples as distinct microspots on the array surface instead of antibodies (86) (**Figure 1**). RPL arrays are then probed with highly specific antibodies that are either phosphorylation state dependent for detecting activation states or independent for measuring abundance. Bound antibodies can then be detected using a secondary antibody that is directly conjugated to a fluorophore

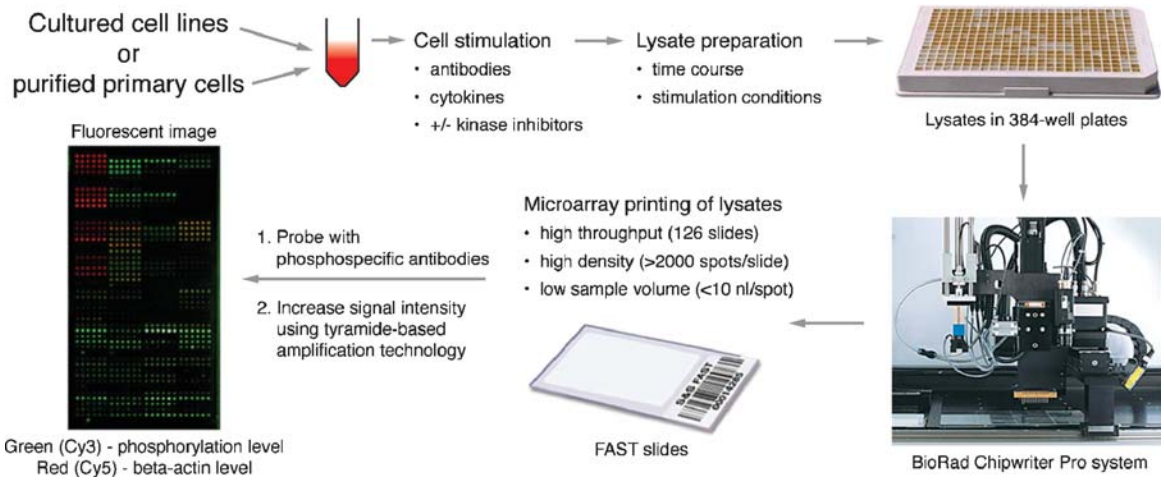

Figure 1

Reverse-phase lysate (RPL) microarrays. A schematic diagram illustrating the steps involved in sample collection, array fabrication, array processing, and data analysis. (Adapted from Reference 86 with permission.)

[e.g., quantum dots (87)] or enzymes (e.g., HRP) for signal amplification. In our experience, enzyme-mediated signal amplification is necessary for detecting low abundance proteins and their phosphorylated epitopes. We employ a tyramide-based amplification technique that deposits biotin molecules adjacent to HRP-conjugated secondary antibodies for signal amplification (88). A fluorescent signal is then generated by incubating the arrays with fluorophore-conjugated streptavidin. The fluorescent intensity of each spot correlates with the abundance or level of phosphorylation of the analyte in question.

The main advantage of this technique is that thousands of samples can be analyzed simultaneously on the same platform, greatly increasing throughput and simplifying quantitative comparisons between samples. Furthermore, an exceedingly small amount of sample (a few micrograms of protein) is required for printing hundreds of RPL arrays, thus permitting the comprehensive analysis of rare cell types and valuable patient samples. Importantly, most if not all commercially available antibodies should work in this format, provided they are not cross-reactive with other proteins. A clear limitation to this approach is that only one analyte can be measured on a single array. To partially overcome this limitation, arrays can be printed in a multisector format (arrays of arrays), allowing up to 16 analytes to be analyzed on a single slide. Although this number is still orders of magnitude smaller than the number of probes on a DNA microarray, it is not as limiting as one might think if the analysis is focused on one functional class of proteins (e.g., signal transduction proteins or apoptosis-related proteins). A particularly successful application of RPL microarrays has been the profiling of phosphorylation states of signaling proteins with the use of phosphorylation-specific antibodies (88). This application has only been possible within the past few years with the commercial availability of an expanding collection of well-characterized phosphorylation-specific antibodies.

The majority of work that has been done using RPL microarrays is in the field of oncology. This body of work, mostly contributed by Lance Liotta's and Emanuel Petricoin's research groups, involves the use of laser capture

microdissection (LCM) to isolate relevant portions of tumor specimens. The isolated cells are then lysed and spotted onto nitrocellulose-coated slides. Each spot contains the entire complement of proteins from a single specimen. RPL microarrays have been used to confirm changes in protein expression levels found using other techniques [e.g., DNA microarrays (89, 90), 2DIGE (91), immunohistochemistry (92, 93), and Western blot analysis (92)]. For instance, RPL microarrays were used to validate the finding that the regulatory signaling protein Rho G-protein dissociation inhibitor is selectively overexpressed in invasive human ovarian cancer versus low malignant potential ovarian cancer (91). More sophisticated studies involving phosphorylation-specific antibodies demonstrate dysregulated signaling in cancer cells (93). A study examining the phosphorylation states of ERK1/2 and Akt in ovarian cancer samples showed a trend toward increasing ERK1/2 phosphorylation with disease stage independent of histological type, but this same trend was not observed with phospho-Akt (94). Discovery of novel and unexpected signaling defects in diseased states (e.g., autoimmunity) is a powerful application of RPL microarrays, as these defects can potentially point to new therapeutic targets. This application is demonstrated in a prostate cancer study in which activation of PKC-α was found to be downregulated in tumor cells compared with normal epithelium by screening a panel of six phosphorylation antibodies using RPL arrays (95). Several studies have also reported early attempts to classify cancer cells into subtypes on the basis of their phosphorylation profiles. In a recent study comparing the phosphorylation states of 26 signaling proteins between primary and matched metastatic ovarian carcinomas, the level of c-Kit phosphorylation alone was found to be sufficient for categorizing samples as being of either primary or secondary origin (96). Interestingly, the phosphorylation-signaling signature appears to change dramatically when a primary tumor transforms into metastatic cancer (96, 97). In another study involving breast cancer specimens from 54 patients, RPL microarrays were used to profile both the activation state and abundance of 11 signaling proteins (97, 98). Hierarchical clustering of the data revealed four distinct tumor subtypes. Although the study did not report whether these subtypes correlated with disease progression and/or response to therapy, we predict that phosphorylation state profiling of clinical specimens will have clinical utility in classifying not only cancer patients, but also autoimmune patients into prognostic subgroups. LCM can be used to isolate relevant cells, including infiltrating lymphocytes, antigen-presenting cells, or cells that are targeted in the immune response (e.g., kidney cells in SLE, beta cells in T1D, or brain tissue in MS).

All the above mentioned work has been done using tumor specimens that are frozen immediately after resection. Thus, the specimen represents a snapshot of the tumor's signaling status at the time of resection. Evidence shows that dysregulation of signaling in cancer cells may only be revealed upon triggering with environmental stimuli (98a). We believe that a similar situation occurs in immune-related cells isolated from autoimmune patients. Therefore, we have focused our attention on stimulating living cells and monitoring their response to stimulation using RPL microarrays. We have successfully applied RPL microarrays to the study of signaling kinetics and pathway delineation in Jurkat T lymphocytes stimulated with phorbol myristate acetate (PMA) and with surface receptor antibodies to CD3 and CD28 (88). Furthermore, by monitoring changes in the phosphorylation state of 62 signaling proteins, we discovered a previously unrecognized link between CD3 crosslinking and dephosphorylation of Raf-1. Because only small amounts of sample are required for printing arrays, RPL microarrays have the potential to analyze rare primary cell populations. As a feasibility study, we profiled the phosphorylation state of 23 signaling proteins

LCM: laser capture microdissection

2DIGE: two-dimensional gel electrophoresis

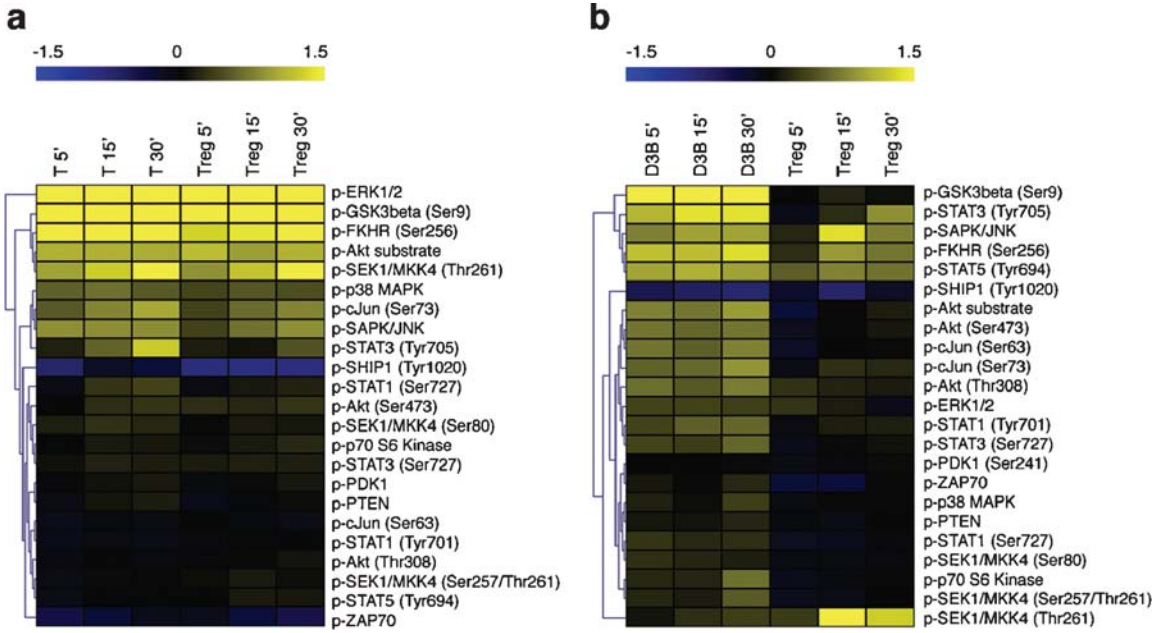

Figure 2

Profiling the phosphorylation state of 23 signaling proteins in Tregs in response to stimulation. (*a*) Naive $CD4^+CD25^-$ T cells (T) or $CD4^+CD25^+$ T cells (Treg) were freshly isolated and stimulated with PMA and ionomycin. (*b*) Day 3 T cell blasts (D3B) and freshly isolated $CD4^+CD25^+$ T cells were purified and stimulated with IL-2. Yellow intensities indicate increases in phosphorylation. Blue intensities indicate decreases in phosphorylation.

in $CD4^+CD25^+$ regulatory T cells (Tregs) isolated from naive mice. We previously reported differential STAT protein phosphorylation in these cells in response to IL-2 stimulation (88). Unlike naive $CD4^+CD25^-$ T cells, Tregs do not normally proliferate in response to T cell receptor (TCR) and CD28 stimulation in the absence of high concentrations of exogenous IL-2 (2000 U/ml) (99). However, some signals must be transmitted through the TCR because stimulation through the TCR is required for Tregs to exert suppression (100). The biochemical mechanism behind this nonproliferative state is not clear. Using RPL microarrays, we found that Tregs responded almost identically to naive $CD4^+CD25^-$ T cells when stimulated with PMA and ionomycin (**Figure 2*a***). This response suggests that the signaling pathways downstream of PMA and ionomycin are intact in Tregs. Alternatively, the pathways that become activated with IL-2 stimulation may be different between Tregs and activated T cell blasts that also express CD25. RPL microarray analysis revealed that the phosphorylation profiles of IL-2-stimulated Tregs and T cell blasts were dramatically different (**Figure 2*b***). Some signaling proteins such as STAT-5 appear to be activated equivalently, but in general the response in Tregs was clearly decreased relative to activated T cell blasts. In particular, Akt was not activated in response to IL-2 stimulation, which may in part explain the lack of proliferative response in Tregs.

In summary, we believe that RPL microarrays represent a new platform that will find numerous applications in the study of immunology. The proper functions of most immune-related cells depend on signals that are transmitted from the cell surface to the

nucleus. In autoimmune states, signal transduction pathways are often dysregulated, leading to inappropriate responses. The advent of RPL microarrays should provide a more complete picture of the dynamic signaling networks that occur in normal and disease states.

EMERGING TECHNOLOGIES

Fluorescence-Activated Cell Sorting (FACS) and FACS Signaling

This review has focused on protein array–based approaches for profiling autoantibody reactivity and intracellular signaling pathways and for measuring cytokine levels. Many of the processes involved in autoimmunity, however, involve interactions between cell-surface molecules. A growing number of therapies target cell-surface molecules in autoimmune disease, including altered peptide ligands, antibodies against adhesion or costimulatory molecules, and antibody depletion. Flow cytometry has been the traditional means of studying cell-surface markers in immunology. Currently, flow cytometric analysis can accommodate 17 fluorescent parameters and 2 scatter parameters (101). By interfacing the cytometer with autosampling instruments (102), the speed of analysis has rapidly increased, whereas the number of cells required has dramatically decreased. This technology will undoubtedly continue to evolve as a proteomics tool, but the complexities in instrumentation, reagent preparation, and fluorescence compensation are formidable obstacles (101). The advent of phospho FACS for studying signaling pathways in heterogeneous cell populations is certain to be a dominant technology platform over the next five years (98a,b) (see FACS Signaling: Emerging Technology). Recent work has combined planar arrays with analysis of whole cells and cell-surface markers. These novel arrays could complement flow cytometry as proteomics tools for the study of autoimmune disease.

FACS SIGNALING: EMERGING TECHNOLOGY

Planar protein arrays and flow cytometry assays have unique sets of advantages and disadvantages, allowing both platforms to contribute in different ways to the study of autoimmune disease. In addition to analyzing cell-surface markers, flow cytometry is now being used to study intracellular signaling events (98a,b). A comparison of RPL microarrays and phospho-flow highlights the complementary nature of the two technologies. The advantages of phospho-flow include the ability to handle heterogeneous cell populations, to perform single cell analysis, and to generate multiparameter data. RPL microarrays, however, have the advantages that more antibodies recognize their epitope in this format and fewer cells and smaller amounts of reagent are needed for parallel experiments than in phospho-flow cytometry. Combining the two technologies to study autoimmunity could circumvent the limitations of either approach, allowing the interpretation of more meaningful data. A similar approach could take advantage of both cell-surface marker arrays and conventional flow cytometry.

Peptide-Major Histocompatibility Complex (MHC) Arrays

Until recently, phenotypic analysis of antigen-specific T lymphocytes was limited to a flow cytometric approach involving single peptide-MHC tetramer staining. To overcome the limitations of this approach, Soen et al. (103) developed peptide-MHC tetramer arrays. This work has been reviewed in depth elsewhere (86). To briefly summarize, various peptide-MHC complexes were spotted on a slide, and the slide was probed with fluorescently labeled T cell populations. Stone et al. (104) advanced this tetramer microarray technology by incorporating more functional assays on the array. The authors successfully characterized the activation status and cytokine production of captured T cells. As proof of principle, the authors spotted anti-IL-2 capture antibodies with HLA-DR1 presenting an influenza hemagglutinin peptide (HA), as well as various other HLA-DR1-peptide complexes. After

FACS: fluorescence-activated cell sorting

incubation with an influenza-specific T cell hybridoma, IL-2 secretion was detectable only on the feature where HLA-DR1-HA had been immobilized, demonstrating the specificity of this technology and the ability to perform a miniaturized ELISPOT on the array. Antibodies against costimulatory molecules were also spotted with the peptide-tetramer complexes and were shown to enhance cytokine production. Similar to the study by Soen et al. (103), Stone and colleagues (104) reported a sensitivity of ~0.1%, putting their technology within the biological range of autoimmune T cell responses. Stone et al. (104) went on to report simultaneous detection of multiple cytokines and staining of bound T cells for activation markers. With this type of technology, investigators might consider profiling and characterizing the autoreactive T cell repertoire in autoimmune disease, in much the same way that array-based autoantibody profiling monitors the humoral immune response.

Although less sophisticated than the reports involving peptide-MHC arrays, arrays composed of antibodies directed against cell-surface markers have been used for immunophenotyping populations of cells. The experimental design involves spotting antibodies directed against cell-surface markers and incubating populations of cells on the array (105, 106). Each 500 μm feature can bind roughly 1500 cells, and the use of a planar surface is believed to enhance the avidity of interactions. The difficulty with interpreting these data is that cell immobilization depends on a poorly characterized function of antibody affinity, molecule copy number, and cell abundance. This developing technology suffers from a number of limitations, including the need for homogeneous populations, such as clonal leukemias and lymphomas. Despite these problems, the main advantage of planar arrays is that a few million cells can be profiled simultaneously for up to 90 surface markers (106). This multiplex screening could help guide more focused flow cytometry and functional studies.

INFORMATICS CHALLENGES IN PROTEOMICS

Informatics, generally defined, is the collection, classification, storage, and analysis of data. Recent developments in experimental methods have changed the nature of each one of these tasks in the context of proteomics. Classification and storage, although important, are relatively straightforward hurdles to scale. What is truly novel about postgenome era proteomics is found in the areas of data collection and analysis. One major change is that, increasingly, a number of methods in proteomics are as much about computationally collecting data as they are about generating data. Once the data have been generated or collected, the sheer scale necessitates new approaches to analysis. In this section, we address current challenges in the pursuits enumerated above, some solutions to these challenges, and computational methods in some commonly used experimental techniques. Our focus on relatively mature techniques should not be interpreted to mean that the experimental methods described above do not present their own set of informatics challenges. Rather, the techniques themselves are so recent that the informatics side of the problem has not yet had a chance to catch up. Storage and analysis of the large data sets generated by techniques like flow cytometry and protein microarrays are areas of active research, sure to yield innovative new approaches in the near future. We conclude by describing some exciting research areas and opportunities that are enabled by these new informatics developments.

Data Classification and Storage

Innovative proteomic techniques, and even some relatively mature ones, have led to the generation of novel types of data sets, unique both in character and in scale. As

both the quantity and variety of proteomic data have increased with recent developments in high-throughput experimental methods, so too have the challenges involved in managing these data. Whereas historically results have been stored predominantly in lab notebooks, gel films, and free text publications, recent years have seen a paradigm shift. The cause for this change is multifold: (*a*) Desktop machines have grown in their capacity for data storage and processing, (*b*) the World Wide Web has facilitated easy data exchange, and (*c*) experimental techniques have become high throughput in nature and as such produce larger quantities of data. In addition, new methods have been devised for automatically and efficiently converting what previously had been stored as qualitative analog data, e.g., 2DIGE images, into digital information, as we discuss below.

Interestingly, although the sheer quantity of proteomic data being produced is novel, these data are still largely of a qualitative nature. Most numeric values reported in proteomics are in somewhat arbitrary, or at least relative, units. Protein microarrays measure fluorescence intensity, but the actual numeric values are determined as much by the laser voltage level at which the slide is scanned as they are by the number of fluorophores present in a given spot. Mass spectrometry measures peak intensity values of different molecular fragments, but the units of intensity are relative to the maximum peak in a given experiment. Qualitative proteomic data may also take the form of protein-protein interaction data, protein presence and post-translational modification state, amino acid sequence, and three-dimensional structure.

Beyond the data themselves, far more challenging to informaticians are issues surrounding metadata, or data about data. For example, along with peak intensity in mass spectroscopy, or fluorescence level in FACS or protein arrays, a researcher is likely to record experimental conditions: sample tissue type, species, method of stimulation, etc. Such metadata must be stored in a structured way for them to be useful at some future date to another scientist trying to retrieve information specific to her/his own areas of interest. The term "structured" in this context can mean many things. At a minimum, it means that the metadata summarizing the experiment are recorded digitally and not in one long entry of prose. Rather, different fields are used to store attributes such as tissue type, stimulation method, duration, etc. Ideally, there is a predefined format or a controlled vocabulary for what may be entered in these fields so that, e.g., a search for all T cell experiments does not miss those described as CTL, $CD8^+$, Th, etc. Notably, these issues are only magnified as one moves toward a clinical setting, where standard practice is free text, paper-based patient records. In the clinic, as in the wet lab, implementation of electronic data entry systems is a constant tug-of-war between structuring data for ease of retrieval and exchange and flexibility to express ideas not easily summed up by, for example, selecting an option from a list.

Data Generation

As mentioned above, collection of data in proteomics often assumes the form of computational methods for generation of digital data from analog media. A prime example is that of image processing for analysis of 2DIGE. A plethora of packages exist for detecting spots on 2D gels, including MELANIE™, PDQuest™, Z3 and Z4000, Phoretix™, and Progenesis™ (107–109). Each package has its own algorithm, but the general method is to scan and digitize a picture, filter the noise from the digital image, and identify individual spots on the gel. Spots are often then manually extracted and analyzed using mass spectrometry to identify the protein of which that spot is composed (see below). Once these gel images have been processed and the information stored in a structured fashion, this information can be deposited into a data repository to be used by other researchers. The ExPASy (Expert

Protein Analysis System) proteomics server, developed by the Swiss Institute of Bioinformatics, hosts the Swiss-2DPAGE database in which a researcher can search for images based on various criteria (110). Once an image has been selected, the researcher can click on a labeled spot and find annotated information about that protein as determined by the original contributor.

Although not discussed at length in this review, mass spectrometry is another popular technique in which computational methods are needed to generate and analyze results. Data from mass spectrometry experiments originate as a stream of relative intensities of molecules displaying a variety of mass-to-charge ratios. These data may be used to identify components of a sample, to compare protein quantities in multiple samples, to sequence a protein de novo, or even to detect post-translational modifications such as phosphorylation or ubiquitination (111). Tens of thousands of spectra may be produced from a single experiment, and output files can be up to megabytes in size, depending on which specific technology is used. Different types of mass spectrometry vary by how the sample is separated prior to analysis, by how it is ionized, and by the method for mass analysis (112). An approach like liquid chromatography mass spectrometry will produce chromatograph data in addition to peak intensities. Tandem mass spectrometry breaks peptides recursively into smaller fragments, generating more peaks with each step.

For protein identification, the mass spectrometry results are compared against a database containing peak information for known proteins that have been fragmented with the same method as the sample. Advanced database search techniques are used to search these existing databases of mass-to-charge ratios of known proteins to identify a sample (113, 114). The success of this method relies on the availability and comprehensiveness of the database being searched.

To sequence a protein de novo, a combinatorial algorithm is used to piece together multiple peptide fragments (115–118) in much the same way stretches of sequenced DNA are assembled in the process of whole genome sequencing. Differences between peaks can correspond to the differences in mass of a single amino acid in the predigestion step. A sequence of differences may then be used to make predictions regarding the sequence of amino acids. Finally, mass spectrometry may also be used to compare both protein abundance and phosphorylation state in two different samples by using labeled isotopes (119). For protein abundance, peaks for fragments of each sample will be only as far apart as the difference in weight of the labeled isotope (multiplied by the number of atoms of the isotope in the fragment). If the protein is purified before being analyzed, then the relative intensities of these closely spaced peaks will reflect the relative abundance of the proteins from which they are derived. Similarly, if comparing the phosphorylation state of two purified peptides, one can compare the relative intensities of the peaks for the phosphorylated and nonphosphorylated amino acid, which will be separated by the weight of the phosphate group.

Standards for Data Storage and Exchange

Once proteomic data have been generated, they must be stored in such a way as to facilitate information retrieval, analysis, and exchange. As the open source model of scientific data grows in popularity and databases like the Database of Interacting Proteins and Biomolecular Interaction Network Database (BIND) grow and flourish, standards must be established as to what format the shared data will assume (120, 121). A number of collaborating organizations have been founded toward this end that focus on different aspects of proteomics standards.

The Proteomics Standards Initiative (PSI) is a working group within the

Human Proteome Organization (**http://www.hupo.org/**) that focuses on data storage standards. Whereas DNA microarrays have MIAME (minimum information about a microarray experiment), which researchers are required to follow to submit DNA microarray data to various journals and repositories (121), PSI has put forth MIAPE (minimum information about a proteomics experiment), based in part on the Proteomics Experiment Data Repository model, or PEDRo, as well as on an XML-based molecular interaction standard (122–124).

With respect to data exchange, the systems biology markup language, or SBML, is an XML-based language that allows researchers to share quantitative information relating to biochemical reaction networks, including cell signaling and gene regulation (125). BioPAX (**http://www.biopax.org/**) was founded in 2002, also for the purpose of creating an exchange format for biological pathway data. It defines an ontology for pathway data, that is, a set of terms and their relationships to each other, and is of a more qualitative nature than is SBML and less geared toward simulation.

Making Sense of the Results

So now that the proteomics data have been generated and stored, what is next? Inspecting a spreadsheet of fluorescence intensity values seldom gives insight into what new knowledge the data hold. Even a detailed set of reaction-based differential equations will generally fail to convey an intuition for what is taking place in the cell. A number of products exist, both commercial and academic, to enable visualization of molecular pathways and interactions. These have been commonly applied to the results of analysis using RNA expression microarrays, but they can be applied to proteomics results as well. At the simplest level, websites such as KEGG (**http://www.genome.jp/kegg/pathway.html**) and Biocarta (**http://www.biocarta.com**) have static images depicting a number of different pathways. More advanced software applications allow the user to create or infer pathways from complex data sets and also to map time series data onto pathway images (126–128). Additional software can find and sort pathways in order of potential interest based statistically on the genes or proteins involved (129, 130).

The underlying motivation for storing data in an organized fashion, in addition to facilitating human comprehension of those data, is to allow computers to operate on these large data sets. That is, computers are able to integrate thousands of data points at one time into a model of the underlying mechanisms far better than can the human mind. This modeling may be done at a high level of abstraction, for example to determine correlation or causality between phosphorylation events in autoreactive lymphocytes, or at a low level, for example using differential equations to model enzymatic reactions in a pathway. In a previous study (131), the authors developed a mathematical model consisting of ordinary differential equations to model 94 different compounds known to be involved in signaling downstream of EGF receptor stimulation. The kinetic parameters used were taken from the literature, determined experimentally, or computed from published time-dependent quantitative observations. Model results were in good agreement with experimental observations.

Clearly, this type of modeling could not be done without fairly detailed knowledge of known pathways, and so one may wonder what knowledge is gained from making such a model. First, researchers would have learned something new had the experimental results not reflected the model's predictions; this would have indicated that there was something amiss in what is currently recognized as truth in this system. Second, as the development and use of such models mature, researchers will be able to perform in silico experiments before doing them at the bench. This will save time and resources by avoiding experiments that the

Bayesian network: a graphical representation of a system in which nodes represent the entities and the edges represent probabilistic dependence relations between the nodes

model shows will not give good results, and it could enable scientists to do virtual experiments that would otherwise be impossible owing to cost, technology, or ethical considerations. But what can informatics techniques offer in areas where not much is known?

Sachs et al. (132) used Bayesian networks, a subfield of graph theory, and flow cytometry data to elucidate known and novel relationships in signaling pathways in human T cells. In computer science, a graph represents some number of nodes, and between these nodes relationships are depicted using edges, or lines. In the context of proteomics, graph nodes represent proteins and edges represent relationships between the proteins. The graph may be further enriched by storing information about these nodes and edges. For example, edges may be labeled as "binds to," "inhibits," "phosphorylates," etc. Bayesian networks are a subset of graph models where edges have a specific directionality to them, and each node reflects an event with a probability distribution determined by its parent nodes. So, an arrow from the "MEK is phosphorylated" node to the "ERK is phosphorylated" node strictly indicates a belief that ERK phosphorylation state is dependent on the MEK phosphorylation state and suggests that MEK phosphorylates ERK.

Sachs et al. (132) used data on the phosphorylation state of 11 different molecules in each of thousands of individual cells under various conditions. Clearly, this goes beyond the capacity of the human mind to process, but creating a Bayesian network from data requires very large data sets. Using these data, the authors predicted 15 known relationships in human T cell signaling as well as two novel connections not previously reported in these types of cells. These results were subsequently confirmed experimentally. This represents an exciting new approach toward elucidation of cell signaling in those pathways about which not much is known.

Challenges Ahead

One of the most exciting informatics challenges facing proteomics researchers is that of integration with other types of high-throughput data in molecular biology. As structured data are created, stored, and shared, both in proteomics and in other areas of biology experiencing similar growth, an opportunity arises to combine these different types of data in ways never before possible. In doing so, we enable exciting new systems approaches that allow researchers to combine proteomic information with other types of biological data, thus enabling new discoveries that rely on these multiple inputs. For example, Ideker et al. (133) used the protein-protein interaction data from the BIND database in conjunction with gene expression data for the proteins in that network. By analyzing which genes are coexpressed and whose protein products are known to interact, they determined coherent functional modules that could not be identified by using either data set alone.

As another illustration of this integrative approach, Mootha et al. (134) identified the gene involved in Leigh syndrome by combining DNA sequence, RNA expression, and mass spectrometry protein expression data. Having previously narrowed the search to a specific region of chromosome 2, and knowing that the disease involved mitochondrial function, they combined these different data sets to identify one gene as the site of mutation in this disease (134). With these types of scenarios as a promising start, it is clear that this data-driven, integrative approach will be key in understanding how molecular mechanisms function at all levels in the cell, which in turn is a key step to understanding why they fail and contribute to the development of autoimmunity (133, 134).

SUMMARY POINTS

1. Proteomics technologies fall into two major categories: unbiased technologies such as mass spectrometry that identify thousands of different proteins or peptides, and biased technology platforms such as FACS, antigen array, and antibody array platforms that use existing capture agents such as purified proteins, peptides, or antibodies.
2. The choice of technology platform is dictated by the hypothesis that is being tested and the availability of specific reagents.
3. Antibody profiles may identify autoimmune disease subsets or predict response to therapy.
4. Cytokine profiling can be employed to identify targets for therapeutic monoclonal antibody development.
5. A major challenge facing protein microarrays is the development of highly specific reagents.
6. Identification of biomarkers and surrogate markers of autoimmunity and tolerance represents a holy grail of clinical immunology.
7. Flow cytometry and emerging planar array technologies have complementary sets of advantages and disadvantages.
8. Various factors, including computing power, the Internet, and the quantity of experimental data being produced, motivate and necessitate novel approaches to data storage, analysis, and visualization.
9. The storage of experimental data in a structured format enables exciting new possibilities for automated and integrative biological research.

FUTURE ISSUES TO BE RESOLVED

1. As promising as protein arrays have become for studying autoimmunity, it is unclear if successful pilot studies in animal models will translate into similar findings in an outbred human population.
2. Which technology platform is correct for a given application is unclear; the current gold standard method for measurement of an analyte (e.g., cytokine or autoantibody measurements by ELISA) often does not correlate with results obtained by other platforms (e.g., bead-based or array-based assays), and it is not clear which data set to believe.
3. Major efforts should be organized to standardize reagents, assays, data storage, and normalization techniques.
4. The most important area for future development in proteomics is not the assay format or instrumentation, but rather creation of computational and statistical tools for analysis of disparate data sets such as transcript profiles, protein profiles, and cell-surface phenotypes.

DISCLOSURE STATEMENT

P.J.U. states the following conflict of interest disclosures, which are all directly relevant to the proteomics technology described in this manuscript: In the past three years he has served as a consultant to Becton Dickinson Biosciences (San Jose, CA) and Genentech, Inc. (South San Francisco, CA), is a member of the Scientific Advisory Board of Monogram Biosciences (formerly ACLARA Biosciences/Virologic) (South San Francisco, CA) and XDx, Inc. (South San Francisco, CA), and is a cofounder and consultant at Bayhill Therapeutics (Palo Alto, CA).

ACKNOWLEDGMENTS

We thank members of the Utz lab and the laboratories of Bill Robinson and Larry Steinman for development of proteomics technologies, some of which are described in this review. I.B. is funded by the Arthritis Foundation Postdoctoral Fellowship. S.M.C. and M.K. are funded by the Stanford MSTP and the Floren Family Foundation. J.D.T. is funded by a fellowship from the National Library of Medicine. A.J.B. is funded by the Lucille Packard Foundation for Children's Health and NIH grants K22LM008261, T32DK063702, and R01DK062948. P.J.U. is the recipient of a Donald E. and Delia B. Baxter Foundation Career Development Award and was supported by the Dana Foundation, the Northern California Chapter of the Arthritis Foundation, the Stanford Program in Molecular and Genetic Medicine, NIH Grants P30 DK56339 (Stanford Digestive Disease Center), DK61934, AI0151614, AI50854, AI50865, AI40093, AR49328, and NHLBI Proteomics Contract N01-HV-28183 during the period when many of these technologies were developed.

LITERATURE CITED

1. Fathman CG, Soares L, Chan S, Utz PJ. 2005. An array of possibilities for understanding the genome and proteome in autoimmunity. *Nature* 435:605–11
2. Anderson AC, Nicholson LB, Legge KL, Turchin V, Zaghouani H, Kuchroo VK. 2000. High frequency of autoreactive myelin proteolipid protein-specific T cells in the periphery of naive mice: mechanisms of selection of the self-reactive repertoire. *J. Exp. Med.* 191:761–70
3. Steinman L. 1996. Multiple sclerosis: a coordinated immunological attack against myelin in the central nervous system. *Cell* 85:299–302
4. Steinman L. 1996. A few autoreactive cells in an autoimmune infiltrate control a vast population of nonspecific cells: a tale of smart bombs and the infantry. *Proc. Natl. Acad. Sci. USA* 93:2253–56
5. Targoni OS, Baus J, Hofstetter HH, Hesse MD, Karulin AY, et al. 2001. Frequencies of neuroantigen-specific T cells in the central nervous system versus the immune periphery during the course of experimental allergic encephalomyelitis. *J. Immunol.* 166:4757–64
6. von Mühlen CA, Tan EM. 1995. Autoantibodies in the diagnosis of systemic rheumatic diseases. *Semin. Arthritis Rheum.* 24:323–58
7. Craft J, Fatenejad S. 1997. Self antigens and epitope spreading in systemic autoimmunity. *Arthritis Rheum.* 40:1374–82
8. Robinson WH, DiGennaro C, Hueber W, Haab BB, Kamachi M, et al. 2002. Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nat. Med.* 8:295–301

9. Robinson WH, Garren H, Utz PJ, Steinman L. 2002. Millennium Award. Proteomics for the development of DNA tolerizing vaccines to treat autoimmune disease. *Clin. Immunol.* 103:7–12
10. Robinson WH, Steinman L, Utz PJ. 2002. Proteomics technologies for the study of autoimmune disease. *Arthritis Rheum.* 46:885–93
11. Kanter J, Narayana S, Ho P, Catz I, Warren K, et al. 2006. Lipid microarrays identify key mediators of autoimmune brain inflammation. *Nat. Med.* In press
12. Joos TO, Schrenk M, Hopfl P, Kroger K, Chowdhury U, et al. 2000. A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics. *Electrophoresis* 21:2641–50
13. Haab BB, Dunham MJ, Brown PO. 2001. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* 2:RESEARCH0004
14. MacBeath G, Schreiber S. 2000. Printing proteins as microarrays for high-throughput function determination. *Science* 289:1760–63
15. Deleted in proof
16. Deleted in proof
17. Deleted in proof
18. James JA, Harley JB. 1996. Human lupus anti-spliceosome A protein autoantibodies bind contiguous surface structures and segregate into two sequential epitope binding patterns. *J. Immunol.* 156:4018–26
19. James JA, Harley JB. 1998. B-cell epitope spreading in autoimmunity. *Immunol. Rev.* 164:185–200
20. James JA, Mamula MJ, Harley JB. 1994. Sequential autoantigenic determinants of the small nuclear ribonucleoprotein Sm D shared by human lupus autoantibodies and MRL lpr/lpr antibodies. *Clin. Exp. Immunol.* 98:419–26
21. James JA, Gross T, Scofield RH, Harley JB. 1995. Immunoglobulin epitope spreading and autoimmune disease after peptide immunization: Sm B/B'-derived PPPGMRPP and PPPGIRGP induce spliceosome autoimmunity. *J. Exp. Med.* 181:453–61
22. Hueber W, Kidd B, Toomoka B, Lee B-J, Bruce B, et al. 2005. Antigen microarray profiling of autoantibodies in rheumatoid arthritis. *Arthritis Rheum.* 52:2645–55
23. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. 1998. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J. Clin. Invest.* 101:273–81
24. Hiller R, Laffer S, Harwanegg C, Huber M, Schmidt WM, et al. 2002. Microarrayed allergen molecules: diagnostic gatekeepers for allergy treatment. *FASEB J.* 16:414–16
25. Jahn-Schmid B, Harwanegg C, Hiller R, Bohle B, Ebner C, et al. 2003. Allergen microarray: comparison of microarray using recombinant allergens with conventional diagnostic methods to detect allergen-specific serum immunoglobulin E. *Clin. Exp. Allergy* 33:1443–49
26. Mezzasoma L, Bacarese-Hamilton T, Di Cristina M, Rossi R, Bistoni F, Crisanti A. 2002. Antigen microarrays for serodiagnosis of infectious diseases. *Clin. Chem.* 48:121–30
27. de Vegvar HEN, Amara RR, Steinman L, Utz PJ, Robinson HL, Robinson WH. 2003. Microarray profiling of antibody responses against simian-human immunodeficiency virus: postchallenge convergence of reactivities independent of host histocompatibility type and vaccine regimen. *J. Virol.* 77:11125–38
28. Routsias JG, Tzioufas AG, Moutsopoulos HM. 2004. The clinical value of intracellular autoantigens B-cell epitopes in systemic rheumatic diseases. *Clin. Chim. Acta* 340:1–25

29. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, et al. 2003. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N. Engl. J. Med.* 349:1526–33
30. Pietropaolo M, Eisenbarth GS. 2001. Autoantibodies in human diabetes. *Curr. Dir. Autoimmun.* 4:252–82
31. Verge CF, Gianani R, Kawasaki E, Yu L, Pietropaolo M, et al. 1996. Prediction of type I diabetes in first-degree relatives using a combination of insulin, GAD, and ICA512bdc/IA-2 autoantibodies. *Diabetes* 45:926–33
32. Hueber W, Utz PJ, Steinman L, Robinson WH. 2002. Autoantibody profiling for the study and treatment of autoimmune disease. *Arthritis Res.* 4:290–95
33. **Robinson WH, Fontoura P, Lee BJ, de Vegvar HE, Tom J, et al. 2003. Protein microarrays guide tolerizing DNA vaccine treatment of autoimmune encephalomyelitis. *Nat. Biotechnol.* 21:1033–39**
34. Quintana FJ, Hagedorn PH, Elizur G, Merbl Y, Domany E, Cohen IR. 2004. Functional immunomics: microarray analysis of IgG autoantibody repertoires predicts the future response of mice to induced diabetes. *Proc. Natl. Acad. Sci. USA* 101:14615–21
35. Wang DN, Liu SY, Trummer BJ, Deng C, Wang AL. 2002. Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nat. Biotechnol.* 20:275–81
36. Hahn BH. 1998. Antibodies to DNA. *N. Engl. J. Med.* 338:1359–68
37. Stafford HA, Anderson CJ, Blalock DB, Reichlin M. 1998. Development of the antiribosomal P autoantibody response. *Clin. Exp. Rheumatol.* 16:119–24
38. Warren SJ, Arteaga LA, Rivitti EA, Aoki V, Hans-Filho G, et al. 2003. The role of subclass switching in the pathogenesis of endemic pemphigus foliaceus. *J. Invest. Dermatol.* 120:104–8
39. Kaufman DL, Clare-Salzler M, Tian J, Forsthuber T, Ting GS, et al. 1993. Spontaneous loss of T-cell tolerance to glutamic acid decarboxylase in murine insulin-dependent diabetes. *Nature* 366:69–72
40. Keech CL, Gordon TP, McCluskey J. 1996. The immune response to 52-kDa Ro and 60-kDa Ro is linked in experimental autoimmunity. *J. Immunol.* 157:3694–99
41. Lehmann PV, Forsthuber T, Miller A, Sercarz EE. 1992. Spreading of T-cell autoimmunity to cryptic determinants of an autoantigen. *Nature* 358:155–57
42. McCarron RM, Fallis RJ, McFarlin DE. 1990. Alterations in T cell antigen specificity and class II restriction during the course of chronic relapsing experimental allergic encephalomyelitis. *J. Neuroimmunol.* 29:73–79
43. Tisch R, Yang XD, Singer SM, Liblau RS, Fugger L, McDevitt HO. 1993. Immune response to glutamic acid decarboxylase correlates with insulitis in non-obese diabetic mice. *Nature* 366:72–75
44. Topfer F, Gordon T, McCluskey J. 1995. Intra- and intermolecular spreading of autoimmunity involving the nuclear self-antigens La (SS-B) and Ro (SS-A). *Proc. Natl. Acad. Sci. USA* 92:875–79
45. Arbuckle MR, Reichlin M, Harley JB, James JA. 1999. Shared early autoantibody recognition events in the development of anti-Sm B/B' in human lupus. *Scand. J. Immunol.* 50:447–55
46. Orth T, Mayet WJ, Deister H, Chang S, Schmitz M, Bachmann M. 1998. Analysis of epitope spreading over an eleven-year period in a patient with systemic lupus erythematosus. *Scand. J. Rheumatol.* 27:461–64

33. Describes a series of elegant experiments that demonstrate the use of arrays to identify antigens to include in a therapeutic vaccine in an animal model of MS.

47. Scofield RH, Zhang F, Kurien BT, Anderson CJ, Reichlin M, et al. 1996. Development of the anti-Ro autoantibody response in a patient with systemic lupus erythematosus. *Arthritis Rheum.* 39:1664–68
48. Naserke HE, Ziegler AG, Lampasona V, Bonifacio E. 1998. Early development and spreading of autoantibodies to epitopes of IA-2 and their association with progression to type 1 diabetes. *J. Immunol.* 161:6963–69
49. Sohnlein P, Muller M, Syren K, Hartmann U, Bohm BO, et al. 2000. Epitope spreading and a varying but not disease-specific GAD65 antibody response in Type I diabetes. The Childhood Diabetes in Finland Study Group. *Diabetologia* 43:210–17
50. Lueking A, Huber O, Wirths C, Schulte K, Stieler K, et al. 2005. Profiling of alopecia areata autoantigens based on protein microarray technology. *Mol. Cell Proteomics* 4:1382–90
51. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, et al. 2004. Self-assembling protein microarrays. *Science* 305:86–90
52. Angenendt P, Glokler J, Murphy D, Lehrach H, Cahill DJ. 2002. Toward optimized antibody microarrays: a comparison of current microarray support materials. *Anal. Biochem.* 309:253–60
53. Angenendt P, Glokler J, Sobek J, Lehrach H, Cahill DJ. 2003. Next generation of protein microarray support materials: evaluation for protein and antibody microarray applications. *J. Chromatogr. A* 1009:97–104
54. Feldmann M. 2002. Development of anti-TNF therapy for rheumatoid arthritis. *Nat. Rev. Immunol.* 2:364–71
55. Feldmann M, Steinman L. 2005. Design of effective immunotherapy for human autoimmunity. *Nature* 435:612–19
56. Vilcek J, Feldmann M. 2004. Historical review: cytokines as therapeutics and targets of therapeutics. *Trends Pharmacol. Sci.* 25:201–9
57. Revel M. 2003. Interferon-β in the treatment of relapsing-remitting multiple sclerosis. *Pharmacol. Ther.* 100:49–62
58. Bresnihan B, Alvaro-Gracia JM, Cobby M, Doherty M, Domljan Z, et al. 1998. Treatment of rheumatoid arthritis with recombinant human interleukin-1 receptor antagonist. *Arthritis Rheum.* 41:2196–204
59. McInnes IB, Gracie JA. 2004. Interleukin-15: a new cytokine target for the treatment of inflammatory diseases. *Curr. Opin. Pharmacol.* 4:392–97
60. Nishimoto N, Yoshizaki K, Miyasaka N, Yamamoto K, Kawai S, et al. 2004. Treatment of rheumatoid arthritis with humanized anti-interleukin-6 receptor antibody: a multicenter, double-blind, placebo-controlled trial. *Arthritis Rheum.* 50:1761–69
61. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, et al. 2003. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl. Acad. Sci. USA* 100:2610–15
62. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, et al. 2003. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J. Exp. Med.* 197:711–23
63. Kirou KA, Lee C, George S, Louca K, Peterson MG, Crow MK. 2005. Activation of the interferon-α pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. *Arthritis Rheum.* 52:1491–503
64. Bengtsson AA, Sturfelt G, Truedsson L, Blomberg J, Alm G, et al. 2000. Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies. *Lupus* 9:664–71

65. Magnano MD, Robinson WH, Genovese MC. 2004. Demyelination and inhibition of tumor necrosis factor (TNF). *Clin. Exp. Rheumatol.* 22:S134–40
66. Huang RP. 2004. Cytokine protein arrays. *Methods Mol. Biol.* 264:215–31
67. Tam SW, Wiese R, Lee S, Gilmore J, Kumble KD. 2002. Simultaneous analysis of eight human Th1/Th2 cytokines using microarrays. *J. Immunol. Methods* 261:157–65
68. Wang CC, Huang RP, Sommer M, Lisoukov H, Huang R, et al. 2002. Array-based multiplexed screening and quantitation of human cytokines and chemokines. *J. Proteome Res.* 1:337–43
69. Moody MD, Van Arsdell SW, Murphy KP, Orencole SF, Burns C. 2001. Array-based ELISAs for high-throughput analysis of human cytokines. *BioTechniques* 31:186–90, 192–94
70. Copeland S, Siddiqui J, Remick D. 2004. Direct comparison of traditional ELISAs and membrane protein arrays for detection and quantification of human cytokines. *J. Immunol. Methods* 284:99–106
71. Nagorsen D, Marincola FM, Panelli MC. 2004. Cytokine and chemokine expression profiles of maturing dendritic cells using multiprotein platform arrays. *Cytokine* 25:31–35
72. Schweitzer B, Roberts S, Grimwade B, Shao WP, Wang MJ, et al. 2002. Multiplexed protein profiling on microarrays by rolling-circle amplification. *Nat. Biotechnol.* 20:359–65
73. Schweitzer B, Wiltshire S, Lambert J, O'Malley S, Kukanskis K, et al. 2000. Inaugural article: immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. *Proc. Natl. Acad. Sci. USA* 97:10113–19
74. Shao WP, Zhou ZM, Laroche I, Lu H, Zong QL, et al. 2003. Optimization of rolling-circle amplified protein microarrays for multiplexed protein profiling. *J. Biomed. Biotechnol.* 2003:299–307
75. Zhou H, Bouwman K, Schotanus M, Verweij C, Marrero JA, et al. 2004. Two-color, rolling-circle amplification on antibody microarrays for sensitive, multiplexed serum-protein measurements. *Genome Biol.* 5:R28
76. Kader HA, Tchernev VT, Satyaraj E, Lejnine S, Kotler G, et al. 2005. Protein microarray analysis of disease activity in pediatric inflammatory bowel disease demonstrates elevated serum PLGF, IL-7, TGF-β1, and IL-12p40 levels in Crohn's disease and ulcerative colitis patients in remission versus active disease. *Am. J. Gastroenterol.* 100:414–23
77. Nolan JP, Sklar LA. 2002. Suspension array technology: evolution of the flat-array paradigm. *Trends Biotechnol.* 20:9–12
78. Szodoray P, Alex P, Brun JG, Centola M, Jonsson R. 2004. Circulating cytokines in primary Sjögren's syndrome determined by a multiplex cytokine array system. *Scand. J. Immunol.* 59:592–99
79. Chan-Hui P-Y, Stephens K, Warnock RA, Singh S. 2003. Applications of eTag(R)(TM) assay platform to systems biology approaches to clinical immunology studies. *Clin. Immunol.* 111:162–74
80. Fredriksson S, Gullberg M, Jarvius J, Olsson C, Pietras K, et al. 2002. Protein detection using proximity-dependent DNA ligation assays. *Nat. Biotechnol.* 20:473–77
81. Gullberg M, Gustafsdottir SM, Schallmeiner E, Jarvius J, Bjarnegard M, et al. 2004. Cytokine detection by antibody-based proximity ligation. *Proc. Natl. Acad. Sci. USA* 101:8420–24

76. Investigated cytokines and growth factors in inflammatory bowel disease using cytokine arrays; one of the best demonstrations of the clinical applications of this technology.

78. Analyzed healthy individuals and patients with primary Sjögren's syndrome using bead-based arrays; found certain cytokine profiles correlate with specific laboratory and clinical parameters.

82. Paweletz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, et al. 2001. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* 20:1981–89

83. Macbeath G. 2002. Protein microarrays and proteomics. *Nat. Genet.* 32:526–32

84. Nielsen UB, Cardone MH, Sinskey AJ, MacBeath G, Sorger PK. 2003. Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc. Natl. Acad. Sci. USA* 100:9330–35

85. Sreekumar A, Nyati MK, Varambally S, Barrette TR, Ghosh D, et al. 2001. Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. *Cancer Res.* 61:7585–93

86. Utz PJ. 2005. Protein arrays for studying blood cells and their secreted products. *Immunol. Rev.* 204:264–82

87. Geho D, Lahar N, Gurnani P, Huebschman M, Herrmann P, et al. 2005. Pegylated, steptavidin-conjugated quantum dots are effective detection elements for reverse-phase protein microarrays. *Bioconjug. Chem.* 16:559–66

88. Chan SM, Ermann J, Su L, Fathman CG, Utz PJ. 2004. Protein microarrays for multiplex analysis of signal transduction pathways. *Nat. Med.* 10:1390–96

89. Nishizuka S, Charboneau L, Young L, Major S, Reinhold WC, et al. 2003. Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc. Natl. Acad. Sci. USA* 100:14229–34

90. Nishizuka S, Chen ST, Gwadry FG, Alexander J, Major SM, et al. 2003. Diagnostic markers that distinguish colon and ovarian adenocarcinomas: identification by genomic, proteomic, and tissue array profiling. *Cancer Res.* 63:5243–50

91. Jones MB, Krutzsch H, Shu H, Zhao Y, Liotta LA, et al. 2002. Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. *Proteomics* 2:76–84

92. Herrmann PC, Gillespie JW, Charboneau L, Bichsel VE, Paweletz CP, et al. 2003. Mitochondrial proteome: altered cytochrome *c* oxidase subunit levels in prostate cancer. *Proteomics* 3:1801–10

93. Zha H, Raffeld M, Charboneau L, Pittaluga S, Kwak LW, et al. 2004. Similarities of prosurvival signals in Bcl-2-positive and Bcl-2-negative follicular lymphomas identified by reverse phase protein microarray. *Lab. Invest.* 84:235–44

94. Wulfkuhle JD, Aquino JA, Calvert VS, Fishman DA, Coukos G, et al. 2003. Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays. *Proteomics* 3:2085–90

95. Grubb RL, Calvert VS, Wulkuhle JD, Paweletz CP, Linehan WM, et al. 2003. Signal pathway profiling of prostate cancer using reverse phase protein arrays. *Proteomics* 3:2142–46

96. Sheehan KM, Calvert VS, Kay EW, Lu Y, Fishman D, et al. 2005. Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol. Cell Proteomics* 4:346–55

97. Petricoin EF 3rd, Bichsel VE, Calvert VS, Espina V, Winters M, et al. 2005. Mapping molecular networks using proteomics: a vision for patient-tailored combination therapy. *J. Clin. Oncol.* 23:3614–21

98. Petricoin E, Zoon K, Kohn E, Barrett J, Liotta L. 2002. Clinical proteomics: translating benchside promise into bedside reality. *Nat. Rev.* 1:683–95

98a. Irish JM, Hovland R, Krutzik PO, Perez OD, Bruserud O, et al. 2004. Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* 118:217–28

82. The original description of the RPL microarray technology in combination with LCM.

88. Employed RPL microarray technology to study dynamic changes in the phosphorylation state of numerous signaling proteins in T cells upon stimulation.

98b. Perez OD, Nolan GP. 2002. Simultaneous measurement of multiple active kinase states using polychromatic flow cytometry. *Nat. Biotechnol.* 20:155–62
99. Tang Q, Henriksen KJ, Bi M, Finger EB, Szot G, et al. 2004. In vitro-expanded antigen-specific regulatory T cells suppress autoimmune diabetes. *J. Exp. Med.* 199:1455–65
100. Sakaguchi S. 2004. Naturally arising $CD4^+$ regulatory T cells for immunologic self-tolerance and negative control of immune responses. *Annu. Rev. Immunol.* 22:531–62
101. Perfetto SP, Chattopadhyay PK, Roederer M. 2004. Seventeen-colour flow cytometry: unravelling the immune system. *Nat. Rev. Immunol.* 4:648–55
102. Edwards BS, Oprea T, Prossnitz ER, Sklar LA. 2004. Flow cytometry for high-throughput, high-content screening. *Curr. Opin. Chem. Biol.* 8:392–98
103. Soen Y, Chen DS, Kraft DL, Davis MM, Brown PO. 2003. Detection and characterization of cellular immune responses using peptide-MHC microarrays. *PLoS Biol.* 1:E65
104. Stone JD, Demkowicz WE Jr, Stern LJ. 2005. HLA-restricted epitope identification and detection of functional T cell responses by using MHC-peptide and costimulatory microarrays. *Proc. Natl. Acad. Sci. USA* 102:3744–49
105. Belov L, de la Vega O, dos Remedios CG, Mulligan SP, Christopherson RI. 2001. Immunophenotyping of leukemias using a cluster of differentiation antibody microarray. *Cancer Res.* 61:4483–89
106. Belov L, Huang P, Barber N, Mulligan SP, Christopherson RI. 2003. Identification of repertoires of surface antigens on leukemias using an antibody microarray. *Proteomics* 3:2147–54
107. Appel RD, Hochstrasser DF, Funk M, Vargas JR, Pellegrini C, et al. 1991. The MELANIE project: from a biopsy to automatic protein map interpretation by computer. *Electrophoresis* 12:722–35
108. Appel RD, Vargas JR, Palagi PM, Walther D, Hochstrasser DF. 1997. Melanie II—a third-generation software package for analysis of two-dimensional electrophoresis images: II. Algorithms. *Electrophoresis* 18:2735–48
109. Marengo E, Robotti E, Antonucci F, Cecconi D, Campostrini N, Righetti PG. 2005. Numerical approaches for quantitative analysis of two-dimensional maps: a review of commercial software and home-made systems. *Proteomics* 5:654–66
110. Sanchez J, Chiappe D, Converset V, Hoogland C, Binz P, et al. 2001. The mouse SWISS-2D PAGE database: a tool for proteomics study of diabetes and obesity. *Proteomics* 1:136–63
111. Gygi S, Corthals G, Zhang Y, Rochon Y, Aebersold R. 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. USA* 97:9390–95
112. Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. *Nature* 422:198–207
113. Gevaert K, Vandekerckhove J. 2000. Protein identification methods in proteomics. *Electrophoresis* 21:1145–54
114. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–67
115. Bruni R, Gianfranceschi G, Koch G. 2005. On peptide de novo sequencing: a new approach. *J. Pept. Sci.* 11:225–34
116. Lin T, Glish GL. 1998. C-terminal peptide sequencing via multistage mass spectrometry. *Anal. Chem.* 70:5162–65

104. Describes planar microarrays of recombinant MHC-peptide complexes, costimulatory molecules, and anti-cytokine capture antibodies.

117. Stults JT. 1990. Peptide sequencing by mass spectrometry. *Methods Biochem. Anal.* 34: 145–201
118. Taka H, Kaga N, Fujimura T, Mineki R, Imaizumi M, et al. 2000. Rapid determination of parvalbumin amino acid sequence from Rana catesbeiana (pI 4.78) by combination of ESI mass spectrometry, protein sequencing, and amino acid analysis. *J. Biochem.* 127:723–29
119. Oda Y, Owa T, Sato T, Boucher B, Daniels S, et al. 2003. Quantitative chemical proteomics for identifying candidate drug targets. *Anal. Chem.* 75:2159–65
120. Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, et al. 2005. NCBI GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res.* 33:D562–66
121. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, et al. 2001. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nat. Genet.* 29:365–71
122. Hermjakob H, Montecchi-Palazzi L, Bader G, Wojcik J, Salwinski L, et al. 2004. The HUPO PSI's molecular interaction format—a community standard for the representation of protein interaction data. *Nat. Biotechnol.* 22:177–83
123. Orchard S, Hermjakob H, Julian RK Jr, Runte K, Sherman D, et al. 2004. Common interchange standards for proteomics data: public availability of tools and schema. *Proteomics* 4:490–91
124. Taylor CF, Paton NW, Garwood KL, Kirby PD, Stead DA, et al. 2003. A systematic approach to modeling, capturing, and disseminating proteomics experimental data. *Nat. Biotechnol.* 21:247–54
125. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, et al. 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19:524–31
126. Demir E, Babur O, Dogrusoz U, Gursoy A, Nisanci G, et al. 2002. PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways. *Bioinformatics* 18:996–1003
127. Nikitin A, Egorov S, Daraselia N, Mazo I. 2003. Pathway studio—the analysis and navigation of molecular networks. *Bioinformatics* 19:2155–57
128. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13:2498–504
129. Doniger SW, Salomonis N, Dahlquist KD, Vranizan K, Lawlor SC, Conklin BR. 2003. MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biol.* 4:R7
130. Draghici S, Khatri P, Bhavsar P, Shah A, Krawetz SA, Tainsky MA. 2003. Onto-Tools, the toolkit of the modern biologist: Onto-Express, Onto-Compare, Onto-Design and Onto-Translate. *Nucleic Acids Res.* 31:3775–81
131. Schoeberl B, Eichler-Jonsson C, Gilles ED, Muller G. 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat. Biotechnol.* 20:370–75
132. Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. 2005. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308:523–29
133. Ideker T, Ozier O, Schwikowski B, Siegel AF. 2002. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18(Suppl. 1):S233–40

121. Put forth standards that have been adopted by numerous journals and repositories as a required format for sharing microarray data.

131. Demonstrated that detailed modeling of a biological system can be accurate and robust as determined by experimental findings and sensitivity analysis.

132. Set a precedent for using a data set with the thousands of replicates commonly desired in Bayesian analysis; led to two experimentally verifiable biological discoveries in a relatively well-known system.

134. Mootha VK, Lepage P, Miller K, Bunkenborg J, Reich M, et al. 2003. Identification of a gene causing human cytochrome *c* oxidase deficiency by integrative genomics. *Proc. Natl. Acad. Sci. USA* 100:605–10
135. Rhyne PW, Scull JD, Stiles LM, Eisinger DP. 2003. Analysis of apoptotic cells using Beadlyte suspension arrays. *Biotechniques* 35:624–29

Annual Review
of Immunology

Volume 24, 2006

Contents